Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Psychometric validation of the Portuguese SCIROCCO tool for assessing the maturity of Local Health Units in providing integrated care: a cross-sectional study

**Maria Carvalho**

mjf.carvalho@ensp.unl.pt

NOVA National School of Public Health (NSPH)

**Henrique Oliveira**

Instituto de Telecomunicações

**Cláudia Almeida**

NOVA National School of Public Health (NSPH)

**João Santos**

School of Economics, Management and Political Science

**Andrea Pavlickova**

NHS National Services Scotland

**Rui Santana**

NOVA National School of Public Health (NSPH)

**Research Article**

**Keywords:** SCIROCCO tool, Delivery of healthcare, Maturity model, Psychometric, Validation study

**Posted Date:** May 26th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9380518/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

**Background**: The transition toward integrated care is crucial for addressing modern healthcare challenges. In Portugal, recent reforms established Local Health Units (LHUs) to integrate primary and hospital care, yet a validated tool to assess their institutional maturity is lacking. This study aimed to translate and validate the Scaling Integrated Care in Context (SCIROCCO) tool into Portuguese and assess its psychometric properties, specifically structural validity and internal consistency.

**Methods**: A cross-sectional study was conducted following a rigorous translation and cultural adaptation process. The sample comprised 105 healthcare professionals from a Portuguese LHU. Structural validity was evaluated using Exploratory Factor Analysis (EFA) based on a polychoric correlation matrix with Robust Diagonally Weighted Least Squares extraction. Item Response Theory (IRT), specifically the Graded Response Model, was applied to estimate item discrimination and difficulty parameters. Reliability was determined using Composite Reliability.

**Results**: The EFA supported a unidimensional structure, explaining 61.5% of the total variance, with excellent model fit indices (RMSEA = 0.057; CFI = 0.996; TLI = 0.995). Factor loadings were strong, ranging from 0.557 to 0.941. The scale demonstrated excellent internal consistency and precision (Composite Reliability = 0.94). IRT analysis revealed strong discrimination parameters across items and a monotonic increase in difficulty thresholds, confirming the tool's effectiveness in capturing the different levels of maturity.

**Conclusions**: The Portuguese version of the SCIROCCO tool is a valid, reliable, and psychometrically robust instrument. By confirming its unidimensional structure and excellent internal consistency, this study establishes a foundational benchmark for the Portuguese healthcare context. The tool empowers policymakers and healthcare managers to accurately assess LHU maturity, fostering evidence-based strategies for successful integrated care implementation.


# Background

Integrated care has emerged as a policy response to the complex challenges that health systems face, including aging populations, the prevalence of chronic diseases, public health emergencies, and resource constraints (Grooten et al., 2018; Read et al., 2019; WHO, 2018). These multifaceted issues often result in inefficiencies, fragmented service delivery, and poor patient outcomes. Integrated care seeks to mitigate these challenges by promoting coordination among providers and sectors, thereby reducing unnecessary care and improving access, quality, and efficiency of care. This is supported by interoperable information systems, shared clinical protocols and multidisciplinary teamwork (Expert Group on Health Systems Performance Assessment, 2017; Health at a Glance: Europe 2024, 2024).

However, integrated care implementation is complex due to inconsistent data and lack of standardized definitions, making it challenging to inform policy development and measure progress (Health at a Glance: Europe 2024, 2024). Furthermore, resource limitations, fragmented governance and institutional immaturity compromise the viability of these initiatives, increasing the risk of project failure (Expert Group on Health Systems Performance Assessment, 2017; Hughes et al., 2022). Therefore, policymakers require structured frameworks to identify gaps and areas for improvement, ensuring that health systems are prepared to provide integrated care and can tailor integrated care approaches to local contexts (Henderson & Pavlickova, 2022).

In response to this need, the SCIROCCO (Scaling Integrated Care in Context) tool was developed to assess the maturity and readiness of healthcare systems to adopt and scale up integrated care initiatives (Henderson & Pavlickova, 2022). The tool evaluates 12 domains essential to integrated care, providing insights into the strengths and weaknesses of healthcare systems, fostering the dialogue among different stakeholders, and guiding policymakers in prioritizing areas for improvement (Henderson et al., 2022; Peytremann-Bridevaux et al., 2021).

Derived from the B3-Maturity Model, SCIROCCO has been successfully applied across 24 regions and countries, demonstrating adaptability to diverse healthcare system contexts (Expert Group on Health Systems Performance Assessment, 2017; Peytremann-Bridevaux et al., 2021). It has supported initiatives such as the development of integrated care in nursing homes in Catalonia, goal-oriented care in Flanders, long-term care reform in Slovenia and Estonia, and digital neighborhood development in Germany (Henderson & Pavlickova, 2022). Moreover, SCIROCCO has supported major policy reforms, including the creation of the Agency of Health and Social Care Integration in Puglia, the “Care Puglia 3.0” initiative (Mingolla, 2020; Wang et al., 2022), and Switzerland’s 2020-2030 Health Policy Strategy (Peytremann-Bridevaux et al., 2021).

