Interdisciplinary
Advances in Health

ORIGINAL

Adaptation and Validation of a Brief Artificial Intelligence Job Performance Scale (BAIJPS) in Nurses

Wilter C. Morales-García[1*], **Liset Z. Sairitupa-Sanchez[2]**

[1]Facultad de Teología, Universidad Peruana Unión. Lima, Perú.
[2]Escuela Profesional de Psicología, Facultad de Ciencias de la Salud, Universidad Peruana Unión. Lima, Perú.

Citation:
Morales-García, W. C., & Sairitupa-Sanchez, L. Z. (2023). Adaptation and Validation of a Brief Artificial Intelligence Job Performance Scale (BAIJPS) in Nurses. *Interdisciplinary Advances in Health*, *1*, 2. https://doi.org/10.69653/iah20232

Submitted: 10-06-2023
Accepted: 19-12-2023
Published: 20-12-2023

***Correspondence**
Facultad de Teología,Universidad Peruana Unión, Km 19, Carretera Central, Lima 15033, Lima, Perú.

Abstract

Background: The integration of artificial intelligence (AI) in healthcare is revolutionizing work practices and improving medical care through technologies such as decision support systems and surgical robots. However, it faces challenges such as cost, accessibility, and the need for specialized training.
Objective: This study aimed to adapt the Brief Job Performance Scale (BJPS) and evaluate the psychometric properties of the Brief Artificial Intelligence Job Performance Scale (BAIJPS) in Peruvian nurses, considering the specific demands and integration of AI in their work practices.
Methods: An instrumental design with convenience sampling was employed, including 199 nurses ($M=35.27$, $SD=8.5$). Analyses included descriptive statistics, Confirmatory Factor Analysis (CFA), and measurement invariance.
Results: The unidimensional factorial structure of the scale showed a good fit ($CFI = 0.97$, $TLI = 0.95$, $RMSEA = 0.08$, $SRMR = 0.03$), with high factor loadings and internal consistency ($\alpha = 0.96$, $\omega = 0.96$). Measurement invariance by gender confirmed that the BAIJPS is applicable equally among men and women.
Conclusions: The BAIJPS is a valid and reliable tool for assessing job performance in nurses in the context of AI integration, reflecting both task and contextual performance. This supports the implementation of policies to improve training and adaptation of nurses to the use of AI technologies, ensuring accurate and culturally relevant measurements.

Keywords:

Nurses, Performance, Artificial Intelligence, Validation.

Introduction

In the current era, the integration of artificial intelligence (AI) in healthcare is profoundly transforming work practices and the quality of medical care. AI technologies, such as clinical decision support systems, surgical robots, and automated diagnostic tools, are revolutionizing healthcare by enhancing diagnostic accuracy and optimizing workflows. These innovations promise not only to elevate the quality of medical care but also to alleviate the workload of healthcare staff, allowing them to focus more on critical interactions and decision-making (Zeng, 2024). However, the implementation of AI in healthcare is not without challenges, including cost, accessibility, and issues of professional liability

(Saleh Ibrahim et al., 2022). Moreover, the perception of primary care professionals regarding AI suggests a generally positive reception, although there is a strong demand for specific training on its practical use (Catalina et al., 2023). Beyond its role in diagnosis and treatment, AI has also significantly improved efficiency in data management and clinical decision-making, reflected in reduced hospital readmission rates and increased treatment adherence. Nevertheless, the potential benefits of these technologies must be balanced with careful attention to ethical and legal aspects, such as data privacy and medical ethics, requiring a robust regulatory framework and effective collaboration among specialists (Hee Lee & Yoon, 2021).

Additionally, AI-driven automation is altering job roles in professions such as radiology and pathology, albeit with the caveat that these technologies are not yet capable of fully replacing physicians, given the global shortage of healthcare personnel and current technological limitations (Tursunbayeva & Renkema, 2023). Lastly, it is crucial not to underestimate the role of emotional intelligence in job performance and burnout prevention among healthcare employees, a factor where AI acts more as a negative moderator, primarily impacting low-level jobs (Prentice et al., 2020).

Recent developments in AI technology, such as platforms for reading and reporting diagnostic tests, have demonstrated high levels of sensitivity and specificity, while also reducing variability in result interpretation—a crucial advancement in settings such as nursing homes and emergency departments. This enables better utilization of human resources in nursing, optimizing processes and ensuring quality care (Bermejo-Peláez et al., 2022). Advances in AI-integrated mathematical models are profoundly impacting nursing research, helping to predict behaviors and health outcomes that directly contribute to improved professional performance (Rea et al., 2020). Furthermore, the application of AI technologies in administrative tasks and clinical decision-making has allowed nurses to dedicate more time to direct patient care, enhancing job satisfaction and staff retention (Rea et al., 2020).

