

# Pengaruh Inkonsistensi Switching Light-Dark Mode Antar Aplikasi Terhadap Mental Workload dan Kinerja Pengguna Smartphone

Layla Ayu Mustika Putri, Alfian Nurlifa\*, Andik Adi Suryanto

Fakultas Teknik, Teknik Informatika, Universitas PGRI Ronggolawe, Tuban, Indonesia  
Email: <sup>1</sup>laylaid51@gmail.com, <sup>2</sup>lifa.nurlifa13@gmail.com, <sup>3</sup>andikadisuryanto@gmail.com  
Email Penulis Korespondensi: lifa.nurlifa13@gmail.com

Submitted 04-02-2026; Accepted 29-04-2026; Published 30-04-2026

## Abstrak

Inkonsistensi visual dalam konteks antarmuka pengguna seperti perubahan tema visual pada mode switching yakni dari Light ke Dark Mode sering dianggap sebagai elemen estetika yang bersifat opsional, namun dampak dari hal ini terhadap beban kognitif dan kinerja dari pengguna belum diteliti secara luas. Penelitian ini bertujuan untuk menganalisis pengaruh mode switching terhadap beban kognitif subjektif, efisiensi waktu penyelesaian tugas serta akurasi dari kinerja pengguna dalam konteks interaksi antarmuka. Metode yang digunakan dalam eksperimen ini adalah Quasi Experimental Design dengan tipe posttest only control group, dengan melibatkan sebanyak 40 responden dan dibagi kedalam dua kelompok dengan kondisi tema statis dan kondisi mode switching. Beban kognitif subjektif diukur dari instrumen NASA-TLX sementara kinerja objektif dievaluasi melalui Task Completion Time dan Error Rate. Hasil dari penelitian ini menunjukkan bahwa kelompok perlakuan mengalami peningkatan beban kognitif, dengan rata-rata skor WWL sebesar 55.23, dibanding dengan kelompok kontrol sebesar 42.50. Dimensi frustrasi dengan hasil tertinggi sebesar 28.0 menjadi bahwa tekanan emosional akibat dari inkonsistensi antarmuka. Secara objektif, penerapan mode switching memperlambat waktu penyelesaian task sebesar 3.65 detik dan meningkatkan error rate hingga dua kali lipat. Hasil pengujian statistik juga menunjukkan perbedaan pada seluruh variabel penelitian, dengan  $p\text{-value} < 0.005$  dan nilai effect size pada beban kognitif sebesar 0.816 (Large Effect). Temuan ini menghasilkan kesimpulan bahwa stabilitas visual merupakan faktor penting dalam desain antarmuka, hal ini diperlukan untuk menjaga efisiensi navigasi, meningkatkan kenyamanan pengguna, serta meminimalkan kesalahan interaksi.

**Kata Kunci:** Inkonsistensi; Light Dark Mode; Mental Workload; NASA-TLX; Desain Quasi Eksperimen

## Abstract

Visual inconsistency in user interface contexts, such as visual theme changes during mode switching (from Light to Dark Mode), is often considered an optional aesthetic element; however, its impact on user cognitive load and performance has not been widely researched. This study aims to analyze the effect of mode switching on subjective cognitive load, task completion time efficiency, and user performance accuracy in an interface interaction context. The method used in this experiment was a Quasi-Experimental Design with a posttest-only control group type, involving 40 respondents divided into two groups: static theme condition and mode switching condition. Subjective cognitive load was measured using the NASA-TLX instrument, while objective performance was evaluated through Task Completion Time and Error Rate. The results of this study indicate that the treatment group experienced an increase in cognitive load, with an average WWL score of 55.23, compared to the control group at 42.50. The frustration dimension was the highest, at 28.0, indicating emotional pressure due to interface inconsistency. Objectively, the application of mode switching slowed down task completion time by 3.65 seconds and increased the error rate by up to two times. Statistical test results also showed differences in all research variables, with a  $p\text{-value} < 0.005$  and an effect size value on cognitive load of 0.816 (Large Effect). These findings lead to the conclusion that visual stability is an important factor in interface design, necessary for maintaining navigation efficiency, increasing user comfort, and minimizing interaction errors.

**Keywords:** Inconsistency; Light Dark Mode; Mental Workload; NASA-TLX; Quasi Experimental Design

## 1. PENDAHULUAN

Revolusi antarmuka pengguna pada perangkat seluler terutama *smartphone* yang begitu cepat menjadikan bidang *Human-Computer Interaction* (HCI) menjadi bidang kajian yang semakin diminati dalam sebuah penelitian, perkembangan dari aplikasi seluler yang selalu *up to date* menuntut kerangka desain antarmuka yang tidak hanya estetis namun juga kuat dalam penggunaannya [1], [2]. Pada konteks ini tema visual menjadi salah satu hal utama dalam pemilihan kenyamanan pengguna, salah satu tema visual yang banyak diminati adalah *Dark Mode* [3]. Pemilihan tema visual *Dark Mode* pada pengguna *smartphone* didasarkan pada kenyamanan, daya tarik visual, dan juga untuk mengurangi *eye strain* di lingkungan dengan potensi cahaya yang minim karena hal tersebut tema visual menjadi pilihan desain yang utama dan memiliki potensi implikasi signifikan terhadap pengalaman pengguna [4], [5]. Perubahan tema visual pada mode *switching* khususnya *Dark* dan *Light Mode* menciptakan tantangan tersendiri dalam menjaga konsistensi desain sebagaimana dikatakan oleh *Shneiderman's* [1]. Pada perangkat modern yang memungkinkan perubahan tema secara dinamis dan lintas aplikasi, proses perpindahan yang terjadi secara tiba-tiba seringkali menimbulkan ketidakselarasan tampilan dalam warna maupun elemen visual, kondisi ini menyebabkan terganggunya harmonisasi antarmuka pengguna dan berpotensi menurunkan kenyamanan serta pengalaman pengguna secara keseluruhan [6].