In Portugal, the SCIROCCO tool is particularly significant due to the healthcare reform in early 2024, which created 31 new Local Health Units (LHUs) in addition to the eight already in existence. The Portuguese healthcare reform merged primary care units and hospitals located in the same regional area into a single organizational and management structure (Goiana-da-Silva et al., 2024; Nunes, 2024). However, no systematic assessment of the maturity of Portuguese LHUs using the SCIROCCO framework has yet been accomplished (Goiana-da-Silva et al., 2024). This lack of assessment represents a serious gap, as it is essential to ensure the reform’s success and sustainability (Peytremann-Bridevaux et al., 2021). The SCIROCCO tool can provide Portuguese policymakers and institutions with valuable insights into priority areas for intervention, supporting the transition to integrated care and improving long-term healthcare outcomes (Henderson & Pavlickova, 2022).

Given the growing interest in integrated care and the lack of validated maturity assessment tools in Portugal, this study aimed to translate and validate the SCIROCCO tool into Portuguese by evaluating its psychometric properties, specifically its structural validity and internal consistency, building upon the foundational measurement properties established by Grooten et al. (2019).

## Methods

# Study design

This cross-sectional study aimed to translate and validate the SCIROCCO tool as an instrument to assess LHUs’ maturity in providing integrated care in the Portuguese context. The study comprised two phases:

- **Phase 1**: Translation of the SCIROCCO tool into Portuguese;
- **Phase 2**: Evaluation of the psychometric properties of the translated Portuguese version.

# Sample and data collection

Participants were recruited through convenience sampling. The Litoral Alentejano LHU Executive Committee sent recruitment emails to all staff members. To be eligible, participants had to be at least 18 years old and currently

employed by that LHU. Data was collected between September 2024 and January 2025, using the LimeSurvey platform. All survey questions were mandatory, and the system was configured to record only completed submissions. The platform's built-in logic automatically discarded incomplete responses, resulting in a complete-case dataset for analysis. A minimum subject-to-item ratio of 10:1 was targeted to ensure the stability of the factor analysis (Hair et al., 2019).

The protocol of this study was approved by the Ethics Committee of the National School of Public Health of the NOVA University of Lisbon (CE-ENSP) on September 23rd, 2024 (Approval No. 05/2024). Informed consent was obtained from all the participants.

# The SCIROCCO tool

The SCIROCCO tool consists of 12 items: "readiness to change", "structure and governance", "digital infrastructure", "process coordination", "funding", "removal of inhibitors", "population approach", "citizen empowerment", "evaluation methods", "breadth of ambition", "innovation management" and "capacity building" (Grooten et al., 2019). Each dimension is described in terms of its objectives and assessed on a scale from 0 (indicating "not fulfilled" or "not present/available") to 5 (equivalent to "fully fulfilled" or "totally present/available").

Background variables were collected to characterize participants based on their personal, academic, and professional profiles. These included age, sex, educational level, and current professional role (Nurse, Physician, Senior Health Officer, Allied Health Professional, Board of Directors Member, Head of Department/ Unit Coordinator, Other healthcare professional, other profession). Participants also specified their primary working setting (hospital, primary care, or long-term care).

# Phase 1: Translation and cultural adaptation of the Portuguese version of the SCIROCCO questionnaire

The SCIROCCO tool, originally developed by the B3 Action Group in Integrated Care and validated by Grooten et al. (2019), was used to develop a Portuguese version through a formal translation and cross-cultural adaptation process, following the guidelines of Cruchinho et al. (2024). The translated version, titled "'*Maturidade das Unidades Locais de Saúde para a prestação de Cuidados Integrados*', is provided as a supplementary file (Additional File 1).

First, formal permission was obtained from the original author to translate and validate the English version of the questionnaire (Grooten et al., 2019) for the Portuguese context. Forward translation was conducted by two independent native Portuguese speakers fluent in English. The first was knowledgeable about the questionnaire's constructs, and the second one was an independent translator, unaware of the tool's objectives. A research team member merged these two versions into a unified translation used for back translation.

The back translation (from Portuguese to English) was conducted by two professional translators, both native English speakers fluent in Portuguese. One of them was unfamiliar with the original version of the questionnaire, and the other was knowledgeable about the tool's construct.

This retroverted version was compared to the original English tool by a fourth independent translator to evaluate semantic, idiomatic, and concept equivalence (harmonization). Technical terms were adjusted to maintain linguistic and cultural relevance, and the original author reviewed the final version to ensure the scale's construct remained consistent with the original.

Finally, a pre-test of the questionnaire was conducted with 10 Portuguese healthcare professionals from the target population. Participants were debriefed on the questionnaire coherence, length, perceptibility, and time to complete it. Their feedback was incorporated to refine the final version of the scale and ensure optimal clarity and cultural appropriateness.

