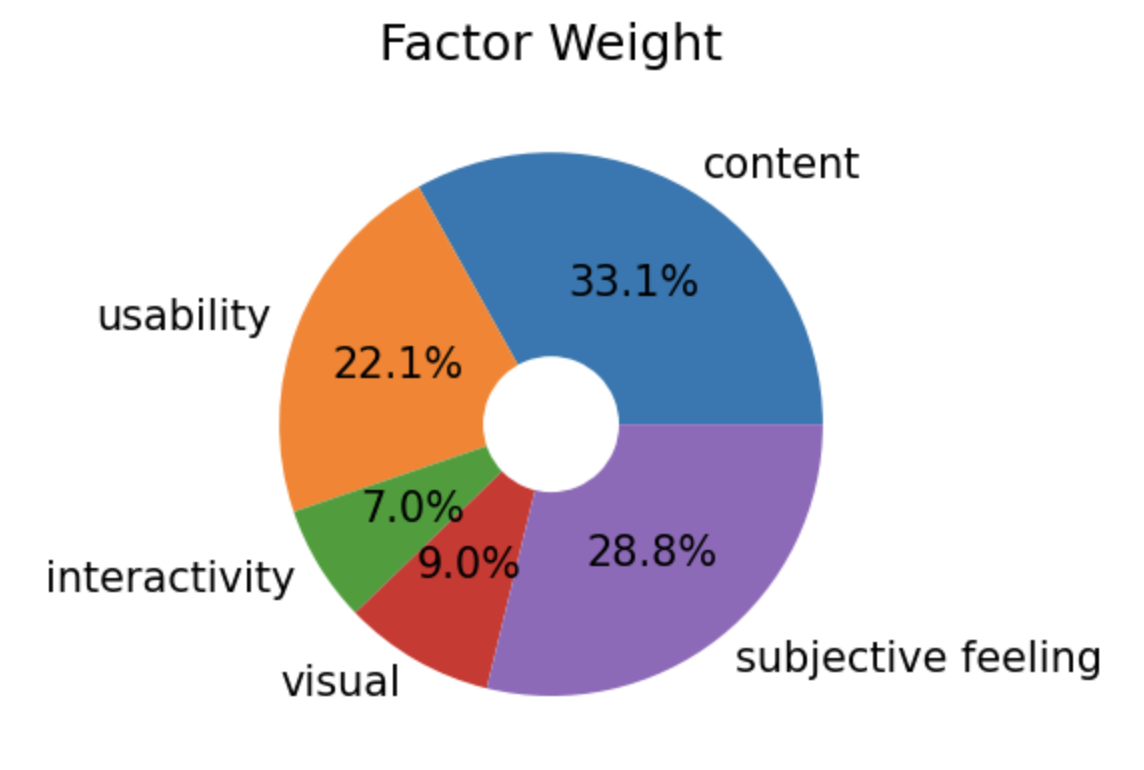
In the website test, XXX once proposed that the key factors to measure website usability are...... . XXX adds to his opinion that the usability factors of a website also include... . In addition, XXX uses... These factors are used to evaluate and test their products. After considering the product characteristics of social networking sites, the quality factors of SafeTweet are determined as follows: content, website availability, website interactivity, visual effects and subjective feelings brought to users.

To further evaluate the influence of influencing factors on the quality of social networking sites, we use AHP (analytic hierarchy Process) to do a weight analysis on the five influencing factors of SafeTweet. (XX supplementary scoring rules) First, we designed a set of questionnaires, including ten Likert scale questions, and compared influencing factors in pairs. The Likert scale is... We then selected the questionnaire for experienced social network users and invited them to answer it. Twenty-two questionnaires were issued, and 20 were recovered, with a recovery rate of 91%.

After the questionnaire was collected, SPSS was used to test the consistency of the questionnaire results. The final CI value was 0.0836<0.1, so the consistency test passed. According to the given results, AHP is used for weight analysis, and the weight values of the five influencing factors are shown in the figure below.



It can be seen that people do not have a high demand for the interaction and visual effects of social networking sites. The current popular social networking sites also tend tothey be a unified coolexcellentand designThe simple UI.

