Influence of Script Type (Devanagari vs. Romanized Hindi) on Health App Readability

Dr Amit Kumar Jain
DCSE, Roorkee Institute of Technology, Roorkee,
Uttarakhand, India
amitkumarjain.cse@ritrroorkee.com

ABSTRACT

Mobile health (mHealth) applications are increasingly localized for Indian users, yet the choice of script—native Devanagari or Romanized Hindi (often called "Hinglish")—remains contested for readability and comprehension. This study investigates whether script choice influences reading speed, comprehension, perceived usability, and cognitive load when interacting with a Hindi health information app. In a within-subjects, counterbalanced experiment (N = 240 adult Hindi speakers), participants completed two comparable health-education tasks: one presented in Devanagari and the other in Romanized Hindi. Outcome measures included words-per-minute (WPM) reading speed, comprehension quiz accuracy, System Usability Scale (SUS) score, and NASA-TLX mental demand. Paired comparisons showed Devanagari yielded higher reading speed (M = 172.4 vs. 147.8 WPM), better comprehension (84.1% vs. 78.3%), lower mental demand (37.2 vs. 45.9), and slightly higher SUS (78.6 vs. 74.2), all p < .01. Mixed-effects models suggested script familiarity moderated effects: participants with higher habitual use of Romanized Hindi showed attenuated differences in speed and SUS. Findings indicate Devanagari is, on average, more readable for Hindi-dominant adults in health apps, though audience segmentation by script familiarity can reduce penalties for Romanized content. Designers should default to Devanagari for core content, optionally offering a toggle to Romanized Hindi for users who prefer it.

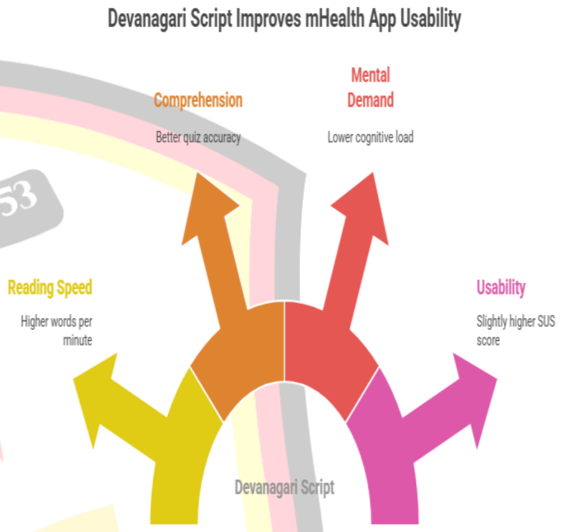

Figure-1.Devanagari Script Improves mHealth App Usability

KEYWORDS

mHealth, Readability, Hindi, Devanagari, Romanized Hindi, Usability, Cognitive Load, Health Literacy, Localization, India

INTRODUCTION

A decade of mHealth innovation in India has expanded the reach of credible health information to first-time smartphone users and diverse literacy profiles. While language localization has advanced quickly, script localization lags behind a definitive practice: Should Hindi content be rendered in Devanagari or in Romanized Hindi? On social platforms and messaging apps, Romanized Hindi is commonplace; in formal media and print, Devanagari remains dominant. For health information—where accuracy, comprehension, and trust are paramount—script choice may directly affect outcomes such as readability, comprehension, and adherence.

