Author Version

A Scale for Measuring Healthcare Service Quality Incorporating Patient-Centred Care and Using a Psychometric Analytics Framework

Jorge Iván Pérez Rave (IDINNOV Research Group, IDINNOV S.A.S, Medellín, Colombia, investigacion@idinnov.com),

Génesis Angélica Sánchez Figueroa (gsanchez@idinnov.com, IDINNOV Research Group, IDINNOV S.A.S, Medellín, Colombia),

Favián González Echavarría (favian.gonzalez@udea.edu.co, Department of Industrial Engineering, Universidad de Antioquia, Medellin, Colombia)

Abstract

Recent developments in healthcare contexts increasingly emphasize patient-centred approaches to service quality measures; however, few studies consider this dimension explicitly. The present study develops and psychometrically validates a scale of healthcare service quality explicitly incorporating a patient-centred care dimension from a communicational perspective. The paper also enriches the traditional content of service quality by including equity items and presents the underlying structure of service quality in an emerging country. The final sample consisted of 869 healthcare users (complete cases in the service quality items derived from 917 surveys received) from Colombia. We used a psychometric analytics framework comprising seven processes incorporating exploratory factor analysis, structural equation modelling, and machine learning methods to examine construct plausibility, reliability, construct validity, equity, and criterion/predictive validity (e.g. explaining/predicting subjective well-being and behavioural intentions). The final scale consists of 17 items and satisfies all psychometric properties. Its validation allows for the discovery and psychometrical confirmation of two essential dimensions: patient-centred communication (eight items) and process quality (nine items). Both dimensions reveal users' relevant needs and complement previous studies that have focused on process aspects of healthcare service quality. We illustrate three practical uses of the scale: the possibility for diagnoses; hypothesis contrast based on confidence intervals; and estimation of the capacity of the service to satisfy specifications.

Deposit licences

Published version

A Scale for Measuring Healthcare Service Quality Incorporating Patient-Centred Care and Using a Psychometric Analytics Framework

Jorge Iván Pérez Rave (IDINNOV Research Group, IDINNOV S.A.S, Medellín, Colombia, investigacion@idinnov.com), *Génesis Angélica Sánchez Figueroa* (gsanchez@idinnov.com, IDINNOV Research Group, IDINNOV S.A.S, Medellín, Colombia), *Favián González Echavarría* (favian.gonzalez@udea.edu.co, Department of Industrial Engineering, Universidad de Antioquia, Medellin, Colombia)

Abstract

Purpose: Recent developments in healthcare contexts increasingly emphasize patient-centred approaches to service quality measures; however, few studies consider this dimension explicitly. The present study develops and psychometrically validates a scale of healthcare service quality explicitly incorporating a patient-centred care dimension from a communicational perspective. The paper also enriches the traditional content of service quality by including equity items and presents the underlying structure of service quality in an emerging country.

Design/methodology/approach: The final sample consisted of 869 healthcare users (complete cases in the service quality items derived from 917 surveys received) from Colombia. We used a psychometric analytics framework comprising seven processes incorporating exploratory factor analysis, structural equation modelling, and machine learning methods to examine construct plausibility, reliability, construct validity, equity, and criterion/predictive validity (e.g. explaining/predicting subjective well-being and behavioural intentions).

Findings: The final scale consists of 17 items and satisfies all psychometric properties. Its validation allows for the discovery and psychometrical confirmation of two essential dimensions: patient-centred communication (eight items) and process quality (nine items).

Originality: Both dimensions reveal users' relevant needs and complement previous studies that have focused on process aspects of healthcare service quality.

Practical implications: We illustrate three practical uses of the scale: the possibility for diagnoses; hypothesis contrast based on confidence intervals; and estimation of the capacity of the service to satisfy specifications.

Keywords: service quality, healthcare, constructs' mining, psychometric analysis, data mining, machine learning.

Introduction

The Institute of Medicine (IoM, 2001) defines healthcare service quality in terms of timeliness, effectiveness, efficiency, patient-centred care, equity, and security. The World Health Organization (WHO) states that healthcare services should be safe, effective, people-centred, and timely (cited by Upadhyai et al., 2019). These domain-specific perspectives have not been considered as often as the service quality model (SERVQUAL, as developed by

Parasuraman et al., 1988) in representations and measurements of service quality in healthcare settings from the user's viewpoint (Al-Damen, 2017; Dean, 1999; Purcărea et al., 2013; Setyawan et al., 2019).

While the SERVQUAL and IoM (2001) dimensions are globally applicable, they must be adapted to the healthcare context (Mathong et al., 2020). Also, SERVQUAL does not cover equity and patient-centred care sufficiently. Equity is essential in the healthcare context to prevent possible discrimination in the services offered. The empirical understanding of patient-centred care is often mistakenly assumed: first, it is common to assume that the mere use of users' viewpoints is a guarantee of patient-centred care. That is, the use of SERVQUAL or another instrument based on the patient's perspective is sufficient to cover any current challenge to their care. Second, the empathy dimension of SERVQUAL covers patient-centred care, but this is insufficient because current versions of SERVQUAL offer a general perspective focused on good treatment of patients (e.g. "XYZ's employees give patients personal attention," and "YZ has patients' best interests at heart"; Babakus and Mangold, 1992, p. 783).

According to the IoM (2001), patient-centred care is defined as being "respectful of and responsive to individual patient preferences, needs, and values, and ensuring that patient values guide all clinical decisions" (p. 3). Thus, patient-centred care should place more attention on domain-specific interactions with care professionals, for example, regarding the causes and treatment of a health issue (Alamo et al., 2002; IoM, 2001; Kitson et al., 2013). Islam and Muhamad (2021) have extended the concept of healthcare service quality by incorporating patient-centred communication. Additionally, recent studies (Behmane et al., 2021; Carter et al., 2021; Jeon et al., 2010; Lin et al., 2021; Makubalo et al., 2020; Skogö et al., 2019) argue that patient-centred care is needed to address current problems of poor user satisfaction, limited understanding of patient needs, lack of patient autonomy/engagement/proactivity in care and clinical services/solutions, and the dehumanization of medical care (Miles & Mezzich, 2011; Ornstein & Kay, 2021).

On the other hand, although recent developments in healthcare contexts increasingly emphasize patient-centred approaches to service quality measures (Islam and Muhamad, 2021; Tate et al., 2020), it is unclear how the dimensions of patient-centred care interact with standard dimensions of service quality. For example, Islam and Muhamad (2021) and Knox et al. (2015) did not examine or discuss the nature of these interrelations, even though patient-centred care has been established as a critical enabler of organizational outcomes (e.g., clinical effectiveness, safety, and resource usage; Australian Commission on Safety and Quality in Health Care, 2011; Doyle et al, 2013).

While measurements of healthcare service quality from the user's viewpoint have been refined, the results cannot be generalized to all countries. Studies have been largely confined to users of healthcare services in developed countries (e.g., Mayo Clinic Arizona, USA, Kennedy et al., 2019; community health centres in the Greater Bay Area of China, Liu et al.,

2021; public/private healthcare sectors in London, Owusu-Frimpong et al., 2010), so the underlying factorial structures based on specific populations in less-developed countries need to be further investigated. Further, the measurements employed need to be psychometrically validated. "Developing countries should try and develop their own models for measuring the quality of healthcare services" (Endeshaw, 2020, p. 106).

The present study considers the IoM (2001) dimensions to develop a measure of the quality of healthcare services based on the following particularities:

(a) Incorporating items about patient-centred care principally from a communicational perspective, covering aspects related to respectful treatment; helpful information and support; and participation (IoM, 2001). Newell and Jordan (2015) pointed out that "communication plays an integral role in service quality in all service professions including healthcare professions" (p. 76), while Perera and Dabney (2020) argued that "providing care that is patient-centred is an important objective in the modern healthcare industry" (p. 551).

(b) Discovering the primary needs (e.g., empirically confirmed dimensions under a reliable and valid underlying factorial structure) of healthcare users in an emerging economy (i.e., Colombia) using a large sample (917 surveys received; 869 complete cases in the service quality items). In this context, psychometrically supported studies on healthcare service quality from patient viewpoints are scarce, as stated by Tobón and Arias (2018), who used SERVQUAL and examined its psychometric properties with a sample of 59 users. Arboleda-Arango, Chernichovsky, and Esperato (2018) found that Colombian surveys have deviated from the international standards of patient satisfaction suggested by international standards (e.g., the WHO). Currently, the Colombian healthcare system faces several problems; these include low numbers of medical and nursing personnel (Organisation for Economic and Cooperative Development [OECD]/The World Bank, 2020); limited availability of medical appointments; poor accessibility and customer service (DANE, 2019); inequity (Ministry of Health and Social Protection, 2013); and generally inadequate service quality (Bonet et al., 2017). Moreover, this paper complements recent qualitative works on service quality in developing countries (e.g., interviews with 61 doctors and 40 inpatients from public/private hospitals in Rio de Janeiro, Guedes & Araujo, 2022) and quantitative studies which used adapted/based versions of SERVQUAL to the health sector (e.g., Mira et al., 1998) in developing countries (e.g., 279 users from an outpatient surgical service in Boyacá, Colombia, Numpaque-Pacabaque, Buitrago-Orjuela, & Pardo-Santamaría, 2019; 121 users in an emergency unit of a hospital in Valparaíso, Chile) by considering a patient-centred care approach, a large sample (917 users), and a psychometric analytics framework.

