JOURNAL
COMPUTER SCIENCES INSTITUTE

JCSI 39 (2026) 102-107
Received: 5 October 2025
Accepted: 12 November 2025

Comparative analysis of user interface quality of mobile operator applications

Darya Benedziktovich*, Oleksii Davydok

Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland

Abstract
This paper presents a comparative analysis of user interface (UI) quality in mobile applications offered by the four leading telecommunication operators in Poland: Orange, Play, Plus, and T-Mobile. The study aims to evaluate the quality of user interfaces in mobile applications and to identify potential differences that may affect user experience. The evaluation employed multiple research methods, including surveys, corridor tests, expert reviews, and statistical analyses. Key UI criteria such as security, functionality, usability, user comfort, intelligibility, availability, and visual attractiveness were assessed. Results indicate that while all applications deliver high-quality functionality and security, minor differences exist in intuitiveness and user comfort. No statistically significant superiority was found for any application across all evaluated criteria.

Keywords: comparative analysis; interface quality; mobile operators

*Corresponding author
Email address: darya.benedziktovich@pollub.edu.pl (D. Benedziktovich)

1. Introduction

The dynamic development of mobile telecommunications has made mobile applications a key element of customer interaction with operators. In Poland, this market is dominated by Orange (17.7 million clients), Play (15.2 million), Plus (13.1 million), and T-Mobile (12.6 million).

Modern mobile applications allow users to manage their accounts, monitor data usage, make payments, contact customer service, and use promotional offers. As competition increases, clarity, security, and ease of use of interfaces become crucial factors directly influencing customer satisfaction. Well-designed interfaces reduce frustration, increase customer loyalty, and strengthen brand image, while poorly designed UIs can decrease user engagement.

Special attention is currently given to user experience (UX) and user interface (UI) design, which directly affect service perception, user involvement, and technical support workloads. Key factors include cross-platform consistency, visual aesthetics, function integration, and data security.

This article aims to evaluate and compare the quality of mobile interfaces offered by the leading telecommunication operators in Poland, considering aspects of functionality, security, and visual aesthetics.

1.1. Literature review

User interfaces in mobile applications have long been widely studied in the scientific literature, both in ergonomic, psychological, technological, and design contexts. Over the decades, numerous approaches and theories have been developed that apply directly to the analysis of telecom applications UI quality.

According to Hoober and Berkman [1], the basis of modern mobile interface design is responsiveness and adaptability to different devices and screen resolutions. Designers should consider not only screen size but also the conditions of use – on the move, with limited attention, or in poor lighting.

Johnson [2] emphasizes the importance of cognitive psychology in UI design, highlighting the need to minimize cognitive load and ensure application predictability. Norman [3] points out the role of affordances and feedback that enable users to intuitively predict the effects of their actions.

Nielsen [4] formulated usability heuristics such as system status visibility, real-world match, error prevention, and consistency. Krug [5] stresses simplicity and clarity, noting that users rarely analyze all available information but mostly scan.

Tidwell [6] highlights the role of design patterns supporting effective content and function organization, increasing predictability. Perea and Giner [7] emphasize iterative prototype testing and continuous feedback from end users.

Wroblewski [8] presents the Mobile First principle, promoting mobile-oriented design that ensures clarity and focus on user tasks.

Recent studies have reinforced the significance of interface aesthetics in mobile app design. For example, Perrig, Ueffing, Opwis, and Brühlmann [9] found in an online experiment that more aesthetically pleasing mobile web app variants significantly improve both perceived usability and objective performance (number of correct answers) compared to less aesthetic versions. Such results suggest that visual appeal is not just cosmetic: it influences how users perceive and interact with the application, affecting satisfaction and task effectiveness.

The literature shows that UI quality cannot be assessed solely through functionality but requires a holistic approach considering ergonomics, intuitiveness, aesthetics, security, accessibility, as well as psychological and cultural aspects.

1.2. Research objective

The purpose of this comparative analysis is to evaluate and identify differences in the quality of mobile application interfaces offered by the four main telecommunication operators in Poland: Orange, Play, Plus, and T-Mobile. The analysis examines how UIs affect customer experience, and which aspects are crucial for improving user satisfaction.

2. Materials and research methods

2.1. Comparative criteria and weights

Each application was evaluated using weighted criteria (Table 1) to select the best one.

