Research Square

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

Psychometric properties of the Chilean version of the Householder Bushfire Risk Perception Scale (HBRPS-4)

José Sandoval-Díaz

jsandoval@ubiobio.cl

University of Bío-Bío

Camila Navarrete-Valladares

Universidad Adventista de Chile

Gonzalo Palomo-Vélez

University of Groningen

Carlos Reyes Valenzuela

Universidad Andina Simón Bolívar, Sede Ecuador

Patricio Neumann Langdon

University of Bío-Bío

Soledad Martínez Labrín

University of Bío-Bío

Carolina Alzugaray Ponce

Universidad Santo Tomás

Loreto Villagrán Valenzuela

University of Concepción

José Ventura-León

Private University of the North

Research Article

Keywords: Risk perception, wildfires, psychometrics, item response theory, measurement invariance, Chile

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11118463/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Psychometric properties of the Chilean version of the Householder Bushfire Risk Perception Scale (HBRPS-4)

José Sandoval-Díaz[1]*, Camila Navarrete-Valladares[2], Gonzalo Palomo-Vélez[3,4], Carlos Reyes Valenzuela[5], Patricio Neumann Langdon[1], Soledad Martínez Labrín[1], Carolina Alzugaray Ponce[6], Loreto Villagrán Valenzuela [7] and José Ventura-León [8]

[1] Centro de Estudios Ñuble, Universidad del Bío-Bío, Chillán, Chile

[2] Universidad Adventista de Chile, Chillán, Chile

[3] Faculty of Behavioral and Social Sciences, University of Groningen, Groningen, Netherlands

[4] Center for Climate and Resilience Research (CR2), Santiago, Chile

[5] Programa Andino de Derechos Humanos, Universidad Andina Simón Bolívar, Quito, Ecuador

[6] Escuela de Psicología, Facultad de Ciencias Sociales, Universidad Santo Tomás, Concepción, Chile

[7] Facultad de Ciencias Sociales, Departamento de Psicología, Universidad de Concepción, Concepción, Chile

[8] Universidad Privada del Norte, Lima, Peru.

* Correspondence: jsandoval@ubiobio.cl

Abstract

Background: Wildfire risk perception is a relevant component of household preparedness and protective behavior. However, brief and specific measures of this construct with psychometric evidence beyond the contexts in which they were originally developed remain scarce, particularly in Latin America. This study adapted and evaluated the psychometric properties of the Chilean version of the *Householder Bushfire Risk Perception Scale* (HBRPS-4).

Methods: A cross-sectional instrumental study was conducted with a probability sample of 1,600 adults living in six municipalities in south-central Chile exposed to moderate to very high wildfire risk. The adaptation process included linguistic and contextual review, cognitive interviews, and pilot testing. Dimensionality, item functioning, measurement precision, measurement invariance, and differential item functioning (DIF) by sex and area of residence were examined using complementary factor-analytic, network, and item response theory approaches.

Results: Findings supported an essentially unidimensional structure and adequate reliability (marginal reliability = 0.85; empirical reliability = 0.90). Item response theory analyses showed generally adequate discrimination and provided evidence on the precision of the scale across the latent trait continuum, with greater information across the central range of the latent trait and lower precision toward its upper end. Measurement invariance analyses supported score comparability by sex and, with greater caution, by

area of residence. Although statistically significant DIF was detected for some items, its magnitude was small and did not reach practical relevance.

Conclusions: The Chilean HBRPS-4 showed favorable psychometric evidence supporting the interpretation of its scores as a brief measure of household wildfire risk perception. The scale appears suitable for population-based research and group-level analyses in adults living in wildfire-prone settings. Further research should examine evidence based on relations with other variables, temporal stability, and cross-cultural comparability.

Keywords: Risk perception; wildfires; psychometrics; item response theory; measurement invariance; Chile.

Background

Wildfire risk perception and protective behavior

Wildfires represent a growing risk, especially in areas where climatic and territorial conditions converge with high human exposure (1, 2). In addition to their ecological and economic effects, these events can generate psychosocial consequences, including impacts on health, livelihoods, and community organization (3, 4). In this context, understanding how people interpret risk is relevant to explaining their preparedness and protective responses.

Risk perception refers to the subjective evaluation of the probability of a threat occurring and the severity of its potential consequences (3, 5). In the case of wildfires, it has

been linked to prior experiences, proximity to vegetation fuels, and residence in rural areas (6–8). Higher levels of perceived risk have also been associated with a greater willingness to adopt preparedness and mitigation measures (7, 8). However, this relationship is not always consistent, and some of this variability may be due to differences in the conceptualization and measurement of the construct (3, 9).

Challenges in measuring wildfire risk perception

Although risk perception occupies a relevant place in wildfire research, its measurement presents conceptual and psychometric limitations. Available measures vary in scope, time horizon, and components assessed, including the likelihood of occurrence, expected consequences, and the concern associated with the event. Hall et al. (3) documented a wide heterogeneity in the indicators used, with frequent use of single questions or ad hoc measures and little consensus on the conceptualization and measurement of the construct.

Structured instruments have been developed to assess different psychological responses to wildfires and other disasters, including the Bushfire Psychological Preparedness Scale (BPSS) (10), the Disaster Risk Perception Scale (11), and the Disaster Adaptation and Resilience Scale (DARS) (12). However, these instruments assess different constructs, such as psychological preparedness, general disaster risk perception, adaptation,

and resilience. Therefore, they are not specifically designed to provide a brief measure of household wildfire risk perception.

This conceptual and methodological heterogeneity hinders comparisons between studies and the accumulation of evidence on the relationship between risk perception, preparedness, mitigation, and wildfire protective behaviors (7, 13). Furthermore, the use of measures with limited coverage of the construct or insufficient psychometric evidence can contribute to divergent results and affect the estimation of these relationships (13). This gap is particularly relevant in Latin America, where evidence on brief and specific measures of wildfire risk perception remains scarce.

The Householder Bushfire Risk Perception Scale (HBRPS-4)

To address the limitations in measuring perceived bushfire risk, Hall et al. (3) developed the Householder Bushfire Risk Perception Scale (HBRPS-4), a brief measure designed to assess perceived bushfire risk to the household. Its development began with a review of 26 studies, which identified 60 items corresponding to four domains: overall risk, probability, consequences, and concern. Through successive stages of selection and expert evaluation, this set was first reduced to 16 items and then to eight (HBRPS-8). After evaluation with 375 Australian residents, the final four-item version (HBRPS-4) was obtained (3).

The HBRPS-4 showed a unidimensional structure, high internal consistency (α = .93), evidence of convergent and discriminant validity, and high temporal stability at six weeks (r = .95) (3). Based on these results, Hall et al. (3) conceptualized household wildfire risk perception as a fuzzy concept, in which probability, consequences, and concern are closely related without constituting clearly differentiated dimensions. This conceptualization is consistent with a global interpretation of the scale scores.

In addition to its brevity, the HBRPS-4 is distinguished by specifically assessing household wildfire risk, differentiating it from related constructs such as psychological preparedness and protective behaviors (10). This distinction is relevant because risk perception can vary depending on the type of threat, prior experiences, and the territorial and sociocultural conditions in which it is configured (9, 13).

However, the evidence obtained from the Australian population does not allow us to assume that the HBRPS-4 will have the same psychometric properties in other sociocultural contexts. Its application in new populations requires examining the internal structure, the functioning of the items, the precision of the scores, and their comparability between relevant groups.

Wildfires in the Chilean context

In central Chile, wildfires have been associated with extreme weather events, changes in land use, and the expansion of inhabited areas into flammable zones (14, 15). In recent

decades, the country has experienced increasingly widespread and severe fires, with impacts on ecosystems, infrastructure, and human safety (16). During the 2024–2025 season, 6,243 wildfires were recorded, affecting more than 95,000 hectares (17).