(c) Performing a comprehensive validation using a psychometrical data mining framework. Few studies about healthcare service quality from patients' viewpoints have examined the concurrent validity of scales considering subjective well-being and behavioural intention (e.g., to recommend the service). Patient well-being is a relevant goal of healthcare systems in all nations (Mody et al., 2020; Stewart et al., 1989). Additionally, although the equity of a

scale is essential to prevent discriminatory decisions driven by data (Pyburn et al., 2008), psychometric studies of service quality have rarely examined this aspect. The present study addresses these considerations by using subjective well-being and behavioural intention as latent variables to be explained/predicted from the service quality scale to be developed. Likewise, it examines equity considering demographic factors (i.e., gender, age, and socioeconomic stratum).

"Service quality has been a matter of concern for public and private healthcare institutions across the world" (Upadhyai et al., 2019, p. 102). The scale developed herein will help improve the diagnosis and care—patient-centred care in particular—and stimulate future studies based on structural relationships involving service quality.

Methods

Item Formulation

We formulated 22 items to represent healthcare service quality, nine of which referred to patient-centred care. They were based on Alamo, Moral, and de Torres (2002); Epstein et al. (2005); and Kitson et al. (2013). One example is "How often do healthcare professionals (doctors, nurses, specialists, psychologists, nutritionists, and other health professionals) encourage you to express your doubts about your health condition (pat.cent1)?" The response format was a 5-point Likert scale: 1 = *never*; 2 = *almost never*; 3 = *sometimes*; 4 = *almost always*; and 5 = *always*. This has been used in several previous studies (e.g., Cuervo et al., 2014 and Thomas et al., 2011). The additional 12 items were based on construct conceptualizations provided by the IoM (2001): Timeliness (3 items); Security (3); Equity (3); and Effectiveness (3). An additional item, relating to tangibles (1), was derived from SERVQUAL. Moreover, we considered user well-being (five items based on Stewart et al., 1989) and recommendation intention (two items based on Reichheld, 2003) to examine criterion validity.

Participants

The initial sample consisted of 917 external users of Colombian health services (e.g., outpatient consultations, drug supply, and diagnostic aids) provided by a large organization (covering more than 65,000 affiliated users in several departments/headquarters). They were recruited via email from March to April 2019 and asked to fill out the questionnaire under study. Considering service quality, well-being, and recommendation intention, the complete observations were 869, 847, and 814, respectively.

Psychometrical Analytics Framework

We used the seven processes of the MinerConstructo interactive methodological framework (observe, explore, confirm, explain, predict, apply, and report; Pérez-Rave, 2021). This framework combines data science capabilities (e.g., automated tasks; machine learning methods; and multiple interactive visualizations of the processes' outputs in real-time) and

psychometric analysis of latent variables in R by integrating resources of several R packages (e.g., "psych", Revelle & Revelle, 2015; "lavaan", Rosseel, 2012; and "semPlot", Epskamp, 2015) into the processes (e.g. see Pérez-Rave et al., 2022a; Pérez-Rave et al., 2022b).

Results and Discussion

This section is organized according to our analytics framework.

Observe

The respondents comprised: gender (female: 66%); age in years (27–39: 18%; 40–59: 57%; and 60 or older: 26%); postgraduate education (64%); and socioeconomic stratum (1 or 2: 19.8%; 3: 55%; 4 to 6: 26%). In Colombia, the stratum is an ordinal variable ranging from 1 (families in the most disadvantaged conditions) to 6 (most advantaged conditions); the government uses this classification to set rates for public services and health, among others (Pérez-Rave et al., 2019).

Table 1 presents a statistical summary of the 21 scale items.

Table 1. Statistical summary of the items (n = 869 complete observations)

	Min	Max	Mean	SD	Median	Q1	Q3	Kurtosis	Symmetry	n
pat.cent1	1	5	2.84	1.3	3	2	4	-1.034	0.072	917
pat.cent2	1	5	3.19	1.134	3	3	4	-0.564	-0.187	917
pat.cent3	1	5	2.689	1.236	3	2	4	-0.891	0.217	917
pat.cent4	1	5	2.888	1.18	3	2	4	-0.824	0.11	917
pat.cent5	1	5	2.427	1.273	2	1	3	-0.879	0.444	917
pat.cent6	1	5	2.833	1.162	3	2	4	-0.776	0.027	917
pat.cent7	1	5	3.205	1.169	3	2	4	-0.677	-0.222	917
pat.cent8	1	5	3.641	1.086	4	3	4	-0.131	-0.617	917
pat.cent9	1	5	2.974	1.207	3	2	4	-0.826	-0.125	917
tang	1	5	3.269	1.109	3	3	4	-0.549	-0.295	917
timely1	1	5	2.35	0.99	2	2	3	-0.354	0.396	907
timely2	1	5	3.266	1.057	3	3	4	-0.332	-0.32	907
timely3	1	5	2.922	1.101	3	2	4	-0.665	-0.113	907
secu1	1	5	3.265	1.057	3	3	4	-0.39	-0.301	901
secu2	1	5	3.405	1.07	4	3	4	-0.237	-0.499	902
secu3	1	5	3.383	1.047	3	3	4	-0.222	-0.467	896
equity1	1	5	3.84	1.098	4	3	5	0.233	-0.879	883
equity2	1	5	3.411	1.164	4	3	4	-0.502	-0.45	883
equity3	1	5	2.857	1.145	3	2	4	-0.697	0.018	883
effectiv1	1	5	2.933	1.11	3	2	4	-0.668	-0.107	876
effectiv2	1	5	2.232	0.999	2	1	3	-0.07	0.602	876
effectiv3	1	5	2.473	0.997	3	2	3	-0.313	0.255	876

The items presented mean values between 2.23 (effectiv2) and 3.84 (equity1), with a global mean value of 3.01. Twelve of the 21 items had means less than 3.0. This suggested that the respondents did not have a favourable view of service quality. The absolute values of univariate skewness and kurtosis of the items were less than 2.0, discounting extreme deviations from the normal distribution. Table 2 presents a Pearson correlation matrix between pairs of items.

Table 2. Pearson's correlation matrix of the items (n: 869 complete cases)

No	Items	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	pat.cent1		0.67	0.68	0.62	0.55	0.6	0.54	0.49	0.55	0.27	0.34	0.29	0.36	0.38	0.27	0.36	0.28	0.27	0.26	0.35	0.3	0.37
2	pat.cent2	0.67		0.7	0.72	0.57	0.68	0.68	0.67	0.66	0.36	0.43	0.38	0.4	0.5	0.42	0.49	0.43	0.41	0.36	0.49	0.35	0.45
3	pat.cent3	0.68	0.7		0.69	0.65	0.6	0.57	0.5	0.6	0.3	0.37	0.29	0.36	0.41	0.27	0.35	0.3	0.33	0.29	0.39	0.3	0.4
4	pat.cent4	0.62	0.72	0.69		0.66	0.68	0.64	0.62	0.65	0.33	0.4	0.36	0.41	0.5	0.35	0.44	0.39	0.38	0.34	0.45	0.34	0.46
5	pat.cent5	0.55	0.57	0.65	0.66		0.62	0.55	0.44	0.61	0.27	0.32	0.28	0.33	0.38	0.29	0.33	0.26	0.3	0.27	0.35	0.28	0.39
6	pat.cent6	0.6	0.68	0.6	0.68	0.62		0.74	0.64	0.72	0.38	0.4	0.35	0.42	0.55	0.39	0.47	0.41	0.41	0.37	0.49	0.38	0.48
7	pat.cent7	0.54	0.68	0.57	0.64	0.55	0.74		0.69	0.72	0.35	0.44	0.41	0.45	0.55	0.44	0.5	0.46	0.41	0.39	0.51	0.37	0.46
8	pat.cent8	0.49	0.67	0.5	0.62	0.44	0.64	0.69		0.69	0.39	0.4	0.41	0.43	0.57	0.45	0.51	0.5	0.46	0.4	0.52	0.35	0.44
9	pat.cent9	0.55	0.66	0.6	0.65	0.61	0.72	0.72	0.69		0.34	0.41	0.42	0.44	0.55	0.37	0.45	0.41	0.4	0.36	0.47	0.35	0.44
10	tang	0.27	0.36	0.3	0.33	0.27	0.38	0.35	0.39	0.34		0.39	0.37	0.39	0.49	0.66	0.47	0.4	0.41	0.54	0.44	0.39	0.46
11	timely1	0.34	0.43	0.37	0.4	0.32	0.4	0.44	0.4	0.41	0.39		0.52	0.54	0.48	0.47	0.47	0.43	0.47	0.47	0.54	0.45	0.57
12	timely2	0.29	0.38	0.29	0.36	0.28	0.35	0.41	0.41	0.42	0.37	0.52		0.75	0.56	0.51	0.53	0.54	0.53	0.49	0.53	0.43	0.52
13	timely3	0.36	0.4	0.36	0.41	0.33	0.42	0.45	0.43	0.44	0.39	0.54	0.75		0.62	0.53	0.54	0.54	0.53	0.51	0.54	0.42	0.51
14	secu1	0.38	0.5	0.41	0.5	0.38	0.55	0.55	0.57	0.55	0.49	0.48	0.56	0.62		0.63	0.66	0.6	0.57	0.51	0.61	0.47	0.57
15	secu2	0.27	0.42	0.27	0.35	0.29	0.39	0.44	0.45	0.37	0.66	0.47	0.51	0.53	0.63		0.64	0.51	0.5	0.58	0.53	0.41	0.51
16	secu3	0.36	0.49	0.35	0.44	0.33	0.47	0.5	0.51	0.45	0.47	0.47	0.53	0.54	0.66	0.64		0.59	0.57	0.53	0.58	0.48	0.59
17	equity1	0.28	0.43	0.3	0.39	0.26	0.41	0.46	0.5	0.41	0.4	0.43	0.54	0.54	0.6	0.51	0.59		0.73	0.55	0.57	0.4	0.49
18	equity2	0.27	0.41	0.33	0.38	0.3	0.41	0.41	0.46	0.4	0.41	0.47	0.53	0.53	0.57	0.5	0.57	0.73		0.62	0.57	0.42	0.54
19	equity3	0.26	0.36	0.29	0.34	0.27	0.37	0.39	0.4	0.36	0.54	0.47	0.49	0.51	0.51	0.58	0.53	0.55	0.62		0.56	0.45	0.55
20	effectiv1	0.35	0.49	0.39	0.45	0.35	0.49	0.51	0.52	0.47	0.44	0.54	0.53	0.54	0.61	0.53	0.58	0.57	0.57	0.56		0.51	0.62
21	effectiv2	0.3	0.35	0.3	0.34	0.28	0.38	0.37	0.35	0.35	0.39	0.45	0.43	0.42	0.47	0.41	0.48	0.4	0.42	0.45	0.51		0.72
22	effectiv3	0.37	0.45	0.4	0.46	0.39	0.48	0.46	0.44	0.44	0.46	0.57	0.52	0.51	0.57	0.51	0.59	0.49	0.54	0.55	0.62	0.72	

