Firman Noor Hasan <firman.noorhasan@uhamka.ac.id>

[INSTEK] Editor Decision

1 message

Muhammad Nur Akbar <journal@uin-alauddin.ac.id> Wed, Oct 21, 2026 at 10:17 AM
To: Riska Amelia Puspita Sari <riska7amelia@gmail.com>, Mizan Aldini Aldini <mizanaldinimalik@anadolu.edu.tr>, Firman Noor Hasan <firman.noorhasan@uhamka.ac.id>

Riska Amelia Puspita Sari, Mizan Aldini Aldini, Firman Noor Hasan:

We have reached a decision regarding your submission to Jurnal INSTEK (Informatika Sains dan Teknologi), "Redevelopment of the Alfagift Application Based on Design Thinking with an Initial Assessment of the Application's Usability Through the System Usability Scale (SUS)".

Our decision is: Revisions Required

Reviewer A:
Recommendation: Revisions Required

---

1. Novelty and Scientific Significance

[2] Incremental: Minor variations or small extensions of well-investigated subjects.

2. Technical Soundness and Methodological Rigor

[2] Flawed: Contains methodological weaknesses or lacks appropriate baselines/controls, affecting reliability.

3. Data Availability and Reproducibility

[3] Partially Reproducible: Experiments are well-explained, but data/code is only available upon request or lacks minor details.

4. Presentation, Structure, and Language

[2] Needs Improvement: Requires substantial language editing (major proofreading) or significant restructuring for clarity.

A. Suggestions for improvements: Basic reporting

1) Draf perlu proofreading bahasa Inggris karena terdapat beberapa frasa yang keliru atau salah ketik, misalnya “the sectore-commerce”, “Score YOUR SUS”, “empathize, ... dan test”, dan “convenience, convenience, and satisfaction”.
2) Cantumkan sumber dan tanggal akses untuk data rating aplikasi (Alfagift 4,3 vs Astro 4,9).
3) Rujukan kebanyakan berupa studi kasus Design Thinking dalam negeri. Sebaiknya tambahkan rujukan

tentang SUS itu sendiri dan tentang UX e-commerce.
4) Tolong perbaiki penulisan daftar pustaka:
- [2] dan [3] belum mencantumkan tahun dan penerbit.
- [5] dan [6] belum mencantumkan nama jurnal.
- Nama penulis pada [10] tidak terbaca, cek kembali dan perbaiki penulisannya.

B. Suggestions for improvements: Experimental design

1) Pengujian awal mengukur aplikasi asli dengan 60 responden, sedangkan pengujian akhir mengukur prototipe Figma dengan 30 responden. Mohon dijelaskan apakah 30 responden tersebut bagian dari 60 responden awal, bagaimana responden mencoba prototipe (skenario tugas, dipandu atau mandiri)
2) Kriteria inklusi tidak konsisten. Pada bagian metode 2.3 dituliskan "minimal 1 bulan", sedangkan pada bagian hasil 3.2.1 dituliskan "minimal tiga bulan terakhir".
3) Abstrak dan tahap Empathize dituliskan ada interview/wawancara mendalam, tetapi pada bagian metode tidak dijelaskan.

C. Suggestions for improvements: Validity of the findings

1) Sebaiknya tidak menuliskan kenaikan 26,29 poin sebagai "signifikan" jika tidak didukung dengan hasil uji statistik.
2) Kalimat "13.3% chose the old design and both, respectively" kurang jelas.

Reviewer B:
Recommendation: Resubmit for Review

---

1. Novelty and Scientific Significance

[2] Incremental: Minor variations or small extensions of well-investigated subjects.

2. Technical Soundness and Methodological Rigor

[2] Flawed: Contains methodological weaknesses or lacks appropriate baselines/controls, affecting reliability.

3. Data Availability and Reproducibility

[2] Hard to Reproduce: Limited transparency regarding datasets or experimental setup makes reproduction difficult.

4. Presentation, Structure, and Language

[2] Needs Improvement: Requires substantial language editing (major proofreading) or significant restructuring for clarity.

A. Suggestions for improvements: Basic reporting

Clarify the scope and title. The title should reflect prototype evaluation rather than only “initial usability evaluation.” Clearly distinguish the existing application, the proposed interface, and the functionality actually tested.

Revise the abstract. Report 60 baseline participants and 30 post-redesign participants explicitly. Identify

whether the comparison is paired, independent, or descriptive. Remove unsupported claims of significant improvement, enhanced transaction transparency, or improved user control.

