Check for updates

Journal of Evaluation in Clinical Practice

WILEY

Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

ORIGINAL PAPER OPEN ACCESS

Cross-Cultural Adaptation and Psychometric Validation of the Hospital Survey on Patient Safety Culture Version 2 (HSOPSv2) in Spanish Hospitals

Julio J. Lopez-Picazo[1,2] | Juan Torres-Ramirez[1] | Juan J. Gascon-Canovas[1] | Julian Alcaraz[3] | Ana M. Seva[1] | Victor Soria-Aledo[1,3] | Pilar Escolar[1]

[1]University of Murcia, Murcia, Spain | [2]Healthcare Quality Unit, Arrixaca University Hospital, Murcia, Spain | [3]Morales Meseguer Hospital, Murcia, Spain

Correspondence: Juan Torres-Ramirez (juan.torresr@um.es)

Received: 2 February 2026 | **Revised:** 2 May 2026 | **Accepted:** 24 June 2026

Funding: Sociedad Murciana de Calidad Asistencial, Grant/Award Number: PI 01/2024

Keywords: hospital | patient safety | psychometric testing | surveys and questionnaires

ABSTRACT

Introduction: Patient safety culture is a key component of healthcare quality. The Hospital Survey on Patient Safety Culture (HSOPS), developed by AHRQ in 2004, has been widely used internationally to assess healthcare professionals' perceptions of safety. In 2019, version 2 (HSOPSv2) was released, with improvements in structure and item clarity. Although a Spanish version for North America exists, it has not yet been validated in the Spanish hospital context.

Objectives: To adapt and validate HSOPSv2 for the Spanish hospital context, ensuring its cultural, linguistic, and psychometric equivalence with the original version, and to assess its suitability for use in Spanish hospitals.

Methods: A cross-cultural adaptation process was conducted, including translation, back-translation, and expert consensus, followed by a cognitive pretest with 20 healthcare professionals. A pilot study was then carried out in four hospitals in Murcia and Alicante, with 369 participants. Internal consistency (Cronbach's α and composite reliability), convergent validity (AVE, factor loadings), and discriminant validity (Fornell–Larcker criterion and HTMT) were assessed through confirmatory factor analysis using the WLSMV estimator.

Results: The 10-factor model showed satisfactory fit indices (CFI and TLI > 0.90; RMSEA < 0.06; SRMR < 0.08). Most dimensions presented acceptable reliability (α and CR > 0.70), except for "Staffing and Work Pace" and "Handoffs and Information Exchange." Convergent validity was adequate in seven of the ten dimensions (AVE ≥ 0.50), with some weaker loadings in negatively worded items. The Fornell–Larcker criterion indicated overlap in some communication-related constructs, although HTMT results supported overall discriminant validity.

Conclusions: The Spanish version of HSOPSv2 shows satisfactory psychometric properties, comparable to other international adaptations. Its availability provides Spanish hospitals with a robust tool to assess patient safety culture, identify areas for improvement, and design interventions aimed at reducing adverse events and strengthening patient safety.

1 | Background

Healthcare organisations are increasingly recognising the need to transform organisational culture to enhance patient safety [1]. Preventable patient harm remains a major concern across all healthcare settings, including medication errors and healthcare-associated infections. Contributing factors often include organisational deficiencies, workload pressures, and communication failures among professionals [2]. Surveys on patient safety

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *Journal of Evaluation in Clinical Practice* published by John Wiley & Sons Ltd.

culture provide a mechanism to capture staff perceptions regarding safety culture within hospitals and to identify areas for improvement. A strong safety culture is essential, as organisational culture influences both the quality of care and the well-being of patients and healthcare staff. When a positive safety culture is established, communication improves, professional errors decrease, and overall quality of care is enhanced [3].

The Hospital Survey on Patient Safety Culture (HSOPS) was developed by the Agency for Healthcare Research and Quality (AHRQ) in 2004 to assess patient safety culture within healthcare organisations [4]. The original version comprised 42 items across 12 dimensions. Since its introduction, HSOPS has become a widely used tool internationally to assess staff perceptions of safety culture in hospitals. The survey supports the identification of strengths and areas for improvement and enables organisations to monitor progress over time.

The first Spanish adaptation of HSOPS version 1 was developed in 2005 [5] and validated in 2009 [6]. This version was pivotal in making the tool applicable in Spanish-speaking hospitals, particularly in Spain and Latin America, thus enabling systematic evaluation of patient safety culture in these contexts. In Spain, HSOPS has been used in several studies [7–10], across different professional groups and hospital services, demonstrating its value as a diagnostic tool for identifying opportunities for improvement.

However, the first version showed variability across adaptations, with weak internal reliability in some dimensions and inconsistent construct validity [11]. The second version of the survey (HSOPSv2) was launched in 2019 in response to advances in patient safety and changes in hospital practice [12]. The revised instrument reduced the number of items from 42 to 32 and the number of dimensions from 12 to 10, removing *"Overall perceptions of patient safety"* and *"Teamwork across units"* [13]. HSOPSv2 places greater emphasis on individual accountability and leadership commitment to patient safety improvement. It also focuses on organisational action—how hospitals address identified problem areas, promote continuous improvement, build resilience, and learn from errors. Furthermore, the updated version facilitates benchmarking across organisations, allowing hospitals to compare their safety culture with others.