However, the implementation of AI also presents significant challenges. AI generative models can lead to quick but unsafe clinical decisions, such as over-triage due to the suggestion of unnecessary diagnostic tests (Saban & Dubovi, 2024). Additionally, there is limited nurse participation in technological development and a clear lack of experimental studies supporting the real impact of AI on clinical practice, highlighting the need to involve nursing staff more actively in all stages of research and development (von Gerich et al., 2022). Thus, while AI has the potential to significantly improve efficiency and quality of nursing care, addressing ethical and training challenges is crucial to ensure the effective and safe implementation of these technologies in healthcare (Clancy, 2020). Therefore, as AI redefines roles and enhances performance in nursing, successful integration of these technologies requires a careful and considered approach towards their development, implementation, and continuous evaluation.

Job performance is defined as a function of skills, motivation, and situational factors (Campbell, 1990). Furthermore, performance can be differentiated into two main dimensions: task performance, which refers to activities directly related to the production of goods and services, and contextual performance, which encompasses behaviors that contribute to the organizational and social environment (Borman & Motowidlo, 1997). In this sense, job performance with artificial intelligence (AI) involves the effectiveness and efficiency in performing tasks and achieving work objectives through the integration and use of AI technologies. This includes the use of AI tools and systems to optimize processes, make informed decisions, improve productivity, update technical knowledge, anticipate outcomes, and propose innovative solutions to emerging problems. Thus, job performance with AI is characterized by the employee's ability to adapt and leverage the opportunities offered by AI, aligning their actions with the organization's expectations and goals.

In Peru, the integration of AI-based learning health systems is revolutionizing medical care, significantly influencing the job performance of nursing staff, particularly in the treatment and prevention of chronic diseases such as kidney disease. These systems comprise health data platforms, intelligent technologies, and integrated medical care, structured in functional layers that include data collection, storage, analysis, and visualization. Implementation in specialized clinics has shown a low error rate in diagnostics and high staff satisfaction, reflecting improvements in the quality and efficiency of medical services, as well as in nurses' job performance, marking significant progress towards the modernization of healthcare in the country (Mita et al., 2023). At the population level, AI has enabled the development of applications and systems like SIAMA in Arequipa and "Perú en tus manos," which have been crucial during health emergencies such as the COVID-19 pandemic, improving public safety and large-scale health management. At the individual level, these technologies allow for deeper analysis of medical data, facilitating early diagnoses and personalized treatments, as seen in the use of deep learning and neural networks to improve the diagnosis of diseases like tuberculosis, overcoming infrastructure and resource limitations (Curioso & Brunette, 2020; Videnza Consultores, 2024). Additionally, at the National Institute of Child Health San Borja, AI is transforming pediatric care through platforms that use algorithms like ChatGPT to answer frequent questions from guardians and machine learning algorithms to diagnose conditions such as scoliosis. These advancements not only improve disease diagnosis and monitoring but also facilitate access to relevant medical information, benefiting both patients and their families, and demonstrating AI's potential to enhance the quality of pediatric healthcare in Peru (Pichihua, 2023).

Various scales have been developed to assess job performance, including the Individual Work Performance Questionnaire (IWPQ) by Koopmans et al. (2012), adapted and validated in different cultural contexts such as Argentina (Gabini & Salessi, 2016), Spain (Ramos-Villagrasa et al., 2019), and Peru (Chalco-Ccapa et al., 2024); the 18-item version by Koopmans et al. (2014); and the Shortened Self-Assessment Work Performance Scale by Azevedo Andrade et al. (2020). From this latter scale, the Brief Job Performance Scale (BJPS) was adapted to the Spanish context (Morales-García, 2024), adapted from its original English version (de Azevedo Andrade et al., 2020). This scale consists of 10 items that evaluate a unidimensional structure assessing both task performance and contextual performance. The Brief Job Performance Scale (BJPS) presents itself as the best option for adaptation to an AI Job Performance Scale due to its concise structure, high reliability, and validation in various cultural contexts. In this sense, the adaptation of the scale is crucial due to the specific characteristics and challenges of the work environment, such as high demands and limited resources. AI allows for more precise and personalized measurements, fundamental in the healthcare sector. This adaptation will improve the assessment of performance in critical tasks and guide the development of effective policies, thereby enhancing the quality of healthcare services in Peru. Validating the scale in the Peruvian context will ensure culturally relevant and accurate measurements.

