



ADVANCE SOCIAL SCIENCE ARCHIVE JOURNAL

Available Online: <https://assajournal.com>

Vol. 05 No. 01. Jan-March 2026. Page#.4675-4688

Print ISSN: [3006-2497](#) Online ISSN: [3006-2500](#)Platform & Workflow by: [Open Journal Systems](#)

Urdu Translation and Psychometric Evaluation of the GELOPH-15 in a Pakistani Sample

Iqra Nawaz

Department of Psychology University of Jhang, Pakistan

Syeda Raiha*

Department of Psychology University of Jhang, Pakistan

*Corresponding Author Syeda Raiha: syedaraiha@hotmail.com**Abstract**

Gelotophobia is the predisposition to fear being laughed at and to read laughter as being ridiculed or as a social threat. The present study involved translation of GELOPH-15 from English into Urdu and investigated the evidence of internal structure and score reliability of GELOPH-15 in Pakistani sample. The translation process consisted of independent forward translation, expert reconciliation, independent back translation, multidisciplinary review and pilot testing with 20 participants. The finalized Urdu version was given online to 503 university students, college students and working adults, who were selected using convenience sampling. The hypothesized unidimensional structure of the instrument was tested by confirmatory factor analysis. Internal consistency was determined by Cronbach's alpha, composite reliability, and maximum reliability. The one-factor model demonstrated an acceptable, although not strong, fit to the data, $\chi^2(90) = 251.69$, $p < .001$, $\chi^2/df = 2.80$, CFI = .926, SRMR = .055, and RMSEA = .060. The test of close fit was statistically significant, PCLOSE = .032, indicating that the hypothesis of close approximate fit was not supported. Standardized factor loadings were modest, ranging from .33 to .50. Cronbach's alpha ($\alpha = .768$), composite reliability (CR = .771), and maximum reliability (MaxR[H] = .778) indicated acceptable score consistency. However, average variance extracted was low (AVE = .186), indicating weak convergence of the items on the latent construct. The results of this study have given some preliminary evidence for the broad unidimensional structure and reliability of Urdu GELOPH-15. Further item-level examination, test-retest assessment, validity testing using external measures, and independent replication are required before the instrument can be recommended for unrestricted use in Pakistan.

Keywords: *gelotophobia, fear of being laughed at, GELOPH-15, Urdu translation, psychometric evaluation, confirmatory factor analysis*

Introduction

The dispositional fear of being laughed at and the tendency to perceive laughter as ridicule, humiliation or social threat is called gelotophobia (Ruch & Proyer, 2008; Ruch, 2009). Individuals with elevated gelotophobia may perceive laughter, smiling, or joking as being directed against them, even when the interaction is neutral or affiliative. Gelotophobia is not just an embarrassing or uncomfortable

situation, it is a more consistent pattern of thoughts, feelings, and behaviors. These responses are often characterized by increased sensitivity to social evaluation, ongoing feelings of shame, interpersonal vigilance and avoidance of situations where laughter might happen (Proyer & Ruch, 2010; Ruch et al., 2014). Therefore, gelotophobia can impact one's perception of normal social interactions and self-assurance in interpersonal, academic, and professional environments.

Gelotophobia overlaps with social anxiety but is conceptually distinct because it specifically concerns the fear of being laughed at. The social anxiety is characterized by a general fear of negative evaluation in a wide variety of social and performance situations, while gelotophobia focuses specifically on the fear that laughter is a form of humiliation, rejection or personal devaluation (Ruch & Proyer, 2008; Titze, 2009). The construct also shares some characteristics with rejection sensitivity and paranoid ideation, but is distinct from these two. Rejection sensitivity refers to an anxious anticipation of rejection from others, and paranoid ideation refers to more general suspicions about the intentions of others. In contrast, gelotophobia is more narrowly organized around laughter and ridicule as indicators of social threat (Downey & Feldman, 1996; Freeman et al., 2002).

From a cognitive perspective, people with high levels of gelotophobia may interpret ambiguous social signals incorrectly and misread the intentions behind laughter or smiling. These interpretations are in line with cognitive models which suggest that dysfunctional self-schemas and biased information processing can affect the way socially threatening situations are perceived (Beck, 1976; Clark & Wells, 1995). Laughter can be perceived as aggression or mockery even if it is not intended to be (Ruch, 2009; Titze, 2009). Such biased interpretations can be exacerbated by previous teasing, bullying, humiliation, or social exclusion, which may lead to negative expectations for future social interactions (Platt et al., 2009; Samson et al., 2011).

