

University Students' Perceived Comprehension of Digital Vocabulary Standardized in the Kamus Besar Bahasa Indonesia (KBBI)

^{*1}Apri Damai Sagita Krissandi, ²Nandita Wana Putri

¹Sanata Dharma University, Indonesia, ²Panca Bhakti School of Health Sciences, Indonesia

*Corresponds email: apridamai@usd.ac.id

Article Info

Received: June 02, 2026

Revised: September 30, 2026

Accepted: October 05, 2026

Keywords:

standardized digital vocabulary
lexical comprehension
language planning
lexical modernization
digital literacy

Abstract

The rapid development of digital communication has generated and popularized technology-related vocabulary in Indonesian, including established Indonesian forms and standardized equivalents of foreign terms. Some of these forms are recorded in the *Kamus Besar Bahasa Indonesia* (KBBI), although their perceived comprehension among university students may vary. This study investigates university students' perceived comprehension of fourteen digital vocabulary items recorded as standardized Indonesian forms in KBBI, including *unggah* (upload), *swafoto* (selfie), *pranala* (link), and *metamesta* (metaverse). A descriptive quantitative survey was conducted with 80 undergraduate students aged 18–24 years. Data were collected using a four-point scale ranging from *Do Not Understand* to *Fully Understand* to measure respondents' perceived comprehension of each lexical item. The overall mean score was 3.1, placing perceived comprehension in the *Understand* category. However, substantial variation was observed across items. The highest mean scores were recorded for *unggah*, *unduh* (download), *virtual*, *warganet* (netizen), and *gawai* (gadget), whereas *metamesta*, *pranala*, and *pelantang* (microphone) received the lowest scores. The findings indicate that lexicographic standardization does not necessarily correspond to uniform perceived comprehension among users. The study contributes to discussions of Indonesian language planning, vocabulary standardization, and digital language use by highlighting the importance of considering user-level comprehension alongside the documentation of standardized lexical forms.

To cite this article: Krissandi, A. D. S., & Putri, N. W. (2026). University Students' Perceived Comprehension of Digital Vocabulary Standardized in the Kamus Besar Bahasa Indonesia (KBBI). *Alinea: Jurnal Bahasa, Sastra Dan Pengajaran*, 6(3), 1822–1833. <https://doi.org/10.58218/alinea.v6i3.3467>.

This article is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/) ©2026 by author/s

Introduction

Language is a dynamic system that changes in response to social, technological, and communicative developments. The emergence of new vocabulary is part of this process, as language users and language institutions create, adapt, borrow, translate, and develop lexical forms to meet changing communicative needs (Chaer, 2009; Kridalaksana, 2008). In the digital era, this process is particularly visible in vocabulary associated with technology, the internet, and online communication.

The rapid expansion of digital technologies, social media, and online communication has increased the circulation of terms such as metaverse, selfie, upload, download, link, server, and gadget. Indonesian language planning has responded to this development through processes of adoption, adaptation, translation, and lexical creation. The Language Development and Fostering Agency has, for example, documented Indonesian equivalents such as *gawai* for gadget, *laman* for homepage, *unduh* for download, *unggah* for upload, and *peladen* for server. These processes are

consistent with the principles of Indonesian terminology development, which accommodate the formation, translation, and adaptation of terms in response to scientific and technological change (Pusat Bahasa, 2007; Badan Pengembangan dan Pembinaan Bahasa, 2015).

For this reason, the present study does not use the term *standardized digital vocabulary* as a strict linguistic category, but rather as a functional term referring to technology-related lexical items that have been recorded in *Kamus Besar Bahasa Indonesia* (KBBI) Daring as Indonesian equivalents or standardized forms of foreign terms. In the earlier version of the manuscript, the fourteen items were collectively described as *standardized digital vocabulary*. However, this label may appear too broad because the items display heterogeneous linguistic characteristics. Some combine Indonesian and foreign elements, some are newly created or morphologically formed Indonesian equivalents, some reuse existing Indonesian lexical items for specialized meanings, and others are adapted or directly adopted forms. Following the terminology of Indonesian language planning, the fourteen items are therefore treated in this study as standardized digital vocabulary because they are recorded in KBBI Daring and function as Indonesian equivalents or standardized forms corresponding to foreign technology-related terms (Badan Pengembangan dan Pembinaan Bahasa, 2026). This terminology acknowledges the linguistic heterogeneity of the dataset while maintaining the study's central focus on users' understanding of standardized vocabulary.

The study is necessary because the standardization of digital and technology-related vocabulary does not automatically ensure that such vocabulary is understood by its intended users. As new technological concepts continue to enter everyday communication, Indonesian language users increasingly encounter both foreign terms and their standardized Indonesian counterparts. A gap between official standardization and actual user comprehension may reduce the effectiveness of language standardization and create difficulties in communication, education, and information access. Therefore, examining university students' understanding of standardized digital vocabulary is important for assessing whether officially recorded forms are recognizable and meaningful to contemporary users. Theoretically, this study contributes empirical evidence to discussions of Indonesian lexical standardization, language planning, and vocabulary comprehension by showing how standardized forms are understood by actual language users. Practically, the findings may benefit language planners and policymakers, educators, textbook and educational-content developers, and digital-content producers by providing information about which standardized terms are readily understood and which may require further explanation, dissemination, or reconsideration. The findings may also help students and other Indonesian language users become more familiar with standardized terminology in the digital domain. If this issue is not investigated, the continued development and standardization of technology-related vocabulary may proceed without sufficient evidence concerning users' actual comprehension, potentially widening the gap between officially standardized terminology and vocabulary used or understood in everyday communication.

