



# Analisis Pengalaman Pengguna Light dan Dark Mode Pada Facebook dan Tokopedia Menggunakan Within-Subject Design

Muhammad Farhan, Chanifah Indah Ratnasari\*

Fakultas Teknologi Industri, Program Studi Informatika, Universitas Islam Indonesia, Sleman  
Jalan Kaliurang No.Km. 14,5, Krawitan, Umbulmartani, Ngemplak, Kabupaten Sleman, Yogyakarta, Indonesia

Email: <sup>1</sup>muhammad.farhan017@students.uii.ac.id, <sup>2\*</sup>chanifah.indah@uii.ac.id

Email Penulis Korespondensi: chanifah.indah@uii.ac.id

Submitted: 28/03/2026; Accepted: 17/04/2026; Published: 26/04/2026

**Abstrak**—Penggunaan light mode dan dark mode pada aplikasi mobile semakin umum, namun pengaruhnya terhadap pengalaman pengguna pada konteks aplikasi yang berbeda masih perlu dibuktikan secara empiris. Penelitian ini bertujuan menganalisis perbedaan pengalaman pengguna antara light mode dan dark mode pada Facebook dan Tokopedia. Penelitian menggunakan pendekatan kuantitatif dengan within-subject design (repeated measures design) yang melibatkan 25 responden. Pengukuran dilakukan menggunakan Time on Task, Error Rate, User Experience Questionnaire (UEQ), serta preferensi pengguna. Hasil menunjukkan bahwa pada Facebook tidak terdapat perbedaan signifikan pada Time on Task, sedangkan pada Tokopedia ditemukan perbedaan signifikan pada T1 ( $p = 0,024$ ) dan T3 ( $p = 0,047$ ). Pada Error Rate, perbedaan signifikan ditemukan pada Facebook T1 ( $p = 0,046$ ) serta Tokopedia T2 ( $p = 0,020$ ) dan T3 ( $p = 0,019$ ). Hasil UEQ menunjukkan kedua mode berada pada kategori positif tanpa perbedaan signifikan secara statistik. Temuan ini menunjukkan bahwa pengaruh mode tampilan lebih terlihat pada metrik performa tugas tertentu dibandingkan pada persepsi pengalaman pengguna secara keseluruhan, sehingga efeknya bersifat kontekstual. Penelitian ini berkontribusi dengan memberikan evaluasi komprehensif yang menggabungkan metrik objektif dan subjektif dalam membandingkan light mode dan dark mode pada dua konteks aplikasi yang berbeda.

**Kata Kunci:** Pengalaman Pengguna; Light Mode; Dark Mode; Within-Subject Design; User Experience Questionnaire; Usability

**Abstract**—The use of light mode and dark mode in mobile applications is becoming increasingly common; however, their effects on user experience across different application contexts still need to be empirically examined. This study aims to analyze differences in user experience between light mode and dark mode in Facebook and Tokopedia. The study employed a quantitative approach using a within-subject design (repeated measures design) involving 25 respondents. Measurements were conducted using Time on Task, Error Rate, the User Experience Questionnaire (UEQ), and user preference. The results showed that Facebook had no significant difference in Time on Task, whereas Tokopedia showed significant differences in T1 ( $p = 0.024$ ) and T3 ( $p = 0.047$ ). For Error Rate, significant differences were found in Facebook T1 ( $p = 0.046$ ) as well as Tokopedia T2 ( $p = 0.020$ ) and T3 ( $p = 0.019$ ). The UEQ results indicated that both modes were in the positive category without statistically significant differences. These findings suggest that the influence of display mode is more evident in specific task performance metrics than in overall user experience perception, indicating that its effects are contextual. This study contributes by providing a comprehensive evaluation that combines objective and subjective metrics to compare light mode and dark mode across two different application contexts.

**Keywords:** User Experience; Light Mode; Dark Mode; Within-Subject Design; User Experience Questionnaire; Usability

## 1. PENDAHULUAN

Perkembangan teknologi smartphone telah mengubah cara masyarakat berkomunikasi, memperoleh informasi, menikmati hiburan, hingga melakukan transaksi digital melalui aplikasi mobile [1]. Aplikasi berbasis smartphone kini semakin dekat dengan aktivitas sehari-hari pengguna. Dalam konteks ini, keberhasilan aplikasi tidak hanya ditentukan oleh kelengkapan fitur, tetapi juga oleh kualitas pengalaman yang dirasakan pengguna saat berinteraksi dengan sistem. Pada aplikasi mobile, user experience menjadi aspek penting karena berkaitan dengan kenyamanan, keterlibatan, kepuasan, dan kecenderungan pengguna untuk terus menggunakan aplikasi dalam jangka panjang [2], [3]. Oleh karena itu, desain antarmuka tidak dapat dipandang hanya sebagai elemen visual, tetapi sebagai bagian yang memengaruhi kualitas interaksi pengguna.

Salah satu keputusan desain antarmuka yang semakin banyak diterapkan pada aplikasi mobile adalah penggunaan dua mode tampilan, yaitu light mode dan dark mode [4]. Light mode menggunakan latar terang dengan teks gelap, sedangkan dark mode menggunakan latar gelap dengan teks terang. Kehadiran dua mode ini memberi fleksibilitas kepada pengguna untuk menyesuaikan tampilan dengan kondisi penggunaan dan preferensi visual masing-masing. Namun, pengaruh kedua mode tersebut terhadap pengalaman pengguna belum menunjukkan hasil yang konsisten. Beberapa studi menunjukkan bahwa perbedaan latar tampilan dapat memengaruhi performa kognitif saat pengguna berinteraksi dengan antarmuka digital [5]. Studi lain menunjukkan bahwa preferensi terhadap dark mode dan light mode dipengaruhi oleh konteks penggunaan, kebiasaan pengguna, dan kebutuhan kenyamanan visual [6], [7]. Di sisi lain, kombinasi warna teks dan latar juga dapat memengaruhi efisiensi pengenalan teks pada perangkat digital, sehingga aspek keterbacaan tetap menjadi pertimbangan penting dalam pemilihan mode tampilan [8]. Temuan tersebut menunjukkan bahwa perbandingan light mode dan dark mode tidak dapat dipahami hanya sebagai pilihan estetika, tetapi juga terkait dengan kualitas interaksi pengguna.



Dalam penelitian pengalaman pengguna, evaluasi terhadap mode tampilan perlu dilakukan secara terstruktur agar tidak hanya menggambarkan preferensi visual, tetapi juga menunjukkan dampaknya terhadap performa penggunaan. Penelitian ini menggunakan within-subject design (repeated measures design), yaitu metode eksperimen di mana responden yang sama menguji lebih dari satu kondisi antarmuka sehingga memungkinkan perbandingan langsung antar kondisi [9]. Selain itu, pengujian berbasis tugas pada aplikasi mobile juga banyak digunakan dalam studi usability untuk menilai bagaimana pengguna berinteraksi dengan aplikasi pada konteks penggunaan yang nyata [10]. Karena itu, evaluasi terhadap light mode dan dark mode sebaiknya menggabungkan metrik objektif dan subjektif agar hasilnya lebih komprehensif [11], [12].