Strengthen the state-of-the-art comparison. Provide a concise comparison with closely related Alfagift and retail application studies, particularly reference [10]. Compare features, design decisions, participants, evaluation procedures, findings, and limitations. Do not characterize previous Alfagift work as merely a general usability assessment without supporting that characterization.

Focus the introduction. Reduce generic definitions of e-commerce, UI, UX, and HCI. Prioritize documented user problems, their consequences, the unresolved research issue, and the study objectives. State the application version and observation period. App ratings should include their platform, retrieval date, and context.

Correct citations and metadata. The DOI listed in reference [6], 10.29100/jipi.v4i1.781, identifies an article about an educational color-recognition game rather than the cited GTA V Roleplay study. Verify and correct this entry. Also correct the author and journal metadata in reference [10], and complete incomplete web references.

Improve tables and figures. Define “Total Score” as the sum of transformed item contributions if that is how it was calculated. Use consistent numbering, readable captions, and clear before–after labels. Consider moving lengthy respondent tables to supplementary materials and presenting distributions and summary statistics in the article.

Edit the manuscript comprehensively. Correct wording such as “Score YOUR SUS,” mixed-language passages, inconsistent headings, and fragmented equations. Align the final manuscript with the current INSTEK template.

B. Suggestions for improvements: Experimental design

Audit the administered instrument. Compare Table 4 with the original questionnaire. Several items replace standard SUS content with recommendation intention, feature findability, confusion, or frustration. If the original questionnaire was standard SUS, correct the reporting using the actual responses. If substantively modified items were administered, do not retrospectively relabel them as standard SUS. Either repeat the evaluation with an appropriate SUS version or justify the adapted instrument and its interpretation.

Explain participant relationships. State whether the 30 post-redesign participants belong to the original 60. Provide anonymized matching identifiers where available. Do not assume that identical eligibility criteria or row numbers establish paired observations.

Resolve eligibility inconsistencies. Clarify the difference between the minimum one-month usage criterion and the three-month usage statement in the results. Describe recruitment, sampling, participant characteristics, and the rationale for restricting ages to 18–34.

Specify the qualitative procedures. Report the number and selection of app reviews and interview participants, collection periods, interview questions, coding procedures, and supporting evidence. Explain how qualitative findings informed design priorities. If in-depth interviews were not conducted, remove that claim.

Document Design Thinking implementation. Present the evidence, decisions, and outputs of each stage, including alternative solutions and prioritization criteria. Explain how quantitative and qualitative components were integrated if the study is described as mixed methods.

Use explicit task scenarios. Evaluate store selection, interpretation of delivery status, and retrieval of transaction history. Record task completion, errors, and completion time alongside perceived usability. State whether participants interacted with the prototype or only viewed screenshots.

Ensure comparable testing conditions. Describe device, task difficulty, instructions, assistance, and evaluation order. Where feasible, compare equivalent representations of the old and new interfaces and counterbalance their order.

Clarify prototype limitations. Distinguish simulated tracking displays from implemented real-time functionality. Discuss operational constraints such as delivery coverage, inventory availability, and data update delays.

Provide reproducibility and ethics information. Include the questionnaire, task protocol, anonymized data, prototype materials, consent procedures, and data protection arrangements.

C. Suggestions for improvements: Validity of the findings

Separate arithmetic accuracy from methodological validity. The reported totals support means of 48.875 and 75.1667, giving a descriptive difference of approximately 26.29 points. Correct arithmetic does not establish that the instrument is standard SUS or that the redesign caused the difference.

Use an appropriate comparison. If participant matching is available, compare baseline and post-redesign scores for the same participants. If samples are genuinely independent, justify that design and account for group differences. If overlap exists but matching cannot be recovered, acknowledge the analytical limitation.

Report uncertainty and variability. Include standard deviations, confidence intervals, and an appropriate effect size. Select statistical tests according to the actual design and assumptions. Remove “significant improvement” unless supported by a valid statistical analysis.

Address incomplete follow-up. Explain why only 30 of the 60 invited respondents completed the second evaluation, where known. Compare baseline characteristics and scores of completers and non-completers where possible, and discuss potential attrition bias.

Limit feature-level interpretation. A global usability score cannot establish the separate effectiveness of store selection, tracking, and order history. Support feature-specific claims through task outcomes or additional evidence. Treat individual questionnaire item analyses as exploratory.

Interpret preference cautiously. Report the participant count and denominator for each preference category and explain what “both” means. Preference for the new interface does not independently establish better task performance.

Expand critical discussion. Consider familiarity, visual appeal, differences between production software and a prototype, participant selection, and testing conditions as alternative explanations. Compare methods and design mechanisms with previous research rather than directly ranking studies by their SUS means.

