

REDESAIN APLIKASI *MOBILE* SUPERVISI OBSERVASI KELAS UNTUK MENINGKATKAN TINGKAT *USABILITY*

Muhammad Ifa Amrillah^{*1}, Retno Indah Rokhmawati², Fitra Abdurrachman Bachtiar³

^{1,2,3}Universitas Brawijaya, Malang

Email: ¹muafalah@student.ub.ac.id, ²retnoindah@ub.ac.id, ³fitra.bachtiar@ub.ac.id

*Penulis Korespondensi

(Naskah masuk: 08 November 2023, diterima untuk diterbitkan: 08 Desember 2023)

Abstrak

Aplikasi *LESON*, yang kini berganti nama menjadi *EDVISOR* merupakan aplikasi berbasis *mobile* yang dikhususkan untuk merekam suasana pembelajaran dan mendukung kegiatan supervisi observasi kelas. Meskipun telah mengalami pembaharuan, pada penelitian sebelumnya mengungkapkan masih terdapat 24 permasalahan pada aplikasi yang perlu perbaikan dan pengembangan lebih lanjut. Metode *Usability Testing* dipilih untuk dapat mengidentifikasi permasalahan secara lebih mendalam dan memahami pengalaman pengguna dalam menggunakan aplikasi. Penelitian ini melibatkan 10 Mahasiswa dan 3 Dosen yang diberikan tugas khusus untuk menguji aplikasi. Pengujian dilakukan sebanyak 2 kali, yaitu sebelum dan setelah dilakukan perbaikan desain. Hasil observasi dan wawancara pada pengujian awal mengungkapkan adanya 42 permasalahan baru, yang kemudian menjadi dasar rekomendasi perbaikan desain. Perbaikan desain menghasilkan *prototype* yang mendekati hasil akhir dari aplikasi. Setelah perbaikan desain, *prototype* diuji kembali kepada pengguna yang sama dengan pengujian awal. Hasil pengujian menunjukkan peningkatan *usability* aplikasi, dengan peningkatan *completion rate* sebesar 13,19%, peningkatan *efficiency* sebesar 0,016 *goals/sec*, dan peningkatan *satisfaction* sebesar 1,21.

Kata kunci: *pengujian aplikasi, observasi kelas, pengalaman pengguna, usability*

REDESIGNING A *MOBILE* CLASSROOM SUPERVISION OBSERVATION APPLICATION TO ENHANCE *USABILITY* LEVEL

Abstract

The *LESON* application, now renamed *EDVISOR*, is a mobile-based application devoted to recording the learning atmosphere and supporting classroom observation supervision activities. Although it has undergone updates, previous research revealed that there are still 24 problems in the application that need further improvement and development. *Usability Testing* method was chosen to be able to identify problems more deeply and understand the user experience in using the application. This research involved 10 students and 3 lecturers who were given a special task to test the application. Testing was carried out twice, namely before and after design improvements were made. The results of observations and interviews in the initial test revealed 42 new problems, which then became the basis for recommendations for design improvements. Design improvements resulted in a *prototype* that was close to the final result of the application. After the design improvements, the *prototype* was tested again with the same users as the initial test. The test results showed an increase in application usability, with an increase in completion rate of 13.19%, an increase in efficiency of 0.016 *goals/sec*, and an increase in satisfaction of 1.21.

Keywords: *application testing, classroom observation, user experience, usability*

1. PENDAHULUAN

Peningkatan kualitas mutu pendidikan adalah salah satu hal yang sangat diharapkan dalam dunia pendidikan (Akhmadjonovich, 2021). Supervisi menjadi salah satu kegiatan yang melibatkan pemantauan dan pelatihan guru untuk meningkatkan kemampuan mengajar (Setia dan Nasrudin, 2020),

sedangkan observasi kelas sendiri merupakan metode pengumpulan informasi yang digunakan untuk mengamati pembelajaran di dalam kelas (Cunningham, Gorman dan Maher, 2020). Proses supervisi observasi kelas seringkali dihadapkan pada tantangan dalam pengumpulan data karena masih dilaksanakan secara manual menggunakan secarik

kertas, menyebabkan potensi kesalahan dalam memberikan masukan atau komentar kepada guru yang diobservasi (Efendi, Fanani dan Supianto, 2020).

Dalam era teknologi informasi saat ini, aplikasi berbasis teknologi menawarkan solusi menarik untuk mendukung proses supervisi observasi kelas. Aplikasi dapat mempermudah pengumpulan data, analisis, dan pelaporan, serta meningkatkan efisiensi dan efektivitas supervisi pendidikan. Salah satu aplikasi yang relevan adalah aplikasi LESON, sebuah aplikasi *mobile* berbasis *Android* yang dikhususkan untuk merekam suasana pembelajaran, mendokumentasikan kegiatan di dalam kelas dan membantu memberikan penilaian yang dapat digunakan sebagai evaluasi pembelajaran (Wibawa, Rokhmawati dan Az-Zahra, 2021).