The fourteen lexical items were selected using three criteria. First, each item was recorded in KBBI Daring as an Indonesian lexical item associated with a technology-related concept or as a standardized equivalent of a foreign term. Second, the items represent concepts that are relevant to digital technology, internet communication, and contemporary online culture. Third, the items were selected to represent forms that differ in their apparent familiarity and linguistic characteristics. The resulting set is intentionally heterogeneous in linguistic formation, allowing the study to examine perceived comprehension across different types of standardized digital vocabulary.

Table 1. Linguistic Characteristics of the Fourteen Standardized Digital Vocabulary Items

No.	Indonesian Term	Source Term	Linguistic Characterization	Relation to Standardization
1	warganet	netizen	Blend combining warga and the English element net.	Indonesian standardized lexical creation

2	gawai	gadget	An established Indonesian lexical item assigned a specialized technological meaning.	Indonesian equivalent through lexical reuse/semantic specialization
3	metamesta	metaverse	Lexical creation combining meta- and semesta.	Indonesian equivalent/creation
4	pranala	link	Indonesian lexical creation used for the concept of a link.	Indonesian equivalent/creation
5	pratayang	preview	Morphological formation with pra- and tayang.	Indonesian equivalent/formation
6	narahubung	contact person	Compound/lexical formation involving nara and hubung.	Indonesian equivalent/creation
7	pelantang	microphone	Morphological formation with peN- and lantang.	Indonesian equivalent/derivation
8	swafoto	selfie	Compound combining swa- and foto.	Indonesian equivalent/creation
9	peladen	server	Morphological formation with peN- and laden.	Indonesian equivalent/derivation
10	laman	page/homepage	Established Indonesian lexical item used as a technical digital term.	Indonesian equivalent through lexical reuse
11	unduh	download	Established Indonesian/related Malay lexical form used as a digital equivalent.	Indonesian equivalent/creation
12	unggah	upload	Established Indonesian/related Malay lexical form used as a digital equivalent.	Indonesian equivalent/creation
13	virtual	virtual	Direct adoption of the international form with no substantial formal change.	Adoption
14	kripto	crypto	Orthographically adapted form of crypto.	Adaptation

The present study aims to examine university students' perceived comprehension of fourteen digital vocabulary items standardized and recorded in KBBI Daring. The selected items represent diverse processes of Indonesian lexical development, including lexical creation, morphological formation, semantic specialization, adaptation, and direct adoption. Previous research has examined Indonesian vocabulary and terminology from several perspectives, but these studies have addressed different populations, linguistic contexts, and analytical concerns. Mutiara et al. (2024), for example, investigated Indonesian vocabulary mastery among Myanmar students in the BIPA context, providing evidence of vocabulary learning among non-native Indonesian language users. However, their study focused on vocabulary mastery in language learning rather than on Indonesian standardized terminology encountered by undergraduate students. Aliyah and Aliwafa (2025) examined vocabulary comprehension in Arabic language learning and demonstrated the relevance of vocabulary understanding to language learning contexts, but their study did not address Indonesian terminology or the relationship between lexical standardization and users' familiarity with standardized forms.

Ningsih and Pallawagau (2025) approached vocabulary from a lexical-semantic perspective, examining how meaning is constructed through semantic relations and context; nevertheless, their study did not investigate users' comprehension of standardized technology-related vocabulary.

Research on Indonesian terminology provides a complementary perspective. Syamsuri (2011) examined the acceptability of Indonesian terms and argued that acceptance is influenced not only by formal standardization but also by factors such as usefulness, efficiency, and circulation in society. However, the study examined terminology more broadly and did not empirically measure university students' perceived comprehension of specific standardized digital terms. Luthfita and Kusumawardani (2022) examined foreign word and term equivalents documented in KBBI, identifying the fields, source languages, and processes involved in the formation of equivalents. Their study provides an important account of terminology development in KBBI, but it focused on the characteristics and formation of lexical equivalents rather than on whether users understand those forms. Similarly, Siregar et al. (2024) investigated the acceptance of language standardization in society and found that some users were unfamiliar with standardized forms. However, their qualitative study involved a relatively small group of participants and addressed language standardization in general rather than digital vocabulary among university students. These studies collectively demonstrate that previous research has addressed vocabulary mastery, lexical semantics, terminology formation, and language standardization, but the users' comprehension of technology-related Indonesian forms formally recorded in KBBI remains insufficiently examined.