Table 2 shows that all correlations were positive (from 0.26 to 0.75), had a mean value of 0.472, and the first and third quartiles were 0.39 and 0.55, respectively. Based on Sloan and Angell's (2015) criterion, 88% of the correlations were moderate (between 0.3 and 0.7) and 3% were high (greater than 0.7). These results inferred possible underlying patterns that would be discovered in subsequent processes using a multivariate perspective.

Explore

The result of a Bartlett test provided a p value < 0.01 (chi-square: 13,822.719, df: 231); hence, the null hypothesis concerning the correlation matrix corresponding to the identity matrix was rejected with a confidence level of 99%. The Kaiser−Meyer−Olkin (KMO) test provided an overall measure of sampling adequacy (MSA) of 0.955, with the MSA for each item ranging from 0.929: "effectiv2" to 0.978: "timely1." Both tests justified the need to perform exploratory factor analysis (EFA) to discover potential multivariate association patterns.

The Kaiser criterion (eigenvalues greater than 1) and Horn parallel analysis (Horn, 1965; corrected eigenvalues greater than 1) suggested that two factors should be retained (Table 3).

Table 3. Analysis of the number of latent factors to retain (n: 869 complete cases)

No	Corrected eigenvalues[a]	Eigenvalues	Eigenvalues from random samples	Bias
1	10.699	10.993	1.294	0.294
2	2.238	2.484	1.246	0.246
3	0.721	0.931	1.209	0.209

[a] It is based on Horn' parallel analysis (eigenvalues - bias)

Table 4 displays the structure assuming two latent factors (including Cronbach's α). All items were conserved because none presented loadings less than 0.45 or greater than 0.45 for more than one factor.

Table 4. Exploratory factorial structure and Cronbach's α (n: 869 complete cases)

Items	PQ	PCC
pat.cent1		0.722
pat.cent2		0.783
pat.cent3		0.781
pat.cent4		0.793
pat.cent5		0.714
pat.cent6		0.764
pat.cent7		0.713
pat.cent8		0.627
pat.cent9		0.747
tang	0.577	
timely1	0.593	

timely2	0.696	
timely3	0.687	
secu1	0.700	
secu2	0.724	
secu3	0.706	
equity1	0.704	
equity2	0.718	
equity3	0.718	
effectiv1	0.690	
effectiv2	0.579	
effectiv3	0.684	
Explained Variance	30.90%	26.70%
Cronbach's α	0.935	0.938

PCC: patient-centred communication; PQ: process quality

Table 4 shows that the "PCC" factor consisted of the nine items concerning patient-centred care (based on a communicational perspective); hence, it was named "Patient-centred communication". The other 13 items cover aspects regarding a factor named "Process quality" (PQ), that is, a process characterized as timely, effective, equitable, safe ("secu"), and tangibly comfortable ("tang"). The latter factor is a reasonable representation, at the item level, of the most traditional dimensions of service quality (excepting items of equity) that are focused on processes other than the user experience. Indeed, as Coulthard (2004) has noted, "SERVQUAL has been criticised for its focus on the processes of service delivery" (p. 3). Thus, the underlying structure extends the common service quality perspective by including items of equity to process quality, and patient-centred care items from a communicational perspective are incorporated into an additional dimension (i.e., patient-centred communication [PCC]). The scale is parsimonious, maintaining only two general dimensions of healthcare service quality. At the same time, it is representative because each dimension comprises nine and 13 items, respectively, so the empirical manifestations of the constructs is clearer.

Confirm

Construct Plausibility, Construct Validity, and Composite Reliability

Table 5 presents the results of confirmatory factor analysis (CFA). Three structures were examined—single (M1), the two latent factors (M2), and M2 refined (M2r)—after five items ("effectiv3," "pat.cent3," "secu2," "timely2," and "equity2") were removed (there were considerable error correlations with other items; visualizing modification indexes). The refined model (M2r, 17 items) preserved the "tang" item and all facets of M2, with two items per facet (i.e., "timely1," and "timely3"). Table 5 includes three additional sample sizes (500, 250, and 150 randomly chosen observations) because the original sample size (869 complete

observations) was larger than traditional studies where the rule of thumb concerning goodness-of-fit measures was established; moreover, chi-square and its complementary measures are highly sensitive to sample size.

Table 5. Goodness-of-fit measurements for the models

Models	Items	n.obs	p-value	Chi-sq	df	Chisq/df	RMSEA	SRMR	CFI	TLI
M1 (single factor)	22	869	0.0	4041.267	209	19.336	0.145	0.098	0.721	0.691
M2 (two factors)	22	869	0.0	1854.424	208	8.916	0.095	0.053	0.88	0.867
M2r (two factors)	17	869	0.0	609.582	118	5.166	0.069	0.039	0.947	0.939
		500	0.0	355.114		3.009	0.069	0.043	0.949	0.941
		250	0.0	291.414		2.47	0.082	0.054	0.926	0.914
		150	0.0	220.727		1.871	0.08	0.06	0.921	0.909

As Table 5 shows, the single-factor model (M1) was discarded in light of the goodness-of-fit measurements (chi-square/*df*: 19.336; root mean square error of approximation [RMSEA]: 0.145; standardized root mean squared residual [SRMR]: 0.098; comparative fit index [CFI]: 0.721; and Tucker−Lewis index [TLI]: 0.691). Likewise, the M2r model presented a better fit than the M2 (e.g., CFI: 0.88 and TLI: 0.867). Thus, the M2r measurement model comprising the two refined factors (17 items) performed better in terms of goodness of fit (e.g., considering the original sample and three subsamples: 150, 250 and 500 observations). In the original sample RMSEA: 0.069; SRMR: 0.039; CFI: 0.947; TLI: 0.939. Likewise, the chi-square/*df* was 5.166, which was greater than the 3.0 ordinarily used as a rule of thumb, but this was due to the large sample size; when the sample sizes were 500, 250, and 150, the indicator was 3.0, 2.47, and 1.87, respectively. This behaviour demonstrates the high sensitivity of the chi-square to sample size. Hence, authors such as Credé and Harms (2015) have suggested caution when rules of thumb in structured equation modelling, arguing that the better models are those with a higher CFI and TLI and lower SRMR, RMSEA, and chi-squares/*df*s.
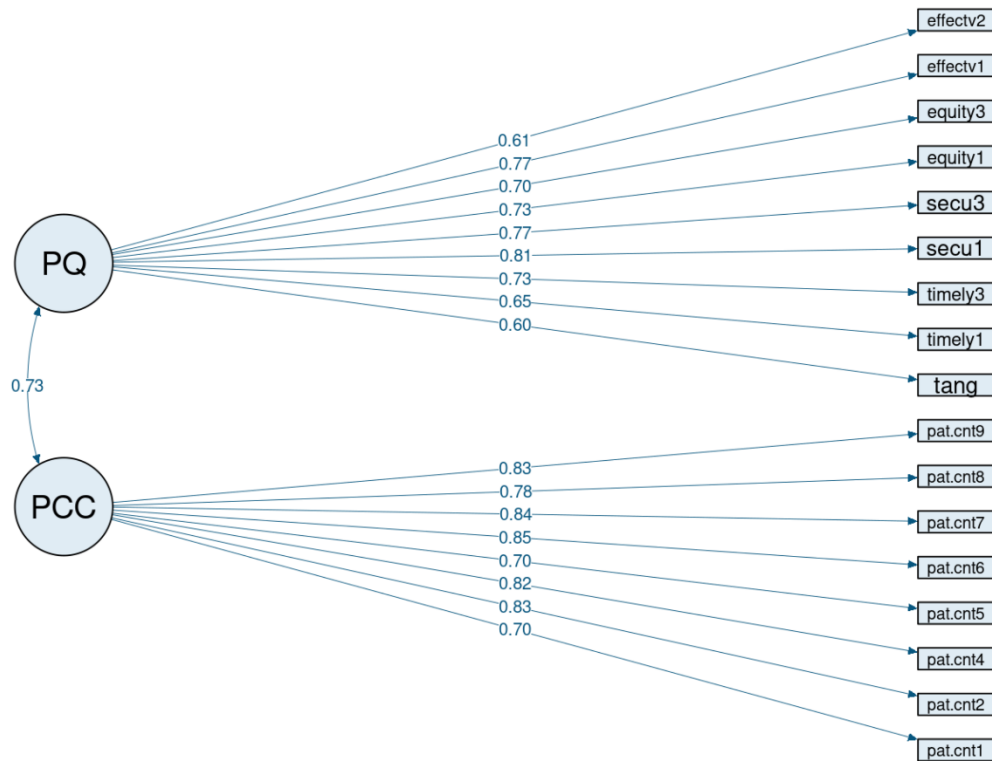
Figure 1. Plot of the application of confirmatory factor analysis to the healthcare service quality scale. *Note(s).* PQ: process quality; PCC: patient-centred communication