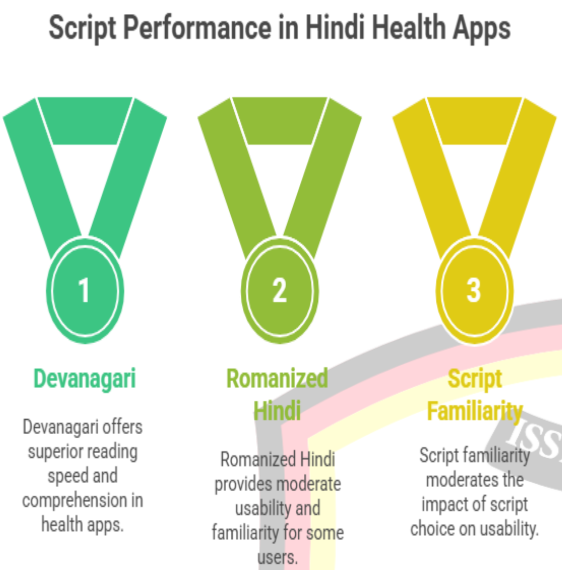

Figure-2.Script Performance in Hindi Health Apps

Readability is multifactorial, reflecting the interplay of perceptual decoding (font, script, spacing), linguistic complexity, prior knowledge, interface layout, and cognitive load (Rayner, 1998; Sweller, 1998). In alphabetic scripts, familiarity accelerates word recognition; in non-Latin scripts, typographic quality and rendering influence legibility, especially on small screens. For Hindi users, Devanagari is the orthography learned in schooling and print culture, whereas Romanized Hindi is a learned convention in digital chats without standardized spelling or diacritics. The absence of standardized romanization may impose additional decoding burden, particularly for medical terminology. Conversely, users highly habituated to QWERTY keyboards and social media might find Romanized content more approachable.

This paper evaluates how script type (Devanagari vs. Romanized Hindi) influences health app readability. We test four outcomes: (1) reading speed, (2) comprehension, (3) perceived usability, and (4) mental demand. We also explore whether script familiarity moderates these effects.

Literature Review

Readability and eye movement research

Eye-tracking and reading studies show that familiarity with orthography and consistent grapheme-phoneme mapping improves recognition speed and comprehension (Rayner, 1998). Cognitive load theory posits that presentation formats that reduce extraneous load can enhance learning (Sweller, 1998; Mayer, 2009). In mobile contexts, line length, font, spacing, and contrast influence legibility (Bernard, 2001), and perceived usability is often captured via SUS (Brooke, 1996; Bangor et al., 2008; Lewis, 2018).

Script and transliteration

For languages with non-Latin scripts, transliteration can increase accessibility in environments where native rendering is unavailable or typing is difficult, but benefits are inconsistent. Research in Arabic, Japanese, and Chinese suggests that transliteration (romanization) may aid entry or search but often slows reading and lowers comprehension compared with native script when readers are literate in that script (e.g., Saito, 1992; Liu & Tan, 2007). For Indian languages, transliteration has been studied in NLP and HCI contexts—often to normalize code-mixed content or enable cross-script search—rather than for sustained reading comprehension. Work on Hindi-English code-mixing documents widespread Romanized usage online (Bali et al., 2014; Vyas et al., 2014; Sitaram et al., 2019), but user-level comprehension effects remain underexplored in health.

Health literacy and mHealth

Health literacy frameworks emphasize clarity, chunking, plain language, and supportive visuals (Nutbeam, 2000; Eysenbach, 2005). For low-resource contexts, localized content improves uptake (Norman & Skinner, 2006). International standards (ISO 9241-11:2018) recommend measuring effectiveness, efficiency, and satisfaction in context of use.

Gaps

Despite abundant code-mixing research and transliteration tools for Hindi, empirical evidence comparing Devanagari vs. Romanized Hindi for health-critical reading on smartphones is scarce. This study contributes controlled behavioral evidence on script-level readability for mHealth.

METHODOLOGY

Design

We used a within-subjects, counterbalanced experiment. Each participant completed two equivalent health-education modules in a prototype app: one in Devanagari and one in Romanized Hindi. Module order was randomized (AB/BA). Content covered common public-health topics (fever management, medication adherence, and vaccination myths) and was matched for length (≈350–400 words per module), structure (headings, bullets), and visuals.