Therefore, this study aims to adapt and evaluate the psychometric properties of an AI Job Performance Scale for nurses.

Methods

Design and Participants

This instrumental study (Ato et al., 2013) used a convenience sampling method to select participants. An electronic sample size calculator proposed by Soper (2024) was employed, considering key factors such as the number of observed and latent variables in the proposed model, the expected effect size (λ=0.10), the statistical significance level (α=0.05), and the desired statistical power (1−β=0.80). The minimum required sample size was calculated to be 87 participants, but 199 nurses aged 24 to 62 years (M=35.27, SD=8.5) were ultimately included. Of these, 61.8% (123 participants) were women, and 38.2% (76 participants) were men. Most participants were single (53.8%), primarily from the coastal region (53.3%). In terms of educational background, 53.3% held bachelor's degrees, and 22.6% had postgraduate studies. Regarding employment status, 31.7% were permanent employees, and 21.1% were under contract (CAS).

Table 1. Descriptive Statistics

Characteristics		n	%
Sex	Female	123	61.8
	Male	76	38.2
Marital Status	Married	65	32.7
	Cohabiting	17	8.5
	Divorced	10	5.0
	Single	107	53.8
Region of Origin	Coast	106	53.3
	Jungle	26	13.1
	Highlands	67	33.7
Level of Education	Specialty	48	24.1
	Bachelor's Degree	106	53.3
	Postgraduate	45	22.6
Employment Status	Contract (CAS)	42	21.1
	Indefinite Contract (728)	29	14.6
	Permanent	63	31.7
	Substitute	15	7.5
	Third-Party	50	25.1

Instruments Job Performance

We used the Spanish version of the Brief Job Performance Scale (BJPS) (Morales-García, 2024), adapted from its original English version (de Azevedo Andrade et al., 2020). This scale consists of 10 items evaluating two key dimensions: task performance and contextual performance. A Likert-type response format is used, ranging from "1= never" to "5= always." Analyses confirmed the factorial structure of the scale, showing good fit indicators and high reliability, with Cronbach's alpha and McDonald's omega values of 0.92. Full measurement invariance between genders was also established. For this study, the BJPS was specifically adapted to evaluate Job Performance through Artificial Intelligence, following the guidelines of the International Test Commission for cross-cultural test adaptation (Muñiz et al., 2013). This adaptation was designed to be particularly applicable to nurses in Peru. Before data collection began, a pilot test was conducted with a group of 10 nurses to ensure the clarity and comprehension of the items. During this preliminary phase, no issues were detected regarding the interpretation of items or the response format of the BJPS-AI (see Table 2).

Table 2. Scale Adaptation

Initial Version	Adapted Version (BAIJPS)	English Version (BAIJPS)
Realizo adecuadamente tareas difíciles.	Utilizo herramientas de inteligencia artificial (IA) para gestionar y ejecutar tareas difíciles de manera eficiente	I use artificial intelligence (AI) tools to manage and execute difficult tasks efficiently.
Intento actualizar mis conocimientos técnicos para hacer mi trabajo.	Me esfuerzo por actualizar mis conocimientos en tecnologías de inteligencia artificial (IA) aplicadas a mi campo de trabajo.	I strive to update my knowledge in artificial intelligence (AI) technologies applied to my field of work.
Realizo mi trabajo de acuerdo con lo que la organización espera de mí.	Integro soluciones basadas en inteligencia artificial (IA) en mi trabajo para cumplir con las expectativas de la organización.	I integrate AI-based solutions into my work to meet organizational expectations.
Planifico la ejecución de mi trabajo definiendo acciones, plazos y prioridades.	Planifico la ejecución de mi trabajo utilizando sistemas de inteligencia artificial (IA) para definir acciones, plazos y prioridades.	I plan the execution of my work using AI systems to define actions, deadlines, and priorities.
Planifico acciones de acuerdo con mis tareas y prácticas habituales de trabajo.	Utilizo plataformas de inteligencia artificial (IA) para alinear mis acciones con las tareas y rutinas organizacionales.	I use AI platforms to align my actions with organizational tasks and routines.
Tomo iniciativas para mejorar mis resultados en el trabajo.	Tomo iniciativas para implementar o mejorar soluciones de inteligencia artificial (IA) que puedan aumentar la eficacia en mi trabajo.	I take initiatives to implement or improve AI solutions that can increase effectiveness in my work.
Busco nuevas soluciones para problemas que puedan surgir en mi trabajo.	Busco y propongo nuevas soluciones de inteligencia artificial (IA) para resolver problemas emergentes en mi entorno laboral.	I seek and propose new AI solutions to address emerging problems in my work environment.
Trabajo duro para realizar las tareas que me han asignado.	Me apoyo en herramientas de inteligencia artificial (IA) para optimizar la ejecución de las tareas que me son asignadas, aumentando mi productividad.	I rely on AI tools to optimize the execution of assigned tasks, increasing my productivity.
Ejecuto mis tareas anticipando sus resultados.	Implemento soluciones de inteligencia artificial (IA) para prever los resultados de mis tareas y ajustar mis estrategias proactivamente.	I implement AI solutions to anticipate the outcomes of my tasks and adjust my strategies proactively.
Aprovecho las oportunidades que pueden mejorar mis resultados en el trabajo.	Exploro y adopto oportunidades emergentes en inteligencia artificial (IA) que puedan potenciar mis resultados y los de la organización.	I explore and adopt emerging AI opportunities that can enhance my results and those of the organization.

