

Face and content validity of the BALANSING mobile application for stunting and anemia screening and education

Apriningsih Apriningsih^{1*}, *Neny Rosmawarni*², and *Firlia Ayu Arini*³

¹Department of Public Health Nutrition, Faculty of Health Science, Universitas Pembangunan Nasional Veteran Jakarta, 16415 Depok City, West Java, Indonesia

²Department of Informatics System, Faculty of Computer, Universitas Pembangunan Nasional Veteran Jakarta, 12450 Jakarta, Indonesia

³Department of Nutrition Science, Faculty of Health Science, Universitas Pembangunan Nasional Veteran Jakarta, Indonesia, 16415 Depok City, West Java, Indonesia

Abstract. This study evaluated the face and content validity of the BALANSING mobile application designed for stunting and anemia screening and education. Six reviewers participated: one Professor of Public Health Nutrition from Universitas Muhammadiyah Jakarta, one Associate Professor from Universiti Teknologi MARA (Malaysia), one nutrition officer from Puskesmas Limo (Depok, West Java), two Posyandu cadres, and one housewife. The mean face validity scores ranged from 3.5 to 4.5. Content validity ranged from 3.0 to 3.75. Item-level Content Validity Index (I-CVI) ranged from 0.83 to 1.00, while Scale-level Content Validity Index (S-CVI/Ave) was 0.92, indicating excellent content validity. Minor revisions were recommended for cultural contextualization and language clarity. The BALANSING application is considered valid and feasible for pilot implementation in community settings.

1 Introduction

Anemia and stunting remain major public health challenges in low- and middle-income countries, including Indonesia [1-3]. Digital health interventions (DHIs), particularly mobile applications, provide scalable opportunities for early screening and preventive health education. Recent studies suggest that DHIs can enhance access to screening tools, facilitate risk stratification, and improve health knowledge when supported by evidence-based content and user-centered design approaches [4-6]. Furthermore, co-design involving end users has been shown to improve the relevance, acceptability, and engagement of digital health applications [4, 5].

Before implementation, health applications require validation to ensure usability and scientific accuracy. Face validity evaluates whether an application appears appropriate and understandable to potential users and stakeholders, while content validity assesses the

* Corresponding author: apriningsih@upnvj.ac.id

relevance and adequacy of the information provided [4, 7]. Both are essential quality indicators for mobile health applications designed for screening and education.

Although various mHealth applications have been developed for nutrition-related problems, few integrate screening and educational functions for both stunting and anemia within a single platform. Existing applications often focus on a single condition and may not undergo rigorous validity assessment before implementation. In addition, culturally contextualized tools aligned with Indonesia's Posyandu system remain limited. To address these gaps, the BALANSING mobile application was developed as an integrated screening and educational tool for stunting and anemia. This study aimed to evaluate the face and content validity of the BALANSING mobile application before community-based implementation.

2 Methods

A methodological validation study with a cross-sectional design was conducted to evaluate face and content validity of the BALANSING mobile application.

The development and validation of BALANSING were guided by the **Technology Acceptance Model (TAM)** integrated with a digital mHealth framework. Key constructs included perceived usefulness, perceived ease of use, and behavioral intention, mapped to usability, clarity, and relevance domains assessed in this study. Six raters were purposively selected, comprising two academic experts in public health nutrition, two primary health care nutrition officers, one community health cadre, and one lay community representative.

Thirteen items covering relevance, clarity, completeness, and usability were evaluated using a four-point Likert scale (1=not relevant to 4=highly relevant). Quantitative analysis included mean score calculation, Item-level Content Validity Index (I-CVI), and Scale-level Content Validity Index (S-CVI/Ave). I-CVI was calculated as the proportion of experts rating an item 3 or 4. S-CVI/Ave was computed as the average of all I-CVI values. An S-CVI/Ave ≥ 0.80 indicates good content validity.

I-CVI was calculated as the proportion of raters assigning scores ≥ 3 . **S-CVI/Ave** was calculated as the average I-CVI across items. **Aiken's V** was calculated using the formula: $V = \frac{\sum(r - l_o)}{[n(c - 1)]}$, where $l_o = 1$, $c = 4$, $n = 6$. Critical values were applied to assess statistical significance.

3 Results and discussion

The face validity assessment revealed that the interface and usability components of BALANSING performed consistently well, with visual design attractiveness and consistency of terms both receiving mean scores of 4.5 (very good), and navigation clarity scoring 4.25 (very good) (Table 1).

