International Journal of Advances in Data and Information Systems
Vol. 7, No. 1, April 2026, pp. 174~187
ISSN: 2721-3056, DOI: 10.59395/ijadis.v7i1.1499

Enhancing E-Insurance Mobile Applications through Cognitive Walkthrough and System Usability Scale (SUS): A Case Study of My Taspen Life

Nur Diarto[1], Harry Budi Santoso[2]
[1,2]Department of Information Technology, Faculty of Computer Science, Universitas Indonesia, Indonesia

Article Info
Article history:
Received Dec 24, 2025
Revised Jan 05, 2026
Accepted Jan 20, 2026

Keywords:
Mobile Application
User Interface (UI)
User-Centered Design
Cognitive Walkthrough
System Usability Scale (SUS)

ABSTRACT

The digital transformation of the insurance sector has necessitated mobile services with superior usability. The My Taspen Life mobile application has experienced low user satisfaction attributed to its suboptimal user interface (UI) and user experience (UX). This study aimed to evaluate the application's usability and propose a validated high-fidelity prototype employing a User-Centered Design (UCD) approach. A mixed-methods evaluation comprised a System Usability Scale (SUS) survey, a Cognitive Walkthrough with experts, and Usability Testing. The identified issues were mapped against Shneiderman's Eight Golden Rules of Interface Design to develop an adequate solution. The initial results showed an average SUS score of 65.51, below the global SUS average, and identified critical constraints in layout, navigation, and visual representation. Following the redesign, the prototype was re-evaluated by 12 participants, equally distributed between employees and external users. The final evaluation demonstrated a significant improvement, achieving a SUS score of 83.54 (Grade A, "Excellent"). Task scenario testing further confirmed that the new interface significantly reduced user dissatisfaction. We conclude that implementing UCD-based intervention can address significant usability problems in e-insurance applications.

This is an open access article under the CC BY-SA license.

Corresponding Author:
Nur Diarto,
Department of Information Technology, Faculty of Computer Science,
Universitas Indonesia,
Kampus UI Depok, Gedung Fakultas Ilmu Komputer, Depok West Java 16424, Indonesia.
Email: nur.diarto@ui.ac.id

1. INTRODUCTION

In the contemporary insurance landscape, digital transformation has evolved from a peripheral trend into a strategic tool for operational sustainability and market survival [1], [2], [3]. Mobile applications, serving as the primary interface for service delivery, have transcended mere digital brand representations to become critical determinants of corporate identity and primary drivers of competition [4]. The literature confirms a direct correlation between platform usability and customer metrics; a quality user interface is a prerequisite for fostering customer loyalty [5], whereas a suboptimal user experience can rapidly erode brand reputation and impede market penetration [6].

In Indonesia, PT Asuransi Jiwa Taspen introduced "My Taspen Life" to facilitate access to its services. However, the application received suboptimal ratings of 3.1 on Google Play Store and 2.6 on Apple App Store. Several issues, ranging from technical instability to unintuitive navigation,

highlighted a fundamental failure in system quality. As defined by Aldholay et al. [7] and Nielsen [8], system quality is a multidimensional construct encompassing both reliability and ease of use. Neglecting these dimensions creates a "usability debt" that hinders the organization's objective of providing a convenient digital platform. Consequently, a comprehensive usability evaluation of My Taspen Life is not merely a technical requirement but a strategic imperative to enhance service quality, strengthen brand positioning, and sustain competitiveness in the modern insurance industry.

To address these issues, the literature advocates the use of robust frameworks. Handasari et al. [9], Sefira et al. [10] and Rinaldi et al. [11], demonstrated the efficacy of the User-Centered Design (UCD) and System Usability Scale (SUS) in improving public apps. However, these evaluation metrics often rely on general user surveys, lacking a multi-perspective diagnostic approach that distinguishes between domain experts and the lay public [12]. This study aims to fill this methodological gap by integrating Cognitive Walkthrough (CW) for expert inspection with a dual-persona Usability Testing strategy, thereby distinguishing between internal employees (domain experts) and external users (the public) to ensure comprehensive evaluation.

The present study addresses the identified knowledge gap by employing a comprehensive UCD framework to answer the following questions:

1) What were the results of the initial interface evaluation of the "My Taspen Life" mobile application?
2) What design changes could be made to improve the interface of the "My Taspen Life" mobile application?
3) What are the results of an evaluation of the redesigned interface of the "My Taspen Life" mobile application?

To answer these questions, we applied a mixed-method UCD approach involving Usability Testing (UT) with a dual-persona strategy, Cognitive Walkthrough (CW), SUS measurement, and heuristic evaluation. The findings are expected to support system improvement at PT Asuransi Jiwa Taspen by applying Shneiderman's Eight Golden Rules [13]. The results will contribute to academic discourse on Human–Computer Interaction (HCI), specifically regarding the optimization of e-insurance platforms.

2. RESEARCH METHOD

We employed an exploratory mixed-methods approach based on a UCD framework to integrate the processes of collecting, analyzing, and interpreting quantitative and qualitative data. This design prioritizes qualitative inquiry to address complex topics, which is subsequently validated by quantitative data [14], [15], [16]. Additionally, we performed a case study to deeply examine specific user experience and usability issues, thereby ensuring that the outcomes stemming from situational analysis can generate practical suggestions for enhancement [17], [18], [19]. Quantitative data were gathered via the System Usability Scale (SUS) to measure perceived performance, while qualitative insights were derived from Cognitive Walkthroughs (CW), Usability Testing (UT), and open-ended questions. An evaluative research approach was employed to diagnose the root causes of consumer dissatisfaction and assess the quality of the "My Taspen Life" interface through expert inspection and task simulation. Finally, a purposive sampling technique is employed to select research participants based on relevant criteria that distinguish between internal domain experts and external public users [16], [17].

