




CPUX-UT

Checklist for the evaluation of the practical examination
for the CPUX-UT certificate

Version 1.15, 1 November 2020



Publisher: UXQB e. V.
Contact: info@uxqb.org

www.uxqb.org

Contents

1	Introduction	3
1.1	The most important reasons for failing the practical examination	4
1.2	Legends.....	5
1.3	Acknowledgments.....	5
2	Formal requirements	6
3	Test script.....	8
4	User profiles and test participants	9
5	Test tasks	11
6	Test session in general	14
7	Briefing and interview.....	16
8	Moderation.....	18
9	Debriefing	21
10	Usability test report – Layout and contents	22
11	Usability test report – Executive summary	23
12	Usability test report - Findings	24
13	Problems that are not covered by the criteria	27
14	Commendable aspects in the submission.....	28
	Appendix. Important changes compared to previous versions	29

1 Introduction

This document contains the evaluation criteria for the practical test leading to the CPUX-UT certificate. The evaluation criteria are used by the evaluator to score deliverables from the usability test conducted by candidates seeking to achieve the CPUX-UT certificate.

This document also informs CPUX-UT candidates about how their CPUX-UT deliverables are scored.

The basic idea behind the rating is that each candidate starts out with 100 points. Points are deducted for each deviation from a good usability test as reflected by the evaluation criteria.

This usability test checklist serves as guidance for the examiner. The examiner is not limited to the items in the usability test checklist. The examiner may deduct points for major or critical problems that are not covered by the checklist.

1.1 The most important reasons for failing the practical examination

ID	Problem	Checklist item
1	<p>Bad audio quality It is difficult to hear what the test participant or the moderator are saying. Use a separate microphone for the recording. Do not use the built-in microphone on a laptop. Check the quality of your audio by making a test recording. Also, check the audio quality after the first usability test session.</p>	2.10
2	<p>Bad video quality It is difficult to read the contents of the screen on the video. Check the quality of your videos by making a test recording. Also, check the video quality after the first usability test session.</p>	2.11, 2.12
3	<p>One or more videos are longer than 30 minutes, or a test session ends abruptly to keep it within 30 minutes.</p>	6.4
4	<p>One or more usability test tasks do not match the requirements in the task description for the practical examination, in particular the prescribed general tasks. The examiners consider it essential that a usability professional is able to follow precise instructions, such as those provided in the task description for the practical examination, exactly.</p>	5.1 to 5.4
5	<p>The task set contains more than four tasks. Exactly 4 tasks are required for the practical examination.</p>	5.5
6	<p>One or more usability test tasks are trivial or have hazy goals. Symptoms: All four tasks are completed within 10 minutes. It's hard to identify more than 10 findings. Most of the findings are trivial.</p>	5.8
7	<p>The moderator is not sufficiently curious</p>	7.8, 8.10
8	<p>The moderator talks too much. The moderator asks questions about issues that can be clearly observed. The moderator turns the usability test session into an interview, for example by repeatedly asking "What would you do?"</p>	8.1, 8.2, 8.7
9	<p>The usability test report describes findings that are not evident from the videos. Such findings could be results from inspections or test participants' opinions</p>	12.14
10	<p>The usability test report does not describe important findings that are evident from the videos</p>	12.15

1.2 Legends

ID	The IDs are used for references.
Evaluation criterion	The criteria reflect key criteria in the CPUX-UT Curriculum.
E1	The maximum number of points that can be deducted if the criterion is violated.
E2	The actual rating (zero or number of points to be deducted) given by the examiner. If the rating is <i>OK</i> , <i>No problem</i> or <i>Good</i> , enter 0.
Evaluator's comment	Justification of the rating, if required. References like <i>V3-05.38</i> may be used in the comments to refer to a sequence in a video submitted by the candidate. The sample reference refers to the video for test participant 3. The referenced event starts approximately 5 minutes and 38 seconds after the start of the video.

1.3 Acknowledgments

This document was created by the following persons:

Rolf Molich (Editor and examiner)

Bernard Rummel (Co-Editor and Examiner)

Susanne Waßerroth (Examiner)

Ludwig Meyer (Examiner)

2 Formal requirements

The examiner starts by checking that the formal requirements for the results in this section are fulfilled.

If one or more of the following 12 formal requirements are not fulfilled, the examiner must stop their assessment and ask the candidate to submit the missing results or resubmit the unsatisfactory results in an improved form within three working days. This resubmission is free of charge. Only one resubmission is allowed. Read more about this situation in the CPUX-UT Examination Regulations.