Penelitian sebelumnya telah mengkaji hubungan antara mode tampilan dan pengalaman pengguna dari berbagai perspektif. Penelitian [5] menunjukkan bahwa perbedaan mode latar tampilan dapat memengaruhi performa kognitif pengguna saat berinteraksi dengan antarmuka digital. Studi literatur pada [6] menjelaskan bahwa preferensi pengguna terhadap light mode dan dark mode bervariasi serta dipengaruhi oleh konteks penggunaan dan kenyamanan visual. Dari sisi implementasi, penelitian [4] menunjukkan bahwa dark mode telah banyak diadopsi pada aplikasi Android dunia nyata. Sementara itu, penelitian [13] menunjukkan bahwa User Experience Questionnaire (UEQ) relevan digunakan untuk mengevaluasi pengalaman pengguna pada aplikasi mobile, sedangkan [9] dan [10] menunjukkan bahwa desain eksperimen berulang dan usability testing berbasis tugas dapat digunakan untuk membandingkan kondisi antarmuka dan menilai efektivitas interaksi pengguna.

Meskipun demikian, hasil penelitian yang ada masih bervariasi tergantung pada konteks aplikasi dan metode evaluasi yang digunakan [5], [6], [7]. Sejumlah studi telah membandingkan light mode dan dark mode, tetapi umumnya dilakukan pada satu jenis aplikasi atau hanya menggunakan satu pendekatan pengukuran [5], [6], [7]. Selain itu, studi yang mengombinasikan metrik objektif seperti Time on Task dan Error Rate dengan metrik subjektif seperti User Experience Questionnaire (UEQ) pada dua aplikasi dengan karakteristik interaksi berbeda masih relatif terbatas, karena kebanyakan penelitian berfokus pada evaluasi satu aplikasi [13], [14], [15]. Kondisi ini menunjukkan bahwa pengaruh mode tampilan belum sepenuhnya dipahami secara komprehensif, terutama apakah bersifat konsisten pada berbagai jenis aplikasi atau justru bergantung pada konteks interaksi masing-masing aplikasi [5], [6], [13], [14], [15]. Kebaruan penelitian ini terletak pada kombinasi evaluasi dua jenis aplikasi dengan karakteristik interaksi berbeda, yaitu media sosial (Facebook) dan e-commerce (Tokopedia), serta penggunaan metrik objektif dan subjektif secara simultan dalam satu within-subject design.

Berdasarkan celah tersebut, penelitian ini mengusulkan evaluasi komparatif antara light mode dan dark mode pada dua aplikasi mobile dengan karakteristik berbeda, yaitu Facebook sebagai aplikasi media sosial dan Tokopedia sebagai aplikasi e-commerce. Kedua aplikasi tersebut dipilih karena mewakili dua pola interaksi yang tidak sama. Facebook lebih banyak digunakan untuk konsumsi konten dan interaksi sosial, sedangkan Tokopedia lebih menuntut ketelitian visual saat pengguna mencari produk, membaca informasi, dan melakukan alur transaksi. Penelitian dilakukan menggunakan desain within-subject design (repeated measures design), sehingga setiap responden menguji kedua mode tampilan pada kedua aplikasi. Untuk memperoleh gambaran yang lebih utuh, evaluasi tidak hanya menggunakan metrik objektif berupa time on task dan error rate, tetapi juga metrik subjektif melalui User Experience Questionnaire (UEQ) dan preferensi pengguna.

Tujuan penelitian ini adalah menganalisis perbedaan pengalaman pengguna antara light mode dan dark mode pada Facebook dan Tokopedia, serta mengidentifikasi apakah pengaruh mode tampilan tersebut bersifat seragam atau kontekstual. Penelitian ini memberikan kontribusi pada dua sisi. Dari sisi akademik, penelitian ini memperkaya pembahasan mengenai hubungan antara mode tampilan dan pengalaman pengguna melalui evaluasi yang menggabungkan metrik objektif dan subjektif. Dari sisi praktis, hasil penelitian dapat menjadi masukan bagi pengembang aplikasi dalam menentukan strategi desain tampilan yang lebih sesuai dengan karakteristik aplikasi dan kebutuhan pengguna. Dengan demikian, penelitian ini tidak hanya membandingkan dua mode tampilan, tetapi juga berupaya menjelaskan bagaimana mode tampilan memengaruhi kualitas interaksi pengguna pada konteks aplikasi yang berbeda.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif untuk membandingkan pengalaman pengguna antara light mode dan dark mode pada dua aplikasi mobile, yaitu Facebook dan Tokopedia. Pengujian dilakukan menggunakan desain within-subject (repeated measures design), di mana dua variasi tampilan dibandingkan pada responden yang sama. Data dikumpulkan melalui penyelesaian skenario tugas, pengukuran time on task dan error rate, pengisian User Experience Questionnaire (UEQ), serta kuesioner preferensi pengguna. Pendekatan ini dipilih agar evaluasi dapat menangkap persepsi pengguna sekaligus performa penggunaan secara langsung melalui kombinasi metrik objektif dan subjektif [10], [11], [12].

Secara umum, tahapan penelitian terdiri atas studi literatur, penentuan objek penelitian, penentuan responden, penyusunan skenario pengujian, pelaksanaan pengujian, pengumpulan data, dan analisis data. Alur tahapan penelitian tersebut dapat dilihat pada Gambar 1.



**Gambar 1.** Tahapan Penelitian

## 2.2 Desain Eksperimen dan Partisipan

Penelitian ini menggunakan desain within-subject, sehingga setiap responden menguji kedua mode tampilan, yaitu light mode dan dark mode, pada aplikasi yang diteliti. Desain ini sesuai untuk studi komparatif karena data diperoleh dari partisipan yang sama pada dua kondisi berbeda, sehingga perbedaan antarindividu dapat ditekan dan perbandingan antarkondisi menjadi lebih langsung [9]. Pendekatan ini selaras dengan desain eksperimen komparatif yang memungkinkan perbandingan dua variasi antarmuka secara langsung [9]. Jumlah responden dalam penelitian ini sebanyak 25 orang. Penggunaan desain within-subject memungkinkan peningkatan sensitivitas analisis karena setiap responden menghasilkan data berpasangan. Jumlah tersebut dinilai memadai untuk studi usability komparatif, meskipun tetap menjadi keterbatasan dalam generalisasi hasil penelitian [16], [17].

Dalam pengujian, responden menggunakan perangkat smartphone masing-masing, baik berbasis Android maupun iOS. Penggunaan perangkat pribadi dipertahankan karena penelitian ini berupaya menangkap pengalaman penggunaan yang lebih dekat dengan konteks nyata. Pelaksanaan pengujian dilakukan pada lingkungan yang bervariasi, baik indoor maupun outdoor, dengan kondisi pencahayaan yang mengikuti situasi penggunaan saat pengambilan data [10]. Pendekatan ini dipilih untuk merepresentasikan kondisi penggunaan nyata (real-world usage). Namun demikian, variasi lingkungan dan pencahayaan menjadi salah satu keterbatasan penelitian yang berpotensi memengaruhi hasil pengujian, terutama karena penelitian ini membandingkan dua mode tampilan yang sensitif terhadap kondisi visual.