Overall, HSOPSv2 represents a refined and updated instrument, aligned with current patient safety principles, designed to enhance item clarity and support more effective implementation. Although AHRQ provides a Spanish version for North America, there is currently no robust and context-specific validation for Spain. This gap limits cross-national comparability and the use of the improved psychometric properties introduced in the updated tool.

A Spanish validation is therefore required to ensure linguistic accuracy, cultural relevance, and psychometric robustness within the Spanish hospital context [14, 15]. Such validation would allow hospitals to monitor their safety culture, perform national and international benchmarking, and inform interventions aimed at reducing adverse events and strengthening patient safety. The aim of this study is to adapt the Hospital Survey on Patient Safety Culture version 2 (HSOPSv2) to the context of the Spanish healthcare system, ensuring that the questionnaire is culturally, linguistically, and metrically equivalent to the original version, and suitable for use in Spanish hospitals.

2 | Methods

2.1 | Instrument

The Hospital Survey on Patient Safety Culture version 2.0 (HSOPS v2), published by the Agency for Healthcare Research and Quality (AHRQ) in 2019 [12], was used as the study instrument. The questionnaire comprises an initial section on job position and work unit, followed by 32 items grouped into 10 dimensions addressing different aspects of patient safety.

Items are rated on five-point Likert scales, which vary depending on the question type: level of agreement ("Strongly disagree", "Disagree", "Neither agree nor disagree", "Agree", "Strongly agree") or frequency ("Never", "Rarely", "Sometimes", "Most of the time", "Always"). Two additional questions assess (a) overall patient safety grade ("Poor", "Fair", "Good", "Very good", "Excellent") and (b) the number of events reported in the previous year ("None", "1–2", "3–5", "6–10", "11 or more"). The survey concludes with general questions about respondents and an open-text field for comments regarding patient safety.

2.2 | Cross-Cultural Adaptation

The adaptation and validation process followed the translation–back translation method with expert validation [14], in accordance with AHRQ recommendations [12] and procedures used in previous survey validation studies [16–18]. A pilot study was subsequently conducted to assess comprehension and feasibility. Two native Spanish translators, both fluent in English and experienced in survey translation, independently produced Spanish versions of the original HSOPS v2. They were instructed to prioritise conceptual equivalence over literal translation, ensuring technical, semantic, and conceptual consistency. Two researchers compared both translations to evaluate linguistic accuracy, naturalness, and clarity. A consensus version of the Spanish questionnaire was then produced. The consensus Spanish version was back-translated into English by a native English translator with prior experience in health survey translation. A meeting between the research team and the translator was held to discuss semantic discrepancies and propose alternative wording. A final consensus back-translated version was then developed.

2.3 | Cognitive Pre-Testing

To evaluate questionnaire comprehension, a qualitative study was conducted with 20 healthcare professionals representing diverse roles. In-depth interviews explored respondents' views on item clarity, wording, relevance of response options, and interpretation of key concepts. Based on participant feedback, the research team reviewed and refined the questionnaire. The resulting version was considered the final Spanish adaptation of the HSOPS v2.

13652753, 2026, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jep.70521 by Ana Myriam Seva Llor - Spanish Cochrane National Provision (Ministerio de Sanidad) , Wiley Online Library on [19/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.4 | Pilot Testing and Assessment of Metric Equivalence

The final version of the questionnaire was administered to a sample of healthcare professionals from four hospitals in the provinces of Murcia and Alicante. The aim was to assess clarity, applicability, and usefulness. The survey was distributed electronically via institutional email or other internal communication channels. A cover letter with study information and appreciation for participation accompanied the survey link. Responses were collected anonymously. All negatively worded items were reverse-coded for statistical analysis. Descriptive statistics were computed to characterise the pilot sample.

Construct validity was examined through confirmatory factor analysis (CFA) [19] using the Weighted Least Squares Mean and Variance adjusted (WLSMV) estimation method, recommended for categorical data and appropriate for five-point Likert scales [20]. Each item was assigned to its corresponding dimension following the original HSOPS v2 structure. Model fit was assessed using the following indices [21, 22]: Root Mean Square Error of Approximation (RMSEA), Standardised Root Mean Square Residual (SRMR), Comparative Fit Index (CFI), Tucker–Lewis Index (TLI) and Chi-square (χ^2) and normalised chi-square (χ^2/df) [23].

Only fully completed questionnaires were included in the confirmatory factor analysis, resulting in an effective analytical sample of 183 participants. This approach was adopted to ensure complete data for model estimation using the WLSMV estimator.

Acceptable model fit was defined as CFI and TLI ≥ 0.90 [24], RMSEA < 0.06, and SRMR < 0.08 [22, 25]. A non-significant chi-square ($p > 0.05$) indicates good fit; alternatively, χ^2 values less than five times the degrees of freedom are acceptable, with values below twice the degrees of freedom considered optimal [26]. Internal consistency was assessed using Cronbach's alpha and Composite Reliability (CR), with 0.70 considered the minimum acceptable threshold [27, 28].

