Usability Testing of a Mobile Application for Alleviating Postpartum Emotional Disorders: A Case of We'll

Wen-Ko Chiou[1], Shih-Chen Lai[2], and Ding-Hau Huang[3(✉)]

[1] Chang Gung University, Taoyuan, Taiwan
[2] Ming Chi University of Technology, New Taipei City, Taiwan
[3] National Taipei University of Business, Taoyuan, Taiwan
hauhuang@ntub.edu.tw

Abstract. Postpartum emotional disorders depending on the severity of symptoms can be divided into postpartum blues, postpartum depression, and postpartum psychosis. Those who develop severe symptoms may intend to commit suicide or harm the newborn. Therefore, these women need professional medical assistance and care services. A previous study (Kao 2017) used social support as the basis to develop a simulation trial version of a mobile application (App), We'll, which helps women suffering from postpartum emotional disorders relieve depression via its function of interaction. The We'll app for the present study is a trial simulation version. The official version of Android We'll system hasn't been evaluated its usability yet. The purpose of the study is for the usability evaluation of Android We'll system. The researchers indicated users' usability, satisfaction, and feedback regarding the interface and functional frameworks of the current app version. This study recruited 30 participants with the inclusion criteria of Taiwanese women aged 20 to 40 years who had experience in giving birth. Task testing comprised 12 items covering all functions provide by the app. All of the participants completed all tasks. The participants were asked to perform thinking aloud during the task testing period. On the basis of the feedback to the thinking aloud process during the test, interface operation issues were categorized into three types of errors, namely, navigation error, presentation error, and control error by the nature of errors and severity regarding how the errors influenced task completion. The overall average score of the SUS was 70.5, higher than the designated standard average of the SUS. Therefore, the satisfaction toward We'll met the standard level. After the aforementioned procedure, usability task testing results were organized to determine the following problems that need to be solved. (a) Redesign the house icons of friends, family, and hospital on the home page (b) Enlarge the click area of friend and family rows (c) Improve the presentation of the EDPS test results, change the histogram into numeral score feedback, and add suggestions and reminders (d) Establish a dynamic tutorial for first-time login users.

Keywords: Postpartum emotional disorder · Postpartum blues ·
Postpartum depression · Usability test · Mobile application

1 Introduction

Nowadays, technology development changes rapidly. Due to the advance of technology and medicine, we are able to cure and control many diseases. However, the flourishing modern development also brings modern plague, depression, which is difficult to be solved. The advanced and ever-changing modern development drives convenience and quality in our life but brings invisible and heavy burden to people in three aspects- physiology, psychology, and society. As a result, depression is gradually developed without being noticed. Currently, there are around 4% of total populations in the world suffering from depression and it has become one of the most common and most difficult to prevent mental diseases all over the world. In addition, the impact comes along with depression is more than what we generally recognize as feeling down in spirits or more seriously, the intention of undervaluing their own lives and committing suicide. It will invade our daily life and slowly cause sufferers losing ability and vitality in life and further lead to the status of unable to work or unable to take care of themselves in daily life. Other than corroding sufferers' psychological and physical functions, the productivity and operation in social and economic levels are reduced due to the growth of depression on a large scale. It makes depression be regarded as one of the most urgent issues that needs to be overcome.

In terms of gender, the percentage of female suffering from depression is twice than male. Pregnant or postpartum women are one of the groups with high-possibility depression suffering because most women experience drastic hormone changes during menstrual period, pregnancy and prenatal period and menopause; it will further affect emotional response. Among them, it is often seen the women at pregnancy and prenatal period suffer from postpartum emotionally handicapped. The most common symptom for postpartum emotionally handicapped is postpartum blues (baby blues) and postpartum depression (PPD). Around 50% to 80% parturient women suffer from emotional instability or temporary postpartum depression, and most parturient women who encounter the symptom will be able to improve in the short time with proper social support (including husband, family, and friends) and sufficient assistance. If the symptom of depression in low spirits lasts more than two weeks, it might cause postpartum depression. Its prevalence rate is around 10–15% (Cox et al. 1993; Darcy et al. 2011; Gavin et al. 2005).

Due to the rapid development of technology, E-mental health (EMH) service has become the application trend for current non-medication treatment. Professional functions and effective information are provided through online social media to allow users improving their mental health problems. The prevalent smart phones and tablet devices contribute to the transformation of EMH into the mode of mobile health (mHealth) service. Combining EMH with social support is one of the important links to effectively improve and reduce the risk of parturient women suffering from postpartum depression. Previous research (Kao 2015) had established an EMH system in the form of APP- We'll, Wishing Well. The system focuses on boosting parturient users' social support and self-esteem (Shaw and Gant 2002) and provides the service of self-check with Edinburgh Postnatal Depression Scale (EPDS) (Shen et al. 2015) in order to

further relieve and prevent postpartum emotionally handicapped and depression problems.

The symptom of postpartum emotionally handicapped will usually disappear automatically in several days after delivery. Few serious cases might be involved with postpartum depression (prevalence rate is around 10–15%) and parturient women might have the intention to commit suicide and the risk of injuring the newborn. However, the symptom usually gets worse and worse because parturient women may not able to be aware of the symptom or are ignored by family, relatives, and friends so that they fail to offer timely proper assistance and support.