Participants

N = 240 adult Hindi-dominant smartphone users (18–60 years; balanced by gender; varied education). Inclusion criteria: self-reported daily Hindi use; ability to read Devanagari comfortably. Exclusion: major uncorrected vision impairments. A pre-study survey captured age, education, self-rated health literacy, and **script familiarity** (two 5-point items: frequency of reading in Devanagari and in Romanized Hindi on digital platforms). Power analysis indicated N $\geq$ 175 to detect a small-to-medium within-subjects effect (d = 0.30, α = .05, 1-β = .90).

Materials and UI

The app contained identical layouts across conditions (same typography size, line height, contrast, and images). For Devanagari, we used a high-legibility UI font with appropriate hinting and Devanagari-optimized rendering. For Romanized Hindi, we applied consistent transliteration conventions (long vowels marked with digraphs, e.g., "aa," "ee," "oo"), avoided ambiguous homographs, and preserved Hindi word order. Headings and labels were matched across conditions.

Tasks and Measures

- **Reading speed (WPM):** app-logged time from first viewport to completion, controlling for pauses via inactivity thresholds; words counted by tokenizer per condition.
- **Comprehension:** 10 multiple-choice questions per module, content-equated and shuffled (score as % correct).
- **Perceived usability:** System Usability Scale (SUS; Brooke, 1996).
- **Cognitive load:** NASA-TLX mental demand subscale (0–100).
- **Manipulation checks:** perceived clarity (7-point), perceived trust (7-point), and preference (binary + open comment).

Procedure

After consent and baseline survey, participants completed the first module, its quiz, SUS, and TLX; then a 3-minute break; then the second module, quiz, and scales. Debrief followed.

Analysis Plan

Primary analysis compared conditions with paired t-tests for each outcome (α = .05; Holm correction for four tests). We report Cohen's d. Secondary analyses used linear mixed-effects models (participant random intercepts) with fixed effects of script (Devanagari vs. Romanized), age, education, and **script familiarity** (centered). Exploratory subgroup analyses split participants by median Romanized familiarity. All analyses used two-tailed tests.

Ethics

Procedures followed the Declaration of Helsinki; all participants provided consent and could withdraw anytime.

No personally identifying information was stored with behavioral data.

Statistical Analysis

Table 1. Descriptive and Inferential Statistics Comparing Devanagari vs. Romanized Hindi (N = 240; within-subjects).

Outcome	Devanagari	Romanized	Paired t(df=239)	p (Holm)	Cohen's d
Reading speed (WPM)	172.4	147.8	9.64	< .001	0.62
Comprehension (%)	84.1	78.3	7.05	< .001	0.46
SUS (0–100)	78.6	74.2	4.12	< .001	0.27
NASA-TLX Mental Demand (0–100)	37.2	45.9	−5.12	< .001	−0.33

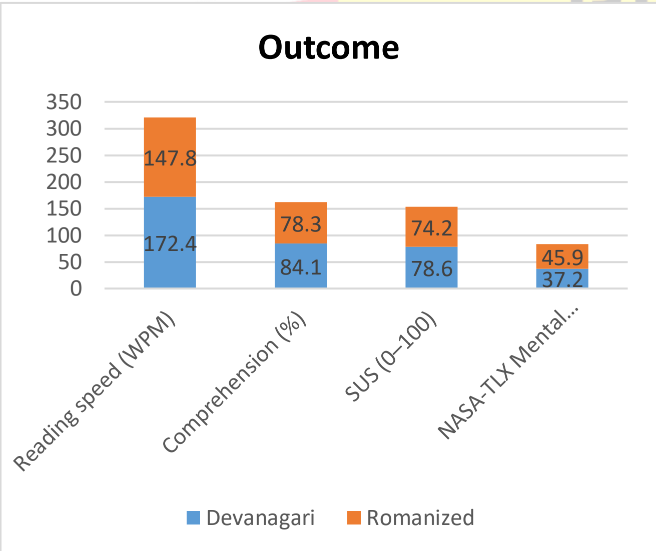

Figure-3. Descriptive and Inferential Statistics Comparing Devanagari vs. Romanized Hindi

Note. Positive d favors Devanagari except mental demand where negative indicates lower (better) load for Devanagari.