In the south-central region, these dynamics converge with the expansion of forest plantations and residential development in traditionally rural areas, favoring the formation of urban-rural interfaces. In these areas, the proximity of homes, infrastructure, and combustible vegetation increases exposure to fire and can facilitate its spread (18–20). In Ñuble and Biobío, where the study was conducted, these conditions coexist with rural and peri-urban communities and accessibility limitations, creating distinct territorial scenarios in the context of wildfire risk (21, 22).

These territorial conditions and prior fire experiences are relevant to risk perception, influencing how people interpret the likelihood and potential consequences of the threat. Studies conducted in Ñuble and Biobío have documented persistent fear, a perception of isolation, emotional distress, and a decreased sense of security in wildfire-affected communities (19, 23). In this context, evaluating the HBRPS-4 in the population of south-central Chile allows examination of its functioning in a scenario characterized by recurrent exposure and territorial heterogeneity.

The present study

Based on the available evidence for the HBRPS-4 and the need to examine its functioning in contexts other than the one in which it was developed, this study aimed to adapt

the scale to the Chilean context and evaluate its psychometric properties in adults residing in south-central Chile.

Specifically, the study aimed to: a) examine item comprehension and the adequacy of response categories during the adaptation process; b) evaluate the scale's internal structure and degree of unidimensionality; c) examine item functioning and score accuracy along the trait continuum; and d) evaluate score comparability by sex and area of residence using measurement invariance analysis and differential item functioning (DIF).

Considering the unidimensional structure and high internal consistency reported by Hall et al. (3), the study expected to find evidence consistent with an essentially unidimensional structure and adequate levels of reliability. Furthermore, the study expanded upon the original psychometric assessment by incorporating analyses of item functioning, measurement precision across the latent trait continuum, and comparability between groups. Taken together, these analyses allow for a more comprehensive examination of the conditions underpinning the interpretation of HBRPS-4 scores in the Chilean population.

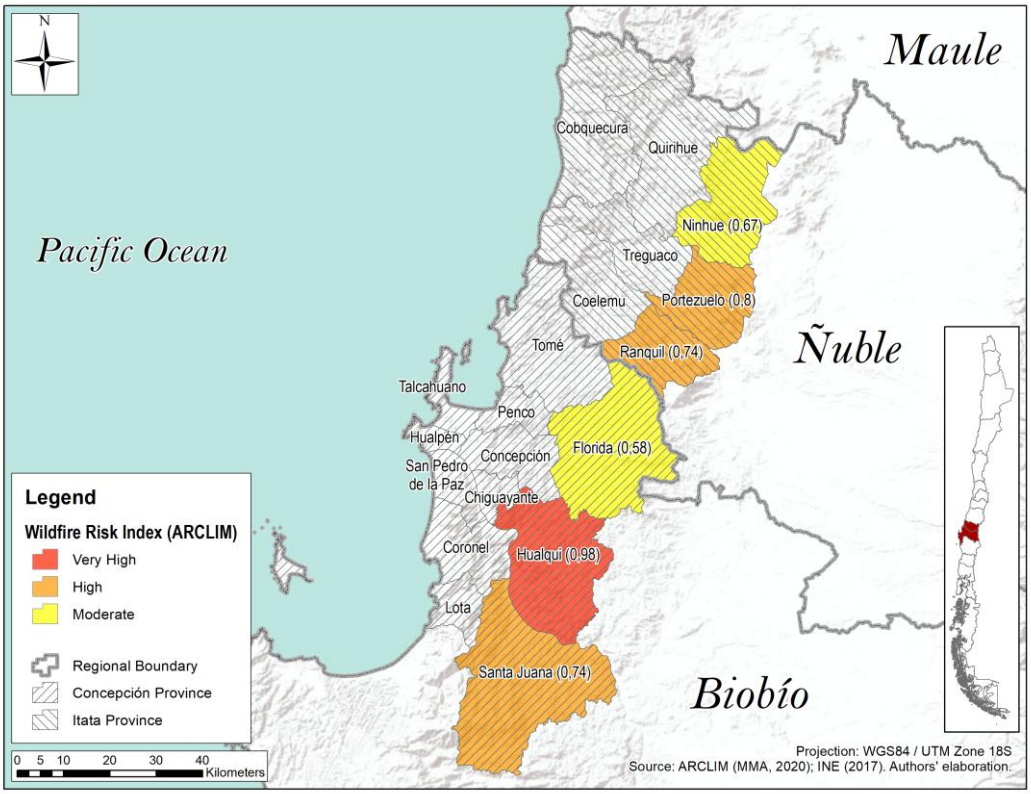
Methods

Study design and participants

A cross-sectional instrumental study was conducted to evaluate the psychometric properties of the Chilean version of the HBRPS-4. The study was carried out in six

municipalities in south-central Chile: Florida, Hualqui, and Santa Juana in the Biobío Region, and Ninhue, Portezuelo, and Ránquil in the Ñuble Region. The municipalities were selected based on the Forest Plantation Fire Risk Index from the Climate Risk Atlas for Chile (ARCLIM), with levels ranging from moderate to very high (24) (Figure 1).

Figure 1. Study area and fire risk index levels of the municipalities included in the Chilean validation of the HBRPS-4.



Note. Prepared by the authors based on the Climate Risk Atlas for Chile (ARCLIM) (24) and the census cartography of the 2017 Population and Housing Census of the National Institute of Statistics of Chile (INE).

A three-stage probability sampling design was used. The sampling frame was constructed from cartographic information from the 2017 Population and Housing Census, supplemented by digital registration of the selected localities. In the first stage, localities were randomly selected within the urban and rural areas of each municipality; in the second, dwellings were selected within each locality; and in the third, one resident aged 18 years or older was selected from each household using the Kish grid.

The analytical sample consisted of 1,600 adults: 300 residents of Florida, 250 of Hualqui, 250 of Santa Juana, 250 of Ninhue, 250 of Portezuelo, and 300 of Ránquil. Data were collected through computer-assisted personal interviewing (CAPI) conducted in participants' homes. The survey was conducted between March 6 and April 23, 2026, with a final response rate of 94%.

The sociodemographic characteristics of the sample are presented in Table 1. No data were missing for any of the four items of the HBRPS-4.

The study was conducted in accordance with applicable national and international ethical standards and was approved by the Scientific Ethics Committee of the Universidad del Bío-Bío (Scientific Ethics Report No. A009/2025, April 22, 2025), in accordance with Chilean Law No. 20,120 and its implementing regulations. All participants provided written informed consent before participating in the study.

Table 1. Sociodemographic characteristics of the analytical sample (N = 1600).

%	*n*	Category	Variable
32.7	523	Male	Sex
67.3	1077	Female	
50.0	800	Biobío	Region
50.0	800	Ñuble	
62.5	1000	Urban	Area of Residence
37.5	600	Rural	
2.3	37	C1–C2	Socioeconomic level
66.3	1061	C3	
31.4	502	D–E	

Note: The area of residence was classified according to the urban-rural definition of the National Institute of Statistics of Chile. Socioeconomic level was estimated using the ESOMAR classification adapted for Chile.

Instrument adaptation and evidence based on response processes

The adaptation of the HBRPS-4 followed recommendations for the contextual adaptation of instruments and the collection of evidence based on response processes (25, 26). Initially, the research team reviewed the items, considering their conceptual equivalence, semantic clarity, cultural relevance, and appropriateness for the target population among adults residing in areas exposed to wildfires.