Emotional characteristics and personality traits have also been linked to Gelotophobia. Individuals with greater fear of being laughed at may find themselves feeling more anxious, experiencing more shame, less self-confidence and struggling to respond appropriately to humor and interpersonal feedback (Ruch & Proyer, 2009; Samson et al., 2011). Previous studies have also shown that the presence of gelotophobia is associated with increased neuroticism, decreased extraversion, diminished emotional stability, and decreased enjoyment of humor or playful social interaction (Ruch & Proyer, 2009; Ruch et al., 2014). Humor is a social glue that helps to regulate stress, build relationships, and foster social bonds, so problems with humor understanding or humor production can exacerbate interpersonal discomfort and social isolation (Martin, 2007).

Gelotophobia can thus have repercussions beyond the discomfort experienced in specific social contexts. Fear of ridicule can be a barrier to self-expression, to confidence in personal relationships, and to participation in groups and public speaking. When people misunderstand the meaning of affiliative laughter as a sign of hostility, it can interfere with communication and contribute to feelings of mistrust, embarrassment, inferiority, or exclusion. People with higher gelotophobia scores might then be less likely to share personal information, playfully interact or form close social bonds. These patterns can over time lead to social isolation and to the lack of supportive relationships, which are essential for psychological adjustment and well-being (Baumeister & Leary, 1995; Cohen, 2004). Assessment of gelotophobia is therefore relevant for research on humor, social cognition, interpersonal functioning, emotional vulnerability, and mental health.

Sociocultural factors also can influence the expression and interpretation of gelotophobia. The norms of emotional expression, interpersonal hierarchy, social harmony, and sensitivity to evaluation can influence the understanding of laughter and social feedback (Matsumoto & Juang, 2016; Triandis, 1995). In collectivist cultures, where harmony and avoiding public embarrassment may be given special importance, people may be especially attuned to signals of ridicule or social disapproval (Kagitcibasi, 2007; Hofstede, 2011). Laughter can be a gesture of warmth and affiliation in one interaction, and a gesture of disrespect and humiliation in another. These differences in context emphasize the need to consider psychological measures in the language and population for which they are designed.

The GELOPH-15 was created as a short self-report measure of differences in fear of being laughed at (Ruch & Proyer, 2008). It consists of 15 items that reflect one underlying construct, with one overall score for gelotophobia. The unidimensionality and reliability of this scale have been confirmed by previous studies on various national and linguistic groups (Proyer et al., 2009; Ruch et al., 2014). However, the results of the psychometric measures collected in one language cannot automatically be generalized to another. Items may be interpreted differently across languages, in terms of sentence structure, idiomatic expression, and social meaning, which can impact the understanding of items by respondents and the degree to which items capture the intended construct.

Translation of a psychological instrument is more than the mere substitution of words with their obvious equivalents in another language. The translation should be linguistically accurate, meaning that it must convey the meaning of the original texts and be understandable to the members of the target group. Brislin (1970, 1976, 1980) and Van de Vijver and Tanzer (2004) have identified four methods for enhancing semantic and conceptual equivalence between the source and target languages: forward translation, back-translation, expert review, and pilot testing. Just because a translated instrument is clear in the target language does not mean that it psychometrically operates in the same way as the original instrument. A translation can seem comprehensible but be different in terms of factor structure, internal consistency, or the representativeness of the items in terms of their factor loadings. Therefore, before using this measure in research, it should be examined empirically in the translated language.

This is especially true in Pakistan, where Urdu is widely understood and used in conjunction with English and several regional languages. The use of English-language psychological instruments might pose challenges for the participants who are more comfortable to read and respond in Urdu. These differences in language can affect understanding, how people react, and how socially sensitive terms like ridicule, embarrassment, shame, and public judgment are understood. They are particularly relevant for evaluating the nature of gelotophobia as laughter is very context-dependent in terms of intention and interpersonal context. Therefore, an Urdu version of the GELOPH-15 may render it more accessible and help to gain insightful understanding of the construct in the context of Urdu speakers.

Although the role of assessment of gelotophobia in social and emotional functioning is important, psychometric evidence regarding Urdu translation of GELOPH-15 is limited. The structural characteristics and reliability of an Urdu version would lay the groundwork for future studies on the investigation of gelotophobia in Pakistani population. It would also help to diminish reliance on English-language administration and further explore the links between gelotophobia and other constructs like social anxiety, shame, interpersonal sensitivity, emotional regulation, and psychological well-being.