Language clarity was rated 4.0 (good), and illustration support was rated 3.75 (good), whereas the educational feature clarity related to stunting received the lowest score, 3.5 (adequate). These findings suggest that the application's design and navigational structure were well received by reviewers, whereas the educational content, particularly the stunting-related features, requires further refinement. This pattern is consistent with evidence from Muhamat et al. [7], who found that interface attractiveness and ease of navigation are critical determinants of initial user acceptability in nutrition-focused mobile applications, whereas content clarity often requires iterative refinement informed by end-user feedback.

Table 1. Face validity results.

Aspect	Mean Score	Interpretation	Reviewer Feedback	Author Decision
Visual design attractiveness	4.5	Very good	Reviewers unanimously agreed the interface was visually appealing, clean, and appropriate for the target community. No revisions suggested.	Accepted as is; design retained in final version.
Navigation clarity	4.25	Very good	Reviewers found the menu flow intuitive and easy to operate. One reviewer suggested adding a back-button shortcut on sub-pages for ease of use.	Minor navigation shortcut added to improve user flow.
Language clarity	4.0	Good	Some reviewers noted that certain terms were too technical for lay users, particularly Posyandu cadres and community members. Simplification of medical terminology was recommended.	Technical terms simplified or accompanied by plain-language explanations throughout the application.
Educational feature clarity (stunting)	3.5	Adequate	Reviewers indicated that stunting-related educational content was adequate but could benefit from more contextual examples, local analogies, and stronger visual support to aid comprehension among non-health-professional users	Content revised to include locally relevant examples and additional infographic-style visuals to improve clarity.
Illustration support	3.75	Good	Reviewers acknowledged the illustrations as helpful but suggested that some images did not fully reflect the local Indonesian context in terms of appearance and setting	Illustrations updated to reflect Indonesian cultural and community contexts.
Consistency of terms	4.5	Very good	All reviewers agreed that terminology was used consistently throughout the application. No revisions required.	Accepted as is; terminology standardization maintained.

Content validity results demonstrated good to excellent validity across all assessed dimensions (Table 2). The relevance of the screening topic and educational materials both achieved mean scores of 3.75 with I-CVI values of 1.00, indicating unanimous expert agreement. Alignment with clinical guidelines also yielded an I-CVI of 1.00, confirming that the application's content is scientifically grounded. Completeness of content and cultural appropriateness received slightly lower I-CVI values of 0.83, which, although still above the accepted minimum threshold of 0.78, indicate areas requiring minor elaboration. The lower score on cultural appropriateness is notable, as Singh Sethi et al. [4] demonstrated that culturally adapted co-designed health applications show significantly higher relevance and

engagement among target communities, underscoring the importance of further contextualizing BALANSING's content for Indonesian community settings.

Table 2. Content validity results with CVI analysis.

Aspect	Mean Score	I-CVI	Interpretation	Reviewer Feedback	Author Decision
Relevance of screening topic	3.75	1.00	Excellent	Experts unanimously confirmed that screening for both stunting and anemia is highly relevant to Indonesia's current public health priorities. No revisions suggested	Accepted as is; screening framework retained.
Relevance of education material	3.75	1.00	Excellent	Reviewers confirmed educational content is appropriate and aligned with the needs of the target population. No revisions suggested	Accepted as is; educational content retained with minor language adjustments.
Completeness of content	3.5	0.83	Good	Some reviewers identified minor gaps, including the absence of information on complementary feeding practices and iron-rich food sources relevant to Indonesian diets.	Additional content on complementary feeding and local iron-rich foods added to the education module.
Cultural appropriateness	3.0	0.83	Good	This was the lowest-rated aspect. Reviewers noted that some content did not adequately reflect local cultural practices, dietary habits, and community norms. The Posyandu cadres specifically highlighted the need for culturally familiar food examples and language style.	Content revised to incorporate culturally appropriate food examples, local dialect-friendly language, and scenarios familiar to Indonesian community settings.
Alignment with guidelines	3.5	1.00	Excellent	All reviewers agreed that content aligns with national and WHO guidelines for stunting and anemia prevention. No major revisions required.	Accepted as is; guideline references retained and cited within the application.