Inkonsistensi dari tampilan visual yang tidak selaras ini menjadi persoalan *usability* yang muncul dalam penerapan suatu mode, hal ini menunjukkan bahwa transisi dari perubahan tema dapat mengganggu konsistensi desain secara keseluruhan sehingga mempengaruhi kenyamanan pengguna dalam berinteraksi dengan antarmuka [1], [7]. Inkonsistensi ini berpotensi menambah *mental workload* yang diperlukan seseorang untuk memproses dan menyelesaikan suatu tugas, untuk mengukur hal tersebut penelitian ini menggunakan instrumen multi dimensi NASA *Task Load Index* (NASA-TLX). NASA-TLX dipilih karena melakukan pengukuran subjektif berdasarkan pada enam dimensi yang relevan yakni *Mental Demand*, *Physical Demand*, *Temporal Demand*, *Performance*, *Effort*, dan *Frustration*, hal ini terbukti efektif dalam mengevaluasi tuntutan kognitif dalam lingkup HCI [4], [8]. Melalui instrumen ini perubahan subjektif yang dirasakan

pengguna akibat mode *switching* dapat diukur secara sistematis dan memberikan data kunci mengenai seberapa besar tuntutan kognitif yang ditimbulkan oleh inkonsistensi visual ini, meskipun fondasi teoritis mengenai kaitan antara inkonsistensi visual dan beban kognitif sudah ada, sebagian besarnya hanya berfokus pada perbandingan kondisi statis saja terhadap metrik kinerja seperti *Task Completion Time* (TCT), akurasi dan beban mental [4], [9]. Beban yang muncul akibat ketidakkonsistenan ini dapat mengurangi kelancaran pengguna dalam menerima dan memproses sebuah informasi, kondisi ini menunjukkan tidak sesuai sebagaimana dikatakan oleh *Sheinderman's* pada salah satu dari prinsip 8 *Golden Rules* mengenai *Reduce Short Term Memory Load* [1], [10]. Untuk menguji hubungan kausal antara inkonsistensi dan beban kognitif dalam lingkungan yang terkontrol, penelitian ini mengadopsi rancangan *Quasi Experimental Design*, penggunaan metode ini untuk memperkirakan dampak kausal suatu perlakuan pada populasi target tanpa penugasan acak penuh [11].

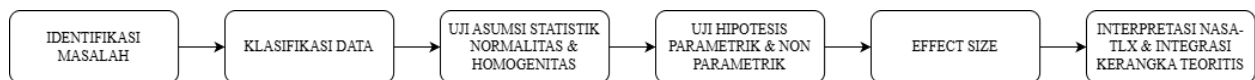
Terdapat temuan empiris yang juga menguatkan persepsi ini dengan menunjukkan bahwa perubahan visual yang terjadi secara tiba-tiba dapat berdampak pada penurunan kinerja pengguna serta kemampuan mereka dalam mempertahankan informasi, sehingga menghasilkan performa yang kurang optimal [12]. Kajian sistematis yang lain juga menegaskan bahwa aspek visual pada antarmuka memberikan peran penting dalam pembentukan respons pengguna, di mana dampaknya sering kali dipengaruhi oleh faktor internal seperti tingkat beban kognitif yang dialami [13]. Kesenjangan utama dalam literatur adalah kurangnya penelitian yang secara eksplisit mengisolasi dan mengukur efek dari proses *switching* itu sendiri, terutama dalam skenario lintas aplikasi pada tingkat sistem operasi yang lebih mencerminkan pola penggunaan *smartphone* sehari-hari [14]. Oleh karena itu, pemodelan kausal yang jelas antara pelanggaran inkonsistensi akibat *switching* dan peningkatan *Short-Term Memory Load* yang terwujud dalam penurunan kinerja objektif TCT dan *Error Rate* masih terbatas. Permasalahan dalam penelitian ini mengidentifikasi adanya perbedaan signifikan pada Beban Kognitif Subjektif yang diukur dari *mental workload* antara kelompok yang mengalami mode *switching* dengan kelompok tema stabil, menilai dampak mode *switching* terhadap aspek Efisiensi dan Efektivitas pengguna melalui indikator TCT dan *Error Rate* selama interaksi lintas aplikasi. Penelitian ini berupaya menjelaskan hubungan kausal mengenai bagaimana pelanggaran terhadap prinsip Konsistensi memengaruhi proses pengolahan informasi serta *Short Term Memory*, juga untuk mengevaluasi bagaimana ketidaksesuaian warna maupun elemen visual pada *smartphone* yang muncul akibat *switching* antara *Light Mode* dan *Dark Mode* pada tingkat sistem berpengaruh terhadap kinerja pengguna serta tingkat beban kognitif yang mereka alami.

## 2. METODOLOGI PENELITIAN

Pada penelitian ini metode yang digunakan adalah *Quasi Experimental Design* (QED) dengan pendekatan kuantitatif dan menggunakan tipe *posttest only control group*. Tipe ini dipilih karena memungkinkan peneliti untuk menguji pengaruh mode *switching* pada kelompok yang sudah tersedia tanpa perlu melakukan pengacakan secara penuh, namun tetap dapat menjaga keabsahan kausalitas melalui perbandingan antara kelompok kontrol dan kelompok perlakuan [15]. Pada pelaksanaannya, penelitian ini membagi responden ke dalam dua kondisi yakni Kelompok Kontrol dengan tema statis yang berada di *Dark Mode* dan Kelompok Perlakuan yang akan melakukan eksperimen pada kondisi *switching* yakni berada di *Light to Dark Mode*.

### 2.1 Prosedur Penelitian

Penelitian ini dilaksanakan dengan pendekatan yang terstruktur untuk memastikan keakuratan data yang diperoleh. Prosedur penelitian yang dilakukan mengikuti prosedur eksperimen kuasi yang terdiri dari beberapa langkah seperti yang ditunjukkan pada Gambar 1 dibawah ini.