Based on the principle of "an ounce of prevention is worth a pound of cure", the early the status of postpartum emotionally handicapped can be found, the better the effectiveness of intervention will be. Previous research (Kao 2017) had developed APP system to support postpartum emotionally handicapped according to the theories and concepts of postpartum emotionally handicapped. A mobile application program structure, "We'll, Wishing Well", has been established and the system uses the interaction mode of "making a wish" as the main activity for users to making a wish to Wishing Well with an expecting and positive attitude. The wishes are solved and fulfilled through the efforts of the online APP users to enhance users' mental status and quality. We'll get you well; therefore, the APP is named as "We'll, Wishing Well" (Fig. 1).
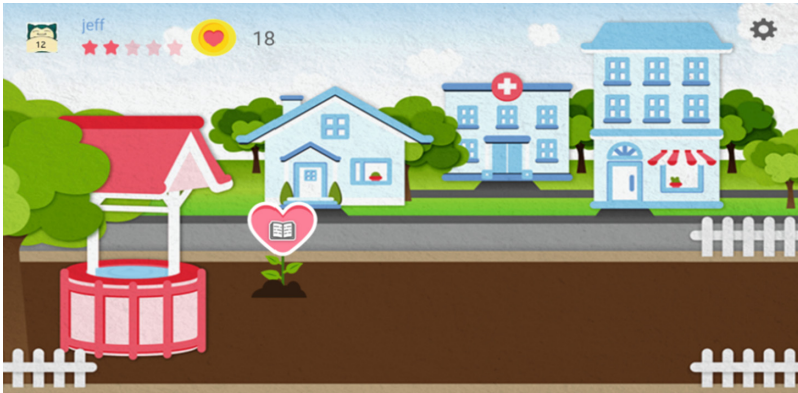
Fig. 1. The homepage of APP "We'll"

The system of the previous research is a version of simulation testing. Therefore, the research evaluates three performance evaluation of effectiveness, efficiency, and satisfaction through usability testing based on the current framework of We'll. The problems related to APP operation or interface are improved and modified through the results of usability testing and further advance its effectiveness, efficiency, and satisfaction.

2   Literature Review

2.1   Postpartum Emotionally Handicapped

Women at pregnancy and prenatal period might have drastic hormone changes before and after delivery and cause the symptom of postpartum emotionally handicapped; its prevalence rate is around 40%–80% (Evins and Theofrastous 1997). Other than the impact from physiological conditions, parturient mental status might also be affected by external pressure, such as pressure of childcare, marital relationship, work and family (Stowe and Nemeroff 1995). These lead to the extension of postpartum emotionally handicapped symptom.

Postpartum emotionally handicapped is the most common issue on women at postpartum period and it can be divided into three stages according to the duration of the illness and the seriousness of the illness. They are postpartum blues (maternity blue or baby blues), postpartum depression (PPD) and postpartum psychosis respectively (Robertson et al. 2004).

Postpartum blues is the most common situation of postpartum emotionally handicapped and its incidence rate is around 30%–75% (O'hara et al. 1984). It usually occurs in a few days after delivery and the symptom includes emotional instability, easy to lose temper, sleep disorder and depraved appetite. Relevant emotional and physiological disorder symptoms are usually minor and they will be self-healed gradually in a couple of days; no medication or other treatment will be required (Kennerley and Gath 1989; Pitt 1973). If the symptom lasts for more than two weeks but is not getting better, it is likely to trigger postpartum depression. It is the major depression under non-psychosis with the prevalence rate of 10%–15% (Cox et al. 1993; O'hara et al. 1991; Wisner et al. 2002). Its symptom is slightly similar to maternity blues and parturient women might feel valueless, lose interests towards life or got bored of life. The cycle of the symptom usually lasts for more than two weeks or several months and the sufferers might have the intention of committing suicide or the risk of injuring the newborn if the level of the illness is serious. There are around 20% parturient women ending their lives by committing suicide due to the postpartum depression (Lindahl et al. 2005), and it requires medical assistance and care. Postpartum psychosis is the most serious and rare illness among the postpartum emotionally handicapped symptom and the probability is one to two in one thousand (around 0.1% to 0.2%) (O'hara et al. 1984). It might cause significant emotional instability on parturient women and even involve with the symptom of reality deviation, including illusion or delusion (Brockington et al. 1981). Therefore, it usually requires medical assistance and hospitalized for observation and treatment.

2.2   Current Status and Development of Mobile Health and E-Mental Health

The original concept of mobile health (mHealth) is to use mobile medical instrument or equipment in the medical clinic to provide patients more convenient medical behaviors of inquiry and health examination. The development until now is that the mobile health has used information and communication technology, such as computer, mobile phone,

and communications satellite, to provide broader medical services and information. The current scope of mobile health includes (1) mobile medical equipment installation and medical services in the hospital and (2) distance medical care and health care outside the hospital.

E-mental health (EMH) is the trend of current mental medical care service. It provides professional functions to solve the problems of mental health. Following four types of medical services are provided through internet via the media of social media, chat room, forum, bulletin board system, and blog: (1) information related to medical care (Lambousis et al. 2002; Santor et al. 2007), (2) illness screening, evaluation, and monitoring (Becker et al. 2008; Chinman et al. 2007; Diamond et al. 2010; Donker et al. 2009; Gringras et al. 2006; Gualtieri 2007; Heron and Smyth 2010; Khazaal et al. 2009), (3) medical intervention (Bergström et al. 2010; Khanna and Kendall 2010; Lindahl et al. 2005), and (4) social support (Scharer 2005).