Table 1. Research Design

Attributes	Descriptions
Classifications	Mixed-method research
Category	Case Study
Paradigm	Evaluative
Purpose	Conducting a usability evaluation of the 'My Taspen Life' mobile application and providing recommendations for improvements in interface design.
Results	Design Recommendations for Improving the My Taspen Life Interface

Attributes	Descriptions
Data Collection	Cognitive walkthrough, usability testing, SUS, and Open-ended Questions
Sampling strategy	Purposive sampling

The research trajectory began with a rigorous problem-identification phase, drawing on data from internal organizational documents, app-review analysis, and stakeholder interviews at PT Asuransi Jiwa Taspen. An Ishikawa diagram was employed as a diagnostic tool to pinpoint the underlying causes of the usability deficiency[20]. Subsequently, a comprehensive literature review was conducted to obtain a solid theoretical framework grounded in usability evaluation and interface design protocols. A suite of research instruments was developed, guided by the principles of UCD, comprising task analysis protocols, Cognitive Walkthrough checklists, SUS, and open-ended questions. The gathered issues aligned with the heuristics principles. Finally, the results of the analysis were mapped to Shneiderman's Eight Golden Rules of Interface Design that served as the foundation for improving the user interface.

The goal of interface design is to devise a bridge between users and systems so that the application can run efficiently and satisfactorily [21]. A high-fidelity prototype is initially created for realistic interactions, ensuring that the proposed design recommendations are testable [22]. The final phase includes a thorough evaluation of this prototype through usability testing and re-measurement using SUS and open-ended questions to validate the improvements. The research method is visually summarized in Figure 1.

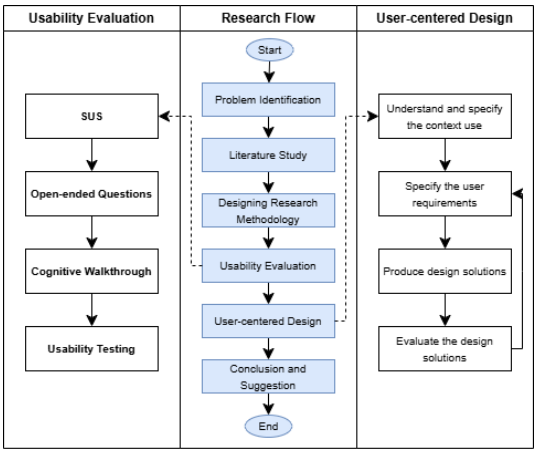


Figure 1. The Research Flow

Data for this research were gathered in two stages. The initial stage focused on diagnosis, involving distributing SUS and open-ended questions to obtain preliminary estimates of user satisfaction. Simultaneously, experts conducted a Cognitive Walkthrough to predict potential interaction problems in specific task scenarios. Following this, Usability Testing was conducted with representative users to empirically validate the issues indicated by the experts [23]. The following stage was dedicated to evaluating the redesigned prototype, using the same instruments to assess differences between usability metrics of the current system and the prototype.

The data processing methods comprised quantitative and qualitative approaches; quantitative data from the SUS were used to calculate the overall usability score. Figure 2 illustrates the interpretation of SUS scores in relation to user satisfaction levels [24].

Grade	SUS score	Percentile Range	Adjective	Acceptability
A+	84.1-100	96-100	Best Imaginable	Acceptable
A	80.8-84	90-95	Excellent	Acceptable
A-	78.9-80.7	85-89		Acceptable
B+	77.2-78.8	80-84		Acceptable
B	74.1-77.1	70-79		Acceptable
B-	72.6-74	65-69		Acceptable
C+	71.1-72.5	60-64	Good	Acceptable
C	65-71	41-59		Marginal
C-	62.7-64.9	35-40		Marginal
D	51.7-62.6	15-34	OK	Marginal
F	25.1-51.6	2-14	Poor	Not Acceptable
F	0-25	0-1.9	Worst Imaginable	Not Acceptable

Figure 2. Interpretation of SUS scores

Meanwhile, a thematic extraction technique was applied to the data from usability tests, cognitive walkthroughs, and open-ended questions. The objective was to systematically categorize prevalent user sentiments [25]. These identified themes served as a foundation for developing targeted design interventions aimed at mitigating specific areas of user frustration [26]. To mitigate subjectivity and ensure methodological rigor, this thematic analysis employed a three-rater approach, where independent experts with relevant domain knowledge validated the findings. This triangulation strategy significantly strengthens the reliability and validity of the analytical results [27]. Instead of simple coding, identified issues are mapped against Nielsen's Usability Heuristics to diagnose specific universal interface violations. These heuristic violations are then translated into design requirements using Shneiderman's principles to ensure the formulated recommendations are theoretically grounded and practical.

3. RESULTS AND DISCUSSION

This study employed a rigorous two-stage data collection process designed to capture both broad user sentiment and specific difficulties encountered in interaction with the program. The initial phase involved distributing online usability questionnaires using two complementary instruments: the System Usability Scale (SUS) and Open-Ended Questions. This phase aimed to establish quantitative baselines for user experience and satisfaction levels while operating the "My Taspen Life" application. A total of 49 respondents participated in this initial survey, providing the primary dataset for the preliminary diagnosis.

The subsequent phase focused on in-depth validation through synchronous usability testing. Using predefined task scenarios, we applied a purposive sampling technique to recruit participants from the initial survey pool who expressed a willingness to engage further. To ensure a balanced perspective, the participants comprised 50% internal users (employees) and 50% external users (lay public). This stratification ensured that the evaluation captured both the domain expertise of internal staff and the fresh perspectives of lay users.

