

Psychometric Analysis of the General Health Questionnaire with Chilean Inmates

Jessica Candia,¹ Yanet Quijada,² Guillermo Sanhueza,³ Liza Zúñiga⁴

¹ Independent researcher. Concepción, Chile.

² Departamento de Psicología, Facultad de Ciencias Sociales, Universidad de Concepción, Chile.

³ School of Social Work, Loyola University Chicago. Chicago, IL, USA.

⁴ Docente part time, Escuela de Ciencia Política, Universidad Diego Portales, Santiago de Chile, Chile.

Correspondence:

Guillermo Sanhueza
School of Social Work, Loyola University Chicago, 1 E Pearson St, Chicago, IL 60611, USA
Phone: +1 (312) 915-7005
Email: gsanhueza@luc.edu

Received: 2 March 2025

Accepted: 17 June 2025

Citation:

Candia, J., Quijada, Y., Sanhueza, G., & Zúñiga, L. (2026). Psychometric Analysis of the General Health Questionnaire with Chilean Inmates. *Salud Mental*, 49(4) 155-162. <https://doi.org/10.17711/SM.0185-3325.2026.21>



ABSTRACT

Introduction. Chilean prisons face various issues that not only are detrimental to human rights but also public safety. The mental health of inmates appears to be one of the most pressing, invisible aspects of incarceration. Nevertheless, this topic remains largely invisible as there are no assessments of mental health on admission or instruments adapted to the prison context, preventing assessment and intervention. **Objective.** to analyze the psychometric properties of the 12-item Goldberg General Health Questionnaire (GHQ-12) in a sample of 1,022 inmates at 20 facilities across Chile. **Method.** we conducted a descriptive cross-sectional study and evaluated the internal structure of the GHQ-12 with both exploratory and confirmatory factor analysis. We also analyzed invariance by length of incarceration to determine the reliability of the instrument. **Results.** our results showed the multidimensionality of the instrument, with a three-factor model. Invariance was demonstrated considering length of incarceration; and a similar factor structure to other studies was obtained: anxiety, social dysfunction, and loss of confidence. **Discussion and conclusion.** Chilean prisons should have higher standards of dignity and effectiveness, including measures to improve their approach to mental health. A critical step to achieving this is ensuring timely detection of cases. The GHQ-12 appears to be a brief, simple, and inexpensive instrument that could be used as a reliable screening test to identify potential cases of mental disorders for subsequent referral and clinical treatment, as well as for prevalence studies of minor mental disorders.

Keywords: Factor analyses, statistics, prisoners, prisons, psychometrics, Chile.

RESUMEN

Introducción. Las cárceles chilenas enfrentan diversos problemas que no sólo son perjudiciales para los derechos humanos sino también para la seguridad pública. La salud mental de los encarcelados, en particular, parece ser uno de los aspectos más urgentes e invisibles del encarcelamiento. Sin embargo, este tema sigue siendo invisible ya que, por ejemplo, no existen evaluaciones de salud mental al ingreso ni instrumentos adecuadamente adaptados al contexto penitenciario, lo que impide la evaluación e intervención. **Objetivo.** Analizar las propiedades psicométricas del Cuestionario de Salud General de Goldberg de 12 ítems en una muestra de 1.022 personas encarceladas de 20 cárceles a lo largo de Chile. **Método.** Realizamos un estudio descriptivo-transversal y evaluamos la estructura interna del GHQ-12 con análisis factorial exploratorio y confirmatorio. También analizamos la invariancia por tiempo de prisión para determinar la confiabilidad del instrumento. **Resultados.** Nuestros resultados muestran la multidimensionalidad del instrumento, con un modelo de tres factores; se demostró invariancia considerando tiempo de prisión; Se demostró una estructura factorial similar a otros estudios: ansiedad, disfunción social y pérdida de confianza. **Discusión y conclusión.** Las cárceles chilenas deberían tener estándares más altos de dignidad y efectividad, incluyendo diversas medidas para mejorar su abordaje de la salud mental. Un paso crítico para esto es contar con una detección oportuna y adecuada de los casos. El GHQ-12 aparece como un instrumento breve, menos complejo y más económico que podría ser utilizado como un tamiz confiable para la identificación de casos potenciales de trastornos mentales para su posterior derivación y abordaje clínico, así como de prevalencia para estudios de trastornos mentales menores.

Palabras clave: Análisis factorial, estadístico, prisioneros, cárceles, psicometría, Chile.