The present study addresses this gap by bringing together two perspectives that have generally been examined separately: the formal standardization of Indonesian digital terminology and the users' perceived comprehension of those standardized forms. Unlike previous studies that primarily examined vocabulary learning, semantic relations, terminology formation, or general language acceptance, this study focuses on fourteen KBBI-recorded digital vocabulary items and examines how undergraduate students perceive their understanding of these forms. The study is also distinctive in its item-level comparison across linguistically heterogeneous forms, including lexical creations, morphological formations, semantic specializations, adaptations, and direct adoptions. By examining these forms among undergraduate students at Sanata Dharma University, the study provides empirical evidence concerning the extent to which formal lexical standardization corresponds with users' perceived familiarity and understanding. The findings are expected to contribute to discussions of Indonesian language planning and terminology development by providing a user-oriented perspective on the implementation and uptake of standardized digital vocabulary.

Accordingly, this study addresses the following research question: How do undergraduate students perceive their understanding of digital vocabulary standardized in KBBI Daring?

Materials and Methods

This study employed a descriptive quantitative survey design. The survey was designed to measure respondents' perceived or self-reported comprehension of standardized digital vocabulary rather than objectively verified semantic knowledge. Participants were selected using purposive sampling according to the study criteria: undergraduate students aged 18–24 years who were actively engaged in digital communication and regularly exposed to technology-related vocabulary in academic and social-media contexts. A total of 80 undergraduate students from Sanata Dharma University, Yogyakarta, Indonesia, participated in the study. Because the sample was purposively drawn from one university, the findings are interpreted as describing the participating respondents rather than university students in general.

The research instrument consisted of fourteen standardized digital vocabulary items recorded in KBBI Daring. Each item was presented with its corresponding source term in English for identification purposes, followed by a statement asking respondents to report how well they understood the meaning of the Indonesian term. The construct measured was therefore perceived lexical comprehension, not objectively demonstrated semantic accuracy. Responses were recorded using a four-point scale: (1) Do Not Understand, (2) Slightly Understand, (3) Understand, and (4) Fully

Understand. The four-point response format was used to avoid a neutral midpoint and to obtain a directional self-assessment of perceived comprehension.

Data collection was conducted in April 2026. The questionnaire was distributed through university digital communication channels and accompanied by instructions regarding completion procedures. The data were tabulated and analyzed descriptively by calculating total and mean scores for each lexical item. For each item, the mean score was obtained by dividing its total score by the 80 respondents. Based on the corrected item totals reported in Table 2, the overall mean was calculated from 1,120 item-level responses (80 respondents × 14 items). The sum of the fourteen item totals was 3,421, producing an overall mean of 3.0545, which was rounded to 3.1. Because all fourteen items had the same number of valid responses, this value is equivalent to the arithmetic mean of the fourteen item means. The mean scores were interpreted using the following categories: 1.00–1.74 = Do Not Understand; 1.75–2.49 = Slightly Understand; 2.50–3.24 = Understand; and 3.25–4.00 = Fully Understand.

Before administration, the instrument underwent expert review by two applied linguistics scholars to assess the relevance of the items and their alignment with the study objectives. A pilot test involving ten university students with characteristics similar to those of the target population was conducted to evaluate item clarity and questionnaire administration procedures. Feedback from the expert review and pilot test was used to refine the wording and administration procedures before the final distribution. These procedures provide evidence of content relevance and item clarity; however, because the instrument measured self-reported comprehension and no objective comprehension test was administered, the study does not claim to establish respondents' actual semantic accuracy.

Results And Discussion

Result

Data were collected through an online survey administered to 80 undergraduate students. Respondents reported their perceived level of understanding of fourteen standardized digital vocabulary items using a four-point scale: *Fully Understand*, *Understand*, *Slightly Understand*, and *Do Not Understand*. Each category was accompanied by an explanation in the questionnaire. *Fully Understand* indicated complete perceived understanding of the lexical item, *Understand* indicated general understanding with some uncertainty, *Slightly Understand* indicated partial understanding with substantial uncertainty, and *Do Not Understand* indicated perceived unfamiliarity with the lexical item.

Table 2 presents the perceived-comprehension scores for the fourteen lexical items. Total scores ranged from 156 to 310, while mean scores ranged from 2.0 to 3.9. The overall mean was 3.1. The total score across 1,120 item-level responses was 3,421, corresponding to an unrounded mean of 3.0545. Because all fourteen items received responses from the same 80 participants, this value is equivalent to the arithmetic mean of the fourteen item means. Based on the classification used in the study, the overall result falls within the *Understand* category.

Table 2. Survey Results on the Perceived Comprehension of Standardized Digital Vocabulary

No.	Lexical Item	Total Score	Mean Score (4-point scale)	Category
1	warganet	289	3.6	Fully Understand
2	gawai	271	3.4	Fully Understand
3	metamesta	156	2.0	Slightly Understand
4	pranala	164	2.1	Slightly Understand
5	pratayang	227	2.8	Understand
6	narahubung	249	3.1	Understand
7	pelantang	171	2.1	Slightly Understand
8	swafoto	266	3.3	Fully Understand
9	peladen	201	2.5	Understand

10	laman	271	3.4	Fully Understand
11	unduh	309	3.9	Fully Understand
12	unggah	310	3.9	Fully Understand
13	virtual	306	3.8	Fully Understand
14	kripto	231	2.9	Understand
	Overall Mean		3.1	Understand

The highest mean scores were obtained by *unggah* and *unduh* ($M = 3.9$), followed by *virtual* ($M = 3.8$). The lowest mean scores were recorded for *metamesta* ($M = 2.0$), *pranala* ($M = 2.1$), and *pelantang* ($M = 2.1$). The remaining items had mean scores between 2.5 and 3.6.