Points may be deducted for minor violations of the formal requirements.

According to the CPUX-UT Examination regulations, the formal requirements are:

ID	Formal requirements	Yes/ No
1	A usability test report must be available	
2	A usability test script must be included in the usability test report	
3	The usability test script must describe the four usability test tasks that were used for the usability test	
4	The usability test report must contain at least 5 substantial findings	
5	A “Declaration of Original Work and Consent Agreement” signed by the candidate must be available	
6	An unsigned copy of the Informed Consent Declaration that the test participants have signed must be available Note: Do not submit signed Informed Consent Declarations. This would violate the data protection rules.	
7	Exactly 3 videos must be available	
8	The videos must be easily viewable on a Windows computer	
9	Each video must last at most 30 minutes	

ID	Formal requirements	Yes/ No
10	<p>It must be easy to understand what is being said on the videos</p> <p>Note: Making a useful recording is a basic qualification. If it is impossible to understand what is being said on the videos, the candidate fails the examination.</p>	
11	<p>Both the test participant and the moderator must be clearly visible on the videos.</p> <p>Note: It should be possible to see whether or not the moderator takes notes during the test session</p>	
12	<p>On the video it must be possible to read text that appears in reasonable size on the screen</p> <p>Note: Making a useful recording is a basic qualification. If it is impossible to read the contents of the screen on the videos, the candidate fails the examination.</p>	

3 Test script

ID	Evaluation criterion	E1	E2
1	Are the criteria used for selecting test participants available and reasonable?	2	
2	Does the test script include a suitable checklist for the briefing?	5	
3	Does the moderator adhere to the checklist for the briefing?	3	
4	Does the test script include a suitable checklist for the interview of the test participant?	5	
5	Does the moderator adhere to the checklist for the interview of the test participant?	3	
6	Does the test script include a suitable checklist for the debriefing? Note1: Limit the debriefing to the questions in “Debriefing”, note 3. Note 2: Do not ask the test participant to rate their experience on a scale.	5	
7	Does the moderator adhere to the checklist for the debriefing?	3	

4 User profiles and test participants

ID	Evaluation criterion	E1	E2
1	<p>Do the test participants match the requirements in the assignment?</p> <p>Note 1: Do not use UX professionals, IT system administrators and people who develop or implement IT systems as test participants.</p> <p>Note 2: If a test participant turns out to be unqualified for the test, a replacement should be scheduled and the test session with the unqualified test participant should not be submitted</p> <p>Note 3: Lack of qualifications includes insufficient command of the language used in the usability test session as well as insufficient experience with the internet.</p>	12	
2	<p>Is brief but adequate information about each test participant available?</p> <p>Note: Observe “List of test participants”, note 3.</p>	2	
3	<p>Are the test participants sufficiently different to constitute an informative sample?</p>	2	
4	<p>Does the usability test report include a reasonable user profile?</p> <p>Note: See the example in the sample usability test report, “Target group for the system”</p>	2	
5	<p>On the basis of their appearance on the video, do test participants seem to match the information about test participants provided in the usability test report?</p>	6	

CPUX-UT – Usability Test Checklist

6	Is the anonymity of test participants preserved? Note: Observe “Sensitive personal information”, note 3.	3	
---	---	---	--

5 Test tasks

ID	Evaluation criterion	E1	E2
1	Does test task 1 match the requirements in the assignment? Note: Only use cases that are covered by the assignment – that is, the task description for the practical examination – may be tested.	5	
2	Does test task 2 match the requirements in the assignment?	5	
3	Does test task 3 match the requirements in the assignment?	5	
4	Does test task 4 match the requirements in the assignment?	5	
5	Are there exactly four usability test tasks?	10	
6	Does each usability test task include a precise success criterion or goal? Note 1: This is particularly important for open-ended tasks. If the task set includes open-ended tasks, the usability test report should describe the goals that each test participant set for the task before they started on it. Note 2: An example of a test task that does not have a precisely defined goal for a car rental website is “Find information about insurance”.	4	
7	Does each test task address exactly one basic function of the interactive system? Note: A test task that tests several use cases should be split.	3	