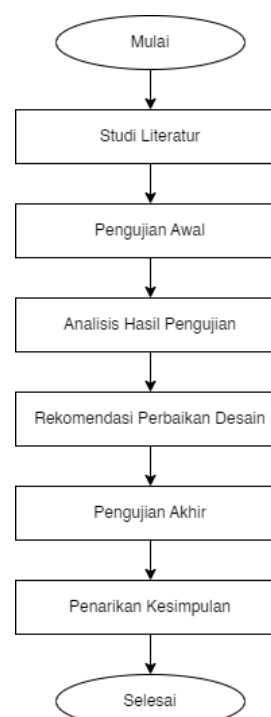
Penelitian sebelumnya yang telah dilakukan oleh Maulana (2022) yang berjudul “Pengembangan Lanjut Aplikasi Supervisi Observasi Pembelajaran Berbasis *Mobile* (LESON)”, aplikasi telah mengalami pembaharuan dari segi antarmuka maupun fitur di dalamnya. Setelah diimplementasikan sebagai artefak aplikasi dan dievaluasi dengan metode Think Aloud kepada 3 pengguna berbeda, masih ditemukan 24 permasalahan pada aplikasi yang perlu diperbaiki dan dikembangkan lebih lanjut, sesuai saran dari penelitian sebelumnya.

Meskipun dalam penelitian terakhir aplikasi telah diuji menggunakan metode Think Aloud dan ditemukan permasalahan, metode Think Aloud mengidentifikasi permasalahan dengan mendengarkan pengguna berbicara tentang penggunaan aplikasi secara luas, sehingga sulit menginterpretasikan data yang dikumpulkan karena pengguna mungkin tidak selalu mengutarakan pemikiran secara akurat atau lengkap, terutama jika pengguna tidak nyaman mengutarakan pemikiran atau kesulitan mengekspresikan diri secara verbal (Fan, Serina dan Truong, 2020). Pengujian penelitian sebelumnya juga masih terbatas pada jumlah pengguna yang digunakan, karena permasalahan yang ditemukan mungkin hanya terjadi pada satu pengguna saja. Oleh karena itu, diperlukan juga pengujian lebih lanjut menggunakan metode yang berbeda untuk menggali lebih dalam permasalahan yang ada pada aplikasi.

Metode *Usability Testing* dipilih karena dapat mengidentifikasi permasalahan secara mendalam, mempelajari perilaku pengguna, dan mengukur seberapa mudah aplikasi digunakan (Moran, 2019). Melalui *Usability Testing* akan dapat diidentifikasi permasalahan dan tantangan yang dihadapi oleh pengguna ketika menggunakan aplikasi. Selain itu, penelitian juga melibatkan 13 pengguna berbeda yang representatif dalam pengumpulan data, menganalisis tantangan yang dihadapi oleh pengguna, dan mengidentifikasi permasalahan yang perlu diatasi.

Berdasarkan dari permasalahan diatas, urgensi penelitian ini adalah perlu adanya pengujian lebih lanjut pada aplikasi menggunakan metode dan jumlah sampel yang berbeda dari penelitian sebelumnya agar dapat mengidentifikasi lebih dalam permasalahan aplikasi dan memperoleh saran perbaikan dari pengguna untuk meningkatkan *usability* aplikasi. Dengan pemahaman yang mendalam tentang pengalaman pengguna yang diperoleh melalui hasil observasi dan wawancara, diharapkan permasalahan pada aplikasi LESON atau setelah adanya penelitian ini disebut sebagai aplikasi EDVISOR, dapat diatasi dengan melakukan perbaikan pada antarmuka pengguna, yang merujuk pada temuan permasalahan dan dapat dijadikan rekomendasi untuk pengembangan aplikasi selanjutnya.

2. METODE PENELITIAN



Gambar 1. Diagram Alur Penelitian

Penelitian ini dilakukan melalui beberapa tahapan sesuai yang dipaparkan pada Gambar 1. Diawali dengan studi literatur yang berhubungan dengan penelitian untuk memahami lebih dalam dasar teori dan mencari referensi literatur lainnya. Dilanjutkan dengan persiapan pengujian yang mencakup menentukan pengguna, merancang skenario tugas, pertanyaan wawancara, dan prosedur pelaksanaan pengujian. Kemudian, dilakukan pengujian bersama dengan pengguna, dan hasil pengujian selanjutnya dianalisis sebagai acuan untuk membuat rekomendasi perbaikan desain. Terakhir, melaksanakan pengujian akhir untuk memvalidasi kembali hasil rekomendasi perbaikan yang telah

dibuat serta penarikan kesimpulan berdasarkan perbandingan hasil pengujian awal dan akhir.

