Deco 7250
Human Computer Interaction

Evaluation Report

Team Blank Frame Only

Member: Tian Shi, Yu Liang,Yingzi Zhuang,Guo Cheng

2022-06-12

1.Executive Summary

1.1 Background

With a growing number of non-traditional users competing with banks and payment service providers, the field of transaction is becoming increasingly crowded. The changes started earlier than the COVID-19 pandemic, it forced established institutions to accelerate the pace of digitization and also keep innovating. The area of digital transaction is becoming more and more important.

1.2 Problem statement

In this design, however, we're targeting users who are struggling with the digitization of traditional transaction methods and have the requirement of managing their property. Applications of established institutions often have complex interfaces, too much functionality and information, which often leads to the non-friendly system and confuses middle-aged and older groups of users. In our project, we distinguished the specific needs and values of the elderly group by analyzing online questionnaires and interviews to set up UX goals, researching and analyzing existing products, and put forward our new design proposal.

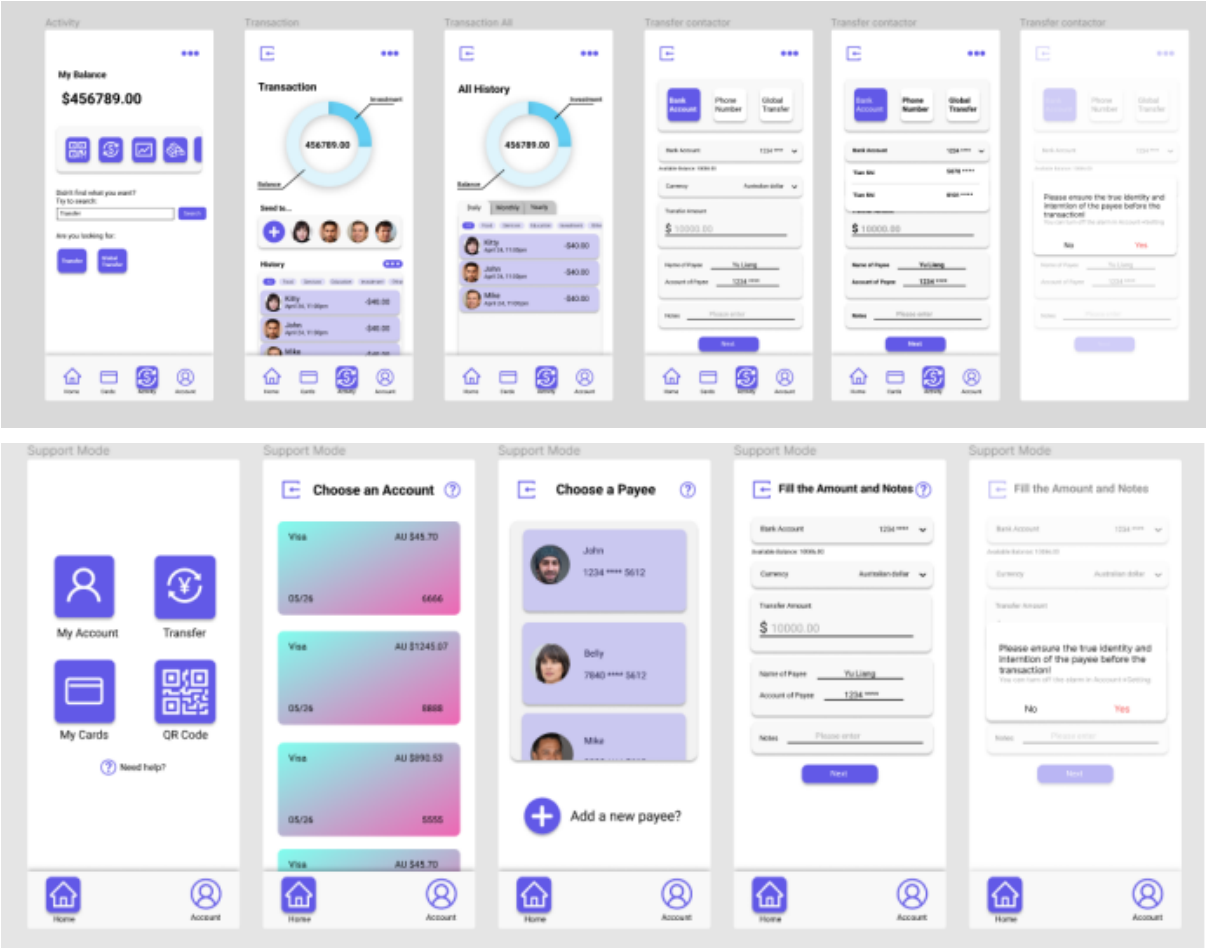
1.3 Evaluation approach

During the evaluation process, our prototype was evaluated via different methods.

