

Development and Psychometric Evaluation of the Migraine Perceptions and Lived Experience Questionnaire: Evidence for Face, Content, and Construct Validity and Reliability

Elnaz. Farmanian¹, Azita. Amir Fakhraei^{1*}, Eghbal. Zarei²

¹ Department of Psychology, B.A.C., Islamic Azad University, Bandarabbas, Iran

² Department of Psychology, University of Hormozgan, Bandarabbas, Iran

* Corresponding author email address: afakhraei2002@iau.ac.ir

E d i t o r

Izet Pehlić
Full professor for Educational sciences, Islamic pedagogical faculty of the University of Zenica, Bosnia and Herzegovina
izet.pehlic@unze.ba

R e v i e w e r s

Reviewer 1: Mahdi Khanjani
Associate Professor, Department of Psychology, Allameh Tabataba'i University, Tehran, Iran.
Email: khanjani_m@atu.ac.ir
Reviewer 2: Zahra Yousefi
Assistant Professor, Department of Psychology, Isfahan Branch (Khorasgan), Islamic Azad University, Isfahan, Iran. Email: Z.yousefi1393@khuif.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

In the item-generation paragraph, the statement that “the integration of phenomenological evidence and relevant literature initially yielded 72 items distributed across the six preliminary conceptual domains” needs a clearer item-generation audit trail. The authors should indicate approximately how many items originated directly from qualitative codes, how many were informed or supplemented by literature, whether any items were adapted from existing instruments, and how duplication across sources was handled. Without such information, readers cannot fully evaluate the content-development process or distinguish genuinely inductive items from theoretically imported content.

The manuscript states that the initial 72 items were distributed across six qualitative domains, with 13, 12, 14, 11, 12, and 10 items assigned to the respective domains. The authors should explain the rationale for this uneven distribution and whether item numbers reflected the richness or frequency of qualitative material, conceptual breadth, investigator judgment, or deliberate oversampling of certain domains. This is relevant because differential domain representation at the initial stage may influence subsequent factor extraction and apparent dimensionality.

The Methods section reports that the psychometric sample of 300 patients was divided into “two independent subsamples” of 150 participants for EFA and CFA, but the manuscript does not explain how this division was performed. The authors should specify whether cases were randomly allocated, split chronologically, or divided by another rule, and they should compare the two subsamples on major demographic and clinical characteristics to demonstrate that the CFA sample was reasonably comparable to the EFA sample. This is particularly important because non-random splitting could introduce systematic differences that affect cross-validation.

The use of “principal components extraction with Varimax rotation” in the exploratory stage warrants reconsideration or explicit justification. Principal components analysis is fundamentally a data-reduction method rather than a common-factor model, whereas the stated objective was to identify latent constructs underlying the MPLEQ. The authors should either reanalyze the data using an appropriate common-factor extraction method, such as principal axis factoring or maximum likelihood factor analysis, or provide a compelling methodological rationale for retaining principal components analysis. In addition, because the CFA model specifies correlated latent factors and the reported inter-factor correlations range from .32 to .64, an oblique rotation such as Promax or Oblimin would arguably have been more theoretically appropriate than orthogonal Varimax rotation.

The manuscript presents CFA fit indices of $\chi^2/df = 2.13$, CFI = .92, TLI = .91, GFI = .88, RMSEA = .062, and SRMR = .054 and concludes that the model has acceptable fit. This interpretation is reasonable, but the manuscript should adopt a more nuanced framework and avoid presenting threshold values as absolute pass/fail criteria. In particular, GFI and AGFI are less commonly emphasized in contemporary structural equation modeling, whereas confidence intervals around RMSEA, standardized residuals, and potentially information criteria provide more informative evidence. The authors should clarify the criteria selected a priori and distinguish conventional heuristic guidelines from definitive evidence of validity.

There is an internal inconsistency between the detailed CFA results and the separate reported ranges of standardized loadings. The detailed item-level results indicate an overall standardized loading range of .53 to .80, whereas the supplied summary of factor-specific ranges elsewhere includes values such as .61–.84, .63–.82, and .60–.83. These values cannot all be simultaneously correct. The authors should carefully audit the CFA output and ensure that every reported loading, range, and summary statistic is derived from the same final model. Such inconsistencies are particularly problematic in a psychometric manuscript because construct-validity conclusions depend directly on the accuracy of these coefficients.

Response: Revised and uploaded the manuscript.

1.2. Reviewer 2

Reviewer:

In the EFA results, the manuscript states that “eight components had eigenvalues greater than 1.00; however, examination of the scree plot, theoretical interpretability, and consistency with the qualitative conceptual framework indicated that a six-factor solution provided the most coherent representation of the data.” This factor-retention decision would be considerably strengthened by reporting parallel analysis, minimum average partial testing, or another empirically supported retention procedure. Reliance on the Kaiser criterion and visual scree inspection alone is relatively weak, particularly when the initial solution yielded eight eigenvalues above 1. The authors should either conduct parallel analysis or acknowledge explicitly that the six-factor decision was partly theory-driven and may therefore require replication.

The manuscript reports a KMO of .91 and a significant Bartlett test, $\chi^2(1540) = 5874.32$, $p < .001$, which support factorability, but the item-to-participant ratio in the EFA sample remains relatively modest because 56 items were analyzed in only 150 participants. The authors should address this issue explicitly and preferably report additional indicators of factor stability, such as communalities, anti-image correlations, determinant of the correlation matrix, or factor overdetermination. Given that some factors ultimately contained only seven items, evidence of strong communalities and well-defined factors would help mitigate concerns about the relatively small EFA sample.

The manuscript states that eight items were removed during EFA because of “loading $< .40$, problematic cross-loading, or conceptual inconsistency,” yet the specific quantitative cross-loading criterion is not reported. The authors should define what constituted a problematic cross-loading, for example, a secondary loading $\geq .30$ or a difference of less than $.20$ between primary and secondary loadings. They should also report the loadings of deleted items, at least in supplementary material, so that readers can evaluate whether deletions were statistically and conceptually justified rather than selectively made to reinforce the expected six-factor structure.

In the CFA section, the sentence “residual covariance was permitted between items concerning sensitivity to light and the need for a dark environment and between items concerning impaired concentration and post-attack mental slowing” requires more exact reporting. The authors should identify the exact item numbers involved, provide the relevant modification indices before modification, report the estimated residual covariances after modification, and explain why shared wording or overlapping content justified correlated residuals beyond their common latent factor. Because post hoc residual correlations can artificially improve model fit, transparent reporting is essential.

The CFA model was estimated using maximum likelihood, but the manuscript does not report whether assumptions relevant to this estimator were examined. Given that all questionnaire items use five-point ordinal Likert responses, the authors should report item distributions, skewness and kurtosis, multivariate normality, and the rationale for treating responses as continuous. If substantial non-normality was present, robust maximum likelihood or an ordinal estimator based on polychoric correlations would be preferable. At minimum, the potential consequences of using conventional maximum likelihood with five-category ordinal data should be acknowledged.

Response: Revised and uploaded the manuscript.

2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.