Menurut Nielsen (2000), untuk menemukan permasalahan pengguna sudah cukup dengan 5 pengguna atau 3 sampai 4 pengguna pada setiap kelompok apabila terdapat lebih dari satu kelompok pengguna, sedangkan berdasarkan penelitian yang dilakukan oleh Faulkner (2003), minimal jumlah 10 pengguna sudah dapat menemukan permasalahan *usability* sebesar 82% dan rata-rata permasalahan sebesar 94%.

Pengambilan pengguna pengujian pada penelitian ini menggunakan teknik *probability sampling* (Zickar dan Keith, 2023), karena peneliti secara sengaja memilih pengguna yang spesifik untuk mendapatkan wawasan mendalam dan mewakili populasi yang diteliti. Pengguna dibagi menjadi 2 kelompok, yaitu kelompok observer yang terdiri dari 10 orang mahasiswa dan kelompok supervisor yang terdiri dari 3 orang dosen pada Program Studi Pendidikan Teknologi Informasi Fakultas Ilmu Komputer Universitas Brawijaya Malang. Pengguna terpilih merupakan pengguna yang belum pernah menggunakan aplikasi LESON dan sudah pernah melakukan kegiatan observasi kelas di sekolah, baik menjadi guru model maupun observer atau supervisor.

Pada pelaksanaan pengujian awal perlu adanya skenario tugas. Skenario tugas merupakan perintah yang diberikan kepada pengguna ketika menggunakan aplikasi (Sukarsa dkk, 2022). Ketika pengujian peneliti tidak memberikan bantuan maupun batasan waktu kepada pengguna agar dapat mengetahui permasalahan *usability* dan lama waktu dalam mengerjakan tugas. Contoh skenario tugas pengujian awal dapat dilihat pada Tabel 1.

Tabel 1. Skenario Tugas Pengujian Awal

Kode	Tugas
T1	Melakukan pendaftaran akun dan <i>login</i> menggunakan akun tersebut.
T2	Mengubah identitas profil dengan menambahkan foto profil, nama instansi, dan nomor induk.

Wawancara diadakan untuk mendalami permasalahan yang dihadapi pengguna selama menggunakan aplikasi. Terdapat 7 pertanyaan pada pengujian awal. Bagi pengguna Mahasiswa hanya 4 pertanyaan pertama yang diajukan dan bagi pengguna Dosen semua pertanyaan akan diberikan. Pemberian pertanyaan didasarkan pada peran mahasiswa sebagai observer dan dosen sebagai supervisor, agar dapat memperoleh umpan balik yang lebih komprehensif dari dua sudut pandang peran yang berbeda. Beberapa pertanyaan wawancara dapat dilihat pada Tabel 2.

Tabel 2. Pertanyaan Wawancara Pengujian Awal

Kode	Tugas
Q1	Apa Anda mengalami permasalahan atau kesulitan saat mengerjakan tugas?
Q2	Menurut Anda, bagian mana yang perlu diubah atau diperbaiki pada aplikasi agar lebih mudah digunakan?

Hasil pengujian yang telah didapatkan selanjutnya dianalisis dan dikelompokkan. Pengelompokan masalah dilakukan menggunakan metode triangulasi teknik dan sumber. Triangulasi teknik mengacu pada penggunaan beberapa teknik atau metode yang berbeda untuk mengumpulkan data yang saling melengkapi, sedangkan triangulasi sumber melibatkan penggunaan beberapa sumber data yang berbeda untuk memperoleh pemahaman yang lebih terpercaya, dan akurat (Sugiyono, 2019).

Analisis masalah yang didapatkan kemudian dicarikan ide solusi dan rekomendasi perbaikan berdasarkan pedoman desain dan saran perbaikan dari pengguna. Ide solusi diwujudkan ke dalam bentuk *prototype* yang mendekati bentuk akhir dari aplikasi. Hasil *prototype* dievaluasi kembali pada tahap selanjutnya. Pengguna yang terlibat dalam pengujian akhir sama dengan pengguna yang telah berpartisipasi dalam pengujian awal. Setelah pengujian akhir, dilakukan penarikan kesimpulan untuk mengetahui adanya peningkatan atau penurunan tingkat *usability* setelah diberikan rekomendasi perbaikan.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengujian *Usability*

Pengujian *usability* dilaksanakan dengan mengikuti skenario tugas yang telah disediakan untuk mengevaluasi kemampuan pengguna dalam menyelesaikan setiap tugas. Pengujian diukur menggunakan 3 aspek sesuai standar (International Standardization Organization, 2018), yaitu aspek *effectiveness*, aspek *efficiency*, dan aspek *satisfaction*. Terdapat 7 tugas yang harus diselesaikan. Selama pengujian peneliti mencatat keberhasilan pengguna dalam menyelesaikan tugas, waktu yang diperlukan untuk mengerjakan tugas dan tingkat kepuasan setelah menjalankan tugas.