Convergent validity was evaluated using Average Variance Extracted (AVE), with acceptable values ≥ 0.50 [29], and by examining standardised factor loadings, considering values ≥ 0.50 as optimal and ≥ 0.30 as acceptable [30]. Discriminant validity was assessed using the Fornell–Larcker criterion [31], which requires the square root of the AVE of each construct to exceed its correlations with other constructs. The Heterotrait–Monotrait ratio (HTMT) was also calculated, with bootstrap confidence intervals (500 replications) [32]. Discriminant validity was supported if HTMT < 0.85 and the confidence interval did not include 1. All statistical analyses were conducted using R version 4.5.1, employing the packages lavaan, psych, semptools, and semTools.

3 | Results

3.1 | Cross-Cultural Adaptation

During the translation process, significant modifications were introduced in the sections concerning job position and work unit to ensure inclusion of all professional categories within Spanish hospitals and to align units and services with the structure of the Spanish healthcare system. Expressions were also adapted, and vocabulary was modified to reflect natural and culturally appropriate language.

3.2 | Cognitive Pre-Test

Participants in the cognitive interviews reported that the questionnaire was clear and easy to complete. Their suggestions mainly concerned the section on job position and work setting. Based on their feedback, questions were added regarding working hours and professional experience. Participants also recommended adopting gender-inclusive language, and minor lexical adjustments were made to reflect commonly used terminology in Spanish hospital settings.

3.3 | Pilot Testing and Assessment of Metric Equivalence

During the pilot phase, a total of 369 healthcare professionals participated. Most respondents were women (74.5%), with a median age of 48 years (IQR = 13). Over half of the sample worked at *Hospital Virgen de la Arrixaca* (52.0%). Regarding professional profile, nursing staff predominated (53.7%), followed by medical staff (20.1%). A total of 13.3% held supervisory or management positions, and 6.5% performed support roles. Four out of five participants (79.4%) reported direct contact with patients in their daily work (Table 1).

For statistical analyses, only responses from participants who completed the entire questionnaire were included, excluding those with missing items or "Does not apply or don't know"

TABLE 1 | Sociodemographic characteristics of the sample.

Variables	*n*	%
Gender		
Male	88	23.8
Female	275	74.5
Hospital		
Virgen de la Arrixaca	192	52.0
Santa Lucía/Santa María del Rosell	36	9.8
Morales Meseguer	86	23.3
Vega Baja	49	13.3
Direct patient care	293	79.4
Professional group		
Nursing	198	53.7
Medical	74	20.1
Supervisory/management	49	13.3
Support roles	24	6.5
Other	18	4.9
Age (years)	**median**	**IQR**
	48	13

Abbreviations: IQR, interquartile range; *n*, sample size. Percentages were calculated using n = 369; totals may not sum to 369 because of missing sociodemographic data.

13652753, 2026, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jep.70521 by Ana Myriam Seva Llor - Spanish Cochrane National Provision (Ministerio de Sanidad) , Wiley Online Library on [19/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

answers. Consequently, 49.6% of the surveys ($n = 183$) were retained for analysis.

A confirmatory factor analysis (CFA) was conducted based on the 10-factor model proposed by AHRQ. Goodness-of-fit indices and their interpretation are presented in Table 2. Standardised factor loadings were significant for all items ($p < 0.001$), with generally high values (> 0.70). Only four items (A2, A3, A5, A9) did not reach the 0.50 threshold (Table 3). Composite reliability (CR) coefficients ranged from 0.57 (*Staffing and Work Pace*) to 0.94 (*Response to Error*), while most ordinal Cronbach's α values exceeded 0.70, except for *Staffing and Work Pace* (0.59) and *Handoffs and Information Exchange* (0.57). Convergent validity, assessed through Average Variance Extracted (AVE), showed that seven out of ten constructs surpassed the recommended 0.50 threshold, with particularly high values for *Reporting Patient Safety Events* (0.85), *Communication about Error* (0.80), and *Response to Error* (0.70). Lower AVE values for *Teamwork*, *Staffing and Work Pace*, and *Handoffs and Information Exchange* suggest weaker convergence for these dimensions. For dimensions with Cronbach's α below or close to 0.70, item-deletion analyses were performed to estimate the impact on reliability; but results did not show any relevant variation.

According to the Fornell–Larcker criterion, the square root of the AVE exceeded the correlations with other constructs in most cases, except for *Teamwork*, *Staffing and Work Pace and Communication Openness* (Table 4). HTMT values ranged between 0.36 and 0.86, remaining below the recommended 0.85 threshold in most cases. In all instances, 95% confidence intervals did not include or exceed 1, supporting adequate discriminant validity among constructs.

4 | Discussion

4.1 | Main Findings

This study represents the first validation in European Spanish of the Hospital Survey on Patient Safety Culture, version 2 (HSOPS v2). A rigorous cross-cultural adaptation process was conducted following international standards, with the involvement of a multidisciplinary panel of experts and professional translators experienced in survey adaptation. The proposed version underwent cognitive pretesting with 20 healthcare professionals, whose feedback was incorporated to produce the final Spanish version of the questionnaire, which was subsequently pilot-tested to assess its validity.