The too rich and complicated interface will make users feel critical to find the crucial point, the use of worse experience. The interaction on social networking sites is relatively fixed and essential, like comments, likes, etc. Therefore, users do not expect too much from the interaction of social networking sites. These fundamental interactions can already meet users' interaction needs on social networking sites. The content and usability of social networking sites have recreated great attention. It can be seen that users are eager to obtain high-quality information from social networking sites. Another point of concern for users' site's usability is often reflected in the interaction between users and threats. Such as the site's fault tolerance... If the user experience is not good in the process of website interaction, it will directly affect the user's evaluation of the website.

Another point worth paying attention to is the subjective feelings of users. Social networking sites are often flooded with information, including users' private information, especially SafeTweet, a private social networking platform dedicated to giving people a voice in the workplace. Therefore, for the main body of this website, that is, users, their subjective feelings are also significant. SafeTweet mainly refers to the feelings of respect, safety and pleasure experienced by users in using SafeTweet when these feelings are satisfied at the same time. They are simultaneously more willing to express their personal opinions and experiences in SafeTweet.

With the determination of SafeTweet influencing factors and their weights, the next step is the user test. We selected 30 volunteers to participate in our user test. They included 28 people aged 19 to 27 who used social networking apps regularly. There were also two participants, aged 37 and 45, who were also on popular social networking sites. There are 20 participants in the Internet industry, including students, developers and software product managers. In addition, there are 10ten non-Internet industry participants, including Japanese, medical, business, public security, etc. We first selected two pilots with experience in testing related products to conduct overall control of the trial, including the introduction before the test, test sequence, test topic and test quantity.

Assessment process

We divided the entire evaluation process into three parts. In the first part, we introduced our testers to the basics of SafeTweet. It was designed to protect free speech in the workplace. In addition, the primary use of SafeTweet and its targeted user groups are introduced to the testers. In addition, participants should be informed that SafeTweet is currently a Prototype, and the definition of a Prototype should be presented to the testers. The inexperienced testers can have a basic understanding of Prototype to improve the accuracy of test results. After completing the basic introduction, ask the tester if they have any other questions and answer them.

The second part is to complete the tasks according to the task list. The task sheet contains all the functions that users can interact with SafeTweet. The task sheet is designed according to the classification of parts of the MoSCoWMoscow table. The categories include... Testers need to complete all tasks under each functional category. The entire test process needs to be recorded to record the completion time of each task and the problems encountered during the task study. And the moderator should not answer any questions related to the test while the participant is performing the test. Participants are required to complete all jobs in the task list by themselves.

The third part is the questionnaire survey after the test. The questionnaire also includes three parts. The first part is the system availability scale. This table was proposed by XXX in 1966 and has been widely used in system usability testing since then. There are ten questions on this scale... . SUS scale is calculated as follows:

According to the research of XXX and XXX, the average score of the current SUS scale is 68. Therefore, we will use this 68 score as a reference value to analyse the SUS result of SafeTweet. The second part of the questionnaire is user experience. This part of the questionnaire was scored according to SafeTweet's influencing factors, and the Likert scale was still used. Influencing factors include content, Usability, Interactivity, Visual presentation and subjective feelings. In each influencing factor, we designed more specific sub-influencing factors to increase the understanding of the influencing factors and improve scoring accuracy. The third part of the questionnaire is the short answer part. The questions include the overall view of SafeTweet, the likes and dislikes of SafeTweet, and the impact of latency on usage when monitoring sensitive information.

Assessment steps

Because of the pandemic, all tests are being conducted online. Each tester was asked to download sunflower Remote Control, a system that could connect two devices remotely. After the tester is paired with the host's computer, the tester can operate on the host's device to complete the test.

1. The test time for all evaluators is 9:00-12:00 am, and the 14:00-12:00-5:00 PMtest is performed by a tester.

2. The host introduces SafeTweet to the testers.

3. The moderator introduces the task list to the tester. If participants have any questions, the facilitator will answer them.

4. Before the test starts, enable screen recording and record the whole process.

5. Evaluators need to complete each task in sequence. The successful completion will be recorded as the task's success; the evaluator was unable to achieve or abandoned registered as a task failure. The moderator should use a stopwatch to record the completion time of each task synchronously on the schedule.