INTRODUCTION

The Chilean prison system faces a series of challenges such as overcrowding and the physical deterioration of its facilities (Espinoza et al., 2014). This is compounded by the limited prospects for the effective reinsertion of incarcerated individuals, which is not only detrimental to human rights but also public safety (Sanhueza, 2023). In this regard, Chile has signed various international treaties for the proper management of incarcerated people (such as the Bangkok and Mandela Rules). In terms of mental health, all of them recognize that inmates have the right to enjoy as good health as that of the general population. They establish that it is the responsibility of the state to facilitate the detection, prevention and appropriate care of mental health issues in prisons through psychosocial support and access to prescribed psychotropic drugs. In the case of serious crises, prisoners are to be temporarily transferred to psychiatric wards or hospitals (Figueroa, 2018). Nevertheless, mental health problems in Chilean prisons remain largely invisible, since there is virtually no assessment of mental health on admission (Sanhueza et al., 2025), which prevents proper management and administration of treatments (Figueroa, 2018).

Various studies have shown that this is not a new issue. A study by Osses-Paredes and Riquelme-Pereira (2013) conducted at a major prison in southern Chile reported that 45% of the prison population interviewed had a diagnosed illness, with a predomination of mental disorders. Using a different methodology, a large study undertaken at seven prisons in Chile used the Composite International Diagnostic Interview (CIDI) to determine the prevalence of mental disorders in prisons. Substance-use disorder was the most common, with a prevalence of 12.2%, followed by anxiety and affective disorders, with prevalences of 8.3% and 8.1% respectively. These three types of disorders had a higher prevalence in the prison population than the general population. Inmates suffering from major depression also had a high risk of comorbidity with other disorders, particularly illicit drug use and anxiety disorders (Mundt et al., 2013).

A study by Baier et al. (2016) designed to conduct a symptomatic follow-up of inmates with suicide risk (linked to major depressive disorder) determined remission rates and associated variables at a prison in Chile. Findings showed that, a year after being incarcerated, 44% of inmates continued to meet the diagnostic criteria for this disorder and that comorbidity with post-traumatic stress disorder was associated with the persistence of the diagnosis. Moreover, 47% of this population presented a high risk of suicide on admission, with only 14% continuing in this situation after a year, with the risk being concentrated in the first months of admission. Another study reported that suicide accounts for 25.2% of deaths and is the second leading cause of death in this context, followed by acts of violence between inmates (Escobar et al., 2008). Ceballos-Espinoza

et al. (2016) reported 132 suicides between 2006–2015 at Chilean penitentiary facilities, with depression being cited as the main trigger (43.2%).

In these studies, length of incarceration was a key factor in the mental health status of inmates. The study by Mundt et al., (2013) determined that those who had been in prison for 12 months or less had higher overall rates of mental disorders. In the case of the study by Ceballos-Espinoza et al. (2016), most suicides occurred during the first year of incarceration (65.15%), coinciding with the findings of Baier et al. (2016).

These figures highlight the importance of disorders that may be less visible, such as major depressive disorder, yet entail significant risk. Having screening tools for diagnoses in the field of mental health can be a step towards guaranteeing the right to mental health in people deprived of liberty. The General Health Questionnaire ([GHQ-12] Goldberg & Williams, 1988) is a screening instrument for detecting non-psychotic or organic mental disorders, widely used in public health (Ministry of Health, 2013). It could be used as a bridge between penitentiary services and health institutions providing a range of health services to inmates.

The psychometric equivalence of an instrument can vary depending on the cultural context and the population due to the meaning and understanding of the construct to be measured (Chahín-Pinzón, 2014). It is a key issue for variables influenced by culture such as mental health (Lewis & Araya, 1995). At the same time, since prisons house a non-representative group of the population, it should not be assumed that the instrument will behave psychometrically in the same way as it does with other populations.

The psychometric properties of the GHQ-12 have been widely studied in various countries around the world and in different populations, with most countries agreeing on the scale's acceptable reliability. However, there is no agreement on its internal structure, since the number and composition of the factors in this instrument are still under debate. In this regard, two meta-analyses of the factor structure of the GHQ-12 conclude that the instrument is essentially one-dimensional (Gnambs & Staufenbiel, 2018).

The bifactorial structure has received significant empirical support by identifying one factor associated with depression, and another with social dysfunction (Werneke et al., 2000). For example, in Argentina, Burrone and colleagues (2016) studied the psychometric properties and reliability of the GHQ-12 using a probabilistic sample of individuals attended at the primary healthcare system in Cordoba, Argentina, finding that a two-factor structure was a good solution in the studied population. Likewise, in a smaller, hospitalized-patient-based sample in Medellín, Colombia, Villa and colleagues (2013) found that a two-factor solution was a good fit. However, its structure could be seen as two parts of the same component (positive and negative) and one of the items should be removed to make

the GHQ-12 suitable for this population as it did not load properly. Others studies have determined that a three-factor structure has the best fit and utility (Campbell et al., 2003; Rivas-Diez & Sánchez-López, 2014). Still others indicate that the four-factor structure makes a better contribution to the study of minor disorders (Griffith & Jones, 2019).