The purpose of the present study was to translate GELOPH-15 from English to Urdu and to assess the psychometric properties of in a Pakistani sample of university students, college students and working adults. The study specifically examined the original unidimensional structure of the Urdu version by conducting confirmatory factor analysis and calculating internal consistency and composite reliability of the Urdu version.

Method

Participants

The validation sample consisted of 503 participants from Pakistan who were recruited using convenience sampling. A total of 250 males and 253 females were included in the sample. The majority of participants were between the ages of 18 and 25 years, with 26–35 years and 36–45 years as the next two age groups. The participants were the university students, college students and working professionals. Participants needed to be able to read and understand Urdu and give informed consent.

Measure

The GELOPH-15 (Ruch & Proyer, 2008) is a 15-item self-report instrument that assesses the dispositional fear of being laughed at (gelotophobia). The instrument is designed as a unidimensional instrument and yields one composite score, with higher scores reflecting more fear of being laughed at. Participants responded to the items on a 4-point Likert-scale from 1 (= strongly disagree) to 4 (= strongly agree). No subscale scores were computed.

Procedure

Institutional authorization to conduct the study was obtained from the relevant university authorities. Before completing the questionnaire, participants received information about the study and provided electronic informed consent. Participants were informed that participation was voluntary and that they could discontinue the survey at any point before submitting their responses and their responses would be kept confidential.

The GELOPH-15 was translated into Urdu using a multistep forward–backward translation process adopted from Brislin (1970, 1976, 1980). The original English items were translated into Urdu independently by two bilingual psychologists. The translators were told about the construct measured by the scale and were asked to focus on the conceptual meaning rather than word-for-word translation.

The two Urdu translations were read by a panel of three experts alongside the original English original English version. The panel assessed semantic correctness, grammatical clarity, conceptual consistency, and comprehensibility. The differences in the translations were discussed until a preliminary Urdu translation was agreed on.

The preliminary Urdu version was independently back-translated into English by three professional linguists who had not participated in the forward-translation stage. The back-translated versions were cross-checked with the original instrument to check for any discrepancies in meaning. Discussions were held to resolve any inconsistencies.

The original, translated, and back-translated versions were reviewed by a final committee comprising two psychologists, one sociologist, one English-language specialist, and one doctoral scholar in psychology. The committee evaluated semantic and conceptual equivalence before approving the prefinal Urdu version.

Pilot Testing

The prefinal Urdu version was pilot-tested with 20 participants, including 10 males and 10 females. Participants were asked to identify words, statements, or response options that were unclear, ambiguous, or difficult to understand. The scale required approximately 7–10 minutes to complete, and no major comprehension difficulties were reported. The pilot study therefore provided preliminary evidence of the linguistic clarity and comprehensibility of the Urdu translation.

Following pilot testing, the finalized Urdu GELOPH-15 was administered through Google Forms. The first section provided information about the study and obtained electronic informed consent. The second section collected demographic information, including age, gender, educational attainment, and family structure. The third section contained the Urdu GELOPH-15 items. The survey link was distributed through WhatsApp, Instagram, and SMS. Participation was voluntary, and respondents could discontinue the survey before submitting their responses. The 20 pilot participants were excluded from the final sample of 503.

Data Analysis

Data were analysed using IBM SPSS Statistics, Version 29, and IBM SPSS Amos, Version 23. Descriptive statistics were calculated to summarize participants' demographic characteristics. Internal consistency was evaluated using Cronbach's alpha. Confirmatory factor analysis was conducted to test the hypothesized one-factor structure of the Urdu GELOPH-15. Model fit was evaluated using the chi-square statistic, chi-square-to-degrees-of-freedom ratio, comparative fit index, standardized root mean square residual, root mean square error of approximation, and the test of close fit. Composite reliability, maximum reliability, and average variance extracted were calculated from the standardized CFA estimates.

Results

The Urdu version of the GELOPH-15 was subjected to psychometric analyses to explore the factorial structure and reliability of the instrument. Confirmatory factor analysis was conducted to test whether the hypothesized unidimensional structure was supported. Internal consistency and model-based reliability estimates were also calculated. The results are reported in standardized factor loadings, model-fit indices and reliability estimates in the subsequent sections.

Table 1*Demographic Characteristics of the Participants (N = 503)*

Characteristic	Category	n	%
Age (years)	18–25	420	83.5
	26–35	75	14.9
	36–45	8	1.6
Gender	Male	250	49.7
	Female	253	50.3
Education	Matriculation	48	9.5
	Intermediate	179	35.6
	Graduation	276	54.9
Family structure	Nuclear	280	55.7
	Joint	223	44.3

Note. N = total sample size; n = frequency.