Pemilihan sampel dilakukan secara purposive sampling. Kriteria responden dalam penelitian ini adalah berusia minimal 18 tahun, aktif menggunakan aplikasi mobile, pernah menggunakan aplikasi Facebook dan Tokopedia, telah mengenal serta menggunakan light mode dan dark mode, menggunakan perangkat smartphone berbasis Android atau iOS, serta tidak memiliki gangguan penglihatan yang signifikan. Pada awal sesi, responden diminta mengisi data demografi melalui Google Form, yang mencakup usia, jenis kelamin, generasi berdasar tahun kelahiran, pekerjaan, perangkat yang digunakan, tipe layar smartphone, kebiasaan penggunaan mode tampilan, serta mode tampilan yang sedang digunakan pada aplikasi. Setelah itu, responden diberikan kebebasan untuk menentukan aplikasi yang akan diuji terlebih dahulu, yaitu Facebook atau Tokopedia. Pada setiap aplikasi, pengujian dilakukan dengan urutan tetap, yaitu light mode terlebih dahulu, kemudian dark mode. Penelitian ini tidak menerapkan teknik counterbalancing pada urutan pengujian, sehingga terdapat potensi bias berupa order effect dan fatigue effect. Oleh karena itu, hasil penelitian ini perlu diinterpretasikan secara hati-hati sebagai salah satu keterbatasan metodologis penelitian. Setelah pengujian pada light mode selesai, responden diberikan jeda sekitar  $\pm 2$  menit sebelum melanjutkan ke dark mode untuk mengurangi kelelahan dan efek adaptasi visual. Selain itu, setelah seluruh rangkaian pengujian pada satu aplikasi selesai, responden juga diberikan jeda sekitar  $\pm 3-5$  menit sebelum melanjutkan ke aplikasi berikutnya. Selama pengujian, moderator hanya memberikan penjelasan pada awal sesi dan tidak terlibat dalam pengerjaan tugas. Namun, apabila responden benar-benar mengalami kesulitan dan tidak dapat melanjutkan tugas, moderator dapat memberikan bantuan seperlunya, dan bantuan tersebut dicatat serta dihitung sebagai error.

## 2.3 Skenario Pengujian

Skenario pengujian disusun untuk merepresentasikan aktivitas umum pengguna pada masing-masing aplikasi. Setiap responden menyelesaikan tiga tugas pada Facebook dan tiga tugas pada Tokopedia. Skenario ini digunakan untuk mengamati performa pengguna secara langsung pada dua mode tampilan yang dibandingkan. Ringkasan skenario pengujian pada Facebook disajikan pada Tabel 1, sedangkan skenario pengujian pada Tokopedia disajikan pada Tabel 2.

**Tabel 1.** Skenario Pengujian Facebook

No	Skenario Tugas	Deskripsi Tugas
1	Menjelajahi halaman beranda (news feed)	Responden membuka aplikasi Facebook dan mengamati tampilan halaman beranda dengan melakukan scroll ke bawah hingga menemukan minimal tiga jenis konten berbeda, yaitu unggahan dari teman, unggahan halaman, dan iklan (sponsored post). Selama proses ini, responden diminta membaca isi unggahan tanpa melakukan interaksi tambahan.

No	Skenario Tugas	Deskripsi Tugas
2	Mencari akun atau halaman dan melakukan interaksi	Responden menggunakan fitur pencarian untuk menemukan satu akun atau halaman tertentu, kemudian membuka salah satu unggahan, memberikan tanda suka (like), dan menuliskan komentar singkat pada unggahan tersebut.
3	Membuat dan memasang postingan	Responden membuat dan memposting konten dengan memilih foto atau video dari galeri, menambahkan keterangan sesuai keinginan, lalu mempublikasikan postingan tersebut.

**Tabel 2.** Skenario Pengujian Tokopedia

No	Skenario Tugas	Deskripsi Tugas
1	Menjelajahi halaman beranda Tokopedia	Responden membuka aplikasi Tokopedia dan mengamati tampilan halaman beranda. Responden kemudian melakukan scroll ke bawah hingga menemukan minimal tiga jenis konten berbeda, seperti produk rekomendasi, promo atau banner diskon, dan kategori produk. Selama proses ini, responden diminta memperhatikan informasi yang ditampilkan tanpa melakukan interaksi tambahan.
2	Mencari produk dan melihat detail produk	Responden menggunakan fitur pencarian untuk menemukan satu produk tertentu. Setelah produk ditemukan, responden membuka halaman detail produk dan membaca informasi utama seperti nama produk, harga, rating, dan deskripsi singkat, tanpa melakukan pembelian.
3	Menambahkan produk ke keranjang dan membuka ringkasan checkout	Responden menambahkan produk tersebut ke dalam keranjang belanja, kemudian membuka halaman keranjang dan melanjutkan hingga ke ringkasan checkout tanpa melakukan pembayaran. Responden diminta memastikan produk berhasil masuk ke keranjang dan memahami alur pembelian yang ditampilkan.

Berdasarkan Tabel 1, skenario pada Facebook dirancang untuk mewakili aktivitas konsumsi konten, pencarian, dan interaksi sosial yang umum dilakukan pengguna. Sementara itu, sebagaimana ditunjukkan pada Tabel 2, skenario pada Tokopedia dirancang untuk mewakili aktivitas pencarian produk, pembacaan informasi, dan simulasi alur pembelian yang menuntut perhatian visual lebih tinggi.

## 2.4 Instrumen dan Parameter Pengukuran

Penelitian ini menggunakan empat jenis pengukuran, yaitu time on task, error rate, User Experience Questionnaire (UEQ), dan kuesioner preferensi pengguna. Kombinasi keempat instrumen tersebut digunakan agar evaluasi tidak hanya menggambarkan persepsi pengguna, tetapi juga menangkap performa penggunaan secara langsung.

### 2.4.1 Time on Task (ToT)

Time on Task digunakan untuk mengukur waktu yang dibutuhkan responden dalam menyelesaikan setiap tugas pengujian. Dalam pengujian mobile usability, waktu penyelesaian tugas merupakan indikator efisiensi yang umum digunakan untuk melihat seberapa cepat pengguna dapat menyelesaikan aktivitas yang diberikan [10]. Parameter ini digunakan untuk melihat efisiensi interaksi pengguna pada masing-masing mode tampilan. Semakin singkat waktu yang dibutuhkan untuk menyelesaikan tugas, semakin efisien interaksi yang terjadi. Perhitungan time on task dilakukan sejak responden mulai mengerjakan tugas hingga tugas dinyatakan selesai. Nilai time on task dihitung menggunakan Persamaan (1).

$$ToT = t_{selesai} - t_{mulai} \tag{1}$$

Nilai rata-rata time on task untuk setiap tugas kemudian dihitung menggunakan Persamaan (2).

$$\overline{ToT} = \frac{\sum_{i=1}^n ToT_i}{n} \tag{2}$$

Variabel yang digunakan dalam penelitian ini meliputi  $t_{mulai}$  yang menunjukkan waktu mulai pelaksanaan tugas, serta  $t_{selesai}$  yang merepresentasikan waktu ketika tugas tersebut diselesaikan. Selanjutnya,  $ToT_i$  mengindikasikan total waktu yang dibutuhkan untuk menyelesaikan tugas ke- $i$ . Adapun  $n$  menyatakan jumlah responden yang terlibat dalam penelitian ini, yang menjadi dasar dalam proses analisis data secara keseluruhan.

### 2.4.2 Error Rate (ER)

Error rate digunakan untuk mengukur tingkat kesalahan responden selama menyelesaikan tugas. Dalam penelitian ini, kesalahan didefinisikan sebagai salah klik (misclick), kesalahan navigasi, kegagalan menemukan fitur, tindakan yang tidak relevan dengan tugas, serta bantuan moderator saat responden tidak dapat melanjutkan tugas. Bantuan moderator tetap dicatat dan dihitung sebagai error agar hasil pengukuran tetap merepresentasikan hambatan yang dialami pengguna selama pengujian. Penggunaan metrik waktu dan kesalahan secara bersamaan



juga lazim dalam evaluasi mobile usability karena keduanya merepresentasikan efisiensi dan efektivitas interaksi secara lebih seimbang [10]. Nilai error rate dihitung menggunakan Persamaan (3).

$$ER = \frac{\sum error}{\sum percobaan} \quad (3)$$

Dalam konteks pengukuran ini,  $\sum error$  merepresentasikan total kesalahan yang dilakukan oleh responden selama pelaksanaan tugas, sedangkan  $\sum percobaan$  menunjukkan jumlah keseluruhan tugas atau percobaan yang dikerjakan. Kedua komponen tersebut digunakan untuk memberikan gambaran tingkat akurasi dan kinerja responden dalam menyelesaikan tugas yang diberikan.