Our aim was to close this knowledge gap by analyzing the psychometric properties of the GHQ-12 in people deprived of liberty at penitentiary centers in Chile. We specifically wanted to evaluate its internal structure through an exploratory and confirmatory factor analysis and subsequently analyze invariance by length of incarceration and determine the reliability of the instrument.

METHOD

Study design

This was a cross-sectional study implemented in the context of a collaborative research and development (R&D) project between [anonymized university] and the Chilean prison service (*Gendarmería de Chile*), conducted at 20 prison facilities in each of the 16 regions in the country.

Participants

Although the study initially included 1,159 respondents, the analytical sample was ultimately reduced to 1,022 participants, housed in 20 prisons across the country. The facilities were visited between June and November 2022.

Measurements

Once consent had been obtained, each participant completed the General Health Questionnaire [GHQ-12]. If an individual had difficulty reading or writing, they were assisted by a trained interviewer.

Procedure

For this study, we used a two-stage sampling procedure. First, we purposively selected a sample of 20 prison facilities, considering the largest prison in each region as the main selection criteria. We then applied a simple, random sampling strategy to select individuals within each predetermined facility, with 95% confidence. We only considered already-sentenced individuals.

Statistical analysis

Evidence exists of different response patterns across positive and negative items, but also of improved discrimination

in Likert vs. dichotomous scoring (Griffith & Jones, 2019). Items were scored using full Likert scoring but reversing positive items so that higher scores always indicated higher symptomatology. All analyses were conducted using a significance level alpha of .05. Data were processed and analyzed in R (Version 4).

To evaluate the factor structure, the sample was split in two. The first subsample ($n = 511$ incarcerated) was used for exploratory factor analysis (EFA), the solution to which was tested using confirmatory factor analysis (CFA) on the second subsample ($n = 511$ incarcerated (e.g. Ledesma et al., 2019)). Subsamples were matched to include similar number of participants from public vs. private prisons, male vs. female prisons, and regions (northern, central, or southern Chile).

EFA

To check the suitability of running EFA on the collected data, we used Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy. Horn's parallel analysis, with nonparametric resampling of 10,000 bootstrapped samples, was used to determine the initial number of factors (Horn, 1965; Ledesma et al., 2019). The initial solution was generated using least-squares factor extraction with Oblimin rotation due to potential correlations between factors. This initial solution was compared with those with one additional factor, to obtain a simple structure in which each item ideally loaded substantially on a single factor. The minimum factor loading was considered as .3 (Bruna et al., 2017; Fasce et al., 2011). Items with factor loadings of less than .3 on all factors were considered indicative of poor solutions. Moreover, determining the number of factors to retain considered VSS results and communalities, indicative of the proportion of variance in each item accounted for by the retained factors, to ensure the overall adequacy of the factor solution.

CFA

The second subsample was used to run a CFA with WLS-MV estimation (suitable for ordinal data). An exhaustive assessment of the overall fit of the EFA model was conducted, including the comparative fit index (CFI), the Tucker-Lewis Index (TLI), and the root mean square error of approximation (RMSEA) as parameters for fit estimation. CFI and TLI values of .95 or higher were considered acceptable, while RMSEA values of less than .05 were considered good and those lower than .08 as acceptable (Cova et al., 2019; Hu & Bentler, 1999). Finally, modification indices were examined to refine the model if necessary. Reliability indicators were calculated based on the final solution using Cronbach's alpha and McDonald's omega.

Measurement invariance (factor structure equivalence)

Metric invariance was assessed for configural, metric, and scalar models. Using the criterion defined by Chen (2007), metric and scalar non-invariance would be indicated by changes in CFI greater than $-.10$, together with a change in RMSEA $> .015$.

Ethical considerations

Preselected candidates were asked about their voluntary participation in the study. Once they had provided their informed consent, they were given the GHQ-12 to be self-administered in rooms equipped for this purpose in the prisons, with only the research team present, who helped participants in case of doubt. The study was approved by the Ethics Committee of the University of [anonymized], code #25-36/2021.

RESULTS

Table 1 gives the descriptive statistics for each item from the GHQ-12, whose values range from 0–3 with the following response options: no, never; average; more than average; much more than average. The highest mean was for item *ghq4*: “I have felt unhappy or depressed”.

Exploratory factor analysis

An overall KMO value of $.83$ (range for individual items: $.75$ – $.89$) showed good (“meritorious”) sampling adequacy,

and Bartlett’s test of sphericity was significant (Chi-square (66) = 2209.46, $p < .001$), indicating adequacy of the correlation matrix. Together they confirmed the appropriateness of the dataset for factor analysis. Details are available in Table 2.

