UNIVERSITY OF TWENTE.

Exam HCI Design & Evaluation
Part II

Module: Intelligent Interaction Design

Bachelor 2 BIT/TCS (EWI)

Module code: 202001032
Date: example
Time: example
Module-coordinator: R. Klaassen
Instructor: R. Klaassen, L. Gatti

Type of exam: Closed book
Allowed aids during the exam: Nothing
Attachments: Multiple Choice form

The test consists of two sections:

- A section with multiple choice questions. Each multiple choice question has one correct answer.

All questions in the multiple choice part have equal weight. For the multiple choice section a standard correction for guessing is applied. For the general case of *n* questions and *p* correct answers the formula for the score of the multiple choice section is:

$1 + 9*\max((p - 0.25*n)/(n - 0.25*n), 0)$

This formula can be changed after the analysis of the answers.

Before leaving the room you must hand in

- this question form
- the answer form for the multiple choice questions

1. A researcher shows people either a horror movie or a comedy movie. After the movie, they give a survey that asks people to rate how much they liked the movie and how frequently they view movies each week. They look at the relationship between the movie type, the liking of the movie and the frequency of watching movies. Consider the following statements.

(i) Movie type (horror or comedy) is a categorical variable.
(ii) Frequency of watching movies is a dependent variable.

Are these statements true?

(a) Both statements (i) and (ii) are true.
(b) Only statement (i) is true.
(c) Only statement (ii) is true.
(d) Both statements (i) and (ii) are false.

2. A researcher wants to do a controlled experiment to test whether people prefer viewing an app on tablet or laptop screen, but wants to achieve the lowest probability for participants to learn the purpose of the study. Which of the following study designs is appropriate?

(a) Have all participants look at the tablet in one study session and the laptop in another study session one week later
(b) Have all participants look at both the tablet and the laptop and rate them within a single study session
(c) Have half participants look at the tablet and half participants look at the laptop and rate it
(d) Allow participants to choose which version to look at and then rate it

3. You run a study in which each participant in the study completes vehicle driver training by either: watching a video, doing a virtual reality task, or driving the real vehicle itself. After training, you measure each participant's performance driving the real vehicle itself. Which statistical test should be used?

(a) Factorial analysis of variance
(b) Independent samples t-test
(c) One-way analysis of variance
(d) Paired-samples t-test

4. I measure the effect of robot color on people's mood. I have a two condition study (robot color: blue or red, between subject design). The data is not normally distributed. Mood is measured on a 100-point Likert scale. Which statistical test should be used?

(a) Independent samples t-test
(b) Mann-Whitney U test
(c) One-way analysis of variance
(d) Paired-samples t-test

5. Consider the following two statements about experimental research:
(i) Between-participant studies require more participants than within-participant studies.
(ii) Between-participant studies need to worry about order effects more than within-participant studies.

Are these statements true?

(a) Both statements (i) and (ii) are true.
(b) Only statement (i) is true.
(c) Only statement (ii) is true.
(d) Both statements (i) and (ii) are false.

6. Consider the following claims about users with impairments

(i) Proxy users are users who have a different impairment than the impairment being studied.
(ii) It is better to schedule users with impairments at their homes.

Are these statements true?

(a) Both statements (i) and (ii) are true
(b) Only statement (i) is true
(c) Only statement (ii) is true
(d) Both statements (i) and (ii) are false

7. For a certain expert-based test it is important to select the right tasks. Consider the two following statements about the type of tasks to be included. Which of the statements is true?

(i) High frequency-tasks should be included
(ii) Rarely occurring but important tasks (such as error recovery) should be included.

Are these statements true?

(a) Statement (i) and statement (ii) are both false
(b) Statement (i) is true; statement (ii) is false
(c) Statement (i) is false; statement (ii) is true
(d) Statement (i) and statement (ii) are both true

8. Name one advantage and one disadvantage of a heuristic review (also called heuristic evaluation) as compared to user testing?

9. The Wilcoxon signed ranks test is the non-parametric alternative for the paired t-test. When do we use the test instead of a paired t-test?

(a) The data is collected by means of a Likert scale
(b) The independent t-test is not applicable
(c) In case the data is collected from a within-group design but it is not from a population that is

normally distributed
(d) In case we cannot use the X^2 test

10. Which of the following is not a typical dependent variable in HCI research?

(a) Subjective satisfaction
(b) Age
(c) Ease of learning
(d) Time to complete a task

11. Given are two statements related to the use of informed consent in HCI research.

(i) Participants should always sign an informed consent before being allowed to participate in a research study

(ii) Participants should always be notified of their participation beforehand and at least give verbal assent before being allowed to participate in a research study

Are these statements true?

(a) Both statements (i) and (ii) are true
(b) Only statement (i) is true
(c) Only statement (ii) is true
(d) Both statements (i) and (ii) are false

12. One specific category of usability testing covered in the book includes cognitive walkthrough. Please answer all of the following three questions.
a) Who is the type of person doing the cognitive walkthrough, what background would they have?
b) What is a good description for explaining the two important elements in doing this activity?
c) And, to what type of understanding does it lead?

Answers

1. a, slides week 6, Lazar Ch. 2.2.2
2. c, Lazar Ch. 3.3.1
3. c, Lazar Ch. 4.4., 4.5
4. b, Lazar Ch. 4.8.2
5. b, Lazar Ch. 3.
6. c, Lazar Ch. 16
7. d, Lazar Ch. 10 (p. 269)
8. See Lazar Ch. 10 (and slides)
9. c, Lazar Ch. 4

10. B, Lazar Ch.2 (in HCI we are interested in effects of system (systems with a purpose) and helping in how to create these. Look at the slides about different kind of variables. Age is thus not only not a dependent variable (outcome) but seldomly an independent variable in most HCI research experiments)

11. D, in some exceptional cases for instance when researching security issues aka covert research, informing the participant beforehand is deemed unfeasible. Sometimes participants are only partially informed or deceived (that is another issue). In all of these cases special care has to be taking in the research protocol and additional time and discussion can be expected in getting an ok for the Ethical Committee assessment as it might not fall under the expedited version.
Assent, giving an approval or ok, often related to children below 13 agreeing to willingly participate, would also have similar issues.

12a. "A cognitive walkthrough is **an expert review method** in which **interface experts simulate users**" and they have a good understanding of who the users are and what tasks they are expected to do.

12b. The usability expert makes a **set of tasks to go through while making notes of usability issues they encounter**, these tasks include **central tasks (recurring or high-frequency)** as well as **critical tasks (rarely occurring but important)**, to find the usability issues the users encounter the expert consistently ask questions about "will the user know what to do to get to the outcome they expect?" (this latter is adapted from Interaction Design Foundation, not in the book/lecture so not needed for the points)

12c. This could lead to **a good overview of how users might interact the first time**, although some people feel it is not as productive as user-based testing.