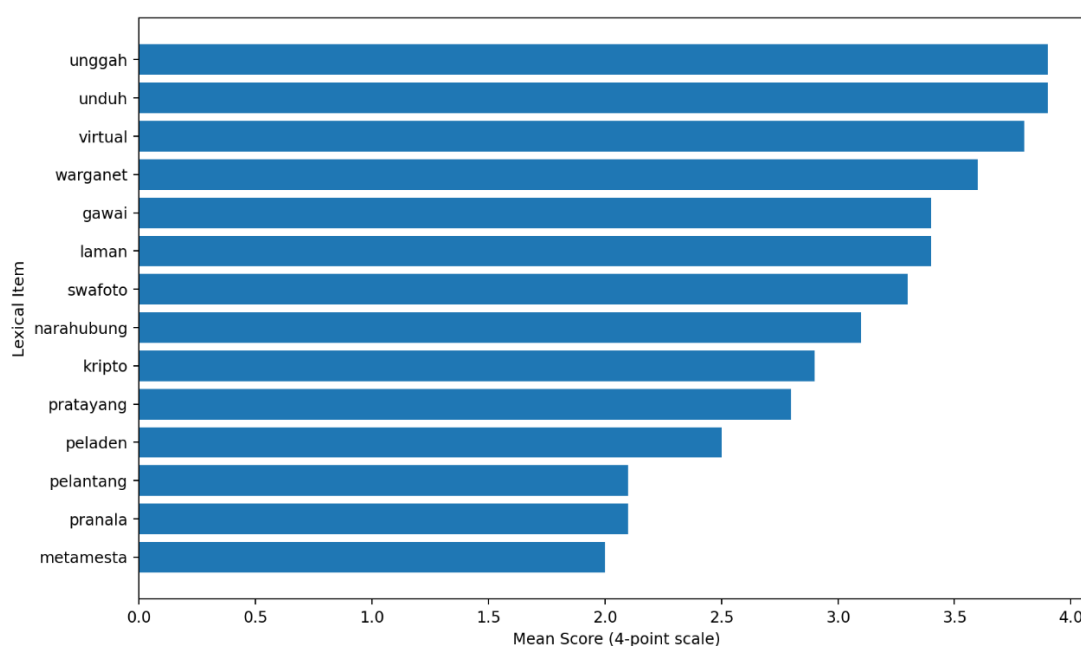


Figure 1. Mean Perceived-Comprehension Scores of the Fourteen Standardized Digital Vocabulary Items. Figure 1 illustrates the mean self-reported comprehension score for each item, allowing comparison of respondents' perceived familiarity with the standardized Indonesian forms.

Figure 1 presents the mean perceived-comprehension score for each of the fourteen lexical items, showing the variation in scores from 2.0 to 3.9. The distribution of the fourteen lexical items across the four perceived-comprehension categories is presented in Figure 2. Seven items were classified as *Fully Understand*, four as *Understand*, and three as *Slightly Understand*. No item was classified as *Do Not Understand* based on its mean score.

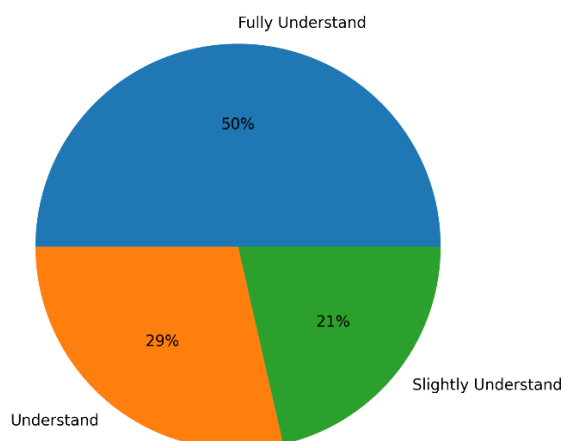


Figure 2. Distribution of Standardized Digital Vocabulary Items Across Perceived-Comprehension Categories. Figure 2 summarizes the proportion of lexical items classified into each category based on their mean scores.

Figure 2 shows the number of lexical items assigned to each perceived-comprehension category based on their mean scores. To provide a more detailed description of the lower-scoring items, *metamesta*, *pranala*, and *pelantang* were examined together with *peladen*. The first three items had the lowest mean scores, while *peladen* had a mean of 2.5, placing it at the lower boundary of the *Understand* category.