Depending on the criteria, EFA yielded adequate models when two or three factors were retained. Horn’s parallel analysis with 1000 iterations, using a 95th centile estimate, suggested retaining two components or three factors. The Very Simple Structure (VSS) analysis yielded a Velicer MAP minimum of $.033$ with two factors, with the Velicer MAP value for three factors at $.036$.

Given the equivocal results regarding the VSS and parallel analyses, we conducted analyses on both 2- and 3- factor solutions, using a polychoric correlation matrix, weighted least squares, and Oblimin rotation. The two-factor solution explained 55% of the variance. There were no items showing cross-loading above $.3$, and the correlation coefficient for the two factors was $.22$. Although the 3-factor solution explained 63% of the variance, the solution included some cross-loadings [*ghq3*]. Factor correlation coefficients were $.12$ (factors 1 / 2), $.41$ (factors 1 / 3), and $.30$ (factors 2 / 3). The item loadings for both solutions, the proportion of variance explained, and model evaluation criteria are shown in Table 3.

Confirmatory factor analysis

The following models were compared: a 1-dimensional model, which assumes that all scale variance can be explained by a single factor. We then tested the fit for two- and three-factor models using CFA. As can be seen in Table 4, CFI and TLI values, as well as the RMSEA, suggested that

Table 1
Descriptive statistics for *ghq* items for full sample ($n = 1022$). (positive items are reversed so that higher scores always indicate higher symptomatology)

	Mean	SD	Median	Min	Max	Range	Skew	Kurtosis	SE
<i>ghq1</i>	1.45	1.08	1	0	3	3	.17	-1.26	.03
<i>ghq2</i>	1.36	1.03	1	0	3	3	.29	-1.05	.03
<i>ghq3</i>	.95	1.03	1	0	3	3	.78	-0.62	.03
<i>ghq4</i>	1.62	.99	1	0	3	3	.16	1.15	.03
<i>ghq5</i>	.59	.96	0	0	3	3	1.47	.87	.03
<i>ghq6</i>	.43	.89	0	0	3	3	1.96	2.51	.03
<i>ghq7</i>	1.21	1.03	1	0	3	3	.10	-1.34	.03
<i>ghq8</i>	1.27	1.15	1	0	3	3	.19	-1.44	.04
<i>ghq9</i>	.96	.98	1	0	3	3	.40	-1.24	.03
<i>ghq10</i>	1.30	1.08	1	0	3	3	.09	-1.35	.03
<i>ghq11</i>	1.15	1.00	1	0	3	3	.14	-1.33	.03
<i>ghq12</i>	1.31	1.03	2	0	3	3	-.06	-1.30	.03

Source: Compiled by the authors.

Table 2

EFA Descriptors (positive items are reversed so that higher scores always indicate higher symptomatology)

Item	tiempo_dic = 1 (< 2y) mean	tiempo_dic = 1 (< 2y) SD	tiempo_dic = 2 (> 2y) mean	tiempo_dic = 2 (> 2y) SD	t-test (Welch)
ghq1	1.54	1.11	1.43	1.05	$t(314.65) = 1.08, p = .28$
ghq2	1.45	1.05	1.38	1.04	$t(328.28) = .68, p = .50$
ghq3	1.16	1.14	.89	1.01	$t(296.78) = 2.65, p < .01$
ghq4	1.72	1.04	1.59	.95	$t(305.53) = 1.32, p = .19$
ghq5	.67	1.00	.57	.96	$t(318.53) = 1.09, p = .28$
ghq6	.52	.98	.40	.86	$t(296.21) = 1.40, p = .16$
ghq7	1.38	1.09	1.13	1.04	$t(318.20) = 2.46, p = .01$
ghq8	1.50	1.15	1.08	1.15	$t(331.76) = 3.83, p < .001$
ghq9	1.10	1.02	.87	.93	$t(304.78) = 2.51, p = .01$
ghq10	1.43	1.16	1.16	1.05	$t(304.40) = 2.51, p = .01$
ghq11	1.32	1.02	1.01	.99	$t(324.75) = 3.29, p < .01$
ghq12	1.40	1.03	1.22	1.04	$t(336.17) = 1.79, p = .07$

Source: Compiled by the authors.