Figure 1 shows that none of the items in the study presented low loadings (of less than 0.5); 12 of 17 items presented high loadings (greater than 0.7), and the loadings of the other three items were moderate (0.61: "effectiv2," 0.65: "timely1," and 0.60: "tang"). All 17 items were statistically significant at least to 0.001 level.

The dimensions of the healthcare service quality scale presented average extracted variance greater than 0.5 (PCC: 0.633 and PQ: 0.507); a composite reliability greater than 0.7 (PCC: 0.932 and PQ: 0.902); and a percentile bootstrap confidence interval of 95% for the correlation between PCC and PQ (0.675; 0.771) that not included 1.0.

In light of the above evidence (i.e., comparing single-factor vs. two-factor models; confidence intervals of correlation between the dimensions; factorial loadings; AVE; and composite reliability) it concluded that the M2r model was plausible and presented good construct validity and reliability.

Construct Equity

This is a relevant property in psychometric latent variable analysis; however, it is not ordinarily examined in management research, personnel selection being the exception. Its

importance lies in the fact that most psychometric scales are affected by the diversity–validity dilemma (Martínez et al., 2006; Pérez-Rave et al., 2021), they are discriminatory in respect of personal factors such as age, gender, and ethnicity. In the present case, gender, age, socioeconomic stratum, and education level were considered for the equity analysis.
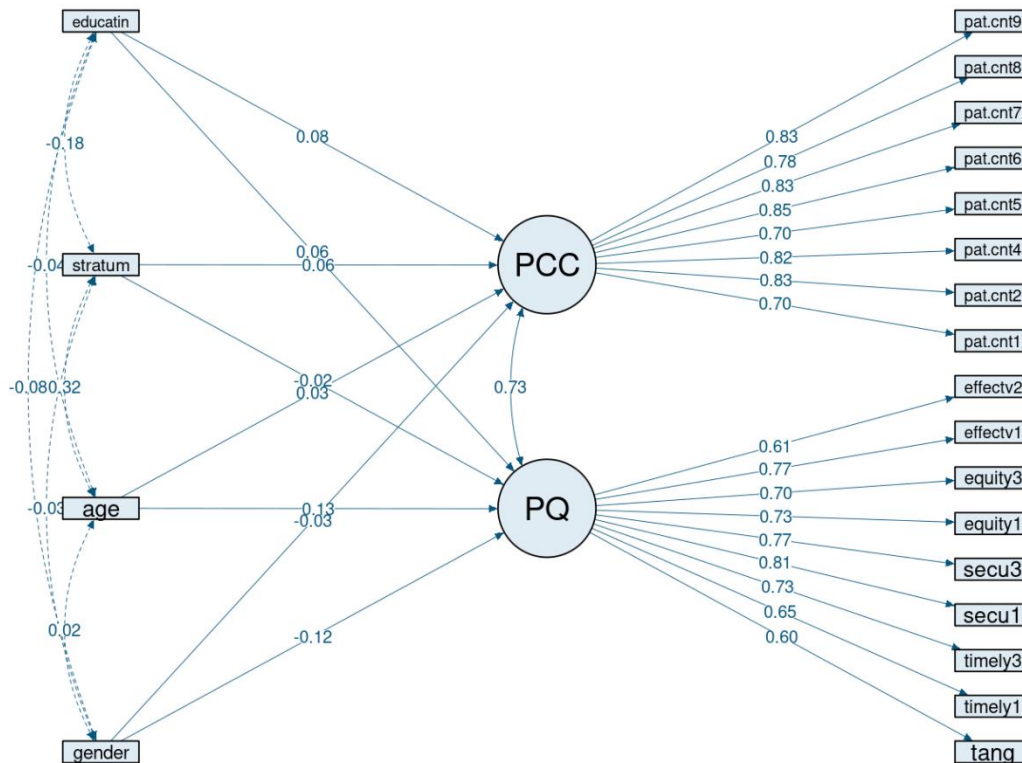
Figure 2. Confirmatory Factor Analysis with Segmentation Variables to Examine Construct Equity. *Note(s).* PQ: process quality; PCC: patient-centred communication.

The goodness-of-fit measures of the CFA model shown in Figure 2 ($n = 869$ complete observations) were as follows: chi-squared/df: 3.911; RMSEA: 0.058; SRMR: 0.035; CFI: 0.945; and TLI: 0.937. The bootstrap confidence intervals for the equity analysis are provided in Table 6.

Table 6. Results of scale equity (869 complete observations)

left.side	direction	right.side	estim	err.est	linfer	lsuper	estand	result
gender	→	PQ	-0.266	0.079	-0.426	-0.119	-0.124	Signif
age	→	PQ	0.204	0.058	0.093	0.318	0.13	Signif
stratum	→	PQ	-0.033	0.059	-0.148	0.084	-0.022	No.Signif
education	→	PQ	0.122	0.075	-0.021	0.269	0.058	No.Signif
gender	→	PCC	-0.054	0.078	-0.206	0.093	-0.026	No.Signif
age	→	PCC	0.05	0.058	-0.06	0.163	0.033	No.Signif

stratum	→	PCC	0.083	0.056	-0.028	0.197	0.055	No.Signif
education	→	PCC	0.175	0.075	0.029	0.321	0.084	Signif

PQ: process quality; PCC: patient-centred communication

Of the four sociodemographic factors, gender and age presented statistically significant relationships with PC, with confidence intervals of (-0.426; -0.119) and (0.093; 0.318), respectively. The females ($M = 3.1$; $SD = 0.781$) tended to perceive PQ more favourably than the males ($M = 2.9$; $SD = 0.847$); likewise, participants aged 27–39 years ($M = 2.856$; $SD = 0.764$) tended to evaluate PQ less favourably than the older participants ("40–59," $M = 3.04$ and $SD = 0.808$; and "60 or more," $M = 3.145$ and $SD = 0.823$). It appears that the older the participants were, the more flexible they became.

Level of education was a significant factor in the participants' evaluation of PCC (confidence interval 0.029; 0.321). Those who had received a postgraduate education tended to be more demanding ($M = 2.99$; $SD = 0.813$) than those without ($M = 3.104$; $SD = 0.798$).

These findings suggest the need to control for sociodemographic factors such as gender, age, and education level to prevent discrimination during decision-making processes driven by the evidence (e.g., allocation of budgets, prizes, letters regarding poor performance, sanctions, and incentives). For example, centres that attend primarily to older people may be more perceived more favourably by their patients.

Explain

Criterion Validity

We compared structural relationships to understand (a) user well-being (WB; five items), for example: "Please rate your ability in the last month to perform physical activities (e.g., sports, carrying groceries, climbing stairs, and walking)." This was rated from 1 (*very bad*) to 5 (*very good*); and (b) service recommendation intention (two items, one of which was "How likely are you to recommend the organization to other people?"). This was rated from 1 (*very unlikely*) to 5 (*very likely*). The items were based on Stewart et al. (1989) and the traditional net promoter score (NPS) measurement, respectively. The purpose of this analysis was to examine whether the empirical evidence supported PQ and PCC as relevant antecedents of the output variables (*n*: 814 complete observations).

We performed the analysis with and without control factors, testing for age, gender, postgraduate, and stratum. We aimed to gain an understanding of the role of healthcare service quality patterns in building subjective well-being and favourable behavioural intentions (i.e., recommending the service).

