Usability Tool Evaluation using Usability Testing Method at PT Media Dokter Investama (Halodoc Indonesia)

Selvi and Rienna Oktarina
Industrial Engineering Department
Bina Nusantara University
Jakarta, Indonesia, 11480
selvi001@binus.ac.id, rienna.oktarina@binus.ac.id

Abstract

Halodoc Indonesia Catalog Division Staff uses a "tool" in carrying out the work process. This "tool" is undoubtedly a very primary and essential tool for the Catalog Division Staff. This high level of need against the "tool" creates a great enthusiasm to observe deeper into everything from this "tool" like how the tool works, how the staff uses the tool, and how the level of usability tool is. About how the tool works and how the staff uses the tool are known by participating in the staff's work. The usability level of the tool is known by using the Usability Testing Method, where the method can calculate the usability level of four components, such as 1) Learnability, 2) Efficiency, 3) Errors, and 4) Satisfaction. Based on the type of work, in general, Catalog Division performs 3 types of work, such as 1) Create doctor data and service products, 2) Create hospital data, and 3) Create and drop schedules also manage doctors and service slots. The Learnability component is represented based on the Success Rate in the amount of 76%. The Efficiency component is represented based on the Time-Based Efficiency, in the amount of 0.002 goals/sec. The Errors component is represented based on the Error Rate, in the amount of 0.6. Based on the results, it turns out that further review is needed and at this stage, found the addition of 20 issues from the tool and proposed 20 improvement recommendations to the tool on these issues. The proposed improvement recommendations are then measured further by calculating the comparison of satisfaction level between before and after the improvement recommendations are conceived. The satisfaction level measurement is carried out to know concretely whether these recommendations have a major effect on the satisfaction level or not. The results will affect future work quality.

Keywords
Usability, Usability Testing, Halodoc, Usability Tool, Tool Improvement.

1. Introduction

The "Tool" used by the Division Halodoc Catalog Indonesia (tool's name can not be mentioned) is a database tool special for internal parties used to input the required data. Data related to the Catalog Division includes data from doctors, hospitals, and merchants (such as COVID-19 and non-COVID-19 vaccines). The reason this tool is the focus of this research is because of the frequent complaints from users (Catalog Division Staff) when they are given assignments by outside parties. This certainly raises a sense of curiosity in finding further about what is the root of this problem.

A review is immediately carried out in advance of finding what the Catalog Division actually does, are there any tools that are usually used by the division, and how the role of this division on the entire membership system in the company. The review results that the Catalog Division do a daily data input (data for doctors, hospitals, and merchants - such as COVID-19 vaccine service products and non-COVID-19 vaccines); then second is inputting doctor's schedules, hospitals/clinics' service hours; and third is inputting changes both on schedules and slots for doctors, hospitals, clinics, and all services.

The tool is used predominantly by the Catalog Division Staff; As well as other supporting software such as Microsoft Excel. The role of this division is also vital, because they have a fairly heavy workload, for example, in one input, there will be many data that needs to be inputted (as evidenced by the number of doctors as many as 44,424 and the number of 1,962 hospitals/clinics/services recorded in the database tool. The staff also stated that they required each member to have a fast work rhythm so that all the work deadline can be done well (e.g. launching of a new product

that must be activated at the particular hour—Because of those things, this division looks has an important role, it is also depicted on the rhythm of demand changes are quite dynamic, frequent and complex). In general, the deadlines given from outside parties (other divisions/hospitals/clinics/companies) for a job of the Catalog Division Staff is a narrow time gap, so it is not uncommon for the Catalog Division Staff to oblige themselves to sleep late and require to wake up in the morning if tomorrow's schedule is their WFO (Working from Office) shifting schedule.

The tool used is also felt does not have an optimal function as it should be, there are still many points that need to be repaired from every complaint of them while operating the tool. The workload is quite heavy, the work time tends to be narrow and the tools that are deemed not optimal in terms of usability certainly make for an unbalanced life. This certainly can disrupt the circadian rhythm of them. Circadian rhythms greatly affect almost every human physiological function (Ekayanti, 2019).

This cannot be tolerated because these bad things can worsen the quality of human resources. In this study, the problem solving is only on the side that aims to shorten the operation time process so that the productivity of each staff further increases, optimize the number of staff, and maintain the work output quality. In the final, the satisfaction level questionnaires are given regarding the prototypes that features have been upgraded. The results will be the cornerstone of the strongest reasons for these prototypes can be forwarded to the Product Division, which, if approved by them, will be forwarded to the core party, namely IT Center.

2. Literature Review

2.1 Usability

"Usability" according to Nielsen (2012) is the process of evaluating or the quality attributes of the user interface related to how easy a user interface can be used by the operator. The meaning of the word "tool" in this study is the Halodoc Indonesia's internal database tool. Therefore the role of evaluating the usability of the tool is so important, then a good design of user interfaces also greatly affects the ease of operating the tool. Usability measurement level consists of three aspects, (Hadi, 2018): 1) Effective is the success of the operator in completing a job correctly. This aspect is measured based on the number of errors made by the operator when doing a job using the tool; 2) Efficient is the effort required by the operator to complete a job. This aspect is measured based on how long the operator does a job using the tool, and 3) Satisfaction is a measure of dissatisfaction and satisfaction in using the tool.