# Phase 2: Psychometric properties of the Portuguese version of the SCIROCCO tool

The psychometric properties of the SCIROCCO tool were evaluated through structural validity and internal consistency. Structural validity assesses the degree to which the score of an instrument adequately reflects the dimensionality of the construct, thus supporting construct validity. Internal consistency is a measure of reliability that assesses a scale's homogeneity, reflecting the degree to which the scale's items are intercorrelated (Boateng et al., 2018).

# Methods for structural validity and internal consistency assessment

Before conducting the Exploratory Factor Analysis (EFA), the polychoric intercorrelation coefficient ($\rho$) was calculated to characterize the strength of association among the instrument items, as all responses were treated as ordinal data (Zumbo et al., 2007). The Kaiser-Meyer-Olkin (KMO) index was calculated to assess sampling adequacy for factor analysis, with a minimum threshold of 0.50 considered acceptable (Watkins, 2018). Bartlett's test of sphericity was used to assess whether the correlation matrix significantly deviated from an identity matrix; p-values <0.05 indicated a favorable matrix for factor analysis (Izquierdo et al., 2014). Additionally, item-level suitability was evaluated using Kaiser's Measure of Sampling Adequacy (MSA), and items were retained if they met the minimum threshold of 0.50 (Lorenzo-Seva & Ferrando, 2021a).

EFA was conducted to assess the factor structure of the SCIROCCO tool items, using the FACTOR software (version 12.06.08; Lorenzo-Seva & Ferrando, 2019). The analysis was carried out based on the polychoric correlation matrix and using the Robust Diagonally Weighted Least Squares (RDWLS) as the factor extraction method, as this method accounts for the sampling variability of the values in the correlation matrix (Ferrando & Lorenzo-Seva, 2014; Forero et al., 2009).

Parallel analysis, with random permutation of the observed data, was used to determine the number of factors to retain (Gaskin & Happell, 2014; Timmerman & Lorenzo-Seva, 2011).

Regarding solution adequacy and item inclusion, the model's total variance explained was evaluated against an average threshold of 56.6%, a threshold that is frequently cited in meta-analyses of psychometric and behavioral studies (Peterson, 2000). Items with factor loadings exceeding 0.50 were retained, following Hair et al. (2019),

who identify such loadings as practically significant and indicative that the factor explains at least 25% of the item's variance.

The scale unidimensionality was assessed using the Unidimensional Congruence (UniCo), the Explained Common Variance (ECV), and the Mean of Item Residual Absolute Loadings (MIREAL) indexes, following the work of Fernando & Lorenzo-Seva (2018). A UniCo value greater than 0.95, an ECV value greater than 0.85, and a MIREAL value less than 0.30 each suggest that the data can be treated as essentially unidimensional (Ferrando & Lorenzo-Seva, 2018).

Following factor extraction, the model fit was assessed using the Root Mean Square of Residuals (RMSR), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI) fit indices (Brown, 2015; Hair et al., 2019; Lorenzo-Seva & Ferrando, 2019). RMSR values should be below 0.08 (Hu & Bentler, 1999), while according to Brown (2015) and Hair et al (2019), RMSEA values should be less than 0.08, with the upper bound of the confidence interval (CI) not exceeding 0.10. CFI and TLI values should be above 0.90, preferably 0.95 (Brown, 2015).

Item response theory (IRT) using the Graded Response Model (GRM) was used to assess the scale's validity, in which the latent trait ($\theta$) represents the LHUs' maturity for integrated care. IRT describes the relationship between the latent trait, the properties of the items in the scale, and respondents' answers to the individual items. For each item, two parameters were estimated: (a) discrimination, which identifies how effectively an item distinguishes between LHUs at different maturity levels; and (b) difficulty, which represents the maturity threshold that an LHU must reach to have a 0.5 probability of endorsing a specific maturity stage (Yang & Kao, 2014).

Lastly, the factor structure's reliability was assessed through Composite Reliability (CR), employing the approach outlined by Fornell and Larcker (1981). This approach is grounded in standardized factor loadings and measurement error variances, thus providing a more accurate estimate of construct reliability (Valentini & Damásio, 2016). CR values above 0.70 were adopted as indicative of adequate construct reliability (Hair et al., 2019).

## Results

# Demographics

Table 1 shows the demographic characteristics of the 105 study participants. The 36-45 age group accounted for the largest proportion (52.4%). Most participants were female (72.4%) and had a master's degree (62.9%). Various professional roles were represented, with nurses comprising the largest group (43.8%). Regarding the work environment, most participants were employed in hospitals (44.8%).

