School of Computer Science

Time Constrained Assessment - Practice

Module Code	**CMP2805M**
Module Title	**User Experience Design**
Module Coordinator	**Dr. Khaled Bachour**
Duration of Assessment	**1 hour 55 minutes**
Date	**24th April 2020**
Release Time	**-**
Submission Time	**-**
Total Wordlimit	**None**

General Instructions to Candidates.
1. You **must** submit your answers as a MS Word Document to Turnitin on Blackboard **before** the submission time: failure to do so will be classified as misconduct in examinations. We strongly recommend you submit 15 minutes prior to the deadline.
2. You **must** also send a copy of your work to the socssubmssions@lincoln.ac.uk at the same time. You must place the Module Code and your Student Id in the Subject Field of the Mail.
3. Hand-written notes or diagrams, **must** be photographed and inserted into the Word Document as an image.
4. This assessment is an open resource format: you may use online resources, lecture and seminar notes, text books and journals.
5. All work will be subject to plagiarism and academic integrity checks. In submitting your assessment you are claiming that it is your own original work; if standard checks suggest otherwise, Academic Misconduct Regulations will be applied.
6. The duration of the Time Constrained Assessment will vary for those students with Learning Support Plans. Extensions do not apply, but Extenuating Circumstances can be applied for in the normal way.

Module Specific Instructions to Candidates:
1. You are free to use a calculator as needed.
2. Show your working where possible.
3. For calculations, you may do this on paper and include a photograph of your workings.
4. You may also use freehand notations for formulas e.g.:
   a. **Sqrt(x)** for the square root of x.
   b. **Sum(x)** for the sum of all values of x.

Answer all of the following questions:

Note: the answers below are complete and would be awarded full marks if given. They are not summaries. In many instances, other answers are also correct. These are just example answers.

1. Usability. [20 marks]

Feedback and *Affordances* are examples of Norman's Design Principles for good usability. For each of those give an example of how a system can fail to provide good usability by not observing these principles, and an example of how a system can achieve good usability by observing them. [5 marks for each correct example].

Answer here:

Feedback:

Bad example: When uploading a submission on Blackboard, after you click submit, the website does not provide any feedback until the upload is complete making it difficult to know if anything is happening.

Good example: When a you push a button on an elevator to indicate the floor you wish to go to, the button lights up to let you know it has been activated.

Affordances:

Bad example: A handle on a door that should be pushed. A handle gives the affordance of pulling, but the door needs to be pushed, so it might be confusing.

Good example: The icon for Microsoft outlook has an envelope on it letting the user intuitively know that this is a mail application.

2. Accessibility. [20 marks]

(a) [10 marks] Colour-Vision Deficiency (CVD) is an impairment that prevents some users from seeing colours correctly. Give two examples of **different** ways in which as system can be designed to support users with CVD [5 marks for each example].

Answer here:

1. Using shapes in addition to colours for example on pedestrian traffic lights, the shape of the green man is different from the shape of the red man, so people with red-green blindness can see what symbol is on.

2. Avoiding the use of problematic colours: in video games, instead of using red for enemy and green for ally, the developers can use orange and blue which are more easy to distinguish

(b) [10 marks] Consider the following technologies, and for each determine if it is an example of Assistive Technology or Accessible Design. Explain your answer. (10 marks).

i.   A car seat with adjustable height.

Answer here:

This is accessible design, the car is designed to be used by people of different heights without the need for additional equipment

ii.   Netflix shows with subtitles included.

Answer here:

This is accessible design, as the show is set up with subtitles in it so people with hearing or cultural impairments can watch it without the need to for additional support.

3. Research Ethics. [20 marks]

Consider the following methods for determining the average waiting time at a local bus stop as part of a university research project. For each, determine if a study using this method should be subject to Ethical Research Principles and explain in one or two sentences why or why not.

a.   Observe people on a bus stop. Mentally count the seconds to estimate the time a person has waited at the stop. Record that time. Repeat as needed choosing a different person at random. Then compute the average of all recorded times. [5 marks]

Answer here:

This study would not be subject to principles of research ethics because the observed humans are not participating in an active capacity, nor is any personal data being collected.

b.   Observe people on a bus stop while holding a stopwatch. Measure the time a person waits at the stop. Record that time. Repeat as needed choosing a different person at random. Then compute the average of all recorded times. [5 marks]

Answer here:

This study would not be subject to principles of research ethics because the observed humans are not participating in an active capacity, nor is any personal data being collected.

c.   Film a bus stop for several hours. Then use the footage to measure the time each person has had to wait. Then compute the average of the measured times. [5 marks]

Answer here:

Research that involves the collection of personal data is always subject to principles of research ethics. This study would therefore be subject to those principles because the video recording would constitute the collection of personal data.

d.   Put up a sign at the bus stop that asks people to measure their own waiting time and then send a text message with their value in seconds to a specified number. Collate responses then compute the average. [5 marks]
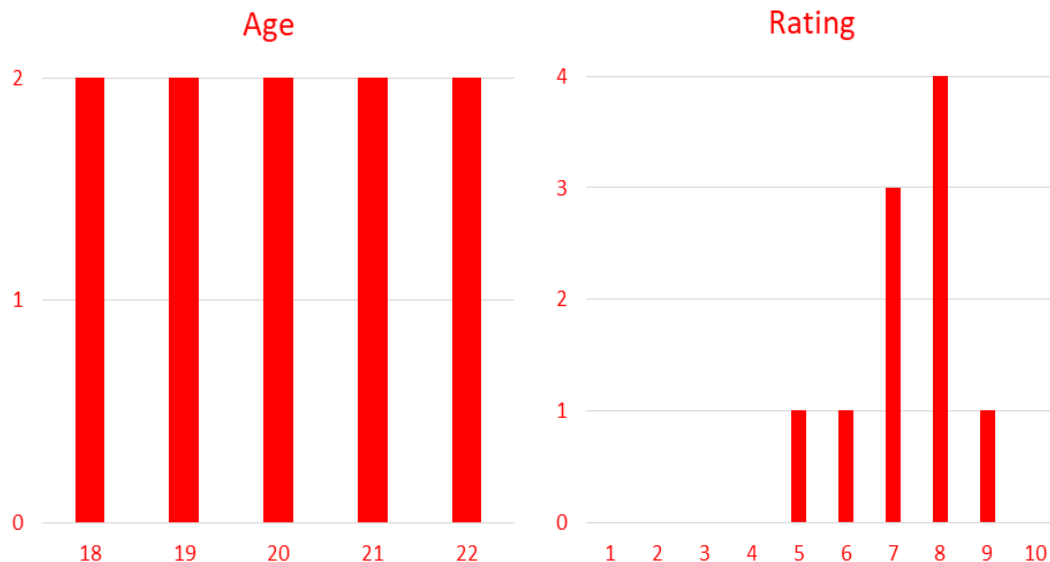
Answer here:

Research that requires the involvement of human participants in an active capacity is subject to research ethics principles. Asking participants to send you a text message makes them active participants, and as such subject to these principles.

4. Data Analysis. [40 marks]

(a) [20 marks] A study is carried out gathering users' opinions of a new website. Ten people were asked to give an overall rating of between 1 (lowest) and 10 (highest) after a brief interaction. In addition they each completed a System Usability Scale questionnaire (SUS). The results are as follows:

Person ID	1	2	3	4	5	6	7	8	9	10
Age	22	19	18	20	18	19	21	21	22	20
Rating	8	8	9	7	7	8	7	8	6	5
SUS	85	70	70	76	72	75	80	74	78	70

1. Below are the distribution graph for the Age and Rating. Name the distributions. [4 marks]

Answer here:

Age: Uniform distribution.
Rating: Skewed distribution.

2. Given that the arithmetic mean Rating is 7.3 and of SUS is 75, calculate the standard deviation of both Rating and SUS scores. Show your working. [6 marks]

Answer here:

Rating:
SD = Sqrt (Sum((x-M)^2)/(n-1)
   = Sqrt ((8 – 7.3)^2 + (8-7.3)^2 + (9-7.3))^2 + (7-7.3)^2 + (7-7.3)^2 + (8-7.3)^2 + (7-7.3)^2 + (8-7.3)^2 + (6-7.3)^2 + (5-7.3)^2)/9
   = Sqrt (0.7^2 + 0.7^2 + 1.7^2 + 0.3^2 + 0.3^2 + 0.7^2+ 0.3^2 + 0.7^2 + 1.3^2 + 2.3^2)/9
   = Sqrt (12.1/9)
   = Sqrt (1.344)
   = 1.159

SUS:
SD = Sqrt ((Sum((x-M)^2)/(n-1))
   = Sqrt (((85-75)^2 + (70-75)^2 + (70-75)^2 + (76-75)^2 + (72-75)^2 + (75-75)^2 + (80-75)^2 + (74-75)^2 + (78-75)^2 + (70-75)^2)/9)
   = Sqrt ((100 + 25 + 25 + 1 + 9 + 0 + 25 + 1 + 9 + 25)/9)
   = Sqrt (24.44)
   = 4.94

3. Given the distribution of Ratings and SUS scores in the results, what would you conclude about the new website? What would you need for a more definitive conclusion? [3 marks]

Answer here:
Even though skewed, the ratings are generally positive, indicating that the users generally liked the website
SUS scores indicate that the website has above average usability (> 68)
More definitive conclusions would require more data (more subjects)

4. There is a correlation of r = +0.822 between the Age and SUS score. What could this mean? What can you do to understand that further? [4 marks]

Answer here:

It means the higher the age is, the higher the SUS score is. This could mean that one causes the other but we cannot know for sure. We would need further investigation possibly with interviews to understand the nature of the relationship between the two and if there really is one.

