Usability indicators + framework

Goal

To establish usability indicators for a platform thereby providing a measurement for user experience of various features. Indicators will support the ongoing product development team to make sure that the new features being developed are centered around user needs, and can enable growth of the user community.

About the usability indicators and framework

We have chosen usability indicators mentioned here. Of these, 4 (numbered 1-4) are high level indicators measuring —

- overall usability of the platform,
- how likely it is to be promoted by its users,
- user satisfaction and
- their 'feeling' of security and privacy.

These 4 indicators will be measured from user responses via surveys.

4 indicators (numbered 5-8) specifically measure the security features on 4 parameters —

- effort required by users,
- success rate,
- efficiency, and
- rate of errors

while using the features. These 4 indicators will be measured via user interviews.

While the indicators will offer a quantitative view of usability, usability testing is conducted for each feature for qualitative analysis which will provide recommendations and suggestions based on pain points and gaps in usability. More on this process can be found here.

How to use the indicators

We will analyze the scores and ask probing questions with the users to discover how the scores can be improved.

The indicators are scalable such that when new features are released beyond the scope of this project, similar survey + usability testing methodology can be applied.

Usability indicators

Legend:
👉 About
📋 How to collect input
🔮 Benchmark score
⏲️ Frequency/Measurement Timing
📐 How to measure

1. System usability score (SUS)
   👉 a general indication of the overall level of usability of a system. This score will provide a global view of usability without identifying specifically what works and doesn't work.
   📋 Survey - *The SUS score is collected through a user survey providing 10 standard statements that respondents evaluate on a 5-point strongly agree- to strongly disagree scale.*
   🔮 A "good" score is 68 or above.
   The choice of 68 as a benchmark is somewhat arbitrary but it has been found to be a good point of reference based on research and experience. In the original work by John Brooke, who developed the SUS, he suggested that a score around 68 could be considered as a "passable" usability. The idea is that a SUS score above 68 generally indicates that the system is perceived as usable, while a score below 68 suggests potential usability issues.
   ⏲️ Quarterly
   📐 Calculator. SUS score is calculated by first summing the score contributions from each item. Each item's score contribution will range from 0 to 4. For items 1,3,5,7,and 9 the score contribution is the scale position minus 1. For items 2,4,6,8 and 10, the contribution is 5 minus the scale position. Multiply the sum of the scores by 2.5 to obtain the overall value of SUS. SUS scores have a range of 0 to 100.

2. Net promoter score (NPS)
   👉 an indication of customer satisfaction and loyalty by asking how likely they are to recommend the product/platform to others
   📋 Survey - *Collected on a scale of 1-10 from not at all likely to extremely likely*
   🔮 Given the available -100 to +100 range, any score above 0 reads as 'good'.
   In the calculation, passive votes are ignored mainly because we're looking at the users who are likely to promote or detract. Promoters add to the score while detractors subtract the score. Since there's a likelihood that there are more detractor users, the score can go negative which is why anything above the neutral 0 reads as good.

   ⏲️ Quarterly
   📐 Calculator. NPS is calculated by subtracting the percentage of customers who answer the NPS question with a 6 or lower (known as 'detractors') from the percentage of customers who answer with a 9 or 10 (known as 'promoters').

3. User satisfaction score (USAS)
   👉 Indication of how happy the users are
   📋 Survey - *Collected on a 1-5 scale: "How satisfied are you with the product/platform?"*
   🔮 a good score will typically fall between 75% and 85%. On a total score of 40, it's good to score in the 30-35 range which corresponds to 75%-87.5%. Above that is exceptional.
   ⏲️ Quarterly
   📐 Calculator. Users are asked to rate the ease of their experience through a numerical 1-5 scale, or using emoticon anger-to-happiness scale. The collected answers are then averaged to give an idea of how much effort a certain process requires of customers.

4. User security score (USS)
   👉 Indication of how safe and secure the users feel
   📋 Survey - *The USS score is collected through a user survey providing 8 statements that respondents evaluate on a 5-point strongly agree to strongly disagree scale.*
   🔮 Higher the better
   ⏲️ Quarterly
   📐 Calculator. USS is a weighted average of user responses i.e. sum of user effort ratings divided by total number of responses.

5. User effort score (UES)
   👉an indication of the amount of effort a user had to exert to use the features
   📋 Survey (pop-up, in-app) - *example: "How easy was it for you to perform [specific feature] action?"*
   🔮 Higher the better on a scale of **1-5** or 1-7 (if the question is how easy it is to perform the task)
   ⏲️ Upon feature release
   📐 Calculator. UES is the total number of customers who agree that their interaction was easy divided by the total number of responses. For example, if 65 customers out of 100 rated you 4, 5 on the 5-point scale, your UES would be 65.