In both scenarios (with and without control factors), the structural model was plausible: (a) without control factors: chi-squared/*df*: 4.511; RMSEA: 0.066; SRMR: 0.04; CFI: 0.938; TLI: 0.93; and (b) with control factors on WB and recommendation intention ("Recom"):

chi-squared/*df*: 3.778; RMSEA: 0.058; SRMR: 0.044; CFI: 0.933; TLI: 0.926. Figure 3 shows the path diagram for the structural model without control factors.
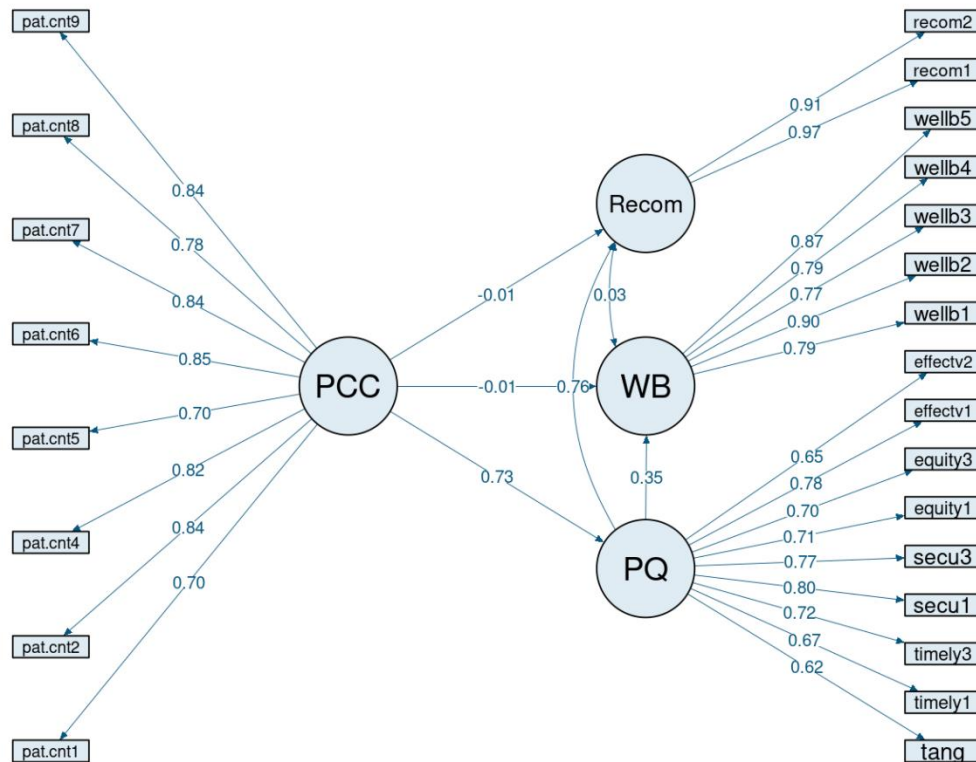

Figure 3. Structural model to examine criterion validity (without control factors). *Note(s)*. PQ: process quality; PCC: patient-centred communication; WB: well-being; Recom: recommendation intention.

We performed the formal contrast of the structural relationships based on bootstrapping with 5,000 replicas. Table 7 shows the percentile confidence intervals (at 95%) considering two scenarios (without and with control factors).

Table 7. Contrasting structural hypotheses to examine the criterion validity of the scale (*n*: 814 complete observations)

Left	Direction	Right	Estim	Std.Err	Lower.L	Upper.L	Stand	Result
Without control factors:								
PCC	→	PQ	1.07	0.079	0.924	1.231	0.731	Signif
PQ	→	WB	0.252	0.051	0.154	0.353	0.347	Signif
PCC	→	WB	-0.007	0.07	-0.145	0.131	-0.006	No.Signif
PQ	→	Recom	0.776	0.063	0.656	0.905	0.755	Signif
PCC	→	Recom	-0.015	0.066	-0.148	0.112	-0.01	No.Signif
With control factors on WB and Recom:								

PCC	→	PQ	1.07	0.081	0.924	1.237	0.731	Signif
PQ	→	WB	0.267	0.053	0.167	0.372	0.365	Signif
PCC	→	WB	-0.015	0.072	-0.158	0.124	-0.014	No.Signif
gender	→	WB	0.141	0.08	-0.011	0.303	0.063	No.Signif
age	→	WB	-0.067	0.064	-0.191	0.053	-0.04	No.Signif
stratum	→	WB	-0.022	0.059	-0.138	0.094	-0.014	No.Signif
education	→	WB	-0.058	0.08	-0.214	0.098	-0.026	No.Signif
PQ	→	Recom	0.769	0.066	0.649	0.907	0.746	Signif
PCC	→	Recom	-0.002	0.067	-0.133	0.127	-0.001	No.Signif
gender	→	Recom	0.066	0.085	-0.101	0.24	0.021	No.Signif
age	→	Recom	0.144	0.065	0.015	0.271	0.062	Signif
stratum	→	Recom	-0.134	0.062	-0.256	-0.013	-0.059	Signif
education	→	Recom	0.078	0.087	-0.092	0.243	0.025	No.Signif

PQ: process quality; PCC: patient-centred communication; WB: well-being; Recom: recommendation intentions.

Table 7 shows that, in both scenarios, the relationships were statistically significant (i.e., none of the intervals included zero). As was expected, PQ was positively related to both service recommendation (e.g., controlling factors; standardized loading: 0.746) and subjective well-being (controlling factors: 0.365). The present study has provided a new understanding of the mechanisms by which PQ and the two criteria ["WB" and "Recom"] were formed, considering patient-centred care from a communicational perspective (PCC). Thus, PCC can stimulate PQ (standardized loading: 0.731 with and without control factors). Then, this last construct is positively related with both "WB" (0.347 without and 0.365 with controlling factors) and recommendation intentions ("Recom"; 0.755 without and 0.746 with controlling factors).

On the other hand, age and socioeconomic stratum resulted in significant differences in recommendation intention; the confidence interval did not include zero. That is, older participants tended to be more likely to recommend the service than younger ones. When participants were from a lower socioeconomic stratum, they tended to recommend the service more than those from higher strata.

In summary, with and without control factors, the results supported the criterion validity of the scale.

Predict

We examined the capability of the healthcare scale to predict user well-being. The two discovered latent factors (PQ and PCC) were compared with the four sociodemographic factors. Additionally, following the MinerConstructo framework (Pérez-Rave, 2021), the test was performed from a machine learning perspective using linear regression, regression tree, random forest, and gradient boosting. The primary performance measure used was the R-

squared, which indicated how far the constructs (or latent factors) explained the variability in the response construct (F3). Thus, the higher the R^2, the greater the predictive capacity of the model. One of the risks inherent in predictive analysis is model overfitting. To prevent this, we first estimated the models and then used the 10-fold cross-validation strategy (Anguita et al., 2012). This preserved 10 data points from each set of random samples (bootstrap). Additional performance measures were root mean square error (RMSE) and mean absolute error (MAE). Figure 4 provides a statistical summary of the observable and latent variables involved in the comparative design (847 complete cases by including the well-being items); F1 corresponds to the standardized scores for PQ, and F2 the standardized scores for PCC.

```
      F1                  F2                  F3              gender
Min.    :-2.57872  Min.    :-2.01631  Min.    :-3.0774  female:554
1st Qu.:-0.62863   1st Qu.:-0.73398   1st Qu.:-0.4853   male  :293
Median : 0.07759   Median :-0.03461   Median : 0.2259
Mean    :-0.00121  Mean    :-0.00247  Mean    : 0.0014
3rd Qu.: 0.70534   3rd Qu.: 0.71987   3rd Qu.: 0.4794
Max.    : 2.39360  Max.    : 2.08265  Max.    : 1.6601
        age            stratum        education
27 to 39  :149    1or2:165    Postg   :542
40 to 59  :486    3   :466    UptoUniv:305
60 or more:212    4to6:216
```

Figure 4. Statistical summary of the variables used in the machine learning approach to examine the predictive validity of the scale. *Note(s).* F1, F2, and F3 represent the standardized scores for process quality, patient-centred communication, and well-being, respectively (*n*: 847 complete observations)

At the top of Table 8, we summarize the performance of the models by taking into consideration the mean values of the metrics in the 10 folds (each comprising 10 observations) used for validation. At the bottom of the table, we present the performance of the winning model ("Boost: Boosting") in each fold.

Table 8. Measures to examine predictive validity (dependent variable: well-being) using the machine learning approach (*n*: 847 complete observations)

Methods:	R-squared	RMSE	MAE
OLS: Linear Regression	0.083	0.96	0.746
Tree: Regression Tree	0.065	0.979	0.751
RF: Random forest	0.089	0.961	0.750
Boost: Boosting	0.099	0.960	0.749
Specific performance of the win model (Boost)*:			
Fold01	0.112	0.845	0.681
Fold02	0.191	0.991	0.743
Fold03	0.012	0.863	0.725
Fold04	0.147	0.921	0.702

Fold05	0.004	1.013	0.794
Fold06	0.134	0.935	0.727
Fold07	0.155	0.985	0.768
Fold08	0.078	0.991	0.757
Fold09	0.158	1.033	0.792
Fold010	0.001	1.022	0.806

*Each fold comprising 10 observations out of training sample

In Table 8, the R^2s range from 0.065 (regression tree) to 0.099 (boosting). The results in the 10 folds range from 0.001 to 0.191, with 1st and 3rd quartiles of 0.0285 - 0.153, respectively.