#### 2.4.3 User Experience Questionnaire (UEQ)

UEQ digunakan untuk mengukur persepsi pengalaman pengguna terhadap antarmuka aplikasi. Instrumen ini terdiri atas 26 pasangan kata bipolar yang merepresentasikan enam skala, yaitu daya tarik, kejelasan, efisiensi, ketepatan, stimulasi, dan kebaruan. UEQ merupakan instrumen yang banyak digunakan untuk evaluasi pengalaman pengguna pada produk interaktif dan dapat diterapkan pada berbagai skenario evaluasi [18]. Hasil UEQ juga dapat dibandingkan dengan data benchmark untuk melihat posisi kualitas pengalaman pengguna secara relatif terhadap produk lain. Benchmark UEQ mengelompokkan hasil ke dalam lima kategori, yaitu excellent, good, above average, below average, dan bad [19]. Pada penelitian ini, UEQ diberikan untuk dua kondisi tampilan pada setiap aplikasi, yaitu light mode dan dark mode. Sesuai alur pengujian, responden terlebih dahulu mengisi UEQ untuk light mode, kemudian mengisi UEQ untuk dark mode pada aplikasi yang sama.

Data UEQ dianalisis melalui beberapa tahap, yaitu pemeriksaan inconsistency, perhitungan nilai rata-rata tiap skala, serta interpretasi berdasarkan benchmark UEQ. Meskipun terdapat beberapa jawaban yang menunjukkan nilai inconsistency di atas batas normal, seluruh data responden tetap digunakan dalam analisis sehingga jumlah data final pada setiap kondisi tetap sebanyak 25 responden. Keputusan ini diambil untuk menjaga konsistensi jumlah sampel pada desain within-subject. Temuan terkait inconsistency disampaikan sebagai bagian dari keterbatasan penelitian pada pembahasan. Tingkat inkonsistensi tersebut menunjukkan adanya variasi dalam konsistensi jawaban responden. Oleh karena itu, interpretasi hasil UEQ dilakukan secara hati-hati dan menjadi bagian dari keterbatasan penelitian.

#### 2.4.4 Preferensi pengguna

Selain UEQ, penelitian ini juga menggunakan kuesioner preferensi pengguna untuk menangkap kecenderungan pilihan responden terhadap light mode dan dark mode. Preferensi disusun berdasarkan enam aspek yang relevan dengan penggunaan mode tampilan, yaitu keterbacaan, kelelahan mata, usability, konsumsi baterai, efisiensi, dan estetika. Pemilihan aspek-aspek tersebut mengacu pada kajian sebelumnya yang menunjukkan bahwa preferensi terhadap light mode dan dark mode dipengaruhi oleh kenyamanan visual, keterbacaan, efisiensi penggunaan, dan karakteristik perangkat [20]. Setiap aspek terdiri atas delapan pernyataan, yaitu empat pernyataan positif dan empat pernyataan negatif, sehingga total item pada kuesioner preferensi adalah 48 pernyataan. Jumlah item yang relatif banyak berpotensi menimbulkan kelelahan responden, sehingga menjadi salah satu keterbatasan penelitian.

Seluruh item preferensi dinilai menggunakan skala Likert 1–7. Pada tahap analisis, pernyataan negatif dibalik (reverse scoring) agar arah interpretasi seluruh item tetap konsisten. Reverse scoring dilakukan menggunakan rumus  $(8 - \text{nilai asli})$  karena skala pengukuran menggunakan rentang 1–7, kemudian seluruh item digabungkan dan dihitung nilai rata-ratanya untuk setiap aspek. Instrumen ini disusun untuk kebutuhan penelitian dan digunakan untuk menggambarkan kecenderungan preferensi pengguna terhadap mode tampilan yang diuji.

### 2.5 Teknik Analisis Data

Analisis data dilakukan secara kuantitatif dengan menggabungkan analisis deskriptif dan inferensial. Sebelum uji perbedaan dilakukan, normalitas data diperiksa menggunakan uji Shapiro–Wilk pada skor selisih pasangan (paired differences) antara light mode dan dark mode untuk setiap tugas, yaitu diff\_T1, diff\_T2, dan diff\_T3. Penggunaan skor selisih pasangan pada uji normalitas dilakukan karena penelitian ini menggunakan desain within-subject, sehingga setiap responden diuji pada dua kondisi yang berbeda [21].

Berdasarkan hasil uji normalitas, distribusi skor selisih pasangan pada data Time on Task dan error rate tidak memenuhi asumsi normalitas secara konsisten. Pada data Time on Task, sebagian skor selisih pasangan menunjukkan nilai signifikansi di bawah 0,05. Sementara itu, pada data error rate, seluruh skor selisih pasangan menunjukkan nilai signifikansi di bawah 0,05. Oleh karena itu, analisis perbandingan antara light mode dan dark mode pada kedua parameter tersebut dilakukan menggunakan uji non-parametrik Wilcoxon Signed-Rank Test [22]. Analisis Time on Task dilakukan untuk membandingkan waktu penyelesaian tugas pada masing-masing mode tampilan untuk setiap tugas pengujian. Analisis error rate dilakukan untuk membandingkan tingkat kesalahan pengguna pada masing-masing mode tampilan untuk setiap tugas. Kedua parameter tersebut dianalisis secara terpisah pada aplikasi Facebook dan Tokopedia.

Data UEQ dianalisis secara deskriptif untuk memperoleh nilai rata-rata tiap skala, pemeriksaan inconsistency, serta interpretasi berdasarkan benchmark UEQ [18], [19]. Selain itu, skor UEQ pada setiap skala dibandingkan antara light mode dan dark mode untuk melihat apakah terdapat perbedaan penilaian antar mode.

Sementara itu, data kuesioner preferensi pengguna dianalisis secara deskriptif dengan menghitung kecenderungan skor pada setiap aspek, yaitu keterbacaan, kelelahan mata, usability, konsumsi baterai, efisiensi, dan estetika. Hasil analisis ini digunakan untuk memberikan gambaran tambahan mengenai kecenderungan pilihan responden terhadap light mode dan dark mode. Seluruh pengujian inferensial dalam penelitian ini menggunakan tingkat signifikansi 0,05.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Karakteristik Responden

Karakteristik responden dalam penelitian ini disajikan pada Tabel 3. Total responden dalam penelitian ini berjumlah 25 orang, dengan komposisi 13 responden laki-laki (52%) dan 12 responden perempuan (48%). Rata-rata usia responden adalah 36,5 tahun dengan rentang usia 19–54 tahun. Berdasarkan kategori generasi, responden didominasi oleh Generasi Z sebanyak 9 orang (36%), diikuti oleh Generasi Milenial sebanyak 8 orang (32%) dan Generasi X sebanyak 8 orang (32%). Dari segi pekerjaan, mayoritas responden merupakan karyawan, sedangkan sisanya berasal dari kategori pelajar/mahasiswa dan beberapa pekerjaan lain. Sebagian besar responden menggunakan perangkat Android (64%) dan layar OLED/AMOLED (72%). Selain itu, mayoritas responden juga lebih sering menggunakan light mode (72%) dibandingkan dark mode (28%).