**Gambar 1.** Prosedur Penelitian

Penelitian ini dimulai dari melakukan identifikasi masalah yang terkait dengan beban kognitif dalam interaksi yang dilakukan manusia, diikuti dengan penetapan variabel penelitian dan menggunakan desain *posttest only control group*. Pada tahap selanjutnya responden akan dibagi menjadi dua kelompok yakni Kelompok Kontrol dan Kelompok Perlakuan, kelompok perlakuan akan diberikan perlakuan berupa eksperimen pada kondisi *switching*, sedangkan pada kelompok kontrol akan menggunakan tema statis yang berada pada *Dark Mode*. Tahap selanjutnya pengambilan data dilakukan dengan menggunakan instrumen NASA-TLX untuk mendapatkan pengukuran secara subjektif, sedangkan pengukuran kinerja secara objektif melalui *Task Completion Time* dan *Error Rate*. Data yang terkumpul nantinya akan diklasifikasikan dan dilakukan Uji Asumsi Statistik, pada tahap uji asumsi dilakukan Uji Normalitas *Shapiro Wilk* untuk melihat distribusi data dan Uji Homogenitas *Levene Test* untuk melihat kesamaan varians antar kelompok. Pengujian tahap selanjutnya yakni Uji Hipotesis yang dilakukan berdasarkan hasil dari uji asumsi, jika data terdistribusi normal akan dilakukan uji parametrik *Independent Samples t-test* dan jika data terdistribusi tidak normal atau tidak homogen akan dilakukan uji non parametrik *Mann Whitney U Test* untuk menentukan signifikansi perbedaan nilai antar kelompok. Lalu pada tahap terakhir akan dilakukan perhitungan *effect size* untuk menentukan besarnya pengaruh perlakuan sebelum dilakukan interpretasi akhir dan penarikan kesimpulan.

## 2.1 Desain Penelitian dan Variabel

Penelitian ini menggunakan pendekatan kuantitatif dengan desain *quasi experimental* untuk mengisolasi dampak mode *switching*. Variabel independen di sini adalah kondisi stabilitas visual yang terdiri dari dua kelompok yakni kelompok kontrol yang menggunakan tema stabil dan kelompok perlakuan yang menggunakan kondisi *switching*. Untuk memecahkannya, dua beban kognitif yang dikategorikan ke dalam beban kognitif subjektif dan kinerja kognitif, diukur dengan NASA-TLX, sedangkan untuk kinerja objektif, diukur dengan *Task Completion Time* dan *Error Rate*. Data primer dalam penelitian ini diperoleh secara langsung dari eksperimen terkontrol menggunakan perangkat Android. Pada penelitian ini, peneliti menggunakan aplikasi X, pengaturan, kalkulator, dan catatan. Variasi aplikasi ini digunakan sebagai representasi dari *cross app tasks*, yang merupakan hal biasa dalam interaksi non-iteratif pengguna smartphone.

## 2.2 Responden dan Prosedur Eksperimen

Penelitian melibatkan total 40 responden dengan penentuan ukuran sampel yang didasarkan pada kebutuhan untuk menciptakan kelompok perbandingan yang seimbang. Hal ini sesuai dengan studi [16] dalam konteks simulasi penerbangan. Dalam penelitian ini, 40 responden akan dibagi secara acak menjadi 2 kelompok, yaitu Kelompok Kontrol dan Kelompok Perlakuan, masing-masing terdiri dari 20 responden, dengan kriteria usia produktif 20-25 tahun, dan tanpa buta warna serta tanpa gangguan penglihatan yang mempengaruhi persepsi luminansi dan kontras. Kelompok Kontrol memiliki peran dasar sebagai pembanding, kelompok ini akan menyelesaikan serangkaian tugas eksperimen menggunakan antarmuka dengan tema yang konsisten secara visual (Mode Gelap), untuk mengukur kinerja pengguna dan beban kognitif dalam kondisi standar. Sementara itu, Kelompok Perlakuan memiliki peran sebagai yang dikenakan perlakuan eksperimen, kelompok ini akan menyelesaikan serangkaian tugas yang sama dengan yang dilakukan oleh Kelompok Kontrol, tetapi dengan kondisi ketidakkonsistenan warna yang nyata. Perlakuan ini melibatkan peralihan mode yang drastis, yaitu transisi mendadak dari Mode Terang ke Mode Gelap di tingkat sistem operasi ketika responden beralih ke aplikasi lain. Tahapan-tahapan eksperimen dimulai dari pemberian instruksi secara detail kepada responden mengenai tugas yang akan dijalankan. Setiap peserta harus menyelesaikan empat tugas, dalam urutan berikut Information Retrieval, untuk menguji transfer data nilai numerik di berbagai aplikasi, Visual Search, untuk mengukur kecepatan identifikasi ikon selama perubahan kontras, Short-Term Numerical Memory, untuk mengukur ingatan terhadap angka setelah perubahan, dan yang terakhir Verbal Memory Recall, untuk mengukur kemampuan mengingat teks setelah perubahan topik. Selanjutnya responden mengisi instrumen NASA-TLX dan pada tahap terakhir eksperimen dilakukan pengumpulan data hasil.

## 2.3 NASA Task Load Index

Instrumen NASA-TLX digunakan dalam penelitian ini untuk mengukur beban kerja mental pengguna secara subjektif. Instrumen ini dipilih didasarkan pada tinjauan [8] yang menempatkan NASA-TLX sebagai standar *multidimensional* dalam penelitian HCI karena dapat menangkap beban kognitif yang tidak terdeteksi oleh metrik kinerja objektif. Selain itu, instrumen ini memiliki validitas psikometri yang kuat dalam mengevaluasi tuntutan proses internal pengguna [17]. *Mental workload* diukur melalui skor rata-rata dari enam dimensi penilaian subjektif seperti pada Tabel 1 dibawah ini.

**Tabel 1.** Instrumen *Subjective*

Dimensi	Deskripsi	Skala
<i>Mental Demand</i>	Seberapa besar aktivitas berpikir, memutuskan, mengingat, dan mencari (memproses informasi) yang dibutuhkan oleh tugas tadi?	0 – 100
<i>Physical Demand</i>	Seberapa banyak aktivitas fisik (menekan tombol, menggeser jari, atau menggerakkan tangan) yang dibutuhkan oleh tugas tadi?	0 – 100
<i>Temporal Demand</i>	Seberapa cepat atau lambat tugas harus diselesaikan? Seberapa tertekan Anda karena keterbatasan waktu?	0 – 100
<i>Performance</i>	Seberapa berhasil Anda dalam mencapai tujuan tugas? Seberapa puas Anda dengan kinerja Anda?	0 – 100
<i>Effort</i>	Seberapa keras Anda harus berusaha secara mental dan fisik untuk mencapai tingkat kinerja Anda?	0 – 100
<i>Frustration</i>	Seberapa tidak aman, kesal, stres, atau terganggu (vs. puas, nyaman, rileks) yang Anda rasakan selama tugas tadi?	0 – 100

Selain itu responden diminta memilih Instrumen Perbandingan (*Pairwise Comparison*) mana yang lebih dominan memengaruhi *mental workload* dalam 15 pasangan perbandingan seperti pada Tabel 2 dibawah ini.