Pengukuran aspek *effectiveness* didasarkan pada data keberhasilan penyelesaian tugas yang dihitung menggunakan *completion rate*, yaitu persentase pengguna berhasil menyelesaikan tugas dengan baik. Setiap pengguna diberi nilai 1 jika berhasil menyelesaikan tugas dengan baik dan nilai 0 jika tidak berhasil menyelesaikan tugas. Hasil perhitungan *completion rate* pada pengujian awal menunjukkan persentase sebesar 86,81% yang dapat dilihat pada Tabel 3.

Pengukuran aspek *efficiency* didasarkan pada data waktu penyelesaian tugas yang dihitung menggunakan *time-based efficiency*, yaitu seberapa upaya pengguna dalam menyelesaikan setiap tugas. Lama waktu dihitung ketika pengguna mulai mengerjakan tugas sampai selesai atau pengguna menyerah. Hasil perhitungan *time-based efficiency* pada pengujian awal menunjukkan hasil sebesar 0,016 *goals/sec* yang dapat dilihat pada Tabel 3.

Pengukuran aspek *satisfaction* didasarkan pada data tingkat kepuasan pengguna yang dihitung menggunakan *single ease question*, yaitu skor yang diberikan pengguna terkait kemudahan penggunaan aplikasi. Pengguna diberikan pertanyaan tunggal mengenai kemudahan penggunaan aplikasi mulai dari sangat mudah (7), mudah (6), tidak sulit (5), cukup (4), tidak mudah (3), sulit (2), sangat sulit (1). Hasil perhitungan *single ease question* pada pengujian awal menunjukkan skor sebesar 5,49 yang dapat dilihat pada Tabel 3.

Tabel 3. Tingkat *Usability* Aplikasi pada Pengujian Awal

Aspek	Metrik	Hasil
<i>Effectiveness</i>	<i>Completion Rate</i>	86,81%
<i>Efficiency</i>	<i>Time-Based Efficiency</i>	0,016 goals/sec
<i>Satisfaction</i>	<i>Single Ease Question</i>	5,49

Data pengujian *usability* juga dianalisis untuk mengidentifikasi faktor-faktor yang mempengaruhi pengguna dalam mengerjakan setiap tugas. Analisis melibatkan evaluasi mendalam terhadap interaksi antara pengguna dan aplikasi. Dengan menganalisis data pengujian *usability*, peneliti dapat melakukan perbaikan yang spesifik dan terarah untuk meningkatkan pengalaman pengguna secara keseluruhan, memastikan bahwa aplikasi yang dihasilkan tidak hanya fungsional secara teknis, tetapi juga nyaman digunakan oleh pengguna.

3.2 Hasil Observasi dan Wawancara

Observasi diperoleh melalui pencatatan perilaku pengguna selama menggunakan aplikasi. Pencatatan observasi bertujuan untuk mengamati dan memperoleh pemahaman tentang interaksi pengguna dengan aplikasi yang sedang diuji. Hasil observasi dianalisis dan dikelompokkan untuk mengidentifikasi permasalahan yang ditemukan. Setiap observasi memiliki kode masalah yang merujuk pada pengelompokan tugas yang sesuai. Contoh hasil observasi dapat dilihat pada Tabel 4.

Wawancara dilakukan kepada pengguna yang telah menyelesaikan pengujian. Tujuan adanya wawancara untuk menggali lebih dalam permasalahan yang dihadapi oleh pengguna ketika mengerjakan tugas. Melalui analisis wawancara, dapat teridentifikasi permasalahan yang sering muncul dan umpan balik pengguna mengenai permasalahan tersebut. Setiap observasi memiliki kode masalah yang merujuk pada pengelompokan tugas yang sesuai. Hasil wawancara disajikan pada Tabel 5.

Tabel 4. Hasil Observasi Pengujian Awal

Kode	Hasil Observasi	Pengguna	Kode Masalah
T1	Pengguna salah memasukkan kata sandi saat <i>login</i> .	R1	M1
T2	Pengguna tidak mengubah foto profil.	R6, R7, R8	M10

Tabel 5. Hasil Wawancara Pengujian Awal

Kode	Hasil Wawancara	Pengguna	Kode Masalah
T1	Banyak area kosong pada halaman Beranda. Tulisan pada <i>modal</i>	R1, R2, R3	M4
T2	berhasil mengubah profil terlalu besar.	R1	M3

3.3 Pengelompokan Masalah dan Rekomendasi Perbaikan

Masalah yang telah ditemukan dari hasil wawancara, observasi, dan pengujian *usability*, kemudian dikelompokkan berdasarkan kesamaan jenis untuk memudahkan dalam mengidentifikasi sumber permasalahan. Setiap masalah dijabarkan dari mana sumbernya dan pengguna yang mengalaminya. Suatu masalah dapat dikatakan valid jika memiliki lebih dari satu sumber atau minimal tiga pengguna menyatakan hal yang sama. Terdapat 42 permasalahan yang berhasil ditemukan. Berikut beberapa permasalahan yang paling sering muncul dapat dilihat pada Tabel 6.