CPUX-UT – Usability Test Checklist

8	<p>Does task 1 have sufficient diagnostic value?</p> <p>Note 1: See “Usability test tasks”, note 3. Each test task must have new diagnostic value. Small but significant differences between usability test tasks are acceptable but should be carefully justified.</p> <p>Note 2: Most often, trivial or unrealistic tasks have little diagnostic value</p>	4	
9	Does task 2 have sufficient diagnostic value?	4	
10	Does task 3 have sufficient diagnostic value?	4	
11	Does task 4 have sufficient diagnostic value?	4	
12	<p>Is each test task description for the test participant short?</p> <p>Note: Do not deduct any points if a mistake of this kind is corrected after the first usability test session.</p>	3	
13	<p>Is each test task description for the test participant unambiguous and clear?</p> <p>Note: Do not deduct any points if a mistake of this kind is corrected after the first usability test session.</p>	3	
14	<p>Is each test task relevant from the test participants’ point of view?</p> <p>Note: Observe “System-oriented tasks”</p>	3	
15	Are test tasks free from clues?	3	

CPUX-UT – Usability Test Checklist

16	Is the purpose of each test task explained or self-evident? Note: A description of the purpose of each task should be included in the task description.	3	
17	Is the starting condition for each test task specified?	1	
18	Are test tasks serious – that is, not humorous?	3	
19	Are pretender tasks avoided? Note: Tasks must not imply political and religious attitudes that test participants might not have.	6	

6 Test session in general

ID	Evaluation criterion	E1	E2
1	<p>Is each test session unaffected by previous test sessions?</p> <p>Note: This could happen, for example, if the cache is not cleared properly between sessions, so input suggestions from previous sessions appear.</p>	2	
2	<p>Are test sessions conducted in a suitable room? Is the room undisturbed?</p>	5	
3	<p>Does the moderator manage time well so as many test tasks as possible are addressed properly within the given time frame for the test session?</p> <p>Note: Cutting off the video during a task or skipping the debriefing in order to keep the video shorter than 30 minutes is not acceptable.</p>	5	
4	<p>Is each test session shorter than 30 minutes?</p> <p>Note: The maximum time for a test session is 30 minutes. The maximum time should be carefully observed even if this means that the test participant does not complete all 4 test tasks.</p>	12	
5	<p>Are there no recurring hardware malfunctions during the test session?</p>	5	

CPUX-UT – Usability Test Checklist

6	<p>Are there no recurring software malfunctions during the test session?</p> <p>Note: This includes annoying and irrelevant pop-ups that could have been easily avoided.</p>	5	
7	<p>If there are hardware or software malfunctions during a test session, are they handled diplomatically and competently by the moderator?</p>	5	

7 Briefing and interview

ID	Evaluation criterion	E1	E2
1	In the briefing, does the moderator orally inform each test participant "We are not testing you"?	5	
2	At the end of the briefing, does the moderator ask "Do you have any questions?"	1	
3	Is the briefing efficient? For example, demonstrations of what think-aloud is are not considered efficient.	5	
4	In the briefing, does the moderator avoid think-aloud instructions like asking for comments, evaluations, and descriptions of actions.	3	
5	In the interview, does the moderator enquire about the test participant's previous knowledge of the interactive system?	3	
6	In the interview, does the moderator enquire about the test participant's previous knowledge of comparable interactive systems?	1	
7	In the interview, does the moderator appear genuinely interested in the test participant? Note 1: Does the moderator have regular eye-contact with the test participant? Note 2: Does the moderator take notes?	5	

CPUX-UT – Usability Test Checklist

8	<p>Is the moderator sufficiently curious during the interview? When appropriate, does the moderator deviate from the interview checklist?</p> <p>Note: Follow-up questions are appropriate for example if the test participant describes an important use case that has not yet been considered.</p>	5	
---	--	---	--

8 Moderation

ID	Evaluation criterion	E1	E2
1	Does the moderator say as little as possible while the test participant is solving test tasks?	5	
2	Does the moderator say as little as possible between test tasks?	5	
3	Are the moderator's remarks and body language free from clues? Note 1: Sample clues are "Don't worry about this", or the body language says "You have almost solved the task. The answer is on the screen." Note 2: The moderator should not explain in any way how the website works.	5	
4	Does the moderator avoid premature confirmation – that is, confirming the solution before the test participant is certain?	2	
5	Does the moderator help or move on to the next task when the usability problem is clear, or the test participant is hopelessly stuck or goes around in circles?	5	
6	If the moderator provides help, is the help limited to the absolute minimum?	5	
7	Does the moderator avoid leading questions?	2	
8	Does the moderator encourage test participants to think aloud without overdoing it?	2	