With the advance of technology and the popularity of smart phone, the service format of e-Health is pushed towards to the service format of mobile health APP. Doctors and patients can carry out self-health management or remote patient information control through relevant application programs in the mobile phone. According to the survey of current mobile medical APP done by WHO, the search of depression is ranked the second, only after diabetes (Martínez-Pérez et al. 2013). Therefore, it will definitely be able to increase the intention of APP use and satisfaction of APP if the functions of mobile medical APP, We'll, Wishing Well, can be effectively integrated as well as the usability of application programs can be advanced.

2.3   Usability Evaluation of Interactive System

ISO 9241 defines usability as a product or service that can be used by specific users efficiently and effectively to achieve a specific target or complete the task. Usability also means the method used during the process of design to advance the level of easy-to-use for products or services.

Nielsen (1994) proposed usability is mainly to explore how to communicate with users, observe users' working environment and conduct scenario analysis in order to find out problems related to product usability. Usability is to allow representative users to operate products or experience services as well as execute a series of tasks in order to test the system. When conducting usability testing, it should follow three indicators of effectiveness, efficiency, and satisfaction as the standards for practical evaluation. Usability evaluation can be divided into the following five elements (Nielsen 1994; Shneiderman and Plaisant 2005): (1) learnability, (2) efficiency, (3) memorability, (4) errors and (5) satisfaction.

Usability scenario testing has been widely applied to the processes of product evaluation and system development. Usability evaluation for system interface must find out the characteristics of the system first before following the characteristics to determine the attributes (such as learnability, errors, and efficiency) and functions that are to be evaluated.

The measurement standard used to execute usability testing can be classified into the following indicators (Albert and Tullis 2013; Sauro and Lewis 2016). Task success: whether users complete the task; task time: the time users spend to complete the task;

error: unexpected errors when users execute the task, such as missing or pressing the wrong button, and errors sometimes delay the execute time or even lead to the failure of the task; efficiency: the dimension for efficiency evaluation is wide and it can be defined as the possession of higher efficiency if the product allows users to complete the task in the shorter time with lower occurrence of error; learnability: beginners knowledge towards products from unfamiliar to professional is the learning process of the user's. The level of difficulty and the speed of familiarity of the process are regarded as learnability.

The evaluation elements for usability mentioned above and the indicators for usability testing vary according to using status of different products or services, content of testing, or target users of the final purpose to make choices of testing standards in order to enhance the effectiveness of usability testing.

2.4   Relevant Research for APP Usability Evaluation

From above, usability testing evaluation can set up the evaluation indicators according to different product demands and the three indicators of effectiveness, efficiency, and satisfaction mentioned in ISO 9241-11 are used as the standards for practical evaluation (Gunter et al. 2016; O'Malley et al. 2014; Mirkovic et al. 2014). According to Gunter's research (2016) on usability evaluation and development of APP: WoundCheck that provides patients after surgery the service of monitoring wound after surgery, main function of the APP is to provide users photo records. The final target for the design is to monitor wounds after users being discharged from hospital and during return visits. The evaluation of usability done in this study is to use user's operation on six tasks and system usability scale (SUS) as standards for satisfaction evaluation. It is conducted in two stages; the functional problems found after the first stage test will be slightly adjusted and the version of the revised function will be tested the usability evaluation in the second stage. The APP version after improvement in the second stage is with better using efficiency and task completion rate than the version in the first stage. Therefore, it approves an APP that combines mobile phone application program with wound monitoring after surgery is the design that meets effectiveness of usability and the expectation from users.

Mirkovic et al. (2014) conducted a research of usability evaluation on APP: Connect Mobile, that implements disease management on patients with cancer. It tests usability evaluation on general smart phone and tablet device to compare APP usability and difference between the two devices. After completing usability task testing and semi-structured interview, the research reveals the difference of usability test on smart phone and tablet device. Users' effectiveness and efficiency in completing the task will be affected by different devices used. Therefore, the research suggests different APP system versions should be designed for different devices.

Lim et al. (2015) aimed to develop a mobile phone APP that is to be used for medical work at the areas with insufficient medical resources for immediate toxemia of pregnancy on the parturient women in the area. The study is a usability evaluation test done in two stages, and it is assessed with scenario testing, think aloud, and questionnaire survey. In order to classify the usability errors found during the testing, the study defines errors at usability evaluation testing as (1) navigation errors,

(2) presentation errors, and (3) control usage errors. The order of priority usability problem-handling will be determined by the frequency and seriousness of the error occurred.

English et al. (2016) investigated APP development for child care at areas with limited medical resources in order to reduce child mortality and implemented usability testing on the APP system. The users of the system are set medical practitioners with different education background as default. The testing process is divided into two stages and time and number of errors to complete the task by respondents are collected during the testing. In order to define the reason of error and improve usability problems effectively, The study classifies the error of the task into three types: (1) navigation errors, (2) control usage errors, and (3) outcome errors. The levels of seriousness according to the impact on the task are (1) low severity: it won't affect users completing the task but will affect the using efficiency and satisfaction; (2) medium severity: users have problems of operation when executing the task and it will affect effectiveness, efficiency, and satisfaction, and high severity: users will fail to complete the task. The study defines and classifies the type of error in order to provide assistance to subsequent improvement and optimization on the APP system.