Table 3. Response Distribution for Lexical Items with the Lowest Perceived-Comprehension Scores

Lexical Item	Fully Understand (%)	Understand (%)	Slightly Understand (%)	Do Not Understand (%)	Mean Score
<i>metamesta</i> (metaverse)	5.0	11.3	57.5	26.2	2.0
<i>pranala</i> (link)	6.3	15.0	56.3	22.5	2.1
<i>pelantang</i> (microphone)	7.5	21.3	48.8	22.5	2.1
<i>peladen</i> (server)	15.0	31.3	43.8	10.0	2.5

For *metamesta*, 5.0% of respondents selected Fully Understand, 11.3% selected Understand, 57.5% selected Slightly Understand, and 26.2% selected Do Not Understand. Thus, 83.7% selected either Slightly Understand or Do Not Understand. For *pranala*, 6.3% of respondents selected Fully Understand, 15.0% selected Understand, 56.3% selected Slightly Understand, and 22.5% selected Do Not Understand. Therefore, 78.8% selected either Slightly Understand or Do Not Understand. For *pelantang*, 7.5% selected Fully Understand, 21.3% selected Understand, 48.8% selected Slightly Understand, and 22.5% selected Do Not Understand. The combined proportion of Slightly Understand and Do Not Understand responses was therefore 71.3%. For *peladen*, 15.0% of respondents selected Fully Understand, 31.3% selected Understand, 43.8% selected Slightly Understand, and 10.0% selected Do Not Understand. The mean score was 2.5, placing the item in the Understand category.

Overall, the fourteen items were distributed across three categories. The Fully Understand category consisted of *warganet*, *gawai*, *swafoto*, *laman*, *unduh*, *unggah*, and *virtual*, with mean scores ranging from 3.3 to 3.9. The Understand category consisted of *pratayang*, *narahubung*, *peladen*, and *kripto*, with mean scores ranging from 2.5 to 3.1. The Slightly Understand category consisted of *metamesta*, *pranala*, and *pelantang*, with mean scores ranging from 2.0 to 2.1. No item had a mean score within the Do Not Understand category.

Discussion

The overall mean of 3.1 indicates that the participating students generally perceived themselves as understanding the selected standardized digital vocabulary. However, this finding should be interpreted specifically as perceived or self-reported comprehension, rather than objectively demonstrated semantic knowledge. Vocabulary knowledge is multidimensional, and recognition of a lexical form does not necessarily indicate precise knowledge of its meaning or the ability to use it appropriately. Recent research distinguishes meaning recognition and meaning recall as related but distinct aspects of vocabulary knowledge, while other studies demonstrate that different measures capture different dimensions of lexical competence (Stewart et al., 2024; Schurz, 2024). In the present study, respondents were not asked to define the terms, provide synonyms, interpret them in context, or use them in sentences. Therefore, the relatively high score for *unggah*, for example, should be interpreted as high perceived familiarity rather than evidence of complete semantic mastery. Similar methodological distinctions have been emphasized in research on vocabulary measurement (Luo et al., 2021; Stewart et al., 2024).

The substantial variation among the fourteen items is therefore more informative than the overall mean alone. *Unggah*, *unduh*, *virtual*, *warganet*, *gawai*, *laman*, and *swafoto* received relatively high scores, whereas *metamesta*, *pranala*, and *pelantang* received the lowest. This variation suggests that being recorded as a standardized Indonesian form does not necessarily result in an equivalent level of familiarity among users. One possible factor is repeated exposure. Research has shown that frequency, lexical coverage, and contextual encounters contribute to vocabulary knowledge and lexical accessibility (Dang, 2022; Ludewig et al., 2023; Sánchez et al., 2023). The relatively high scores for *unggah* and *unduh* are compatible with this explanation because these terms are frequently encountered in common digital activities. However, because exposure and frequency were not measured in the present study, they cannot be treated as causal explanations. Research on digital environments likewise indicates that repeated digital input can support vocabulary development (Behforouz & Frumuselu, 2021; Hanafiah et al., 2022; Zhu et al., 2024), but these studies provide contextual support rather than direct evidence for the causes of the present pattern.

The lower scores for *metamesta*, *pranala*, and *pelantang* present a different pattern. These terms are formally established Indonesian forms, yet they were less familiar to the respondents. This finding points to a distinction between formal standardization and user-level lexical uptake. Indonesian technology-related terminology can be formed through different processes, including translation, adaptation, morphological formation, and absorption (Fitria, 2023). Consequently, the establishment of an Indonesian equivalent does not necessarily mean that the form has achieved the same level of familiarity as the concept it represents. This distinction is particularly relevant in digital communication, where Indonesian users encounter standardized Indonesian forms alongside English forms, informal expressions, abbreviations, and newly emerging vocabulary (Bajuri et al., 2025; Pratama et al., 2026).

The result for *pranala* illustrates this distinction particularly clearly. Although the concept represented by *pranala* is common in digital communication, the term itself received a relatively low mean score. Thus, familiarity with a digital concept does not necessarily imply familiarity with the specific lexical form used to represent it. This interpretation is compatible with research showing that lexical knowledge involves the establishment and accessibility of form–meaning relationships rather than simple recognition of isolated forms (Alshehri & Zhang, 2022; Pan et al., 2023). Similarly, the low score for *metamesta* should not be interpreted as evidence that respondents were unfamiliar with the concept of the metaverse. Rather, it indicates limited perceived familiarity with the Indonesian lexical form *metamesta*. Only 16.3% of respondents reported that they *Understood* or *Fully Understood* the term, while 83.7% reported limited or no understanding. The finding therefore suggests that establishing a lexical equivalent may represent only one stage of terminology development; subsequent circulation and contextual use may also be relevant to users' familiarity. Digital platforms may provide one possible means of such dissemination, as illustrated by research on introducing Indonesian lexical equivalents through Instagram (Khoiriyah et al., 2025).