**Tabel 2.** Instrumen Perbandingan (*Pairwise Comparison*)

No	Perbandingan	Beban Dominan
1.	<i>Mental Demand vs. Physical Demand</i>	MD / PD
2.	<i>Mental Demand vs. Temporal Demand</i>	MD / TD

3.	<i>Mental Demand vs. Performance</i>	MD / PE
4.	<i>Mental Demand vs. Effort</i>	MD / EF
5.	<i>Mental Demand vs. Frustration</i>	MD / FR
6.	<i>Physical Demand vs. Temporal Demand</i>	PD / TD
7.	<i>Physical Demand vs. Performance</i>	PD / PE
8.	<i>Physical Demand vs. Effort</i>	PD / EF
9.	<i>Physical Demand vs. Frustration</i>	PD / FR
10.	<i>Temporal Demand vs. Performance</i>	TD / PE
11.	<i>Temporal Demand vs. Effort</i>	TD / EF
12.	<i>Temporal Demand vs. Frustration</i>	TD / FR
13.	<i>Performance vs. Effort</i>	PE / EF
14.	<i>Performance vs. Frustration</i>	PE / FR
15.	<i>Effort vs. Frustration</i>	EF / FR

Setelah pengumpulan data, penjumlahan akhir dilakukan menggunakan rumus *Total Weighted Workload* (WWL). Rumus ini dihitung untuk memperoleh satu skor yang mewakili beban kerja kognitif pengguna. Seperti yang diterapkan pada studi psikometri NASA-TLX [17], skor akhir ini berfungsi sebagai metrik komparatif utama untuk menentukan signifikansi perbedaan beban mental antara Kelompok Kontrol dan Kelompok Perlakuan.

$$WWL = \frac{\sum_{i=1}^6 (Rating_i \times Weight_i)}{Total\ Weight} \quad (1)$$

## 2.4 Teknik Analisis Data

Proses analisis dilakukan dalam beberapa tahap untuk memastikan bahwa hasil yang diperoleh valid secara empiris dan dapat menguji perbedaan antara kelompok kontrol dan kelompok perlakuan. Analisis ini mencakup pengujian asumsi yang terdiri dari uji normalitas dan uji homogenitas, uji perbedaan (hipotesis), dan effect size. Metode-metode ini umum digunakan dalam penelitian HCI modern untuk mengevaluasi kinerja pengguna dan beban kognitif [4], [9], [14]. Pada tahap pertama dilakukan uji asumsi, uji ini terdiri dari uji normalitas dan homogenitas, menurut [15], [18] uji asumsi berfungsi sebagai prasyarat mutlak untuk menentukan jenis analisis statistik yang valid (Parametrik vs Non-Parametrik), uji normalitas (*Shapiro Wilk*) merupakan uji yang harus dilakukan sebelum melakukan analisis parametrik guna memastikan data berdistribusi normal, sedangkan uji homogenitas (*Uji Levene*) memastikan varians antara kelas eksperimen dan kontrol bersifat homogen.

Untuk selanjutnya dilakukan uji hipotesis, Uji hipotesis digunakan untuk membuktikan adanya perbedaan yang signifikan akibat perlakuan (intervensi) yang diberikan. Menurut [18] uji hipotesis parametrik *Independen T-test* dilakukan untuk melihat atau mengetahui apakah ada perbedaan secara signifikan antara data nilai sebelum dan sesudah perlakuan pada kelompok eksperimen dan kontrol. Jika nilai signifikansi  $< 0,05$  ( $H_0$  ditolak), hal ini membuktikan adanya pengaruh nyata dari media yang diterapkan. Sedangkan menurut [15] uji hipotesis Non-parametrik *Mann-Whitney* digunakan untuk menentukan apakah terdapat perbedaan nilai yang signifikan antara kelas kontrol dan eksperimen pada variabel terikat. Penolakan  $H_0$  menunjukkan bahwa intervensi memberikan dampak yang berbeda secara statistik dibanding kelompok kontrol.

Untuk uji terakhir adalah analisis *effect size*, menggunakan *Cohen's d* untuk uji parameterik (2) dan *Rank Biserial Correlation* untuk uji non-parameterik (3), uji ini digunakan untuk mengetahui seberapa besar pengaruh mode switching, bukan hanya signifikan atau tidak. Menurut [18] Effect size dihitung untuk mengetahui "nilai keefektifan" penggunaan intervensi. Nilai effect size yang tinggi menunjukkan bahwa perbedaan atau efek yang diamati sangat signifikan dan besar, serta mengindikasikan bahwa variabel bebas memiliki pengaruh yang cukup tinggi terhadap variabel terikat.

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s_p} \quad (2)$$

$$r_{rb} = 1 - \frac{2u}{n_1 n_2} \quad (3)$$

## 3. HASIL DAN PEMBAHASAN

Bagian ini menjelaskan analisa dari penelitian mengenai efek dari setiap inkonsistensi *switching* terhadap *mental workload* user yang diawali dengan analisa statistik deskriptif yang berfungsi untuk menggambarkan secara umum mengenai data, dilanjutkan dengan analisa untuk uji prasyarat yaitu analisa normalitas dan analisa homogenitas, untuk mengetahui jenis uji yang akan digunakan, serta analisa uji hipotesis ditambah perhitungan ukuran efek (*effect size*) analisa untuk mengukur signifikansi serta estimasi besaran dampak dari suatu intervensi.

### 3.1 Analisis Statistik Deskriptif

Pada analisis statistik deskriptif akan menampilkan gambaran umum data dari hasil perhitungan ketiga variabel, yakni *Weighted Workload*, *Error Rate*, dan *Task Completion Time*.

### 3.1.1 NASA TASK LOAD INDEX

Hasil dari perbandingan rata-rata pada skor dimensi NASA-TL ditampilkan dalam Tabel 3 dibawah ini.