**Tabel 3.** Karakteristik Responden

Karakteristik	Kategori	Jumlah	Persentase
Jenis kelamin	Laki-laki	13	52%
	Perempuan	12	48%
Usia	Rata-rata	36,5 tahun	-
	Rentang	19–54 tahun	-
Generasi	Gen Z	9	36%
	Milenial	8	32%
	Gen X	8	32%
Pekerjaan	Karyawan	11	44%
	Pelajar/Mahasiswa	3	12%
	Lainnya	11	44%
Perangkat	Android	16	64%
	iOS	9	36%
Tipe layar	OLED/AMOLED	18	72%
	IPS LCD	7	28%
Kebiasaan mode tampilan	Light mode	18	72%
	Dark mode	7	28%

#### 3.2 Hasil Pengujian Objektif

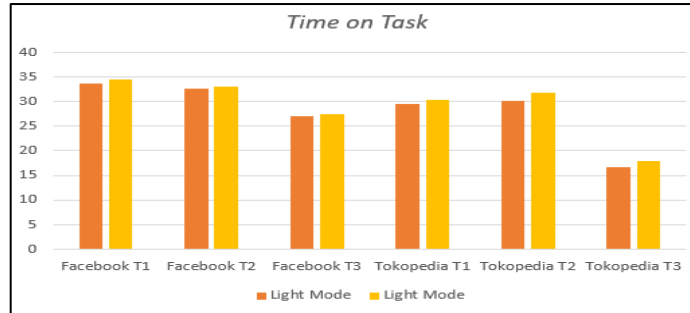
##### 3.2.1 Hasil Time on Task

Hasil statistik deskriptif dan uji Wilcoxon untuk parameter Time on Task disajikan pada Tabel 4. Pada aplikasi Facebook, median waktu penyelesaian tugas antara light mode dan dark mode menunjukkan perbedaan yang relatif kecil pada seluruh tugas. Hasil uji Wilcoxon juga menunjukkan bahwa tidak terdapat perbedaan signifikan pada T1, T2, maupun T3. Pada aplikasi Tokopedia, median Time on Task pada dark mode cenderung lebih tinggi dibandingkan light mode pada seluruh tugas. Hasil uji Wilcoxon menunjukkan bahwa perbedaan tersebut signifikan pada T1 dan T3, tetapi tidak signifikan pada T2.

**Tabel 4.** Hasil Time on Task pada Facebook dan Tokopedia

Aplikasi	Tugas	Light Mode (Median ± IQR)	Dark Mode (Median ± IQR)	Z	p-value	Keputusan
Facebook	T1	33 ± 9	36 ± 12	-1,239	0,215	Tidak signifikan
Facebook	T2	33 ± 8	32 ± 8	-0,838	0,402	Tidak signifikan
Facebook	T3	26 ± 6	27 ± 6	-0,852	0,394	Tidak signifikan
Tokopedia	T1	29 ± 5,5	31 ± 8	-2,255	0,024	Signifikan
Tokopedia	T2	30 ± 4	31 ± 8	-1,841	0,066	Tidak signifikan
Tokopedia	T3	16 ± 5	18 ± 6	-1,985	0,047	Signifikan

Untuk memperjelas perbandingan time on task antara light mode dan dark mode pada setiap tugas, visualisasi data disajikan pada Gambar 2.



**Gambar 2.** Perbandingan Time on Task pada Facebook dan Tokopedia

Berdasarkan Tabel 4 dan Gambar 2, pada Facebook perbedaan Time on Task antara light mode dan dark mode relatif kecil dan tidak signifikan. Sebaliknya, pada Tokopedia, dark mode menunjukkan waktu penyelesaian yang lebih tinggi pada seluruh tugas, dengan perbedaan signifikan pada T1 dan T3. Temuan ini tetap perlu diinterpretasikan secara hati-hati karena penelitian tidak menerapkan counterbalancing dan dilakukan pada lingkungan pengujian yang bervariasi.

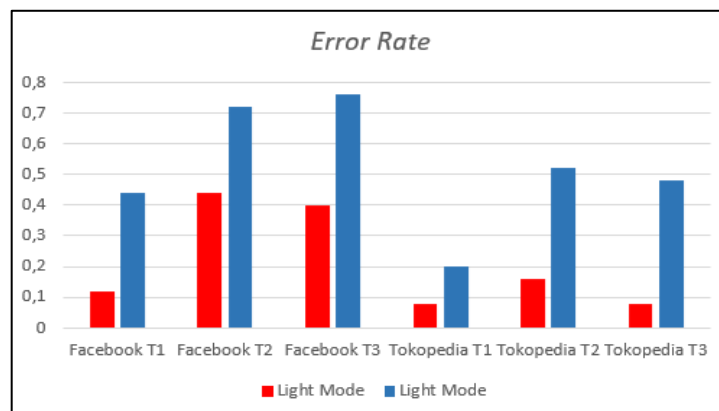
### 3.2.2 Hasil Error Rate

Hasil statistik deskriptif dan uji Wilcoxon untuk parameter Error Rate disajikan pada Tabel 5. Pada aplikasi Facebook, median error rate pada T1 adalah 0 pada light mode dan 0 pada dark mode, pada T2 sebesar 0 pada light mode dan 1 pada dark mode, serta pada T3 sebesar 0 pada kedua mode. Hasil uji Wilcoxon menunjukkan bahwa perbedaan error rate pada Facebook signifikan pada T1, tetapi tidak signifikan pada T2 dan T3. Pada aplikasi Tokopedia, median error rate pada seluruh tugas bernilai 0 baik pada light mode maupun dark mode. Namun, nilai IQR pada dark mode lebih besar pada T2 dan T3 dibandingkan light mode. Hasil uji Wilcoxon menunjukkan bahwa perbedaan error rate pada Tokopedia signifikan pada T2 dan T3, tetapi tidak signifikan pada T1.

**Tabel 5.** Hasil Error Rate pada Facebook dan Tokopedia

Aplikasi	Tugas	Light Mode (Median ± IQR)	Dark Mode (Median ± IQR)	Z	p-value	Keputusan
Facebook	T1	0 ± 0	0 ± 1	-2,000	0,046	Signifikan
Facebook	T2	0 ± 1	1 ± 1	-1,538	0,124	Tidak signifikan
Facebook	T3	0 ± 1	0 ± 2	-1,789	0,074	Tidak signifikan
Tokopedia	T1	0 ± 0	0 ± 0	-1,134	0,257	Tidak signifikan
Tokopedia	T2	0 ± 0	0 ± 1	-2,324	0,020	Signifikan
Tokopedia	T3	0 ± 0	0 ± 1	-2,352	0,019	Signifikan

Untuk memperjelas perbandingan error rate antara light mode dan dark mode pada setiap tugas, visualisasi data disajikan pada Gambar 3.



**Gambar 3.** Perbandingan Error rate pada Facebook dan Tokopedia

Berdasarkan Tabel 5 dan Gambar 3, nilai error rate pada kedua aplikasi secara umum rendah. Perbedaan signifikan hanya muncul pada T1 di Facebook serta T2 dan T3 di Tokopedia. Hasil ini menunjukkan bahwa pengaruh mode tampilan terhadap kesalahan pengguna tidak muncul pada seluruh tugas dan tetap perlu dibaca hati-hati karena potensi bias akibat urutan pengujian yang tidak diacak serta kondisi lingkungan yang tidak sepenuhnya terkontrol.