Table 1: Comparative criteria

Criterion	Weight (pt)
K1 Security	22
K2 Functionality	18
K3 Usability	17
K4 User comfort	12
K5 Intelligibility	11
K6 Availability	10
K7 Visual attractiveness	10
Total	100

The final score for each application was calculated according to weighted formulas:

$$o_F = \left[\sum_{i=1}^{7} w_i\, o_i\right] \quad (1)$$

where o_F is the final score of the application, w_i is the weight of the i-th criterion and o_i is the score of the application in the given criterion.

The score o_i for each criterion is the average of the ratings from different evaluation methods, according to formula:

$$o_i = \frac{1}{5n_i}\sum_{j=1}^{n_i} m_{ij}\,, \quad m_{ij} \in [1,5] \quad (2)$$

where n_i is the number of methods evaluating the i-th criterion, m_{ij} is the score in the i-th criterion based on the j-th research method.

2.2. Research objects

The research covered four mobile applications of telecom operators: Orange, Play, Plus, and T-Mobile. All were available for Android and iOS. The latest versions were tested: Orange 5.108.0, Play 10.12.0, Plus 2.372.000, and T-Mobile 25.2.31.

2.3. Research methods

A set of diverse methods was used to evaluate the quality of the mobile application interfaces:

1. Surveys enabled the collection of opinions from a wide range of mobile app users. A questionnaire was developed and distributed electronically using Google Forms.
2. Corridor tests allowed for practical verification of app functionality in conditions simulating everyday use. This method involves spontaneous involvement of random users performing specific tasks in the app. It makes it possible to quickly identify obvious interface problems and obtain immediate feedback.
3. Expert reviews involved individuals with experience using the analyzed apps. Each expert had used the app for at least several months. They evaluated the apps according to predefined criteria, allowing identification of more complex design problems.

Each method enabled analysis of different interface quality aspects. Depending on the characteristics of the method, different groups of criteria were evaluated. Detailed assignment of criteria to methods is shown in Table 2.

Table 2: Criteria analyzed by method

Method	Criteria analyzed
Surveys	K1, K2, K3, K4, K5, K7
Corridor tests	K2, K3, K4, K7
Expert reviews	K1, K2, K3, K4, K6, K7

2.4. Statistical method

To ensure reliable analysis of empirical data from surveys, corridor tests, and expert reviews, it was first necessary to check whether the analyzed variables followed a normal distribution. This is key when choosing the appropriate statistical test – parametric or non-parametric.

For small samples (n < 50), as in this study, the Shapiro-Wilk test was used, which is characterized by high sensitivity and effectiveness. It verifies the null hypothesis:

H_0: the sample comes from a normally distributed population,

H_1: the sample does not come from a normally distributed population.

The analysis was conducted in R using the shapiro.test() function, separately for each set of responses (e.g., Play app security rating). The function returns statistic WWW and p-value. If $p < \alpha$ (where α = 0.05), the hypothesis of normality is rejected.

If the data did not meet the normality condition, the Kruskal-Wallis non-parametric test was applied. This test is used to compare more than two independent groups and is the non-parametric equivalent of one-way ANOVA.

The test statistic H is calculated as:

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3(N+1) \quad (3)$$

where N is the total number of observations, k is the number of groups (applications), R_i is the sum of ranks in group i and n_i is the number of observations in group i.

The statistic H is compared with the chi-squared distribution (χ^2) with k-1 degrees of freedom. If $p < \alpha$, at least one group differs significantly.

Calculations were also performed in Python using scipy.stats.kruskal(). The significance level was set at α = 0.05. If p-value < 0.05, it was concluded that at least one pair of apps differed significantly.

3. Results

For each study, dedicated questionnaires were prepared with questions corresponding to the examined criteria. The surveys primarily employed closed-ended questions in the form of a Likert scale, ranging from 1 (very poor) to 5 (very good). This format enabled respondents to express their evaluations in a simple and intuitive way, ensuring clarity and comparability of the collected data.

3.1. Survey research

For the Orange mobile application, men accounted for 75% of all respondents (12 out of 16), and women – 25% (4 out of 16). The majority of respondents (9 people) were between 20 and 22 years old; the youngest users were 18, the oldest – 24 years.

The Play mobile application was assessed by 73% men (16 out of 22) and 27% (6 out of 22) women. Respondents were between 17-25 years of age.