CPUX-UT – Usability Test Checklist

9	<p>Are test tasks communicated properly to the test participant without any risk for misunderstanding?</p> <p>Note 1: Encourage the test participant to paraphrase the task.</p> <p>Note 2: The moderator should not read the task aloud to the test participant.</p>	2	
10	<p>Is the moderator sufficiently curious during moderation? Does the moderator follow up on any important unclarities?</p>	3	
11	<p>Does the moderator appear interested in what the test participant is saying and doing?</p> <p>Note 1: The moderator should have regular eye-contact with the test participant</p> <p>Note 2: The moderator should take notes</p> <p>Note 3: The moderator should not engage in side activities during the usability test session. such as checking their smartphone.</p>	5	
12	<p>Is the moderator unbiased?</p> <p>The moderator must not defend the interactive system. The moderator must not express their own views on the interactive system. The moderator must not have a secret agenda such as a pet peeve or theory about the interface being tested and try to get the test participant to articulate it.</p>	2	

CPUX-UT – Usability Test Checklist

13	<p>Is the moderator respectful towards the test participant?</p> <p>Note 1: The moderator should pay attention to test participants' suggestions even if they appear unreasonable.</p> <p>Note 2: The moderator should not take over the keyboard or the mouse without explicit permission from the test participant. The usability test session should be designed so no intervention from the moderator is required, for example to enter passwords, etc.</p>	5	
14	<p>Is the moderator respectful towards the manufacturer or provider of the test object?</p> <p>Note: The moderator should not badmouth the interactive system or the development team, even if it is done to calm the test participant.</p>	2	
15	<p>If the test participant can't solve a task and wants to see how it can be done, does the moderator diplomatically refuse to reveal the intended way of solving the task?</p> <p>Note: It's OK to show the solution as part of the debriefing, but only if the test participant insists.</p>	5	
16	<p>Is the moderator sufficiently familiar with the product that is being tested?</p>	10	

9 Debriefing

ID	Evaluation criterion	E1	E2
1	<p>Is the debriefing efficient?</p> <p>Note: The debriefing should take less than 3 minutes. Using more than 3 minutes for the debriefing requires a justification.</p>	9	
2	<p>Does the debriefing mainly focus on experience that is not evident from the test session?</p>	2	
3	<p>Does the debriefing ask what the test participant considers the most important usability problems?</p>	2	
4	<p>Does the debriefing ask what the test participant considers the most important positive impressions?</p> <p>Note: It's acceptable if one question addresses both positive and negative aspects.</p>	2	

10 Usability test report – Layout and contents

ID	Evaluation criterion	E1	E2
1	Does the usability test report contain an executive summary?	10	
2	Is the usability test report free from unnecessary information? Note: The usability test report must be as short as possible. Up to 25 pages are reasonable. Examples of unnecessary information are biographies of the usability specialists involved in the test, transcripts and explanations of what usability is.	10	
3	Does the usability test report contain a version date?	1	
4	Is the usability test report free from spelling mistakes and grammatical mistakes?	5	

11 Usability test report – Executive summary

ID	Evaluation criterion	E1	E2
1	Does the executive summary appear at the very beginning of the usability test report?	5	
2	Is the executive summary contained on one page only?	2	
3	Does the executive summary contain a brief description of the most important positive findings?	2	
4	Are the positive findings that are listed in the executive summary also listed in the Findings section of the usability test report?	2	
5	Does the executive summary contain a brief description of the most important usability problems?	2	
6	Does the executive summary prioritize problems correctly, for example by listing problems that are rated “major” or “minor” only if all problems that are rated “critical” are listed?	2	
7	Are the usability problems that are listed in the executive summary also listed in the Findings section of the usability test report? Note: It is commendable to combine several problems into an overall theme in the executive summary	2	

12 Usability test report - Findings

ID	Evaluation criterion	E1	E2
1	<p>Are there more than 10 substantial findings?</p> <p>Note 1: Examples of unsubstantial findings are: “Test participants liked the website”, “Test participants found no spelling errors”</p> <p>Note 2: “Good ideas” usually do not count as findings, in particular if they are essentially recommendations related to another finding.</p>	20	
2	Are all findings easy to understand?	12	
3	<p>Are all findings unique?</p> <p>Note: Two separately reported findings must not describe what is essentially the same usability finding.</p>	2	
4	<p>Are all findings actionable?</p> <p>Note 1: Generally phrased findings such as "The website does not speak the language of the user" should be accompanied by at least two valid examples.</p> <p>Note 2: Findings must be analysed in sufficient depth to derive a remediation plan. Example: It is insufficient to report that a test participant got lost in navigation. Instead, the report should detail where exactly and why the navigation problem happened.</p>	8	
5	Are all findings rated?	5	