Currently, a lot of APP systems use usability testing to evaluate its system and verify system effectiveness, efficiency, and satisfaction through the combination of task evaluation and questionnaire. From literature review, it is found defining the characteristics of operational errors and the seriousness will be able to effectively contribute to the subsequent functional improvement on the system. Therefore, evaluation methods mentioned on above literatures will be referred for the usability evaluation on our research.

3   Methods

3.1   Experimental Design

The research uses usability task testing to provide respondents user experience of participating in real situation. Through data of usability testing and respondents' feedback on the interface and usability problems of the current We'll APP version, they are used as the basis of subsequent modification on APP in the future.

In order to meet the standardization of the equipment used, the usability testing uses unified ASUS ZenPad 8.0 Tablet for testing. Canon 760D single lens reflex camera and Lumix LF-7 bridge camera are used for behind-the-scenes footage during usability testing and recording interface process of operational scenario tasks. The research is divided into four stages. The first stage is research introduction to explain the procedure and purpose of usability testing, basic operation of the interface, and scenario task as well as allows respondents to get familiar with the operation of ASUS ZenPad 8.0 Tablet. The second stage is the usability scenario test done by respondents and respondents are asked to think aloud during the process in order to encourage respondents to express their thinking process during the testing. In the third stage, satisfaction evaluation questionnaire will be carried out after completing scenario testing, including demographic questionnaire to record the respondent's basic

information and system usability scale (SUS) to evaluate the usability satisfaction on APP. In the final stage, an interview with five post-testing qualitative questions will be done on respondents.

Process of usability testing includes the following four stages: (1) respondent recruitment, (2) usability task testing, (3) satisfaction questionnaire survey, and (4) post-testing qualitative interview; as shown on Fig. 2.
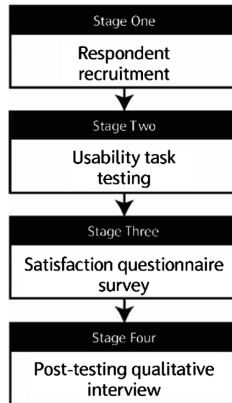
Fig. 2. Usability testing flow chart

The target of usability testing is (1) to test under the condition of scenario task to understand the common operational errors during the process; (2) to find out satisfaction on APP through the questionnaire survey; (3) to guide respondents providing feedback on usability and satisfaction on the APP during interview. The problems on operational process or interface of the APP that are found through the data and feedback collected from respondents during the testing will be modified and improved immediately.

During the implementation of the scenario task, the number of times that respondents need to be reminded and the type of error during the task implementation will be classified and recorded. The reminding standards and definition of each type of error are explained as below:

(1) Remind: The delay of operational process caused by the same step operated by the respondent for more than two times during the testing or the assistance requested by the respondent actively will generate the action of reminding; the reminded operational problems and types should be recorded.

(2) Navigation error: The function provided by the interface that fails to guide the respondent to achieve the function or feedback pre-set during the implementation of the scenario task and causes time delay of implementation will be classified as navigation error.

(3) Presentation error: The icon problems on the interface that cause the respondent fails to judge the function or makes wrong judgment and further delay the task will be classified as presentation error.

(4) Control usage error: If the time delay is caused by respondent being not able to enter the demand message during the testing or executing wrong functions and might further affect the completion of the task, it will be defined as control usage error.

3.2 Respondent

The research recruits 30 respondents for the experiment. The recruitment conditions are (1) women that have pregnant and delivery experience and have children, (2) women with age between 18 and 45 years who are still planning for pregnancy or have opportunities to give birth, and (3) women who have smart phone and experience of using it for more than four years as well as the habit of using tablet device in the daily life.

Exclusion criteria on the respondent includes (1) women with age more than 45 years old, (2) women who have pregnant experience but without delivery experience, and (3) pregnant women with foreign nationality or who cannot read Chinese due to the current system and scale are all presented in Chinese.

3.3 Introduction of Usability Scenario Task Testing

In order to test the usability and effectiveness of We'll, Wishing Well, respondents are recruited for scenario task testing. Respondents must complete 12 preset tasks during the test, and the content of the tasks covers all of the functions provided by APPP. The step and content of the scenario tasks are as shown on Table 1.

3.4 System Usability Scale (SUS)

The study uses system usability scale (SUS) developed by John Brooke in 1986 to measure the respondent's subjective perception on operational interface and usability.

4 Results and Discussion

4.1 Analysis of the Respondents

The study recruited 30 respondents in total, and case enrollment was from April 30th, 2018 to September 11th, 2018. 90% of the respondents were primipara, 60% of them used aided APP during pregnancy, and all of the respondents had the experience of using smart phone for more than 4 years.