Berdasarkan pengelompokan masalah pada Tabel 6, kemudian dirumuskan rekomendasi perbaikan yang didapatkan dari pedoman *Google Material Design* dan saran dari pengguna pada saat wawancara. Hasil rekomendasi perbaikan lebih lengkap dijabarkan pada Tabel 7.

Tabel 6. Pengelompokan Masalah

Kode	Permasalahan	Sumber	Pengguna
M1	<i>Text input</i> hurufnya terlalu kecil, warnanya kurang terbaca dan tidak konsisten dengan <i>text input</i> halaman lain.	Wawancara Observasi	R1, R2, R3, R4, R5, R6, R7, R8, R10, R13
M2	Tombol banyak yang tidak konsisten baik dari segi warna, tulisan maupun ukuran.	Wawancara Observasi	R1, R2, R3, R6, R7, R8, R13
M3	<i>Modal</i> pemberitahuan tulisannya terlalu besar dan <i>background</i> terlalu terang.	Wawancara	R1, R7, R8, R9
M4	Pengguna mengeluhkan banyaknya area kosong pada beberapa halaman.	Wawancara	R1, R2, R3, R7, R8, R10

Tabel 7. Rekomendasi Perbaikan

Kode	Rekomendasi Perbaikan	Sumber
F1	Perbaikan ukuran teks pada <i>input</i> , pastikan warna teks kontras dengan latar belakang dan gunakan konsistensi dalam ukuran dan gaya teks di seluruh halaman. Terapkan konsisten dalam gaya dan ukuran tombol di seluruh aplikasi.	M1
F2	Gunakan palet warna yang konsisten dan pastikan teks pada tombol mudah terbaca.	M2
F3	Perbaiki <i>modal</i> pemberitahuan dengan ukuran teks yang lebih sesuai dan latar	M3

Kode	Rekomendasi Perbaikan	Sumber
	belakang yang lebih lembut atau kontras. Pastikan teks dan latar belakang memiliki kontras yang memadai agar pesan mudah dibaca.	
F4	Periksa tata letak dan penempatan elemen di halaman untuk meminimalkan ruang kosong yang tidak perlu.	M4

3.4 Pembuatan Antarmuka Perbaikan Desain

Pembuatan antarmuka berfokus untuk mengimplementasikan hasil rekomendasi perbaikan ke dalam bentuk tampilan antarmuka dengan tetap berpedoman pada prinsip-prinsip desain yang relevan. Pertama, merancang *wireframe*. *Wireframe* diperlukan untuk menggambarkan tata letak dan struktur visual dari antarmuka yang dikembangkan (Shania, Handayani dan Asih, 2023). Proses pembuatan *wireframe* meliputi identifikasi fungsi utama di setiap halaman, pengaturan hierarki konten, serta penentuan posisi dan ukuran elemen antarmuka. Gambar 2 merupakan contoh tampilan *wireframe* pada halaman Beranda antarmuka baru. Kedua, membuat *prototype*. *prototype* dibuat dengan mengacu pada pedoman *Google Material Design* dan saran dari pengguna. Pembuatan *prototype* melibatkan perbaikan dan penyempurnaan komponen visual berdasarkan masalah yang diidentifikasi dengan tujuan untuk meningkatkan pengalaman pengguna dalam menggunakan aplikasi. Terdapat 50 antarmuka baru yang dihasilkan dari pembuatan *prototype*. Contoh Gambar 3 menunjukkan tampilan halaman Beranda sebelum mengalami perbaikan, sementara Gambar 4 menunjukkan perubahan desain setelah perbaikan telah diterapkan. Hasil pembuatan *prototype* juga dievaluasi kembali pada pengujian akhir.

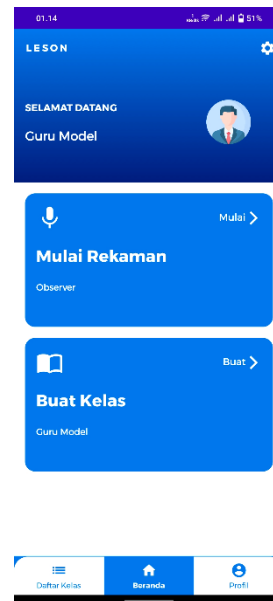
Aplikasi LESON juga telah mengalami perubahan yang signifikan dan kini hadir dengan identitas baru sebagai aplikasi EDVISOR. Perubahan nama didasari oleh permasalahan yang muncul pada identitas aplikasi LESON dan dalam upaya untuk memberikan kesan yang lebih profesional dan relevan dengan tujuan aplikasi. Peneliti memutuskan untuk mengganti nama menjadi EDVISOR.