Table 1 presents the demographic characteristics of the 503 participants. Most participants were aged 18–25 years (83.5%), followed by those aged 26–35 years (14.9%) and 36–45 years (1.6%). Male and female participants were represented in nearly equal proportions, at 49.7% and 50.3%, respectively. More than half of the participants were classified in the graduation category (54.9%), followed by intermediate education (35.6%) and matriculation (9.5%). Regarding family structure, 55.7% of participants belonged to nuclear families and 44.3% belonged to joint families.

Table 2*Internal Consistency Reliability of the Urdu GELOPH-15*

Measure	No. of items	Cronbach's α
Urdu GELOPH-15	15	.768

Note. α = Cronbach's alpha.

Cronbach's alpha was used to determine the internal consistency reliability of Urdu GELOPH-15. The 15-item Urdu GELOPH-15 demonstrated acceptable internal consistency, Cronbach's α = .768. This coefficient indicates that the items showed an acceptable level of score consistency in the present sample. Evidence concerning the instrument's internal structure was examined separately through confirmatory factor analysis.

Confirmatory Factor Analysis (CFA)

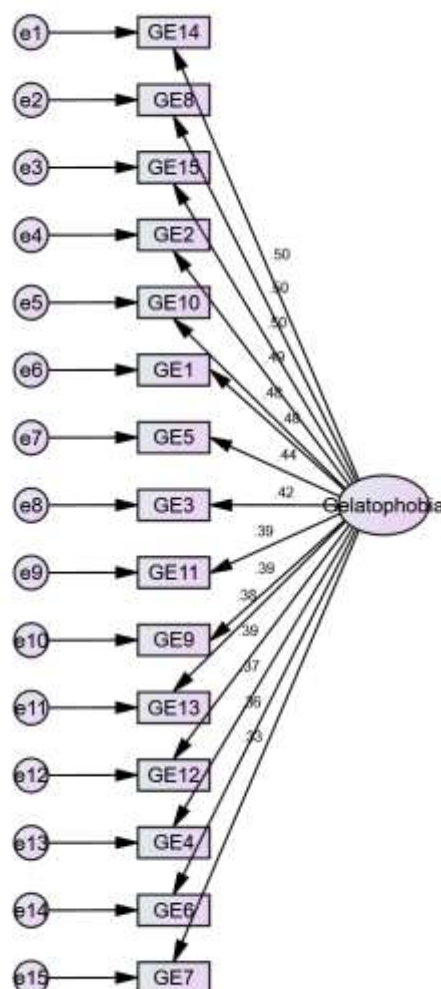
To test the unidimensionality of the Urdu GELOPH-15 scale, confirmatory factor analysis (CFA) was carried out. All 15 items were hypothesized to measure one underlying factor, the "gelotophobia", as originally conceptualized by the instrument (Ruch & Proyer, 2008; Ruch et al., 2014). Data from 503 participants were used for analysis.

The CFA model was estimated using maximum likelihood estimation. Before analysis, the data were screened for incomplete responses, duplicate submissions, univariate outliers, and departures from normality. Standardized factor loadings were examined to determine the strength of the relationships between the individual items and the latent construct. Model fit was evaluated using χ^2/df ,

CFI, SRMR, RMSEA, and PCLOSE. The total GELOPH-15 score was calculated as mean. The final CFA model and standardized factor loadings are shown in Figure 1. Table 3 shows the model-fit indices.

Figure 1

Standardized One-Factor Confirmatory Factor Analysis Model of the Urdu GELOPH-15



Note. Rectangles represent the observed items, the oval represents the latent factor, and circles represent item-specific residuals. Values shown on the single-headed paths are standardized factor loadings.

Figure 1 presents the standardized one-factor CFA model of the Urdu GELOPH-15. All 15 items were specified as indicators of one latent factor labelled gelotophobia. Standardized factor loadings ranged from .33 to .50. The highest loadings were observed for GE8, GE14, and GE15, each of which loaded at .50, followed by GE2 (.49), GE1 (.48), and GE10 (.48). The lowest loadings were observed for GE7 (.33), GE6 (.36), and GE4 (.37). Although all items loaded positively on the hypothesized factor, the

loadings were generally modest, indicating that the latent construct explained limited variance in several items.