**Table 1**: Sociodemographic characterization of the sample.

| Variable | n | % |
|---|---|---|
| Age group (years) | | |
| < 25 years | 1 | 1.0 |
| 25 – 35 | 7 | 6.7 |
| 36 – 45 | 55 | 52.4 |
| 46 – 55 | 32 | 30.5 |
| 56 – 65 | 6 | 5.7 |
| > 66 | 4 | 3.8 |
| Sex | | |
| Female | 76 | 72.4 |
| Male | 29 | 27.6 |
| Education level | | |
| Secondary education | 7 | 6.7 |
| Bachelor’s degree | 29 | 27.6 |
| Master’s degree | 66 | 62.9 |
| PhD | 3 | 2.9 |
| Professional role | | |
| Nurse | 46 | 43.8 |
| Physician | 13 | 12.4 |
| Senior Health Officer | 14 | 13.3 |
| Allied Health professionals | 8 | 7.6 |
| Board of Directors member | 6 | 5.7 |
| Head of Department/ Unit Coordinator | 6 | 5.7 |
| Other healthcare professional | 10 | 9.5 |
| Other profession | 2 | 1.9 |
| Work setting | | |
| Hospital | 47 | 44.8 |
| Primary care | 42 | 40.0 |
| Long-term care | 13 | 12.4 |
| Other | 3 | 2.9 |

*Note*. *N* = 105

# Summary of statistics and item distribution

Before assessing the scale's structural validity and internal consistency, a descriptive analysis of the items was conducted to examine their preliminary psychometric properties. As shown in Table 2, the mean scores ranged from 1.31 to 2.32, with median values ranging from 1 to 2.

The SCIROCCO items showed positive skewness (ranging from 0.14 to 1.35) and varying levels of kurtosis (ranging from -1.01 to 1.19). This combination, together with the ordinal nature of the data, typically results in a non-normal multivariate distribution. Mardia's (1970) multivariate skewness and kurtosis analysis confirmed this scenario, as the null hypothesis that the data follow a multivariate normal distribution was rejected, thereby justifying the use of a polychoric correlation matrix rather than a Pearson matrix, as well as the need to employ robust extraction methods in EFA.

**Table 2**: Distribution of items and summary statistics.

| Items | Item Distribution (%) | | | | | | *M* ± *SD* | Mdn (IQR) | *Sk* | *Ku* |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | | | | |
| 1 | 2.9 | 52.4 | 15.2 | 18.1 | 7.6 | 3.8 | 1.87±1.21 | 1 (1-3) | 0.95 | 0.01 |
| 2 | 13.3 | 35.2 | 33.3 | 4.8 | 7.6 | 5.7 | 1.75±1.31 | 2 (1-2) | 0.94 | 0.50 |
| 3 | 12.4 | 43.8 | 13.3 | 7.6 | 18.1 | 4.8 | 1.90±1.47 | 1 (1-3) | 0.66 | -0.82 |
| 4 | 8.6 | 22.9 | 36.2 | 21.9 | 8.6 | 1.9 | 2.05±1.14 | 2 (1-3) | 0.22 | -0.21 |
| 5 | 39.0 | 26.7 | 13.3 | 7.6 | 8.6 | 4.8 | 1.34±1.51 | 1 (0-2) | 1.02 | -0.03 |
| 6 | 22.9 | 54.3 | 4.8 | 8.6 | 6.7 | 2.9 | 1.31±1.26 | 1 (1-1) | 1.35 | 1.19 |
| 7 | 7.6 | 19 | 35.2 | 17.1 | 13.3 | 7.6 | 2.32±1.34 | 2 (1-3) | 0.31 | -0.51 |
| 8 | 12.4 | 49.5 | 16.2 | 10.5 | 9.5 | 1.9 | 1.61±1.23 | 1 (1-2) | 0.94 | 0.16 |
| 9 | 30.5 | 22.9 | 17.1 | 12.4 | 13.3 | 3.8 | 1.67±1.54 | 1 (0-3) | 0.54 | -0.89 |
| 10 | 21.9 | 10.5 | 26.7 | 17.1 | 13.3 | 10.5 | 2.21±1.62 | 2 (1-3) | 0.14 | -1.01 |
| 11 | 8.6 | 39.0 | 22.9 | 17.1 | 7.6 | 4.8 | 1.91±1.28 | 2 (1-3) | 0.71 | -0.13 |
| 12 | 14.3 | 41.9 | 14.3 | 17.1 | 4.8 | 7.6 | 1.79±1.42 | 1 (1-3) | 0.83 | -0.13 |

*Note*: *N* = 105. *M* = Mean, *SD* = Standard Deviation, Mdn = Median, IQR = Interquartile Range (25% - 75%), *Sk* = Skewness, *Ku* = Kurtosis

# Structural validity and internal consistency

The suitability of the data for EFA was established through the analysis of the polychoric correlation matrix, Bartlett's test of sphericity, and KMO statistics. Item correlations ranged from 0.288 to 0.887, thus supporting the scale's homogeneity. Bartlett's test of sphericity was significant ($\chi^2(66) = 981.9$, $p < 0.001$), suggesting that the correlation matrix was statistically distinguishable from the identity matrix. The KMO statistic was 0.898,

indicating meritorious sampling adequacy, with all individual MSA values above the recommended threshold of 0.50, ranging from 0.869 to 0.962.