6. Task success rate (TSR)
   👉 an indication of how successfully users can perform tasks
   📋 User interview (mod/unmod)
   🔮 higher the better since this is in %
   ⏲️ Upon feature release
   📐 Calculator. Users will be scored upon completion of the tasks. TSR for each feature/task is a percentage of users who were able to complete the task or use the feature.

7. Efficiency (time to complete tasks)
   👉 an indication of how long it takes a user to complete a specific task
   📋 User interview (mod/unmod), *audit report if possible, or using a tool*
   🔮 lesser the time it takes, the better
   ⏲️ Upon feature release
   📐 This is calculated by observing the amount of time user spends accomplishing the task.

8. Error rate and data loss in managing data
   👉 an indication of how many errors a user encounters or error logs the system generates while using the features
   📋 User tests and audit logs as per the defined error states

🔮 lesser the better
⏲️ Upon feature release
📐 Error rate is calculated by summing the number of times a user encounters an error

User data to be collected:

This is required to view the above metrics with a wider context and run evaluations of the metrics, for example: viewing the system usability score for different user roles.

- User roles, to analyze the user experience indicators across different user roles and determine which ones need enhancements
- Organization name, to make sure feedback collected is representative of the product/platform's user base, and not the result of 1-2 organizations. Frequency of use, to contextualize the dependence of users on the application
- Session duration, same as frequency of use
- Device type, to debug faulty UX and understand which device needs to be prioritized for a better experience
- Browser type, same as device type

Usability framework

1. Usability indicators dashboard —
    a. We will track measurements after each survey.
    b. A dashboard will showcase graphical trends of change in indicators over time.
    c. All the indicator calculations will be presented separately.
2. Process for implementing the usability framework
    a. Collect responses via surveys
        i. Create a database of users with the following information: name, user role, organization, email.
        ii. Send the survey form over email every quarter. Using an automated mailer tool makes it easier to send the survey form every quarter or if needed on an adhoc basis.

iii. Allow 2 weeks to collect all the responses. Follow up with users who have not responded 1 week after sending the survey.
iv. Allow for 75% user responses to the survey for effective indicator calculation

b. Recording and calculating indicators
    i. Use the indicators calculator to calculate and record metrics into the tracker after each survey.

3. Setting up usability testing guidelines to generate recommendations and suggestions
    a. Based on the score, carry out in-depth user testing sessions to explore why the user gave a certain response, what works and doesn't work for them.
    b. The session duration can last 15 min to 45 mins depending on the complexity of the feature/task.
    c. Use the testing template from Appendix A to run the testing sessions.
    d. The direct responses received from users will feed into product improvements.

To be noted

1. The indicators themselves do not offer insights to improve the user experience. Usability testing will have to be performed to generate insights.
2. Since the indicators are directly based on user responses, connections with users are likely to influence the scores.
3. User security score must not be taken at its face value and should be cross-referenced with a technical security audit. Both the users' feelings and safety protocols are valuable representations of the state of security of the application.
4. This framework is applicable for all use cases that HROs use the product/platform for. Users will score as per their experience, and provide feedback as per their use case which may or may not feed into the product roadmap, depending on the product vision and the use cases.
    a. Case: a user responds poorly to the user satisfaction survey, we may choose to ignore the response if their use case and feedback is not aligned with the product vision. The inference might be that their use case is not the best fit for the product/platform and not that the product/platform doesn't have a good user experience.
5. If the product/platform users are too few, a small sample size may lead to an inconclusive outcome for indicators.

6. Ideally the features and the testing guidelines shouldn't change till user testing with all the users is complete. For example, if the steps of installation are further optimized while user testing is underway, and the UX is different for different users, then the indicators will lead to inconsistent outcomes.

Framework guidelines

1. The usability indicators dashboard can be extended when new features are introduced. For this, start by adding the feature to Appendix A along with a testing guide for the new feature followed by including it in the task success rate and frequency tracker.
2. Terminology "customer" and "user" have been used interchangeably.

Appendix-A: Tasks/Features and testing guide

Task/Feature	Testing guide (WIP)	Timing for User effort score(UES)	Comments
Task 1	Testing guide link		
Task 2	Testing guide link		
Task 3	Testing guide link		
Task 4	Testing guide link		
Task 5	Testing guide link		
Task 6	Testing guide link		
Task 7	Testing guide link		

Task 8	Testing guide link		
Task 9	Testing guide link		
Task 10	Testing guide link		