Table 3

Model-Fit Indices for the One-Factor Urdu GELOPH-15

Fit index	Estimate	Recommended criterion	Interpretation
χ^2	251.690	—	—
df	90	—	—
χ^2/df	2.797	≤ 3.00	Acceptable
CFI	.926	$\geq .90$	Acceptable
SRMR	.055	$\leq .08$	Good
RMSEA	.060	$\leq .06$	Acceptable
pclose	.032	$> .05$	Close fit not supported

Note. CFI = comparative fit index; PCLOSE = probability of close fit; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

The hypothesized one-factor model demonstrated an acceptable fit to the data, $\chi^2 (90) = 251.69$, $p < .001$, $\chi^2/df = 2.80$. The comparative fit index was .926 and was still below the more stringent benchmark of .95, but was still considered as being within the conventional range of acceptable fit. The SRMR was .055, which was good for residual fit, while the RMSEA was .060, which was at the high end of the acceptable range. The test of close fit was statistically significant, PCLOSE = .032, which means that the hypothesis of close approximate fit was rejected. Taken together, the fit indices provided reasonable, but not strong support for the hypothesized unidimensional structure of the Urdu GELOPH-15.

Table 4

Composite Reliability and Average Variance Extracted for the Urdu GELOPH-15

Construct	CR	AVE	MaxR(H)
Gelotophobia	.771	.186	.778

Note. AVE = average variance extracted; CR = composite reliability; MaxR(H) = maximum reliability.

The Urdu GELOPH-15 had acceptable composite reliability (CR = .771) and maximum reliability (MaxR(H) = .778). The average variance extracted was, however, well below the recommended level of .50 (AVE = .186). This means that the latent factor of gelotophobia accounted for only 18.6% of the variance in its indicators. Thus, score reliability was acceptable, but item-level convergence on the latent construct was weak.

Discussion

The present study translated the GELOPH-15 into Urdu and evaluated its internal structure and score reliability in a Pakistani sample. The hypothesized one-factor model demonstrated acceptable global fit, while Cronbach's alpha and model-based reliability estimates indicated acceptable score consistency. However, standardized item loadings were modest and average variance extracted was low, suggesting that the items shared limited variance with the latent gelotophobia factor. Accordingly, the findings provide preliminary evidence for the structural plausibility and reliability of the Urdu GELOPH-15, but they do not establish comprehensive construct validity.

The one factor structure is consistent with the original conceptualization of the gelotophobia as a dispositional fear of being laughed at. Previous empirical studies found that while there are cognitive, emotional and behavioral components to the phenomenon of gelotophobia, the most parsimonious view is that it is essentially a unidimensional phenomenon (Ruch & Proyer, 2008a). The GELOPH-15 was then developed as a brief measure of the fear and its intensity, based on one composite score (Ruch & Proyer, 2008b). A large multinational study also showed that gelotophobia could be measured in various national and linguistic settings, but that cultural differences were still relevant to the interpretation of the responses (Proyer et al., 2009). Theoretical research that followed also maintained the presence of a coherent concept of gelotophobia and focused on the fact that it can manifest in different ways, depending on interpersonal and situational factors (Ruch et al., 2014).

The present results expand this literature by showing that the broad unidimensional structure of the GELOPH-15 can be translated to Urdu. However, the replication of the general factor structure does not imply that all translated items perform in the same manner as the original ones. Structural plausibility is related, but not synonymous with item quality. A translated scale might be expected to have the same number of factors, but the individual indicators will have lower correlations with the desired construct. The current results should thus be understood as an endorsement of the overall structure of the scale, and not as an absolute validation of the psychometric performance of each item.

The item-level estimates should also be taken into account when considering the global model-fit evidence. Fit indices measure the degree of fit between the covariance structure predicted by the hypothesized model and the observed data, while standardized loadings show the extent to which each item is related to the latent construct. A model can have good overall fit even if some measures have only weak correlations with the latent variable (Brown, 2015). Furthermore, model adequacy should be assessed by multiple complementary indices as each index measures a different aspect of model-data correspondence and there is no single cut-off that can be used to establish model fit (Hu & Bentler, 1999). Therefore, the evidence from the present study is consistent with the plausibility of the unidimensional model, but the evidence for the items is not considered to be strong.