**Tabel 3.** Perbandingan rata-rata skor dimensi NASA-TLX

Dimensi	Kelompok Perlakuan ( <i>Mean</i> )	Kelompok Kontrol ( <i>Mean</i> )	Selisih
<i>Mental Demand</i>	46.5	35	11.5
<i>Physical Demand</i>	40.5	23	17.5
<i>Temporal Demand</i>	43	41.5	1.5
<i>Performance</i>	73	75.5	2.5
<i>Effort</i>	57	41	16.0
<i>Frustration</i>	46	18	28.0

Dari Tabel 3 diatas dapat diketahui bahwa dimensi Frustrasi memiliki selisih terbesar yakni 28.0, yang selaras dengan intervensi mode-switching, dibandingkan dengan kondisi statis, yang secara signifikan meningkatkan ketidaknyamanan emosional pengguna. Ini sejalan dengan [18] yang menjelaskan perbedaan skor besar sebagai respon yang lebih kuat dari variabel independen untuk dimensi afektif pengguna. Selanjutnya, terjadi peningkatan masing-masing sebesar 17.5 dan 16.0 pada Physical Demand dan Effort, yang berarti bahwa visual yang berubah untuk dimensi lebih berdampak. Di sisi lain, Kinerja mengalami penurunan sedikit sebesar -2.5. Ini menunjukkan persepsi yang lebih besar terhadap penyelesaian tugas yang sukses di antara peserta dalam perlakuan. Secara keseluruhan, peningkatan di hampir semua dimensi NASA-TLX meningkatkan hipotesis awal yang konsisten tentang hubungan kausal bahwa inkonsistensi visual akibat mode yang dialihkan dan beban kognitif subjektif, seperti yang dinyatakan dalam [18] sebelum pengujian signifikansi statistik lebih lanjut

### 3.1.2 Error Rate

Berdasarkan pada Tabel 4 dibawah ini menampilkan bahwa kelompok perlakuan memiliki rata-rata kesalahan yang lebih tinggi dengan Mean 1.35 dan nilai maksimum yang juga lebih besar, sedangkan kelompok kontrol dengan Mean 0.65. Hal ini menandakan bahwa mode switching meningkatkan tingkat kesalahan pengguna jika dibandingkan dengan penggunaan tema statis.

**Tabel 4.** Hasil *Mean Error Rate*

Kelompok	<i>Mean (Fault)</i>	Min	Max
Perlakuan	1.35	0	3
Kontrol	0.65	0	2

Hasil ini sejalan dengan penelitian [19] yang menyatakan bahwa desain antarmuka yang tidak intuitif dan sulit diakses fiturnya secara langsung menurunkan tingkat keberhasilan tugas pengguna. Ketidakkonsistenan visual memaksa pengguna melakukan trial and error yang lebih sering, sehingga probabilitas terjadinya kesalahan menjadi lebih tinggi.

### 3.1.3 Task Completion Time

Pada Tabel 5 dibawah ini terdapat hasil dari TCT dimana mengukur waktu penyelesaian tugas menunjukkan bahwa kelompok perlakuan mengalami peningkatan waktu penyelesaian tugas dengan *mean* 11.95 detik dibandingkan kelompok kontrol dengan *mean* 8.30 detik. Perbedaan ini menunjukkan bahwa penggunaan tema statis jauh lebih meningkatkan penyelesaian tugas dibandingkan dengan penggunaan Mode Switching.

**Tabel 5.** Hasil *Mean TCT*

Kelompok	<i>Mean (Detik)</i>	Min	Max
Perlakuan	11.94	5.25	22.25
Kontrol	8.30	4.75	16.75

Dalam hal ini, peningkatan waktu penyelesaian tugas yang terjadi pada kelompok perlakuan mengindikasikan bahwa terdapat *hurdle* atau hambatan dalam efisiensi penyelesaian tugas yang disebabkan oleh desain antarmuka yang tidak konsisten. Hal ini sejalan dengan prinsip evaluasi pengalaman pengguna, di mana dalam suatu platform digital, efisiensi navigasi harus menjadi fokus utama untuk mencapai tingkat keberhasilan yang lebih [20].

### 3.1.4 Perbandingan Variabel

Dari hasil Tabel 6 dibawah ini menunjukkan bahwa kelompok perlakuan secara konsisten memiliki skor lebih tinggi pada semua variabel studi dibandingkan kelompok kontrol, yang mencakup *mental workload* NASA-TLX, TCT, dan *Error Rate*.

**Tabel 6.** Perbandingan Total Mean Variabel

Variabel	Perlakuan ( <i>Mean</i> )	Kontrol ( <i>Mean</i> )	Selisih
NASA-TLX (WWL)	55.23	42.50	12.73
TCT	11.94	8.30	3.64
<i>Error Rate</i>	1.35	0.65	0.7

Perbedaan hasil skor seperti pada Tabel 6 diatas menunjukkan bahwa intervensi pergantian mode meningkatkan beban kognitif dan menurunkan efisiensi serta akurasi pengguna jika dibandingkan dengan menggunakan tema statis.

### 3.2 Hasil Uji Asumsi

Hasil uji normalitas dari Uji *Shapiro Wilk* dalam Tabel 7 dibawah ini, menunjukkan bahwa variabel NASA-TLX menunjukkan distribusi data normal di kelompok perlakuan dengan  $p = 0.491$  dan di kelompok kontrol dengan  $p = 0.841$ , karena nilai signifikansinya lebih besar dari 0.05. Sebaliknya, variabel TCT dan Error Rate tidak memenuhi asumsi normalitas karena ada setidaknya satu kelompok dengan nilai signifikansi  $p < 0.05$ .

**Tabel 7.** Uji Normalitas *Shapiro Wilk*

Kelompok Responden	<i>Statistic</i>	<i>df</i>	<i>Sig. (p-value)</i>
TCT Perlakuan	.948	20	.333
TCT Kontrol	.879	20	.017
<i>Error Rate</i> Perlakuan	.881	20	.018
<i>Error Rate</i> Kontrol	.734	20	<.001
NASA-TLX Perlakuan	.957	20	.491
NASA-TLX Kontrol	.974	20	.841

Selanjutnya, hasil uji homogenitas menggunakan uji *Levene* pada Tabel 8 dibawah ini, menunjukkan bahwa variabel NASA-TLX memiliki varians yang homogen  $p = 0.538 > 0.05$ . Mengacu pada kriteria pemilihan metode analisis statistik yang dikemukakan oleh [15], analisis data NASA-TLX dilanjutkan menggunakan uji parametrik, sedangkan variabel TCT dan Error Rate dianalisis menggunakan uji non-parametrik.