The quality of usability according to (Hadi, 2018) consists of five components, namely: learnability, efficiency, memorability, errors, and satisfaction; Meanwhile, the required usability in this study includes only four components, namely:

1. Learnability is measuring operator ease of using the tool. Measured based on success rate or the percentage of successful operators in completing a job correctly with the tool (Nielsen, 2001). This is the formula of learnability:

$$Success\ Rate = \frac{S+(P \times 0,5)}{Total\ Task} \times 100\% \qquad (1)$$

2. Efficiency is measuring the operator speed of doing a job on the tool. Efficiency can be calculated by the time spent in doing a job (Situmorang, 2019)

$$Task\ time = End\ time - Start\ time \qquad (2)$$

3. Error is measuring the error rate made by operators in doing a job on the tool.

$$Time\ Based\ Efficiency = \frac{\Sigma_{j=1}^{R} \Sigma_{i=1}^{N} \frac{n_{ij}}{t_{ij}}}{NR} \qquad (3)$$

4. Satisfaction is measuring the satisfaction level in doing a job. Satisfaction was calculated using the System Usability Scale (SUS) questionnaire score.

$$Skor = (Q1 - 1) + (5 - Q2) + \ldots \times 2,5 \qquad (5)$$

Memorability is a component to measure how the operator maintains their learning knowledge in operating the tool after a certain period of time. This research does not calculate the memorability component because it is not relevant because the staff who will be the research subject has memorized every single placement and all tool operations. This suitable tool has just been released.

2.2 Usability Testing

Usability Testing is one of several ways to evaluate the tool (Rubin, 2008). (Babich, 2018) states that usability testing is the best way to know the real experience of tool users by seeing firsthand how a process is carried out by the operators concerned. According to Nielsen (2012), the advantage of using this method is that it is not a complicated process and does not need to spend a lot of money in the process of activities, only relying on operators as respondents and then observed. In this process, tool evaluation involves all users (staff). Tullis & Albert (2008) argue that there are two ways in the usability measurement process, first is by using assumptions, and the second way is given the usability metric or answering questions such as, how do division staff think about the tools used at this time, what kind of constraints are felt on the tool, and so on. This stages in testing usability metrics consist of several steps, such as:

1. Choose the questionnaire method (In this study the usability metric is used because the results obtained will be more precise so that the improvement approach is right on target).
2. Select participants (In this study, all staff).
3. Determine the number of samples (Due to the total number of staff are 9 members, the sample number used is all of them), and
4. Process and interpret the data.

The questionnaire used in the previous explanation is known as the System Usability Scale (SUS). In addition to the questionnaires SUS, the next author did an interview against each operator using the tool. Data will be obtained from interviews. SUS questionnaires will be in use in the search for the occurred problems and become the foundation approach to recommendations for improvements in the tool. The SUS questionnaire consists of 10 questions with a Likert Scale of 1-5 for each question; Where score (1) means strongly disagree, score (2) means disagree, score (3) means neutral, score (4) means agree, and score (5) means strongly agree. The scale that can be used for the results of the SUS questionnaire is illustrated in Figure 1:

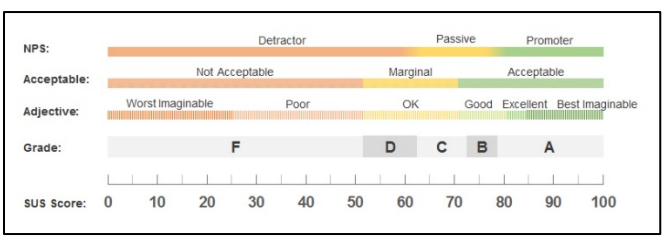

Figure 1. SUS Score Scale

3. Research Methodology

3.1 Usability Study Literature

Deepening the theories related to research (usability, usability testing, usability on the tool, tool improvement, and the SUS questionnaire). Then these theories are used as a basis for knowledge to plan the research process, the variables to be taken during the research process, how to process data, analyze data results, conduct further interview processes if the calculation results are deemed not good, how to solve problems, and drawing conclusions at the end of the study.

3.2 Planning

Planning required preparation. The preparation referred to in this research is determining scenarios that will be carried out by participants to be tested, planning the research implementation schedule, and preparing the SUS questionnaire. The scenario used in the calculation is doing 3 main jobs, such as 1) Create doctor data and service products; 2) Create hospital data; and 3) Create and delete schedules, manage doctor and service slots.

3.3 Implementation

The stage of this research implementation. The first research implementation is to ensure that the scenarios that have been arranged to be run sequentially and according to what has been planned. Then always monitor at each stage of the process. The first data taken is the success level. This stage uses the symbol "P" for partially completed jobs, and the "S" for tasks that are perfect/successful in their entirety. Task 1 (T1) will be said to be successful if it is successfully completed within 180 seconds, task 2 (T2) will be said to be successful if it is successfully completed in 120 seconds, and task 3 (T3) will be said to be successful if it is successfully completed in 480 seconds. The second data taken is the time to complete the task. At this stage, it requires the use of a stopwatch in calculating the duration of each given task. The third data that was taken is the error. At this stage, describe in advance the possible errors that can occur in each task. Once described, count how many errors occurred by comparing the list of possible mistakes to the work that has been done. This data can be retrieved together at the first data collection. Then prepare a System Usability Scale (SUS) questionnaire, monitor the filling of the questionnaire so that the obtained results are valid data and in accordance with the actual situation (avoid filling the questionnaire randomly).