Procedure

This study was approved by the Ethics Committee of a Peruvian university, under code CEUPeU-0537. Participants were invited to complete an online questionnaire via Google Forms between January and March 2023. Confidentiality of information was ensured, and ethical principles dictated by the Declaration of Helsinki were followed before data collection. All participants were duly informed about the study's objectives and provided informed consent before participating in the survey.

Data Analysis

Initially, a descriptive analysis of the BJPS-AI items was performed, evaluating measures such as mean, standard deviation, skewness, and kurtosis. Additionally, a correlation analysis between items was conducted. Following the recommendations of George & Paul Mallery (2003), skewness (g1) and kurtosis (g2) values within the range of ±1.5 were considered acceptable. Corrected item-total correlation analysis was applied to discard items with a correlation r(i-tc) ≤ 0.2 (Kline, 2016).

Subsequently, a Confirmatory Factor Analysis (CFA) was performed to validate the unifactorial structure of the scale, using the MLR estimation method, recommended for samples that do not assume normality (Muthen & Muthen, 2017). Model fit criteria included chi-square ($\chi 2$), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI) with recommended thresholds ≥ 0.90, as well as the Root Mean Square Error of Approximation (RMSEA) and the Standardized Root Mean Square Residual (SRMR) with values ≤ 0.08 (Kline, 2016; Schumacker & Lomax, 2016). The reliability of the scale was established through Cronbach's alpha and McDonald's omega, with values above 0.70 indicating adequate internal consistency (McDonald, 1999).

To examine measurement invariance (MI) by gender, a multigroup confirmatory factor analysis was conducted, considering four levels of invariance: configural, metric, scalar, and strict. Invariance was confirmed if differences in ΔCFI were less than 0.010 (Chen, 2007). Additionally, an explanatory model was developed using structural equation modeling, applying the same fit criteria and the MLR estimator.

Statistical analyses were performed using RStudio (Allaire, 2018) with R version 4.1.1 (R Foundation for Statistical Computing, Vienna, Austria; http://www.R-project.org). The "lavaan" package (Rosseel, 2012) was used for confirmatory factor analysis and structural equation modeling, and the "semTools" package (Jorgensen et al., 2022) was employed for measurement invariance analysis.

Results

Descriptive Statistics of the Items

The descriptive analysis of the scale items showed means ranging from 2.22 to 2.65, with standard deviations between 1.09 and 1.15, indicating moderate variability in the responses. Skewness (g1) values ranged from 0.17 to 0.71, suggesting a slight skew towards higher scores for some items. Kurtosis (g2) ranged from -0.86 to -0.16, indicating that most item distributions do not have heavy tails and tend to be flatter than the normal distribution. The corrected item-total correlation (r.cor) showed high values, ranging from 0.73 to 0.89, reflecting a strong association of each item with the total scale, suggesting that all items adequately contribute to the measurement of the construct. These results support the internal consistency and unidimensional structure of the scale.