The methods we used are: SUS, Think Loud, Time of Tasks, and also Heuristic Evaluation methods based on Nielsen’s 10 principles. We firstly completed the tests with our team members, and then we invited students from other groups to participate, and they did provide plenty of valuable evaluating results.

1.4 Outcomes

We iterated our prototype based on the results we collected from the evaluation.



We designed our support mode based on the requirement raised by our evaluation participants, we considered the evaluated results such as adding explanations to the icons, and making the texts easier to read. The user interface of support mode is fixed to meet the requirement of the older users.
And it shows improved satisfaction in our further test.

2. Evaluation Plan

2.1 Purpose

Reviewing existing systems can help us find current problems and improve them. We collected qualitative data to diagnose and identify the problems and causes in the design, such as SUS, Time on Tasks, and Heuristic, and combined them with the Think Loud method. The data and results can form specific optimization suggestions

and our future direction. In the assessment activities, we want to have a clear understanding of the following aspects of the system:

2.1.1 Whether the functional structure of the system is clear:

Our system has multiple secondary and tertiary pages which have their own architecture, therefore it may be difficult for users to understand, and redundant pages may exist. We want our system to have a clear structure with no unnecessary pages, and help users quickly understand the main pages and functions of the system, tell users where they are currently located, and access all the main features through the home page.

2.1.2 Whether the use process of the system is clear:

Intersection applications often have a variety of complex functions, and to protect the property security of users, it is particularly important to reduce irrelevant branches and prevent users from misoperation processes. We want to show every process in a simple, clear and clear way, give users enough time to think and modify, and provide feedback on their operations.

2.1.3 Whether the controls of the interface are used accurately:

The accurate use of interface controls can help users to better understand the role of each function. We want the icons, tabbar and button etc, used in the system to be stylistically consistent, and easily understood and accepted by the users, and have accurate state representation functions, such as indicating the location of the user.

2.2 SUS Evaluation

The SUS method is a questionnaire-based way of assessing whether a system is suitable for use by users. It's a quick questionnaire that is scored on a five-point scale after respondents have used the app. 1 is strongly disagree and 5 is strongly agree. We analyzed data based on how respondents scored each question.
For the reverse rating mechanism in the SUS method, we need to transform the original data after waiting to obtain the data we need.

2.2.1 Target Users

We invited some young people who are proficient in using mobile phones to test our application. We also invited some seniors as our special user group to test the

features we developed for seniors. This time the test population is mainly young people, and through the data they give, developers iterate and modify.

2.2.2 Task Settings

(1) Respondents are invited to test and use the app.

(2) When the tester has used it for a period of time, it is graded on the questions listed by the developer.

(3) After the original data is obtained, the developer converts the data through calculation, which has achieved the purpose of being usable.

(4) The developer tabulates the data.

2.2.3 Other Setting

The test was hosted by a moderator, who was mainly responsible for explaining the application to testers and collecting and distributing questionnaires.

2.2.4 Results Report

We convert the data of the mobile phone and make it in the form of a table. The mean of the data was analyzed, and each set of data was rated according to the SUS method. Whether the output system for the level meets the user's needs.

2.3 Time of Tasks Evaluation

Efficiency is one of the main aspects of product availability, and the time of tasks is one of the best predictors of efficiency. In addition to reviewing the efficiency of the system, the proposed method is also very effective for diagnosing the problems existing in the interaction interface, because the longer task times may be caused by the problems of interacting with the interface.

2.3.1 Target User

Supportive mode is one of the main functions in our system. Since the target user of this mode is people who potentially need help, or have low sensitivity or vision, we will not conduct time on tasks for the user group. Therefore, our testing subjects will be youth groups who need to use transaction applications. At the same time, developers will also participate in the test, and the resulting time will be used as a reference result.

2.3.2 Task Settings

(1) Apply for a new bank card in your current account.
(2) Use the bank card to complete a single transfer.
(3) Delete the bank card in account management.

2.3.3 Starting Time

The moment when the initial interface appears in the user-shared screen.

2.3.4 Ending Time

In the user-shared screen, the moment when the ‘delete successfully’ notification appears.

2.3.5 Other Settings

(1) Since users may be influenced by sound and thinking which cause changes in task time, we (including users and host) will keep quiet during the test to avoid interference.
(2) The activity will be hosted by one developer, while the time and questions users encounter will be recorded by another developer.

2.3.6 Results report

Due to the small number of samples for this test, we will report the time of completion for each user and the problems they encounter which may lead to a time extension.

2.4 Think Loud Evaluation

Think Loud is a basic technique for usability which requires participants to think aloud when they are doing tasks. It can be used to understand participants’ mental processes and find problems with the interface and interaction process.

2.4.1 Target users

We divided the users into two groups and set up different tasks for them. One is the youth group who need to use a transaction application. They can often receive new technologies faster and have good self-learning skills. The other is for groups that may need help and are less receptive, such as people with less vision or older people with lower sensitivity.