3.4 Data Processing

The first component of usability level is learnability. The first component is calculated based on the first data: success rate. The overall average obtained in this study was 76%. The second usability level is the efficiency component. This second component is calculated based on the second data: the task completion time. The average time for completing the first task (T1) is 216 seconds, the second task (T2) is 202.33 seconds, and the third task (T3) is 461.89 seconds. Once done calculate the Time-Based Efficiency. TBE for T1 is 0.002 goals/sec; for T2 is 0.003 goals/sec; and for T3 is 0.001 goals/sec. Then the calculation of the task completion ratio of staff to experts is carried out (In this study the role of an expert is the Division Head). The third usability data is a component of errors. This third component is calculated based on the third data: the total errors of the three tasks that have been completed by all staff. The total number of errors committed by all staff was 63 and the error rate obtained was 0.6. The fourth level of usability data is the satisfaction component. This fourth component is calculated based on the fourth data: the results based on the SUS Questionnaire.

3.5 Data Analyzing Result

The analysis process is carried out based on the knowledge obtained from the literature study stage, through case experiences that have been previously applied to usability research journals, and ensure that you have read many reliable sources so that the analysis results quality can be accounted for. Analysis of the learnability component is that the average success rate is compared to the ideal average success rate according to research references that use Sauro theory, which is 78%, so the average success rate of the staff is in the bad category (2% below Sauro's average success rate). The analysis of the efficiency component is that the ideal average ratio according to research references that use Sauro is 3.20 or 3 times longer than the expert. Therefore, it can be concluded that the working time for the staff was categorized as normal. The analysis of the errors component is that the error rate is included in the fair category but still needs attention because this figure is almost touching the ideal error rate according to research references that use Sauro. Analysis of the satisfaction component is that the average is in the "Poor" Adjective Rating category, the obtained Grade Scale is the "F" value, and the Acceptability Ranges is in the " Not Acceptable " category. This condition is very apprehensive so that the tool really needs to be examined more deeply regarding common problems, this can be done by further interviewing the division team, where this will be discussed in the next section.

3.6 Further Observation

The interview is a further research stage in this research process. This stage is carried out in the form of an interview process with all staff in identifying further problems with the tool that have not been previously identified. This is done because the results obtained at the data processing and analysis stage are theoretically considered very bad. This stage is done to determine whether the results really actually happened, and make sure things or other problems occur on the tool used. The results of the interview raised 30 problems that had not been known before (from 30 problems, can be grouped into 20 problem categories).

3.7 Recommend Improvements

Recommend improvements is the stage of the formulation of recommendations to address the root problems. Recommendations for improvement are proposed based on an industrial engineering mindset, where each recommendation for improvement is encouraged to have the goal of minimizing processing time, increasing work productivity, workers or tools, and not increasing any costs. A total of 20 recommendations are detailed and described in the next section.

3.8 Analyzing Satisfaction Level

Analysis of the level of satisfaction is the stage of finding out the level of satisfaction of the division staff with the conditions of pre-repair recommendations and post-improvement recommendations, as well as knowing the comparison of satisfaction levels pre and post improvement recommendations. This stage is carried out by distributing questionnaires to staff to be filled in honestly (using a Likert scale), then the data obtained is analyzed to be used as a reference for the success scale of all recommendations for improvement made.

3.8 Conclusions and Suggestions

Formulating conclusions for research carried out and providing suggestions that can help further research.

4. Results

4.1 Learnability Component

This component is calculated based on the success rate. Table 1 is the results of calculating the *success rate* (in percent) of all staff using equation 1 (where T1 is the 1st task, T2 is the 2nd task, and T3 is the 3rd task):

Table 1 Success Rate

Staff	Tak			Success Rate
	T1	T2	T3	
S1	P	P	S	67%
S2	S	S	S	100%
S3	S	S	P	83%
S4	S	S	P	83%
S5	P	P	S	67%
S6	S	P	S	83%
S7	S	S	S	100%
S8	P	P	P	50%
S9	P	P	P	50%
Average				**76%**

4.2 Efficiency Component

The time for completing the task (calculated with Equation 2) along with the average calculation for each task that has been completed is illustrated in Table 2:

Table 2 Completion Time

Staff	Completion Time (Second)		
	T1	T2	T3
S1	211	122	457
S2	179	101	429
S3	164	117	481
S4	177	119	490
S5	237	124	449
S6	166	121	441
S7	167	125	432
S8	262	152	497
S9	258	146	481
Average	**202.33**	**125.22**	**461.89**

Table of Time-Based Efficiency (TBE) calculation (in goals/sec) :

Table 3 Time-Based Efficiency

Time Based Efficiency		
T1	**T2**	**T3**
0,002	0,003	0,001

The table of the staff to expert task completion ratio is illustrated in Table 4:

Table 4 Task Completion Ratio

Task	Avarage Completion Time		Ratio
	Staff	Expert	
T1	202	172	0,9
T2	125	92	0,7
T3	462	423	0,9

4.3 Errors Component

The total number of errors for the three tasks that have been completed by all staff is illustrated in Table 5:

Table 5 Total Errors

Staff	Total Errors		
	T1	T2	T3
S1	5	2	0
S2	2	3	0
S3	4	2	0
S4	2	4	0
S5	3	5	1
S6	2	2	0
S7	2	4	0
S8	6	4	1
S9	6	3	0
Total Errors	**63**		
Error Rate	**0.6**		

4.4 Satisfaction Component

Here are the results of the SUS questionnaire calculation in Table 6:

Table 6 SUS Questionnaire Calculation

Staff	Questions										Sum	Sum × 2,5
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10		
S1	4	0	1	0	4	4	1	1	3	0	18	45
S2	4	0	1	0	4	4	1	1	3	0	18	45
S3	4	0	1	0	4	4	1	1	3	0	18	45
S4	3	0	1	1	3	4	0	0	3	1	16	40
S5	4	1	0	1	3	4	0	0	3	1	17	42,5
S6	4	1	0	1	4	3	1	1	4	0	19	47,5
S7	4	0	0	1	4	3	1	1	4	0	18	45
S8	3	1	0	0	3	3	0	0	4	0	14	35
S9	3	1	1	0	3	3	0	0	4	1	16	40
											Sum	**385**
											Avg	**42,7**

5. Disussion

5.1 Learnability Component

The average success rate amounted to 76%, where this illustrates that if there are staff who obtain the success rate above 76% then the success rate is categorized into the "S" category or complete success, and vice versa. Therefore, S1 and S5 had a poor success rate of 67% (9% below the average success rate). S2 and S7 had a good success rate, 100% (24% above the average success rate). S3, S4, and S6 had a good success rate, namely 83% (7% above the average success rate). S8 and S9 had a poor success rate of 50% (26% below the average success rate). If the average success rate of the staff compared to the average success rate is ideal according to Sauro, 78%, the average success rate of the staff is in the bad category (2% below the average success rate Sauro).

5.2 Efficiency Component

The efficiency component calculates the time to complete each task. According to Table 2, Staff 1 (S1) completed the first task (T1) for 211 seconds, the second task (T2) for 122 seconds, and the third task (T3) for 457 seconds. Staff 2 (S2) completed the first task (T1) for 179 seconds, the second task (T2) for 101 seconds, and the third task (T3) for 429 seconds. Staff 3 (S3) completed the first task (T1) for 164 seconds, the second task (T2) for 117 seconds, and the third task (T3) for 481 seconds. Staff 4 (S4) completed the first task (T1) for 117 seconds, the second task (T2) for 119 seconds, and the third task (T3) for 490 seconds. Staff 5 (S5) completed the first task (T1) for 237 seconds, the second task (T2) for 124 seconds, and the third task (T3) for 449 seconds. Staff 6 (S6) completed the first task (T1) for 166 seconds, the second task (T2) for 121 seconds, and the third task (T3) for 441 seconds. Staff 7 (S7) completed the first task (T1) for 167 seconds, the second task (T2) for 125 seconds, and the third task (T3) for 432 seconds. Staff 8 (S8) completed the first task (T1) for 262 seconds, the second task (T2) for 152 seconds, and the third task (T3) for 497 seconds. Staff 9 (S9) completed the first task (T1) for 258 seconds, the second task (T2) for 146 seconds, and the third task (T3) for 481 seconds.

Time-Based Efficiency is the time that can be used by staff to complete each task. The Time-Based Efficiency calculation for the first task (T1) based on Table 3 is 0.002 goals/sec, this data shows that the first task (T1) can be completed by staff in 500 seconds. The Time-Based Efficiency calculation for the second task (T2) is 0.003 goals/sec, this data shows that the second task (T2) can be completed by staff in 333 seconds. The Time-Based Efficiency calculation for the third task (T3) is 0.001 goals/sec, this data shows that the third task (T3) can be completed by the staff in 1,000 seconds. When compared with experts, who completed the first task (T1) for 172 seconds, the second task (T2) for 92 seconds, and the third task (T3) for 423 seconds, then the ratio of completing the first task (T1) and the third task (T3) staff was 0.9 times slower than experts, while the third task completion ratio (T2) of staff was

0.7 times slower than experts. The average ideal ratio according to Sauro is 3.20 or 3 times longer than the expert. Therefore, it can be concluded that the working time for the staff was categorized as normal.

5.3 Errors Component

The first task (T1) has some risk of errors, which are: uppercase and lowercase typo in writing the name of the doctor, meta keywords, name of the hospital, filling meta keywords section takes a long time, looking for a doctor's specialty, and uppercase and lowercase typo in medical education background. The second task (T2) has some risk of errors, which are: finding the type of hospital category more than once, city, subdistrict, region, and zone of territories takes a long time. The third task (T3) has one risk of error: slip when choosing a schedule.

The total errors of all assigned tasks made by the staff were 63 times. Where the errors made by Staff 1 (S1) on the first task (T1) were 5 pieces, in the second task (T2) were 2 pieces. There were 2 errors made by Staff 2 (S2) in the first task (T1), 3 in the second task (T2). There are 4 errors made by Staff 3 (S3) in the first task (T1), 2 in the second task (T2). There were 2 errors made by Staff 4 (S4) in the first task (T1), 4 in the second task (T2). There were 3 errors made by Staff 5 (S5) in the first task (T1), 5 in the second task (T2), and 1 in the third task (T3). There are 2 errors made by Staff 6 (S6) in the first task (T1), 2 in the second task (T2). There are 2 errors made by Staff 7 (S7) in the first task (T1), 4 in the second task (T2). There are 6 errors made by Staff 8 (S8) in the first task (T1), 4 in the second task (T2), and 1 in the third task (T3). The errors made by Staff 9 (S9) in the first task (T1) were as many as 6, in the second task (T2) were 3.