3.1. Participants

The questionnaire was distributed openly to both internal and external groups. Based on the results, 49 respondents completed the usability questionnaire, comprising 38 internal and 11 external participants. Furthermore, 12 individuals were selected for usability testing based on their willingness to participate. This group, consisting of 6 internal and 6 external participants, subsequently took part in the current system evaluation and prototype evaluation phase.

3.2. Current System Evaluation

The evaluation of the current system interface served as the critical diagnostic baseline for this study. The analysis uncovered the specific barriers hindering user adoption by integrating quantitative metrics from the SUS with qualitative insights derived from open-ended questions, cognitive walkthroughs, and dual-persona usability testing. The analysis aimed to measure current satisfaction levels and identify issues using standard usability heuristics, thereby establishing empirical grounds for the subsequent design interventions guided by Shneiderman's principles.

First, the SUS comprising 10 standard items was administered to 49 respondents, following the adaptation guidelines [28]. The quantitative analysis yielded an overall average score of 65.51. When benchmarked against the industrial standard, this value was below the global average score of 68 [29]. The score indicated the My Taspen Life application's usability as average, suggesting that improvements could be made to achieve higher levels of system efficiency and enhance the usability of the My Taspen Life application. The detailed breakdown of the SUS results is shown in Table 2.

Table 2. SUS calculation results: current System

Respondents	Calculation result per question										Final Scores
	1	2	3	4	5	6	7	8	9	10	(Total*2,5)
QRC-01	3	3	2	2	2	1	0	2	3	0	45
QRC-02	4	4	4	4	4	4	4	4	0	0	80
QRC-03	4	4	4	4	4	4	4	4	4	2	95
QRC-04	2	3	3	3	3	3	3	3	3	1	67.5
QRC-05	3	3	3	3	3	3	3	3	3	2	72.5
QRC-06	4	4	4	4	4	4	4	4	4	0	90
QRC-07	4	4	4	4	4	4	4	4	4	4	100
QRC-08	2	2	2	3	3	2	2	1	3	2	55
QRC-09	3	3	3	3	1	1	3	1	3	1	55
QRC-10	3	3	3	3	3	1	1	1	3	1	55
QRC-11	2	1	2	3	1	3	1	3	3	1	50
QRC-12	1	1	1	3	1	3	1	1	1	0	32.5
QRC-13	4	4	4	3	4	3	4	4	4	4	95
QRC-14	3	2	2	4	3	2	2	2	3	2	62.5
QRC-15	2	2	2	2	2	2	2	2	2	2	50
QRC-16	3	3	3	3	3	3	3	3	3	1	70
QRC-17	2	2	3	3	2	0	1	1	2	2	45
QRC-18	1	0	1	1	2	1	1	1	1	0	22.5
QRC-19	1	3	1	1	2	2	1	2	2	2	42.5
QRC-20	3	3	3	3	3	3	3	3	3	3	75
QRC-21	4	4	3	2	3	3	3	4	3	2	77.5
QRC-22	2	2	2	2	2	2	2	2	2	2	50
QRC-23	3	3	3	4	3	2	3	3	3	3	75
QRC-24	4	4	4	4	4	4	4	4	4	0	90
QRC-25	3	3	3	3	3	3	3	3	2	0	65
QRC-26	4	1	4	4	4	4	4	4	4	0	82.5
QRC-27	4	4	4	4	4	4	4	4	4	4	100
QRC-28	3	3	3	1	2	2	3	3	1	1	55
QRC-29	4	2	3	2	3	2	3	2	3	2	65
QRC-30	3	3	3	3	3	2	3	3	3	1	67.5
QRC-31	4	3	4	1	3	3	1	3	2	1	62.5
QRC-32	3	3	3	3	3	3	3	4	1	3	72.5
QRC-33	3	3	3	3	3	3	3	3	3	1	70
QRC-34	3	1	3	1	3	3	3	4	3	1	62.5
QRC-35	4	1	3	1	3	2	4	2	3	3	65
QRC-36	3	3	3	4	1	1	3	3	3	4	70
QRC-37	2	2	2	1	2	1	1	2	2	2	42.5
QRC-38	2	3	3	3	3	2	3	3	3	3	70
QRC-39	4	4	4	4	4	4	4	4	4	4	100
QRC-40	4	1	3	4	4	4	4	4	4	0	80
QRC-41	2	1	2	1	3	3	2	2	3	1	50
QRC-42	2	0	2	2	3	1	2	2	3	1	45
QRC-43	3	1	1	1	1	1	1	1	1	1	30
QRC-44	3	3	3	1	3	1	1	3	3	1	55
QRC-45	4	4	4	3	4	4	4	4	4	4	97.5
QRC-46	3	3	3	1	3	3	3	1	3	1	60
QRC-47	3	2	2	1	2	2	2	3	2	2	52.5
QRC-48	4	4	4	3	3	3	3	3	3	1	77.5
QRC-49	3	3	3	4	1	0	3	3	3	1	60
										Average Score	65,51

Second, open-ended questions served as the primary source of qualitative data, focusing on users' frustrations and recommendations to enhance the usability of the My Taspen Life application. The participants identified interface complexity and usability hurdles as their primary concerns, suggesting that process simplification and UI upgrades were necessary. The process was conducted with the support of three raters. The thematic analysis yielded 19 issues categorized by usability guidelines; these are listed in Table 3.