Table 3. Descriptive Statistics

Items	**M**	**DS**	**g1**	**g2**	**r.cor**
1	2.50	1.13	0.31	-0.64	0.73
2	2.65	1.15	0.17	-0.86	0.73
3	2.41	1.09	0.39	-0.59	0.85
4	2.29	1.12	0.57	-0.45	0.85
5	2.28	1.12	0.57	-0.55	0.87
6	2.40	1.15	0.46	-0.55	0.87
7	2.22	1.09	0.71	-0.16	0.80
8	2.46	1.14	0.42	-0.60	0.89
9	2.35	1.13	0.56	-0.41	0.86
10	2.35	1.1	0.50	-0.39	0.89

Confirmatory Factor Analysis and Reliability

A confirmatory factor analysis (CFA) was conducted following the established guidelines for the BAIJPS (Morales-García, 2024). The goodness-of-fit indices showed satisfactory performance: $\chi 2$ = 75.410, df = 33, p = 0.000, CFI = 0.97, TLI = 0.95, RMSEA = 0.08 (90% CI: 0.06 - 0.10), and SRMR = 0.03. All factor loadings reached significant values, exceeding the threshold of 0.50, highlighting the construct validity of the scale in measuring mental well-being. These results robustly validate the unidimensional factorial structure of the scale, confirming its applicability in evaluating the BAIJPS. Additionally, the internal consistency analysis of the scale showed highly positive results, with reliability coefficients, including Cronbach's alpha (α) and McDonald's omega (ω), both at 0.96, indicating excellent internal consistency.

Invariance

To evaluate measurement invariance by gender, a sequence of hierarchical variance models was adopted, progressively increasing the restriction: configural invariance (reference model), metric invariance (equality of factor loadings), scalar invariance (equality of factor loadings and intercepts), and strict invariance (equality of all factor loadings, intercepts, and residuals). Changes in CFI (ΔCFI) less than 0.010 and in RMSEA (ΔRMSEA) less than 0.015 (Chen, 2007) confirmed the invariance of the model across gender groups, with ΔCFI ranging from -0.004 to 0.003. This indicates that the BAIJPS measures mental well-being equivalently in men and women, which is crucial for its application in research and clinical practice involving both sexes.

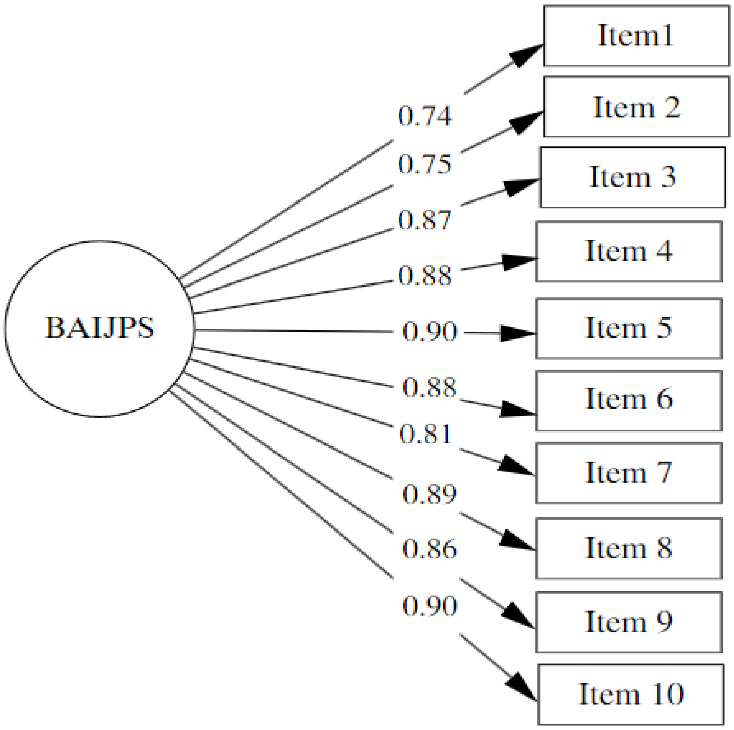

Figure 1. CFA

Table 4. Invariance by Gender

Invariance	χ2	df	p	TLI	RMSEA	SRMR	CFI	ΔCFI
Configural	108.585	66	0.001	0.956	0.080	0.035	0.968	
Metric	121.393	75	0.001	0.958	0.079	0.056	0.965	0.003
Scalar	131.341	84	0.001	0.961	0.075	0.057	0.964	0.001
Strict	135.656	94	0.003	0.970	0.067	0.057	0.968	-0.004

Discussion

In the modern era, the integration of artificial intelligence (AI) in the healthcare sector is revolutionizing work practices and the quality of medical care. Technologies such as clinical decision support systems, surgical robots, and automated diagnostic tools are enhancing diagnostic accuracy and optimizing workflows. These technologies not only improve the quality of medical care but also alleviate the workload of healthcare staff, allowing them to focus more on decision-making and critical interactions. However, the implementation of AI faces challenges such as cost, accessibility, and issues of professional liability. Although the perception of primary care staff towards AI is generally positive, there is a notable demand for specific training on its practical use. Additionally, AI is transforming data management and clinical decision-making, reducing hospital readmission rates, and increasing treatment adherence. However, it is essential to balance the benefits of these technologies with rigorous attention to ethical and legal aspects, such as data privacy and medical ethics. AI-driven automation is altering job roles in fields like radiology and pathology, although these technologies are not yet ready to completely replace doctors. Therefore, the aim of this study was to adapt and evaluate the psychometric properties of an AI Job Performance Scale in nurses in Peru.