**Tabel 8.** Uji Homogenitas *Levene*

Total Skor Nasa	<i>Levene Statistic</i>	<i>df 1</i>	<i>df 2</i>	<i>Sig.</i>
<i>Based On Mean</i>	.386	1	38	.538
<i>Based On Median</i>	.294	1	38	.591
<i>Based On Median and Adjusted df</i>	.294	1	34.345	.591
<i>Based On Median and with Adjusted df</i>	.396	1	38	.533

### 3.3 Hasil Uji Hipotesis

Uji hipotesis dilakukan untuk mengetahui signifikansi perbedaan antara kedua kelompok. Berdasar pada hasil uji asumsi variabel NASA-TLX diuji menggunakan statistik parametrik *Independent Samples T-test*, sedangkan variabel TCT dan Error Rate menggunakan statistik non-parametrik *Mann-Whitney U Test*. Berdasarkan Tabel 9 diketahui nilai *Sig. (2-tailed)* sebesar 0.014. Karena nilai  $p < 0.05$ , sehingga  $H_0$  ditolak dan  $H_1$  diterima. Hal ini menunjukkan bahwa terdapat perbedaan beban kognitif yang signifikan antara pengguna yang terpapar mode *switching* dengan pengguna pada kondisi kontrol.

**Tabel 9.** Uji Hipotesis *Independent Samples T-test*

Total Skor NASA-TLX	<i>t</i>	<i>df</i>	<i>Sig One Sided P</i>	<i>Sig Two Sided P</i>	<i>Mean Diff</i>	<i>Std Error Diff</i>	<i>Lower</i>	<i>Upper</i>
<i>Equal variances assumed</i>	2.580	38	.007	.014	12.73300	4.93612	2.74034	22.72566
<i>Equal variances not assumed</i>	2.580	36.906	.007	.014	12.73300	4.93612	2.73061	22.73539

Hasil uji *Mann-Whitney U* Tabel 10 pada variabel TCT memiliki nilai *Asymp. Sig. (2-tailed)* sebesar 0.002. Karena  $0,002 < 0,05$ , maka terdapat pengaruh signifikan mode *switching* terhadap lamanya waktu penyelesaian tugas. Sedangkan untuk variabel *Error Rate* memiliki nilai *Asymp. Sig. (2-tailed)* sebesar 0,014. Karena  $0,014 < 0,05$ , maka terdapat pengaruh signifikan mode *switching* terhadap jumlah kesalahan yang dilakukan pengguna.

**Tabel 10.** Uji Hipotesis *Mann-Whitney U*

Kelompok Responden	<i>N</i>	<i>Mean Rank</i>
<i>Mann-Whitney U</i>	87.500	114.000

<i>Wilcoxon W</i>	297.500	324.000
<i>Z</i>	-3.045	-2.453
<i>Asymp. Sig. (2-tailed)</i>	0.002	0.014
<i>Exact Sig. [2*(1-tailed Sig.)]</i>	0.002	0.020

Hasil uji hipotesis ini menunjukkan bahwa mode switching berpengaruh signifikan ditunjukkan dengan nilai  $p < 0,05$  terhadap peningkatan *mental workload*, TCT, dan *Error Rate*, sehingga  $H_0$  ditolak dan  $H_1$  diterima. Hal ini membuktikan adanya hubungan kausal di mana ketidakkonsistenan visual secara nyata menurunkan performa dan kenyamanan mental pengguna. Pada analisis di tahap terakhir, akan mengukur seberapa besar dampak yang signifikan tersebut melalui uji effect size.

### 3.4 Hasil Analisis Effect Size

Analisis dari Tabel 11 dibawah ini memperlihatkan beban kognitif NASA-TLX memiliki nilai effect size 0.816 yang masuk dalam kategori Besar (*Large Effect*).

**Tabel 11.** Effect Size NASA-TLX

Total Skor NASA-TLX	Standardidzer	Point Estimate	Lower	Upper
<i>Cohen's d</i>	15.60939	0.816	0.165	1.457
<i>Hedges' correction</i>	15.92615	0.800	0.161	1.428
<i>Glass's delta</i>	16.89982	0.753	0.081	1.409

Sedangkan menurut Tabel 12 pada variabel TCT menunjukan nilai 0.48, *Error Rate* 0.39 yang keduanya berada dalam kategori Sedang (*Moderate Effect*). Hasil ini bisa diartikan bahwa intervensi mode *switching* secara efektif berpengaruh pada *mental load* yang dirasakan oleh responden, dan berpengaruh secara moderat pada area yang berkenaan dengan efisiensi dan akurasi, meskipun dari sisi statistik pengukuran ini cukup berarti. Klasifikasi ini sejalan dengan kriteria yang digunakan oleh [18] dalam uji analisis pengaruh perlakuan terhadap variabel dependen.

**Tabel 12.** Effect Size TCT dan Error Rate

Variabel	Z Value	N	$r = \frac{Z}{\sqrt{N}}$	R Value	Effect
TCT	-3.045	40	$r = \frac{3.045}{\sqrt{40}}$	0.48	Sedang – Besar
Error Rate	-2.453	40	$r = \frac{2.453}{\sqrt{40}}$	0.39	Sedang

### 3.5 Pembahasan

#### 3.5.1 Analisis Inkonsistensi Visual terhadap Performa Objektif (TCT dan Error Rate)

Penelitian ini secara empiris menunjukkan bahwa dampak negatif dari hambatan kognitif yang ditimbulkan oleh inkonsistensi visual akibat dari mode *switching* sangat berdampak bagi kinerja keseluruhan pengguna. Data eksperimen menunjukkan adanya perbedaan signifikan antara pengguna yang mengalami inkonsistensi visual akibat mode *switching* dengan pengguna tema statis. Pada *Task Completion Time*, kelompok kontrol dengan tema statis menyelesaikan tugas dengan lebih efisien dengan *mean* 8.30 detik. Sebaliknya, kelompok perlakuan memerlukan waktu lebih lama dengan *mean* 11.95 detik. Perbedaan waktu sebesar 3.65 detik merupakan hambatan kognitif yang disebabkan oleh inkonsistensi dalam desain antarmuka. Temuan ini berkorelasi dengan studi [20] yang menekankan bahwa kesuksesan platform digital tergantung pada navigasi yang efisien. Hasil dari *Error Rate*, juga menunjukkan adanya penurunan akurasi tugas yang disebabkan oleh beban mental yang tinggi. Ini secara khusus mendukung studi [19] yang menyatakan bahwa kurangnya desain antarmuka yang intuitif mengurangi kemungkinan keberhasilan pengguna dalam menyelesaikan tugas. Dengan demikian, ketidakkonsistenan visual membuat pengguna harus beradaptasi ulang dan meningkatkan risiko human error.