2.4.2 Task Settings

(1) For the youth group, we will set the same task as the task time, namely to add new cards, complete the transfer, and delete the new card.
(2) For groups who need help, we will set the following tasks:
    ① Open the support mode in the Settings page
    ② Complete a transfer in support mode

2.4.3 Other Settings

(1) The activity will be hosted by one host and the record of user operations and psychologically thinking will be recorded by another one.
(2) The host and the recorder will remain silent during the youth group test.
(3) the host can provide some help to the users during the support group test.

2.4.4 Results report

We will report and analyze the operational process and verbal mental activity of all the users involved in the test.

2.5 Heuristic Evaluation

Heuristic evaluation is the use of a relatively simple, generic, and inspiring set of usability rules to evaluate the usability of a product. It is quick and cheap and can be done using existing internal staff. Heuristics are not limited to a specific time period or product phase and can be used wherever interaction experts are needed to evaluate and validate solutions.

2.5.1 Evaluators

According to an experiment conducted by Nielsen in 2000 (Nielsen, 2000), only about five evaluators can find an average of 85% of usability problems by using heuristic evaluation. The evaluators need to have a background in HCI, which greatly affects the results. Therefore, we decided to use 5 evaluators for this heuristic evaluation, 4 internal team members and 1 external member, all with the same academic background and all beginners in the HCI field.

2.5.2 Heuristic evaluation methods

Since our evaluation target is a mobile application, according to G. Joyce's research in 2016(Joyce, 2016), heuristic evaluation of mobile applications using the SMART method can find more usability issues than using the Nielsen method, but the Nielsen method is easier to use. Therefore, we decided to use a combination of Nielsen(Nielsen, 1994) and SMART (Joyce, 2014) methods for this heuristic evaluation, expecting to get more comprehensive evaluation results.

H1	Visibility of system status
H2	Match between system and the real world
H3	User control and freedom
H4	Consistency and standards
H5	Error prevention
H6	Recognition rather than recall
H7	Flexibility and efficiency of use
H8	Aesthetic and minimalist design
H9	Help users recognize, diagnose, and recover from errors
H10	Help and documentation
H11	Display an overlay pointing out main features when appropriate or requested to help first-time users.
H12	Each interface should focus on one task so that it's glanceable to users who are interrupted frequently.
H13	Allow configuration options and shortcuts.
H14	Facilitate easier input by displaying keyboard buttons that are as large as possible, supporting, multimodal input, and keeping form fields to a minimum.

2.5.3 Level of usability issues

Level	Description
0	No usability issues at all
1	UI issues only, can be solved if there is enough time
2	Smaller usability issues that should be prioritized
3	Major usability issues that are very important and should be given high priority
4	Serious usability issues, which must be resolved immediately before product release

3. Evaluation Activities

3.1 SUS Evaluation Activities

Background & Rationale

The SUS method was published by John Brooke in 1986. Its full name is the System Usability Scale. It was originally used as a usability test for VT00 terminal applications. After that, this method has been widely used for applicability testing of various hardware, consumer software, websites, and mobile applications. It has become an industry standard and is a very convenient and quick application test in the form of a questionnaire. This questionnaire has ten questions and five answer options. After the testers complete the test of the application, they fill in the questionnaire, and then the developers convert the data and make a form.
This test invited 15 testers, including five students from other groups for testing, eight ordinary users for testing, and two special users (senior users) for testing. After testers use the features of the app, testers fill out a questionnaire. After the developers analyze the data, they can modify and iterate on the existing problems of the application.