Subsequently, cognitive interviews were conducted with 12 community leaders from the Ñuble region, selected for their knowledge of the local context and ensuring balanced representation by sex, between October 13 and 15, 2025. *Think-aloud* and *verbal probing* techniques (27) were used to examine the comprehension and interpretation of the items, the response process, and the appropriateness of the response categories.

The cognitive interviews identified difficulties in distinguishing between some of the extreme response categories of the original seven-option format. Consequently, the format was reduced to five ordinal categories, without modifying the content of the four items.

The revised version was subsequently evaluated through a pilot study conducted in the municipality of Ninhue between November 27 and 30, 2025. Sixty adults aged 18 years or older participated and were selected using a probability sampling design. The pilot assessed the feasibility of application, the instructions, the support materials, the administration protocol, and response times. The questionnaire was administered via computer-assisted personal interviewing (CAPI), with an average duration of 27 minutes and a response rate of 92%. No difficulties in comprehension or response were identified in the four items of the HBRPS-4.

Based on these stages, the version used in the main study was established, with four items and five ordinal response categories.

Instrument: Chilean version of the HBRPS-4

The HBRPS-4 is a four-item scale developed by Hall et al. (3) to assess household wildfire risk perception. The Chilean version retained the four items of the original instrument and used five ordinal response categories, established during the adaptation process described above.

Responses were coded from 1 to 5, with higher scores indicating higher perceived wildfire risk to the household. The wording of the items, the content assessed, and the response categories used in the study are presented in Table 2.

Table 2. Items and response categories of the Chilean version of the HBRPS-4

Code	Item	Evaluated Content	Response categories
RP1	What do you think will be the level of wildfire risk to your home in the upcoming fire season (hot season)?	Perceived risk level for the household	1 = Insignificant; 2 = Minor; 3 = Moderate; 4 = Major; 5 = Catastrophic
RP2	What do you think is the likelihood that your home will be threatened by a forest fire?	Perceived possibility of threat	1 = Insignificant; 2 = Minor; 3 = Moderate; 4 = Major; 5 = Catastrophic
RP3	If there were a forest fire near your home, how dangerous could it be to you and your household members?	Perceived danger	1 = Not dangerous at all; 2 = Low danger; 3 = Moderately dangerous; 4 = Mostly dangerous; 5 = Very dangerous
RP4	How often do you think about the threat of a forest fire to your home?	Frequency of thinking about the threat	1 = Never; 2 = Almost never; 3 = Occasionally; 4 = Almost every day; 5 = Every day

Note: The items correspond to the formulation used in the study. Response categories were coded from 1 to 5, with higher scores indicating a greater perception of the risk of wildfires in the home.

Analytical strategy

The analyses were performed in R, version 4.4.1 (28). Initially, the distribution of responses for each item was examined using category plots created with the ggstats package (29). Due to the ordinal nature of the responses, associations between items were estimated using polychoric correlations (30).

Dimensionality was examined using complementary procedures. First, a parallel analysis was performed on the polychoric correlation matrix using the psych package (31). Subsequently, exploratory graph analysis (EGA) was conducted using the walktrap algorithm on a regularized network with the graphical least absolute shrinkage

and selection operator (glasso) (32). The combination of factorial and network approaches allowed for the examination of dimensionality from complementary analytical perspectives (33). As further evidence of essential unidimensionality, a bifactor model with two group factors was estimated using the Schmid-Leiman transformation on the polychoric correlation matrix. From this model, the explained common variance (ECV), the percentage of uncontaminated correlations (PUC), hierarchical omega (ωH), and total omega (ωT) were obtained. A unidimensional confirmatory factor model for ordered categorical variables was also estimated using weighted least squares mean and variance adjusted (WLSMV), from which standardized factor loadings and their confidence intervals were obtained. The construct replicability index (34) was calculated using the IRTools package (35). Due to the small number of items, the indices derived from the bifactor model were interpreted only as supplementary evidence of the predominance of a general factor.

Following the dimensionality assessment, the items were calibrated using item response theory (IRT) with the mirt package (36). A graded response model (GRM) (37), appropriate for items with ordinal categories, was fitted. Discrimination parameters (a) and thresholds (b) were estimated using marginal maximum likelihood, along with their standard errors derived from the information matrix.

Overall GRM fit was assessed using the limited-information M_2 statistic, together with RMSEA, CFI, TLI, and SRMSR. CFI and TLI values ≥ 0.95 and SRMSR values ≤ 0.08

were considered indicative of a good fit (38). Since RMSEA and TLI can exhibit unfavorable behavior in models with few degrees of freedom, these indices were interpreted in conjunction with the other fit indicators.

Given the sensitivity of chi-square statistics to sample size, item fit was evaluated using the generalized S-X^2 statistic for polytomous items, together with its associated RMSEA and inspection of the item characteristic curves. Local independence was assessed using Yen's Q_3 residuals (40), following the relative criterion proposed by Christensen et al. (41).

At the scale level, measurement precision was assessed using marginal reliability, empirical reliability, conditional reliability, and the test information function across the latent trait continuum. As a sensitivity analysis, the GRM was re-estimated excluding the item with the lowest discrimination, and the marginal reliability and test information function were compared with those obtained in the main four-item model.

Finally, group comparability was examined through measurement invariance and DIF analyses by sex and area of residence. Invariance was assessed using multigroup confirmatory factor analysis with WLSMV, theta parameterization, and category identification according to the procedure of Wu and Estabrook (42), using the lavaan (43) and semTools (44) packages. Configural, threshold, metric, scalar, and strict models were evaluated sequentially. Invariance was considered supported when the changes between nested models met the criteria $|\Delta CFI| \leq 0.010$ and $|\Delta RMSEA| \leq 0.015$ (45).

Due to empty or sparsely populated cells in some subgroups, categories 1 and 2 were collapsed before the multigroup analyses. Category 1 had no responses in the rural area for RP2 or RP3 and only one response among men for RP2. Therefore, invariance was assessed on a four-category representation, not on the five categories used in the GRM. Differential item functioning (DIF) was assessed using ordinal logistic regression with the lordif package (46), using sex and area of residence as grouping variables. DIF detection was based on the chi-square statistic, and its magnitude was assessed by the change in McFadden's pseudo-R^2. DIF was considered practically relevant when Δpseudo-$R^2 \geq 0.02$ (47).

Results

Distribution of responses

The distribution of responses to the HBRPS-4 items is shown in Figure 2. For items RP1–RP3, between 59.6% and 66.3% of responses were concentrated in the two highest response categories. In contrast, RP4 showed a different pattern, with a higher concentration in the middle category (53.5%) and a low frequency of responses in the highest category (3.8%). All response categories were used, with no empty categories observed.



Figure 2. Distribution of response categories for HBRPS-4 items. Percentages correspond to the relative frequency of responses in each category. Higher values indicate higher perceived wildfire risk.

Evidence based on internal structure

Parallel analysis supported the retention of one factor, and exploratory graph analysis identified a single community comprising all four HBRPS-4 items (Figure 3). Complementary indices supported an essentially unidimensional structure: ECV = 0.80, PUC = 0.67, ωH = 0.76, ωT = 0.88, and H = 0.95. The unidimensional model estimated with WLSMV achieved adequate incremental indices, $\chi^2(2) = 70.86$, $p < 0.001$, CFI = 0.996, TLI = 0.988, and SRMR = 0.035; RMSEA was high (0.147, 90% CI [0.119, 0.177]), which was interpreted cautiously given the small number of degrees of freedom. Standardized factor loadings were high for RP1 (λ = 0.867, 95% CI [0.852, 0.882]), RP2 (λ = 0.966, 95% CI [0.955, 0.978]), and RP3 (λ = 0.786, 95% CI [0.766, 0.807]), while RP4 had the lowest (λ = 0.381, 95% CI [0.339, 0.423]). The upper bound of the interval

for RP2 did not reach unity, so the solution did not include improper estimates. Overall, these results supported the subsequent fitting of a one-dimensional GRM.