Align conclusions with evidence. Present the results as preliminary findings concerning the evaluated interface and participant group. Claims about transaction transparency, user control, operational efficiency, loyalty, or business outcomes require direct measurement. Explicitly acknowledge instrument uncertainty, incomplete follow-up, limited participant diversity, and prototype-based evaluation.

Reviewer A:
Recommendation: Revisions Required

---

1. Novelty and Scientific Significance

[2] Incremental: Minor variations or small extensions of well-investigated subjects.

2. Technical Soundness and Methodological Rigor

[2] Flawed: Contains methodological weaknesses or lacks appropriate baselines/controls, affecting reliability.

3. Data Availability and Reproducibility

[3] Partially Reproducible: Experiments are well-explained, but data/code is only available upon request or lacks minor details.

4. Presentation, Structure, and Language

[2] Needs Improvement: Requires substantial language editing (major proofreading) or significant restructuring for clarity.

A. Suggestions for improvements: Basic reporting

1) Draf perlu proofreading bahasa Inggris karena terdapat beberapa frasa yang keliru atau salah ketik, misalnya “the sectore-commerce”, “Score YOUR SUS”, “empathize, ... dan test”, dan “convenience, convenience, and satisfaction”.
2) Cantumkan sumber dan tanggal akses untuk data rating aplikasi (Alfagift 4,3 vs Astro 4,9).
3) Rujukan kebanyakan berupa studi kasus Design Thinking dalam negeri. Sebaiknya tambahkan rujukan tentang SUS itu sendiri dan tentang UX e-commerce.
4) Tolong perbaiki penulisan daftar pustaka:
- [2] dan [3] belum mencantumkan tahun dan penerbit.
- [5] dan [6] belum mencantumkan nama jurnal.
- Nama penulis pada [10] tidak terbaca, cek kembali dan perbaiki penulisannya.

B. Suggestions for improvements: Experimental design

1) Pengujian awal mengukur aplikasi asli dengan 60 responden, sedangkan pengujian akhir mengukur prototipe Figma dengan 30 responden. Mohon dijelaskan apakah 30 responden tersebut bagian dari 60 responden awal, bagaimana responden mencoba prototipe (skenario tugas, dipandu atau mandiri)
2) Kriteria inklusi tidak konsisten. Pada bagian metode 2.3 dituliskan "minimal 1 bulan", sedangkan pada bagian hasil 3.2.1 dituliskan "minimal tiga bulan terakhir".
3) Abstrak dan tahap Empathize dituliskan ada interview/wawancara mendalam, tetapi pada bagian metode tidak dijelaskan.

C. Suggestions for improvements: Validity of the findings

1) Sebaiknya tidak menuliskan kenaikan 26,29 poin sebagai "signifikan" jika tidak didukung dengan hasil uji statistik.
2) Kalimat "13.3% chose the old design and both, respectively" kurang jelas.

Reviewer B:
Recommendation: Resubmit for Review

---

1. Novelty and Scientific Significance

[2] Incremental: Minor variations or small extensions of well-investigated subjects.

2. Technical Soundness and Methodological Rigor

[2] Flawed: Contains methodological weaknesses or lacks appropriate baselines/controls, affecting reliability.

3. Data Availability and Reproducibility

[2] Hard to Reproduce: Limited transparency regarding datasets or experimental setup makes reproduction difficult.

4. Presentation, Structure, and Language

[2] Needs Improvement: Requires substantial language editing (major proofreading) or significant restructuring for clarity.

A. Suggestions for improvements: Basic reporting

Clarify the scope and title. The title should reflect prototype evaluation rather than only “initial usability evaluation.” Clearly distinguish the existing application, the proposed interface, and the functionality actually tested.

Revise the abstract. Report 60 baseline participants and 30 post-redesign participants explicitly. Identify whether the comparison is paired, independent, or descriptive. Remove unsupported claims of significant improvement, enhanced transaction transparency, or improved user control.

Strengthen the state-of-the-art comparison. Provide a concise comparison with closely related Alfagift and retail application studies, particularly reference [10]. Compare features, design decisions, participants, evaluation procedures, findings, and limitations. Do not characterize previous Alfagift work as merely a general usability assessment without supporting that characterization.

Focus the introduction. Reduce generic definitions of e-commerce, UI, UX, and HCI. Prioritize documented user problems, their consequences, the unresolved research issue, and the study objectives. State the application version and observation period. App ratings should include their platform, retrieval date, and context.

Correct citations and metadata. The DOI listed in reference [6], 10.29100/jipi.v4i1.781, identifies an article about an educational color-recognition game rather than the cited GTA V Roleplay study. Verify and correct this entry. Also correct the author and journal metadata in reference [10], and complete incomplete web references.