Now, replacing well-being by recommendation intention as the response variable (n: 814 complete cases), the R^2s ranged from 0.519 (regression tree) to 0.631 (random forest; winning model). This is consistent with other evidence regarding the positive relationship between service quality and behavioural intention (e.g., word of mouth and repurchase; Binnawas et al., 2020; Monoarfa & Usman, 2020).

Additionally, to examine whether the scale factors under study (PQ and PCC) had a relevant role as predictors of the response variables ("WB" and "Recom," respectively) from a machine learning perspective, consider Figure 5.
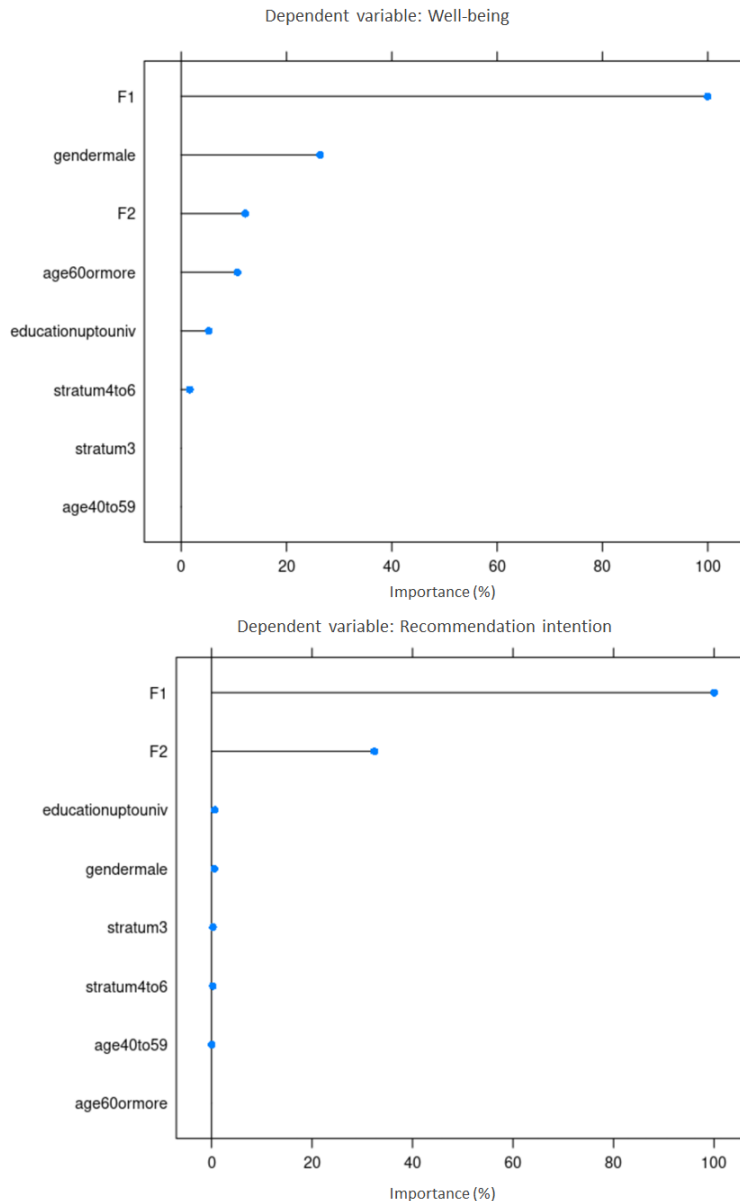
Figure 5. Predictors' importance from a non-parametric perspective based on the boosting model. *Note(s)*. The importance of the variable is scaled to have a maximum value of 100; F1 represents the standardized scores for process quality and F2 represents the standardized scores for patient-centred communication.

Figure 5 shows that the most relevant predictor of user well-being (F3, upper plot in Figure 5) from a machine learning approach was F1 (PQ), followed by gender and F2 (PCC). In the case of recommendation intention (bottom plot in Figure 5), PQ (F1) and PCC (F2) were the better predictors.

In summary, the results presented in this section favour the predictive validity of the developed scale and, along with the evidence of criterion validity (see Figure 3), highlight

the value of incorporating the patient-centred care approach in the healthcare service quality model.

Apply

Figure 6 shows the density graph of the global scores for each construct (PQ and PCC).
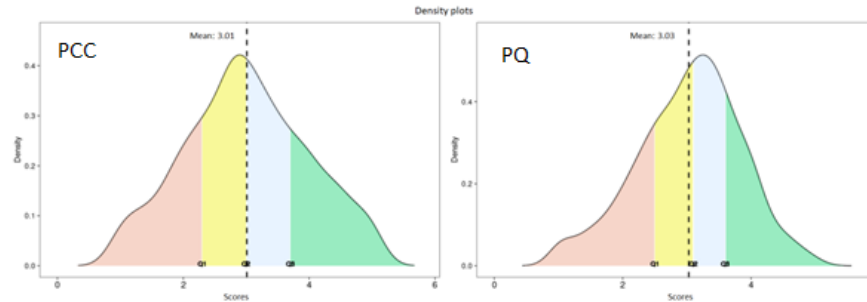
Figure 6. Density plot of patient-centred communication. *Note(s)*. Screen image (translated into English) from the MinerConstruct platform (Pérez-Rave, 2021). PQ: process quality; PCC: patient-centred communication; *n*: 869 complete observations.

The scores ($M = 3.01$, $SD = 0.976$) for PCC indicated that the healthcare services in the present study did not satisfy favourably the needs of external users in terms of PCC; more than 75% of the participants scored less than 4.0 on the PCC dimension. The scores presented a high coefficient of variation (32.4%). The PQ dimension presented a similar pattern ($M = 3.03$, $SD = 0.809$; coefficient of variation 26.7%).

The distribution of the dimensions' scores, which were notably symmetrical (Figure 1), allowed us to observe that service quality was stable. This makes it possible to predict PCC and PQ performances in the near future (Gutiérrez & Salazar, 2009) and enable quantitative characterizations of these dimensions by using several approaches such as hypothesis contrast of means and service capability analysis (from the statistical process control perspective). Hence, the scale could be employed to generate practical insights for decision-making processes in organizations by (a) diagnosing system quality levels using PCC and PQ scores; (b) establishing confidence intervals to contrast future hypotheses concerning service quality levels; and (c) defining the capacity of a service to satisfy quality specifications. These are useful for self-diagnostics and comparison of performance between services, centres, and countries. Appendix 1 illustrates some applications of these approaches. Moreover, the scale is presented in Appendix 2 so it can be shared with other researchers.

Conclusions

We propose a measure of the quality of healthcare services incorporating 17 items. The scale validation allowed for the discovery and psychometric confirmation of two essential dimensions relating to healthcare services in a developing country (Colombia): PCC (eight items) and PQ (nine items). The scale satisfies several psychometrical properties examined using a psychometric analytics framework and a sample of 917 healthcare users from

Colombia. The discovered dimensions represent relevant needs of users, contribute to an understanding of the conceptualization of service quality from a patient-centred care perspective, and complement previous studies that have focused on process aspects of healthcare service quality. Moreover, we illustrate three practical uses of the developed scale: diagnoses, hypothesis contrast, and service capacity analysis.

We have thereby addressed the calls of the IoM; WHO; Newell and Jordan (2015), and Perera and Dabney (2020) regarding the importance of patient-centred care in modern healthcare organizations. Moreover, the present study complements the work of previous researchers who used SERVQUAL. These include Tobón and Arias (2018; $n = 51$) and Asiamah et al. (2021; $n = 610$); the latter extended SERVQUAL by adding items relating to hospital hygiene and sanitation in Ghana. Others have adapted traditional items to new contexts, for instance, Almuhanadi et al. (2020) in Bahrein ($n = 520$); Lin et al. (2021), in Taiwan ($n = 200$); and Hou et al. (2019), in Datuan ($n = 350$). Moreover, given that the PQ dimension comprises items relating to timeliness, effectiveness, security, tangibles, and equity, and the PCC dimension addresses patient-centred care, we have addressed Arboleda-Arango, Chernichovsky, and Esperato's (2018) call for Colombian surveys that take users' viewpoints into consideration.

The present study has also provided empirical quantitative evidence that is consistent with some of the problems reported by Bonet, Guzmán, and Hahn (2017); DANE (2019); and the Ministry of Health and Social Protection (2013). These include poor customer service, timeliness, accessibility, inequity, and service quality in general.

With respect to the scale validation process, we took a comprehensive perspective by using a psychometrical data mining framework (seven processes incorporating exploratory factor analysis, structural equation modelling, and machine learning methods), which allowed for the provision of original evidence about construct plausibility, reliability, construct validity, equity, and criterion/predictive validity. These last consider the scale's capability of explaining/predicting subjective well-being and behavioural intention (i.e., recommending the service). The discovered/confirmed dimensions satisfy all these properties, which permits us to suggest that the scale is adequate for use in other research and in practical settings. The scale shows that both PCC and PQ are capable of significantly contributing to users' well-being and recommendation intention.