The parallel analysis suggested retaining only one factor, as it was the only one with an empirical eigenvalue higher than that of the permuted real data. This single-factor solution accounted for 61.5% of the model's variance explained, which is above the 50th percentile reported by Peterson (2000) in his meta-analysis (see Table 1 in Peterson, 2000). The factor loadings ranged from 0.557 to 0.941, proving the model's factorial validity (Table 3).

The suggestion that the data can be treated as essentially unidimensional was confirmed by the UniCo (0.989), ECV (0.918), and MIREAL (0.196) indices, in accordance with the recommendations of Ferrando & Lorenzo-Seva (2018).

The exploratory factor model presented a very good fit (robust mean and variance-adjusted $\chi^2(54) = 72.411$; $p = 0.048$), with CFI = 0.996 and TLI = 0.995, both exceeding the desirable threshold of 0.95. The RMSEA was 0.057 ($CI_{95\%}$ [0.025, 0.069]) and the RMSR was 0.062 ($CI_{95\%}$ [0.054, 0.065]), both within the criteria established for a well-fitted factor model.

The analysis of GRM parameters for the 12 scale items, evaluated on a 6-point Likert scale (0 to 5), demonstrated robust psychometric properties (Table 4). Regarding item discrimination (a), values ranged from 0.670 (item 6: "Removal of inhibitors") to 2.778 (item 12: "Capacity Building"). High discrimination ($a > 2.0$) was observed for the items 11 ("Innovation management") and 12 ("Capacity building"), demonstrating superior ability to differentiate between latent trait levels. Although item 6 ("Removal of inhibitors") presented lower discrimination value ($a < 0.75$), most scale items exceeded the 1.0 threshold, thereby contributing to the construct's overall accuracy.

Regarding the difficulty parameter *b*, five ordered thresholds *(b1-b5*) were extracted for all items, thereby confirming the monotonous nature of the response categories (De Ayala, 2022). The thresholds covered a wide range of the latent trait, from -2.314 to +3.417. More specifically, a considerable proportion of *b1* values (transition between categories 0 and 1) were below -1.5, indicating that the lowest category is selected only at low levels of maturity. On the other hand, the *b5* values (the transition between categories 4 and 5) were remarkably high, particularly for item 6 (*b5* = 3.417), suggesting that the maximum score represents an advanced stage of the construct. This distribution confirms the scale's effectiveness in capturing the entire maturity continuum.

Finally, the factor model of the SCIROCCO tool demonstrated high reliability, exceeding the desirable Composite Reliability (CR) threshold of 0.70 recommended by Hair et al. (2019), specifically 0.94.

**Table 3**: Factor loadings and communalities.

| Items | Factor Loadings | Communality ($h^2$) |
|---|---|---|
| 1. Readiness to change | **0.822** | **0.676** |
| 2. Structure & Governance | **0.787** | **0.619** |
| 3. Digital Infrastructure | **0.738** | **0.544** |
| 4. Process Coordination | **0.770** | **0.593** |
| 5. Funding | 0.678 | 0.459 |
| 6. Removal of Inhibitors | 0.557 | 0.310 |
| 7. Population Approach | 0.651 | 0.423 |
| 8. Citizen Empowerment | **0.781** | **0.610** |
| 9. Evaluation Methods | **0.759** | **0.575** |
| 10. Breadth of Ambition | **0.794** | **0.631** |
| 11. Innovation Management | **0.908** | **0.825** |
| 12. Capacity Building | **0.941** | **0.885** |

*Note*: N = 105. The extraction method was *Robust Diagonally Weighted Least Squares* (RDWLS) with an oblique (*Robust Promin*) rotation. Factor loadings above 0.70 are in bold. Communalities above 0.50 are in bold.

**Table 4**: Discrimination and threshold parameters for SCIROCCO tool.

| | **Discrimination parameter** | **Difficulty parameters for each threshold** | | | | |
|---|---|---|---|---|---|---|
| | **a** | **b1** | **b2** | **b3** | **b4** | **b5** |
| 1. Readiness to change | 1.444 | -2.314 | 0.160 | 0.655 | 1.465 | 2.157 |
| 2. Structure & Governance | 1.275 | -1.412 | -0.046 | 1.159 | 1.412 | 2.007 |
| 3. Digital Infrastructure | 1.093 | -1.567 | 0.211 | 0.692 | 1.008 | 2.261 |
| 4. Process Coordination | 1.207 | -1.776 | -0.628 | 0.594 | 1.630 | 2.693 |
| 5. Funding | 0.921 | -0.410 | 0.597 | 1.193 | 1.639 | 2.463 |
| 6. Removal of Inhibitors | 0.670 | -1.336 | 1.336 | 1.638 | 2.352 | 3.417 |
| 7. Population Approach | 0.857 | -2.200 | -0.958 | 0.466 | 1.242 | 2.200 |
| 8. Citizen Empowerment | 1.251 | -1.480 | 0.388 | 0.993 | 1.542 | 2.655 |
| 9. Evaluation Methods | 1.164 | -0.673 | 0.110 | 0.709 | 1.250 | 2.338 |
| 10. Breadth of Ambition | 1.308 | -0.976 | -0.575 | 0.288 | 0.897 | 1.580 |
| 11. Innovation Management | 2.168 | -1.506 | -0.066 | 0.593 | 1.273 | 1.837 |
| 12. Capacity Building | 2.778 | -1.135 | 0.166 | 0.572 | 1.229 | 1.521 |