Table 3

EFA factor loadings for initial 2- and 3-factor solutions

Item	2-factor		3-factor		
	1	2	1	2	3
ghq1: My worries have made me lose sleep.	-.08	.74	.75	.02	.05
ghq2: I have felt constantly overwhelmed or under great stress.	-.03	.83	.84	.07	.08
ghq3: I have had the feeling that I cannot overcome my difficulties.	.13	.67	.41	.10	.38
ghq4: I have felt unhappy or depressed.	-.16	.80	.72	-.09	.16
ghq5: I have lost confidence in myself.	.19	.63	.18	.05	.68
ghq6: I have thought that I am good for nothing.	.22	.66	.06	.00	.97
ghq7: I have been able to concentrate well on what I do. (-)	.63	-.07	-.16	.58	.11
ghq8: I have felt that I play a useful/important role in life. (-)	.66	-.03	-.21	.58	.24
ghq9: I have felt capable of making decisions. (-)	.82	-.09	-.23	.75	.18
ghq10: I have been able to enjoy my daily activities. (-)	.68	.12	.20	.76	-.11
ghq11: I have been able to cope adequately with my problems. (-)	.81	.03	-.02	.80	.06
ghq12: All things considered, I feel reasonably happy. (-)	.70	.11	.22	.80	-.16
% var	27%	27%	20%	27%	16%
% cumul var	27%	55%	20%	46%	63%
Model evaluation	Chi square = 659.12 Tucker-Lewis = .716 RMSEA [CI] = .167 [.16, .18] BIC = 390.96		Chi-square = 172.85 Tucker-Lewis = .916 RMSEA [CI] = .091 [.08, .11] BIC = -32.95		

Source: Compiled by the authors.

both models were inadequate, with the three-factor solution providing a marginally better fit. We therefore used the modification indices of the 3-factor model to attempt to improve the model, by allowing item 3 to split across factors 1 and 3, yielding an acceptable solution.

We then tested a recent model from the literature using ESEM on a large sample (Griffith & Jones, 2019) to see how our solutions compared. This model showed improved goodness-of-fit indices. However, it provided a complex and therefore impractical solution with many items loading on more than one factor. Accordingly, we determined that the modified three-factor solution was the best solution.

The reliability of the modified 3-factor solution showed acceptable alphas and omegas for all subscales (f1: $\alpha = .82$, $\omega = .82$; f2: $\alpha = .81$, $\omega = .81$; f3: $\alpha = .73$, $\omega = .74$).

Measurement invariance

Results for measurement invariance for the modified three-factor model are given in Table 5. The baseline model showed an acceptable fit, with RMSEA = .074 and CFI = .964, thereby corroborating configural invariance. Although the absolute metric and scalar invariance were rejected using scaled difference chi-square tests, both were supported using the relative indices of differences in CFI (-.004 and .002, respectively) and RMSEA (.000 and -.003, respectively).

DISCUSSION AND CONCLUSION

Our main goal was to analyze the psychometric properties of the 12-item version of the Goldberg General Health Questionnaire (GHQ-12) in Chilean prisoners. The GHQ-12 showed a better solution by retaining three factors, with good internal consistency with Alpha and Omega greater than .7 for each of these, in the Chilean prison population. The three factors explained 63% of the variance. However, this solution includes a cross-loading in factors 1 (anxiety) and 3 (loss of confidence) of item ghq3: "I have had the feeling that I have not been able to overcome difficulties."

The results are in line with several authors reporting the multidimensionality of the instrument. The factors themselves are similar in structure to those reported in other studies (Graetz, 1991; Rivas-Diez & Sanchés-López, 2014). Factor 1 contains six positive items and includes elements of anxiety and depression. Factor 2, with four negative items, includes aspects of anhedonia and social dysfunction, while Factor 3, with just two items, is related to loss of confidence or self-esteem. We can therefore summarize the three factors as anxiety, social dysfunction, and loss of confidence.

Authors have attributed the diversity of factors obtained to methodological decisions such as the criteria for factor retention (Daradkeh, et. al. 2001; López-Castedo & Domínguez, 2010), and the inadequate treatment of variables measured

Table 4
Model fit indices for 2- and 3- factor CFA solutions

CFA Sample inmates						
Sample Model	<i>n</i> = 511 Chi-sq (scaled)	df	<i>P</i>	CFI scaled	TLI scaled	RMSEA [90%CI] scaled
1-factor	13765.761	54	< .001	.669	.569	.219 [.209, .229]
2-factor	272.874	53	< .001	.942	.928	.090 [.080, .101]
3-factor	217.165	52	< .001	.957	.945	.079 [.068, .090]
3-factor modified	166.537	50	< .001	.969	.960	.068 [.056, .079]
4-factor ESEM	57.557	38	< .05	.995	.991	.032 [.012, .048]

Source: Compiled by the author.

Table 5
GHQ-12 invariance fit indices by length of incarceration

CFA Sample, 3-factor modified model									
Sample Model	<i>n</i> = 511 Chi-sq	Diff Chi-sq	df	<i>p</i>	TLI scaled	CFI scaled	Diff CFI	RMSEA [90%CI] scaled	Diff RMSEA
Configural	240.23		100	< .001	.953	.964	n/a	.074 [.062, .086]	
Metric	264.93	24.69	110	< .001	.953	.961	-.004	.074 [.063, .086]	.000
Scalar	279.23	14.307	131	< .001	.962	.962	.002	.067 [.056, .078]	-.003

Source: Compiled by the author.