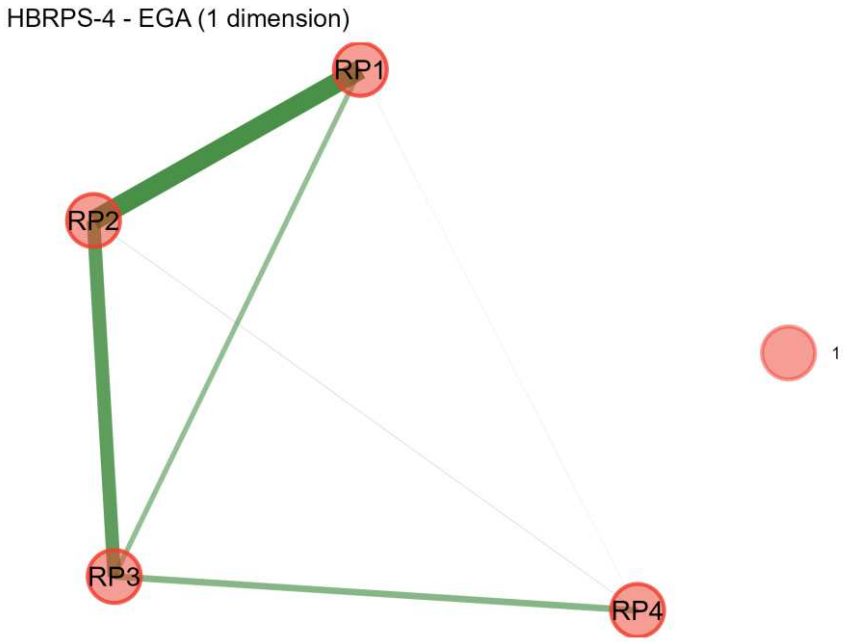


Figure 3. Structure estimated using exploratory graph analysis of the HBRPS-4. Nodes represent items and edges represent their regularized partial associations. The community estimated using EGA included all four items.

Item functioning

The overall fit of the graded response model is summarized in Table 3. Fit indices were $M_2(2) = 54.07$, $p < 0.001$, CFI = 0.981, TLI = 0.943, SRMSR = 0.042, and RMSEA = 0.128, 90% CI [0.100, 0.158]. The TLI fell slightly below the prespecified criterion and the RMSEA exceeded the conventional threshold; given the very small number of degrees of freedom, these indices were interpreted alongside the CFI and SRMSR, which met the prespecified criteria. RP2 showed the highest discrimination parameter (a =

16.192), followed by RP1 (a = 3.330) and RP3 (a = 2.343), while RP4 showed the lowest value (a = 0.705), indicating a lower capacity to discriminate between different levels of the latent trait. However, the RP2 estimate was markedly imprecise: its standard error reached 5.528, with a coefficient of variation of 34.1% and a 95% confidence interval between 5.36 and 27.03, compared to coefficients of variation of 4.4%, 4.5%, and 8.2% for RP1, RP3, and RP4. Therefore, the point estimate for RP2 indicates very high discrimination but substantial estimation uncertainty and should not be directly compared with those of the other items. The thresholds were ordered across the four items; in RP4, the extremes were located outside the practical range of the latent trait (b_1 = −6.234 and b_4 = 4.898), indicating that transitions involving the extreme response categories occurred at very low or very high levels of the latent trait.

Table 3. Parameters of the graded response model for HBRPS-4 items.

b_4	b_3	b_2	b_1	SE (a)	a	Item
1.090	-0.299	-1.314	-2.886	0.146	3.330	RP1 (P1.1)
1.103	-0.213	-1.190	-2.778	5.528	16.192	RP2 (P1.2)
0.864	-0.482	-1.687	-3.024	0.106	2.343	RP3 (P1.3)
4.898	1.300	-2.552	-6.234	0.058	0.705	RP4 (P1.4)

Note. a = discrimination parameter; SE (a) = standard error of the discrimination parameter; b_1 to b_4 = thresholds of the response categories. The parameters were estimated using marginal maximum likelihood. The standard error of RP2 implies a coefficient of variation of 34.1%, so its discrimination parameter should be interpreted as an imprecise estimate.

The item characteristic curves showed greater separation between the response probabilities of the categories in RP2 and greater overlap in RP4 (Figure 4). The RMSEA values derived from the S-X^2 statistic ranged from 0.041 to 0.086 (Table 4). No evidence

of local dependence was observed. The Q_3 residuals of the six pairs ranged from −0.571 (RP1 and RP2) to 0.139 (RP3 and RP4), with a mean of −0.183, close to the expected value of −0.333 for a four-item scale, and the relative criterion of Christensen et al. (41) placed the cutoff point at $Q_3^* = 0.322$, which no pair exceeded.

Table 4. Individual fit of the HBRPS-4 items

p	RMSEA	df	S-X²	Item
< 0.001	0.086	13	166.097	RP1 (P1.1)
< 0.001	0.044	8	32.769	RP2 (P1.2)
< 0.001	0.041	15	54.965	RP3 (P1.3)
< 0.001	0.051	23	118.729	RP4 (P1.4)

Note. df = degrees of freedom. Individual fit was assessed using the S-X² statistic and its associated RMSEA; the chi-square statistic was significant for all four items, an expected result with N = 1,600, so the interpretation relied on the RMSEA. Local independence was assessed using the Q_3 residuals of the six pairs of items, which ranged from −0.571 to 0.139, with a relative cutoff point of $Q_3^* = 0.322$ that no pair exceeded (41).

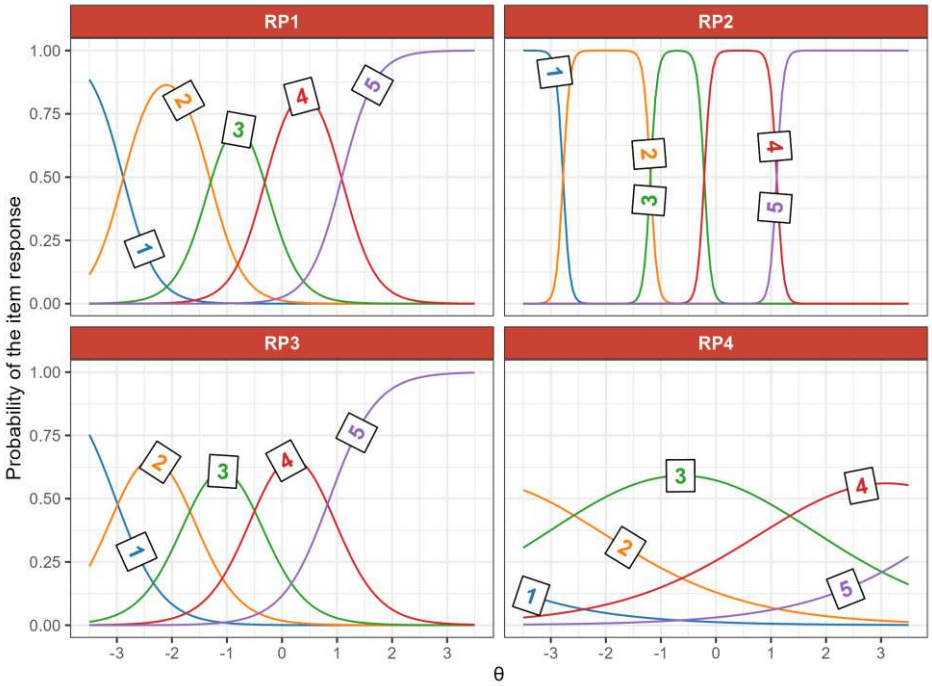


Figure 4. Characteristic curves of the items of the Chilean version of the HBRPS-4. The curves represent the probability of selecting each response category as a function of the level of the latent trait, θ. Each line corresponds to a response category of the item.