4.2 Analysis of Scenario Task Testing

Table 2 shows the number of times of reminding when executing the task and the number of total errors of each task during the usability testing of APP- We'll, Wishing

Table 1. Content of usability scenario task

Variable	Content of the task
Task 1	Log in APP with FB account and enter personal details
Task 2	Take a screenshot of the first emotional state test result done on APP (Edinburgh questionnaire)
Task 3	Click houses of friends' and search friends from the left column to add one good friend
Task 4	Click your house and search family members from the left column to add one family member
Task 5	Throw a gold coin to the wishing well, make a wish for action assistance, and choose to open to the public
Task 6	Throw a gold coin to the wishing well, make a wish for expressing emotion, and choose not to open to the public. (Choose to make the wish to either family member or friend)
Task 7	Click houses of friends' and search friends from the left column and select to deliver the gold coin to the friend in the first priority
Task 8	Click houses of friends' and search friends from the left column. Select the friend in the first priority and select the function of visiting good friends to reply his/her wish
Task 9	Clink your house and search family members from the left column. Select the family member in the first priority to enter setup to remove family relationship
Task 10	Click knowledge of health education in the hospital, select notice of postpartum care center, answer questions after reading the notice to obtain the gold coin, and take a screenshot
Task 11	Throw a gold coin to the wishing well, make a wish for specific experience and knowledge, and choose to make it public. Upload the screenshot from task 2 and task 10
Task 12	Log out App

Well. The 30 respondents recruited at the stage all completed the entire scenario task testing during the tested period, and the final statistics revealed 93 times of task reminding were executed as well as 8 times of navigation error, 50 times of presentation error, and 6 times of control error were accumulated.

The reminding standard used for the usability testing in the research was the action of reminding would be done when the delay of the operational process was caused by the respondent operating the same step for more than two times or respondent requesting assistant actively during the task testing. During the testing, most of the respondents were not able to distinguish the houses of friends' and their own smoothly and it caused the delay of the task execution while some respondents said they forgot which icon of house on EDPS questionnaire should be clicked to guide the function. Through the observation during the testing, it revealed one of the key reasons that respondents need reminding is the difficulty in identifying the icon of house function.

Presentation error was the more frequent type of error occurred during the testing, and it could be imputed to task (3) click houses of friends' and search friends from the left column to add one good friend and task (4) click your house and search family

Table 2. Record of times of each type of error

Task	Remind	Navigation error	Presentation error	Control error
1	10	0	2	3
2	11	0	7	0
3	18	3	15	0
4	15	0	9	0
5	10	2	4	1
6	0	0	0	0
7	4	1	1	0
8	7	0	5	0
9	4	0	0	1
10	2	0	4	0
11	11	2	3	1
12	1	0	0	0
Total	93	8	50	6

members from the left column to add one family member. There were 18 respondents reflected the houses of friends and their own on the homepage of APP were with very low identifiability when executing the two tasks. It caused confusion and led to misjudgment during task execution. Homepage of We'll, Wishing Well is as Fig. 3.
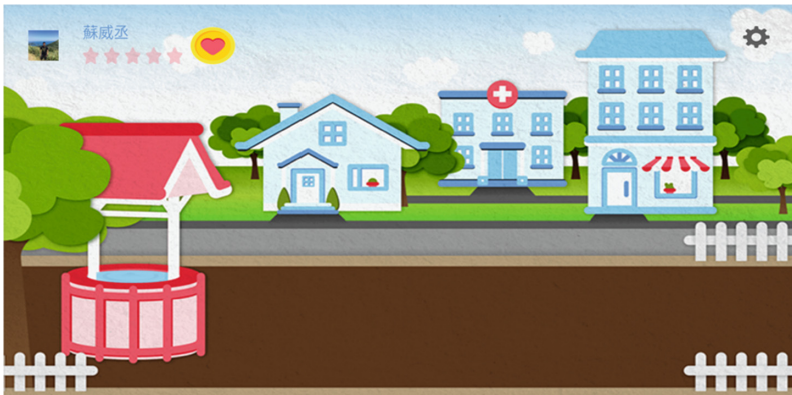
Fig. 3. Homepage of We'll APP (the first one on the left is user's own house, middle one is hospital, and the first one on the right is friend's house)

There were 8 times of navigation error in the testing. Respondents expressed the design for menu clicking among friends and family members were not intuitive enough at task (3) click houses of friends' and search friends from the left column to add one good friend. The design on APP was to click the name of the friend or family member to show the window of functional option, but respondents repeat-clicked the icon of

friends or family members with intuition during the operation. Therefore, APP system was not able to show the window of functional option. The interface of friend list on We'll, Wishing Well is as shown on Fig. 4.

Fig. 4. Interface of friends on We'll APP (the left column is the list of friends)

There were 6 times of control error, including 4 of the respondents were students but there was no option of student on the menu design for occupation on APP when they first logged in and required to enter the personal details. As a result, they used service industry as an alternative and it was classified as control error because the correct information was not able to be entered. In addition, many respondents reflected after completing Edinburgh Questionnaire (EDPS) for task (2) that the test result used bar chart and color for identification (green: low severity; yellow: medium severity; red: high severity) and they were not able to precisely interpret the testing results and reminding of the questionnaire; as shown on Fig. 5.
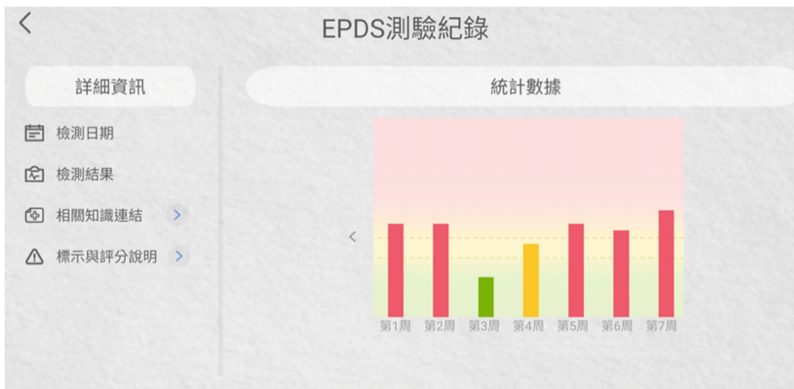

Fig. 5. Interface of post-testing records for Edinburgh Questionnaire (EDPS) on We'll APP (Color figure online)

4.3   Analysis of Scenario Task Execution Time

Average time used to complete each task is shown on Table 3.