## Discussion

This study aimed to adapt the SCIROCCO tool for the Portuguese context, resulting in the instrument: “Maturidade das Unidades Locais de Saúde para a prestação de Cuidados Integrados” (Maturity of LHUs to provide integrated care) (Additional File 1). As far as we know, this is the first scientifically validated tool for the Portuguese context, specifically designed to assess LHUs’ maturity in providing integrated care. The reported outcomes confirm the scale’s unidimensionality and high reliability and structural validity. In addition, this is the first study to apply the combination of IRT and EFA to investigate the measurement properties of SCIROCCO, which shows that the tool is robust for assessing the maturity of the health system for integrated care.

The translation and cultural adaptation process followed a rigorous and systemic method, to ensure content equivalence and cultural context relevance (Cruchinho et al., 2024), which was reinforced by pre-testing and a final review by the original author. Although our sample size ($n$ = 105) yielded an observation-to-variable ratio of 8.75:1, slightly below the frequently suggested 10:1 threshold (Zheng et al., 2010), the factor structure remained unchanged. In fact, this sample size provides a more robust subject-to-variable ratio than the original validation study of the SCIROCCO tool, which utilized a 5.5:1 ratio ($n$ = 67; Grooten et al., 2019). As noted in other psychometric studies, lower ratios do not necessarily compromise factor stability when the underlying tool structure is highly stable and has high communalities (Morgado et al., 2017).

Inter-item correlations in the polychoric matrix (0.288–0.887) suggested a high degree of scale’s homogeneity without redundancy (Streiner et al., 2014). Despite presenting the lowest correlation (0.288), the 6th dimension (“Removal of inhibitors”) was retained based on the sample’s robust adequacy (KMO > 0.80; Bartlett’s test $p$ < 0.001) and high factor loadings (all greater than 0.55) (Marôco, 2024). The EFA confirmed a single-factor solution above the 50th percentile of the factor analyses examined by Peterson (2000). This result suggests greater explanatory power than the 55.57% reported in the original instrument validation (Grooten et al., 2019), supported by excellent fit indices (RMSR = 0.062; RMSEA = 0.057; CFI = 0.996; TLI = 0.995) (Lorenzo-Seva & Ferrando, 2021b).

The SCIROCCO Portuguese version presented robust internal consistency (CR > 0.90), thereby suggesting that the internal consistency already reported by the authors of the tool’s original version was maintained.

The results from the IRT analysis showed that the SCIROCCO tool’s factor structure is robust. All items demonstrated strong discrimination power and covered the entire range of response categories, thus ensuring adequate accuracy across all the maturity levels. The monotonic increase in difficulty parameters suggests that the provided response options adequately captured the respondents’ latent trait. As expected, individuals working in LHUs with lower maturity chose lower Likert scale levels, while those in LHUs with higher maturity consistently chose higher Likert scale levels. In sum, the SCIROCCO tool can provide stable, adequate, and discriminative scores that allow effective distinction among LHUs at different maturity levels.

- Practical implications for healthcare

Despite the growing complexity of healthcare systems, maturity matrices remain underused by managers and policymakers. This study fills a crucial gap by providing a culturally appropriate, scientifically validated tool that establish a benchmark for the Portuguese context. Healthcare managers can use this tool as a roadmap to identify weaknesses and design tailored strategic actions and policies to enhance the quality of care. For

policymakers, it offers the opportunity to integrate outcomes into contracting processes and monitor the evolution of LHUs, using performance indicators based on maturity.

- Study limitations and future research

This study presents several limitations that should be acknowledged. Firstly, the use of convenience sampling, namely from a single LHU (Litoral Alentejano), is a limiting factor in terms of generalizing the results to a national level (the whole country). Future studies should include a more heterogeneous population across different geographical regions. Secondly, although the factorial model analyzed via EFA has shown the robustness of the SCIROCCO tool, using larger samples and achieving a ratio greater than 10:1 would further improve the accuracy of the model parameter estimates. Third, the tool's self-report nature may introduce social desirability bias despite the study's confidential design. Finally, future research should investigate measurement invariance across professional categories and geographical regions to ensure the tool's stability regardless of the participants' backgrounds.

## Conclusions

The Portuguese version of the SCIROCCO tool is a valid and reliable instrument for assessing LHU maturity. Strong psychometric properties of the SCIROCCO tool were reported, more specifically, high internal consistency and sound construct validity, strengthening its use in the Portuguese healthcare context. By providing a robust metric for integrated healthcare implementation, this tool could enable healthcare managers and policymakers to accurately monitor the development of LHUs and tailor evidence-based strategic actions and policies to support their continued improvement.

## Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki. The study protocol was approved by the Ethics Committee of the National School of Public Health of the NOVA University of Lisbon (CE-ENSP) on September 23rd, 2024 (Approval No. 05/2024). Informed consent was obtained from all the participants.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

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

Funding

Not applicable.

**Clinical Trial Number**

Not applicable.

Authors' contributions

Maria Carvalho contributed to all aspects of the study, including conceptualization, methodology, data analysis, and manuscript writing. Henrique Oliveira contributed to the methodology, results and discussion sections and performed the final revision of the manuscript. Cláudia Almeida contributed to the study design, supervised the research process, and participated in manuscript revision. João Santos collaborated in the translation process and was responsible for the linguistic refinement of the English manuscript. Andre Pavlickova reviewed the manuscript. Rui Santana contributed to the study design, provided mentorship, and critically revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

# References

1. Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. In *Frontiers in Public Health* (Vol. 6). Frontiers Media S.A. https://doi.org/10.3389/fpubh.2018.00149
2. Brown, T. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). The Guilford Press. www.guilford.com/MSS
3. Cruchinho, P., López-Franco, M. D., Capelas, M. L., Almeida, S., Bennett, P. M., da Silva, M. M., Teixeira, G., Nunes, E., Lucas, P., & Gaspar, F. (2024). Translation, Cross-Cultural Adaptation, and Validation of Measurement Instruments: A Practical Guideline for Novice Researchers. *Journal of Multidisciplinary Healthcare*, *17*, 2701–2728. https://doi.org/10.2147/JMDH.S419714
4. De Ayala, R. J. (2022). *The Theory and Practice of Item Response Theory* (2nd ed.). The Guilford Press.
5. Expert Group on Health Systems Performance Assessment. (2017). *Blocks: Tools and methodologies to assess integrated care in Europe*. https://doi.org/10.2875/69305
6. Ferrando, P. J., & Lorenzo-Seva, U. (2014). Exploratory Item Factor Analysis: Some additional considerations. *Anales de Psicología*, *30*(3), 1170–1175.
7. Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the Quality and Appropriateness of Factor Solutions and Factor Score Estimates in Exploratory Item Factor Analysis. *Educational and Psychological Measurement*, *78*(5), 762–780. https://doi.org/10.1177/0013164417719308
8. Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor Analysis with Ordinal Indicators: A Monte Carlo Study Comparing DWLS and ULS Estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 625–641. https://doi.org/10.1080/10705510903203573
9. Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research* , *18*.
10. Gaskin, C. J., & Happell, B. (2014). On exploratory factor analysis: A review of recent evidence, an assessment of current practice, and recommendations for future use. In *International Journal of Nursing Studies* (Vol. 51,

Number 3, pp. 511–521). https://doi.org/10.1016/j.ijnurstu.2013.10.005

11. Goiana-da-Silva, F., Sá, J., Cabral, M., Guedes, R., Vasconcelos, R., Sarmento, J., Morais Nunes, A., Moreira, R., Miraldo, M., Ashrafian, H., Darzi, A., & Araújo, F. (2024). The Portuguese NHS 2024 reform: transformation through vertical integration. *Frontiers in Public Health*, *12*. https://doi.org/10.3389/fpubh.2024.1389057
12. Grooten, L., Borgermans, L., & Vrijhoef, H. J. M. (2018). An instrument to measure maturity of integrated care: A first validation study. *International Journal of Integrated Care*, *18*(1). https://doi.org/10.5334/ijic.3063
13. Grooten, L., Vrijhoef, H. J. M., Calciolari, S., Ortiz, L. G. G., Janečková, M., Minkman, M. M. N., & Devroey, D. (2019). Assessing the maturity of the healthcare system for integrated care: Testing measurement properties of the SCIROCCO tool. *BMC Medical Research Methodology*, *19*(1). https://doi.org/10.1186/s12874-019-0704-1
14. Hair, J. F., L.D.S. Gabriel, M., da Silva, D., & Braga Junior, S. (2019). Development and validation of attitudes measurement scales: fundamental and practical aspects. *RAUSP Management Journal*, *54*(4), 490–507. https://doi.org/10.1108/RAUSP-05-2019-0098
15. *Health at a Glance: Europe 2024*. (2024). OECD. https://doi.org/10.1787/b3704e14-en
16. Henderson, D., Klonowska, K., Pavlickova, A., & Alexandru, C. (2022). *Capacity Building for Integrated Care: from the Maturity Assessment to Improvement Planning*.
17. Henderson, D., & Pavlickova, A. (2022). *D2.5 Layman Version of the Final Report WP 2-Dissemination*.
18. Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.
19. Hughes, G., Shaw, S. E., & Greenhalgh, T. (2022). Why doesn't integrated care work? Using Strong Structuration Theory to explain the limitations of an English case. *Sociology of Health and Illness*, *44*(1), 113–129. https://doi.org/10.1111/1467-9566.13398
20. Izquierdo, I., Olea, J., & Abad, F. J. (2014). El análisis factorial exploratorio en estudios de validación: Usos y recomendaciones. *Psicothema*, *26*(3), 395–400. https://doi.org/10.7334/psicothema2013.349
21. Lorenzo-Seva, U., & Ferrando, P. J. (2019). Robust Promin: A method for diagonally weighted factor rotation. *Liberabit: Revista Peruana de Psicología*, *25*(1), 99–106. https://doi.org/10.24265/liberabit.2019.v25n1.08
22. Lorenzo-Seva, U., & Ferrando, P. J. (2021a). MSA: The Forgotten Index for Identifying Inappropriate Items Before Computing Exploratory Item Factor Analysis. *Methodology*, *17*(4), 296–306. https://doi.org/10.5964/meth.7185
23. Lorenzo-Seva, U., & Ferrando, P. J. (2021b). Not Positive Definite Correlation Matrices in Exploratory Item Factor Analysis: Causes, Consequences and a Proposed Solution. *Structural Equation Modeling*, *28*(1), 138–147. https://doi.org/10.1080/10705511.2020.1735393
24. Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519530.
25. Marôco, J. (2024). Factor Analysis of Ordinal Items: Old Questions, Modern Solutions? *Stats*, *7*(3), 984–1001. https://doi.org/10.3390/stats7030060
26. Mingolla, S. (2020). *Assessing the digital health transformation with the SCIROCCO Exchange tool in Puglia Region*.
27. Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S., & Ferreira, M. E. C. (2017). Scale development: Ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexao e Critica*, *30*(1), 1–20. https://doi.org/10.1186/s41155-016-0057-1