The translation procedure gave preliminary evidence of the understandability and general consistency of the Urdu items with the intended meaning of the source instrument. Established translation procedures include independent forward translation, reconciliation, back-translation, multidisciplinary review, and pilot testing (International Test Commission, 2018; Sousa & Rojjanasrirat, 2011). These procedures are useful for identifying semantic errors, vague language, grammatical errors and phrases that do not convey effectively into the target language. The current guidance for scale development also focuses on the importance of pre-testing, content evaluation, structural analysis, reliability assessment, and validity testing as related, but different, phases of instrument evaluation (Boateng et al., 2018).

However, the completion of the translation procedures does not guarantee complete cross-cultural equivalence. A conceptually suitable item can be grammatically correct and be evaluated as such by experts, but have a different psychological effect on the target population. Linguistic equivalence is related to whether the wording has the intended meaning, while measurement equivalence is related to whether the translated items have similar relationships with the intended construct. Construct bias, method bias, and item bias are all potential sources of bias in cross-cultural assessment, and cannot be

eliminated by forward and backward translation (Van de Vijver & Tanzer, 2004). International Test Commission (2018) also separates test development and linguistic translation and review from the empirical evidence required to support score interpretation in different languages and cultures.

This is important for understanding the small factor loadings. The results do not necessarily indicate that the translation process failed. Instead, they suggest that the latent construct was not well represented by some Urdu items in the current sample. This may be due to residual ambiguity, item-specific content, limited variation in responses, response style, or measurement error. A few respondents also may have taken the terms "laughter," "embarrassment," "mockery," and "social evaluation" to mean something other than the specific fear of being laughed at. The possibilities are only hypotheses since CFA shows the items' statistical performance but does not show the cognitive processes that underlie the performance.

The results of cognitive interviewing would give a better foundation for determining the sources of weak item functioning. This approach can be used to explore participants' understanding of key words, their interpretation of the situations described, their recall of relevant experiences, and their choice of response categories (Willis, 2005). The respondents might be asked to rephrase the items that were weaker, describe the social contexts in which they had experienced them, and clarify their interpretation of the items in terms of ridicule, general embarrassment, social anxiety, or other kinds of experiences. Such evidence would be useful in deciding whether an item needs to be linguistically revised, culturally contextualized, or kept as is since it is a less common expression of gelotophobia and of theoretical importance.

A bilingual equivalence study would provide an additional source of evidence. Future research should administer the Urdu and English versions to bilingual participants using an appropriate counterbalanced design and examine whether the two versions produce comparable scores and item-response patterns. Previous methodological studies have demonstrated the ability of bilingual designs to detect translated items that may have different functions and to guide item revision and translation decisions (Sireci & Berberoğlu, 2000). Such evidence would provide a stronger basis for evaluating equivalence than expert review and back-translation alone.

The results for reliability and convergent validity should be interpreted separately. Internal consistency relates to the extent to which the items are consistent within the measurement, while composite reliability estimates consistency based on parameters of the fitted measurement model. Average variance extracted is a measure that tackles another problem by assessing the amount of variance in the indicators that is caused by the latent construct, not by item variance and measurement error (Fornell & Larcker, 1981). Inadequate convergent validity cannot be overcome by acceptable reliability. The present results indicate that although the individual indicators of the Urdu GELOPH-15 do not have a strong enough association with gelotophobia, the total score may be obtained in a reasonably consistent manner.

Cronbach's alpha should also be understood as a measure of limited evidence of score consistency, not as evidence of unidimensionality or validity. The number of items, the pattern of item relationships, and assumptions that may not be valid for congeneric psychological measures, are factors that affect alpha. In the modern psychometric literature, therefore, it is recommended to take into account model-based reliability estimates and item-level structural evidence, not only alpha (McNeish,

2018; Raykov, 1997). The results of the alpha and composite-reliability analyses support the overall finding of good score consistency, but do not negate the concerns raised by the modest factor loadings and low shared variance.

The modest factor loadings do not, by themselves, justify deleting items solely to improve numerical fit indices. Mechanical item removal may yield a more statistically pure solution, but narrow the conceptual scope of the construct and result in a scale that works well only in the development sample. The first step in examining each weak item should be to consider its theoretical importance, the wording of the item, how it is distributed, how it correlates with the total test when it is corrected, how much variance is left in the item after removing its correlation with the total test, and the interpretation by the target population. When deciding to keep or remove items, scale-development guidance suggests using a combination of theory, qualitative, and quantitative evidence, but not as a rule of automatic deletion of items based on a loading threshold (Boateng et al., 2018). The revised version should then be reviewed again by experts, cognitively pretested and independently psychometrically evaluated.