Results confirmed that the ten-factor structure proposed by AHRQ demonstrated an adequate model fit in the Spanish hospital context. The fit indices met internationally accepted thresholds (CFI and TLI > 0.90; RMSEA < 0.06; SRMR < 0.08; χ^2/df < 2), supporting the construct validity of the adapted instrument.

Internal reliability was also generally acceptable, with Cronbach's α and Composite Reliability (CR) values above the 0.70 threshold for most dimensions. However, three factors did not reach this level: *Teamwork* ($\alpha = 0.71$; CR = 0.61), *Staffing and work pace* ($\alpha = 0.59$; CR = 0.57), and *Handoffs and information exchange* ($\alpha = 0.57$; CR = 0.64). These results indicate that the items within these dimensions are not fully consistent with one another. This finding is consistent with previous international validation studies, in which the same dimensions have shown relatively lower internal consistency, likely due to item heterogeneity or contextual variation in how these practices are perceived across healthcare systems. When assessing whether Cronbach's α would improve by removing any individual item, no substantial increase was observed. Therefore, the decision was made to retain all items, in line with other national validations, to preserve the integrity of the original AHRQ model [12].

Convergent validity, measured through the Average Variance Extracted (AVE), was adequate for seven out of the ten factors, particularly those related to communication and incident reporting. However, three dimensions (*Teamwork*, *Staffing and work pace*, and *Handoffs and information exchange*) exhibited AVE values below the 0.50 threshold, suggesting weaker convergence within these constructs. This result may also be partially related to the reduced effective sample size available for confirmatory factor analysis, which can affect the stability of parameter estimates in multidimensional models with a relatively large number of items. Regarding factor loadings, although most were acceptable, four items showed weaker results: A2, A3, A5, and A9. Specifically, *"A3 – Staff in this unit work longer hours than is best for patient care"* had a standardised loading of 0.27, suggesting that this item may not correlate adequately with its intended dimension (*Staffing and work pace*). Notably, three of these four items were negatively worded, which may have affected response validity [33]. These results suggest heterogeneity in responses within these dimensions, possibly due to differences in professional roles or challenges in item interpretation. It is therefore advisable to review the items composing the *Teamwork*, *Staffing and work pace*, and *Handoffs and information exchange* dimensions in future Spanish adaptations.

TABLE 2 | Model fit indices for the confirmatory factor analysis.

Index	Value	Acceptability criterion (interpretation)
χ^2 (*p*-value)	572.492 ($p < 0.001$)	$p > 0.05$ (a significant p is expected with large samples)
df	419	—
χ^2/df	1.366	< 2 (good fit)
CFI	0.995	≥ 0.95 (good fit)
TLI	0.994	≥ 0.90 (good fit)
RMSEA (90% CI)	0.045 (0.035–0.054)	≤ 0.05 (good fit)
SRMR	0.064	≤ 0.08 (good fit)

Abbreviations: CFI, Comparative Fit Index; df, degrees of freedom; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardised Root Mean Square Residual; TLI, Tucker–Lewis Index; χ^2, chi-square.

13652753, 2026, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jep.70521 by Ana Myriam Seva Llor - Spanish Cochrane National Provision (Ministerio de Sanidad) , Wiley Online Library on [19/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 | Reliability indices for the questionnaire factors.

Construct	Item	Factor Loading	α	CR	AVE
1. Teamwork	A1	0.78	0.71	0.61	0.44
	A8	0.72			
	A9(r)	0.46			
2. Staffing and Work Pace	A2	0.47	0.59	0.57	0.31
	A3(r)	0.27			
	A5(r)	0.40			
	A11(r)	0.87			
3. Organizational Learning—Continuous Improvement	A4	0.78	0.77	0.82	0.61
	A12	0.77			
	A14(r)	0.78			
4. Response to Error	A6(r)	0.86	0.87	0.94	0.70
	A7(r)	0.89			
	A10	0.73			
	A13(r)	0.86			
5. Supervisor, Manager, or Clinical Leader Support for Patient Safety	B1	0.90	0.82	0.88	0.74
	B2(r)	0.75			
	B3	0.91			
6. Communication About Error	C1	0.80	0.92	0.90	0.80
	C2	0.93			
	C3	0.95			
7. Communication Openness	C4	0.77	0.84	0.81	0.59
	C5	0.78			
	C6	0.92			
	C7(r)	0.56			
8. Reporting Patient Safety Events	D1	0.92	0.91	0.88	0.85
	D2	0.92			
9. Hospital Management Support for Patient Safety	F1	0.84	0.78	0.84	0.60
	F2	0.92			
	F3(r)	0.51			
10. Handoffs and Information Exchange	F4(r)	0.80	0.57	0.64	0.40
	F5(r)	0.52			
	F6	0.54			

Abbreviations: (r), reverse-coded items; α, Cronbach's alpha (ordinal); AVE, Average Variance Extracted; CR, Composite Reliability.