This paper enables researchers and managers to comprehensively understand and address service quality in the health sector by considering the traditional service quality approach (based on PQ) and explicitly incorporating a PCC dimension (and its items) from a communicational perspective. Thus, researchers might use the developed scale to capture/produce original data about PCC and PQ in several contexts and quantitatively contrast structural relationships comprising antecedents or consequents of such dimensions. Moreover, managers might use the developed scale to measure and improve service quality in healthcare organizations by considering three practical alternatives: the possibility for

diagnoses, hypothesis contrast based on confidence intervals, and estimation of the capacity of the service to satisfy specifications. Additionally, society can find in this paper a comprehensive basis (service quality in function of PQ and PCC) for recognizing and demanding a health service with quality from the user viewpoint.

Limitations

Although the addressed organization serves a huge number of users (more than 65,000 across several centres and cities), and the sample was larger (917 surveys received and 869 complete cases in the service quality items) than has hitherto been the case in similar research, the study used data from a single organization and a unique cultural context—Colombia. Hence, the sample limits the generalizations of our findings to other cultures and contexts. Future studies are required to validate the scale's psychometric properties by considering other countries, organizations, and populations.

Conflict of Interest

All authors declare that they have no conflict of interest.

Compliance with Ethical Standards Statement

The manuscript does not contain clinical studies. No personal data was collected from the participants. All surveys were completed voluntarily and anonymously. The healthcare organization was informed about the purpose of the study and the data to be collected. It authorized the application of the anonymous online questionnaire.

References

Alamo, M., Moral, R., & de Torres, L. (2002). Evaluation of a patient-centred approach in generalized musculoskeletal chronic pain/fibromyalgia patients in primary care. *Patient education and counseling*, 48(1), 23-31.

Al-Damen, R. (2017). Health care service quality and its impact on patient satisfaction "case of Al-Bashir Hospital". *International Journal of Business and Management*; 12(9).

Almuhanadi, S., Alhammadi, H., Suresh, A., & Alawi, S.A. (2020) Assessing service quality dimensions and their effect on patients satisfaction in Bahrain primary healthcare using a modified version of the general practice assessment questionnaire. *Patient Preference and Adherence*, 14, 2541-2549.

Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., and Ridella, S. (2012). The 'K'in K-fold cross validation. In 20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN) (pp. 441-446). i6doc. com publ.

Arboleda-Arango, A. M., Chernichovsky, D., & Esperato, A. (2018). Patient Satisfaction Surveys in Colombia: Scope for Improvement. *Revista Salud Uninorte*, 34(1), 33-46.

Asiamah, N., Frimpong Opuni, F., Aggrey, M., & Adu-Gyamfi, K. (2021) ADAPTED SERVQUAL: A Health Service Quality Scale Incorporating Indicators of Sanitation and Hygiene. *Quality Management in Health Care*, 30(3), 184-193

Australian Commission on Safety and Quality in Healthcare (2012). National Safety and Quality Health Service Standards (ACSQHC). Sydney.

Babakus, E., & Mangold, W. G. (1992). Adapting the SERVQUAL scale to hospital services: an empirical investigation. *Health Services Research*, 26(6), 767.

Behmane, D., Rutitis, D., & Batraga, A. (2021). Conceptual Framework for Attracting Foreign Patients to Health Care Services. In *Eurasian Business and Economics Perspectives* (pp. 259-275). Springer, Cham.

Binnawas, M. S. H., Khalifa, G. S., & Bhaumik, A. (2020). Antecedents of student's behavioral intentions in higher education institutions. *International Journal of Psychosocial Rehabilitation*, 24(03), 1949-1962.

Bonet, J., Guzmán, K. y Hahn, L. (2017). La salud en Colombia: una perspectiva regional. Bogotá: Banco de la República.

Carter, S., Ng, R., El-Den, S., & Schneider, C. (2021). Low perceived service quality in community pharmacy is associated with poor medication adherence. *Patient Education and Counseling*, 104(2), 387-394.

Coulthard, L. J. M. (2004). A review and critique of research using SERVQUAL. *International Journal of Market Research*, 46(4), 479-497.

Credé, M., & Harms, P. D. (2015). 25 years of higher-order confirmatory factor analysis in the organizational sciences: A critical review and development of reporting recommendations. *Journal of Organizational Behavior*, 36(6), 845-872.

Cuervo, A. A. V., Martínez, E. A. C., Quintana, J. T., & Amezaga, T. R. W. (2014). Differences in types and technological means by which Mexican high schools students perform cyberbullying: Its relationship with traditional bullying. *Journal of Educational and Developmental Psychology*, 4(1), 105-113.

Dean, A. M. (1999). The applicability of SERVQUAL in different health care environments. *Health Marketing Quarterly*, 16(3), 1-21.

Departamento Administrativo Nacional de Estadística (DANE). (2019). Encuestas calidad de vida (ECV). Obtained from: https://www.dane.gov.co/index.php/estadisticas-por-tema/salud/calidad-de-vida-ecv/encuesta-nacional-de-calidad-de-vida-ecv-2019

Doyle C., Lennox L., Bell, D. (2013). A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open*, 3(1): 1–17.

Endeshaw, B. (2020). Healthcare service quality-measurement models: a review. *Journal of Health Research,* 35(2), 106-117.

Epstein, R., Franks, P., Fiscella, K., Shields, C., Meldrum, S., Kravitz, R., & Duberstein, P. (2005). Measuring patient-centered communication in patient–physician consultations: theoretical and practical issues. *Social Science & Medicine*, 61(7), 1516-1528.

Guedes, M., & Araujo, C. (2022). Perceived Quality of Hospital Services from the Perspective of Doctors and Patients: An Integrative Model. Latin American Business Review, 23(1), 1-20.

Gutiérrez, H. and Salazar, V. (2009), Control estadístico de calidad y seis sigma, 2nd ed., McGraw-Hill, México DF.Horn, J. (1965). A rationale and test for the number of factors in factor analysis. *Psychométrika*, 30(2), 179-85.

Hou, J., Cai, L., Kang, J., & Tang, H. (2019) Service Quality Evaluation of Community-based TCM Health Management Delivered by the Family Doctor-led Team Using the SERVQUAL Scale. *Chinese General Practice*, 22(28), 3441-3445.

Institute of Medicine (IOM). (2021). Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, D.C: National Academy Press.

Islam, S., & Muhamad, N. (2021). Patient-centered communication: an extension of the HCAHPS survey. *Benchmarking: An International Journal*, 28(6), 2047-2074.

Jeon, Y. H., Kraus, S. G., Jowsey, T., & Glasgow, N. J. (2010). The experience of living with chronic heart failure: a narrative review of qualitative studies. BMC health services research, 10(1), 1-9.

Kennedy, D. M., Anastos, C. T., & Genau, M. C. (2019). Improving healthcare service quality through performance management. *Leadership in Health Services,* 32(3), 477-492.

Kitson, A., Marshall, A., Bassett, K., & Zeitz, K. (2013). What are the core elements of patient-centred care? A narrative review and synthesis of the literature from health policy, medicine and nursing. *Journal of Advanced Nursing*, 69(1), 4-15.

Knox, K., Kelly, F., Mey, A., Hattingh, L., Fowler, J. L., & Wheeler, A. J. (2015). Australian mental health consumers' and carers' experiences of community pharmacy service. *Health Expectations*, 18(6), 2107-2120.

Lin, C. H., Siao, S. F., Tung, H. H., Chung, K. P., & Shun, S. C. (2021). The gaps of healthcare service quality in nurse practitioner practice and its associated factors from the patients' perspective. *Journal of Nursing Scholarship*, 53(3), 378-386.

Liu, R., Shi, L., Meng, Y., He, N., Wu, J., Yan, X., & Hu, R. (2021). The institutional primary healthcare service quality and patients' experiences in Chinese community health centres:

results from the Greater Bay Area study, China. *International journal for equity in health*, 20(1), 1-9.

Makubalo, T., Scholtz, B., & Tokosi, T. O. (2020, April). Blockchain Technology for Empowering Patient-Centred Healthcare: A Pilot Study. In Conference on e-Business, e-Services and e-Society (pp. 15-26). Springer, Cham.

Martínez, M. R., Hernández, M. J., & Hernández, M. V. (2006). Psicometría, Alianza Editorial. Madrid, España.

Mathong, P., Sureeyatanapas, P., Arunyanart, S., & Niyamosoth, T. (2020). The assessment of service quality for third-party logistics providers in the beverage industry. *Cogent Engineering*, 7(1), 1785214.

Miles, A., & Mezzich, J. E. (2011). Person-centred medicine: Advancing methods, promoting implementation. *International Journal of Person Centred Medicine*, 1, 423–428.

Ministerio de Salud y Protección Social. (2013). Plan decenal de Salud Pública 2012-2021: la salud en Colombia la construyes tú. Obtained from: https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/PSP/PDSP.pdf

Mira, J. J., Aranaz, J., Rodriguez-Marín, J., Buil, J. A., Castell, M., & Vitaller, J. (1998). SERVQHOS: un cuestionario para evaluar la calidad percibida de la atención hospitalaria. Medicina preventiva, 4(4), 12-8.