Measurement precision

The HBRPS-4 showed a marginal reliability of 0.850 and an empirical reliability of 0.900 (Table 5). Conditional reliability exceeded 0.80 in 71.6% of the interval between $\theta = -3.17$ and $\theta = 1.43$. Conditional reliability showed a multimodal pattern, with peaks up to 0.986 and intermediate lows of 0.737, 0.811, and 0.758, indicating that measurement precision was not uniform across this range. The test information function indicated greater measurement precision across the central range of the latent trait, approximately between $\theta = -2$ and $\theta = 2$, reaching its maximum around $\theta = 0$ and decreasing toward the upper end of the continuum (Figure 5).

Table 5. Measurement precision of the Chilean version of the Householder Bushfire Risk Perception Scale (HBRPS-4).

Rural	Urban	Females	Males	Total	Measurement
—	—	—	—	0.850	Marginal reliability
0.899	0.882	0.905	0.896	0.900	Empirical reliability

Note: Marginal reliability applies to the total sample. Empirical reliability was estimated using the graded response model for the total sample and by subgroup. Higher values indicate greater average precision of the latent trait estimates.

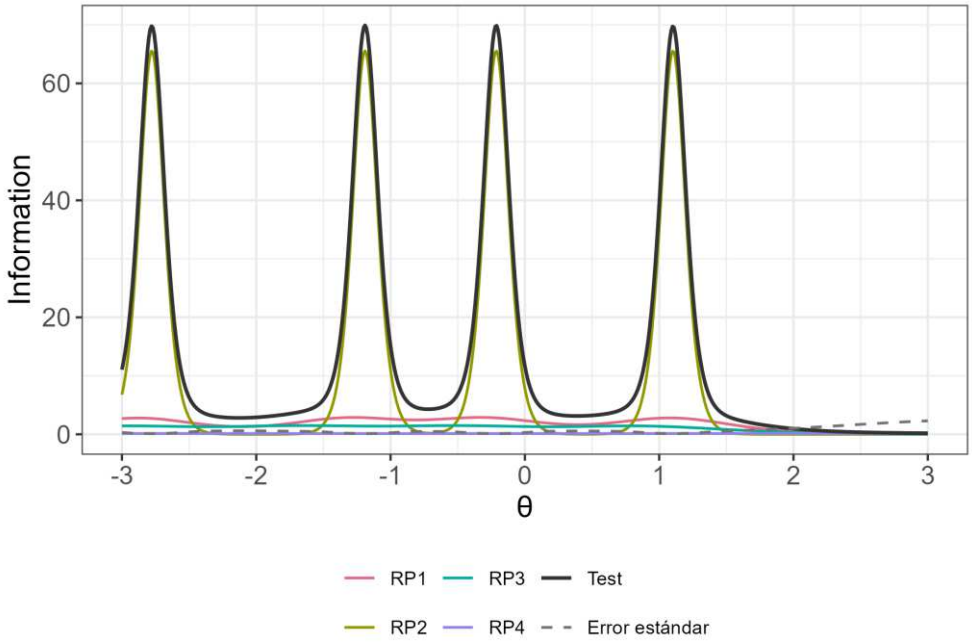


Figure 5. Item and test information functions and standard error of measurement of the HBRPS-4.

As a sensitivity analysis, the effect of excluding RP4 was examined. Compared with the four-item model, excluding RP4 resulted in lower marginal reliability and reduced test information at all three examined levels of θ, with the largest difference at θ = 0 (Table 6).

Table 6. Sensitivity analysis after the exclusion of item RP4 in the Chilean version of the Householder Bushfire Risk Perception Scale (HBRPS-4).

I(θ = 2)	I(θ = 0)	I(θ = −2)	Marginal reliability	Model
0.959	11.631	2.899	0.850	With RP4 (4 items)
0.823	8.261	2.757	0.834	Without RP4 (3 items)

Note. I(θ) = test information at the indicated level of the latent trait. Marginal reliability and test information are compared at θ = −2, θ = 0 and θ = 2 for the full model and the model without RP4.

Measurement invariance and differential item functioning

The multigroup measurement invariance analyses supported strict invariance of the HBRPS-4 across sex (Table 7). Changes in the fit indices between the configural, threshold, metric, scalar, and strict models remained within the established criteria. It should be noted that the configural model exhibited a high RMSEA (0.152), attributable to the small number of degrees of freedom; therefore, the sequence was interpreted based on incremental changes in the CFI and RMSEA (45) rather than absolute indices.

Table 7. Multigroup factorial invariance of the Chilean version of the Householder Bushfire Risk Perception Scale (HBRPS-4) according to sex.

ΔSRMR	ΔRMSEA	ΔCFI	SRMR	RMSEA	TLI	CFI	Model
-	-	-	0.039	0.152	0.986	0.995	Configural
0.000	-0.064	0.002	0.039	0.088	0.995	0.997	Thresholds
0.005	-0.014	0.000	0.044	0.074	0.997	0.997	Metric
-0.001	-0.001	-0.001	0.043	0.073	0.997	0.996	Scalar
0.006	-0.002	-0.001	0.049	0.071	0.997	0.995	Strict

Note. The fit indices for the configural, threshold, metric, scalar, and strict models are presented. ΔCFI, ΔRMSEA, and ΔSRMR correspond to the changes in fit indices between nested models. Invariance was assessed using the criteria $|\Delta CFI| \leq 0.010$ and $|\Delta RMSEA| \leq 0.015$ (45); ΔSRMR is reported for descriptive purposes.

For area of residence, the results were also compatible with strict invariance (Table 8), with maximum changes of $\Delta CFI = -0.008$ and $\Delta RMSEA = 0.011$. However, ΔSRMR reached 0.038 at the metric step and 0.031 at the strict step, and the configural model showed an RMSEA of 0.217. Thus, measurement invariance by area of residence was supported according to Chen's incremental criteria (45), but not under a joint criterion incorporating ΔSRMR, and should therefore be interpreted with caution.

Table 8. Multigroup factorial invariance of the Chilean version of the Householder Bushfire Risk Perception Scale (HBRPS-4) according to area of residence.

ΔSRMR	ΔRMSEA	ΔCFI	SRMR	RMSEA	TLI	CFI	Model
-	-	-	0.043	0.217	0.974	0.991	Configural
0.000	-0.048	-0.001	0.043	0.169	0.984	0.990	Thresholds
0.038	0.010	-0.006	0.081	0.179	0.982	0.984	Metric
0.000	-0.015	-0.001	0.081	0.164	0.985	0.983	Scalar
0.031	0.011	-0.008	0.112	0.175	0.983	0.975	Strict

Note. The fit indices for the configural, threshold, metric, scalar, and strict models are presented. ΔCFI, ΔRMSEA, and ΔSRMR correspond to the changes in fit indices between nested models. Invariance was assessed using the criteria $|\Delta CFI| \leq 0.010$ and $|\Delta RMSEA| \leq 0.015$ (45); ΔSRMR is reported for descriptive purposes.

DIF analyses detected statistically significant effects for RP4 by sex ($p < 0.001$) and for RP3 and RP4 by area of residence ($p < 0.001$ in both cases) (Table 9). However, none of these effects reached the prespecified threshold for practical relevance, as McFadden's pseudo-R^2 values were below 0.02 in all cases.