The Fornell–Larcker criterion indicated some issues with discriminant validity. In particular, *Teamwork* and *Staffing and work pace* showed correlations with other factors that exceeded their √AVE, suggesting conceptual overlap. Similarly, *Communication openness* exhibited strong correlations with *Communication about error*, indicating that respondents may not clearly distinguish between these two constructs. However, the HTMT criterion yielded more favourable results, with all values within acceptable limits, suggesting that discriminant validity was not compromised. These findings support the view that the dimensions maintain conceptual distinctiveness, each capturing specific aspects of patient safety culture.

It is important to note that the reduced number of items per dimension [2–4] increases the sensitivity of psychometric indices to one or two items with low factor loadings [11, 34]. Furthermore, the high intercorrelations observed between some dimensions may warrant further examination in future revisions—potentially by merging related constructs—to determine whether this enhances both convergent and discriminant validity.

4.2 | Comparison with Other Countries

Similar validation processes of the HSOPS v2 have been conducted in Brazil [35], Chile [36], China [34], South Korea [37], Indonesia [38], Italy [39], Malaysia [40], and Norway [41]. In Brazil, the United States, Italy, Malaysia, and Norway, the pilot validation surveys were distributed to all hospital healthcare professionals, following the same approach as in the present study. In contrast, studies from Chile, China, South Korea, and Indonesia included only nursing staff in their validation

13652753, 2026, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jep.70521 by Ana Myriam Seva Llor - Spanish Cochrane National Provision (Ministerio de Sanidad) , Wiley Online Library on [19/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 | Discriminant validity (Fornell–Larcker criterion).

Latent correlation matrix with square root of AVE on the diagonal (in bold)										
Construct	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	**0.67**									
2	0.52	**0.55**								
3	0.67	0.55	**0.78**							
4	0.70	0.51	0.62	**0.84**						
5	0.70	0.51	0.73	0.61	**0.86**					
6	0.71	0.63	0.78	0.58	0.76	**0.90**				
7	0.70	0.52	0.65	0.60	0.78	0.88	**0.77**			
8	0.48	0.49	0.63	0.53	0.61	0.84	0.81	**0.92**		
9	0.55	0.72	0.75	0.55	0.65	0.71	0.69	0.53	**0.78**	
10	0.41	0.60	0.36	0.32	0.39	0.45	0.55	0.34	0.59	**0.63**

TABLE 5 | Cronbach's α values across countries.

Dimension	Spain	USA	S. Korea	Indonesia	Brazil	China	Malaysia	Norway	Italy
Teamwork	0.71	0.76	0.77	0.76	0.68	0.75	0.61	0.58	0.70
Staffing and Work Pace	0.59	0.67	0.61	0.73	0.47	0.75	0.60	0.74	0.53
Organizational Learning—Continuous Improvement	0.77	0.76	0.71	0.76	0.60	0.87	0.64	0.71	0.70
Response to Error	0.87	0.83	0.72	0.68	0.76	0.82	0.61	0.74	0.77
Supervisor, Manager, or Clinical Leader Support for Patient Safety	0.82	0.77	0.75	0.74	0.71	0.68	0.60	0.74	0.72
Communication About Error	0.92	0.89	0.83	0.73	0.87	0.83	0.71	0.85	0.89
Communication Openness	0.84	0.83	0.73	0.67	0.76	0.75	0.63	0.76	0.71
Reporting Patient Safety Events	0.91	0.75	0.73	0.81	0.81	0.82	0.80	0.86	0.82
Hospital Management Support for Patient Safety	0.78	0.77	0.72	0.75	0.62	0.87	0.61	0.79	0.69
Handoffs and Information Exchange	0.57	0.72	0.72	0.76	0.50	0.93	0.66	0.80	0.67

samples. Furthermore, while Brazil, Chile, and Malaysia recruited participants from a single hospital, the remaining countries conducted multi-centre studies, with China including the largest number of participating hospitals (five).

In Chile, researchers performed an exploratory factor analysis (EFA) prior to validation and decided to reduce and reorganise the questionnaire, resulting in a version with seven dimensions and 23 items, which improved internal consistency. In South Korea, one item was removed (*"A5 – In my unit, there are too many temporary, float, or contract staff"*), as the expert panel considered it not applicable to the Korean context, where such employment arrangements are uncommon—unlike in Spain [42]. Other countries, like the present study, retained the original AHRQ structure without item elimination.

Across all countries, the model fit indices reported were acceptable. However, internal reliability showed some variability, with China being the only country where all Cronbach's α values exceeded the 0.70 threshold across every dimension.

When compared with the international average, the Spanish version showed higher internal consistency in eight out of ten dimensions, with lower reliability observed only for *Staffing and work pace* and *Handoffs and information exchange*. These two dimensions also displayed lower reliability in Brazil, the United States, and Italy, indicating that the present findings are consistent with international evidence (Table 5).