#### 3.5.2 Dampak Mode Switching terhadap Mental Workload

Analisis beban kognitif subjektif dengan instrumen NASA-TLX menunjukkan bahwa mode *switching* memberi tekanan psikologis yang jauh lebih besar dibandingkan kondisi statis. Dengan rata-rata WWL dari kelompok perlakuan adalah 55.23, sedangkan kelompok kontrol hanya 42.50. Selisih skor 12.73 ini menunjukkan adanya beban tambahan yang signifikan secara statistik. Temuan ini sesuai dengan metodologi dalam penelitian [18], yang menyatakan bahwa selisih skor besar menandakan pengaruh kuat variabel bebas terhadap aspek kognitif dan afektif pengguna. Data juga memperlihatkan lonjakan besar pada dimensi *Frustration* sebesar 28.0, diikuti *Physical Demand* sebesar 17.5 dan *Effort* sebesar 16.0. Hasil skor dimensi frustrasi yang tinggi menunjukkan tekanan emosional responden akibat perubahan visual yang terjadi secara tiba-tiba. Berdasarkan hasil tersebut kriteria ini sesuai pada penelitian [18] dengan hasil nilai *effect*

size 0.816, hasil ini termasuk ke dalam kategori *Large Effect*, sehingga dapat menegaskan bahwa mode *switching* sangat meningkatkan beban mental pengguna.

### 3.5.3 Analisis Korelasi

Secara keseluruhan, hasil data subjektif dan objektif dalam penelitian ini menunjukkan bahwa tampilan visual yang stabil sangat penting dalam desain antarmuka. Hasil penelitian menunjukkan adanya hubungan kausal, dalam hal ini peningkatan beban kognitif untuk kelompok perlakuan dengan nilai WWL 55.23, sejalan dengan penurunan efisiensi dan akurasi. Hasil ini selaras dengan penelitian [18] yang menunjukkan beban kerja mental yang lebih tinggi, akan berbanding terbalik dengan akurasi dalam mempertahankan informasi. Keterbatasan konsistensi visual akibat mode *switching* menuntut pengguna untuk beradaptasi dengan elemen yang secara visual mengurangi beban kerja. Ini juga berkontribusi kepada kelelahan visual, terbukti dari hasil wawancara singkat dengan responden, dan hasil eksperimen juga menunjukkan peningkatan waktu respon yang lebih lama, pada TCT dengan *mean* 3.65 detik dan menghasilkan *Error Rate* sebesar 0.70. Penurunan efisiensi navigasi akibat desain yang tidak konsisten juga mendukung hasil penelitian [20] yang menekankan pentingnya konsistensi visual untuk menjaga kegunaan sistem. Sebagai solusi, temuan kualitatif menyarankan penggunaan efek transisi bertahap atau *fading effect* yang halus untuk membantu adaptasi mata. Implementasi transisi yang tepat tidak hanya menjaga kenyamanan psikologis, tetapi juga penting untuk mempertahankan konsistensi visual sesuai prinsip *Golden Rules Strive for Consistency* dan mengurangi risiko kesalahan teknis sesuai prinsip *Golden Rules Reduce Short term Memory Load*. Hal ini dipertegas oleh [19] bahwa antarmuka yang tidak ramah pengguna dan sulit diakses akan secara langsung menurunkan tingkat keberhasilan pengguna dalam melaksanakan tugas. Dengan demikian, dapat disimpulkan bahwa penggunaan tema visual statis atau transisi yang terkelola dengan baik jauh lebih efektif dalam menjaga kestabilan kognitif dan performa kerja pengguna dibandingkan perubahan tema yang terjadi secara instan dan kontras.

## 4. KESIMPULAN

Hasil eksperimen mode *switching* terbukti meningkatkan beban kognitif dan penurunan performa kerja dari pengguna. Berdasarkan hasil evaluasi, kelompok yang menerapkan mode *switching* memiliki *mental workload* yang lebih tinggi jika dibandingkan dengan kelompok yang menerapkan mode statis. Seperti hasil instrumen NASA-TLX skor *WWL* yang didapat kelompok perlakuan memiliki rata-rata 55.23 sedangkan kelompok statis hanya 42.50. Hal ini juga diperkuat dengan hasil dari analisis *effect size* sebesar 0.816 yang masuk ke dalam kategori *Large Effect*. Dengan ini perubahan yang tidak konsisten secara visual akibat dari mode *switching* menjadi penyebab paling dominan timbulnya stress kognitif. Peningkatan beban mental didorong juga oleh peningkatan pada dimensi frustrasi dengan 28.0 poin, dan peningkatan pada dimensi *effort* sebesar 16.0 poin. Bersama dengan peningkatan beban mental, analisis kerja objektif juga mengungkapkan bahwa mode *switching* secara signifikan mengurangi efisiensi dan akurasi dari pengguna. Hal ini dibuktikan dari hasil TCT yang didapat oleh kelompok perlakuan dengan rata-rata perlambatan sebesar 3.65 detik dan terjadinya *Error Rate* pada kelompok perlakuan yang meningkat hingga dua kali lipat. Meskipun pengaruh pada efisiensi dan akurasi berada pada tingkat moderat, hasil ini tetap membuktikan bahwa inkonsistensi visual memaksa sistem kognitif pengguna untuk terus melakukan proses penyesuaian ulang, yang secara langsung hal tersebut mengganggu stabilitas fokus pada pengguna dan memicu terjadinya *human error*. Secara keseluruhan, penelitian ini menegaskan bahwa tema visual statis jauh lebih efektif dalam menjaga stabilitas kognitif, kenyamanan psikologis, dan performa kerja pengguna dibandingkan dengan penerapan mode *switching* yang tidak terkelola dengan baik. Terlepas dari hasil itu, penelitian ini juga memiliki keterbatasan pada jumlah responden dan juga konteks tugas yang spesifik, sehingga hasil penelitian perlu digeneralisasi dengan hati-hati pada populasi yang lebih luas. Pada penelitian selanjutnya disarankan untuk melibatkan lebih banyak partisipan dengan latar belakang yang lebih beragam juga dengan mengeksplorasi konteks penggunaan pada platform yang berbeda. Sebagai rekomendasi bagi pengembang antarmuka, sangat disarankan untuk mengimplementasikan fitur mode *switching* dengan memperhatikan aspek kenyamanan visual salah satunya dengan penerapan *fading effect* yang cepat dan halus. Dengan harapan lonjakan beban mental dan risiko kesalahan operasional dapat diminimalkan untuk tercapainya pengalaman pengguna yang lebih optimal.