**3.3 Hasil Evaluasi UEQ**

Hasil evaluasi pengalaman pengguna menggunakan UEQ disajikan pada Tabel 6 dan Tabel 7. Dalam penelitian ini, hasil UEQ dipaparkan berdasarkan nilai mean setiap skala, kategori benchmark, serta hasil uji perbandingan antara light mode dan dark mode. Penggunaan UEQ untuk mengevaluasi pengalaman pengguna pada berbagai skenario produk interaktif telah banyak digunakan [18], sedangkan interpretasi kategori benchmark mengacu pada kategori excellent, good, above average, below average, dan bad [19].

**Tabel 6.** Hasil UEQ pada Aplikasi Facebook

Skala	Mean Light Mode	Benchmark Light Mode	Mean Dark Mode	Benchmark Dark Mode	p-value	Hasil
Daya tarik	1,313	Above Average	1,093	Below Average	0,5254	Tidak signifikan
Kejelasan	1,190	Below Average	1,110	Below Average	0,8136	Tidak signifikan
Efisiensi	1,280	Above Average	0,930	Below Average	0,3311	Tidak signifikan
Ketepatan	1,260	Above Average	1,030	Below Average	0,4803	Tidak signifikan
Stimulasi	1,030	Above Average	0,870	Below Average	0,6698	Tidak signifikan
Kebaruan	0,540	Below Average	0,430	Below Average	0,7433	Tidak signifikan

Pada aplikasi Facebook, seluruh nilai mean UEQ berada pada rentang positif (>0). Meskipun terdapat perbedaan nilai secara deskriptif antara Light Mode dan Dark Mode, hasil uji inferensial menunjukkan bahwa tidak terdapat perbedaan signifikan antara kedua mode pada seluruh skala UEQ ( $p > 0,05$ ), sehingga tidak dapat disimpulkan adanya perbedaan yang bermakna. Berdasarkan benchmark, light mode berada pada kategori above average pada daya tarik, efisiensi, ketepatan, dan stimulasi, sedangkan dark mode berada pada kategori below average pada seluruh skala.

**Tabel 7.** Hasil UEQ pada Aplikasi Tokopedia

Skala	Mean Light Mode	Benchmark Light Mode	Mean Dark Mode	Benchmark Dark Mode	p-value	Hasil
Daya tarik	1,360	Above Average	0,953	Below Average	0,2979	Tidak signifikan
Kejelasan	1,580	Above Average	1,040	Below Average	0,1738	Tidak signifikan
Efisiensi	1,450	Above Average	1,040	Below Average	0,2649	Tidak signifikan
Ketepatan	1,380	Above Average	0,890	Below Average	0,1668	Tidak signifikan
Stimulasi	1,150	Above Average	1,020	Above Average	0,7352	Tidak signifikan
Kebaruan	0,600	Below Average	0,940	Above Average	0,3096	Tidak signifikan

Pada aplikasi Tokopedia, seluruh nilai mean UEQ juga berada pada rentang positif (>0). Meskipun light mode menunjukkan nilai lebih tinggi pada lima dari enam skala, terutama pada kejelasan, efisiensi, dan ketepatan, hasil uji inferensial menunjukkan bahwa tidak terdapat perbedaan signifikan antara kedua mode ( $p > 0,05$ ). Berdasarkan benchmark, light mode berada pada kategori above average pada lima skala, sedangkan dark mode hanya mencapai kategori above average pada stimulasi dan kebaruan. Dengan demikian, perbedaan yang terlihat pada Tokopedia hanya menunjukkan kecenderungan deskriptif dan tidak dapat diinterpretasikan sebagai perbedaan pengalaman pengguna yang bermakna secara statistik.

Jika dilihat berdasarkan pengelompokan skala UEQ [17], pada aplikasi Facebook kualitas pragmatis yang terdiri atas kejelasan, efisiensi, dan ketepatan menunjukkan nilai yang lebih tinggi dibandingkan kualitas hedonis, baik pada light mode maupun dark mode. Pada Facebook, kualitas pragmatis tercatat sebesar 1,24 pada light mode dan 1,02 pada dark mode, sedangkan kualitas hedonis sebesar 0,79 pada light mode dan 0,65 pada dark mode. Pola yang sama juga terlihat pada Tokopedia, di mana light mode menunjukkan kualitas pragmatis yang lebih tinggi dibandingkan kualitas hedonis, masing-masing sebesar 1,47 dan 0,88. Sementara itu, pada dark mode Tokopedia, nilai kualitas pragmatis dan hedonis relatif berdekatan, yaitu 0,99 dan 0,98. Temuan ini hanya menunjukkan pola penilaian secara deskriptif pada aspek kegunaan dan pengalaman visual, bukan perbedaan yang terbukti signifikan secara statistik.

Selain hasil mean, benchmark, dan perbandingan skor UEQ antara light mode dan dark mode, tingkat inconsistency  $\geq 3$  pada jawaban UEQ dengan total responden pada masing-masing kondisi sebanyak 25 orang disajikan pada Tabel 8. Tingkat inkonsistensi tertinggi terdapat pada kondisi Facebook light mode sebesar 32%, sedangkan yang terendah terdapat pada Tokopedia dark mode sebesar 8%. Seluruh data tetap digunakan dalam analisis untuk menjaga kelengkapan data pada desain within-subject. Namun demikian, tingkat inkonsistensi yang relatif tinggi, khususnya pada Facebook light mode, menunjukkan bahwa kualitas data subjektif perlu diperhatikan, sehingga hasil UEQ pada penelitian ini lebih tepat dipahami sebagai kecenderungan umum dan diinterpretasikan secara hati-hati.

**Tabel 8.** Jumlah Responden Tidak Konsisten pada Setiap Kondisi UEQ (n = 25)

Kondisi	Tidak Konsisten ( $\geq 3$ )	Persentase
Facebook Light Mode	8	32%
Facebook Dark Mode	6	24%
Tokopedia Light Mode	7	28%
Tokopedia Dark Mode	2	8%

### 3.4 Hasil Preferensi Pengguna

Hasil rekapitulasi preferensi pengguna terhadap light mode dan dark mode disajikan pada Tabel 9. Secara umum, preferensi pengguna tidak sepenuhnya mengarah pada satu mode tampilan tertentu, melainkan bergantung pada aspek yang dinilai. Sebagian besar aspek menunjukkan kecenderungan pada light mode, yaitu keterbacaan, usability, efisiensi kinerja, serta estetika dan daya tarik. Sebaliknya, aspek kelelahan mata dan konsumsi baterai menunjukkan kecenderungan pada dark mode. Temuan ini menunjukkan kecenderungan preferensi terhadap light mode pada aspek yang berkaitan dengan performa tugas dan kemudahan penggunaan, sedangkan dark mode lebih disukai pada aspek kenyamanan visual dan efisiensi energi.

**Tabel 9.** Rekapitulasi Preferensi Pengguna pada Setiap Aspek

Aspek	Mean	Preferensi
Keterbacaan (Readability)	2,955	Light Mode
Kelelahan Mata (Eye Strain)	4,135	Dark Mode
Usability	3,180	Light Mode
Konsumsi Baterai	4,610	Dark Mode
Efisiensi Kinerja	3,360	Light Mode
Estetika dan Daya Tarik	3,430	Light Mode

Berdasarkan Tabel 9, light mode lebih dipilih pada empat aspek, yaitu keterbacaan, usability, efisiensi kinerja, serta estetika dan daya tarik. Sementara itu, dark mode lebih dipilih pada aspek kelelahan mata dan konsumsi baterai. Hasil ini menunjukkan bahwa preferensi pengguna bersifat kontekstual dan tidak mutlak pada satu mode tertentu. Dengan kata lain, light mode cenderung lebih mendukung aspek yang berkaitan dengan performa tugas dan kemudahan penggunaan, sedangkan dark mode lebih berkaitan dengan kenyamanan visual dan efisiensi energi. Temuan preferensi ini bersifat subjektif dan tidak secara langsung merepresentasikan performa penggunaan, sehingga tidak dapat disimpulkan bahwa mode yang lebih disukai memberikan performa yang lebih baik.