The result for *pelantang* should be interpreted more cautiously. Its mean score was 2.1, with 71.3% of respondents reporting either *Slightly Understand* or *Do Not Understand*. The present data do not establish whether this pattern resulted from limited exposure, competition with the English form *microphone*, morphological characteristics, or another factor. These explanations would require additional evidence because exposure, frequency, lexical transparency, and objective semantic knowledge were not independently measured.

The *peladen* result further demonstrates why perceived comprehension should not be equated with precise semantic knowledge. Although its mean score of 2.5 placed it in the *Understand* category, 53.8% of respondents still reported *Slightly Understand* or *Do Not Understand*. The form may invite interpretation through its Indonesian morphological structure, but the present instrument cannot determine whether respondents actually relied on morphological information when judging its meaning. Morphological awareness may contribute to vocabulary knowledge, but it does not necessarily guarantee precise semantic comprehension (Alshehri & Zhang, 2022; Pan et al., 2023).

The fourteen items should also not be treated as a homogeneous morphological category. They represent different relationships between Indonesian lexical formation and their source concepts, including lexical creation, morphological formation, semantic specialization, adaptation, and direct adoption. Research on Indonesian ICT terminology similarly demonstrates that technology-related terms may enter Indonesian through multiple processes (Fitria, 2023), while studies of loanword adaptation show that borrowed forms can undergo phonological and orthographic modification in the receiving language (Glewwe, 2021; Zeng, 2022). This linguistic heterogeneity supports the use of standardized digital vocabulary as a functional analytical label rather than *hybrid vocabulary*. The term refers here to digital-domain lexical items recorded as standardized Indonesian forms in KBBI, without claiming that they share a single mechanism of lexical formation.

This distinction also has implications for language planning. Language standardization provides an institutional mechanism for establishing and documenting preferred lexical forms, but standardization alone cannot guarantee immediate recognition or use by all language users. This is particularly relevant in digital communication, where lexical forms circulate rapidly across different platforms and registers (Herring, 2022; Bajuri et al., 2025; Pratama et al., 2026). The *Pedoman Umum Pembentukan Istilah* provides a framework for Indonesian terminology formation (Pusat Bahasa, 2007), while the Badan Pengembangan dan Pembinaan Bahasa has also addressed terminology and language-use issues in institutional contexts (Badan Pengembangan dan Pembinaan Bahasa, 2015). KBBI Daring serves as a contemporary lexicographic reference for Indonesian vocabulary (Badan Pengembangan dan Pembinaan Bahasa, 2026). The present study adds a user-oriented perspective by showing that lexicographic inclusion and perceived user comprehension are not necessarily equivalent.

This finding is relevant to broader research on vocabulary and language comprehension. Studies have demonstrated relationships between vocabulary breadth, vocabulary depth, and comprehension, while also showing that these relationships depend on how vocabulary knowledge is measured (Luo et al., 2021; Chen & Zhang, 2023; Dagnaw, 2023; Tong & Tong, 2022; Wang & Zhang, 2025). Research involving Indonesian vocabulary learning also shows the importance of examining vocabulary knowledge in relation to learners and learning contexts (Mutiarra et al., 2024; Aliyah & Aliwafa, 2025). In this respect, the present study differs from studies focused primarily on vocabulary learning or reading comprehension because it examines students' perceived comprehension of standardized Indonesian digital vocabulary.

The findings also have a methodological implication. The questionnaire presented the English source term alongside the Indonesian lexical item. Although this was intended to identify the target concept, the source-language equivalent may have provided an additional semantic cue. Research on glossing indicates that providing equivalent information can influence vocabulary processing and learning (Glossing and Vocabulary Learning, 2022; Zhu et al., 2024). Accordingly, future studies should consider presenting standardized forms without source-language equivalents when the objective is to measure unaided lexical recognition. Future research could also combine self-report measures with

meaning-recognition, meaning-recall, and contextual-use tasks. Such designs would provide a stronger basis for distinguishing familiarity with a lexical form from accurate semantic knowledge, consistent with recent research distinguishing recognition and recall as different dimensions of vocabulary knowledge (Stewart et al., 2024).

Overall, the findings indicate that the uptake of standardized digital vocabulary is uneven. The higher scores for several frequently encountered forms contrast with the lower perceived comprehension of *metamesta*, *pranala*, and *pelantang*. The results therefore point to a potential gap between lexicographic standardization and user-level lexical familiarity. Importantly, the study does not establish that exposure, frequency, lexical formation, or semantic transparency caused the observed differences, because these variables were not directly measured. Its contribution is more specific: it provides empirical evidence that standardized digital vocabulary recorded in KBBI is not perceived as equally comprehensible by university students. Further research should therefore examine the relationship among standardization, lexical exposure, semantic knowledge, and actual use to determine how standardized terminology becomes established in users' linguistic repertoires.

Conclusions

This study examined the perceived comprehension of fourteen digital vocabulary items standardized and recorded in KBBI Daring among 80 undergraduate students at Sanata Dharma University. The overall mean score was 3.1 on a four-point scale, placing the participating respondents within the *Understand* category. However, perceived comprehension varied considerably across individual lexical items. *Unggah*, *unduh*, and *virtual* received the highest scores, whereas *metamesta*, *pranala*, and *pelantang* received the lowest scores.