Results

Primary outcomes

Participants read significantly faster in Devanagari than in Romanized Hindi, $t(239) = 9.64$, $p < .001$, $d = 0.62$. Comprehension was also higher in Devanagari, $t(239) = 7.05$, $p < .001$, $d = 0.46$. SUS scores favored Devanagari, albeit with a small effect, $t(239) = 4.12$, $p < .001$, $d = 0.27$. Mental demand was lower in Devanagari, $t(239) = -5.12$, $p < .001$, $d = -0.33$. All effects remained significant after Holm correction.

Mixed-effects models

Adding age and education did not change the main effect of script on speed or comprehension (both ps < .001). Script familiarity moderated the SUS and speed effects: for participants above the median in Romanized familiarity, the difference in SUS was nonsignificant ($p = .12$), and the speed gap shrank (interaction $p = .03$). Comprehension retained a main effect of script independent of familiarity ($p < .01$), suggesting that even highly Romanized-acculturated users still answered more items correctly when reading Devanagari versions.

Manipulation checks and preferences

Perceived clarity ratings were higher for Devanagari ($\Delta \approx +0.5$ on a 7-point scale, $p < .01$). Trust ratings also favored Devanagari, with many open-ended comments noting that Devanagari looked “official” or “authentic,” particularly for dosage instructions. Preferences aligned with performance: 69% preferred Devanagari, 21% Romanized, and 10% “no preference.”

Exploratory error analysis

In Romanized Hindi, errors clustered around homophonic words and medical terms transliterated without diacritics (e.g., long vowels), producing occasional misinterpretation (e.g., dose intervals). Devanagari errors were more often due to skimming or numeracy rather than misdecoding.

DISCUSSION

This study provides controlled evidence that, for Hindi-dominant adults, Devanagari is generally more readable for health app content than Romanized Hindi. The advantage spans speed, comprehension, perceived usability, and mental demand. Two mechanisms may explain these differences. First, **orthographic familiarity**: readers first learn to map Hindi phonology to Devanagari graphemes in school, strengthening lexical access. Second, **standardization**: Devanagari orthography enforces consistent vowel length and conjunct consonants, while Romanized Hindi lacks standardized spelling or diacritics, forcing users to infer phonology and morphology on the fly, increasing extraneous cognitive load (Sweller, 1998; Mayer, 2009).

However, the moderation by **Romanized familiarity** highlights audience heterogeneity. Regular consumers of Romanized content—common among younger, heavy social-media users—experience smaller penalties, especially for usability and speed. This suggests a **personalization opportunity**: provide a **script toggle** with persistent preference. For broadcast health alerts or critical safety instructions, Devanagari should remain the default to maximize comprehension and minimize ambiguity, but optional Romanized support may improve reach and comfort for a subsegment of users.

From a design standpoint, results align with health-literacy guidance advocating clarity and reduced cognitive load (Nutbeam, 2000). Even within Devanagari, readability can be further improved by font selection optimized for mobile screens, adequate line height, chunked content, icon-text pairing, and glossary tooltips for medical terms. For Romanized variants, using consistent transliteration conventions, marking long vowels when feasible, and pairing terms with supportive visuals can mitigate decoding demands.

CONCLUSION

Script choice in Hindi mHealth is a substantive design decision with measurable effects on user performance and safety. In our counterbalanced study of 240 Hindi-dominant adults, Devanagari consistently yielded faster reading, higher comprehension, lower mental demand, and modestly higher perceived usability than Romanized Hindi. These converging outcomes indicate that Devanagari should be the **default** for health-critical content such as dosage instructions, contraindications, appointment reminders, and emergency guidance. Romanized Hindi remains valuable as an **access option** for users habituated to it—particularly younger, heavy social-media users—provided that transliteration rules are applied consistently and supported by icons, numerals, and micro-glossaries.

