

QUESTIONNAIRES FOR USER EXPERIMENTS

Introduction

When you plan an experiment, try to look for a standard tool, method, or questionnaire to assess your variables. When you try to assess latent variables, such as workload, stress, anxiety, usability, sickness, pleasantness, etc., a standard questionnaire is fundamental. Such questionnaires have undergone psychometric qualification, including assessment of reliability, validity, and sensitivity, making them more valuable tools than those made ad-hoc.

Here in this document, there are examples of questionnaires (standard and ad-hoc) used for user experiments in Virtual Reality and Human-Computer Interaction. The tips provided in this document should not keep the reader from study more about standards for design and application of experiments.

1. Subject Characterization

Before the experiment even starts, information about the subject is collected with a simple form, often referenced as “Demographic Questionnaire”. The questionnaire should aim in collecting as much as information the experimenter judge necessary to understand the behaviour of the subject during the experiment and the data collected. It should not be just a collection of random questions about the user, but rather based on the literature about the variables assessed. For instance, there are experiments that shows an effect of age and sex in tactile perception, so such information must be addressed in the data analysis. Moreover, there are experiments that present an effect of the dominant hand on veering for blind and blindfolded pedestrians. Therefore, in such scenario, the experiment want to be sure about the subject handedness.

1.1 Suggested Questions

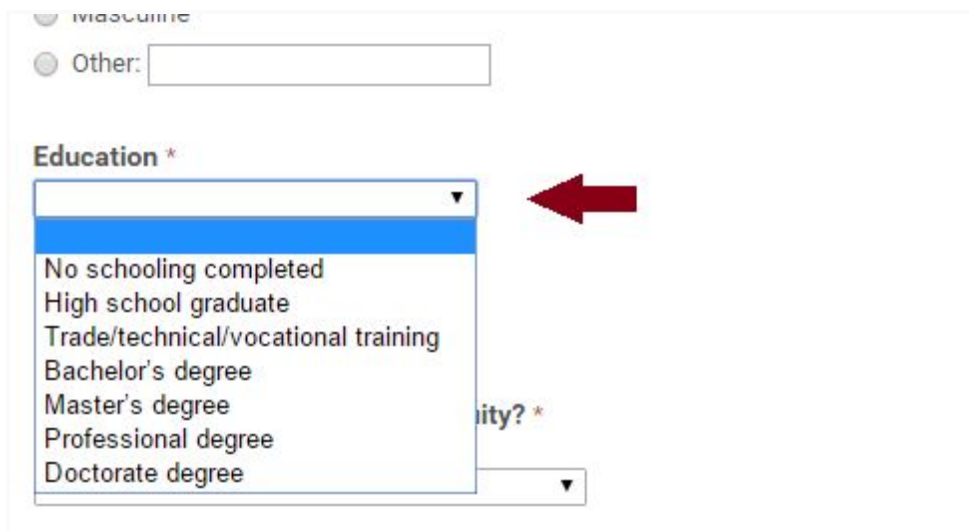
The following items are not part of a template, but are suggestions based on previous user experiments, related work and other questionnaires.

Item	Description
User ID	Some subjects do not like to provide their names as identifier. The ID can be an integer provided by the tester.
Timestamp	Date and time when the subject fulfilled the questionnaire. That item is automatic when using online tools, such as Google Forms.
Age	In decimal. If you use the format DD/MM/AAAA, make sure you registered the Timestamp to calculate the age of the subject in the day of the test.
Primary Language	In your Institute or University, it may be common to use your native language, but you may have a subject that do not speak it natively.
Education	Highest degree or level of school the subject has completed. If currently enrolled, highest degree received. Find a suggestion in Section 1.2.

Gender/Sex	Feminine-Masculine-Other for gender, or Female-Male-Other for sex.
Handedness	Dominant Hand. Can be left, right, or ambidextrous. There is a better way to assess handedness rather than ask the user. See Section 1.3.
Previous Experience	Does the subject have experience with techniques and applications related or similar to the one tested? See more in Section 1.4.
Naivety to the Protocol	Did the subject perform a previous version of the current setup? Does the subject have any knowledge about the test and protocol?
Disability/Medical Condition	Any condition that may affect the user's performance, like vision/hearing disability, skin problem (in tactile setups), attention disorder, etc.
Contact	Name and email of the subject can be provided optionally in case he/her want to receive information about the test after.