Table 9. Differential Item Functioning (DIF) Analysis of the Chilean version of the Householder Bushfire Risk Perception Scale (HBRPS-4).

Statistical DIF	p (χ^2)	Total Δpseudo-R^2	Non-uniform Δpseudo-R^2	Uniform Δpseudo-R^2	Group	Item
No	0.396	0.0005	0.0004	0.0001	Sex	RP1
No	1.000	0.0000	0.0000	0.0000	Sex	RP2
No	0.853	0.0001	0.0001	0.0000	Sex	RP3
Yes	< 0.001	0.0048	0.0016	0.0032	Sex	RP4
No	0.742	0.0001	0.0000	0.0001	Area of Residence	RP1
No	1.000	0.0000	0.0000	0.0000	Area of Residence	RP2
Yes	< 0.001	0.0106	0.0067	0.0039	Area of Residence	RP3
Yes	< 0.001	0.0143	0.0125	0.0018	Area of Residence	RP4

Note. Differential item functioning (DIF) was assessed using ordinal logistic regression. The p (χ^2) column reports the p-value of the likelihood ratio test comparing the model without DIF to the model with uniform and non-uniform DIF, and the DIF statistic column summarizes the result. The magnitude of DIF was estimated by the change in McFadden's pseudo-R^2 (Δpseudo-R^2). DIF was considered to have practical relevance when $\Delta\text{pseudo-}R^2 \geq 0.02$ (47). No item met this criterion, either by sex or by area of residence.

Discussion

The findings provide favorable evidence for interpreting the scores of the Chilean version of the Householder Bushfire Risk Perception Scale (HBRPS-4) as a brief measure of household wildfire risk perception. Taken together, the results support its use among adults living in wildfire-prone areas of south-central Chile. The study expands upon the original psychometric evaluation by incorporating evidence on item functioning through item response theory, measurement precision across the latent trait continuum, and comparability of the measure by sex and area of residence.

The adaptation process showed that retaining the original items does not guarantee that the response format will function equivalently in a new context. Cognitive interviews identified difficulties between some extreme categories of the original seven-option format, which justified its reduction to five ordinal categories. This adjustment was made without modifying the content of the four items and did not generate comprehension or response difficulties during the pilot study. In this regard, the incorporation of cognitive interviews and pilot tests provided relevant evidence based on response processes and was consistent with recommendations for the adaptation and evaluation of psychometric instruments (25–27).

The essentially unidimensional structure observed in the Chilean sample is consistent with the solution reported for the original version of the HBRPS-4 (3). In the present study, parallel analysis, exploratory graph analysis, and complementary indicators converged on the identification of a predominant general factor, supporting the interpretation of a global score for household wildfire risk perception. The convergence between

factorial and network approaches strengthens this interpretation by reducing dependence on a single analytical technique (32, 33).

This unidimensionality does not imply that the four items are interchangeable or that they represent exactly the same content. Rather, it indicates that the perceived level of risk to the home, the likelihood of a threat, the danger of a nearby fire, and the frequency of thinking about that threat all contribute to a common overall dimension. This interpretation is consistent with Hall et al.'s original conceptualization (3) of a fuzzy concept, in which related components are integrated without constituting clearly differentiated dimensions.

However, the structural similarity between the Chilean and Australian versions does not imply cross-cultural equivalence. Risk perception can vary according to norms, values, prior experiences, and sociocultural characteristics, so comparability between countries must be assessed directly through cross-cultural designs and measurement invariance analyses (48, 49). Consequently, these results should be interpreted as evidence of structural similarity with the solution reported in the original study, and not as proof of equivalence between cultural contexts.

Although the HBRPS-4 showed an essentially unidimensional structure, the items differed in their relationship to the latent trait and in their discriminative capacity. RP2 presented the highest factor loading and the highest discrimination parameter, while RP4 showed the lowest values in both indicators. These differences do not contradict

essential unidimensionality, but rather reflect differences in the strength of the items' relationships with and discrimination along the same latent dimension (50).

RP2 showed a particularly high discrimination parameter, although its estimation was markedly imprecise. The large standard error and wide confidence interval indicate that the point estimate should be interpreted with caution and should not be directly compared with those of the other items. RP2 and RP1 assess closely related content—likelihood of threat and level of risk to the household—and share the same response categories, characteristics that could contribute to the observed pattern. This explanation is tentative and requires examination in future applications. Furthermore, the local dependence analyses showed no evidence of relevant residual dependence between the two items.

RP4 showed a different pattern, characterized by lower discrimination, greater overlap between categories, and extreme thresholds. The frequency with which a person thinks about the threat of a wildfire does not necessarily correspond to a direct estimate of the likelihood of the threat or the danger posed by the event. In this sense, RP4 could incorporate a component of cognitive salience of the threat, potentially sensitive to prior experience, perceived proximity, and available information (9, 13). Likewise, the literature shows that wildfire risk perception can be shaped by territorial experiences, meanings attributed to the landscape, and memories associated with fire (51, 52). Since these

factors were not directly assessed in the present study, this interpretation should be considered hypothetical.

Despite its lower discrimination and the presence of extreme thresholds, the results do not necessarily support the removal of RP4. The sensitivity analysis showed that its exclusion slightly reduced marginal reliability and decreased test information at different levels of the latent trait. In this sense, RP4 provides additional information for measurement, although less efficiently than the other items. Considering that it also assesses distinct content—the frequency of thinking about the threat—its retention is reasonable in the Chilean version of the HBRPS-4. However, its functioning should continue to be examined in future applications, particularly in samples with different levels of exposure to and experience with wildfires.

Regarding measurement precision, the HBRPS-4 showed adequate levels of marginal and empirical reliability. However, measurement precision was not uniform across the latent trait continuum, as expected in IRT-based measurement, where precision depends on the amount of information provided at different levels of the latent trait (53). Test information was concentrated mainly across the central range of the latent trait and decreased toward its upper end.

This pattern of measurement precision has implications for the potential uses of the scale. Its four-item format facilitates brief administration and its incorporation into population studies or protocols that simultaneously assess multiple constructs. However,

lower measurement precision in some regions of the latent trait continuum requires caution when precise individual-level estimates are needed. Consequently, the HBRPS-4 appears particularly suitable for research and group-level analyses, while its use for individual decision-making would require additional evidence.

The invariance analyses provided favorable evidence across sex. Changes in CFI and RMSEA remained within the established criteria across the evaluated levels of invariance, while the DIF detected for RP4 was of small magnitude and did not reach the established criterion for practical relevance. Overall, no practically meaningful differential item functioning was observed that would compromise score comparability between men and women (45, 49).

Evidence of invariance by area of residence was less consistent. Although changes in CFI and RMSEA remained within the established incremental criteria, some absolute fit indices and changes in SRMR showed a less favorable pattern. Similarly, DIF analyses identified statistically significant effects for RP3 and RP4, although neither reached the established criterion for practical relevance. Overall, the results provide favorable evidence of comparability based on the incremental criteria, but comparability by area of residence should be interpreted with caution given the less consistent pattern observed for other fit indices.

It should also be considered that the invariance analyses were performed after collapsing the two lowest response categories due to empty or sparsely populated cells in some

subgroups. Consequently, the evidence obtained corresponds to a four-category representation and does not directly apply to the five-category format used in the graded response model. This limits the scope of conclusions regarding group comparability and should be examined in new samples that allow the full category structure to be evaluated without collapsing categories.