3.5 Pengujian Akhir

Pengujian akhir dilakukan untuk mengevaluasi peningkatan pengalaman pengguna setelah penerapan rekomendasi perbaikan. Pengguna yang berpartisipasi pada pengujian akhir merupakan pengguna yang sama dengan pengguna pada pengujian awal guna menghindari bias atau perbedaan yang signifikan, serta untuk mengukur sejauh mana peningkatan pengalaman yang dialami oleh pengguna. Beberapa alur pengujian akhir juga telah disesuaikan agar dapat menguji hasil rekomendasi perbaikan dengan maksimal.



Gambar 2. *Wireframe* Halaman Beranda



Gambar 3. Antarmuka Halaman Beranda Sebelum Perbaikan

Skenario pengujian akhir dirancang berdasarkan alur terbaru dari aplikasi EDVISOR. Perubahan skenario telah mempertimbangkan perubahan antarmuka, sehingga pengguna dapat mengakses fitur-fitur terbaru maupun yang telah diperbaiki. Pemberian skenario tugas berfokus untuk menguji interaksi pengguna, mengidentifikasi potensi masalah, serta mengukur sejauh mana pengguna dapat menyelesaikan tugas yang diberikan. Skenario pengujian terbaru terdiri dari 9 tugas yang beberapa diantaranya dapat dilihat pada Tabel 8.



Gambar 4. Antarmuka Halaman Beranda Setelah Perbaikan

Tabel 8. Skenario Tugas Pengujian Akhir

Kode	Tugas
TA8	Melihat detail kelas yang telah dibuat, mengunggah dokumen kelas, dan melihat hasil observasi.
TA9	Melakukan kegiatan refleksi dan mengunggah laporan kelas.

Pertanyaan wawancara pada pengujian akhir telah mengalami beberapa perubahan dari pertanyaan wawancara pada pengujian awal. Perubahan dilakukan berdasarkan evaluasi hasil pengujian awal yang telah dilakukan sebelumnya. Tujuan perubahan pertanyaan wawancara adalah untuk memastikan bahwa pertanyaan wawancara yang diajukan kepada responden baru dapat mengevaluasi tampilan antarmuka baru dengan lebih relevan. Berikut pertanyaan wawancara pengujian akhir dapat dilihat pada Tabel 9.

Tabel 9. Pertanyaan Wawancara Pengujian Akhir

Kode	Tugas
QA3	Bagaimana pengalaman Anda ketika melakukan skenario tugas yang diberikan?
QA4	Bagaimana pendapat Anda tentang perubahan yang telah dilakukan pada antarmuka aplikasi setelah perbaikan?

Hasil pengujian *usability* pada pengujian akhir masih dilakukan dengan cara yang sama pada pengujian awal. Pengujian masih menggunakan aspek *effectiveness*, aspek *efficiency*, dan aspek *satisfaction*. Data pengujian diambil dari keberhasilan penyelesaian tugas, waktu pengerjaan tugas, dan tingkat kepuasan pengguna. Hasil pengujian *usability* pada pengujian akhir dapat dilihat pada Tabel 10.

Tabel 10. Tingkat *Usability* Aplikasi pada Pengujian Akhir

Aspek	Metrik	Hasil
<i>Effectiveness</i>	<i>Completion Rate</i>	100%
<i>Efficiency</i>	<i>Time-Based Efficiency</i>	0,032 goals/sec
<i>Satisfaction</i>	<i>Single Ease Question</i>	6,7

Pengujian akhir dilakukan menggunakan *prototype* yang mendekati hasil akhir aplikasi, sedangkan pengujian awal menggunakan artefak aplikasi yang sudah jadi, sehingga terdapat beberapa keterbatasan yang perlu diperhatikan. Perbedaan karakteristik antara aplikasi nyata dan *prototype* dapat mempengaruhi hasil pengukuran. Aplikasi nyata sudah melewati tahap pengembangan yang lebih lanjut, sehingga mungkin memiliki optimalisasi kinerja dan fitur yang lebih matang dibandingkan dengan *prototype* yang hanya berupa representasi visual.

3.6 Penggunaan Gambar

Uji statistik pada penelitian ini digunakan untuk penggambaran atau membuat kesimpulan lebih lanjut berdasarkan data yang sudah ada. Diawali dengan uji normalitas data yang dilakukan pada jumlah sampel 13 orang menggunakan Shapiro-Wilk. Uji normalitas data memanfaatkan program IBM SPSS V.25.0 dan didapatkan hasil pada Tabel 11.