### 3.5 Pembahasan

Secara umum, hasil penelitian menunjukkan bahwa pengaruh perbedaan mode tampilan antara light mode dan dark mode tidak bersifat seragam pada seluruh kondisi pengujian. Pada aplikasi Facebook, perbedaan mode tampilan tidak menunjukkan pengaruh yang signifikan terhadap Time on Task, dan hanya memengaruhi error rate pada satu tugas. Sebaliknya, pada aplikasi Tokopedia, pengaruh mode tampilan terlihat lebih jelas, baik pada Time on Task maupun error rate, meskipun tidak terjadi pada seluruh tugas. Pola ini menunjukkan bahwa efek mode tampilan sangat bergantung pada karakteristik aplikasi dan jenis tugas yang dilakukan pengguna [6], [20]. Namun demikian, temuan pada tugas-tugas tertentu tetap perlu diinterpretasikan secara hati-hati karena penelitian ini tidak menerapkan counterbalancing dan dilakukan pada lingkungan pengujian yang bervariasi, sehingga masih terdapat kemungkinan pengaruh order effect dan kondisi lingkungan terhadap hasil yang diperoleh.

Perbedaan hasil antara Facebook dan Tokopedia menunjukkan bahwa pengaruh mode tampilan bersifat kontekstual. Pada Facebook, aktivitas pengguna lebih banyak berupa konsumsi konten, pencarian akun atau halaman, dan interaksi sosial yang relatif familiar, sehingga perubahan mode tampilan tidak cukup kuat untuk mengubah efisiensi dan akurasi secara berarti. Sebaliknya, pada Tokopedia pengguna dihadapkan pada antarmuka yang lebih padat informasi, seperti harga, rating, deskripsi, dan tombol aksi, sehingga perbedaan mode tampilan lebih mungkin memengaruhi kecepatan dan ketepatan pengguna. Temuan ini sejalan dengan penelitian yang menunjukkan bahwa perbedaan mode latar dapat memengaruhi performa kognitif [5] dan bahwa kombinasi warna teks-latar dapat memengaruhi keterbacaan serta efisiensi pengenalan teks [8].

Hasil UEQ juga menunjukkan pola yang sejalan dengan temuan tersebut. Pada kedua aplikasi, light mode cenderung memiliki nilai deskriptif lebih tinggi dibandingkan dark mode, terutama pada skala pragmatis seperti kejelasan, efisiensi, dan ketepatan. Namun, karena hasil uji inferensial tidak menunjukkan perbedaan signifikan, temuan ini tidak cukup kuat untuk menyimpulkan adanya perbedaan pengalaman pengguna secara statistik. UEQ dan benchmark tetap relevan untuk membaca kecenderungan umum pengalaman pengguna [18], [19].

Jika dilihat secara terpadu, kombinasi hasil Time on Task, Error Rate, dan UEQ menunjukkan bahwa perbedaan mode tampilan lebih terlihat pada performa tugas tertentu dibandingkan pada persepsi pengalaman pengguna secara keseluruhan. Pada Tokopedia, misalnya, perbedaan pada metrik objektif muncul pada beberapa tugas, tetapi tidak diikuti oleh perbedaan signifikan pada seluruh skala UEQ. Hal ini menunjukkan bahwa evaluasi



objektif dan subjektif dapat memberikan informasi yang saling melengkapi, tetapi tidak selalu menunjukkan pola yang sama. Dengan demikian, penggunaan lebih dari satu jenis metrik menjadi penting agar pengalaman pengguna dapat dipahami secara lebih utuh [11], [12].

Temuan penelitian ini juga mendukung kajian sebelumnya yang menunjukkan bahwa pengaruh light mode dan dark mode bersifat kontekstual dan dipengaruhi oleh kenyamanan visual, keterbacaan, karakteristik perangkat, serta situasi penggunaan [6], [20]. Dalam penelitian ini, mayoritas responden memang lebih terbiasa menggunakan light mode, sehingga faktor kebiasaan penggunaan dapat menjadi salah satu alasan mengapa light mode cenderung memberikan hasil yang lebih baik. Namun demikian, hasil pada Tokopedia dark mode yang masih mencapai kategori above average pada skala stimulasi dan kebaruan menunjukkan bahwa dark mode tetap dapat memberikan pengalaman yang dianggap lebih menarik atau lebih segar secara visual, meskipun belum tentu lebih efisien untuk semua jenis tugas [6], [20].

Temuan dari kuesioner preferensi pengguna memperjelas hasil pengujian objektif dan UEQ. Pengguna cenderung memilih light mode pada aspek keterbacaan, usability, efisiensi kinerja, serta estetika dan daya tarik, sedangkan dark mode lebih dipilih pada aspek kelelahan mata dan konsumsi baterai. Pola ini menunjukkan bahwa preferensi terhadap mode tampilan tidak bersifat mutlak, tetapi bergantung pada kebutuhan penggunaan [6], [20]. Dengan demikian, light mode tampak lebih mendukung performa tugas dan kemudahan penggunaan, sedangkan dark mode lebih berkaitan dengan kenyamanan visual dan efisiensi energi. Namun, hasil preferensi ini tetap perlu dibaca secara hati-hati karena berasal dari data subjektif dan dikumpulkan melalui kuesioner yang relatif panjang, sehingga berpotensi dipengaruhi oleh kelelahan responden.

Dari sisi praktis, hasil ini menunjukkan bahwa pengembang aplikasi tidak sebaiknya mengasumsikan bahwa dark mode dan light mode akan memberikan efek yang sama pada semua konteks. Pada aplikasi yang menekankan pencarian informasi, pembacaan detail, dan alur keputusan bertahap, light mode menunjukkan kecenderungan mendukung efisiensi dan mengurangi potensi kesalahan. Sebaliknya, pada aplikasi yang lebih berorientasi pada konsumsi konten dan interaksi sosial, perbedaan performa antara kedua mode tampilan dapat menjadi lebih kecil. Dengan demikian, keputusan desain terkait mode tampilan sebaiknya mempertimbangkan karakteristik tugas utama yang didukung oleh aplikasi, bukan hanya preferensi estetika semata [5], [6], [8], [20].

## 4. KESIMPULAN

Penelitian ini menunjukkan bahwa pengaruh light mode dan dark mode terhadap pengalaman pengguna tidak selalu bersifat seragam, tetapi bergantung pada karakteristik aplikasi, jenis tugas, dan aspek yang dievaluasi. Pada aplikasi Facebook, perbedaan mode tampilan tidak menunjukkan pengaruh yang signifikan terhadap sebagian besar parameter, sedangkan pada aplikasi Tokopedia perbedaan hanya muncul pada tugas-tugas tertentu pada Time on Task dan Error Rate. Hasil ini menunjukkan bahwa pengaruh mode tampilan bersifat kontekstual dan tidak dapat digeneralisasi secara langsung pada semua jenis aplikasi. Dari sisi persepsi pengalaman pengguna, hasil UEQ menunjukkan bahwa kedua mode tampilan berada pada kategori positif dan tidak menunjukkan perbedaan signifikan secara statistik, sehingga perbedaan yang muncul hanya bersifat deskriptif. Hasil preferensi pengguna juga menunjukkan bahwa pilihan terhadap mode tampilan bergantung pada aspek yang dinilai, di mana light mode lebih dipilih pada aspek keterbacaan, usability, efisiensi kinerja, serta estetika dan daya tarik, sedangkan dark mode lebih dipilih pada aspek kelelahan mata dan konsumsi baterai. Namun demikian, hasil penelitian ini perlu diinterpretasikan secara hati-hati mengingat adanya keterbatasan metodologis, seperti tidak diterapkannya teknik counterbalancing dalam desain within-subject, variasi lingkungan pengujian, jumlah sampel yang relatif terbatas, serta tingkat inkonsistensi jawaban UEQ. Oleh karena itu, hasil penelitian ini lebih tepat dipandang sebagai indikasi awal yang masih memerlukan validasi lebih lanjut melalui desain eksperimen yang lebih terkontrol. Penelitian selanjutnya disarankan untuk menerapkan teknik counterbalancing, mengontrol kondisi pengujian secara lebih ketat, serta melibatkan jumlah responden yang lebih besar agar hasil yang diperoleh lebih kuat dan dapat digeneralisasikan dengan lebih baik.