A Confirmatory Factor Analysis (CFA) for the BAIJPS was conducted, aligning with and distinguishing itself from previous studies in its approach and results. Following the guidelines established by Morales-García (2024), the goodness-of-fit indices indicated satisfactory performance with a $\chi^2 = 75.410$, df = 33, p = 0.000, CFI = 0.97, TLI = 0.95, RMSEA = 0.08, and SRMR = 0.03. These indices suggest that the unidimensional structure of the BAIJPS is robust, providing a good fit of the theoretical model to the observed data. In comparison, Morales-García's (2024) study on the original version of the BAIJPS reported a $\chi^2 = 139.820$, df = 35, p < .001, CFI = 0.94, TLI = 0.93, RMSEA = 0.07, and SRMR = 0.03. While both studies showed good fit indices, our research presented better values in CFI and TLI, implying greater precision in adapting the scale to specific contexts where AI is applied. These findings contribute to the existing body of knowledge by validating the applicability of the BAIJPS in environments influenced by emerging technologies and highlight the need to adapt and revalidate psychometric instruments as work dynamics evolve. In this sense, the results obtained in the study underline the construct validity of the BAIJPS, with all factor loadings exceeding the threshold of 0.50. This threshold reflects each item's sufficient capacity to represent the underlying theoretical construct of job performance with AI, emphasizing the relevance of considering technological specificities in performance evaluation. This factor is critical since the incorporation of AI in the workplace can significantly alter both the required competencies and the tasks performed by employees.

In analyzing the reliability of the BAIJPS in nurses, our study revealed highly positive results in terms of internal consistency. Specifically, the reliability coefficients obtained, including Cronbach's alpha (α) and McDonald's omega (ω), both yielded a value of 0.96. These results indicate excellent internal consistency. When comparing and contrasting these findings with previous research, particularly the study conducted by Morales-García (2024), notable similarities and differences are observed. Morales-García reported reliability coefficients of α = 0.92 and ω = 0.92, which also denote robust internal consistency, though slightly lower than our study's findings. This contrast may be attributed to methodological differences, such as sample composition, the context in which the scale was administered, or even slight variations in the scale's question formulation. Our findings align with the existing body of knowledge by confirming that high levels of reliability can be achieved in job performance assessment scales within the healthcare field, particularly in studies incorporating the AI dimension.

The BAIJPS demonstrated factorial invariance by gender, using a hierarchical four-level approach: configural, metric, scalar, and strict. This approach aligns with previous studies, such as Morales-García's (2024), which evaluated the invariance of the same scale in a unidimensional structure by gender. Both studies found that the scale maintains structural and functional consistency between men and women, which is crucial for the validity of applications in clinical and research settings. However, there are notable differences in the details of the results. In our study, ΔCFI values ranged from -0.004 to 0.003, indicating excellent stability in gender comparisons across all levels of invariance. Morales-García (2024) also reported similar stability, although with slight variations in ΔCFI values, which remained within acceptable limits proposed by Chen (2007). This underscores the robustness of the BAIJPS in consistently measuring mental well-being across different populations. In this regard, the sequence of factorial invariance models adopted in this study provided a rigorous framework for evaluating the invariance of the BAIJPS. Maintaining ΔCFI and ΔRMSEA within the recommended limits (Chen, 2007) at all stages suggests that the scale is equitable for measuring mental well-being in men and women, regardless of the structure of factor loadings, intercepts, and residuals. This supports the validity of the scale for use in longitudinal and cross-sectional studies comparing different demographic groups.

Implications

The integration of artificial intelligence (AI) in nursing practice, as demonstrated by the adaptation and validation of the BAIJPS, suggests significant improvements in the efficiency and effectiveness of nurses. This advancement enables more precise task management and better data-driven decision-making. Professional practice must adapt to incorporate continuous training in AI technologies, ensuring that professionals are not only skilled in using these tools but also understand their impact on clinical and administrative outcomes. The findings highlight the need for policies that promote investment in AI in the healthcare sector, not only in terms of technology but also in staff training. Policies should aim at creating regulations that ensure the ethical and safe use of AI, as well as establishing standards for evaluating and validating AI tools in different clinical contexts. This study broadens the theoretical understanding of job performance by including the dimension of AI, highlighting how technology can influence and enhance performance in specific tasks and contexts. It contributes to the existing literature by providing a framework to evaluate the impact of AI on job performance, suggesting that future theories should consider emerging technologies as integral components of performance models.