Note: χ^2 Scaled Chi-Square test statistic, $\Delta\chi^2$ Scaled difference Chi-Square test statistic.

TLI Tucker Lewis index.

CFI Comparative fit index, Δ CFI difference in CFI between models.

RMSEA root mean square error of approximation, Δ RMSEA Difference in RMSEA between models.

with interval scales (Rivas-Diez & Sanchés-López, 2014). This study used the Likert score, inverting the positive items, so that higher scores always indicate greater symptoms. It has also been attributed to the particular characteristics of the sample, as in the case of this study, corresponding to inmates.

The invariance demonstrated by the scale considering the length of incarceration factor, more or less than two years, provides evidence of the relevance of using the GHQ-12 considering this factor, since prevalence studies have shown a greater presence of mental disorders, particularly during the first months of incarceration (Mundt et al., 2013; Baier et al. 2016).

We argue that Chilean prisons should have higher standards of dignity and effectiveness, including various measures to improve their approach to mental health. A critical step here is the timely detection of cases. This requires suitable instruments that will allow screening processes, given the high prevalence of mental health disorders found in various studies on the prison population in Chile (Osses-Paredes & Riquelme-Pereira, 2013; Mundt et al., 2013). In this regard, the GHQ-12 is a brief, simple and inexpensive instrument that could be used as a reliable screener for identifying potential cases of mental disorders for subsequent referral and clinical treatment, as well as for prevalence studies of minor mental disorders.

One of the strengths of this study is its suitable sample size for achieving the proposed objectives. It provides evidence of the validity of the instrument for screening a population as specific as that found in the context of incarceration, and for serving as a bridge between health services providing care to people deprived of liberty. At the same time, the instrument proved to be invariant over time, in view of the evidence showing that mental health is more severely affected during the first months in prison.

Limitations include the fact that given the high prevalence of mental health problems in prisons reported in national studies, both clinical and non-clinical populations may have been included, which could have influenced results. Nevertheless, the GHQ-12 has demonstrated psychometric invariance between these two types of populations in previous studies (Ruiz et al., 2017). This study does not provide information on test-retest reliability, which could be important in the prison population, given the alterations in mental health that occur during the first year of incarceration (Baier et al. 2016; Ceballos-Espinoza et al., 2016).

However, this instrument has proven to be consistent and reliable in annual application periods in the general population (Pevalin, 2000), with our findings indicating that the instrument is invariant regardless of the length of incarceration. Future studies could explore test-retest reliability by applying repeated measures in this population. Moreover, in Chile, it was reported that the GHQ-12 detected a higher number of false positive cases the lower the educational attainment of a population (Araya et al., 1992). Determining a cut-off score for this instrument through an appropriate diagnostic accuracy study, including the use of gold standard measures,

could contribute to more accurate, contextualized care referrals. This is particularly true if the majority of this population has lower educational attainment than the national average (Ministry of Justice and Human Rights, 2024). International studies using this instrument to determine the prevalence of symptoms and psychopathological risk have used different cut-off scores. However, none of them have empirically examined the diagnostic accuracy of the instrument in this population, instead using a generic recommendation from the authors of the questionnaire (Boothby et al., 2010; Smith & Borland, 1999; McGilloway & Donnelly, 2004).

This manuscript explored an often-overlooked issue of mental health among inmates in a large sample in Chile to contribute to the literature in both the criminological and mental health fields in the region. By evaluating the psychometric properties of the GHQ-12 we sought to provide a novel, meaningful contribution to raising awareness about mental health conditions in the Chilean prison system. In addition, we used both exploratory and confirmatory factor analyses, together with invariance testing, to add methodological rigor to this important discussion.

Funding

This study was partially funded by the Agencia Nacional de Investigación y Desarrollo de Chile [ANID] through FONDEF Project #id21i10255 as well as the Loyola University Chicago Startup Fund #104502.

Conflict of interest

The authors declare they have no conflicts of interest.

Acknowledgments

This study was made possible by the collaboration of the Chilean prison service.