Table 3. Table showing time used for each scenario task

Task	Mean	SD	Min	Max
1	123.83	47.36	66	280
2	64.9	18.95	36	129
3	51.33	17.81	30	103
4	35.17	12.99	15	78
5	44.57	13.23	26	75
6	39.23	10.57	16	56
7	34.17	18.44	9	118
8	42.23	27.83	16	121
9	29.6	10.55	13	53
10	50.6	14.09	23	82
11	51.53	15.72	32	87
12	21.93	12.43	6	58
Total	49.09	10.31	6	280
				Time (sec)

4.4   Analysis of Satisfaction on System Usability

The average point of the system usability scale (SUS) questionnaire done by the 30 respondents in the study is 70.5 and it is greater than the average SUS point of 68, which shows the satisfaction on We'll meets the standard. The simulation testing version of SUS questionnaire points on the previous research is 78.54. The points for both simulation version and the Android version used for the study fall in the interval of level C. Therefore, there is no significant difference of usability evaluation between the two systems and they both achieve the point better than average, which represent the system is affirmed and recognized by users.

4.5   Discussion

The study conducts usability evaluation on We'll mobile APP version to verify the usability of its APP version and find out the operational problems on the interface or functions. According to the data of the testing results, the problems to be improved on We'll are analyzed and summarized as well as correction proposals are proposed and executed in order to achieve the target of APP interface and functional design optimization.

The previous developed We'll system framework according to social support and established the version of simulator. The research only tested usability testing on scenario task on the three functions provided: (1) real-time testing: providing Edinburgh questionnaire testing, (2) formal support: providing professional health education

knowledge, such as notice for maternal women, and (3) non-formal support: providing the function of wishing well to advance social support through making wishes. The research focuses on more details and designs 12 tasks for usability scenario testing in terms of the main functions provided by APP, including basic information, navigation function, interface operation, and We'll. A lot of usability data and interface operational problems that are not known from the previous research are found out through scenario task test.

Depression issues and social impact are emphasized more and more in recent academic research. "Depression Monitor" APP established by Nasser (2012) has completed the development and can be downloaded free of charge on APP store provided by Apple; as shown on Fig. 6. Depression Monitor APP provides depression evaluation scale (Patient Health Questionnaire-9, PHQ-9) to allow users using PHQ-9 to test their depression status through APP. Furthermore, the previous research on depression issues focused on postpartum depression women and developed We'll APP to introduce the most common Edinburgh postnatal depression scale to screen out postnatal women's mental state. Both Depression Monitor and We'll APP provide scale screening tool for the area but Depression Monitor only provides scale screening, integration, and transmission through one-way information (that is, users fill out the scale and submit it to the system for system to provide the feedback to users). However, We'll establishes an interactive social media platform between puerperal and supporters based on social support framework. It presents interactive interface of making wishes through play as well as provides professional child-rearing health education knowledge to puerperal so that they can obtain evaluation screening, professional information and emotional support through We'll.
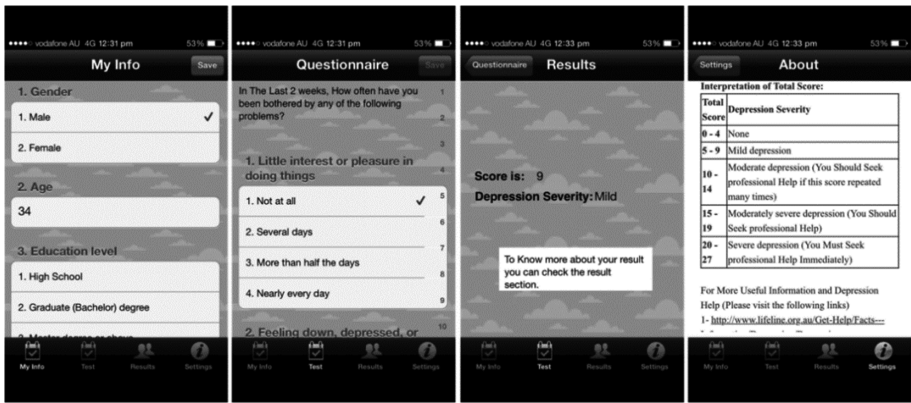

Fig. 6. "Depression Monitor" APP (BinDhim et al. 2014)

Overall, respondents completed all the scenario tasks without any major error and it revealed the competence of technology on the current systematic framework of We'll meets the standard (O'Malley et al. 2014). The mutual verification by the number of times of reminding and the type of errors during task execution found out the usability

problems of the interface (Hong et al. 2014; Lim et al. 2015; Gunter et al. 2016) while the operational problems defined by navigation error, presentation error, and control error during task execution were effective for the subsequent modification proposals proposed (Mirkovic et al. 2014; Hong et al. 2014; English et al. 2016). Besides, the overall average point of system usability scale (SUS) is 70.5 and it is clearly higher than the average point of the scale (Brooke 1996), showing a certain level of satisfaction on the usability of We'll from respondents.