Improve tables and figures. Define “Total Score” as the sum of transformed item contributions if that is how it was calculated. Use consistent numbering, readable captions, and clear before–after labels. Consider moving lengthy respondent tables to supplementary materials and presenting distributions and summary statistics in the article.

Edit the manuscript comprehensively. Correct wording such as “Score YOUR SUS,” mixed-language passages, inconsistent headings, and fragmented equations. Align the final manuscript with the current INSTEK template.

B. Suggestions for improvements: Experimental design

Audit the administered instrument. Compare Table 4 with the original questionnaire. Several items replace standard SUS content with recommendation intention, feature findability, confusion, or frustration. If the original questionnaire was standard SUS, correct the reporting using the actual responses. If substantively modified items were administered, do not retrospectively relabel them as standard SUS. Either repeat the evaluation with an appropriate SUS version or justify the adapted instrument and its interpretation.

Explain participant relationships. State whether the 30 post-redesign participants belong to the original 60. Provide anonymized matching identifiers where available. Do not assume that identical eligibility criteria or row numbers establish paired observations.

Resolve eligibility inconsistencies. Clarify the difference between the minimum one-month usage criterion and the three-month usage statement in the results. Describe recruitment, sampling, participant characteristics, and the rationale for restricting ages to 18–34.

Specify the qualitative procedures. Report the number and selection of app reviews and interview participants, collection periods, interview questions, coding procedures, and supporting evidence. Explain how qualitative findings informed design priorities. If in-depth interviews were not conducted, remove that claim.

Document Design Thinking implementation. Present the evidence, decisions, and outputs of each stage, including alternative solutions and prioritization criteria. Explain how quantitative and qualitative components were integrated if the study is described as mixed methods.

Use explicit task scenarios. Evaluate store selection, interpretation of delivery status, and retrieval of transaction history. Record task completion, errors, and completion time alongside perceived usability. State whether participants interacted with the prototype or only viewed screenshots.

Ensure comparable testing conditions. Describe device, task difficulty, instructions, assistance, and evaluation order. Where feasible, compare equivalent representations of the old and new interfaces and counterbalance their order.

Clarify prototype limitations. Distinguish simulated tracking displays from implemented real-time functionality. Discuss operational constraints such as delivery coverage, inventory availability, and data update delays.

Provide reproducibility and ethics information. Include the questionnaire, task protocol, anonymized data, prototype materials, consent procedures, and data protection arrangements.

C. Suggestions for improvements: Validity of the findings

Separate arithmetic accuracy from methodological validity. The reported totals support means of 48.875 and 75.1667, giving a descriptive difference of approximately 26.29 points. Correct arithmetic does not establish that the instrument is standard SUS or that the redesign caused the difference.

Use an appropriate comparison. If participant matching is available, compare baseline and post-redesign scores for the same participants. If samples are genuinely independent, justify that design and account for group differences. If overlap exists but matching cannot be recovered, acknowledge the analytical limitation.

Report uncertainty and variability. Include standard deviations, confidence intervals, and an appropriate effect size. Select statistical tests according to the actual design and assumptions. Remove “significant improvement” unless supported by a valid statistical analysis.

Address incomplete follow-up. Explain why only 30 of the 60 invited respondents completed the second evaluation, where known. Compare baseline characteristics and scores of completers and non-completers where possible, and discuss potential attrition bias.

Limit feature-level interpretation. A global usability score cannot establish the separate effectiveness of store selection, tracking, and order history. Support feature-specific claims through task outcomes or additional evidence. Treat individual questionnaire item analyses as exploratory.

Interpret preference cautiously. Report the participant count and denominator for each preference category and explain what “both” means. Preference for the new interface does not independently establish better task performance.

Expand critical discussion. Consider familiarity, visual appeal, differences between production software and a prototype, participant selection, and testing conditions as alternative explanations. Compare methods and design mechanisms with previous research rather than directly ranking studies by their SUS means.

Align conclusions with evidence. Present the results as preliminary findings concerning the evaluated interface and participant group. Claims about transaction transparency, user control, operational efficiency, loyalty, or business outcomes require direct measurement. Explicitly acknowledge instrument uncertainty, incomplete follow-up, limited participant diversity, and prototype-based evaluation.

Jurnal INSTEK (Informatika Sains dan Teknologi)
http://journal.uin-alauddin.ac.id/index.php/instek

2 attachments

C-69935_review.docx
1660K

B-Hasil Review Naskah INSTEK - UIUX Redesign of the Alfagift Application Based on Design Thinking and Initial Usability Evaluation Using the System Usability Scale.pdf
308K