1.2 Education

There are a lot of different examples of how to add this item on a demographic questionnaire. A good way to do it is as a list. The user should chose the highest degree or level of school the subject has completed. If currently enrolled, it should be the highest degree received.



The screenshot shows a portion of a questionnaire. At the top, there are radio buttons for 'Masculine' and 'Other:'. Below this is the 'Education' field, which is a dropdown menu. The dropdown is open, showing a list of education levels. A red arrow points to the dropdown menu. The list contains the following items: No schooling completed, High school graduate, Trade/technical/vocational training, Bachelor's degree, Master's degree, Professional degree, and Doctorate degree. To the right of the dropdown, there is a label 'ity? *'.

Figure 1. Education item.

The list can contain:

- No schooling completed
- High school graduate
- Trade/technical/vocational training
- Bachelor's degree
- Master's degree
- Professional degree
- Doctorate degree

1.3 Handedness

Handedness is not just a trivial variable. When you ask the user about his/her dominant hand or hand of preference, you may receive an answer based on what the user thought would be the best in that given time. To better assess handedness, there are standard questionnaires for it. A very used and well received tool is the Edinburgh Handedness Inventory. A 20-item scale that can be found in Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.

There is also a tested 8-item version of the Edinburgh Inventory, that can be used instead:

Please mark the box that best describes which hand you use for the activity in question: *

	Always Left	Usually Left	No Preference	Usually Right	Always Right
Writing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Throwing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scissors	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Toothbrush	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Knife (without fork)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Spoon	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Match (when striking)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Computer mouse	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 2. Eight-item Edinburgh Inventory.

In addition, there is a 4-item version that deserves attention: Veale, J. F. (2014). Edinburgh Handedness Inventory–Short Form: a revised version based on confirmatory factor analysis. *Laterality: Asymmetries of Body, Brain and Cognition*, 19(2), 164-177.

1.4 Previous Experience

That item is not related to the previous experience with the research question and applied protocol, to which the subjects should be naive (to avoid biases, and as a matter of fairness). This item concerns the ability and experience of the subject with related technology and methods. We want to know, for instance, if the performance of a subject when manipulating a 3D object was affected by his/her skills in playing 3D games, or his/her research on 3D manipulation.

Some useful questions in this topic:

- [In the last six months, (optional)] how often have you been playing video games?
- When you play, how often do you use a joypad rather than the keyboard?
- Which of the following game controllers have you used before?
- Have you ever tried a Head-mounted Display?
- Have you ever experienced Virtual Reality?
- Have you ever used a device for cutaneous stimulation or force feedback?

Those questions are not standard, but address important characteristics of the subject tested.

1.5 Perceptual Conditions

It is important to address vision and hearing disabilities and other kinds of perceptual conditions before the test starts. Even if the experiment should be designed to be accessible to different user profiles, there are profiles that we do not want to address as a variable. For instance, if you assess a method for visual feedback you have to guarantee that the subject is able to see the feedback.

1.5.1 Vision

Examples:

How do you classify your visual acuity? (Degree of vision loss)

- Normal Vision (≥ 0.8)
- Mild Vision Loss (< 0.8 and ≥ 0.3)
- Moderate Vision Loss (< 0.3 and ≥ 0.125)
- Severe Vision Loss (< 0.125 and ≥ 0.05)
- Profound Vision Loss (< 0.05 and ≥ 0.02)
- Near-total Vision Loss (< 0.02 and $\geq \text{NLP}$)
- Total Vision Loss (NLP)

How do you classify your field of view? (Degree of field loss)

- Range of Normal Vision (average field radius: 60°)
- Mild Visual Impairment (50° to 40°)
- Moderate Visual Impairment (30° to 20°)
- Severe Visual Impairment (10° to 8°)
- Profound Visual Impairment (6° to 4°)
- NearBlindness/Blindness (2° to 0°)

Do you have any other vision problem?

e.g. Myopia, Hyperopia, Color Blindness, Astigmatism...

How many years since you were diagnosed with that vision problem/loss?

Only if you have loss of vision acuity, field of view or other vision problem

1.5.2 Hearing

Example:

How do you classify your hearing acuity? (Degree of hearing loss)

- Normal (less than 15dB)
- Slight (16dB to 25dB)
- Mild (26dB to 40dB)
- Moderate (41dB to 55dB)
- Moderately Severe (56dB to 70dB)
- Severe (71dB to 90dB)
- Profound (more than 91dB)

1.5.3 Touch

Example:

Do you have any different characteristic on the skin of the <stimulated loci>?

e.g. Loss of sensibility, Scar, Hyperhidrosis...