So that the calculation of the error rate is calculated by the formula for the total error divided by the total chance of an error. Where the total error is 63, while the total chance of an error occurring is by multiplying 12 possibilities by the total staff working on the task, which is 9. Then 63 is divided by 12 times 9. Then the resulting error rate is 0.6. If these results are compared with the ideal error rate according to Sauro, 0.7; So the error rate generated by the staff is in the fair category, but it still needs attention because this figure is almost touching the ideal error rate according to Sauro.

5.4 Satisfaction Component

The first SUS questionnaire statement which reads "I think I will use this tool often" results in 33.3% agree and 66.7% strongly agree, if it is concluded these points get an average vote to agree. The second SUS questionnaire statement which reads "I feel this tool is too complicated even though it can be made simpler" gets the results 44.4% agree and 55.6% strongly agree if it is concluded this point gets an average vote. The third SUS questionnaire statement that says "I feel the tool is easy to use" gets 44.4% strongly disagree and 55.6% do not agree if it is concluded at this point to get an average sound does not agree. The fourth SUS questionnaire statement which reads "I think I need help from a technical person to be able to use this tool" results in 44.4% agree and 55.6% strongly agree if it is concluded that this point gets an average vote to agree.

The fifth SUS questionnaire statement which reads"I find that there are various kinds of things that are well integrated into the tool system" get the results 44.4% agree and 55.6% strongly agree if it is concluded this point gets an average vote. The sixth SUS questionnaire statement which reads"I feel that there are many inconsistencies in this tool" gets the results of 55.6% strongly disagree and 44.4% disagree if it is concluded this point gets an average vote of disagreement. The seventh SUS questionnaire statement which reads"I feel the majority of users will be able to learn this tool quickly" get the results of 44.4% strongly disagree and 55.6% disagree, if it is concluded this point gets an average vote disagree. The statement of the eight SUS questionnaire which reads"I find this tool very impractical when used" results in 55.6% agree and 44.4% strongly agree if it is concluded that this point gets an average vote agree. The statement of the ninth SUS questionnaire which reads"I really believe I can use this tool" gets the results of 55.6% to agree and 44.4% strongly agree if it is concluded that this point gets an average vote to agree. The tenth SUS questionnaire statement which reads "I have to learn a lot of things before I can use this tool" gets the results of 33.3% to agree and 66.7% strongly agree if it is concluded that this point gets an average vote agree.

The System Usability Scale (SUS) questionnaire for this tool gets an average value of 42.7. If the value is reviewed within the SUS assessment scale, then this average has Adjective category Ratings "Poor", Grade Scale obtained is

the value "F", and Acceptability Ranges in the category "Not Acceptable". This condition is very concerning, so the tool really needs to be examined more deeply regarding common problems, this can be done by further interviewing the division team, where this will be discussed in the next section.

5.5 Recommend Improvements

After the interview process further towards team divisions for observing deeper problems, in the end, the author found a variety of problems that occur on the tool that is often experienced by division. Table 7 are detailed issues and recommendations to improve the tool:

Table 7 Detailed Issues and Recommendation**s**

No.	Issues	Recommendations
1	Cannot delete the current slot.	Redesign the system from the website tool to add a deletion system/ slot removal at the current time.
2	Cannot add "add leave" and delete "add leave" (leave) on the current day (current time).	Rearrange the “scheduling” section so that it can add and remove add leaves at the current time.
3	Schedule closings often occur suddenly due to customers making bookings on the same day.	Reset the system in the application so that it cannot be booked if the order occurs on the same day (the book button is inactive).
4	Adding hours to the schedule is very time-consuming.	Input features straight hours you want to create (type "08: 00-17: 00").
5	Frequent double slots occur in one hour.	Raising the issue to the Party of IT in order to be in surgery there are any issues in the system of the tool that are thus possible double slots.
6	It is very time-consuming to enter “National Holidays” on any active schedule.	Create a "database" on the tool (data on national holidays in the year), so that the tool immediately marks these days as red dates (the operator does not need to input one by one) like Sundays in the tool.
7	The same operational schedule must be input one by one back in the following year.	Add the “ repeat schedule for next year” feature.
8	There is no operational license number at the provider location.	Add section "Operational License Number" on Provider Location so that data can be a point of trust of the customer to conduct tests in the relevant medical partner.
9	When clicking "Personnel Details" and then wanting to return to the previous page (click "back"), the display will go to "General Info" instead of "Personnel List".	Rearrange the system tools to return to the "Personnel List" instead of "General Info".
10	Auto-resize & remove photo backgrounds are done using external applications, upload photos one by one, re-delete photos if the order is wrong.	Features in the photo "upload" section with the ability to auto-resize when the operator inserts a photo, create a "framing" feature so that the operator knows what kind of photo snippet fits each photo placement (slide left-right, zoom in-out to fit into the frame that has been provided), hold feature "remove background" for doctors who have a colorful background, move the location of the sequence enables the photographs, and can perform uploading video.
11	The maximum character limit for writing the title and name of the university (especially overseas universities) is not long enough.	Increase the allowed character limit to 150 characters from the previous 50 characters.
12	In the " Last Name, Education, Experience, Photo, Description, and Email" section cannot be deleted (if there is a writing error). When deleted, the display will continue to fill that section with its previous condition.	Add a "trash bin button" feature so that the data is wrong and should not need to be filled in, not disturbing the display.