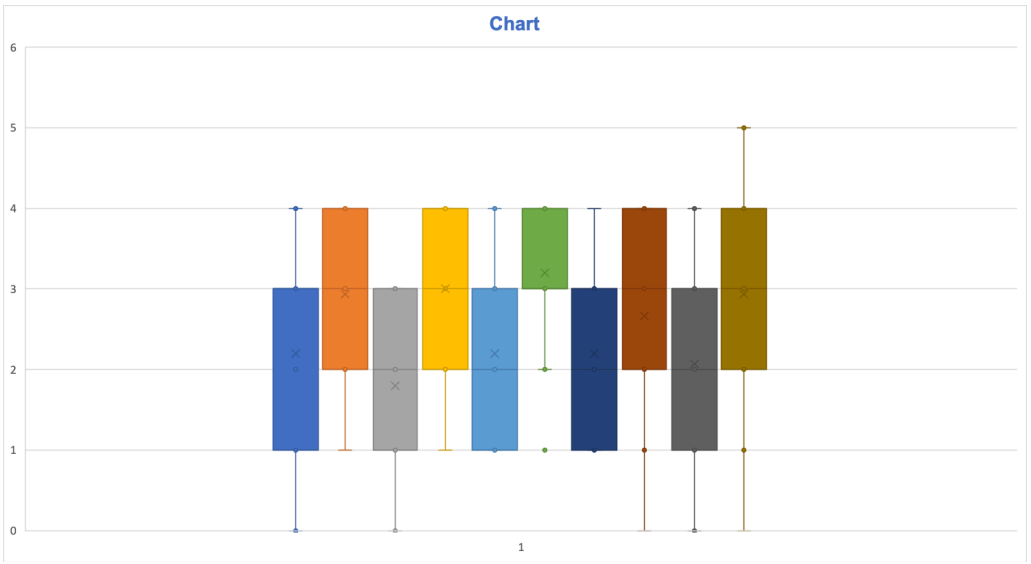
Evaluation Protocol

Task	Duration
Fifteen testers were invited to participate in the test, and testers were free to use the features of the application.	10 mins
Testers begin to test the function of each page as required by the moderator and consider whether the function matches the interface design.	40 mins
After completing the experience of using the product, the tester will score the problems of the application on a scale of 1-5 points. 1 means strongly disagree, 5 means strongly agree.	5mins
Developers collect the data and transform the data using the SUS method, and finally make a table. It would facilitate data analysis.	60mins

Row data

	Resp1	Resp2	Resp3	Resp4	Resp5	Resp6	Resp7	Resp8	Resp9	Resp10	Resp11	Resp12	Resp13	Resp14	Resp15
I think that I would like to use this system frequently.	3	1	4	1	4	4	2	5	4	5	2	3	4	2	4
I found the system unnecessarily complex.	3	2	2	4	1	3	3	2	2	1	1	2	3	1	1
I thought the system was easy to use.	2	1	3	2	3	2	1	4	4	4	3	2	4	4	3
I think that I would need the support of a technical person to be able to use this syster	1	3	1	3	2	1	4	1	2	1	2	3	2	3	1
I found the various functions in this system were well integrated.	3	2	3	2	4	2	3	5	3	5	3	4	4	3	2
I thought there was too much inconsistency in this system.	2	4	1	3	1	3	1	1	2	2	1	2	1	1	2
I would imagine that most people would learn to use this system very quickly.	4	2	3	2	4	2	2	3	4	5	2	3	4	4	4
I found the system very cumbersome to use.	2	4	1	4	2	2	3	1	3	2	1	2	1	2	1
I felt very confident using the system.	3	1	4	1	5	3	2	4	2	5	3	3	3	4	3
I needed to learn a lot of things before I could get going with this system.	2	5	1	3	1	1	4	0	4	0	2	2	2	2	2

	Resp1	Resp2	Resp3	Resp4	Resp5	Resp6	Resp7	Resp8	Resp9	Resp10	Resp11	Resp12	Resp13	Resp14	Resp15	Average
I think that I would like to use this system frequently.	2	0	3	0	3	3	1	4	3	4	1	2	3	1	3	2.2
I found the system unnecessarily complex.	2	3	3	1	4	2	2	3	3	4	4	3	2	4	4	2.93333333
I thought the system was easy to use.	1	0	2	1	2	1	0	3	3	3	2	1	3	3	2	1.8
I think that I would need the support of a technical person to be able to use this syster	4	3	4	2	3	4	1	4	3	4	3	2	2	2	4	3
I found the various functions in this system were well integrated.	2	1	2	1	3	1	2	4	2	4	2	3	3	2	1	2.2
I thought there was too much inconsistency in this system.	3	1	4	2	4	2	4	4	3	3	4	3	4	4	3	3.2
I would imagine that most people would learn to use this system very quickly.	3	1	2	1	3	1	1	2	3	4	1	2	3	3	3	2.2
I found the system very cumbersome to use.	3	1	4	1	3	3	2	4	2	3	4	3	0	3	4	2.66666667
I felt very confident using the system.	2	0	3	0	4	2	1	3	1	4	2	2	2	3	2	2.06666667
I needed to learn a lot of things before I could get going with this system.	3	0	4	2	4	4	1	5	1	5	3	3	3	3	3	2.93333333
Scores	62.5	25	77.5	27.5	82.5	57.5	37.5	90	60	95	65	60	62.5	70	72.5	
Rank	B	D	B	D	B	C	C	A	B	A	B	B	B	B	B	



Data Analysis

The data obtained through the SUS Law shows that the app has many shortcomings. From the questionnaire, it is known that users feel that the functions of the application are too complicated, which leads to a decrease in the frequency of users' use, which reduces the user's dependence on the application. The use of the application is not easy enough, which brings learning costs to users. The settings of the application page are not well integrated with the functions, so it is necessary to adjust the interface design of the application to better adapt to the functions. However, after the user learns, the proficiency of the application will be improved. Testers expressed concerns about the need to learn how to operate the app before using it. These problems revealed by the investigation are the main content of subsequent software iterations and revisions. The SUS method is great for helping developers test their applications.