REFERENCES

- Araya, R., Wynn, R., & Lewis, G. (1992). Comparison of two self administered psychiatric questionnaires (GHQ-12 and SRQ-20) in primary care in Chile. *Social Psychiatry and Psychiatric Epidemiology*, 27(4), 168–173. <https://doi.org/10.1007/bf00789001>
- Baier, A., Fritsch, R., Ignatyev, Y., Priebe, S., & Mundt, A. P. (2016). The course of major depression during imprisonment – A one year cohort study. *Journal of Affective Disorders*, 189, 207–213. <https://doi.org/10.1016/j.jad.2015.09.003>
- Boothby, M. R. K., Mulholland, I., Cases, A., Carrington, K., & Bolger, T. (2010). Towards mental health promotion in prisons: the role of screening for emotional distress. *Procedia - Social and Behavioral Sciences*, 5, 90–94. <https://doi.org/10.1016/j.sbspro.2010.07.056>
- Bruna, D., Pérez, M., Bustos, C., & Núñez, J. (2017). Psychometric Properties of the Self-Regulated Learning Inventory in Chilean University Students. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica. RIDEP*, 44(2), 77–91. <https://doi.org/10.21865/RIDEP44.2.07>
- Burrone, M. S., Abeldaño, A., Lucchese, M., Susser, L., Enders, J., Alvarado, R., Valencia, E., & Fernandez, R. (2016). Evaluación psicométrica y estudio de fiabilidad del cuestionario general de salud (GHQ-12) en consultantes adultos del primer nivel de atención en Córdoba - Argentina. *Revista de la Facultad de Ciencias Médicas de Córdoba*, 72(4), 236–242. <https://doi.org/10.31053/1853.0605.v72.n4.13827>