13	"Meta keywords" must be retyped word for word (the source of the word mostly comes from "name" and "specialty"), "Specialty list" is searched one by one with English terms, "Facilities" is entered one by one even though it is the same for all branches.	Add "suggestion words" feature to "meta keywords" which are copied based on "name" and "specialty"; In the "specialties" section, "suggestion words" comes from the meaning of "post salutation"; and in the "facilities" section under the same management provider, given the same "suggestion facilities" as other branches.
14	It is difficult to search for Personnel, Provider Location, and Department names because the writing must be exactly the words that have been stored in the tool.	Set the search feature is enough to mention one word of the entire name so that the search process does not take long.
15	Region, District, Type of hospital, and Department list are not sorted ascending (AZ).	Tool structure is changed so that Region, District, Type of hospital, and Department list are sorted ascending.
16	Uppercase-lowercase typo in writing personnel names and education.	Auto-uppercase in writing names and education (every time after the "space" character).
17	The quite unfamiliar time zone naming.	Naming the time zone as WIB, WIT, WITA and designing a database on the tool so that region A is directly categorized into time zone A, region B into timezone B, and so on.
18	The calculation of the number of doctors in the personnel list by the department is not by name.	The inactive "Department" can be hidden in "Personnel Management", so it does not disturb operators in sorting data (minimizing errors data input).
19	Doctors who are already "inactive" still appear on the personnel list.	Separate doctors with "active" and "inactive" status to minimize the amount of data on the display, as well as to minimize errors in data input (with a dropdown feature, to be able to choose whether to display active/inactive/all).
20	The personnel list appears only 10 lines at the initial default view.	Reset the default display so that it can display the entire data because the search process that is always done by the operator is for "all data".

5.6 Interference and Satisfaction Level to the Tool

Interference level of the pre-recommendation in this study is also defined as the level of disturbance of each issue before the recommendation for improvement is carried out, obtained based on a questionnaire using a Likert Scale of 1-5, where the value "1" which means very disturbing to the value "5" which means very unobtrusive. The satisfaction level of post-recommendation is obtained based on a questionnaire using a Likert Scale of 1-5, where the value "1" means very dissatisfied up to the value "5" which means very satisfied. The questionnaires were given to 9 staff.

Table 8 Interference & Satisfaction Level

Issue No.	Interference Level					Recommendation	Satisfaction Level				
	1	2	3	4	5	No.	1	2	3	4	5
1	88.9%	-	11.1%	-	-	**1**	-	-	-	11.1%	88.9%
2	55.6%	44.4%	-	-	-	**2**	-	-	-	33.3%	66.7%
3	66.7%	22.2%	11.1%	-	-	**3**	-	-	-	33.3%	66.7%
4	44.4%	44.4%	11.1%	-	-	**4**	-	-	-	44.4%	55.6%
5	22.2%	55.6%	22.2%	-	-	**5**	-	-	-	55.6%	44.4%
6	22.2%	66.7%	11.1%	-	-	**6**	-	-	-	44.4%	55.6%
7	22.2%	66.7%	11.1%	-	-	**7**	-	-	-	33.3%	66.7%
8	44.4%	44.4%	11.1%	-	-	**8**	-	-	-	55.6%	44.4%
9	33.3%	44.4%	22.2%	-	-	**9**	-	-	-	22.2%	77.8%
10	55.6%	33.3%	11.1%	-	-	**10**	-	-	-	33.3%	66.7%
1	33.3%	33.3%	33.3%	-	-	**1**	-	-	-	33.3%	66.7%
12	11.1%	55.6%	33.3%	-	-	**12**	-	-	-	44.4%	55.6%
13	33.3%	33.3%	33.3%	-	-	**13**	-	-	-	66.7%	33.3%
14	11.1%	55.6%	33.3%	-	-	**14**	-	-	-	55.6%	44.4%
15	44.4%	33.3%	22.2%	-	-	**15**	-	-	-	66.7%	33.3%
16	33.3%	55.6%	11.1%	-	-	**16**	-	-	-	66.7%	33.3%
17	55.6%	44.4%	-	-	-	**17**	-	-	-	66.7%	33.3%
18	33.3%	55.6%	11.1%	-	-	**18**	-	-	-	11.1%	88.9%
19	55.6%	33.3%	11.1%	-	-	**19**	-	-	-	66.7%	33.3%
20	44.4%	22.2%	33.3%	-	-	**20**	-	-	-	44.4%	55.6%
Average	**40.5%**	**42.2%**	**17.2%**	**-**	**-**	**Average**	**-**	**-**	**-**	**44%**	**56%**

The results of the interference level questionnaire have been illustrated in Table 8. In the table, it can be seen clearly the level of disturbance for each issue (from issue number 1 to issue number 20). For example, in issue number 1, there are as many as 88.9% of the respondents consider this issue to be a very disturbing issue, and the remaining respondents (as many as 11.1%) consider this issue to be a bit disturbed. Broadly speaking, the data above informs that all issues fall into the category of moderate to severe disturbance. It is concluded that the highest percentage is in the category "Enough to disturb" (Value 2), amounting to 42.2%; The second highest percentage was in the "Very Annoying" category (Value 1), amounting to 40.5%; The third highest percentage was in the "Slightly Disturbing" category (Value 3), amounting to 17.2%; and in the remaining categories, there is no evaluation at all from any staff.