The findings indicate that the formal standardization and dictionary inclusion of digital vocabulary do not necessarily correspond to uniform perceived familiarity among users. In particular, several Indonesian standardized forms received relatively low scores even when their corresponding digital concepts are widely encountered. The study therefore highlights a potential gap between formal lexical standardization and user-level lexical familiarity.

The findings have implications for Indonesian language planning and vocabulary dissemination. For less familiar standardized digital terms, future research could examine whether contextualized exposure through educational materials and digital communication supports their recognition and use. However, the present study does not establish causal relationships involving exposure, frequency of use, lexical transparency, or competition with foreign forms, as these variables were not directly measured.

The study has three principal limitations. First, the data represent self-reported comprehension rather than objectively tested semantic knowledge. Second, the sample was limited to 80 undergraduate students from one university and therefore cannot be generalized to university students nationally. Third, the presentation of English source terms may have provided additional semantic cues. Future research should employ objective meaning-recognition and meaning-recall tasks, measure actual exposure and frequency of use, and include participants from multiple universities and age groups.

Declaration of Conflicting Interest

The authors declare that there is no conflict of interest in this work.

References

- Alshehri, M. G., & Zhang, D. (2022). The lexical basis of second language reading comprehension: From (sub)lexical knowledge to processing efficiency. *Language Learning*, 72(2), 325–364. <https://doi.org/10.1111/lang.12478>
- Aliyah, M., & Aliwafa, A. (2025). Implementasi metode Total Physical Response (TPR): Strategi peningkatan pemahaman kosakata bahasa Arab. *Alinea: Jurnal Bahasa, Sastra dan Pengajaran*, 5(3), 919–933. <https://doi.org/10.58218/alinea.v5i3.1978>

- Badan Pengembangan dan Pembinaan Bahasa. (2015). *Pembahasan materi dalam forum diskusi tenaga ahli bahasa di Kepolisian dan DPR*. Kementerian Pendidikan dan Kebudayaan Republik Indonesia. <https://badanbahasa.kemendikdasmen.go.id/kegiatan-detail/1803/pembahasan-materi-dalam-forum-diskusi-tenaga-ahli-bahasa-di-kepolisian-dan-dpr>
- Badan Pengembangan dan Pembinaan Bahasa. (2026). *KBBI VI Daring*. Kementerian Pendidikan Dasar dan Menengah Republik Indonesia. <https://kbbi.kemendikdasmen.go.id/>
- Bajuri, M. Z. J., Rahman, F., & Ilma, A. A. (2025). Perkembangan bahasa di media sosial: Dari bahasa gaul hingga singkatan populer. *Pujangga*, 10(2), 150–166. <https://doi.org/10.47313/pujangga.v10i2.3910>
- Behforouz, B., & Frumuselu, A. D. (2021). The effect of text messaging on EFL learners' lexical depth and breadth. *Journal of Language and Education*, 7(4). <https://doi.org/10.17323/jle.2021.11469>
- Chaer, A. (2009). *Linguistik umum*. Rineka Cipta.
- Chen, T., & Zhang, D. (2023). Different aspects of vocabulary depth knowledge in L2 Chinese reading comprehension: Comparing higher- and lower-proficiency readers. *Foreign Language Annals*, 56(3), 786–806. <https://doi.org/10.1111/flan.12687>
- Dagnaw, A. T. (2023). Revisiting the role of breadth and depth of vocabulary knowledge in reading comprehension. *Cogent Education*, 10(1), 2217345. <https://doi.org/10.1080/2331186X.2023.2217345>
- Dang, T. N. Y. (2022). Relationships between lexical coverage, learner knowledge, and teacher perceptions of the usefulness of high-frequency words. *Foreign Language Annals*. <https://doi.org/10.1111/flan.12663>
- Fitria, T. N. (2023). Word formation analysis of foreign terms in information and communication technology based on General Guidelines of Formation Terms in Indonesia. *Journal of Language Intelligence and Culture*, 5(2), 159–180. <https://doi.org/10.35719/jlic.v5i2.116>
- Glewwe, E. (2021). The phonological determinants of tone in English loanwords in Mandarin. *Phonology*. <https://doi.org/10.1017/S0952675721000154>
- Hanafiah, W., et al. (2022). The impact of CALL on vocabulary learning, speaking skill, and foreign language speaking anxiety: The case study of Indonesian EFL learners. *Education Research International*. <https://doi.org/10.1155/2022/5500077>
- Herring, S. C. (2022). Grammar and electronic communication. In *The International Encyclopedia of Language and Social Interaction*. <https://doi.org/10.1002/9781405198431.wbeal0466.pub2>
- Khoiriyah, A. M., Salsa, E. E., Lontoh, R. R., & Ertinawati, Y. (2025). Pembinaan Bahasa Indonesia di ruang digital: Studi pemanfaatan Instagram untuk pengenalan padanan kata. *CENDEKIA: Jurnal Ilmu Sosial, Bahasa dan Pendidikan*, 5(4), 57–66. <https://doi.org/10.55606/cendekia.v5i4.7539>
- Koslow, A. (2022). Meaning change and changing meaning. *Synthese*, 200, 94. <https://doi.org/10.1007/s11229-022-03563-8>
- Kridalaksana, H. (2008). *Pembentukan kata dalam bahasa Indonesia*. PT Gramedia Pustaka Utama.
- Ludewig, U., Hübner, N., & Schroeder, S. (2023). Vocabulary, text coverage, word frequency and the lexical threshold in elementary school reading comprehension. *Reading and Writing*, 36, 2409–2431. <https://doi.org/10.1007/s11145-022-10385-0>
- Luo, Y., Song, H., Wan, L., & Zhang, X. (2021). The effect of vocabulary depth and breadth on English listening comprehension can depend on how comprehension is measured. *Frontiers in Psychology*, 12, 657573. <https://doi.org/10.3389/fpsyg.2021.657573>
- Mutiara, P., Pebriani, Y., & Dwinitia, S. (2024). Penguasaan kosakata bahasa Indonesia pada mahasiswa Myanmar dalam BIPA di Universitas Negeri Padang. *Alinea: Jurnal Bahasa, Sastra dan Pengajaran*, 4(1), 62–71. <https://doi.org/10.58218/alinea.v4i1.812>
- Ningsih, S., & Pallawagau, B. (2025). Semantik kata *rahmah* sebagai kasih sayang universal: Relevansi terhadap krisis ekologis dan pemanasan global. *Alinea: Jurnal Bahasa, Sastra dan Pengajaran*, 5(3), 813–824. <https://doi.org/10.58218/alinea.v5i3.1894>