Limitations

While the current study provides valuable data on the adaptation and validation of the BAIJPS in Peruvian nurses, it is limited by its cross-sectional design. This design involves collecting data at a single point in time, restricting the ability to infer causal relationships between AI implementation and job performance. Additionally, the long-term effects of AI on nursing practice cannot be evaluated using this approach. To address these limitations, longitudinal studies are recommended to observe the evolution of job performance with the use of AI over time. Another significant limitation of the study is the homogeneity of the sample, focused on nurses from urban areas in a single country. Differences in technological infrastructure, access to AI training, and work culture between different regions and countries can significantly influence the outcomes of AI implementation. Therefore, future research should include more diverse samples, encompassing different geographical contexts and types of healthcare institutions, to validate the generalization of the BAIJPS. Furthermore, including additional variables that may influence job performance, such as job stress, organizational support, and digital competence, could provide a more comprehensive and explanatory model of the impact of AI on job performance.

Conclusion

The adaptation and validation of the BAIJPS in nurses represent a significant advancement in evaluating the impact of technology on clinical practice and administrative management in the healthcare sector. This study underscores the crucial integration of artificial intelligence (AI) tools in job performance, highlighting not only improved efficiency in task execution but also in strategic decision-making and process management. The findings suggest that AI-mediated job performance can significantly increase productivity and job satisfaction among nurses, providing a robust metric for evaluating both task performance and contextual performance.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

WCM-G and LS-S participated in the conceptualization of the idea, WCM-G were in charge of the methodology and software. For validation, formal analysis, and research, LS-S and WCM-G. Data curation and resources were commissioned by WCM-G and LS-S. The writing of the first draft, review and editing, visualization and supervision were carried out by WCM-G. All authors have read and approved the final version of the manuscript.

References

Ato, M., López, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicologia*, *29*(3), 1038–1059. https://doi.org/10.6018/analesps.29.3.178511

Bermejo-Peláez, D., Marcos-Mencía, D., Álamo, E., Pérez-Panizo, N., Mousa, A., Dacal, E., Lin, L., Vladimirov, A., Cuadrado, D., Mateos-Nozal, J., Galán, J. C., Romero-Hernandez, B., Cantón, R., Luengo-Oroz, M., & Rodriguez-Dominguez, M. (2022). A Smartphone-Based Platform Assisted by Artificial Intelligence for Reading and Reporting Rapid Diagnostic Tests: Evaluation Study in SARS-CoV-2 Lateral Flow Immunoassays. *JMIR Public Health and Surveillance*, *8*(12). https://doi.org/10.2196/38533

Borman, W. C., & Motowidlo, S. J. (1997). Task performance and contextual performance: The meaning for personnel selection research. *Human Performance*, *10*(2). https://doi.org/10.1207/s15327043hup1002_3

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In *Handbook of Industrial and Organizational Psychology*.

Catalina, Q. M., Fuster-Casanovas, A., Vidal-Alaball, J., Escalé-Besa, A., Marin-Gomez, F. X., Femenia, J., & Solé-Casals, J. (2023). Knowledge and perception of primary care healthcare professionals on the use of artificial intelligence as a healthcare tool. *Digital Health*, *9*. https://doi.org/10.1177/20552076231180511

Chalco-Ccapa, I., Torres-Mamani, G., Morales-García, M., Flores-Saenz, A. A., Sairitupa-Sanchez, L. Z., Paredes-Saavedra, M., & Morales-García, W. C. (2024). Validation and invariance of an Individual Work Performance Questionnaire (IWPQ-P) in Peruvian Nurses. *Data and Metadata*, *3*, 259. https://doi.org/10.56294/dm2024259

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Clancy, T. R. (2020). Artificial Intelligence and Nursing: The Future Is Now. *Journal of Nursing Administration*, *50*(3). https://doi.org/10.1097/NNA.0000000000000855

Curioso, W. H., & Brunette, M. J. (2020). Inteligencia artificial e innovación para optimizar el proceso de diagnóstico de la tuberculosis. *Revista Peruana de Medicina Experimental y Salud Pública*, *37*(3). https://doi.org/10.17843/rpmesp.2020.373.5585

Gabini, S., & Salessi, S. (2016). Validación de la escala de rendimiento laboral individual en trabajadores argentinos. *Revista Evaluar*, *16*(1), 31–44. https://doi.org/10.35670/1667-4545.V16.N1.15714

George, D., & Paul Mallery, with. (2003). SPSS for Windows Step by Step A Simple Guide and Reference Fourth Edition (11.0 update) Answers to Selected Exercises. *A Simple Guide and Reference*, *63*.