The results of the satisfaction level questionnaire also have been illustrated in Table 8. In the table, it can be seen clearly the level of satisfaction with each recommendation for improvement (recommendations for improvement number 1 to number 20). For example, the recommendation for improvement number 1, there are 11.1% of the respondents who are satisfied with the recommendation for improvement, and the remaining respondents (as many as 88.9 %) are very satisfied. Broadly speaking, the data above informs that all recommendations for improvement are in the satisfied category. This is concluded in the highest percentage in the "Very Satisfied" category (Value 5), amounting to 56 %; percentage the second highest in the category of "Satisfied" (Value 4), by 44 %; and in the remaining categories there is no evaluation at all from any staff.

5.7 Comparison of Satisfaction Levels Pre and Post Improvement Recommendations on Tool

Figure 2 is a comparison of the satisfaction levels of pre and post repair on the tool (red for before and blue for after):

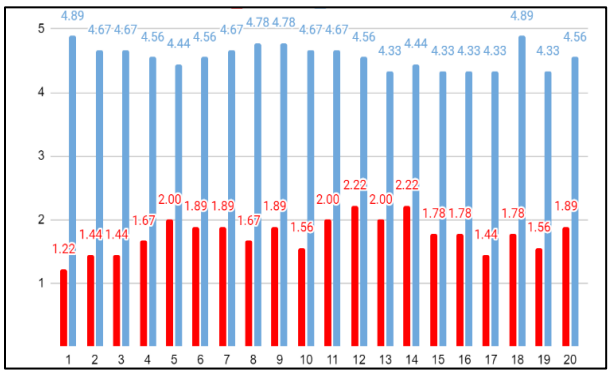

Figure 2 Comparison of Pre and Post Satisfaction Level

Figure 2 illustrates the level of difference between pre and post improvement recommendations. The red graph shows the average satisfaction before the recommendation for improvement, while the blue graph shows the average satisfaction after the recommendation for improvement. Broadly speaking, the graph informs that each tool improvement recommendation increases the satisfaction level of division staff. Regarding the percentage increase in the satisfaction level of the division staff, it is illustrated in the Table 9.

Table 9 Percentage of Increase in Satisfaction Level of Improvement Recommendations

Recommedation No.	% Increase Satisfaction Level	Recommedation No.	% Increase Satisfaction Level
1	73.4%	**11**	53.4%
2	64.4%	**12**	46.6%
3	64.4%	**13**	46.6%
4	57.8%	**14**	44.4%
5	48.8%	**15**	51.2%
6	53.4%	**16**	51.2%
7	55.6%	**17**	57.8%
8	62.2%	**18**	62.2%
9	57.8%	**19**	55.6%
10	62.2%	**20**	53.4%

6. Conclusion and Suggestions

6.1 Conclusion

Based on data analysis and processing, then k esimpulan who successfully obtained are :

1. The level of usability to the tool reviewed based components Learnability, Efficiency, Errors, and Satisfaction. For the Learnability component, the average success rate of the staff is compared with the average success rate according to Sauro, which is 78%, then the staff's average success rate is in the bad category (2% below Sauro's average success rate). For the Efficiency component, The average ideal ratio according to Sauro is 3.20 or 3 times longer than the expert. Therefore, it can be concluded that the working time for the staff was categorized as normal. For the Errors component, the error rate is compared with the ideal error rate according to Sauro, 0.7; So the error rate generated by the staff is in the fair category, but it still needs attention because this figure is almost touching the ideal error rate according to Sauro. For the Satisfaction component, the System Usability Scale (SUS) questionnaire is reviewed within the SUS assessment scale, then the average has Adjective category Ratings "Poor", Grade Scale obtained is the value "F", and Acceptability Ranges in the category "Not Acceptable". This condition is very concerning, so the tool really needs to be reviewed more deeply regarding common problems, this can be done by further interviewing the division team.
2. After further interviewing, there are found 20 issues and the 20 recommendation improvements are given to the tool to optimize the level of usability are redesigning the system from the website tool to add a deletion system/slot removal at the current time; Set back section "scheduling" to be able to add and remove add leave at the current time; Set on the application system to not be able to book when the booking occurs on the same day (book button inactive); The added feature is to directly input the clock you want to create; Added the feature "repeat schedule for next year"; Add section "Operational License Number"; Rearrange system tools so that they return to "Personnel List" instead of "General Info", and so on.
3. The 20 improvement recommendations have been proven to be successful as seen from the calculation of the percentage increase in the satisfaction level. It can be seen that the smallest percentage increase value is 44.4% and the largest percentage increase value is 73.4%.