The interpretation is also limited by the characteristics of the sample. The majority of the participants were young and educated, and older adults and those with other levels of education and socioeconomic status were under-represented. The results should not be assumed to be representative of rural populations, individuals with lower levels of literacy, clinical populations, or individuals from other regional and linguistic backgrounds. The online self-report administration may also have resulted in uncontrolled variation in privacy, attention, reading conditions and response environments. These limitations do not invalidate the findings but restrict the populations and purposes to which they can reasonably be generalized.

The Urdu GELOPH-15 should be considered a research tool, but not a diagnostic or screening tool. The present study provides evidence for linguistic translation and review, composite reliability, internal consistency, and internal structure. It does not validate the diagnostic accuracy, criterion validity, stability of scores, or the suitability of the original classification cut points for Pakistani populations. The original cut-off criteria were derived from the source version and cannot be directly translated to an Urdu version without evidence that the scores have the same meaning and implications (Ruch & Proyer, 2008b). Validity evidence is also required to be tied to the specific interpretation and proposed use of scores, not the instrument in the abstract, in professional testing standards (American Educational Research Association et al., 2014).

The validity argument needs to be extended in future studies by relationships with external variables. Theoretically related constructs (such as fear of negative evaluation, social anxiety, shame proneness, interpersonal sensitivity, and rejection sensitivity) could be used to examine convergent evidence. Constructs to be used in evaluating discriminant evidence should be ones that are expected to stay conceptually separate from gelotophobia. The test-retest analysis is also required to ensure the Urdu scores are stable over time. These analyses would give evidence that is not possible to get from internal consistency and CFA.

In addition, the test's measurement invariance should be assessed for relevant demographic and linguistic groups prior to the interpretation of group differences. Invariance testing assesses whether a construct is measured in a similar way and has a similar meaning across groups or occasions; otherwise, the differences may not be real differences in the underlying construct (Putnick & Bornstein, 2016).

Comparisons can be made on the basis of gender, age, educational level, geographical region, urban–rural background, Urdu administration versus English administration. Differential item-functioning analysis also could be used to detect specific items that function differently among subgroups.

In general, validation is a process of evidence collection for specific interpretations and uses of scale scores. It cannot be determined by a translation process, one reliability coefficient or a single confirmatory factor analysis. The standards for contemporary tests focus on evidence related to test content, response processes, internal structure, relationships with external variables, fairness, and consequences of score use (American Educational Research Association et al., 2014). In the same way, in the existing scale-development frameworks, dimensionality, reliability, and various types of validity are not mutually exclusive but are complementary steps (Boateng et al., 2018). The present study is thus an important first step in the process of Urdu translation, however, more evidence is needed to make a comprehensive claim for validation.

Finally, the Urdu GELOPH-15 demonstrated a broad unidimensional structure as suggested in the original instrument and acceptable score reliability. The low item loadings, however, and the low shared variance suggest that the translated indicators were not sufficiently strong to represent gelotophobia. The results are thus preliminary evidence of structural plausibility and reliability rather than complete validity. The instrument should therefore be regarded as a preliminary research measure rather than a diagnostic or screening tool. The instrument needs to be further evaluated cognitively, bilingually, item-level refined, assessed for external validity, tested for measurement invariance, and replicated independently before it can be recommended for unrestricted research or application in Pakistan.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, Article 149. <https://doi.org/10.3389/fpubh.2018.00149>
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1(3), 185–216. <https://doi.org/10.1177/135910457000100301>
- Brislin, R. W. (1976). Comparative research methodology: Cross-cultural studies. *International Journal of Psychology*, 11(3), 215–229. <https://doi.org/10.1080/00207597608247359>
- Brislin, R. W. (1980). Translation and content analysis of oral and written materials. In H. C. Triandis and J. W. Berry (Eds.), *Handbook of cross-cultural psychology: Vol. 2. Methodology* (pp. 389–444). Allyn & Bacon.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.

- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, and F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69–93). Guilford Press.
- Cohen, S. (2004). Social relationships and health. *American Psychologist*, 59(8), 676–684. <https://doi.org/10.1037/0003-066X.59.8.676>
- Downey, G., & Feldman, S. I. (1996). Implications of rejection sensitivity for intimate relationships. *Journal of Personality and Social Psychology*, 70(6), 1327–1343. <https://doi.org/10.1037/0022-3514.70.6.1327>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., & Bebbington, P. E. (2002). A cognitive model of persecutory delusions. *British Journal of Clinical Psychology*, 41(4), 331–347. <https://doi.org/10.1348/014466502760387461>
- Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture*, 2(1), Article 8. <https://doi.org/10.9707/2307-0919.1014>
- Hu, L.-T., & 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>
- International Test Commission. (2018). ITC guidelines for translating and adapting tests (second edition). *International Journal of Testing*, 18(2), 101–134. <https://doi.org/10.1080/15305058.2017.1398166>
- Kağıtçıbaşı, Ç. (2007). *Family, self, and human development across cultures: Theory and applications* (2nd ed.). Lawrence Erlbaum Associates.
- Martin, R. A. (2007). *The psychology of humor: An integrative approach*. Elsevier Academic Press.
- Matsumoto, D., & Juang, L. (2016). *Culture and psychology* (6th ed.). Cengage Learning.
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412–433. <https://doi.org/10.1037/met0000144>
- Platt, T., Proyer, R. T., & Ruch, W. (2009). Gelotophobia and bullying: The assessment of the fear of being laughed at and its application among bullying victims. *Psychology Science Quarterly*, 51(2), 135–147. <https://doi.org/10.5167/uzh-19396>
- Proyer, R. T., & Ruch, W. (2010). Enjoying and fearing laughter: Personality characteristics of gelotophobes, gelotophiles, and katagelastics. *Psychological Test and Assessment Modeling*, 52(2), 148–160. <https://doi.org/10.5167/uzh-34689>
- Proyer, R. T., Ruch, W., Ali, N. S., Al-Olimat, H. S., Amemiya, T., Adal, T. A., Ansari, S. A., Arhar, Š., Asem, G., Baudin, N., Bawab, S., Bergen, D., Brdar, I., Brites, R., Brunner-Sciarra, M., Carrell, A., Carretero-Dios, H., Celik, M., Ceschi, G., ... Yeun, E. J. (2009). Breaking ground in cross-cultural research on the fear of being laughed at (gelotophobia): A multi-national study involving 73 countries. *HUMOR: International Journal of Humor Research*, 22(1–2), 253–279. <https://doi.org/10.1515/HUMR.2009.012>

- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*(2), 173–184. <https://doi.org/10.1177/01466216970212006>
- Ruch, W. (2009). Fearing humor? Gelotophobia: The fear of being laughed at—Introduction and overview. *HUMOR: International Journal of Humor Research, 22*(1–2), 1–25. <https://doi.org/10.1515/HUMR.2009.001>
- Ruch, W., & Proyer, R. T. (2008a). The fear of being laughed at: Individual and group differences in gelotophobia. *HUMOR: International Journal of Humor Research, 21*(1), 47–67. <https://doi.org/10.1515/HUMOR.2008.002>
- Ruch, W., & Proyer, R. T. (2008b). Who is gelotophobic? Assessment criteria for the fear of being laughed at. *Swiss Journal of Psychology, 67*(1), 19–27. <https://doi.org/10.1024/1421-0185.67.1.19>
- Ruch, W., & Proyer, R. T. (2009). Who fears being laughed at? The location of gelotophobia in the Eysenckian PEN-model of personality. *Personality and Individual Differences, 46*(5–6), 627–630. <https://doi.org/10.1016/j.paid.2009.01.004>
- Ruch, W., Hofmann, J., Platt, T., & Proyer, R. T. (2014). The state-of-the art in gelotophobia research: A review and some theoretical extensions. *HUMOR: International Journal of Humor Research, 27*(1), 23–45. <https://doi.org/10.1515/humor-2013-0046>
- Samson, A. C., Huber, O., & Ruch, W. (2011). Teasing, ridiculing and the relation to the fear of being laughed at in individuals with Asperger’s syndrome. *Journal of Autism and Developmental Disorders, 41*(4), 475–483. <https://doi.org/10.1007/s10803-010-1071-2>
- Sireci, S. G., & Berberoğlu, G. (2000). Using bilingual respondents to evaluate translated-adapted items. *Applied Measurement in Education, 13*(3), 229–248. https://doi.org/10.1207/S15324818AME1303_1
- Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice, 17*(2), 268–274. <https://doi.org/10.1111/j.1365-2753.2010.01434.x>
- Titze, M. (2009). Gelotophobia: The fear of being laughed at. *HUMOR: International Journal of Humor Research, 22*(1–2), 27–48. <https://doi.org/10.1515/HUMR.2009.002>
- Triandis, H. C. (1995). *Individualism and collectivism*. Westview Press.
- Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology, 54*(2), 119–135. <https://doi.org/10.1016/j.erap.2003.12.004>
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. SAGE Publications.