4.3 | Strengths and Limitations

One of the main strengths of this study lies in its rigorous process of translation and cross-cultural adaptation, conducted in accordance with international standards. This ensures that the Spanish version of the HSOPS v2 is conceptually equivalent to the original and retains its psychometric integrity. Unlike in some previous validations, this study included a diverse range of healthcare professional profiles, thereby providing a sample more representative of the hospital work environment. Additionally, the use of robust confirmatory factor analysis methods with estimation techniques appropriate for ordinal data (WLSMV) adds further strength and methodological rigour to the results.

Among the study's limitations, it should be noted that although the initial sample included 369 healthcare professionals, only

13652753, 2026, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jep.70521 by Ana Myriam Seva Llor - Spanish Cochrane National Provision (Ministerio de Sanidad) , Wiley Online Library on [19/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

fully completed questionnaires were retained for confirmatory factor analysis, resulting in an effective analytical sample of 183 participants. While this sample size is considered acceptable for CFA using the WLSMV estimator with ordinal data, it may have reduced the stability of some parameter estimates, particularly in dimensions with weaker factor loadings and lower AVE values. Furthermore, excluding incomplete questionnaires may have introduced some degree of bias if missing responses were not completely random. This reduced proportion of fully completed questionnaires could be related to the limited distribution period (the survey was available online for 1 month), lack of time to complete the questionnaire, or reduced motivation among staff.

The study sample was drawn from four hospitals located in two provinces in southeastern Spain. Therefore, caution should be exercised when extrapolating these findings to the entire Spanish hospital system until further validation studies with broader national samples are conducted.

5 | Conclusions

This study successfully adapted and validated the Hospital Survey on Patient Safety Culture, version 2 (HSOPS v2), for use in the Spanish hospital context, following a rigorous process of translation, back-translation, expert review, cognitive pretesting, and pilot testing among healthcare professionals. The findings confirm that the ten-dimension structure proposed by AHRQ demonstrates an adequate model fit, with good internal reliability and convergent validity across most factors, and overall acceptable discriminant validity. However, the dimensions *Teamwork*, *Staffing and Work Pace*, and *Handoffs and Information Exchange* showed lower internal consistency, suggesting the need for further refinement in future adaptations.

The availability of this validated version in European Spanish provides a robust tool for assessing patient safety culture in hospitals across Spain. It facilitates international comparisons and enables longitudinal monitoring of safety culture over time. Its use will support hospital managers and quality leaders in identifying areas for improvement, designing targeted interventions to reduce adverse events, and strengthening patient safety within the Spanish healthcare system.

Author Contributions

JJLPF: conceptualization, study design, supervision, manuscript drafting; coordinated pretest and hospital participation. JJGC: expert review of adaptation process; designed and coordinated the statistical analysis strategy. JAM and AMSL: coordinated pretest and hospital participation. JTR: support in data collection and preliminary analysis; executed the statistical analysis and drafted the initial manuscript. All authors participated in successive versions of the translation, contributed to manuscript revisions, and approved the final version.

Acknowledgements

The authors acknowledge the Agency for Healthcare Research and Quality (AHRQ) for granting permission to translate the Hospital Survey on Patient Safety Culture 2.0 into Spanish. We also thank all healthcare professionals from the participating hospitals in Murcia and Alicante for their valuable time and contributions to the survey. This research was supported by the Murcian Society for Healthcare Quality (Sociedad Murciana de Calidad Asistencial, SOMUCA) through its Research Grant Program, project PI 01/2024 (2024 call).

Ethics Statement

The study was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained from the Clinical Research Ethics Committee of the University of Murcia (Spain) (ACTA14/2024/CEI) and from the local Research Ethics Committees of each participating hospital. Participation was voluntary, and informed consent was obtained from all respondents prior to completing the survey. Data were collected anonymously to ensure confidentiality.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. S. Singer, S. Lin, A. Falwell, D. Gaba, and L. Baker, "Relationship of Safety Climate and Safety Performance in Hospitals," *Health Services Research* 44, no. 2p1 (2009): 399–421.

2. M. Panagioti, K. Khan, R. N. Keers, et al., "Prevalence, Severity, and Nature of Preventable Patient Harm Across Medical Care Settings: Systematic Review and Meta-Analysis," *BMJ* 366 (July 2019): l4185.

3. C. E. Sammer, K. Lykens, K. P. Singh, D. A. Mains, and N. A. Lackan, "What Is Patient Safety Culture? A Review of the Literature," *Journal of Nursing Scholarship* 42, no. 2 (2010): 156–165.

4. J. S. Sorra and V. F. Nieva, "Hospital Survey on Patient Safety Culture. Prepared by Westat under Contract No. 290-96-0004. Rockville, MD: Agency for Healthcare Research and Quality; September 2004. AHRQ Publication No. 04-0041. (2004), https://doi.org/10.1037/e563322011-001.

5. J. J. Gascón-Cánovas and P. J. Saturno-Hernández, *Versión española del Hospital Survey on Patient Safety* (Madrid: Ministerio de Sanidad y Consumo, 2005).

6. J. Palacio-Vieira, S. Moncada, and L. Artazcoz, "Validación De La Encuesta Sobre La Cultura De La Seguridad Del Paciente (HSOPS) En Hospitales Españoles," *Rev Esp Salud Pública* 84, no. 5 (2010): 577–584.