In the group of T-Mobile users, 86% were men (6 out of 7), and 14% (1 out of 7) – women. The age range of respondents was 19-27 years.

Among users of the Plus application, 100% of the 8 respondents were men, and the age of the respondents ranged from 21 to 31 years, with most users aged 21-25 years.

Table 3 presents the correlation between the questions included in the questionnaire and their corresponding evaluation criteria.

Table 3: Comparative criteria

Questions	Criterion
How do you evaluate the level of security of the application during login?	K1 Security
How do you evaluate the application in terms of the number of functions offered?	K2 Functionality
How do you evaluate the ease of use of the application?	K3 Usability
Do you consider to top-up via the mobile application convenient?	
How do you evaluate the comfort of using the application?	K4 User comfort
How do you evaluate the intelligibility of the application?	K5 Intelligibility
How do you evaluate the graphic interface of the application?	K7 Visual attractiveness

3.2. Corridor tests

In the corridor tests, 11 volunteers participated, randomly selected from various age groups and professional environments, who were able to provide comprehensive responses in a questionnaire regarding the functionality and features of the tested mobile telecommunication applications. Table 4 presents the correlation between the survey questions and their corresponding evaluation criteria.

Table 4: Comparative criteria

Questions	Criterion
Account balance check	K2 Functionality
Account top-up	
Tariff change	
Responsiveness	K4 User comfort
Visual component	K7 Visual attractiveness

3.3. Expert research

As part of the expert research, 8 experienced users were invited – two for each of the analyzed mobile applications. Participants were active users of the given applications for at least several months and had sufficient knowledge to make an informed assessment of functionality, performance, and aspects of the user interface. Table 5 presents the correlation between the survey questions and the corresponding evaluation criteria.

Table 5: Comparative criteria

Questions	Criterion
Application security	K1 Security
Account balance check	K2 Functionality
Account top-up	
Tariff change	
Settings change	
Additional actions occurring in applications	
Responsiveness	K4 User comfort
Availability of the application on other operating systems	K6 Availability
Visual component	K7 Visual attractiveness

4. Data analysis and interpretation

This section presents the analysis of surveys, corridor tests, and expert evaluations to assess the quality of user interfaces in the four mobile applications. Statistical tests were applied to support the interpretation of the results.

4.1. Survey research

The Shapiro-Wilk test was performed for Orange, Play, T-Mobile, and Plus separately. Tables 6-9 show results. The value W_{crit} was read from tables for the sample size n and the significance level $\alpha = 0.05$.

Table 6: Shapiro-Wilk test results for the Orange application

$W_{crit} = 0.82$	Value, W	Normal distribution
Security	0.78	No
Top-up	0.80	No
Graphic interface	0.88	Yes
Intelligibility	0.75	No
Ease of use	0.80	No
Comfort	0.88	Yes
Number of functions	0.82	Yes

Table 7 presents the results of the Shapiro-Wilk test conducted for the examined features of the Play application.

Table 7: Shapiro-Wilk test results for the Play application

W_{crit} = 0.83	Value, W	Normal distribution
Security	0.85	Yes
Top-up	0.83	No
Graphic interface	0.81	No
Intelligibility	0.88	Yes
Ease of use	0.82	No
Comfort	0.79	No
Number of functions	0.84	Yes

Table 8 presents the results of the Shapiro-Wilk test performed for the examined features of the T-Mobile application.

Table 8: Shapiro-Wilk test results for the T-Mobile application

W_{crit} = 0.85	Value, W	Normal distribution
Security	0.79	No
Top-up	0.83	No
Graphic interface	0.87	Yes
Intelligibility	0.73	No
Ease of use	0.82	No
Comfort	0.92	Yes
Number of functions	0.96	Yes

The results of the Shapiro-Wilk test for the examined features of the Plus application are presented in Table 9.

Table 9: Shapiro-Wilk test results for the Plus application

W_{crit} = 0.85	Value, W	Normal distribution
Security	0.80	No
Top-up	0.93	Yes
Graphic interface	0.92	Yes
Intelligibility	0.80	No
Ease of use	0.76	No
Comfort	0.89	Yes
Number of functions	0.83	No

Since none of the examined features followed a normal distribution across the four groups, the nonparametric Kruskal-Wallis test was applied, as it allows comparison of more than two independent groups without assuming normality. The results revealed no statistically significant differences between applications, which may stem from the limited sample size, high rating variability, or genuinely minor differences in user interfaces. Overall, the findings suggest that none of the applications clearly stands out in the evaluated features.