The summary of validity indices (Table 3) provides strong aggregate evidence of the application's psychometric soundness. The Mean Item Score of 3.47 out of 4 indicates overall high-quality perception across all evaluated dimensions. The S-CVI/Ave of 0.92 exceeds the

recommended threshold of 0.90, confirming excellent scale-level content validity and high inter-rater agreement.

The Aiken's V coefficient of 0.76 falls within the acceptable-to-strong range; for six raters using a four-point scale, values above 0.70 are interpreted as statistically acceptable agreement beyond chance. Collectively, these three indices corroborate one another and confirm that BALANSING meets established validity standards for health applications. These results are aligned with findings from Brewer et al. [5], who reported that mHealth applications developed through structured expert validation achieve higher trust and program integration within community health systems.

Tabel 3. Summary of validity indices.

Indicator	Value	Interpretation
Mean Item Score	3.47 / 4	Very Good
S-CVI/Ave	0.92	Excellent
Overall Aiken's V	0.76	Acceptable–Strong
Number of Items	13	

From a digital health perspective, early-stage psychometric validation is essential for ensuring adoption and sustainability. Evidence from peer-reviewed sources indicates that applications subjected to rigorous content and face validity procedures demonstrate greater usability and user trust [6]. The minor recommendations raised by reviewers, particularly around terminology simplification and cultural contextualization, are consistent with user-centered design principles and can be addressed through iterative development before pilot testing [4, 7]. Overall, the results support the readiness of BALANSING for progression to usability testing, construct validation, and effectiveness trials in real-world community settings.

This study has several limitations that should be taken into account when interpreting the findings. The expert panel was small (n=6) and recruited through purposive sampling from a single geographic region, which may limit the representativeness and generalizability of the results. Only face and content validity were assessed; psychometric properties, such as construct validity, reliability, and criterion validity, remain to be established in subsequent studies. Additionally, validation was conducted under structured review conditions rather than real-world community settings, and end-user testing with the actual target population was not included in this phase. Behavioral constructs from the Technology Acceptance Model, including perceived usefulness and ease of use, were not empirically measured, and future usability trials are needed to assess actual adoption and effectiveness of the BALANSING application in practice.

4 Conclusion

The BALANSING mobile application demonstrated strong face validity and excellent content validity, with an S-CVI/Ave of 0.92 and Aiken's V of 0.76, confirming its readiness for pilot implementation in community-based settings. Minor revisions regarding cultural contextualization, language clarity, and content completeness are recommended before implementation. Future studies should prioritize usability testing with the target population, reliability assessment, construct validity evaluation, and pilot effectiveness trials to comprehensively establish the application's psychometric soundness and real-world impact on stunting and anemia prevention in Indonesia.

Reference

1. E. Meiyetrian, B. Utomo, Anemia and stunting in Indonesian children. *Narra J.* **5**, 1–15 (2025).
2. A. Apriningsih, L. Hanifah, N. Nasrulloh, Exclusive breastfeeding practice during COVID-19 pandemic in West Java, Indonesia: A cross-sectional study. *PLoS One* **19**, e0303386 (2024). <https://doi.org/10.1371/journal.pone.0303386>
3. A. Apriningsih, S. Madanijah, C.M. Dwiriani, R. Kolopaking, Improving school readiness for weekly iron folic acid supplementation program through school readiness training and technical assistance intervention. *J. Gizi Pangan* **16**(Suppl. 1), 139–146 (2021).
4. M.I. Singh Sethi, T. Manickam, T. Chakraborty, S. Bada Math, Co-designed mental health screening app (Here for You) for university students: Pilot feasibility mixed methods study. *JMIR Form. Res.* **10**, e75616 (2026). <https://doi.org/10.2196/75616>
5. L.C. Brewer, S.N. Hayes, A.R. Caron, D.A. Derby, N.S. Breutzman, A. Wicks, et al., Promoting cardiovascular health and wellness among African-Americans: Community participatory approach to design an innovative mobile-health intervention. *PLoS One* **14**, e0218724 (2019). <https://doi.org/10.1371/journal.pone.0218724>
6. S. Pan, L.Q. Rong, Mobile applications in clinical and perioperative care for anesthesia: A narrative review. *J. Med. Internet Res.* **23**, e25115 (2021). <https://doi.org/10.2196/25115>
7. N.A. Muhamat, R. Hasan, N. Saddki, M. Arshad, M. Ahmad, Development and usability testing of mobile application on diet and oral health. *PLoS One* **16**, e0257035 (2021). <https://doi.org/10.1371/journal.pone.0257035>