3.2 Time of Tasks Evaluation Activities

1. Activity Protocol

Process	Time
Step 1 Share the interface and explain the tasks to the evaluator, requiring the evaluator to share their screen when operating.	5min
Step 2 At the moment evaluators share their screen and open the homepage, start timing.	Depends on the evaluators.
Step 3 When the evaluators are doing the tasks, the recorder writes down all the operations and the problems they encounter.	
Step 4 At the moment evaluators finish the tasks and the 'delete successfully' page appear, stop timing	

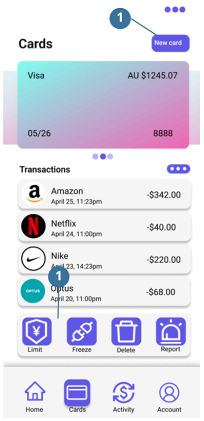
The target evaluators joined in the time of task activity were the classmates in the studio, they were users who had experience using transaction applications and were receptive to new applications. During the activity process, the host and the recorder recorded user operations and problems in strict compliance with the protocol requirements, and remained silent throughout.

2. Raw Data

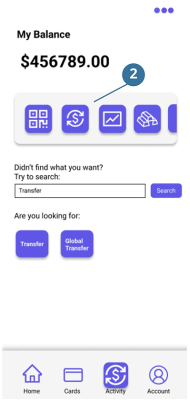
User	Time to finish the tasks	Problems encounter
One of the group members	40s	/
User1	1min10s	1. Couldn't find the add card function 2. Couldn't recognize the transaction icon 3. Hesitant to answer the caution window

User2	48s	1. Couldn't jump from card page to transaction page

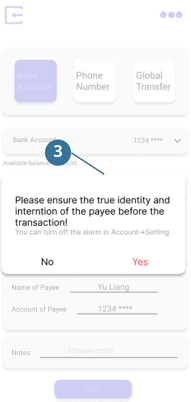
3. Data Analysis



The functional structure of the card page is not arranged well and fits users experience and habit. The add card function is also one of the card management functions, however, in our system it is seperated with other management like freeze and delete.



The controls of the activity page are difficult for users to understand, which require text explanation under the icon.



The information is not conveyed accurately in the alarm window, and the options setting is going against users' usage habits.

3.3 Think Loud Evaluation Activities

1. Activity Protocol

Process	Time
Step 1 Share the interface and explain the tasks to the evaluator, requiring the evaluator to share their screen when operating.	5min
Step 2 When the evaluators are doing tasks, the recorder writes down what they are thinking and the problems they encounter.	Depends on the evaluators.

We seperated the evaluators into 2 different groups, which were youth group and group team, and each group had their own tasks to be finished. During the youth group process, the recorder and host were silent in strict compliance with the protocol requirements, while in the elderly group, the host was ready to provide the evaluators with certain help because they were less receptive to the application and may have eyesight problems etc.

2. Raw Data

User1: an adult with experience about using transaction application

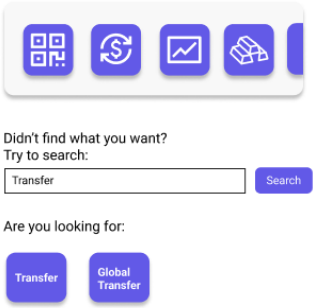
Operation	Quote

Click ‘Card’ in the bottom menu	/
Click ‘Add new card’	‘Where is the new card? Oh on the top. I thought it was on the management bar like delete or freeze card’
Fill in the chart and click ‘Apply’, then return to the homepage	/
Click ‘transaction’ on the homepage	‘Nothing happen’
Enter ‘activity’ page and click ‘transaction’	‘I have to search for it? Is the search engine the only way to find other activities? That’s troublesome.’
Click the icon under the ‘Send to…’ and finish the transaction	/
Enter the card page and delete the card	/

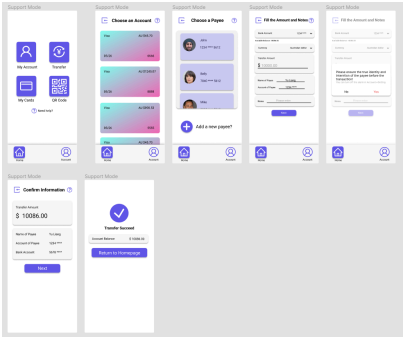
User2: an old man who gets used to go to the offline bank

Operation	Quote
Click ‘My account’ in the bottom menu	‘I have no idea where it is. No one will know about the function if you don’t tell them!’
Click ‘Setting’	
Turn on ‘Support Mode’	
Back to the homepage	/
Click ‘Transfer’	‘Do I have other activities to do like the ordinary mode? I would also like to invest in something!’
Choose an account and a payee	/
Fill the amount and click ‘next’	‘I don’t need to check any of the information, do I?’
Read the alarm and click ‘yes’	‘The font has become smaller here.’
Confirm the information and transfer succeed	/

3. Data Analysis



During the youth group evaluation, the users believed it was troublesome to search for every activity when they needed to, and it would be better if they could customize their own activity bar.



