

## Sampling Design and Procedure for Evaluating User Experience and Usability using UEQ and SUS methods

Prepared by Simona Stojanova and Argene Superina.

With these guidelines we want to present the whole sampling design and procedure for testing the User Experience of a newly developed Augmented Reality (AR) mobile application. We will provide the general rules and steps, explaining them through an example.

1. **Define the population:** Define the group of individuals or objects from which you will be selecting the sample (e.g. faculty's employees and students in our case).
2. **Determine the sampling frame:** A sampling frame is a list of all the individuals or objects in the population. It represents a criteria for specifying the population of interest (e.g. gender, age, employment status, mobile phone software).
3. **Choose the sampling method:** There are several sampling methods to choose from. In general, sampling techniques can be divided into two types:
  - *Probability or random sampling* including simple random sampling, stratified sampling, cluster sampling, systematic sampling, multistage sampling.
  - *Non-probability or non-random sampling* including quota sampling, snowball sampling, judgement sampling, convenience sampling.

The method chosen will depend on the size of the population, the resources available, and the research question. In this case we used stratified sampling which ensures that subgroups (strata) of a given population are each adequately represented within the whole sample population of a research study. This sampling technique is a probability type of sampling. The "stratas" were formed based on the following characteristics: age, gender, employment status and type of mobile phone software. Then, we randomly selected proportional amounts of people from each group. This technique is a good choice because it allows to further study the subgroups within a population, to observe relationships between two or more subgroups, or to examine the rare extremes. Main disadvantage of stratified sampling is that it can be difficult to identify appropriate strata for a study and it can be more complex to organize and analyze the results compared to simple random sampling.

4. **Determine the sample size:** The sample size is the number of individuals or objects that will be included in the sample. It should be large enough to represent the population, but small enough to be manageable. Try to include about the same number of group representatives.
5. **Select the sample:** Once you have determined the sampling method and sample size, you can select the sample. This involves randomly choosing individuals or objects from the sampling frame.
6. **Collect the data:** After selecting the sample, you can collect the data through surveys, observations, or experiments. Since this article is focused on the two main methods UEQ and SUS, we continue with their short explanations as well as the demographic questions.

**Demographic questionnaire** – Demographic variables are examined through questions about age, gender, level of education, standard of living, employment status, profession/occupation and similar. Choose the ones applicable for your research. In our case, we also included questions related to the use of such AR apps (e.g: "How often do you use AR/VR/MR solutions?"). Question related to the type of their mobile phone software is also included here.

**Define tasks that the user should complete.** Define tasks in terms of questions which will determine the success rate of testing for each participant, based on their responses to these questions.

**System Usability Scale (SUS)** - This scale represents a standardized metric for **measuring the usability** of a digital product (tool, website, application, etc.). It is a 7-point Likert-type scale consisting of ten statements measured by the level of agreement for each statement (1- Strongly disagree; to 7- Strongly Agree). It is a widely used scale which covers three important aspects while measuring usability: effectiveness, efficiency, and satisfaction. SUS scores range from 0 to 100, but they do not represent percentiles. The average value for SUS is 68. Every value above 68 is perceived as an acceptable score (Stojmenova et al. ).

URL: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html> - choose your language.

**User Experience Questionnaire (UEQ)** – This scale represents a reliable questionnaire to measure the user experience of interactive products. It is available in more than 30 languages. Easy to use due to rich supplementary material. The questionnaire consists of 26 questions, which can be answered using a 7-point Likert scale (1 – Strongly disagree; to 7 – Strongly agree). The answers are used to provide scores on six aspects of user experience – Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation and Novelty. The scale is widely used since it allows data comparison between the scale variables and collected data can be easily compared with other, existing data or benchmarking (UEQ, online).

URL: <https://www.ueq-online.org/> - choose your language.

**Other details/Tips:**

- Make sure to provide the participants with proper instructions before and during the testing.
- If needed, assure that someone is controlling their testing process.
- Ask them to complete the questionnaire at the end, following the same order of questionnaires.
- Make sure that you provide a statement of consent that each participant and the institution will agree that the collected data can be further used for scientific research purposes.
- Provide incentives for participants (e.g. free lunch at the faculty restaurant in our case).

7. **Analyze the data:** Once you have collected the data, you can analyze it using statistical techniques to draw conclusions about the population.

For data analysis of the SUS, a simple formula has to be applied. For the statements that are odd (1, 3, 5, 7 and 9), you have to follow the rule “Response – 1” (to subtract 1). Let’s suppose that statement 5 has been evaluated as 3,  $3 - 1 = 2$ , you'll compute 2. For the even statements (2, 4, 6, 8 and 10), you have to follow the rule “5 – Response” (to subtract from 5). Let’s suppose that the value for the 4th statement was 4. Then we have  $5 - 4 = 1$ , so you'll compute 1. This will give you the the individual values of SUS for each partner and for each statement. The sum of these individual values from each respondent are then multiplied by a factor of 2.5 than summed together and divided with the number of participants.

The UEQ survey has their Data analysis tool. It is an excel tool, which does all the required calculations. You only need to insert the dataset into the “Data” sheet placing each participant in a different row, then all the statistical analysis is done automatically. You can download the Excel-Sheet file named “*Data Analysis Tool (UEQ-S)*” from here: <https://www.ueq-online.org/>

## References:

1. Stojmenova, K., Jakus, G., Tomažič, S., Sodnik, J. (2022). Is less really more? A user study on visual in-vehicle information systems in automated vehicles from a user experience and usability perspective. In: Tareq Ahram and Christianne Falcão (eds) Usability and User Experience. AHFE (2022) International Conference. AHFE Open Access, vol 39. AHFE International, USA. <http://doi.org/10.54941/ahfe1001741>
2. UEQ online. Data Analysis Tools. Available from <https://www.ueq-online.org/>