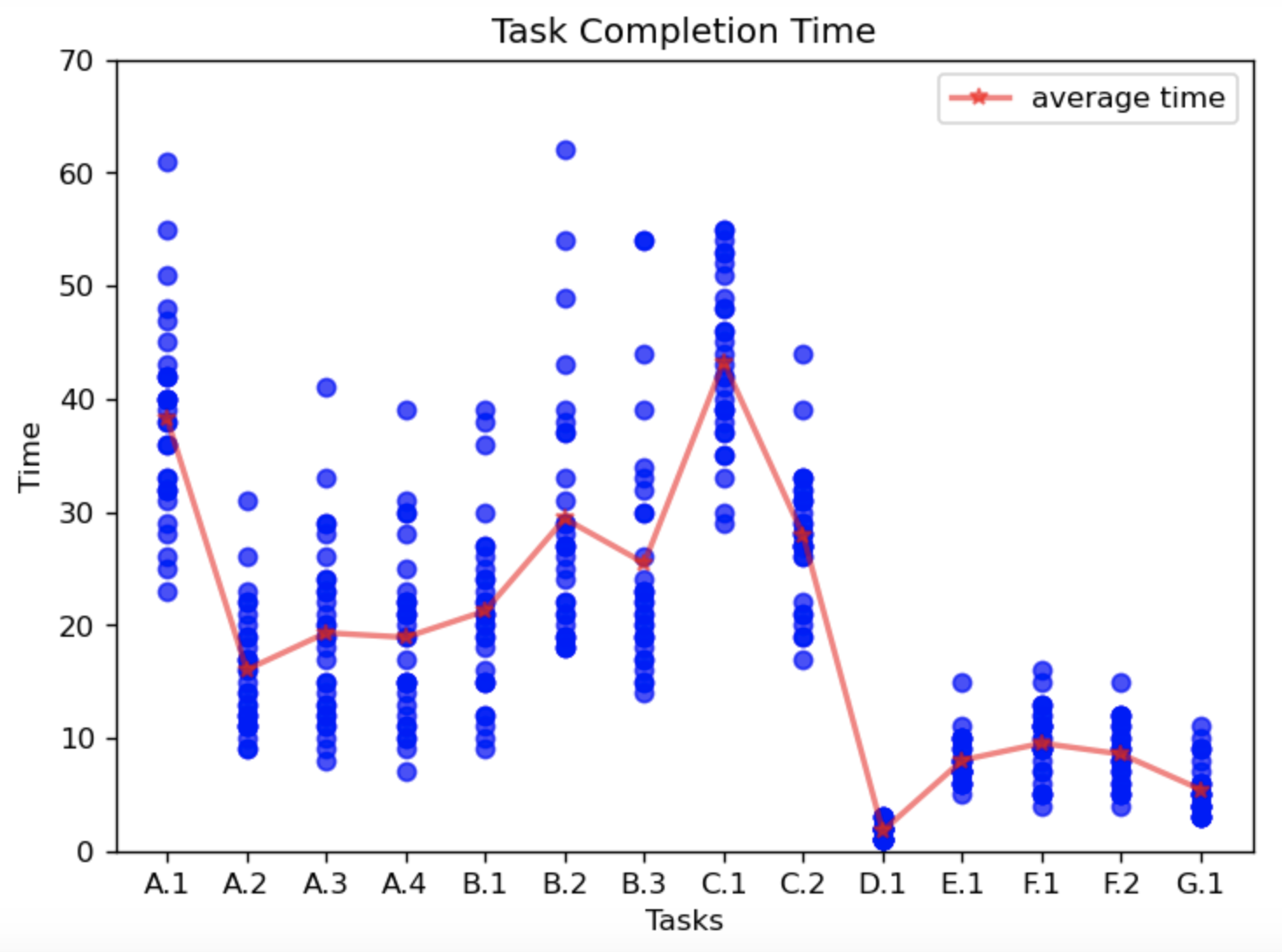
6The moderator shall not instruct the tester to complete the job during the tests.