## REFERENCES

- [1] A. D. W. Sumari, F. E. Perdana, D. Nugraheny, dan S. Lovrencic, "Improving the User Interface and Experience of a Student Portal Through the Eight Golden Rules," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 24, no. 3, hlm. 555–576, Jul 2025, doi: 10.30812/matrik.v24i3.4542.
- [2] M. Ahmad Faudzi, Z. Che Cob, R. Omar, S. A. Sharudin, dan M. Ghazali, "Investigating the User Interface Design Frameworks of Current Mobile Learning Applications: A Systematic Review," 1 Januari 2023, *MDPI*. doi: 10.3390/educsci13010094.
- [3] Đ. Damnjanović, D. Stojić, D. Vujičić, dan M. Milošević, "Minimalistic User Interface Design and Dark Mode Usage in Human-Computer Interaction," dalam *10th International Scientific Conference Technics, Informatic, and Education*, University of Kragujevac, Faculty of Technical Sciences, Čačak, 2024, hlm. 124–128. doi: 10.46793/TIE24.124D.
- [4] T. J. Ettl, D. Steinmann, K. Bektaş, dan A. Abbad-Andaloussi, "An Eye Tracking Study on the Effects of Dark and Light Themes on User Performance and Workload," dalam *Eye Tracking Research and Applications Symposium (ETRA)*, Association for Computing Machinery, Mei 2025. doi: 10.1145/3715669.3725879.

- [5] J. Virtanen, “Dark Mode Preferences: Exploring User Motivations in Interface Theme Selection,” 2023.
- [6] S. Andrew, C. Bishop, dan G. W. Tigwell, “Light and Dark Mode: A Comparison Between Android and iOS App UI Modes and Interviews with App Designers and Developers,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 8, no. 1, Mar 2024, doi: 10.1145/3643539.
- [7] Lappeenranta-, “USABILITY CHALLENGES OF DARK MODE IN EDUCATIONAL APPS: A USER-CENTERED DESIGN FRAMEWORK,” 2025.
- [8] E. Babaci, T. Dingler, B. Tag, dan E. Velloso, “Should we use the NASA-TLX in HCI? A review of theoretical and methodological issues around Mental Workload Measurement,” *International Journal of Human Computer Studies*, vol. 201, Jun 2025, doi: 10.1016/j.ijhcs.2025.103515.
- [9] T. Gazit, T. Tager-Shafir, H. X. Zhong, P. C. K. Hung, dan 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.
- [10] F. Alshehri, M. Freeman, dan A. E. Freeman, “Are you smart enough for your smart phone? A cognitive load comparison Are you smart enough for your smart phone? A cognitive load comparison Recommended Citation Recommended Citation,” 2013. [Daring]. Tersedia pada: <https://ro.uow.edu.au/eispapershttps://ro.uow.edu.au/eispapers/2354>
- [11] M. Gardesevic dkk., “Brighter Time: A Smartphone App Recording Cognitive Task Performance and Illuminance in Everyday Life,” *Clocks Sleep*, vol. 4, no. 4, hlm. 577–594, Des 2022, doi: 10.3390/clockssleep4040045.
- [12] A. Skulmowski, “When color coding backfires: A guidance reversal effect when learning with realistic visualizations,” *Educ. Inf. Technol. (Dordr)*, vol. 27, no. 4, hlm. 4621–4636, Mei 2022, doi: 10.1007/s10639-021-10796-6.
- [13] L. Yang, B. Qi, dan Q. Guo, “The Effect of Icon Color Combinations in Information Interfaces on Task Performance under Varying Levels of Cognitive Load,” *Applied Sciences (Switzerland)*, vol. 14, no. 10, Mei 2024, doi: 10.3390/app14104212.
- [14] M. Sandesara dkk., “Design and Experience of Mobile Applications: A Pilot Survey,” 1 Juli 2022, *MDPI*. doi: 10.3390/math10142380.
- [15] L. Nurtamara, S. Amintarti, A. Ajizah, D. A. Widiyastuti, N. Noorhidayati, dan A. Rezeki, “The argumentation of the students about plant biodiversity: Quasi-experimental research,” *BIO-INOVED : Jurnal Biologi-Inovasi Pendidikan*, vol. 7, no. 2, hlm. 202, Jun 2025, doi: 10.20527/bino.v7i2.22369.
- [16] Z. Shen dkk., “Human-computer interaction interface design of flight simulator based on situation awareness,” *Sci. Rep.*, vol. 14, no. 1, Des 2024, doi: 10.1038/s41598-024-78043-9.
- [17] J. W. Hutson, A. E. Franklin, B. A. Rogers, dan D. Walker, “Psychometric Testing of NASA-TLX to Measure Learners’ Cognitive Load in Individual and Group Nursing Simulations,” *Clin. Simul. Nurs.*, vol. 95, Okt 2024, doi: 10.1016/j.ecns.2024.101607.
- [18] Chairatul Umamah, Nurul Hasanah, dan S. Suprianto, “Analisis Effect Size Penggunaan E-magazine Berbasis Inkuiri Terbimbing terhadap Hasil Belajar Siswa,” *JURNAL PENDIDIKAN MIPA*, vol. 14, no. 4, hlm. 1042–1048, Des 2024, doi: 10.37630/jpm.v14i4.1982.
- [19] W. T. Pambudi dan A. C. Wardhana, “Design Development of the JekNyong Application Using the Design Thinking Method,” *JURIKOM (Jurnal Riset Komputer)*, vol. 12, no. 3, hlm. 149–160, Jun 2025, doi: 10.30865/jurikom.v12i3.8539.
- [20] M. Widya, T. Cahyatama, dan A. Cahya Wardhana, “Redesign Of BUMDes And MSMEs Marketplace Platform Using User Centered Design Method,” *Jurnal Riset Komputer*, vol. 12, no. 6, hlm. 2407–389, 2025, doi: 10.30865/jurikom.v12i6.8829.