The support mode seems to have more problems:

1. Users are not informed about the function when they first download the app.
2. In the support mode the interfaces are clearer and simpler, however, the functions are also reduced, which causes some discontent among the evaluators.

3.4 Heuristic Evaluation Activities

Evaluation guidelines

- Use the principles as a basis for evaluation and avoid using personal preferences.
- Try to put yourself in the perspective of an inexperienced initial user.
- Identifying potential usability problems in the product is to improve the product.

Evaluation protocol

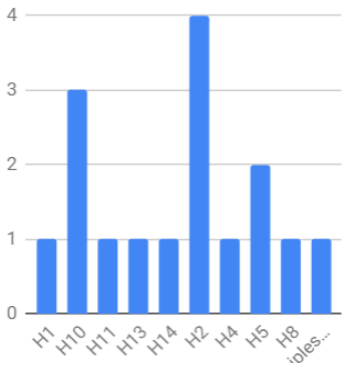
Task	Duration
Evaluators will be free to use the product to understand the basic user interface and the range of interactions.	10 mins
Evaluators evaluate each interface individually using the heuristics given in the plan.	30 mins
Evaluators use a specific *Evaluation Record Sheet* to record the issues identified on a one-by-one basis.	10 mins

The team should summarizes and analyzes the collected data and holds a group meeting to discuss.	30 mins

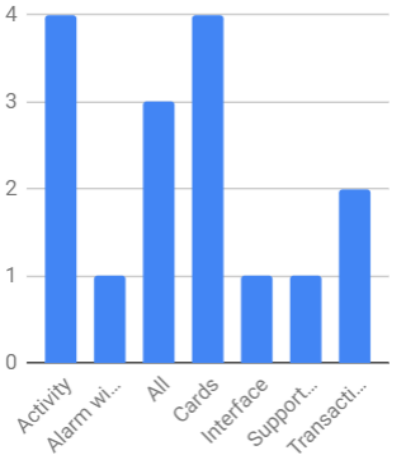
Row data

Issues Number	Evaluator	Principles violated	Level of usability issues	Interface	Problem Description
1	E1	H2	2	Activity	Need to explanin whtat icons means.
2	E1	H13	3	Activity	Add/delete common used function in activity
3	E1	H10	4	All	No any help icon or tips
4	E1	H5	2	Card	Card management need a confirm button
5	T1	H14	2	Transaction	Overly complex form interactions in a single interface
6	T1	H11	4	All	Lack of guidance for new users
7	T1	H2	2	Cards	The meaning of an icon is ambiguous with respect to its function
8	T2	H2	2	Alarm window	The statement and option of the alarm window is contrary to the users habits
9	T2	H5	3	Cards	Vital operation of the cards should be confirm twice
10	T2	H10	4	Support Mode	Introduce support mode to the users who need
11	T3	H4	2	Activity&Account	The system is deficient in error prevention, and some prompt information is not clear enough.
12	T3	H1	1	Cards	The system uses some words and sentence patterns that users are familiar with, which is more in line with the usage habits of the real world.
13	T3	H8	3	Cards&Transaction	Using familiar language to inform users of problems and solutions, rather than showing professional information such as code.
14	T4	H2	2	Activity	Meaning of each icons are not clear enough
15	T4	H10	4	Support Mode	Put the entrance of support mode on a obvious area - easy to access

Data analysis



According to the statistics, a total of 15 usability issues were tested in the Heuristics evaluation. This product has more usability problems in the Heuristics of H2,H5,H10. And it basically conforms to H3,H6,H7,H9,H12.



According to the statistics, the product has more usability problems in Activity and Cards sections, which should be improved as the focus of the next iteration.

4. Results

Our team collected plenty of data during the evaluation activities, some good points were generated from the evaluation, and there are also some problems raised. Results of our evaluation show that our applications had satisfied the majority of our evaluating participants, however, there are some problems raised from different evaluation activities.

Our team found some shortcomings of this app via data obtained through the SUS Laws. The complexity of functions of the application is one reason that leads to decreased use of bank applications, and it reduces users' dependence on this app. Users need to spend more time adapting the system and learning to use different functions.
Based on the SUS result, our team tried to adjust the app's interface design, and tried to combine similar functions.

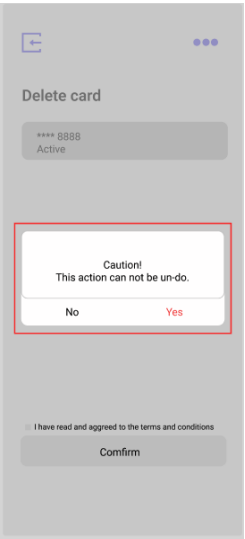
During the 'Time of Task' and 'Think loud', we found that the time that each participant spent adapting the system was different, and some participants were just waiting without moving and speaking, as our host and recorders did not provide any tips. Result of these evaluation actions shows that we should pay more attention to the information that is provided to the users, such as guiding information and function introduction. The old users tend to ask the host for help, as they might be less receptive to the application, which means we should consider how we could make the supporting mode more useful and acceptable to the old users.