7. After the task, participants should complete three questionnaires: SUS, UEQ and short answer sections.

8. After the test participants complete the questionnaire, the host should retrieve the task sheet and questionnaire and adequately store it together with the screen recording of the renamed participant. A complete task list ensures that 80% of the tasks have been completed and that the analysis data is valid.

Results analysis

Duration analysis

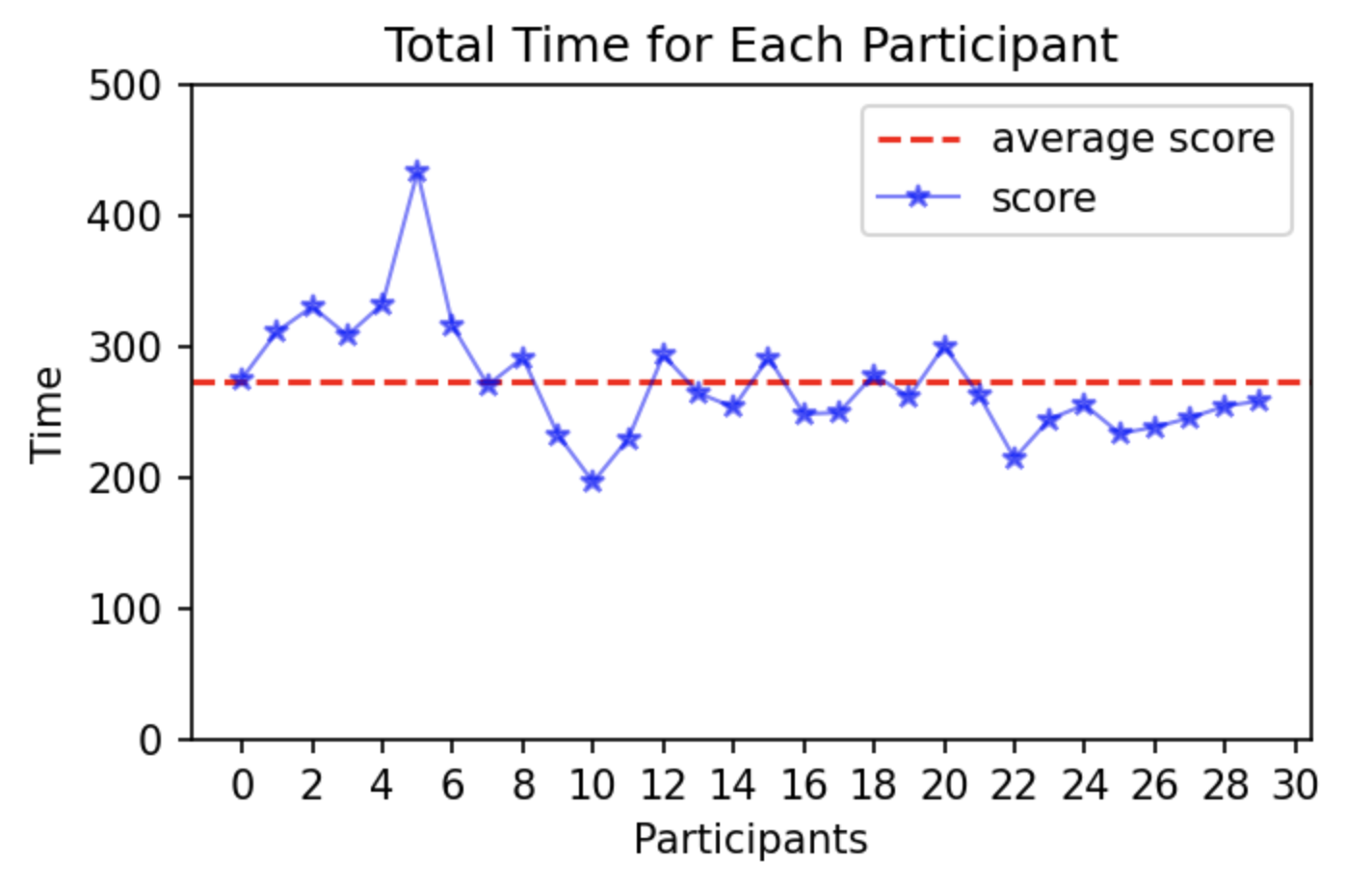


This section summarises how long it took each participant to complete each task, the total time it took to complete all tasks, and the average time for each task. As can be seen from the figure, there are 21 sub-tasks in total, and the average time to complete all tasks is 274 seconds. Thus, the task completion process is smooth. Evaluator No. 6 took the longest to most extended 434 seconds. Participant No. 6 was basiperienced in social networking sites, and the total time consuming was higher than other experienced participants. It takes time to understand the purpose of the task, figure out what each button does, and ion carefully. However, all tasks in the task list are completed. Participant 6 is also in XXX... The most time-consuming evaluator of the task, so there are seemingly anomalous time nodes in Figure X. The c.1 task, which took the longest average time of all tasks, was 43 seconds. The task of C.1 is to input sensitive content and use the sensitive information monitoring function of the system to detect it. The main problems are as follows:

1. The tester has no concept of the function of detecting sensitive information and cannot be accurate. It cannot increase the sensitivity to the system, so it will spend some time thinking about sensitive content.

2. Some evaluators ignore that the SafeTw social networking site is mainly based on the work scene and when they use sensitivity. When utterly unrelated to the work scene, the system may not be able to identify it.

3. The sensitive information system can detect various sensitive information and prompt users on which kind of information is leaked. Some testers spontaneously hope to test other types of sensitive information the method recognised by inputting different sensitive information data.

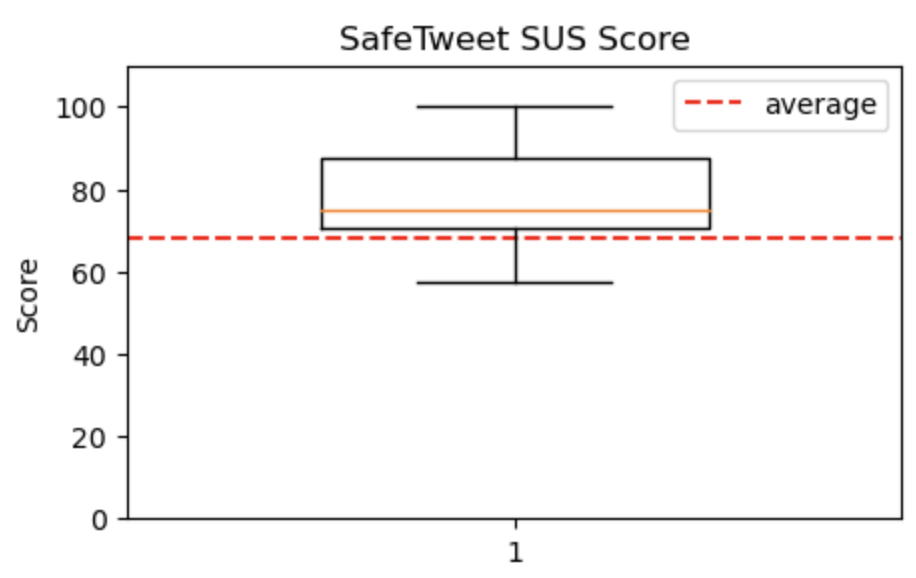
In addition, the other long task is A.1, which takes 38 seconds on average. The main content of this task is to ask the user to register an account. As five items of personal information need to be filled in, and when setting the password, most of the testers have to think about it, so it generally takes a long time on average; the shortest taskrage was D.1, which lasted only 1.8 seconds. The main content of this task is that testers select the content they are interested in to give a thumbs-up, which is an interactive function.