The practical utility of the HBRPS-4 should be interpreted considering that risk perception constitutes a relevant, but not sufficient, component of preparedness for wildfires, which also depends on broader psychological, social, and contextual processes (54). Its usefulness lies, rather, in its integration into broader assessments that incorporate other psychological, social, and contextual variables related to preparedness and protective behavior. In this sense, the scale can contribute to identifying profiles or groups with different levels of perceived risk and to examining their relationship with other relevant variables, without purporting to provide a comprehensive assessment of preparedness or protective behavior.

From this perspective, the HBRPS-4 can be integrated into broader models that jointly examine risk perception, psychological preparedness, coping, and community resilience (55). These constructs are conceptually related but not equivalent, and their joint assessment can contribute to a more comprehensive understanding of the psychological and community processes involved in wildfire risk. In this framework, the HBRPS-4

provides a specific measure of household wildfire risk perception, without purporting to provide a comprehensive measure of these broader constructs.

This study has several limitations. First, although the probability sampling design strengthens the representativeness of the sample in the selected territories, the study was conducted in six municipalities in south-central Chile, which limits the generalizability of the findings to other geographic and sociocultural contexts. Second, the cross-sectional design does not allow the temporal stability of the scores to be assessed. Third, although the study incorporated evidence based on response processes, internal structure, item functioning, measurement precision, and group comparability, other sources of validity evidence, particularly those based on relationships with external variables, should be examined in future studies.

In addition, the reduction from seven to five response categories was based on evidence obtained from cognitive interviews and the pilot study conducted in the Chilean context. Although this adaptation improved the distinguishability of the response categories, it implies that the Chilean version does not reproduce the original response format exactly. Future studies should therefore examine whether the five-category structure performs similarly in other populations and cultural contexts.

The measurement invariance results should also be interpreted considering that the two lowest response categories had to be collapsed because of empty or sparsely populated cells in some subgroups. Therefore, the invariance analyses were based on a four-

category representation rather than the five-category format used in the main psychometric analyses. Future studies with larger and more heterogeneous samples should evaluate measurement invariance while retaining the complete response-category structure.

Despite these limitations, the convergence of evidence based on response processes, internal structure, item functioning, measurement precision, and group comparability provides favorable support for the Chilean version of the HBRPS-4. The results indicate that the scale can be used as a brief measure of household wildfire risk perception in population-based and group-level research, while further evidence is needed to support its use for individual-level assessment and in other cultural contexts.

Conclusions

The Chilean version of the HBRPS-4 showed favorable psychometric evidence supporting its use as a brief measure of household wildfire risk perception among adults living in south-central Chile. Evidence based on response processes, internal structure, item functioning, and measurement precision supported an essentially unidimensional interpretation of the scale and adequate reliability, with greater measurement precision across the central range of the latent trait and lower precision toward its upper end. The results also provided favorable evidence of score comparability across sex and, with greater caution, across areas of residence. Although statistically significant DIF

was detected for some items, none of these effects reached the established criterion for practical relevance. These findings support the use of the HBRPS-4 as a brief measure for population-based and group-level research on household wildfire risk perception. Further research should examine additional sources of validity evidence, the temporal stability of scores, and the performance of the scale in other sociocultural and wildfire-risk contexts. Cross-cultural studies are also needed to determine whether the HBRPS-4 retains comparable measurement properties across different cultural settings.

Abbreviations

HBRPS-4: Householder Bushfire Risk Perception Scale; BPSS: Bushfire Psychological Preparedness Scale; DARS: Disaster Adaptation and Resilience Scale; ARCLIM: Atlas de Riesgo Climático para Chile; CAPI: Computer-Assisted Personal Interviewing; EGA: Exploratory Graph Analysis; ECV: Explained Common Variance; PUC: Percentage of Uncontaminated Correlations; WLSMV: Weighted Least Squares Mean and Variance Adjusted; IRT: Item Response Theory; GRM: Graded Response Model; DIF: Differential Item Functioning; CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; RMSEA: Root Mean Square Error of Approximation; SRMSR: Standardized Root Mean Square Residual; SRMR: Standardized Root Mean Square Residual.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with applicable national and international ethical standards and was approved by the Scientific Ethics Committee of the Universidad del Bío-Bío (Scientific Ethics Report No. A009/2025, April 22, 2025), in accordance with Chilean Law No. 20,120 and its implementing regulations. All participants provided written informed consent before participating in the study.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available and cannot be shared due to institutional and ethical restrictions established by the Scientific Ethics Committee of the Universidad del Bío-Bío.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the National Agency for Research and Development of Chile (ANID), Fondecyt Regular Project No. 1250716.

Authors' contributions

Conceptualization, J.S.-D. and C.R.-V.; Methodology, J.S.-D.; Software, J.S.-D. and J.V.-L.; Validation, J.S.-D.; Formal analysis, J.S.-D. and J.V.-L.; Investigation, J.S.-D.; Resources, J.S.-D.; Data curation, J.S.-D.; Writing—original draft preparation, J.S.-D.

and J.V.-L.; Writing—review and editing, J.S.-D., J.V.-L., C.R.-V., C.N.-V., G.P.-V., P.N.L., S.M.L., C.A.P., and L.V.V.; Visualization, J.S.-D. and C.N.-V.; Supervision, J.S.-D.; Project administration, J.S.-D.; Funding acquisition, J.S.-D. All authors read and approved the final manuscript.

Acknowledgements

The authors thank all participants and community leaders from the study territories for their valuable collaboration. The authors also acknowledge Inmetrics Consulting, which was contracted to conduct the fieldwork, including the cognitive interviews, pilot study, and main survey data collection.

References

1. McCaffrey S. Community wildfire preparedness: a global state-of-the-knowledge summary of social science research. Current Forestry Reports. 2015;1(2):81–90.

2. McCaffrey S, Toman E, Stidham M, Shindler B. Social science research related to wildfire management: an overview of recent findings and future research needs. *International Journal of Wildland Fire*. 2013;22(1):15–24.

3. Hall A, McLennan J, Marques MD, Bearman C. Conceptualising and measuring householder bushfire (wildfire) risk perception: The householder bushfire risk perception scale (HBRPS-4). International Journal of Disaster Risk Reduction. 2022;67:102667.

4. Labra Ó, Maltais D. Consecuencias de los desastres naturales en la vida de las personas: síntesis de la literatura y orientaciones para la intervención psicosocial. Revista de Trabajo Social. 2013(85):53-67. doi:10.7764/rts.85.53-67.

5. Meldrum JR, Champ PA, Brenkert-Smith H, Warziniack T, Barth CM, Falk LC. Understanding Gaps Between the Risk Perceptions of Wildland-Urban Interface (WUI) Residents and Wildfire Professionals. Risk Analysis. 2015;35(9):1746-61.

6. Champ PA, Brenkert-Smith H. Is Seeing Believing? Perceptions of Wildfire Risk Over Time. Risk Analysis. 2016;36(4):816-30.

7. Koksal K, McLennan J, Every D, Bearman C. Australian wildland-urban interface householders' wildfire safety preparations: 'Everyday life' project priorities and perceptions of wildfire risk. *International Journal of Disaster Risk Reduction*. 2019;33:142–154.

8. McNeill IM, Dunlop PD, Heath JB, Skinner TC, Morrison DL. Expecting the Unexpected: Predicting Physiological and Psychological Wildfire Preparedness from Perceived Risk, Responsibility, and Obstacles. Risk Analysis. 2013;33(10):1829-43.

9. Wachinger G, Renn O, Begg C, Kuhlicke C. The Risk Perception Paradox-Implications for Governance and Communication of Natural Hazards. Risk Analysis. 2013;33(6):1049-65.

10. Boylan JL, Lawrence C. The development and validation of the bushfire psychological preparedness scale. International Journal of Disaster Risk Reduction. 2020;47:101530.