4.2. Corridor tests

The result of the Shapiro-Wilk test for the features examined in the Orange application is presented in Table 10. The value W_{crit} was read from tables for the sample size n and the significance level $\alpha = 0.05$.

Table 10: Shapiro-Wilk test results for the Orange application

W_{crit} = 0.77	Value, W	Normal distribution
Account balance check	0.77	Yes
Account top-up	0.86	Yes
Tariff change	0.88	Yes
Responsiveness	0.79	Yes
Visual content	0.79	Yes

Table 11 presents the results of the Shapiro-Wilk test conducted for the examined features of the Play application.

Table 11: Shapiro-Wilk test results for the Play application

W_{crit} = 0.77	Value, W	Normal distribution
Account balance check	0.73	No
Account top-up	0.81	Yes
Tariff change	0.72	No
Responsiveness	0.79	Yes
Visual content	0.87	Yes

Table 12 presents the results of the Shapiro-Wilk test performed for the features of the T-Mobile application.

Table 12: Shapiro-Wilk test results for the T-Mobile application

W_{crit}= 0.77	Value, W	Normal distribution
Account balance check	0.79	Yes
Account top-up	0.80	Yes
Tariff change	0.77	Yes
Responsiveness	0.65	No
Visual content	0.77	Yes

Table 13 presents the results of the Shapiro-Wilk test for the examined features of the Plus application.

Table 13: Shapiro-Wilk test results for the Plus application

W_{crit} = 0,77	Value, W	Normal distribution
Account balance check	0.79	Yes
Account top-up	0.63	No
Tariff change	0.76	No
Responsiveness	0.62	No
Visual content	0.79	Yes

For the feature “Visual content”, which has a normal distribution, a Student’s t-test was performed. The t-statistic is calculated according to formula:

$$T = \frac{x_1 - x_2}{S} \tag{4}$$

The formula for pooled variance:

$$S = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \times \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \tag{5}$$

The test was conducted at the significance level α. The results are presented in Table 14.

Table 14: Student's t-test for features with normal distribution

Compared applications	Visual content
Play vs Orange	-0.86
Play vs Plus	-1.16
Play vs T-Mobile	-1.49
Orange vs Plus	-0.32
Orange vs T-Mobile	-0.67
Plus vs T-Mobile	-0.34

For the Student's t-test, considering the sample size in each group, the critical value was adopted as $t = \pm 1.725$ for the significance level $\alpha = 0.1$. None of the compared applications exceeded this threshold in the analysis of the "Visual content" feature, which allows the conclusion that the differences between applications are not statistically significant. The results indicate that users rated the graphical layer of all four applications comparably, which may indicate a similar level of aesthetics, readability, and visual consistency.

For features that do not have a normal distribution, the Kruskal-Wallis test was performed.

Based on the Kruskal-Wallis test results, it can be stated that none of the analyzed features showed statistically significant differences between the application ratings. This may indicate a large diversity of individual user ratings or a similar level of interface quality of the analyzed applications. Therefore, it can be concluded that within the studied sample there are no significant differences between the applications in terms of the studied functionalities.

4.3. Expert research

Based on the opinions of experts who evaluated the functionalities of the Play, Orange, Plus, and T-Mobile applications, the following conclusions can be formulated:

Account balance check was rated very highly in all applications. Experts highlighted quick and intuitive access to the balance. A slightly lower score in the case of T-Mobile may result from the need to perform additional steps before displaying the information.

Account top-up worked efficiently and was rated maximally in Play, Orange, and Plus applications. The T-Mobile application received a score of 4, which may indicate less intuitive handling or additional confirmations in the top-up process.

Tariff change was rated best in Orange, where experts appreciated the clear structure of options and simple access to the offer. The other applications received a score of 4, which suggests less transparent navigation or the need to go through additional sections.

In terms of responsiveness and speed of application operation, T-Mobile and Play were rated highest. Orange received a lower score, which may indicate noticeable slowdowns or less efficient rendering of interface components.