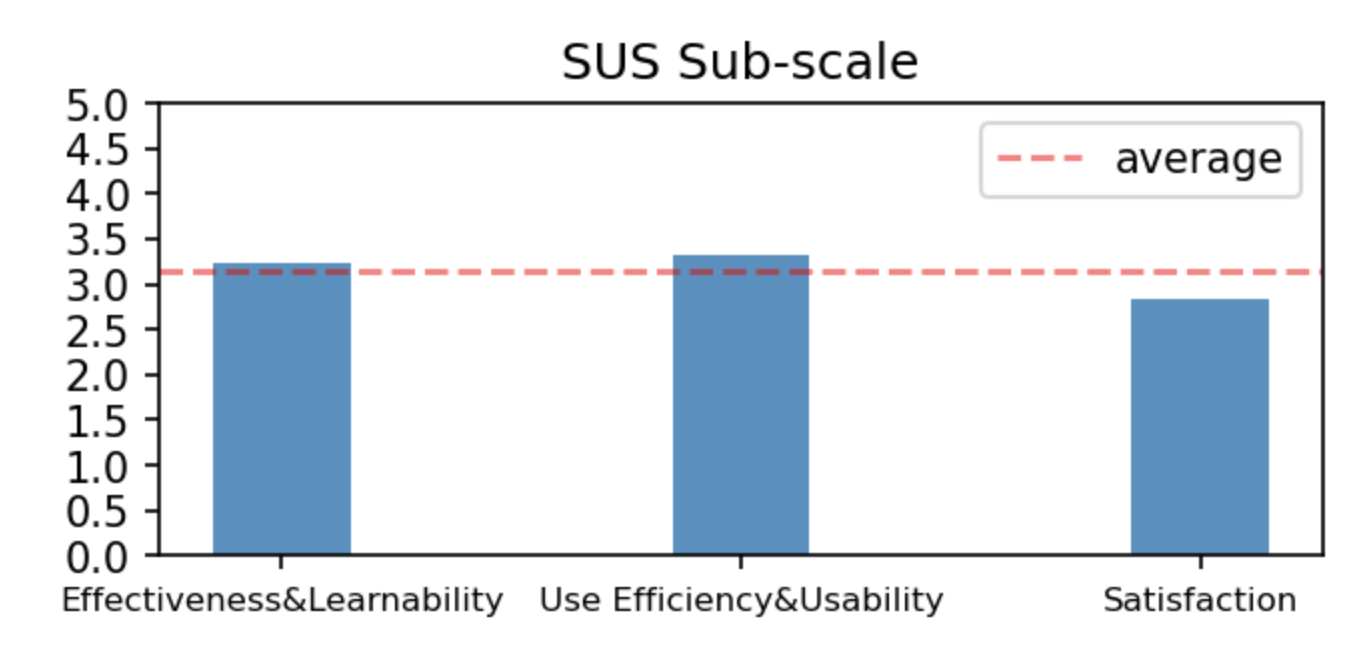
Throughout the test, younger participants with experience using social networking sites coucoucoucouldt time. In the case of Evaluator 11, the time to complete all tasks is 197 seconds. At the end of all tests, both experienced and inexperienced social network users completed all jobs in the task list smoothly. In addition, SafeTweet is not limited to a user's age, occupation, or even social network experience. To sum up, SafeTweet, as a prototype, has the characteristics of simple operation and easy learning.