28. Nunes, A. (2024). Vertical Integration in Healthcare and Patient Satisfaction: An Exploratory Analysis of Portuguese Reforms. *Sustainability (Switzerland)*, *16*(3). https://doi.org/10.3390/su16031078
29. Peterson, R. A. (2000). A Meta-Analysis of Variance Accounted for and Factor Loadings in Exploratory Factor Analysis. In Marketing Letters (Vol. 11, Number 3). Kluwer Academic Publishers.
30. Peytremann-Bridevaux, I., Schusselé Filliettaz, S., Berchtold, P., Grossglauser, M., Pavlickova, A., & Gilles, I. (2021). Healthcare system maturity for integrated care: Results of a Swiss nationwide survey using the SCIROCCO tool. *BMJ Open*, *11*(2). https://doi.org/10.1136/bmjopen-2020-041956
31. Read, D. M. Y., Dalton, H., Booth, A., Goodwin, N., Hendry, A., & Perkins, D. (2019). Using the project INTEGRATE framework in practice in central coast, Australia. *International Journal of Integrated Care*, *19*(2). https://doi.org/10.5334/ijic.4624
32. Streiner, D. L., Norman, G. R., & Cairney, J. (2014). *Health Measurement Scales: A practical guide to their development and use* (Oxford, Ed.; 5nd ed.).
33. Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, *16*(2), 209–220. https://doi.org/10.1037/a0023353
34. Valentini, F., & Damásio, B. F. (2016). Average Variance Extracted and Composite Reliability: Reliability Coefficients. *Psicologia: Teoria e Pesquisa*, *32*(2), 1–7. https://doi.org/10.1590/0102-3772e322225
35. Wang, S., Marshall, L., Zabala, I., Marti, T., & Evan, S. (2022). *CAPACITY-BUILDING SUPPORT FOR INTEGRATED CARE: IMPROVEMENT PLANNING*.
36. Watkins, M. W. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/10.1177/0095798418771807
37. WHO. (2018). *Integrating health services: Technical series on primary care*.
38. Yang, F. M., & Kao, S. T. (2014). *Item response theory for measurement validity*. https://doi.org/10.3969/j.issn.1002-0829.2014.03.010
39. Zheng, J., You, L. M., Lou, T. Q., Chen, N. C., Lai, D. Y., Liang, Y. Y., Li, Y. N., Gu, Y. M., Lv, S. F., & Zhai, C. Q. (2010). Development and psychometric evaluation of the Dialysis patient-perceived Exercise Benefits and Barriers Scale. *International Journal of Nursing Studies*, *47*(2), 166–180. https://doi.org/10.1016/j.ijnurstu.2009.05.023
40. Zumbo, B D, Zumbo, Bruno D, Gadermann, A. M., & Zeisser, C. (2007). Ordinal Versions of Coefficients Alpha and Theta For Likert Rating Scales. In *Journal of Modern Applied Statistical Methods* (Vol. 6, Number 1). http://educ.ubc.ca/faculty/zumbo/cv.htmhttp://educ.ubc.ca/faculty/zumbo/papers/ordinal_alpha_reprint.pdf

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1.docx