- Pan, D. J., Nakayama, M., McBride, C., Cheah, Z. R. E., Zheng, M., & Yeung, C. C. L. (2023). Cognitive-linguistic skills and vocabulary knowledge breadth and depth in children's L1 Chinese and L2 English. *Applied Psycholinguistics*, 44(1), 77–99. <https://doi.org/10.1017/S0142716422000480>
- Pratama, I. U., Sari, R. P., Setianingsih, D., & Ariana, D. (2026). Perubahan penggunaan bahasa akibat perkembangan teknologi digital. *Dialektika: Jurnal Kependidikan, Kebahasaan, dan Kesastraan Indonesia*, 3(1). <https://doi.org/10.21154/dialektika.v3i1.13152>
- Pusat Bahasa. (2007). *Pedoman umum pembentukan istilah* (Edisi ketiga, cetakan keempat). Departemen Pendidikan Nasional.
- Sánchez, A., Carreiras, M., & Paz-Alonso, P. M. (2023). Word frequency and reading demands modulate brain activation in the inferior frontal gyrus. *Scientific Reports*, 13, 17217. <https://doi.org/10.1038/s41598-023-44420-z>
- Schurz, A. (2024). Measuring teenage learners' automatized, explicit, and/or implicit knowledge: A question of context? *Language Learning*, 74(2), 506–541. <https://doi.org/10.1111/lang.12624>
- Stewart, J., Gyllstad, H., Nicklin, C., & McLean, S. (2024). Establishing meaning recall and meaning recognition vocabulary knowledge as distinct psychometric constructs in relation to reading proficiency. *Language Testing*, 41(1). <https://doi.org/10.1177/02655322231162853>
- Syamsuri, A. S. (2011). Keberterimaan istilah bahasa Indonesia. *Kajian Linguistik dan Sastra*, 23(2), 109–122. <https://doi.org/10.23917/cls.v23i2.4307>
- Tan, L. Y., McLean, S., Kim, Y. A., et al. (2024). Rasch modelling vs. item facility: Implications on the validity of assessments of Asian EFL/ESL vocabulary knowledge and lexical sophistication modelling. *Language Testing in Asia*, 14, 55. <https://doi.org/10.1186/s40468-024-00327-8>
- Tong, X., & Tong, S. X. (2022). How vocabulary breadth and depth influence bilingual reading comprehension: Direct and indirect pathways. *Learning and Individual Differences*, 100, 102227. <https://doi.org/10.1016/j.lindif.2022.102227>
- Wang, T., & Zhang, H. (2025). Reciprocal effects of vocabulary breadth, vocabulary depth, and reading comprehension: A cross-lagged panel analysis in Chinese-speaking EFL learners. *Humanities and Social Sciences Communications*, 12, 410. <https://doi.org/10.1057/s41599-025-04694-2>
- Zeng, H. (2022). Loanwords adaptation: A corpus study of Sino–Japanese *kan-on* initials. *Studies in Language Sciences*, 20(1), 9–31. https://doi.org/10.34609/sls.20.1_9
- Zhu, T., Zhang, Y., & Irwin, D. (2024). Second and foreign language vocabulary learning through digital reading: A meta-analysis. *Education and Information Technologies*, 29, 4531–4563. <https://doi.org/10.1007/s10639-023-11969-1>

About the Authors

Apri Damai Sagita Krissandi from the Sanata Dharma University, Indonesia. His research interests include Indonesian Language Education, with a focus on children's literature.

Email: apridamai@usd.ac.id

Nandita Wana Putri from the Panca Bhakti School of Health Sciences, Bandar Lampung, Indonesia. Her research interests include Language Education and Linguistics.

Email: nanditawana@gmail.com