The visual component, responsible for aesthetics and readability of the interface, obtained a consistent score in all applications. According to experts, each of the applications presents a good graphical level, with a readable layout and modern color scheme.

Application security was rated maximally in all cases. High ratings resulted from the use of standards such as biometrics, two-step login, or local data encryption.

Availability of applications on different operating systems (Android/iOS) also received the highest possible scores, which confirms that each of the applications works reliably regardless of the platform.

Regarding additional functions, such as special offers, discounts, or service point location, Orange and Plus were rated highest. T-Mobile and Play received lower scores, which may result from a limited number of conveniences or less extensive promotional sections.

Settings change was rated maximally in Play, Orange, and Plus applications. T-Mobile received a slightly lower score, which nevertheless keeps it at a high level, though it may indicate minor unreadabilities in the layout.

5. Final evaluation

Based on the conducted analyses (including the results of survey research, corridor tests, and expert evaluations), a final classification of the mobile applications of the four operators: Orange, Play, Plus, and T-Mobile, was developed. A point-based summary is presented in Table 15.

Table 15: Point-based summary

Criterion	k_i			
	Orange	Play	T-Mobile	Plus
K1 Security	19.5	19.7	19.8	18.4
K2 Functionality	15.3	14.4	13.6	14.5
K3 Usability	13.4	13.4	11.9	11.9
K4 User comfort	9.9	9.3	8.6	8.8
K5 Intelligibility	7.8	8.3	8.2	8.0
K6 Availability	10	10	10	10
K7 Visual attractiveness	7.6	7.7	7.1	7.7
k_F	83.6	82.8	79.3	79.2

The results indicate that the Orange application achieved the highest overall score – 83.6/100, slightly ahead of Play – 82.8/100. The next positions were taken by Plus (79.3 points) and T-Mobile (79.2 points). However, the differences between the applications are minor, which confirms their comparable overall level. It should be noted that such small discrepancies may partly result from the use of the five-point scale (1-5), which – although intuitive and simple for respondents – limits the precision of assessments compared to a ten-point scale, where greater detail in evaluations could be captured.

6. Conclusions

A multi-stage comparative analysis of the user interface quality of mobile applications offered by Poland's four largest telecommunication operators – Orange, Play, Plus, and T-Mobile – revealed both strengths and weaknesses of each solution. Based on surveys, corridor tests, expert evaluations, and multi-criteria statistical analysis, Orange achieved the highest overall score, slightly ahead of Play. The differences between the applications were

relatively small, confirming the generally high and uniform level of interface development in the sector.

Key evaluation criteria included security, functionality, usability, user comfort, intelligibility, availability, and visual attractiveness, enabling a multidimensional assessment that reflected both technical parameters and user perceptions. Statistical methods, such as the Shapiro-Wilk and Kruskal-Wallis tests, provided objective grounds for comparison, showing that many features did not differ significantly between applications.

The findings may be valuable for operators as a basis for further interface improvements, as well as for designers seeking inspiration and best practices. Personalization, responsiveness, and aesthetics emerged as decisive factors in positive user evaluations. Future development aligned with trends such as AI, automation, dark mode, and loyalty programs may further enhance user satisfaction, engagement, and brand competitiveness.

References

[1] S. Hoober, E. Berkman, Designing Mobile Interfaces: Patterns for Interaction Design, O'Reilly Media, Sebastopol, 2011.

[2] J. Johnson, Designing with the Mind in Mind, Morgan Kaufmann, San Francisco, 2014.

[3] D. A. Norman, The Design of Everyday Things, MIT Press, Cambridge, 2013.

[4] J. Nielsen, Usability Engineering, Academic Press, San Francisco, 1993.

[5] S. Krug, Don't Make Me Think, New Riders, Berkeley, 2014.

[6] J. Tidwell, Designing Interfaces, O'Reilly Media, Sebastopol, 2020.

[7] P. Perea, P. Giner, UX Design. Projektowanie aplikacji dla urządzeń mobilnych, Helion, Gliwice, 2020.

[8] L. Wroblewski, Mobile First, A Book Apart, New York, 2011.

[9] S. A. C. Perrig, D. Ueffing, K. Opwis, F. Brühlmann, Smartphone app aesthetics influence users' experience and performance, Front. Psychol. 14 (2023) 1113842, https://doi.org/10.3389/fpsyg.2023.1113842.