Table 3. Usability Issues from Open-Ended Questions: current system

Category	Issue Description	Code
Content organization	Report data content is inconsistent or invalid.	IOC-01
	Critical information is difficult to locate, incomplete, or obscured by images.	IOC-02
Navigation	Convoluted menu structure hinders feature discovery.	IOC-03
	Difficulty in locating specific menus or information.	IOC-04
Layout	Data tables are too long and illegible on mobile phone screens.	IOC-05
	Poor text arrangement negatively affects readability.	IOC-06
	Disorganized homepage layout.	IOC-07
Visual representation	The application interface appears rigid, monotonous, and unappealing.	IOC-08
	The interface is cluttered with excessive, distracting images.	IOC-09
	The visual design lacks professional credibility, which reduces perceived security.	IOC-10
Consistency and standards	Design feels outdated compared to modern application standards.	IOC-11
Help and Feedback	Frequent system errors and functional failures.	IOC-12
	Delayed data synchronization results in inaccurate information presentation.	IOC-13
Interaction	Application workflow is confusing and unintuitive.	IOC-14
	Password complexity rules are burdensome; there are few quick login features.	IOC-15
	Performance is sluggish and heavily dependent on high-speed connectivity.	IOC-16
Customization	Absence of data filtering mechanisms for searching.	IOC-17
Learning	Steep learning curve for new users requiring significant adaptation time.	IOC-18
Accessibility	Access difficulties and burdensome account security protocols.	IOC-19

In addition, the diagnostic evaluation employed the Cognitive Walkthrough (CW) method, conducted in three distinct stages: preparation, execution, and problem analysis. In the preparation phase, a comprehensive set of task-based scenarios was developed and rigorously evaluated to ensure relevance to the application's functions and alignment with specific user roles. The design of these scenarios centered on core features related to the insurance essence within the My Taspen Life application architecture. Subsequently, during the execution phase, an assessment was conducted by three independent evaluators with specialized expertise: an internal Business Analyst from My Taspen Life, a professional Web Developer, and a UI/UX Designer. All three assessed potential usability issues based on the defined scenarios and documented their findings using a standardized evaluation checklist [30]. The analysis identified 13 potential usability issues (Table 4).

Table 4. Potential Usability Issues from CW

Category	Issue Description	Code
Consistency and Standards	Placing the back button at the bottom of the page is inconsistent with standard applications (which are usually top-aligned), affecting its reachability.	ICC-01
Content Organization	Placing "Payment Information" (transaction) under the "Profile" (personal data) menu is unrelated and confusing to users.	ICC-02
Help and Feedback	Bank selection options appear as large cards obscuring the information below, with no visual feedback after selection.	ICC-03
Interaction	An excessive number of actions (clicks/steps) is required to reach Claim History information.	ICC-04
Layout	The absence of a search function, extensive lists, and oversized product visuals necessitate excessive scrolling.	ICC-05
	Tab menu display is truncated (invisible) on devices with specific resolutions.	ICC-06
Layout and Help & Feedback	The account Statement table is too long, pushing the back button too far down; the print button is unresponsive.	ICC-07
Navigation	Absence of a Back button to switch products traps users if an incorrect selection is made.	ICC-08
	Redundant navigation (duplicate menus) in different locations creates inefficiency and confusion.	ICC-09
	Lack of download or return buttons forces users to close the application to exit the information page.	ICC-10
Selection Based	The 'Register/Login' button design is confusing because a dual-function split-button within a single visual, causing potential click errors (slips).	ICC-11
Visual Representation	The button design fails to provide clear signifiers that the element is clickable, leading users to overlook it.	ICC-12
	Input forms lack visual distinction, making it hard to tell which fields are dropdowns and which are free-text.	ICC-13

The final diagnostic phase comprised Usability Testing, conducted using 13 specific task scenarios designed to cover the application's core insurance features(Table 5). A detailed distribution of these scenario completion processes is presented in Table 6. During the testing sessions, a think-aloud protocol was employed to capture real-time feedback, and post-task

interviews were conducted to confirm observed difficulties. The identified problems ranged from inefficient navigation to confusing layout and interface elements, factors that significantly affected task completion and reduced overall user satisfaction.

Table 5. Task Scenario from CW

No	Task Description	Code
1	Logging into the system	T1
2	Viewing product catalog	T2
3	Checking product simulations	T3
4	Purchasing a product	T4
5	Viewing Life Insurance Application (SPAJ) information	T5
6	Viewing policy information	T6
7	Viewing policy benefits	T7
8	Viewing account statement information	T8
9	Viewing Pension Fund (DPLK) statement	T9
10	Viewing payment information	T10
11	Viewing claim history	T11
12	Adding a payment method	T12
13	Updating claim bank account details	T13

Table 6. Problems mapping matrix: current system

Respondents	Completion Task												
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13
QRC-01													
QRC-02													
QRC-03													
QRC-04													
QRC-05													
QRC-06													
QRC-07													
QRC-08													
QRC-09													
QRC-10													
QRC-11													
QRC-12													
Problems Frequency	3	4	7	8	12	1	1	12	0	3	4	0	2
Percentage	25	33.3	58.3	66.7	100	8.3	8.3	100	0	25	33.3	0	16.7

: Successful task execution
: Tasks were completed with noted issues
: Task failed to execute
: Task is not executed

The usability issues identified were categorized into four severity levels based on an established framework, reflecting their varying degrees of impact on the user experience [29]. Level 1 (Low Frequency), defined by a failure rate of ≤10%, was observed in scenarios T6 and T7 (8.3%). Level 2 (Medium Frequency), ranging from 11% to 50%, encompassed a broader set of issues occurring in T1, T2, T10, T11, and T13, indicating that these features posed difficulties for a significant subset of respondents. Level 3 (High Frequency), with rates between 51% and 89%, highlighted significant hurdles in T3 (58.3%) and T4 (66.7%), indicating that most users struggled with these tasks. Level 4 (Very High Frequency), characterized by failure rates ≥90%, identified two critical points in T5 and T8, where the rates reached 100%, indicating that every respondent was unable to complete these tasks without assistance. Scenario T9 could not be executed due to participant limitations regarding specific program enrollment. This classification(Table 7) was instrumental in prioritizing immediate design interventions for the critical usability issues.