Mody, M., Suess, C., & Dogru, T. (2020). Restorative servicescapes in health care: Examining the influence of hotel-like attributes on patient well-being. *Cornell Hospitality Quarterly*, 61(1), 19-39.

Monoarfa, T., & Usman, O. (2020). How does service quality impact the behavioral intentions of full-service airline passengers through a hierarchical model?. *Management Science Letters,* 10(16), 4011-4022.

Newell, S., & Jordan, Z. (2015). The patient experience of patient-centered communication with nurses in the hospital setting: a qualitative systematic review protocol. *JBI Evidence Synthesis*, 13(1), 76-87.

Numpaque-Pacabaque, A., Buitrago-Orjuela, L. Á., & Pardo-Santamaría, D. F. (2019). Calidad de la atención en el servicio de cirugía ambulatoria desde la percepción del usuario. Revista de la Facultad de Medicina, 67(2), 235-239.

OECD/The World Bank (2020), Panorama de la Salud: Latinoamérica y el Caribe 2020, OECD Publishing, Paris, https://doi.org/10.1787/740f9640-es.

Ornstein, P. H., & Kay, J. (2021). Enduring Difficulties in Medical Education and Training: Is There a "Cure"?. In The Annual of Psychoanalysis (pp. 155-172). Routledge.

Owusu-Frimpong, N., Nwankwo, S., & Dason, B. (2010). Measuring service quality and patient satisfaction with access to public and private healthcare delivery. *International Journal of Public Sector Management*, 23(3), 203-220.

Parasuraman, A., Zeithaml, V. A., & Berry, L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing,* 64(1), 12-40.

Perera, S., & Dabney, B. W. (2020). Case management service quality and patient-centered care. *Journal of Health Organization and Management*, 34(5), 551-568.

Pérez-Rave, J. (2021). MinerConstructo: marco inteligente para aprender, actualizarse y practicar minería de constructos con rigor científico [internal technical guide], IDINNOV, nref: 10-986-55, 13-ago.-2021, Medellín, 20p.

Pérez-Rave, J., Jaramillo-Álvarez, P., & González Echavarría, F. (2021). A psychometric data science approach to study latent variables: a case of class quality and student satisfaction. Total Quality Management & Business Excellence.

Pérez-Rave, J., Correa-Morales, J. and González-Echavarría, F. (2022b). Organizational social commitment and employee well-being: illustrating a construct mining approach in R, DYNA (in press).

Pérez-Rave, J., Fernández-Guerrero, R., Salas-Vallina, A. and González-Echavarría, F. (2022a). A measurement model of dynamic capabilities of the continuous improvement project and its role in the renewal of the company's products/services. Operations Management Research (in press).

Purcărea, V. L., Gheorghe, I. R., & Petrescu, C. M. (2013). The assessment of perceived service quality of public health care services in Romania using the SERVQUAL scale. *Procedia Economics and Finance*, 6, 573-585.

Pyburn Jr, K. M., Ployhart, R. E., & Kravitz, D. A. (2008). The diversity–validity dilemma: Overview and legal context. *Personnel Psychology*, 61(1), 143-151.

Setyawan, F. E. B., Supriyanto, S., Tunjungsari, F., Hanifaty, W. O. N., & Lestari, R. (2019). Medical staff services quality to patients satisfaction based on SERVQUAL dimensions. *International Journal of Public Health Science (IJPHS),* 8(1), 51-57.

Skogö Nyvang, J., Hedström, M., Iversen, M. D., & Andreassen Gleissman, S. (2019). Striving for a silent knee: a qualitative study of patients' experiences with knee replacement surgery and their perceptions of fulfilled expectations. International Journal of Qualitative Studies on Health and Well-being, 14(1), 1620551.

Sloan, L., & Angell, R. (2015). Learn about Pearson's Correlation Coefficient in SPSS with Data from the UK Living Cost and Food Survey (2010). SAGE Publications, https://bit.ly/3pnoh4u

Stewart, A., Greenfield, S., Hays, R., Wells, K., Rogers, W., ... & Ware, J. (1989). Functional status and well-being of patients with chronic conditions: results from the Medical Outcomes Study. *Jama*, 262(7), 907-913.

Tobón, L. A. L., & Arias, J. A. C. (2018). Evaluación de la calidad de los servicios médicos según las dimensiones del SERVQUAL en un hospital de Colombia. Archivos de medicina, 14(4), 1-4, https://dialnet.unirioja.es/servlet/articulo?codigo=6804255

Thomas, H. M., Bryce, C. L., Ness, R. B., & Hess, R. (2011). Dyspareunia is associated with decreased frequency of intercourse in the menopausal transition. *Menopause* (New York, NY), 18(2), 152-157.

Upadhyai, R., Jain, A. K., Roy, H., & Pant, V. (2019). A review of healthcare service quality dimensions and their measurement. *Journal of Health Management*, 21(1), 102-127.

Appendix 1

Illustrating Contrasts in Means and Service Capability Analysis Using Results From the "Apply" Section

Equation 1 describes the functions relating to the PCC (CI.PCC) and PQ (CI.PQ) scores, where K is a constant regarding the confidence level (e.g., 1.96 in the case of a 95% confidence level based on the normal distribution), $\bar{X}$ is the average score, S is the standard deviation, and n is the sample size.

$$CI = \bar{X} \pm K \times S/\sqrt{n} \tag{1}$$

$$CI.PCC = 3.01 \pm K \times 0.976/\sqrt{869} = 3.01 \pm 0.0649$$

$$CI.PQ = 3.03 \pm K \times 0.809/\sqrt{869} = 3.03 \pm 0.0538$$

Based on Equation 1, the true mean score of PCC in the present context ranged from 2.945 to 3.075 with a 95% confidence level. Likewise, the confidence interval for the true mean score of PQ was 2.976, 3.084 at 95% of confidence. In other words, it is expected that 95 of 100 randomly extracted samples of patients of the studied service present a mean score of the service (in PCC and PQ) that fall into the estimated intervals.

Equation 2 is the capacity ratio (C) of the service considering service desirable specifications. US and LS are upper and lower desired specifications, respectively (in this illustrating case, assume 3.5 and 5), and sigma (σ) is the service standard deviation; in this case, given the large size of the sample, we used the sample deviation ($\hat{\sigma} = S$). Thus, based on Gutiérrez and Salazar (2009), "C" represents the proportion of the specification band (5 − 3.5) that is occupied by the service ($6\hat{\sigma}$). Hence, "C" values less than 100% are desired.

$$C = \frac{6\hat{\sigma}}{US-LS} \times 100\% \tag{2}$$

$$C_{PCC} = \frac{6 \times 0.976}{5-3.5} = 390.4\%$$

$$C_{PQ} = \frac{6 \times 0.809}{5-3.5} = 323.6\%$$

The numerator represents the amplitude of the real score variation; the denominator represents the amplitude of the desirable specifications (3.5 to 5). Therefore, the service variation (numerator) in PCC exceeds the specification band by 390.4% and in PQ by 323.6%. These values reinforce the need for actions that favour service quality from the viewpoint of the user. One initiative should be orientated towards centring service quality near the middle point (4.25) of the desirable specifications (3.5 to 5). After that, the variation (standard deviation) around this point should be reduced. Using the proposed scale (PCC and PQ) and considering the functions described in Equations 1 and 2, organizations could gain insights into competitive referencing and decision-making processes. For example, both functions' confidence intervals and service capability are useful resources for benchmarking between services, centres, countries, and so on. In conclusion, diagnosis, confidence intervals, and service capability are just some of the many practical uses of the proposed scale.

Appendix 2

The Healthcare Service Quality Scale

- **PQ dimension**

Please indicate how often…

… it is easy for you to get a health services appointment when you need one ("timely1").

… healthcare professionals strive to keep your waiting time as short as possible ("timely3").

… healthcare professionals make you feel safe while delivering services ("secu1").

… healthcare professionals provide error-free care (in medication, diagnoses, surgeries, or other procedures) ("secu3").

… healthcare professionals provide the same treatment to users (i.e., patients and their families/companions), regardless of income, ethnicity, region, culture, colour, or gender ("equity1").

… the organization offers the same level of comfort in all facilities, services, and regions ("equity3").

… healthcare professionals provide understandable information to users (i.e., patients and their families/companions) ("effectiv1").

… medications, vaccines, tests, technology tools, and equipment are available when needed ("effectiv2").

… the facilities (cleaning, lighting, and so on) are comfortable for you ("tang").

- **PCC dimension**

How often do healthcare professionals (doctors, nurses, specialists, psychologists, nutritionists, and other health professionals)…

… encourage you to express your doubts about your health condition ("pat.cent1").

… clearly explain the reasons for your health condition to you ("pat.cent2").

… allow you to go into detail about your medical symptoms ("pat.cent4").

… ask for your opinion on the treatment to be carried out ("pat.cent5").

… struggle to understand how you feel ("pat.cent6").

… encourage you to develop self-care ("pat.cent7").

… clearly explain to you how to carry out your treatment ("pat.cent8").

… reassure you about your health condition ("pat.cent9").