Other than the discussion above, respondents also provided suggestions and feedback to We'll through post-test qualitative interview, such as (1) proactive dynamic interface-using teaching can be established for the first-time logging in to strengthen users' familiarity towards interface; (2) adding pregnancy records to the function of hospital icon, including baby's ultrasonography images and records of prenatal visit; (3) establishing online professional counseling for obstetrics & gynecology to solve puerpera's questions in real time on line; (4) producing dynamic feedback when delivering the gold coin to friend or family member in order to strengthen the link between the user and friend or family member; (5) designing the linking function for dad so that both puerperal users and their husband can link to the function of reminding with the APP account for husbands to be reminded the using status (making a wish, the point after emotional state test, and its evaluation) after completion; (6) function of wishes made can be added to the function of wishing well to allow users to review the status and record of the wishes in the past.

The next stage of the research will conduct modification on We'll based on the usability testing result this time as well evaluate users suggestions to add new functions in order to enhance the usability and effectiveness of We'll so that the target users in the future will have higher evaluation and satisfaction towards We'll.

5   Conclusions and Suggestions

The research verifies whether the design of interface and function meet usability and satisfaction through the data of effectiveness and satisfaction collected during usability testing execution, and the testing data collected and respondents' feedback will be used as the reference to optimize APP function and interface. 30 respondents all successfully completed 12 tasks during scenario testing and the problems occurred during operation were smoothly solved after task reminding. Therefore, it can be seen that the usability and effectiveness of APP functional framework has achieved a certain level of use. The result measured by system usability scale (SUS) also revealed the respondents' satisfaction towards current version also meets the standard of system satisfaction. The usability testing for this stage has been completed and the testing data & suggestions provided by respondents during testing will be applied for problem modification and function optimization on the current version. A revised version will be launched to enhance APP usability and satisfaction from target users.

The results of usability task testing show the problems of We'll APP interface or functions, and it is suggested to make the following improvement and modification for subsequent research:

1. Re-designing the appearance of house icon for friends, family member, and the hospital on the home page to strengthen user's identifiability and effectively enhance user's recognition towards icon function.
2. Expanding reflection area on the list of friends and the list of family members to make sure the window of functional option will be shown when users click the icon or name of their friends' or family members'.
3. Improving the presentation of EDPS testing result to change the bar chart to actual points as well as add suggestion reminding in order to strengthen users' emotional states and understanding towards questionnaire feedback.
4. Establishing dynamic interface functional using teaching when users log in for the first time to strengthen users' connection and functional familiarity towards the APP.

References

Albert, W., Tullis, T.: Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics. Newnes, Oxford (2013)

Becker, J., et al.: Functioning and validity of a computerized adaptive test to measure anxiety (A-CAT). Depress. Anxiety **25**(12), E182–E194 (2008)

Bergström, J., et al.: Internet-versus group-administered cognitive behaviour therapy for panic disorder in a psychiatric setting: a randomised trial. BMC Psychiatry **10**(1), 54 (2010)

BinDhim, N.F., Shaman, A.M., Trevena, L., Basyouni, M.H., Pont, L.G., Alhawassi, T.M.: Depression screening via a smartphone app: cross-country user characteristics and feasibility. J. Am. Med. Inform. Assoc. **22**(1), 29–34 (2014)

Brockington, I.F., Cernik, K.F., Schofield, E.M., Downing, A.R., Francis, A.F., Keelan, C.: Puerperal psychosis: phenomena and diagnosis. Arch. Gen. Psychiatry **38**(7), 829–833 (1981)

Brooke, J.: SUS-A quick and dirty usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)

Chinman, M., Hassell, J., Magnabosco, J., Nowlin-Finch, N., Marusak, S., Young, A.S.: The feasibility of computerized patient self-assessment at mental health clinics. Adm. Policy Ment. Health Ment. Health Serv. Res. **34**(4), 401–409 (2007)

Cox, J.L., Murray, D., Chapman, G.: A controlled study of the onset, duration and prevalence of postnatal depression. Br. J. Psychiatry **163**(1), 27–31 (1993)

Darcy, J.M., Grzywacz, J.G., Stephens, R.L., Leng, I., Clinch, C.R., Arcury, T.A.: Maternal depressive symptomatology: 16-month follow-up of infant and maternal health-related quality of life. J. Am. Board Fam. Med. **24**(3), 249–257 (2011)

Diamond, G., et al.: Development, validation, and utility of internet-based, behavioral health screen for adolescents. Pediatrics (2009)

Donker, T., van Straten, A., Marks, I., Cuijpers, P.: A brief Web-based screening questionnaire for common mental disorders: development and validation. J. Med. Internet Res. **11**(3), e19 (2009)

English, L.L., et al.: The paediatric risk assessment (PARA) mobile app to reduce postdischarge child mortality: design, usability, and feasibility for health care workers in Uganda. JMIR mHealth uHealth **4**(1), e16 (2016)

Evins, G.G., Theofrastous, J.P.: Postpartum depression: a review of postpartum screening. Prim. Care Update Ob/Gyns **4**(6), 241–246 (1997)