Table 7. Classification summary of problems: current system

Level	Classification	Frequency Range	Associated Tasks
1	Low	≤10%	T6, T7
2	Middle	11%–50%	T1, T2, T10, T11, T13
3	High	51%–89%	T3, T4
4	Very High	≥90%	T5, T8

To conclude the diagnostic phase, the data from all evaluation instruments were synthesized to provide a holistic view of the application's usability landscape. By combining the quantitative baseline derived from SUS scores with the qualitative depth of open-ended questions, cognitive walkthroughs, and usability testing, this study identified a consistent pattern of interface failures. Table 8 summarizes the key findings from each evaluation method, providing a comprehensive empirical foundation to guide the subsequent interface redesign in accordance with Shneiderman's principles.Table 8. Overall evaluation results summary: current system

No.	Methods	Indicators	Results	Summary
1	SUS	Average SUS Score	Achieved an average score of 65.51, which is below the global industrial benchmark of 68	The "My Taspen Life" application demonstrates suboptimal interface usability across multiple critical dimensions. Consequently, significant improvements are required, specifically in the interface design to address these deficiencies and elevate the system to an acceptable usability standard.
2	Open-ended question	Feedback themes	Identified 19 specific issues, primarily centering on interface complexity, rigid visual representation, and confusing navigation flow.	
3	CW	CW evaluation checklist	Detected 13 potential usability violations, highlighting inconsistencies in button design, layout truncation on mobile devices, and a lack of system feedback.	
4	Usability testing	Task completion level	Uncovered 24 operational issues. Critical failures (Level 4 severity) were found in task tracking and specific interactions, with 100% failure rates in two key scenarios.	

3.3. Designing Prototypes

Following the diagnostic evaluation, 56 usability issues were aggregated and rigorously filtered, with 19 technical constraints excluded to focus on actionable improvements to the user interface. The remaining valid issues were consolidated and analyzed using a heuristic framework, revealing critical violations primarily in Consistency and Standards (e.g., inconsistent button designs), Aesthetic and Minimalist Design (visual clutter), and Match Between Real World and System (illogical hierarchy). Additional deficiencies in Flexibility and Efficiency of Use, and in User Control and Freedom were also noted, serving as the direct empirical requirements for the interface remediation process.

Based on this consolidated list of heuristic issues, further analysis was undertaken to develop precise recommendations grounded in Shneiderman's Eight Golden Rules of Interface Design. A supplementary list of recommended improvements was formulated to complement these theoretical solutions (Table 9), thereby ensuring that the proposed solutions were both theoretically robust and grounded in users' actual needs and experiences.

Table 9. Design recommendations based on Shneiderman's Rules: current system

Rec. Code	Recommendation Description	Shneiderman's Rule
KP-01	Ensure consistent visual signifiers for interactive elements.	Strive for consistency
KP-02	Standardize terminology application-wide.	Strive for consistency
KP-03	Modernize visuals for responsiveness and clarity.	Reduce short-term memory load
KP-04	Add navigation buttons (back/cancel) for user control.	Permit easy reversal of actions
KP-05	Reorganize menu structure to align with user expectations.	Support internal locus of control
KP-06	Remove duplicate menus to simplify navigation.	Reduce short-term memory load
KP-07	Optimize data tables with pagination.	Reduce short-term memory load
KP-08	Standardize header navigation placement.	Strive for consistency
KP-09	Implement visual feedback for system interactions.	Offer informative feedback
KP-10	Implement search and filtering mechanisms.	Enable frequent users to use shortcuts
KP-11	Streamline workflows by reducing steps.	Enable frequent users to use shortcuts
KP-12	Provide onboarding guides for new users.	Offer informative feedback
KP-13	Improve error messages and button responsiveness.	Offer simple error handling

To visually demonstrate the practical application of the recommendations, Figure 3 presents a comparative transformation of key interface screens, specifically implementing recommendations KP-03, KP-05, KP-06, and KP-11. These improvement recommendations were translated into low-fidelity formats, which were subsequently transformed into high-fidelity designs, marking an evolution from the initial problematic design to the optimized high-fidelity prototype.

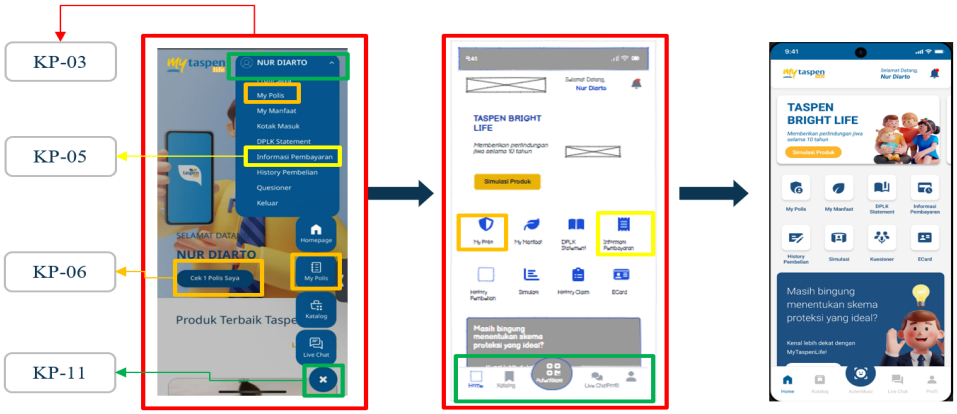


Figure 3. UI Design Transformation

3.4. Prototype Evaluation

An analysis of the high-fidelity prototype was performed to determine the effectiveness of the refined My Taspen Life application. This second assessment phase mirrored the methodology of the initial baseline study, using a tripartite approach: task-based scenarios, SUS, and open-ended qualitative feedback. Maintaining the same procedural mechanism as in the current system evaluation ensured a consistent comparison between the existing conditions and the improved prototype. Thirteen scenarios were used to cover the core features of the insurance essence. A detailed distribution of the respondents' mapping is illustrated in Table 10.