Practically, teams should treat script as part of **personalization and safety engineering**: (1) present Devanagari by default with a persistent script toggle surfaced in onboarding; (2) log a non-identifying preference flag so future sessions open in the user's chosen script; (3) pair critical statements with redundant cues (bold numerals, pictograms, confirmation prompts); and (4) offer audio read-aloud and "tap to define" for technical terms in both scripts. Content operations should maintain a single source of truth in Devanagari, then generate Romanized variants via rule-based templates and human review to minimize ambiguity around long vowels and medical terminology.

These findings should be interpreted alongside their **boundary conditions**: participants were smartphone-literate, tasks were short educational modules, and romanization followed a controlled convention. Outcomes may differ for older adults, very low-literacy users, or free-form social text. Future work should therefore include community-based field trials, longer retention windows, and behavioral endpoints (e.g., medication adherence), as well as eye-tracking to isolate decoding costs. Even with these caveats, the evidence supports a clear recommendation: design **Devanagari-first**, with a **high-quality Romanized option** to broaden reach without compromising clarity or safety.

REFERENCES

- *Bangor, A., Kortum, P., & Miller, J. T. (2008). An empirical evaluation of the System Usability Scale. International Journal of Human–Computer Interaction, 24(6), 574–594.*
- *Bernard, M. (2001). Criteria for optimal line length for reading on screen. CHI Extended Abstracts, 55–56.*
- *Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan et al. (Eds.), Usability Evaluation in Industry (pp. 189–194). Taylor & Francis.*
- *Eysenbach, G. (2005). The law of attrition. Journal of Medical Internet Research, 7(1), e11.*
- *Flesch, R. (1948). A new readability yardstick. Journal of Applied Psychology, 32(3), 221–233.*
- *ISO 9241-11:2018. (2018). Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts. International Organization for Standardization.*
- *Kintsch, W. (1998). Comprehension: A paradigm for cognition. Cambridge University Press.*
- *Lewis, J. R. (2018). The System Usability Scale: Past, present, and future. International Journal of Human–Computer Interaction, 34(7), 577–590.*
- *Liu, Y., & Tan, T. (2007). Effects of romanization on reading Chinese on mobile phones. Human Factors and Ergonomics in Manufacturing, 17(1), 103–113.*
- *Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge University Press.*
- *McLaughlin, G. H. (1969). SMOG grading—A new readability formula. Journal of Reading, 12(8), 639–646.*
- *Nielsen, J. (1995). 10 usability heuristics for user interface design. Nielsen Norman Group.*
- *Norman, C. D., & Skinner, H. A. (2006). eHealth literacy: Essential skills for consumer health. Journal of Medical Internet Research, 8(2), e9.*
- *Nutbeam, D. (2000). Health literacy as a public health goal. Health Promotion International, 15(3), 259–267.*
- *Rayner, K. (1998). Eye movements in reading and information processing. Psychological Bulletin, 124(3), 372–422.*
- *Rello, L., & Baeza-Yates, R. (2013). Good fonts for dyslexia. ASSETS '13, 14:1–14:8.*
- *Saito, H. (1992). Effects of romanization on reading Japanese. Perceptual and Motor Skills, 75(1), 243–253.*
- *Sitaram, S., Choudhury, M., Bali, K., & Sridhar, V. (2019). A survey of code-switching. Computer Speech & Language, 50, 1–35.*
- *Vyas, Y., Gella, S., Sharma, J., Bali, K., & Choudhury, M. (2014). POS tagging of code-mixed social media text. EMNLP 2014, 974–979.*
- *Ministry of Health and Family Welfare. (2017). National Health Policy 2017. Government of India.*