In the heuristic evaluation activities, we found that the main problems raised are related to H2, H5 and H10. Also, there are some problems in activity and cards, which means our application still has some defects that should be improved to fit the requirements of real world use. This represents a lack of guidance for new users, which can lead to user loss. It is also more difficult for users to solve errors on their own and they cannot find effective help information. This is a very serious usability issue and should be a high priority in the iteration process.

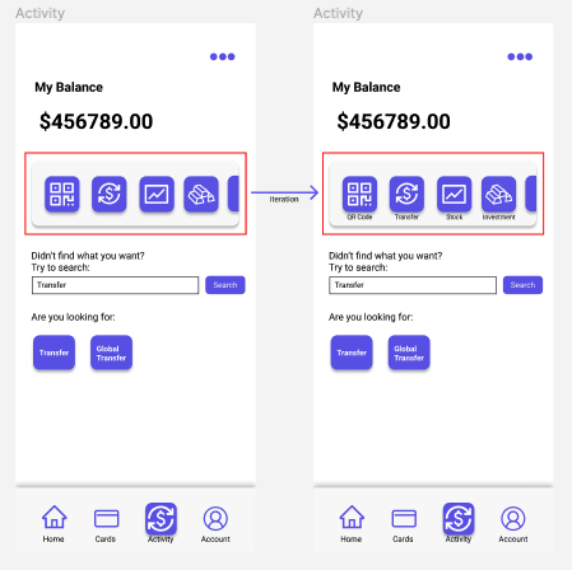
5. Future Directions

Future work on the application will focus on improving and iterating on the original interface and functionality. Iterations of the application are performed based on the data obtained from the evaluation. Firstly, making improvements to existing features. The data collected in the evaluation shows that some functions of the application are not perfect.

1. In the principle of error prevention, the text information prompts used by the application are not clear enough, which is easy to confuse users.

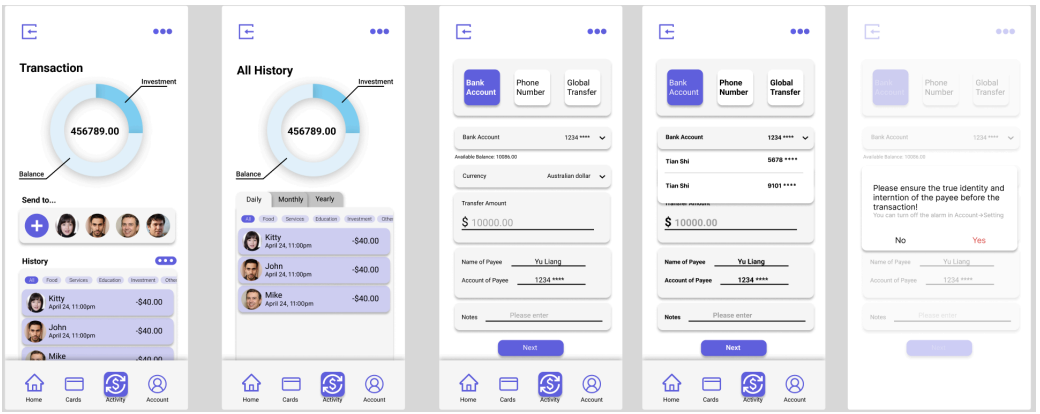


2. There are also some icons that lack text prompts, and users cannot find function icons when using them.



3. Furthermore, in the app evaluation, we invited testers to rate the aesthetics and simplicity of the app. They think the interface design of the application is

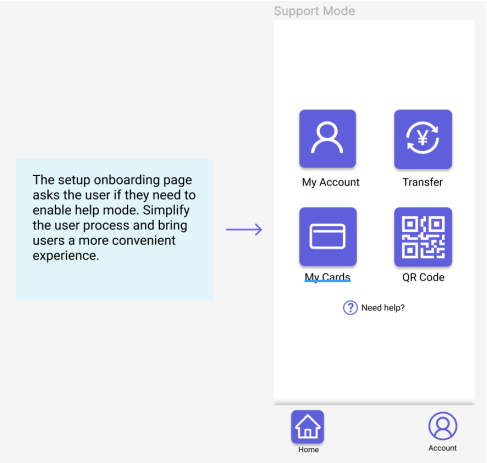
too complicated. Pages need to be optimized for simplicity, and they need to be more detailed and clearly divided.