Hee Lee, D., & Yoon, S. N. (2021). Application of artificial intelligence-based technologies in the healthcare industry: Opportunities and challenges. *International Journal of Environmental Research and Public Health*, *18*(1). https://doi.org/10.3390/ijerph18010271

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). semTools: Useful tools for structural equation modeling. In *The Comprehensive R Archive Network* (R package version 0.5-6). https://cran.r-project.org/package=semTools

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (Cuarta Ed.). Guilford Press.

Koopmans, L., Bernaards, C., Hildebrandt, V., Van Buuren, S., Van Der Beek, A. J., & de Vet, H. C. w. (2012). Development of an individual work performance questionnaire. *International Journal of Productivity and Performance Management*, *62*(1), 6–28. https://doi.org/10.1108/17410401311285273

Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., De Vet, H. C. W., & Van Der Beek, A. J. (2014). Construct validity of the individual work performance questionnaire. *Journal of Occupational and Environmental Medicine*, *56*(3). https://doi.org/10.1097/JOM.0000000000000113

McDonald, R. P. (1999). *Test Theory: A United Treatment*. Lawrence Erlbaum.

Mita, V., Castillo, L., Castillo-Sequera, J. L., & Wong, L. (2023). A Learning Health-Care System for Improving Renal Health Services in Peru Using Data Analytics. *International Journal of Online and Biomedical Engineering*, *19*(14). https://doi.org/10.3991/ijoe.v19i14.41949

Morales-García, W. (2024). Adaptación y validación de una escala de Escala de Autoevaluación del Desempeño en el Trabajo en el personal de enfermería. *In Press*.

Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: Segunda edición. In *Psicothema* (Vol. 25, Issue 2). https://doi.org/10.7334/psicothema2013.24

Muthen, L., & Muthen, B. (2017). *Mplus Statistical Analysis with latent variables. User's guide* (8th ed.). Muthén & Muthén.

Pichihua, S. (2023). *Así se usa la inteligencia artificial en el Instituto Nacional de Salud del Niño San Borja*. El Peruano.

Prentice, C., Dominique Lopes, S., & Wang, X. (2020). Emotional intelligence or artificial intelligence– an employee perspective. *Journal of Hospitality Marketing and Management*, *29*(4). https://doi.org/10.1080/19368623.2019.1647124

Rea, T., Guillari, A., Sergi, C., & Serra, N. (2020). Mathematical models in nursing research. *Journal of Public Health Research*, *9*(4). https://doi.org/10.4081/jphr.2020.1952

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/JSS.V048.I02

Saban, M., & Dubovi, I. (2024). A comparative vignette study: Evaluating the potential role of a generative AI model in enhancing clinical decision-making in nursing. *Journal of Advanced Nursing*. https://doi.org/10.1111/jan.16101

Saleh Ibrahim, Y., Khalid Al-Azzawi, W., Hamad Mohamad, A. A., Nouri Hassan, A., & Meraf, Z. (2022). Perception of the Impact of Artificial Intelligence in the Decision-Making Processes of Public Healthcare Professionals. *Journal of Environmental and Public Health*, *2022*. https://doi.org/10.1155/2022/8028275

Schumacker, R. E., & Lomax, R. G. (2016). *A Beginner´s Guide to Structural Equation Modeling* (4th ed.). Taylor & Francis.

Soper, D. (2024). *A-priori Sample Size Calculator for Structural Equation Models [Software]*. Software.

Tursunbayeva, A., & Renkema, M. (2023). Artificial intelligence in health-care: implications for the job design of healthcare professionals. *Asia Pacific Journal of Human Resources*, *61*(4). https://doi.org/10.1111/1744-7941.12325

Videnza Consultores. (2024). *Inteligencia artificial para la salud en el Perú*. Gestión.

von Gerich, H., Moen, H., Block, L. J., Chu, C. H., DeForest, H., Hobensack, M., Michalowski, M., Mitchell, J., Nibber, R., Olalia, M. A., Pruinelli, L., Ronquillo, C. E., Topaz, M., & Peltonen, L. M. (2022). Artificial Intelligence -based technologies in nursing: A scoping literature review of the evidence. In *International Journal of Nursing Studies* (Vol. 127). https://doi.org/10.1016/j.ijnurstu.2021.104153