5. Suppose this website is for a care home for the elderly. What problem can you point out with data that was collected that would jeopardize our ability to evaluate its usability from that data? [3 marks]

Answer here:

Our sample of participants are aged between 18 and 22, and are likely not the target audience for this website. It's possible that while they found its usability good, this does not represent what our target population might find.

(b) [20 marks] A new study is carried out to compare completing a specific task on the previous question's website ('new website') - with completing the same task on a competitor's website ('competitor's website'). One hundred people are asked to evaluate the new website, and one hundred people are asked to evaluate the competitor's website. All participants fill in the SUS questionnaire, with the following results:
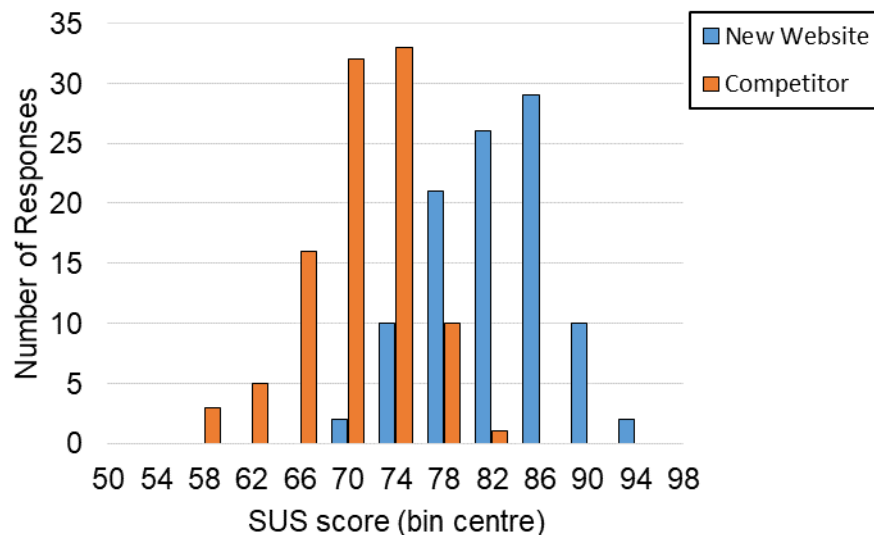
New website Mean: 80.73 (Standard Deviation.: 5.04)
Competitor website Mean: 69.3 (Standard Deviation: 4.81)

1. We want to know whether the new website is better than the competitor's website, and so want to perform a **significance test.** Given the data in the graph above, **what test** would be appropriate for this, and **why**? [4 marks]

   Answer here:

   Data look approximately normal, and can be considered continuous data, so assume that a parametric t-test would be appropriate.

2. What is the **"null hypothesis"** for this significance test? [2 marks]

   Answer here:

   That the two sets are drawn from the same distribution, i.e. that there is no difference between the two websites.

3. Using this test, we obtain a **p-value of 0.001**. What does this mean in terms of significance (i.e. is it a significant result)? [2 marks]

Answer here:

This is below the standard threshold of 0.05. Therefore this can be considered as a significant result.

4. Mention and discuss **two** limitations of only using p-values to interpret the results? [6 marks]

Answer here:

Does not give indication of effect size: the p-value tells us that there is a difference between the two data sets, but it doesn't tell us how big the difference is.
P-values are variable under replication: collecting a different set of data from the a similar experiment would likely result in a different p-value.

5. Calculate the **effect size (Cohen's d)** between the two sets of data. [4 marks]

Answer here:

d = (M1 – M2) / Spooled SD
  = (M1-M2) / Sqrt ((SD1^2*(n1-1) + SD2^2*(n2-1)/(n1+n2-2))
  = (80.73 – 69.3) / Sqrt ((5.04^2*99 + 4.81^2*99)/198)
  = 11.43 / Sqrt ((2514 + 2290)/198)
  = 11.43 / Sqrt (24.3)
   = 11.43/4.93
   = 2.32

6. The **95% Confidence Interval** of the difference of means for this data is **(10.06, 12.9).** What does this mean? [2 marks]

Answer here:

This means we can be 95% certain that the actual difference in means between the two data sets lies between 10.06 and 12.9.