## REFERENCES

- [1] D. S. Waruwu, A. Harefa, B. P. Lase, and A. T. Harefa, "Dampak Perkembangan Teknologi Smartphone pada Masyarakat," *Jiip - Jurnal Ilmiah Ilmu Pendidikan*, vol. 7, no. 10, pp. 11237–11247, Oct. 2024, doi: 10.54371/jiip.v7i10.5937.
- [2] G. Lu, S. Qu, and Y. Chen, "Understanding user experience for mobile applications: a systematic literature review," *Discover Applied Sciences*, vol. 7, no. 6, Jun. 2025, doi: 10.1007/s42452-025-07170-3.
- [3] W.-T. Wang, W.-M. Ou, and W.-Y. Chen, "The impact of inertia and user satisfaction on the continuance intentions to use mobile communication applications: A mobile service quality perspective," *Int. J. Inf. Manage.*, vol. 44, pp. 178–193, Feb. 2019, doi: 10.1016/j.ijinfomgt.2018.10.011.
- [4] S. Ma, C. Chen, H. Khalajzadeh, and J. Grundy, "A First Look at Dark Mode in Real-world Android Apps," *ACM Transactions on Software Engineering and Methodology*, vol. 33, no. 1, Nov. 2023, doi: 10.1145/3604607.
- [5] T. Gazit, T. Tager-Shafir, H. X. Zhong, P. C. K. Hung, and V. Cheung, "The dark side of the interface: examining the influence of different background modes on cognitive performance," *Ergonomics*, 2025, doi: 10.1080/00140139.2025.2483451.



- [6] M. R. Atsani, I. A. Mukaromah, and M. H. C. Anugerah, “The Role of Color in User Experience: A Systematic Literature Study of User Preferences for Dark and Light Mode,” *Transactions on Informatics and Data Science*, vol. 2, no. 1, pp. 41–54, Jun. 2025, doi: 10.24090/tids.v2i1.13903.
- [7] A. Shrestha et al., “An Exploration of Effects of Dark Mode on University Students: A Human Computer Interface Analysis,” Sep. 2024, [Online]. Available: <http://arxiv.org/abs/2409.10895>
- [8] H. Guo, W. Wang, F. Song, Y. Liu, H. Liu, and Z. Bo, “The Effect of Color Combinations on the Efficiency of Text Recognition in Digital Devices,” *Int. J. Hum. Comput. Interact.*, vol. 40, no. 5, pp. 1317–1326, 2024, doi: 10.1080/10447318.2022.2143768.
- [9] S. M. Van Hoecke, I. Schrijver, and I. S. Robert, “Methodological Preparation of a Within-Subject Audiovisual Cognition, Reception and Perception Study,” *Journal of Audiovisual Translation*, vol. 5, no. 1, Apr. 2022, doi: 10.47476/jat.v5i1.2022.163.
- [10] P. Weichbroth, “Usability Testing of Mobile Applications: A Methodological Framework,” *Applied Sciences*, vol. 14, no. 5, p. 1792, Feb. 2024, doi: 10.3390/app14051792.
- [11] M. W. Wronikowska et al., “Systematic review of applied usability metrics within usability evaluation methods for hospital electronic healthcare record systems,” *J. Eval. Clin. Pract.*, vol. 27, no. 6, pp. 1403–1416, Dec. 2021, doi: 10.1111/jep.13582.
- [12] A. I. Nur, H. B. Santoso, and P. O. H. Putra, “The Method and Metric of User Experience Evaluation: A Systematic Literature Review,” in *2021 10th International Conference on Software and Computer Applications*, Malaysia: ACM, Feb. 2021, pp. 307–317. doi: 10.1145/3457784.3457832.
- [13] M. A. Kushendriawan, H. B. Santoso, P. O. H. Putra, and M. Schrepp, “Evaluating User Experience of a Mobile Health Application ‘Halodoc’ using User Experience Questionnaire and Usability Testing,” *Jurnal Sistem Informasi*, vol. 17, no. 1, pp. 58–71, Apr. 2021, doi: 10.21609/jsi.v17i1.1063.
- [14] E. Prafiantini et al., “Usability testing of EatsUp®: mobile application for monitoring balanced dietary practices and active lifestyle among adolescents—a study in Jakarta, Indonesia,” *Front. Digit. Health*, vol. 7, 2025, doi: 10.3389/fgth.2025.1506952.
- [15] D. A. P. Wulandari, “Analysis and Evaluation of SIDUN Mobile Application in UEQ-Based User Experience Perspective,” *Journal of Computer Networks, Architecture and High Performance Computing*, vol. 7, no. 2, pp. 554–561, May 2025, doi: 10.47709/cnahpc.v7i2.5932.
- [16] L. Faulkner, “Beyond the five-user assumption: Benefits of increased sample sizes in usability testing,” *Behavior Research Methods, Instruments, & Computers*, vol. 35, no. 3, pp. 379–383, Aug. 2003, doi: 10.3758/BF03195514.
- [17] M. Schrepp, *User Experience Questionnaire Handbook: All You Need to Know to Apply the UEQ Successfully in Your Projects - Version 11*. 2023. Accessed: Mar. 25, 2026. [Online]. Available: <https://www.ueq-online.org/Material/Handbook.pdf>
- [18] M. Schrepp, A. Hinderks, and J. Thomaschewski, “Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios,” in *Lecture Notes in Computer Science*, A. Marcus, Ed., Heraklion, Crete, Greece: Springer, Cham, Jun. 2014, pp. 383–392. doi: 10.1007/978-3-319-07668-3\_37.
- [19] M. Schrepp, A. Hinderks, and J. Thomaschewski, “Construction of a Benchmark for the User Experience Questionnaire (UEQ),” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 4, pp. 40–44, Jun. 2017, doi: 10.9781/ijimai.2017.445.
- [20] A. Yang, C. H. Goh, and L. J. Yi, “The Research into Dark Mode: A Systematic Review Using Two-Stage Approach and S-O-R Framework,” *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 15, no. 4, Jan. 2024, doi: 10.17762/ijcnis.v15i4.6301.
- [21] X. Manfei, D. Fralick, J. Z. Zheng, B. Wang, X. M. Tu, and C. Feng, “The Differences and Similarities Between Two-Sample T-Test and Paired T-Test,” *Shanghai Arch. Psychiatry*, vol. 29, no. 3, pp. 184–188, Jun. 2017, doi: 10.11919/j.issn.1002-0829.217070.
- [22] E. Whitley and J. Ball, “Statistics review 6: Nonparametric methods,” *Crit. Care*, vol. 6, no. 6, pp. 509–513, 2002, doi: 10.1186/cc1820.