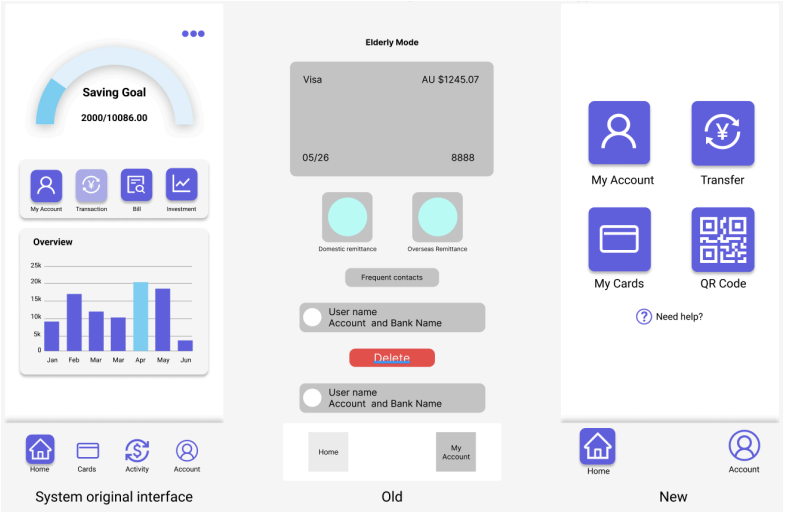
The data collected through the SUS method shows that some interface designs are not well integrated with functions, and users will be cumbersome to use. Therefore, in the subsequent product iterations, it is necessary to modify these problems, increase the language description of the application, simplify the user prompt text, and optimize the page design to better adapt to the function. Reduce the cost of learning and use for users, and users can quickly use it proficiently after installing the application.

In addition, we also invited a special user population of the app for evaluation. Our special user population is for people who may need help and are less receptive, such as people with poor vision or older adults with low sensitivity. Our app has features developed specifically for these groups of people. However, due to the lack of research and data, the model developed for the elderly in the early stage could not bring a convenient experience to users. Therefore, in-depth development of this mode is required in the later stage.

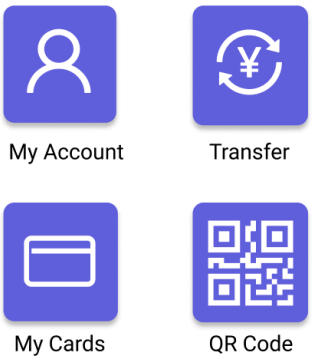
4. For example, when the users use the application for the first time, ask the users or provide an indicator bar, explain the function to them, and ask them whether to turn on this function to make it more convenient for the user to find.



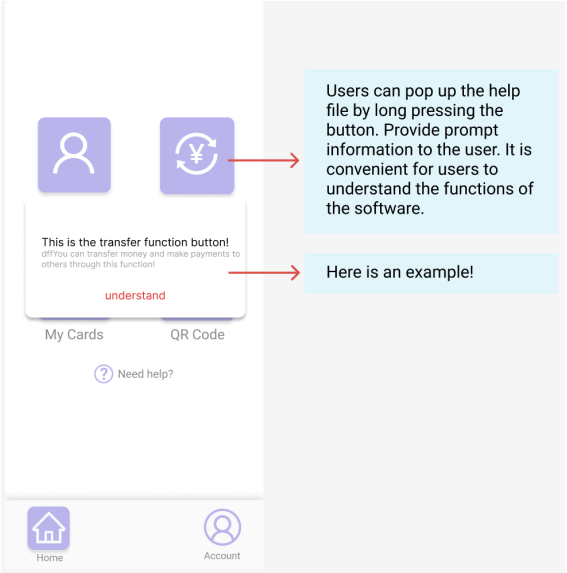
5. And simplify some complex and unnecessary functions.



6. Enlarge the font so that the elderly can see the text on the screen clearly. And text prompts for icons.



7. Add the function of long pressing the button to pop up a prompt window to explain to elderly users and users who are not skilled in mobile phone operation. It is convenient for users to get used to and learn how to use our software.



In the future development direction, we will continue to revise and iterate the application based on the feedback from users and the results of this evaluation to meet the needs of users.

Reference

Jakob Nielsen. (2000). Why You Only Need to Test with 5 Users. Nielsen Norman Group. https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

Jakob Nielsen. (1994). 10 Usability Heuristics for User Interface Design. Nielsen Norman Group. https://www.nngroup.com/articles/ten-usability-heuristics/

Joyce, G., & Lilley, M. (2014, June). Towards the development of usability heuristics for native smartphone mobile applications. In International Conference of Design, User Experience, and Usability (pp. 465-474). Springer, Cham.

Joyce, G., Lilley, M., Barker, T., & Jefferies, A. (2016). Mobile application usability: heuristic evaluation and evaluation of heuristics. In Advances in human factors, software, and systems engineering (pp. 77-86). Springer, Cham.

P. Jeff Sauro. (2011). "Measuring Usability with the System Usability Scale (SUS) – MeasuringU", *Measuringu.com*,. https://measuringu.com/sus/

"Understanding and interpreting box plots | Wellbeing@School", *Wellbeingatschool.org.nz*, 2022. https://www.wellbeingatschool.org.nz/information-sheet/understanding-and-interpreting-box-plots.

Appendix