7. A. Gil-Aucejo, S. Martínez-Martín, P. Flores-Sánchez, et al., "Assessment of the Patient Safety Culture in the ICU of a Second Level Hospital at the End of the Third COVID-19 Wave," *Enferm Intensiva Engl Ed* 33, no. 4 (October 2022): 185–196.

8. M. Mella Laborde, M. T. Gea Velázquez, J. M. Aranaz Andrés, G. Ramos Forner, and A. F. Compañ Rosique, "Análisis De La Cultura De Seguridad Del Paciente En Un Hospital Universitario," *Gaceta Sanitaria* 34, no. 5 (September 2020): 500–513.

9. F. Pozo Muñoz and V. Padilla Marín, "Evaluación De La Cultura De Seguridad Del Paciente En El Ámbito De Un Área Sanitaria," *Revista de Calidad Asistencial* 28, no. 6 (2013): 329–336.

10. F. Roqueta Egea, S. Tomás Vecina, M. Chanovas Borrás, "Cultura De Seguridad Del Paciente En Los Servicios De Urgencias: Resultados De Su Evaluación En 30 Hospitales Del Sistema Nacional De Salud Español," *Emerg Rev Soc Esp Med Urgenc Emerg* 23, no. 5 (2011): 356–363.

11. P. Waterson, E. M. Carman, T. Manser, and A. Hammer, "Hospital Survey on Patient Safety Culture (Hspsc): a Systematic Review of the

13652753, 2026, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jep.70521 by Ana Myriam Seva Llor - Spanish Cochrane National Provision (Ministerio de Sanidad) , Wiley Online Library on [19/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Psychometric Properties of 62 International Studies," *BMJ Open* 9, no. 9 (September 2019): e026896.

12. J. Sorra, N. Yount, and T. Famolaro AHRQ Hospital Survey on Patient Safety Culture Version 2.0: User's Guide. June 2021.

13. J. Sorra, T. Famolaro, and N. D. Yount, "Transitioning to the SOPSTM Hospital Survey Version 2.0: What's Different and What To Expect, Part I: Main Report," Prepared by Westat under Contract No. HHSP233201500026I/HHSP23337004T. Rockville, MD: Agency for Healthcare Research and Quality; September 2019. AHRQ Publication No. 19-0076-1-EF. (September 2019).

14. P. Cruchinho, M. D. López-Franco, M. L. Capelas, et al., "Translation, Cross-Cultural Adaptation, and Validation of Measurement Instruments: A Practical Guideline for Novice Researchers," *Journal of Multidisciplinary Healthcare* 17 (May 2024): 2701–2728.

15. T. V. Perneger, A. Staines, and F. Kundig, "Internal Consistency, Factor Structure and Construct Validity of the French Version of the Hospital Survey on Patient Safety Culture," *BMJ Quality & Safety* 23, no. 5 (May 2014): 389–397.

16. S. Arias-Rivera, M. M. Sánchez-Sánchez, M. P. Fraile-Gamo, et al., "Adaptación Transcultural Al Castellano Del Nursing Activities Score," *Enfermería Intensiva* 24 24, no. 1 (2013): 12–22.

17. P. Garcimartin, Y. Pardo-Cladellas, J. M. Verdú-Rotellar, P. Delgado-Hito, M. Astals-Vizcaino, and J. Comin-Colet, "Adaptación Transcultural Al Español Del Cuestionario Patient Empowerment in Long-Term Conditions," *Atención Primaria* 51, no. 1 (2019): 24–31.

18. A. Ocaña-Ortiz, J. J. Paredes-Carbonell, R. Peiró-Pérez, E. Pérez-Sanz, V. Gea-Caballero, et al., "Evaluación Participativa Del Territorio Con Enfoque De Equidad: Adaptación Y Validación De La Place Standard Al Contexto Español," *Gaceta Sanitaria* 36, no. 4 (August 2022): 360–367.

19. J. Herrero, "El Análisis Factorial Confirmatorio en el Estudio de la Estructura y Estabilidad de los Instrumentos de Evaluación: Un Ejemplo Con El Cuestionario De Autoestima CA-14," *Psychosocial Intervention* 19, no. 3 (December 2010): 289–300.

20. E. Sánchez-Moreno, L. P. Gallardo-Peralta, and J. L. Gálvez-Nieto Validación psicométrica de la Escala de Malestar Psicológico Kessler 10 (K-10) en personas mayores en España. Gac Sanit [Internet]. January 1, 2025 [September 8, 2025];39; 00.

21. R. B. Kline, *Principles and Practice of Structural Equation Modeling* (New York, NY, US: The Guilford Press, 2005). 2nd ed.

22. Á. Morata-Ramirez, F. P. Holgado-Tello, M. I. Barbero-García, and G. Mendez, "Análisis Factorial Confirmatorio: Recomendaciones Sobre Mínimos Cuadrados No Ponderados En Función Del Error Tipo I de Ji-Cuadrado y RMSEA," *Acción Psicológica* 12, no. 1 (June 2015): 79–90.