Questionnaire analysis

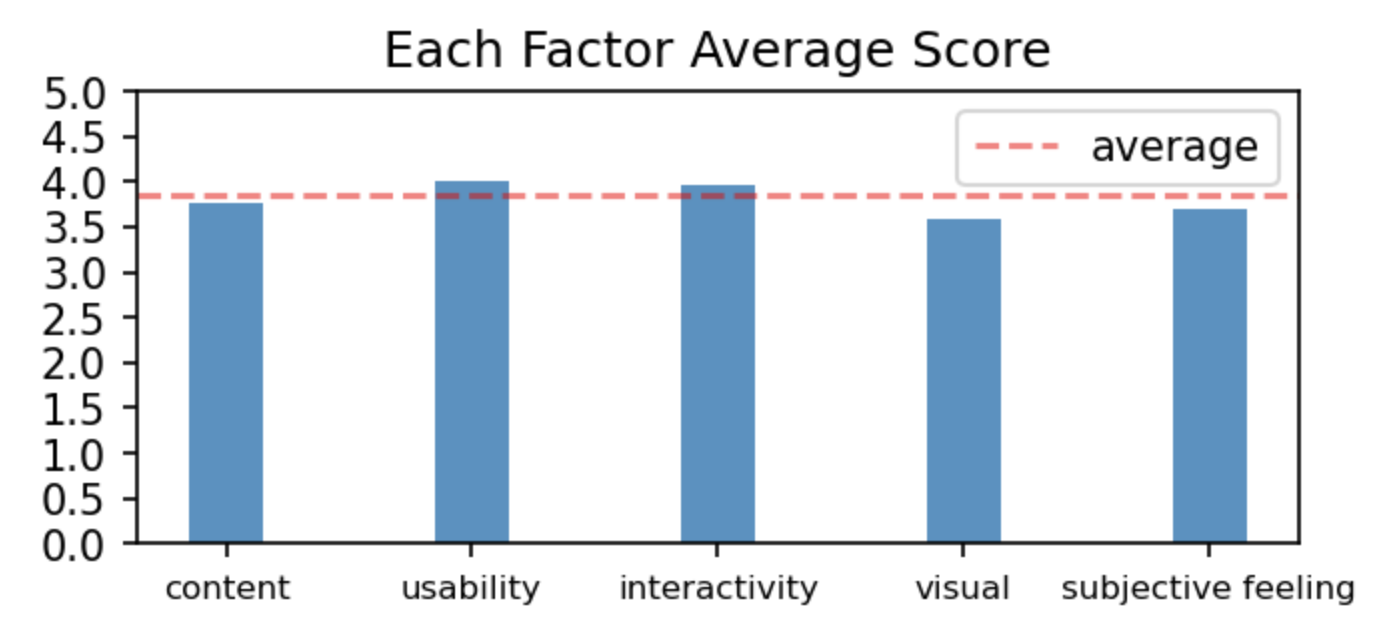
We used the SUS scale to collhonestreal user feed in the first part of the questionnaire. Brooke first compiled SUSBrooke in 1986. It can scientifically quantify user experience and measure the overall perceived availability of products or systems after completing a series of task scenarios, avoiding premature attention to details early. SUS provides an overall usability measurement consisting of 10 items with odd numbers for positive statements and even numbers for negative comments. The subscale "Effectiveness & Learnability" is formed by items 4, 5, and 10. The subscale composed of items 2, 3, 7, and 8 is "Use Efficiency & Usability". The first three items constitute the subscale "Satisfaction".

According to the box graph presented by the SUS questionnaire, the median line and the lower quartile line of the SUS score of SafeTweet are both above the mean streak, proving that SafeTweet's user experience is qualified. The bar chart shows SafeTweet's SUS subscale score, and the average SUS score is 3.15. SafeTweet has the best performance in Use Efficiency & Usability, with; r than the average score, indicating that there are specific problems in users' Satisfaction with SafeTweet. Therefore, we will analyse users' satisfaction with SafeTweet through the second and third parts of the questionnaire survey.





The second part of the questionnaire is based on the user experience questionnaire-customise dd by SafeTweet. The questions are designed around the quality influencing factors conducted before the start of the test, and there are 2-9 questions under each influencing factor. SafeTweet's quality factors are mainly divided into five parts: content, usability, interactivity, visual effects and subjective feelings.Thirty0 participants answered the UEQ questionnaire truthfully and accurately within 10 minutes after the test, with an average score of 3.84 for all final results. SafeTweet performed above average in usability and interactivity and nearly average in content with the score bar chart of influencing factors. The visual performance of the influencing factors is poor in THE score of UEQ, followed by the subjective feelings brought to the user, which are also lower than the mean.



Because SafeTweet was only a prototype for a social network, there was not enough front-end design. Users in the test can perceive this point, so users will be stricter in scoring this part. Secondly, the user experience with a lower-than-average score also reflects that the prototype product is not attractive enough to users. As for content influencing factors, SafeTweet is only a prototype at present, and its size is small, so it cannot fully display its quality potential content. So, its score is roughly in line with the average. However, in SafeTweet's usability and interactivity section, the above-average well and meeting perform meeting and meeting users' needs.

Visual Presentation and Subjective feelings in the second part of the questionnaire directly impact users' satisfaction with SafeTweet; especially the Visual effects are intuitive. Therefore, the low score of these two parts embodies the user satisfaction of SUS in the first part of the questionnaire. Subsequent improvements to the prototype should focus on visual effects to improve user experience and satisfaction.