Table 10. Problems mapping matrix: prototype

Respondents	Completion Task												
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13
QRC-01													
QRC-02													
QRC-03													
QRC-04													
QRC-05													
QRC-06													
QRC-07													
QRC-08													
QRC-09													
QRC-10													
QRC-11													
QRC-12													
Problems Frequency	0	2	0	4	0	1	0	0	0	1	1	0	2
Percentage	0	16.67	0	33.33	0	8.333	0	0	0	8.333	8.333	0	16.67

: Successful task execution
: Tasks were completed with noted issues
: Task failed to execute
: Task is not executed

The results of prototype testing indicated that the most of challenges identified were of low to medium severity. Nevertheless, specific minor issues persisted and required further refinement to ensure optimal user support. A comprehensive summary of these findings is presented in Table 11.

Table 11. Classification summary of problems: prototype

Level	Classification	Frequency Range	Associated Tasks
1	Low	≤10%	T1, T3, T5, T6, T7, T8, T10, T11, T12
2	Middle	11%–50%	T2, T4, T13
3	High	51%–89%	
4	Very High	≥90%	

Following the qualitative assessment, the SUS was used to re-evaluate the quantitative usability performance. Based on responses from all participants to the 10 standard items, the SUS

score for the high-fidelity prototype was 83.54. According to Sauro and Lewis's grading scale, this score corresponds to a Grade A and is classified as Excellent. This substantial increase from the initial baseline of 65.51 demonstrates that the user-centered design interventions have notably improved the application's usability compared to the initial state. The detailed breakdown of the final SUS results is shown in Table 12.

Table 12. SUS calculation results: prototype

Respondents	Calculation result per question										Final Scores
	1	2	3	4	5	6	7	8	9	10	(Total*2,5)
QRC-01	3	3	3	3	4	4	3	4	4	3	85
QRC-02	3	4	3	4	4	4	3	4	4	4	92.5
QRC-03	3	3	4	3	3	3	3	4	4	2	80
QRC-04	4	3	3	3	4	3	3	3	3	3	80
QRC-05	3	4	3	1	3	3	2	2	3	0	60
QRC-06	4	3	3	2	4	3	3	3	3	2	75
QRC-07	3	3	3	3	3	3	3	3	3	3	75
QRC-08	3	3	3	3	3	3	3	3	3	3	75
QRC-09	4	4	4	4	4	4	4	4	3	4	97.5
QRC-10	4	4	3	4	4	4	3	4	4	4	95
QRC-11	4	3	4	3	4	3	4	4	4	3	90
QRC-12	4	4	4	4	4	4	4	4	4	3	97.5
										Average Score	83.54

To provide a detailed view of the usability improvements at the individual level, a direct comparison of the SUS scores for each of the 12 participants is presented in Table 13. These participants, representing a subset of the initial 49 respondents, evaluated both the current system and the prototype.

Table 13. SUS comparison

Respondents	Current system score	Prototype score	Difference
QRC-01	55	85	30
QRC-02	55	92.5	37.5
QRC-03	50	80	30
QRC-04	32.5	80	47.5
QRC-05	50	60	10
QRC-06	45	75	30
QRC-07	22.5	75	52.5
QRC-08	55	75	20
QRC-09	60	97.5	37.5
QRC-10	52.5	95	42.5
QRC-11	77.5	90	12.5
QRC-12	60	97.5	37.5
Average	51.25	83.54	32.29

Table 14. Classification summary of problems: prototype

System Condition	Respondents(N)	Mean(M)	Std. Deviation (SD)
Current System	12	51.25	13.56
Prototype	12	83.54	11.45

A paired-samples t-test was conducted to evaluate the impact of the prototype on usability scores among 12 participants who tested both systems (Table 14). The results showed a statistically significant increase in SUS scores from the current system ($M = 51.25$, $SD = 13.56$) to the developed prototype ($M = 83.54$, $SD = 11.45$), with $t = 8.555$, $p < 0.001$. These findings provide strong evidence that the prototype significantly improves the user usability compared to the current system.

Complementing these metrics, qualitative analysis of open-ended questions revealed a significant reduction in negative sentiment. The feedback indicated a marked shift in user perception, with nearly all respondents expressing satisfaction with the improved efficiency and visual clarity, thereby confirming the redesign's effectiveness. A thematic summary of these responses is provided in Table 15.

Table 15. Usability Issues from Open-Ended Questions: prototype

Category	Issue Description	Code
Layout	Users perceive the font size as too small, affecting readability.	IOP-01
Visual Representation	Icon and menu colors lack visual appeal and are not sufficiently eye-catching.	IOP-02
Customization	Absence of auto-fill functionality necessitates manual data entry in forms.	IOP-03

In the spirit of continuous improvement inherent to the User-Centered Design framework, a final remediation phase was undertaken to address several minor issues identified in the prototype evaluation. Focusing on the feedback detailed in Table 16, design adjustments were targeted at increasing the base font size to enhance legibility across devices (Layout) and refining the color saturation of icons to improve visual hierarchy and appeal (Visual Representation). The feedback regarding auto-fill functionality (Customization) was acknowledged but excluded from this phase, as it would require developing a new feature rather than an interface adjustment. Consequently, these two actionable issues regarding layout and visual representation were mapped against Nielsen's heuristic principles to guide the final design refinement (Table 16). These targeted refinements eliminated key sources of user friction, resulting in a polished final interface that is fully aligned with user needs and ready for implementation.