- Campbell, A., Walker, J., & Farrell, G. (2003). Confirmatory Factor Analysis of the GHQ-12: Can I See that Again?. *Australian & New Zealand Journal of Psychiatry*, 37(4), 475–483. <https://doi.org/10.1046/j.1440-1614.2003.01208.x>
- Ceballos-Espinoza, F., Chávez-Hernández, A. M., Padilla-Gallegos, G. M., & Leenaars, A. A. (2016). Suicidio en las cárceles de Chile durante la década 2006-2015. *Revista Criminalidad*, 58(3), 101–118. <http://www.scielo.org.co/pdf/crim/v58n3/v58n3a09.pdf>
- Chahín-Pinzón, N. (2014). Aspectos a tener en cuenta cuando se realiza una adaptación de test entre diferentes culturas. *Psychologia. Avances de la disciplina*, 8(2), 109–112.
- Cova, F., Bustos, C., Rincón, P., Grandón, P., Saldivia, S., & Inostroza, C. (2019). Propiedades Psicométricas de una Forma Breve del Cuestionario de Parentalidad Alabama en Familias de Preescolares Chilenos. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 51(2). <https://doi.org/10.21865/ridep51.2.03>
- 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>
- Daradkeh, T. K., Ghubash, R., & El-Rufaie, O. E. F. (2001). Reliability, Validity, and Factor Structure of the Arabic Version of the 12-Item General Health Questionnaire. *Psychological Reports*, 89(1), 85–94. <https://doi.org/10.2466/pr0.2001.89.1.85>
- Escobar, J., González, B., Quiroz, M., & Rodríguez, G. (2008). Decesos en el sistema penitenciario: una descripción preliminar por tipo de muertes. *Revista de Estudios Criminológicos y Penitenciarios*, 13, 35–54.
- Espinoza, O., Martínez, F., & Sanhueza, G. (2014). El Sistema penitenciario y su Impacto en los Derechos Humanos de las Personas Privadas de Libertad (pp. 245–289). *En Informe sobre Derechos Humanos en Chile*. Ediciones Universidad Diego Portales.
- Fasce H, E., Pérez V, C., Ortiz M, L., Parra P, P., & Matus B, O. (2011). Estructura factorial y confiabilidad de la escala de aprendizaje autodirigido de Fisher, King & Tague en alumnos de medicina chilenos. *Revista médica de Chile*, 139(11), 1428–1434. <https://doi.org/10.4067/s0034-98872011001100006>
- Figuerola, F. (2018). Discapacidad psicosocial, salud mental y prisión en Chile. Aproximaciones hacia una sensibilización con enfoque de derechos humanos. *Revista de Derecho*, 35(2), 11–26. <https://doi.org/10.21703/issn0717-0599/2018.n35-01>
- Gnams, T., & Staufienbiel, T. (2018). The Structure of the General Health Questionnaire (GHQ-12): Two Meta-Analytic Factor Analyses. *Health Psychology Review*, 12(2), 179–194. <https://doi.org/10.1080/17437199.2018.1426484>
- Goldberg, D. P., & Williams, P. (1988). *A users' guide to the general health questionnaire*. GL Assessment.
- Graetz, B. (1991). Multidimensional properties of the General Health Questionnaire. *Social Psychiatry and Psychiatric Epidemiology*, 26(3), 132–138. <https://doi.org/10.1007/bf00782952>
- Griffith, G. J., & Jones, K. (2019). Understanding the population structure of the GHQ-12: Methodological considerations in dimensionally complex measurement outcomes. *Social Science & Medicine*, 243, Article 112638. <https://doi.org/10.1016/j.socscimed.2019.112638>
- Horn, J. L. (1965). A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/bf02289447>
- Hu, L., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Ledesma, R., Ferrando, P., & Tosi, J. (2019). Uso del análisis factorial exploratorio en RIDEP. Recomendaciones para autores y revisores. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 52(3). <https://doi.org/10.21865/ridep52.3.13>
- Lewis, G., & Araya, R. I. (1995). Is the General Health Questionnaire (12 item) a culturally biased measure of psychiatric disorder? *Social Psychiatry and Psychiatric Epidemiology*, 30(1), 20–25. <https://doi.org/10.1007/bf00784430>
- López-Castedo, A., & Domínguez, J. (2010). Exploratory Factor Analysis and Psychometric Properties of the General Health Questionnaire in Spanish Adolescents. *Psychological Reports*, 107(1), 120–126. <https://doi.org/10.2466/06.08.13.pr0.107.4.120-126>
- McGilloway, S., & Donnelly, M. (2004). Mental illness in the UK criminal justice system: A police liaison scheme for Mentally Disordered Offenders in Belfast. *Journal of Mental Health*, 13(3), 263–275. <https://doi.org/10.1080/09638230410001700899>
- Ministerio de Salud (2013). *Guía Clínica AUGe Depresión en personas de 15 años y más* (2ª ed.). https://diprece.minsal.cl/wrdprss_minsal/
- Ministerio de Justicia y Derechos Humanos (2025). *Caracterización de Personas Privadas de Libertad*. Gendarmería de Chile. https://www.gendarmeria.gob.cl/car_personas_pp.html
- Mundt, A. P., Alvarado, R., Fritsch, R., Poblete, C., Villagra, C., Kastner, S., & Priebe, S. (2013). Prevalence Rates of Mental Disorders in Chilean Prisons. *PLoS ONE*, 8(7), e69109. <https://doi.org/10.1371/journal.pone.0069109>
- Osses-Paredes, C., & Riquelme-Pereira, N. (2013). Situación de salud de reclusos de un centro de cumplimiento penitenciario, Chile. *Revista Española de Sanidad Penitenciaria*, 15(3), 98–104.
- Pevalin, D. J. (2000). Multiple applications of the GHQ-12 in a general population sample: an investigation of long-term retest effects. *Social Psychiatry and Psychiatric Epidemiology*, 35(11), 508–512. <https://doi.org/10.1007/s001270050272>
- Sanhueza, G., Candia, J., & Zúñiga, L. (2025). Mental Health in the Chilean Incarcerated Population: A Screening Approach. *International Journal of Environmental Research and Public Health*, 22(2), Article 268. <https://doi.org/10.3390/ijerph22020268>
- Sanhueza, G. E. (2023). Human Rights in Chilean Prisons: Advances and Pending Challenges from a Social Work Perspective. *Journal of Human Rights and Social Work*, 8(4), 439–448. <https://doi.org/10.1007/s41134-023-00276-x>
- Smith, C., & Borland, J. (1999). Minor psychiatric disturbance in women serving a prison sentence: The use of the General Health Questionnaire in the estimation of the prevalence of non-psychotic disturbance in women prisoners. *Legal and Criminological Psychology*, 4(2), 273–284. <https://doi.org/10.1348/135532599167905>
- Rivas-Diez, R., & Sánchez-López, M. d. P. (2014). Propiedades psicométricas del Cuestionario de Salud General (GHQ-12) en población femenina chilena. *Revista Argentina de Clínica Psicológica*, 23(3), 251–260.
- Ruiz, F. J., García-Beltrán, D. M., & Suárez-Falcón, J. C. (2017). General Health Questionnaire-12 validity in Colombia and factorial equivalence between clinical and nonclinical participants. *Psychiatry Research*, 256, 53–58. <https://doi.org/10.1016/j.psychres.2017.06.020>
- Villa G, I. C., Zuluaga Arboleda, C. & Restrepo Roldán, L. F. (2013). Psychometric Properties of the General Health Goldberg GHQ-12 Questionnaire Applied at a Hospital Facility in the City of Medellín. *Avances en Psicología Latinoamericana*, 31(3), 532–545. <https://revistas.urosario.edu.co/index.php/apl/article/viewFile/1656/2206>
- Werneke, U., Goldberg, D. P., Yalcin, I., & Üstün, B. T. (2000). The Stability of the Factor Structure of the General Health Questionnaire. *Psychological Medicine*, 30(4), 823–829. <https://doi.org/10.1017/s0033291799002287>