6.2 Suggestions

Based on the basis of the conclusions above, the suggestions can be in put forward are:

1. A review of the problem with the tool reflecting the experience of the Catalog Division Staff can be done further carefully.
2. Recommendations for tool improvements can still be discussed carefully both with the Head of the Catalog Division and with the Product Division before making visualizations or prototypes so that the resulting recommendations for improvements are recommendations for fixes that solve the root cause.

References

Babich, N. (2018). A Comprehensive Guide To Mobile App Design.

Firmansyah, R. (2018). Usability Testing Dengan Use Questionnaire Pada Aplikasi Sipolin Provinsi Jawa Barat. Jurnal Swabumi, 6(1), 1-7.

Halodoc Indonesia. (2016). Halodoc Indonesia. Dipetik September 17, 2020, dari https://www.halodoc.com/media

Huda, N. (2019). Implementasi Metode Usability Testing Dengan System Usability Scale Dalam Penilaian Website RS Siloam Palembang. KLIK, 6(1).

IE Foreword. (t.thn.). Dipetik Januari 17, 2021, dari https://ie.binus.ac.id/

Kevin Ryan Hadi, H. M.-Z. (2018). Analisis dan Perbaikan Usability Aplikasi Mobile KAI Access Dengan Metode Usability Testing dan Use Questionnaire. 2(9), 2742-2750.

Merry Septemi Ekayanti, M. F. (2019). Irama Sirkadian pada Stroke Akut. 2(1), 9-18.

Nielsen. (2001). Success Rate: The Simplest Usability Metric. [online] Nielsen Norman Group.

Nielsen J. (2012). Usability 101: Introduction to Usability.

Ningrum, S. W. (2019). Evaluasi dan Perbaikan Usability Aplikasi Mobile Ojesy Menggunakan Metode Usability Testing dan Use Questionnaire. JPTIIK, 3(5).

Rubin, J. &. (2008). Handbook of Usability Testing (2nd ed.). Boulevard Wiley.

Sabandar, V. P. (2018). Evaluasi Aplikasi Media Pembelajaran Statistika Dasar Menggunakan Metode Usability Testing. TEKNIKA, 7(1).

Sejarah BINUS University. (t.thn.). Dipetik January 17, 2021, dari http://web.binus.edu/About.BINUS/Sejarah.BINUS/Indonesia

Setiawan, A. (2018). Evaluasi Website Perguruan Tinggi Menggunakan Metode Usability Testing. Jurnal Informatika: Jurnal Pengembangan IT (JPIT), 03(03), 5.

Setiawati, A. (2018). Pengembangan dan Pengujian Aspek Usability Pada Sistem Informasi Perpustakaan (Studi Kasus: Stikom Dinamika Bangsa Jambi). Processor, 13(1).

Suyanto, S. (2019). Pengujian Usability dengan Teknik System Usability Scale pada Test Engine Try Out Sertifikasi. MATRIK, 19(1).

Theresia Karina Situmorang, H. M.-Z. (2019). Evaluasi Usability Pada Aplikasi m-KantorPos dengan Menggunakan Metode Usability Testing. Malang: Universitas Brawijaya.

Tullis, T. a. (2008). Measuring the User Experience. Morgan Kaufman.

Utami, N. W. (2020). Evaluasi Usability Pada E-Learning Universitas Pendidikan Ganesha Dengan Metode Usability Testing. Janapati, 9(1), 1-12.

Wedayanti, N. L. (2019). Evaluasi Aspek Usability pada Aplikasi Simalu Menggunakan Metode Usability Testing. Jurnal Ilmiah Merpati, 7(2), 1-12.

Wibowo, D. S. (2016). Usability Testing Sistem pada E-Academic Politeknik Harapan Bersama. emitor, 16(1), 1-7.

Willy Arief Pramono, H. M.-Z. (2019). Evaluasi Usability pada Aplikasi MyTelkomsel dengan Menggunakan Metode Usability Testing. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 3(3), 2951-2959.

Wiratama, L. S. (2017). Evaluasi Antarmuka Website Menggunakan Metode Usability Testing (Studi Kasus Smk Muhammadiyah 2 Sragen). Simetris, 8(1).

Yahya Dwi Wijaya, I. B. (2018). Evaluasi Aplikasi Pencarian Bakat "APIK" Dengan Menggunakan Usability Testing. Jurnal Ilmiah IT CIDA, 4(1), 1-12.

Bina Husada Palembang. *Jurnal FIK*, 1-12.

Biographies

Selvi is an undergraduate student from BINUS University. She was the student chapter's president from the world's largest professional society that supports the industrial engineering profession and individuals involved with improving quality and productivity, called the Institute of Industrial and Systems Engineers (IISE). She achieved Gold Award from IISE Center in Georgia, because of her chapter improvements. She improved member growth to 16%, member engagement, relationship, value proposition, and loyalty. She sustained the chapter, constructed a strategic plan, increased financial development to 20.8% than before. Her chapter also awarded 3rd place for Most Members between United Arab Emirates, Jordan, Mexico, and Ecuador.

Rienna Oktarina is a Faculty Member and Subject Content Coordinator for Supply Chain Engineering at Bina Nusantara University. She earned PhD in Industrial Engineering from Institut Teknologi Bandung Her research interests include logistics, supply chain management, humanitarian logistics, warehousing, transportation and distribution.