23. R. B. Kline, *Principles and Practice of Structural Equation Modeling* (Guilford Publications, 2023).

24. R. Schumaker and R. Lomax. A beginner's Guide to Structural Equation Modeling. 2022.

25. L. Hu and P. M. Bentler, "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling: A Multidisciplinary Journal* January 1 6, no. 1 (1999): 1–55.

26. M. Alavi, D. C. Visentin, D. K. Thapa, G. E. Hunt, R. Watson, and M. Cleary, "Chi-Square for Model Fit in Confirmatory Factor Analysis," *Journal of Advanced Nursing* 76, no. 9 (2020): 2209–2211.

27. H. C. Oviedo and A. Campo-Arias, "Aproximación Al Uso Del Coeficiente Alfa De Cronbach," *Rev Colomb Psiquiatr* 34, no. 4 (December 2005): 572–580.

28. Á. Roco-Videla, S. V. Flores, M. Olguín-Barraza, N. Maureira-Carsalade, et al., "[Cronbach's Alpha and Its Confidence Interval]," *Nutricion Hospitalaria* 41, no. 1 (February 2024): 270–271.

29. G. W. Cheung, H. D. Cooper-Thomas, R. S. Lau, and L. C. Wang, "Reporting Reliability, Convergent and Discriminant Validity With Structural Equation Modeling: A Review and Best-Practice Recommendations," *Asia Pacific Journal of Management* 41, no. 2 (June 2024): 745–783.

30. J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis* (Andover, Hampshire: Cengage, 2019). Eighth Edition.

31. J. Henseler, C. M. Ringle, and M. Sarstedt, "A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling," *Journal of the Academy of Marketing Science* 43, no. 1 (January 2015): 115–135.

32. E. Roemer, F. Schuberth, and J. Henseler, "HTMT2–an Improved Criterion for Assessing Discriminant Validity in Structural Equation Modeling," *Industrial Management & Data Systems* 121, no. 12 (September 2021): 2637–2650.

33. V. D. Rada and de, "Influye El Diseño De Las Preguntas En Las Respuestas De Los Entrevistados?," *Rev Esp Sociol* 31, no. 1 (2022): a83, https://doi.org/10.22325/fes/res.2022.83.

34. Y. Wu, W. Hua, D. Zhu, R. Onishi, Y. Yang, and T. Hasegawa, "Cross-Cultural Adaptation and Validation of the Chinese Version of the Revised Surveys on Patient Safety Culture™ (Sops®) Hospital Survey 2.0," *BMC Nursing* 21, no. 1 (December 2022): 369.

35. C. T. Reis, J. Laguardia, P. Bruno de Araújo Andreoli, C. Nogueira Júnior, and M. Martins, "Cross-Cultural Adaptation and Validation of the Hospital Survey on Patient Safety Culture 2.0 – Brazilian Version," *BMC Health Services Research* 23, no. 1 (January 2023): 32.

36. P. Hurtado-Arenas, M. R. Guevara, and V. M. González-Chordá, "Patient Safety Culture From a Nursing Perspective in a Chilean Hospital," *Nursing Reports* 14, no. 2 (June 2024): 1439–1451.

37. S. E. Lee and V. S. Dahinten, "Adaptation and Validation of a Korean-Language Version of the Revised Hospital Survey on Patient Safety Culture (K-HSOPSC 2.0)," *BMC Nursing* 20, no. 1 (January 2021): 12.

38. L. Suryani, S. Letchmi, and F. Binti Moch Said, "Cross-Culture Adaptation and Validation of the Indonesian Version of the Hospital Survey on Patient Safety Culture (HSOPSC 2.0)," *Belitung Nursing Journal* 8, no. 2 (April 2022): 169–175.

39. A. Bagnasco, G. Catania, M. T. Loiudice, et al. Validation of the Hospital Survey on Patient Safety Culture 2.0 in Italian Hospitals: A Cross-Sectional Study of Healthcare Personnel Perceptions. J Adv Nurs [Internet]. 2025 [September 9, 2025].

40. D. S. H. Imran Ho, M. H. Jaafar, and A. Mohammed Nawi, "Revised Hospital Survey on Patient Safety Culture (HSOPSC 2.0): Cultural Adaptation, Validity and Reliability of the Malay Version," *BMC Health Services Research* 24, no. 1 (October 2024): 1287.

41. E. Olsen, S. A. J. Addo, S. S. Hernes, M. H. Christiansen, A. S. Haugen, and A. C. L. Leonardsen, "Psychometric Properties and Criterion Related Validity of the Norwegian Version of Hospital Survey on Patient Safety Culture 2.0," *BMC Health Services Research* 24, no. 1 (May 2024): 642.

42. P. Galbany-Estragués and P. Millán-Martínez [Shortage of Nurses in Spain: From the Global Case to Particular Situation. SESPAS Report 2024]. Gac Sanit. 2024;38 Suppl 1:102376.

13652753, 2026, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jep.70521 by Ana Myriam Seva Llor - Spanish Cochrane National Provision (Ministerio de Sanidad) , Wiley Online Library on [19/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License