11. Kiymis I, Kaya AA. Development of the disaster risk perception scale: evaluation of its impact on disaster preparedness. Disaster medicine and public health preparedness. 2025;19:e38.

12. First JM, Yu M, Houston JB. The Disaster Adaptation and Resilience Scale: development and validation of an individual-level protection measure. Disasters. 2021;45(4):939-67.

13. Wilson RS, Zwickle A, Walpole H. Developing a broadly applicable measure of risk perception. Risk Analysis. 2019;39(4):777-91.

14. Bowman DMJS, Moreira-Muñoz A, Kolden CA, Chávez RO, Muñoz AA, Salinas F, et al. Human–environmental drivers and impacts of the globally extreme 2017 Chilean fires. Ambio. 2019;48(4):350-62.

15. Sarricolea P, Serrano-Notivoli R, Fuentealba M, Hernández-Mora M, de la Barrera F, Smith P, et al. Recent wildfires in Central Chile: Detecting links between burned areas and population exposure in the wildland urban interface. Science of The Total Environment. 2020;706:135894.

16. Global Forest Watch. Chile: fires [Internet]. Washington, DC: World Resources Institute; [cited 2026 Sep 1]. Available from: https://www.globalforestwatch.org/dashboards/country/CHL/?category=fires.

17. Corporación Nacional Forestal (CONAF). *Informe Diario Estadístico: período de incendios forestales 2024–2025*. Santiago: CONAF; 2025.

18. Miranda A, Carrasco J, González M, Pais C, Lara A, Altamirano A, et al. Evidence-based mapping of the wildland-urban interface to better identify human communities threatened by wildfires. Environmental Research Letters. 2020;15(9):094069.

19. Toro M, Sandoval-Díaz J, Loyola C. Dimensiones Psicosociales de la Vulnerabilidad de Lugar ante el Riesgo de Incendios Forestales en Chile. Revista de Estudios Latinoamericanos sobre Reducción del Riesgo de Desastres REDER. 2026;10(1):59-74.

20. González ME, Sapiains R, Gómez-González S, Garreaud R, Miranda A, Galleguillos M, et al. *Incendios forestales en Chile: causas, impactos y resiliencia*. Santiago: Centro de Ciencia del Clima y la Resiliencia (CR)2; 2020. Available from: https://www.cr2.cl/incendios/.

21. CONAF. Manual con Medidas para la Prevención de Incendios Forestales VIII Región. Chile: CONAF; 2006.

22. Nuñez MA, Ocampo-Salazar C. Dimensiones de la gobernanza del riesgo de desastres: Una revisión sistemática. Revista de Estudios Latinoamericanos sobre Reducción del Riesgo de Desastres REDER. 2025;9(1):51-66.

23. Villagran Valenzuela L, Bilbao Ramírez MÁ, Antileo Reiman PT, Aguilera Pinto AB, Oliveira Lima IA, Lagos Castro MI, et al. Tormenta de fuego en La Palma (Chile): Significados compartidos surgidos de una intervención comunitaria postdesastre. Investigación y Desarrollo. 2022;30(2):67-95.

24. Ministerio del Medio Ambiente. Atlas de Riesgos Climáticos para Chile (ARCLIM) [Internet]. Santiago: Ministerio del Medio Ambiente, Gobierno de Chile; 2020 [cited 2026 Sep 7]. Available from: https://arclim.mma.gob.cl/.

25. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association; 2014.

26. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Frontiers in Public Health. 2018;6:149.

27. Willis G. Cognitive Interviewing: A Tool For Improving Questionnaire Design: Sage Publications; 2005.

28. R Core Team. R: A language and environment for statistical computing. Version 4.4.1. Vienna: R Foundation for Statistical Computing; 2024.

29. Larmarange J. *ggstats: Extension to 'ggplot2' for Plotting Stats*. Version 0.8.0. CRAN; 2025.

30. Özdemir HF, Toraman Ç, Kutlu Ö. The use of polychoric and Pearson correlation matrices in the determination of construct validity of Likert type scales. Turkish Journal of Education. 2019;8(3):180-95.

31. Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research. Evanston, IL, USA: Northwestern University; 2024.

32. Golino HF, Epskamp S. Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. PLOS ONE. 2017;12:e0174035.

33. Epskamp S, Rhemtulla M, Borsboom D. Generalized network psychometrics: Combining network and latent variable models. Psychometrika. 2017;82:904-27.

34. Hancock GR, Mueller RO. Rethinking construct reliability within latent variable systems. In: Cudeck R, du Toit S, Sörbom D, editors. *Structural equation modeling: Present and future—A Festschrift in honor of Karl Jöreskog*. Lincolnwood, IL: Scientific Software International; 2001. p. 195–216.

35. Ventura-León J. *IRTools (Version 1.0)* [computer software]. GitHub; 2024. Available from: https://github.com/jventural/IRTools.

36. Chalmers RP. mirt: A multidimensional item response theory package for the R environment. Journal of Statistical Software. 2012;48(6):1-29.

37. Samejima F. Estimation of latent ability using a response pattern of graded scores. *Psychometrika*. 1969;34(Suppl 1):1–97. doi:10.1007/BF03372160.

38. Maydeu-Olivares A, Joe H. Assessing approximate fit in categorical data analysis. Multivariate Behavioral Research. 2014;49:305-28.

39. Orlando M, Thissen D. Likelihood-based item-fit indices for dichotomous item response theory models. Applied Psychological Measurement. 2000;24:50-64.

40. Yen WM. Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. Applied Psychological Measurement. 1984;8:125-45.

41. Christensen KB, Makransky G, Horton M. Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. Applied Psychological Measurement. 2017;41:178-94.

42. Wu H, Estabrook R. Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. Psychometrika. 2016;81:1014-45.

43. Rosseel Y. lavaan: An R package for structural equation modeling. Journal of Statistical Software. 2012;48:1-36.

44. Jorgensen TD, Pornprasertmanit S, Schoemann AM, Rosseel Y. semTools: Useful tools for structural equation modeling. Version 0.5-6. CRAN; 2022.

45. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling. 2007;14:464-504.

46. Choi SW, Gibbons LE, Crane PK. lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. Journal of Statistical Software. 2011;39:1-30.

47. Jodoin MG, Gierl MJ. Evaluating Type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. Applied Measurement in Education. 2001;14:329-49.

48. Renn O, Rohrmann B. Cross-Cultural Risk Perception: State and Challenges. In: Renn O, Rohrmann B, editors. Cross-Cultural Risk Perception: A Survey of Empirical Studies. Boston, MA: Springer US; 2000. p. 211-33.

49. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Developmental Review. 2016;41:71-90.

50. de Ayala RJ. The Theory and Practice of Item Response Theory. New York, NY, USA: Guilford Press; 2009.

51. Reid K, Beilin R. Where's the Fire? Co-Constructing Bushfire in the Everyday Landscape. Society & Natural Resources. 2014;27(2):140-54.

52. Calvani S, Paoloni R, Foderi C, Frassinelli N, Kirschner JA, Menini A, et al. Wildfire risk perception survey: Insights from local communities in Tuscany, Italy. Fire Ecology. 2025;21:38.

53. Reise SP, Revicki DA, editors. *Handbook of item response theory modeling: Applications to typical performance assessment*. New York: Routledge; 2015.

54. Paton D. Disaster risk reduction: Psychological perspectives on preparedness. Australian Journal of Psychology. 2019;71:327-41.

55. Sandoval-Díaz J, Navarrete-Valladares C, Vega-Ortega J, Suazo-Muñoz C, Gallegos-Riquelme JP, Sandoval-Obando E, et al. Community resilience to wildfires: A

systematic review of impacts, coping strategies, indicators, and governance challenges. Progress in Disaster Science. 2025;27:100447.