Table 16. Problem mapping based on Nielsen's heuristic principles

Issue Code	Issue Description	Usability Heuristic Principle
KMP-01	The color selection for inactive icons is too pale, resulting in poor visibility for users.	Aesthetic and Minimalist Design
KMP-02	Users perceive the font size as too small, compromising visual clarity.	Aesthetic and Minimalist Design

The final design recommendations were formulated by strictly applying Shneiderman's Eight Golden Rules. These targeted improvements, aimed at optimizing visual hierarchy and legibility, are outlined in Table 17.

Table 17. Design recommendations based on Shneiderman's Rules: prototype

Rec. Code	Recommendation Description	Shneiderman's Rule
KPP-01	Adjust inactive icon color contrast for better visibility.	Reduce short-term memory load
KPP-02	Increase global font size to enhance readability.	Reduce short-term memory load

The prototype underwent a targeted design refinement to operationalize these final recommendations. Figure 4 illustrates the comparative transformation resulting from this second iteration, specifically highlighting the improvement in contrast ratio for inactive icons (KPP-01). Figure 5 shows the scaling of global typography (KPP-02) to maximize legibility. These adjustments marked the culmination of the iterative design process, delivering a fully optimized interface that met both heuristic standards and user expectations.

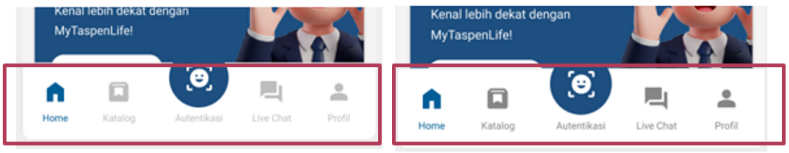


Figure 4. UI Design Transformation for KPP-01

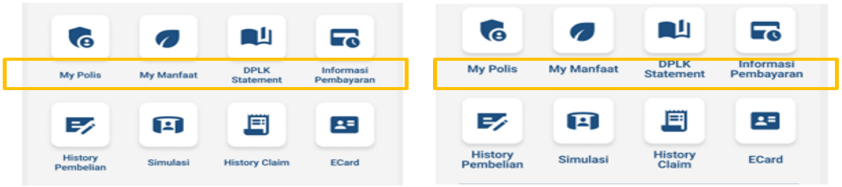


Figure 5. UI Design Transformation for KPP-02

The results across all three methods suggest that this design intervention effectively mitigates most of the challenges identified in the current system. Table 18 lists the results from all evaluative instruments, demonstrating the successful enhancement of overall usability.

Table 18. Overall evaluation results summary: prototype

No.	Methods	Indicators	Results	Summary
1	Usability testing	Task completion level	Task completion rates improved drastically with no critical (Level 4) failures persisting. Identified obstacles were reduced to low-to-moderate severity, confirming operational efficiency.	Synthesizing the cumulative findings from these evaluation phases, the redesigned "My Taspen Life" application demonstrates a robust level of interface usability across all critical dimensions. Consequently, the implemented improvements in interface design have effectively addressed the prior deficiencies, elevating the system to a high standard of usability that supports both user satisfaction and operational reliability.
2	SUS	Average SUS Score	Achieved a final score of 83.54 (Grade A, Excellent), representing a significant improvement from the baseline (65.51) and exceeding global industrial benchmarks.	
3	Open-ended question	Feedback themes	Feedback shifted predominantly to positive sentiment regarding visual modernity and ease of use. The remaining minor issues (legibility/contrast) were identified and subsequently resolved.	

4. CONCLUSION

In this study, we successfully diagnosed and remediated usability deficiencies in the "My Taspen Life" application. The initial evaluation revealed a system rife with usability issues, with a marginal SUS score of 65.51 and critical failures in task completion. Utilizing a User-Centered Design process grounded in Shneiderman's Eight Golden Rules, a high-fidelity prototype was developed to address 56 identified heuristic problems.

Post-implementation evaluation confirmed significant performance improvements, with the SUS score increasing to 83.54 (Grade A, Excellent). The redesign eliminated all critical (Level 4) operational failures, enabling unobstructed completion of essential tasks. Qualitative feedback reflected this improvement, showing increased user satisfaction with the modernized visuals and navigation tools. These findings will provide PT Asuransi Jiwa Taspen with a validated solution that resolves technical issues while enhancing user trust and competitive positioning in the digital insurance sector.

ACKNOWLEDGEMENTS

We thank PT Asuransi Jiwa Taspen, resource personnel, the participants, and the experts who contributed to this research, as well as the Faculty of Computer Science, University of Indonesia.

REFERENCES

[1] Otoritas Jasa Keuangan, "Peta Jalan Pengembangan dan Penguatan Perasuransian Indonesia 2023-2027," Jakarta, 2023.

[2] Deloitte, "2022 Insurance Outlook," 2022.

[3] M. Eling and M. Lehmann, "The Impact of Digitalization on the Insurance Value Chain and the Insurability of Risks," *Geneva Pap. Risk Insur. Issues Pract.*, vol. 43, no. 3, pp. 359–396, Jul. 2018, doi: 10.1057/s41288-017-0073-0.

[4] P. Wahyu Setyaningsih and A. Yakobus Chandra, "Analysis of User Experience Usage on the Sardjito Hospital and Yogyakarta Regional Public Hospital Websites Using the User Experience Questionnaire (UEQ)," *Int. J. Adv. Data Inf. Syst.*, vol. 6, no. 2, pp. 496–512, Aug. 2025, doi: 10.59395/ijadis.v6i2.1411.

[5] Y. Skaf, C. Eid, A. Thrassou, S. El Nemar, and K. S. Rebeiz, "Technology and service quality: achieving insurance industry customer satisfaction and loyalty under crisis conditions," *EuroMed J. Bus.*, 2024, doi: 10.1108/EMJB-01-2024-0027.