Tabel 11. Hasil Uji Normalitas

Aspek	df	Sig.	Status
<i>Effectiveness</i> Pengujian Awal	13	0,006	Tidak Normal
<i>Effectiveness</i> Pengujian Akhir	13	0	Tidak Normal
<i>Efficiency</i> Pengujian Awal	13	0,006	Tidak Normal
<i>Efficiency</i> Pengujian Akhir	13	0,594	Normal
<i>Satisfaction</i> Pengujian Awal	13	0,267	Normal
<i>Satisfaction</i> Pengujian Akhir	13	0,0192	Normal

Berdasarkan Tabel 11, hasil uji normalitas menyatakan bahwa ada 3 kelompok data yang memiliki nilai Sig. lebih dari 0,05 dan 3 kelompok data yang memiliki nilai Sig. kurang dari 0,05. Dapat disimpulkan hasil uji normalitas data mengindikasikan ada 3 kelompok data terdistribusi tidak normal dengan nilai Sig. kurang dari 0,05 dan 3 kelompok data terdistribusi normal dengan nilai Sig. lebih dari 0,05 (Kim dan Park, 2019).

Setelah dilakukan uji normalitas, selanjutnya dilakukan perhitungan statistik non-parametrik dengan menggunakan uji Wilcoxon Signed Rank. Uji Wilcoxon adalah uji *non-parametric* yang digunakan untuk membandingkan perbedaan antara dua sampel terkait atau data berpasangan (Setyawan, 2017). Penggunaan uji Wilcoxon bertujuan untuk mengetahui perbedaan kelompok data pada pengujian awal dan pengujian akhir. Hipotesis dapat dikatakan diterima apabila nilai Asymp.Sig. kurang dari 0,05, sedangkan jika nilai Asymp.Sig. lebih dari 0,05, maka hipotesis dikatakan ditolak. Berikut hasil uji Wilcoxon dapat dilihat pada Tabel 12.

Tabel 12. Hasil Uji Wilcoxon

	<i>Effectiveness</i>	<i>Efficiency</i>	<i>Satisfaction</i>
Z	-2,972	-3,184	-3,182
Asymp.Sig. (2-tailed)	0,003	0,001	0,001
Status	Diterima	Diterima	Diterima

Berdasarkan Tabel 12, hasil uji Wilcoxon Signed Rank menyatakan bahwa pada aspek *Effectiveness* nilai Asymp.Sig. menunjukkan sebesar 0,003, aspek *Efficiency* nilai Asymp.Sig. menunjukkan sebesar 0,001, dan aspek *Satisfaction* nilai Asymp.Sig. menunjukkan sebesar 0,001, hal tersebut mengindikasikan ketiga aspek memiliki nilai Asymp.Sig. kurang dari 0,05 (Murakami dan Lee, 2023), yang berarti terdapat perbedaan *usability* pada aplikasi *mobile* supervisi observasi kelas dari hasil pengujian awal dan pengujian akhir.

3.7 Perbandingan Pengujian Awal dan Akhir

Perbandingan ditujukan untuk mengetahui adanya peningkatan atau penurunan tingkat *usability* terhadap rekomendasi antarmuka baru. Hasil pengujian dibandingkan meliputi 3 aspek, yaitu aspek *effectiveness*, aspek *efficiency* dan aspek *satisfaction*. Perbandingan dikhususkan hanya pada tugas 1, 2, 3, 4, 5, 6 dan 7 karena pada pengujian awal tidak ada tugas 8 dan 9. Hasil perbandingan pengujian awal dan pengujian akhir dapat dilihat pada Tabel 13.

Tabel 13. Perbandingan Pengujian Awal dan Pengujian Akhir

Aspek	Pengujian Awal	Pengujian Akhir
<i>Effectiveness</i>	86,81%	100%
<i>Efficiency</i>	0,016 <i>goals/sec</i>	0,032 <i>goals/sec</i>
<i>Satisfaction</i>	5,49	6,7

Berdasarkan Tabel 20, setelah adanya perbaikan antarmuka dan dilakukan pengujian pada antarmuka baru. Hasil pengujian akhir menunjukkan peningkatan signifikan pada aspek *effectiveness* dari 86,81% menjadi 100%, aspek *efficiency* meningkat dari 0,016 *goals/sec* menjadi 0,032 *goals/sec*, dan aspek *satisfaction* meningkat dari 5,49 menjadi 6,7.

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan pada aplikasi LESON atau yang sekarang disebut sebagai aplikasi EDVISOR dapat disimpulkan hasil pengujian awal aplikasi menggunakan metode *Usability Testing* ditemukan adanya 42 permasalahan yang bersumber dari observasi dan wawancara oleh 13 pengguna. Temuan permasalahan juga telah dilakukan 42 rekomendasi perbaikan antarmuka yang bersumber dari pedoman *Google Material Design* dan saran pengguna, sehingga menghasilkan adanya *prototype* yang mendekati hasil akhir aplikasi bernama EDVISOR.