Gavin, N.I., Gaynes, B.N., Lohr, K.N., Meltzer-Brody, S., Gartlehner, G., Swinson, T.: Perinatal depression: a systematic review of prevalence and incidence. Obstertrics Gynecol. **106**(5, part 1), 1071–1083 (2005)

Gringras, P., Santosh, P., Baird, G.: Development of an Internet-based real-time system for monitoring pharmacological interventions in children with neurodevelopmental and neuropsychiatric disorders. Child Care Health Dev. **32**(5), 591–600 (2006)

Gualtieri, C.T.: An internet-based symptom questionnaire that is reliable, valid, and available to psychiatrists, neurologists, and psychologists. Medscape Gen. Med. **9**(4), 3 (2007)

Gunter, R., et al.: Evaluating patient usability of an image-based mobile health platform for postoperative wound monitoring. JMIR mHealth uHealth **4**(3), e113 (2016)

Heron, K.E., Smyth, J.M.: Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br. J. Health. Psychol. **15**(1), 1–39 (2010)

Hong, Y.: Testing usability and acceptability of a web application to promote physical activity (iCanFit) among older adults. JMIR Hum. Factors **1**(1), e2 (2014)

Kennerley, H., Gath, D.: Maternity blues: I detection and measurement by questionnaire. Br. J. Psychiatry **155**(3), 356–362 (1989)

Khanna, M.S., Kendall, P.C.: Computer-assisted CBT for child anxiety: the coping cat CD-ROM. Cogn. Behav. Pract. **15**(2), 159–165 (2008)

Khanna, M.S., Kendall, P.C.: Computer-assisted cognitive behavioral therapy for child anxiety: results of a randomized clinical trial. J. Consult. Clin. Psychol. **78**(5), 737 (2010)

Khazaal, Y., et al.: Brief DISCERN, six questions for the evaluation of evidence-based content of health-related websites. Patient Educ. Couns. **77**(1), 33–37 (2009)

Lambousis, E., Politis, A., Markidis, M., Christodoulou, G.N.: Development and use of online mental health services in Greece. J. Telemed. Telecare **8**(2_suppl), 51–52 (2002)

Lim, J., et al.: Usability and feasibility of PIERS on the Move: an mHealth app for pre-eclampsia triage. JMIR mHealth uHealth **3**(2), e37 (2015)

Lindahl, V., Pearson, J.L., Colpe, L.: Prevalence of suicidality during pregnancy and the postpartum. Arch. Women's Ment. Health **8**(2), 77–87 (2005)

Martínez-Pérez, B., De La Torre-Díez, I., López-Coronado, M.: Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis. J. Med. Internet Res. **15**(6), e120 (2013)

Mirkovic, J., Kaufman, D.R., Ruland, C.M.: Supporting cancer patients in illness management: usability evaluation of a mobile app. JMIR mHealth uHealth **2**(3), e33 (2014)

Nasser, F.B.: Health Monitor Project. Smart Health Project (2012)

Nielsen, J.: Usability Engineering. Elsevier, Amsterdam (1994)

O'hara, M.W., Neunaber, D.J., Zekoski, E.M.: Prospective study of postpartum depression: prevalence, course, and predictive factors. J. Abnorm. Psychol. **93**(2), 158 (1984)

O'hara, M.W., Schlechte, J.A., Lewis, D.A., Varner, M.W.: Controlled prospective study of postpartum mood disorders: psychological, environmental, and hormonal variables. J. Abnorm. Psychol. **100**(1), 63 (1991)

O'Malley, G., Dowdall, G., Burls, A., Perry, I.J., Curran, N.: Exploring the usability of a mobile app for adolescent obesity management. JMIR mHealth uHealth **2**(2), e29 (2014)

Robertson, E., Grace, S., Wallington, T., Stewart, D.E.: Antenatal risk factors for postpartum depression: a synthesis of recent literature. Gen. Hosp. Psychiatry **26**(4), 289–295 (2004)

Santor, D.A., Poulin, C., LeBlanc, J.C., Kusumakar, V.: Online health promotion, early identification of difficulties, and help seeking in young people. J. Am. Acad. Child Adolesc. Psychiatry **46**(1), 50–59 (2007)

Sauro, J., Lewis, J.R.: Quantifying the User Experience: Practical Statistics for User Research. Morgan Kaufmann, Burlington (2016)

Scharer, K.: An Internet discussion board for parents of mentally ill young children. J. Child Adolesc. Psychiatr. Nurs. **18**(1), 17–25 (2005)

Shaw, L.H., Gant, L.M.: In defense of the internet: the relationship between Internet communication and depression, loneliness, self-esteem, and perceived social support. CyberPsychology Behav. **5**(2), 157–171 (2002)

Shen, N., et al.: Finding a depression app: a review and content analysis of the depression app marketplace. JMIR mHealth uHealth **3**(1), e16 (2015)

Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Pearson Education Inc., Upper Saddle River (2005)

Stowe, Z.N., Nemeroff, C.B.: Women at risk for postpartum-onset major depression. Am. J. Obstet. Gynecol. **173**(2), 639–645 (1995)

Wisner, K.L., Parry, B.L., Piontek, C.M.: Postpartum depression. N. Engl. J. Med. **347**(3), 194–199 (2002)

高鈞瑩: 以社會支持發展產後憂鬱症及產後情緒低落輔助APP之研究與設計 (未出版之碩士論文) 長庚大學,桃園市 (2017)