CPUX-UT – Usability Test Checklist

6	Are all ratings justifiable? Note: The rating “catastrophic” of a finding in a practical CPUX-UT test is most likely incorrect.	5	
7	Is the meaning of each rating explained in a useful and usable way?	2	
8	Do at least 3 of the findings contain useful and usable recommendations? Note: Recommendations are particularly important for findings that are rated “critical”.	3	
9	Are at least 25% of the findings positive?	1	
10	Are positive findings completely positive? Note: A positive finding should not contain any criticism	1	
11	Are positive findings substantial (as opposed to flimsy or ironic)?	2	
12	Are findings that are not obvious from the textual description illustrated with appropriate screenshots?	2	
13	Are findings supplemented with quotes from test participants, where appropriate? Note: Three reasonable quotes are sufficient.	2	

CPUX-UT – Usability Test Checklist

14	<p>Are findings in accordance with test participants' actual behaviour as recorded on the videos?</p> <p>Note 1: Findings must not be based on inspection or an observation by the moderator of a potential usability problem or a bug that did not cause any problems for test participants</p> <p>Note 2: Findings must not be based on the moderator's opinions.</p> <p>Note 3: Findings must not be based on the test participant's opinions</p> <p>Note 4: Findings must only be based on test sessions that are submitted for evaluation</p>	12	
15	<p>Unreported findings: All major or critical usability problems that are evident from one or more submitted videos must be reported</p>	12	
16	<p>Does each finding reflect a single usability issue?</p> <p>Note: Findings must not be conglomerates of several independent but possibly related issues.</p>	2	

13 Problems that are not covered by the criteria

In this section, the examiner can address problems that are not foreseen in the criteria, but which the examiner still considers relevant.

Problems listed in this section represent professional, ethical or economic risks for the candidate, the test participant, the client of the usability test or the operator of the interactive system.

In other words, this section is about the small unbelievabilities of everyday testing

Examples of problems that are not foreseen in the criteria:

- A large number of careless errors in the report;
- Use of a giant monitor, resulting in very large video files;
- The moderator leaves the test room to get a drink without locking the computer.

ID	Problem that is not covered by the criteria	E2
1		
2		
3		
4		

14 Commendable aspects in the submission

At their discretion, evaluators may list commendable aspects in the submission in this section.

Examples of commendable aspects:

- A table that shows how each test participant fared with each task
- A list of significant results from the interview of the test participant
- Well-considered and well-described recommendations for improvement
- A candid description of how the moderator improved their moderation practice or the tasks from test participant 1 to 2, or from test participant 2 to 3.

ID	Commendable aspect
1	
2	

Appendix. Important changes compared to previous versions

Date, version	Change
30-06-2015, Version 1.01	First version.
20-01-2017, Version 1.06	<p>2.i: OK to submit blank NDA for data protection reasons</p> <p>5.b: Number of tasks in test script must match assignment exactly</p> <p>5.g: Increased from 1 to 4 points. 5.h: Increased from 2 to 4 points.</p> <p>8.g: Avoid “confirmation questions”</p>
31-08-2018, Version 1.08	<p>Many items have been updated to reflect the examiners’ experiences from evaluating a substantial number of practical examinations.</p> <p>A list of the eight most important reasons for failing the practical examination has been added in section 1.2.</p> <p>A list of formal requirements that are checked before the detailed evaluation starts has replaced the previous list of “Initial check of deliverables”. The list, which appears in section 2, now matches a similar list in the CPUX-UT Examination Regulations.</p> <p>The possibility of awarding positive points for commendable aspects of the submission has been added in section 13.</p> <p>All items in the checklist have been renumbered.</p> <p>New items are: 3.7, 4.1 notes, 4.3, 4.4, 5.1 to 5.6, 5.13, 5.14, 5.19 note, 6.4, 8.10 notes, 8.12 notes, 9.5 notes, 11.6 notes, 12.1, 12.6 note 2, 12.16 note 2+3</p>
01-11-2020, Version 1.15	<p>Many notes in the items have been updated to reflect the examiners’ experiences from evaluating a substantial number of practical examinations.</p> <p>Section 1.1: The list of the most important reasons for failing the practical examination has been updated.</p> <p>The possibility of deducting points for serious problems that are not covered by the criteria has been added.</p> <p>Examples of positive comments have been added. The possibility of awarding positive points for commendable aspects of the submission has been removed.</p> <p>New items are: 5.8 – 5.11, 11.6</p>