[6] B. Szajna, R. W. Scamell, U. Expectations, and B. M. Szajna Neeley, "The Effects of Information System User Expectations on Their Performance and Perceptions," 1993.

[7] A. H. Aldholay, O. Isaac, Z. Abdullah, and T. Ramayah, "The role of transformational

leadership as a mediating variable in DeLone and McLean information system success model: The context of online learning usage in Yemen," *Telemat. Informatics*, vol. 35, no. 5, pp. 1421–1437, Aug. 2018, doi: 10.1016/j.tele.2018.03.012.

[8] J. Nielsen, *Designing web usability: The practice of simplicity*. Indianapolis, IN: New Riders Publishing, 2000.

[9] S. P. Handasari, R. Wulandari, and Haikal, "Evaluation of the Usability and User Experience of the Jaminan Kesehatan Nasional Mobile Application in Indonesia," *Healthc. Inform. Res.*, vol. 30, no. 4, pp. 324–332, Oct. 2024, doi: 10.4258/hir.2024.30.4.324.

[10] P. R. Sefira, I. L. Sardi, and D. R. Anbiya, "Redesigning the User Interface and User Experience (UI/UX) for the Claims Section Website Using a User-Centered Design Approach (Case Study: PT XYZ)," in *2025 International Conference on Data Science and Its Applications, ICoDSA 2025*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 13–18. doi: 10.1109/ICoDSA67155.2025.11157436.

[11] S. Rinaldi, A. Gandhi, and N. Selviandro, "Usability Evaluation and Recommendation of UserInterface Design for e-HAC Application by UsingUser-Centered Design Method," *Int. Conf. Adv. Commun. Technol.*, 2022.

[12] L. Perotti, O. Stamm, M. Dietrich, I. Buchem, and U. Müller-Werdan, "The usability and user experience of an interactive e-learning platform to empower older adults when using electronic personal health records: an online intervention study," *Univers. Access Inf. Soc.*, vol. 24, no. 2, pp. 1193–1208, Jun. 2025, doi: 10.1007/s10209-024-01124-z.

[13] B. Shneiderman and C. Plaisant, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 4th ed. Pearson Addison Wesley, 2004.

[14] S. Dawadi, S. Shrestha, and R. A. Giri, "Mixed-Methods Research: A Discussion on its Types, Challenges, and Criticisms," *J. Pract. Stud. Educ.*, vol. 2, no. 2, pp. 25–36, Feb. 2021, doi: 10.46809/jpse.v2i2.20.

[15] M. R. Indra, T. Raharjo, B. Hardian, and A. Wahbi, "Challenges in Agile IT Project: A Case Study of Banking Company," in *2021 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ICACSIS53237.2021.9631352.

[16] J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th ed. SAGE Publications, 2014.

[17] M. Saunders, P. Lewis, and A. Thornhill, *Research Methods for Business Students Seventh Edition*, 7th ed. Pearson , 2016.

[18] U. Sekaran and R. Bougie, *Research Method for Business Textbook: A Skill Building Approach*. John Wiley & Sons Ltd., 2016.

[19] R. K. Yin, *Case Study Research and Applications: Design and Methods* , 6th ed. Sage.

[20] J. T. Marchewka, *Information Technology Project Management: Providing Measurable Organizational Value*, 5th ed. Wiley, 2015.

[21] J. Jhonson, *Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Guidelines*. Morgan Kaufmann, 2010.

[22] O. D. Alao, E. A. Priscilla, R. C. Amanze, S. O. Kuyoro, and A. O. Adebayo, "User-Centered/User Experience Uc/Ux Design Thinking Approach for Designing a University Information Management System," *Ing. des Syst. d'Information*, vol. 27, no. 4, pp. 577–590, Aug. 2022, doi: 10.18280/isi.270407.

[23] F. Bagheri, F. Abbasi, M. Sadeghi, and R. Khajouei, "Evaluating the usability of a cancer registry system using Cognitive Walkthrough, and assessing user agreement with its problems," *BMC Med. Inform. Decis. Mak.*, vol. 23, no. 1, Dec. 2023, doi: 10.1186/s12911-023-02120-8.

[24] Fachri Munandar and Harry Budi Santoso, "User Interface Improvement by Evaluating Usability and User Experience: Case Study of Indonesia's Government Financial Management Information System," *Indones. J. Comput. Sci.*, vol. 14, no. 1, Feb. 2025, doi: 10.33022/ijcs.v14i1.4630.

[25] J. Marenkov, T. Robal, and A. Kalja, "A tool for design-time usability evaluation of web

user interfaces," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2017, pp. 394–407. doi: 10.1007/978-3-319-66917-5_26.

[26] B. A. Kumar, M. S. Goundar, and S. S. Chand, "Usability guideline for Mobile learning applications: an update," *Educ. Inf. Technol.*, vol. 24, no. 6, pp. 3537–3553, Nov. 2019, doi: 10.1007/s10639-019-09937-9.

[27] N. McDonald, S. Schoenebeck, and A. Forte, "Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice," Nov. 01, 2019, *Association for Computing Machinery*. doi: 10.1145/3359174.

[28] H. B. Santoso, M. Schrepp, and R. Y. Kartono, "Measuring User Experience of the Student-Centered e-Learning Environment," *J. Educ. Online*, vol. 13, p. 59, 2016, doi: 10.9743/JEO.2016.1.5.

[29] J. Sauro and J. R. Lewis, *Quantifying the user experience: Practical Statistics for User Research*. Morgan Kaufmann, 2012.

[30] Ridwan Afandi and H. B. Santoso, "Improving the Usability and User Experience of a Ticketing and Knowledge Base Web Application: A Case Study of the ESDM 136 Contact Center," *Indones. J. Comput. Sci.*, vol. 14, no. 1, Feb. 2025, doi: 10.33022/ijcs.v14i1.4622.