2. Usability

Standardized usability questionnaires are questionnaires designed for the assessment of perceived usability, typically with a specific set of questions presented in a specified order using a specified format with specific rules for producing scores based on the answers of respondents.

For usability testing, standardized questionnaires are available for assessment of a product at the end of a study (post-study) and after each task in a study (post-task).

2.1 Post-study Questionnaires

Examples of post-study questionnaires:

- SUS - System Usability Scale (10-item) *is one of the most used
- QUIS - User Interface Satisfaction (6 to 27 items)
- SUMI - Software Usability Measurement Inventory (50-item)
- PSSUQ - Post-Study Usability Questionnaire (13 to 19 items)
- UMUX - Usability Metric for User Experience (4-item)
- UMUX-LITE (2-item)

2.2 Post-task Questionnaires

Examples of post-task questionnaires:

- ASQ - After Scenario Questionnaire (3-item)
- SEQ - Single Ease Question (1-item)
- SMEQ - Subjective Mental Effort Questionnaire (1-item)
- UME - Usability Magnitude Estimation (1-item)

2.3 Website Usability

Standardized questionnaires are also available for the general assessment of website usability. For example:

- WAMMI - Website Analysis and Measurement Inventory (20-item)
- SUPR-Q - Standardized User Experience Percentile Rank Questionnaire (8-item)

3. Workload

There are several questionnaires for assessment of perceived workload, but the most popular seems to be the NASA-TLX (NASA Task Load Index). The NASA TLX has been developed by NASA to assess the relative importance of six factors in determining how much workload the subject experienced: Mental Demand, Physical Demand, Temporal Demand, Effort, Performance, and Frustration.

The NASA-TLX is a two step evaluation tool. First there is a weighting process (Sources of Load) and then the 6-item scale (Magnitude of Load).

The instructions to apply the NASA TLX can be found here:
<http://humansystems.arc.nasa.gov/groups/tlx/paperpencil.html>

A version as Google Forms can be found here:
<http://goo.gl/forms/mduSIWkQrSVcaswg1>

4. Anxiety

The State-Trait Anxiety Inventory (STAI) is a commonly used measure of trait and state anxiety. The Form Y, its most popular version, has 20 items for assessing trait anxiety and 20 for state anxiety. State anxiety items include: “I am tense; I am worried” and “I feel calm; I feel secure.” Trait anxiety items include: “I worry too much over something that really doesn’t matter” and “I am content; I am a steady person.” All items are rated on a 4-point scale (e.g., from “Almost Never” to “Almost Always”). Higher scores indicate greater anxiety. A version as Google Forms can be found in here:
<http://goo.gl/forms/yFvdffs3WIPCUu8D2>

There is also a 6-item version: Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31(3), 301-306.

5. Sickness and Cybersickness

A popular questionnaire used to assess sickness using Virtual Reality devices is the SSQ (Simulator Sickness Questionnaire). The SSQ is a 27-item scale correspondent to a list of 27 symptoms which are commonly experienced by users of virtual reality systems. Each item is rated with the scale from none, slight, moderate to severe.

With it, four representative scores can be found: Nausea-related subscore, Oculomotor-related subscore, Disorientation-related subscore are the scores for the symptoms for the specific aspects. Then, the total Score is the score representing the overall severity of cybersickness experienced by the users of virtual reality systems.

The SSQ is applied before and after the procedure as a Pre-exposure SSQ and Post-exposure SSQ.

6. Presence

A popular questionnaire used to assess the sense of presence in a virtual environment is the IPQ (Igroup Presence Questionnaire). The IPQ is a 14-item scale. To apply it, more information can be found in:
<http://www.igroup.org/pq/ipq/index.php>

7. ...

Feel free to add more standard questionnaires, such as for the assessment of accessibility, and for collaborative tasks.

More Informations

Recommended readings:

- ETS Standards for Quality and Fairness: <https://www.ets.org/s/about/pdf/standards.pdf>

- Diretrizes e Normas Regulamentadoras de Pesquisas Envolvendo Seres Humanos:
<https://www.ufrgs.br/bioetica/res19696.htm>
- Declaração de Helsinque:
http://www.amb.org.br/_arquivos/_downloads/491535001395167888_DoHBrazilianPortugueseVersionRev.pdf

Contributions

Author	Last Edited
Victor Adriel de J. Oliveira	May 30, 2016