Setelah perbaikan dan pengujian akhir pada aplikasi, ditemukan adanya peningkatan *usability* yang diukur dengan membandingkan hasil pengujian

usability awal dan akhir. Pengujian *usability* aspek *effectiveness* yang diukur menggunakan *completion rate* menunjukkan peningkatan sebesar 13,19% dengan hasil pengujian awal sebesar 86,81% dan pengujian akhir sebesar 100%. Pengujian *usability* aspek *efficiency* yang diukur menggunakan *time-based efficiency* menunjukkan peningkatan sebesar 0,016 *goals/sec* dengan hasil pengujian awal sebesar 0,016 *goals/sec* dan pengujian akhir sebesar 0,032 *goals/sec*. Pengujian *usability* aspek *satisfaction* yang diukur menggunakan *single ease question* menunjukkan peningkatan skor sebesar 1,21 dengan hasil pengujian awal sebesar 5,49 dan pengujian akhir sebesar 6,7.

Saran yang dapat peneliti berikan untuk pengembangan penelitian selanjutnya yaitu dapat dilakukan pengujian langsung ketika kegiatan supervisi observasi kelas di sekolah menggunakan metode yang berbeda, sehingga dapat memberikan pemahaman yang lebih tentang pengalaman pengguna dan mendapatkan data yang lebih bervariasi untuk perbaikan aplikasi. Selain itu, penelitian berikutnya juga dapat melakukan implementasi hasil antarmuka baru ke dalam bentuk artefak aplikasi jadi yang sudah dapat digunakan.

DAFTAR PUSTAKA

- AKHMADJONOVICH, U.D., 2021. Integration Of Science dan Education As An Important Factor In Improving The Quality Of Education. The American Journal of Management dan Economics Innovations, 03(10), pp.21–27.
- CUNNINGHAM, K., GORMAN, M. dan MAHER, J., 2020. The value of using classroom observations as part of a multi-methodological approach to evaluate student engagement in vocational agricultural education. Educational Research dan Evaluation, 26(1–2), pp.4–29.
- EFENDI, F.S., FANANI, L. dan SUPIANTO, A.A., 2020. Rancang Bangun Aplikasi Pendukung untuk Observasi Kelas berbasis Mobile. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 4(6), pp.1828–1840.
- FAN, M., SERINA, S. dan TRUONG, K., 2020. Practices dan Challenges of Using Think-Aloud Protocols in Industry: An International Survey. Journal of Usability Studies, 5(2), pp.85–102.
- FAULKNER, L., 2003. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. Behavior Research Methods, Instruments, & Computers, 35(3), pp.379–383.
- INTERNATIONAL STANDARDIZATION ORGANIZATION, 2018. ISO 9241-11:2018, Ergonomics of human-system interaction — Part 11: Usability: Definitions dan concepts. [online] Tersedia di:

- <<https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en>> [Diakses 9 September 2022].
- KIM, T.K. dan PARK, J.H., 2019. More about the basic assumptions of t-test: normality dan sample size. *Korean Journal of Anesthesiology*, 72(4), pp.331–335.
- MAULANA, F.D., DEWI, R.K. dan ROKHMAWATI, R.I., 2022. Pengembangan Lanjut Aplikasi Supervisi Pembelajaran berbasis Mobile (LESON). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(1), pp.384–393.
- MORAN, K., 2019. Usability Testing 101. [online] Tersedia di: <<https://www.nngroup.com/articles/usability-testing-101/>> [Diakses 6 April 2023].
- MURAKAMI, H. dan LEE, S., 2023. On unbiasedness dan biasedness of the Wilcoxon dan some nonparametric tests. *WIREs Computational Statistics*, 15(3).
- NIELSEN, J., 2000. Why You Only Need to Test with 5 Users. [online] Tersedia di: <<https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>> [Diakses 9 September 2022].
- SETIA, R. dan NASRUDIN, D., 2020. Teacher Supervision As An Improvement In The Quality Of Education. *International Journal of Education dan Social Science Research*, 03(03), pp.11–22.
- SHANIA, M., HANDAYANI, P.W. dan ASIH, S., 2023. Designing High-Fidelity Mobile Health for Depression in Indonesian Adolescents Using Design Science Research: Mixed Method Approaches. *JMIR Formative Research*, 7, p.e48913.
- SUGIYONO, 2019. *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: Alfabeta.
- SUKARSA, I.M., BUANA, I.P.W., ARYA UTAMA, I.P.J. dan WISSWANI, N.W., 2022. Evaluasi Usability dan Perbaikan Antarmuka untuk Meningkatkan User Experience Menggunakan Metode Usability Testing (Studi Kasus: Aplikasi Warga Bali). *Jurnal Teknologi Informasi dan Ilmu Komputer*, 9(5), p.1003–1009.
- WIBAWA, A.S., ROKHMAWATI, R.I. dan AZ-ZAHRA, H.M., 2021. Perancangan Antarmuka Pengguna Aplikasi Leson menggunakan Metode Human-Centered Design (HCD). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 5(12), pp.5497–5504.
- ZICKAR, M.J. dan KEITH, M.G., 2023. Innovations in Sampling: Improving the Appropriateness dan Quality of Samples in Organizational Research. *Annual Review of Organizational Psychology dan Organizational Behavior*, 10(1), pp.315–337.