

CPSC 444: User Interface Design
2006-07/Term 2: Midterm Exam Review Questions (Answer Key)

Question 1: True False

For each question, circle one of either true or false. You do not have to provide a justification for the answer you have given.

- (a) Construct validity refers to the causal relationship between independent and dependent variables. **True or false?** *False – construct validity refers to whether you are measuring the correct entity.*
- (b) If two distributions are normally distributed and have the same standard deviation, they will have the same variance. **True or false?** *True – SD is just the square root of the variance.*
- (c) Sending designers to interview workers in an office setting is an example of contextual inquiry. **True or false?** *True – the central feature of CI is to intensely interview people while they work.*

Question 2: Empirical Laws

You are working on improving the usability of a horizontal toolbar in a word processing application. The toolbar consists of a single row of twelve 16x16-pixel icons. Using Fitts' Law, answer the following question:

- (a) **In no more than four sentences**, list two ways to make the toolbar faster and easier to use and use Fitts' Law to explain why these will make the toolbar more usable.

By making the individual icons **bigger** you make the toolbar easier to use because Fitts' Law predicts that larger items take less time to hit. By placing the toolbar **closer to where users' pointing cursors are most likely to be** you also make the toolbar easier to use because Fitts' Law predicts that shorter movements take less time to complete.

- (b) From Fitts' Law, you learned that circular popup menus are more efficient when compared to linear popup menus. Using Fitts' Law, **explain in no more than two sentences** what you can do to linear popup menus to better balance the access times for all items.

One thing that can be done to better balance the access times for all items in linear pop-up menus is to make items that are farther away from the starting point **larger** than those that are closer. By doing so, you can at least partially negate the movement penalty associated with larger item amplitude (i.e. distance).

Question 3: Models of the User

- (a) What role does attention play in human memory?

Attention is what is required for information to move from our sensory stores (image, iconic & haptic) into working memory.

- (b) What are the two main differences between working memory and long-term memory?

Working memory has a small capacity and rapid access and decay, whereas long term memory is very large (if not unlimited), but has slower access and little or no decay.

- (c) Describe the interference model of forgetting?

One item in long-term memory reduces a person's ability of retrieving a second item. For example, changing your telephone number makes it hard to remember your old number [retroactive interference].

- (d) Why are large areas of highly saturated colours hard on users?

Because they cause the user to refocus which is tiring.

Question 4: Experimental Design and Statistical Analysis

Buggy Inc. is a company that specializes in the development of visual debugging interfaces for large, complex software engineering projects. They have recently developed a software package that uses an innovative scheme to highlight syntax errors in code. Because their package has the exact same functionality as a software package sold by their competitors, they want to run a controlled user study to see which of the two software packages is more usable (hoping they can say their package is more usable than their competitor's). They are most interested in comparing how rapidly users can find syntactic errors in their code and how quickly their interface allows them to make corrections.

- (a) What are the independent and dependent variables for this user study? In terms of a statistical test, what is the relationship between the independent and dependent variables?

The most obvious independent variable is software package (the company wants to compare performance of the two designs). Dependent variables will be time-based: time to find syntactic code errors, and time required to make a correction once an error is found.

A statistical test measures whether the independent variable (the 'input') has a significant influence on the value of the dependent variable (the 'output').

(b) What kind of statistical test might be most appropriate to compare how quickly users can find syntactic errors in their code? What does the statistical test tell you? Specify one assumption that must be satisfied before you can use the test.

This experiment's goal is to compare the performance obtainable for two designs, given a single independent variable. The simplest statistical test (and the simplest is usually the best!) in this case is a t-test (2-tailed since we have no experimental reason to expect a difference in one direction). The t-test assumes a normal distribution of the measured data.

(c) When designing a user study such as this one, what are two possible threats to validity that need to be considered? **Explain** how they threaten validity.

1. ecological (face) validity: are the results realistic, by virtue of the experimental task being representative of what a user would actually do?
2. statistical validity: could the results found been obtained by chance?
3. construct validity: do the performance metrics employed actually have bearing on product usability?
(other answers possible).

(d) What is a nuisance variable? Give one example of a nuisance variable in this kind of situation and explain how you might accommodate for it.

Nuisance variables are parameters which although the experimenter does not care about their effect, are likely to affect experiment outcome and are difficult to hold constant and thus eliminate as an experiment factor. In this study, *user* will be an important nuisance variable: the experimenters are not concerned with individual differences yet individual performance will probably vary. Other nuisance variables in a coding test like this might be time of day, subject energy level, or the difficulty / nature of the bugs used in the experiment task.

Question 5: Controlled User Studies

Design an experiment to test **whether the horizontal application menu bars found in Microsoft Windows are faster to access than the horizontal application menu bars found on the Mac OS** by defining each of the following experimental components. These are the menu bars that have items like "File, Edit, and View" in every application. Note that there are many possible experiments: we are just looking for reasonable answers!

Subject Pool:

Since expert vs novice performance is not specified, assume a range is needed. We should therefore recruit n Mac and n Windows users, representing the range of expertise that would be found in the product's target population; where n should be at least 10.

Hypotheses:

Null: *There is no difference in access time for Windows and Mac horizontal application menu bars.*

Alternative: *There is a significant difference in access time for Windows and Mac horizontal application menu bars.*

Independent Variable: *Operating system (Mac or Windows).*

Dependent Variable: *Access time, measured from the time a target is announced to the subject, until the target is acquired.*

Experiment Task (Be concise but specific – include details like instructions to subjects, time limits, etc.):

[many possible variants on this answer]

Users are seated in front of either a Mac or Windows computer, with a blank screen. For a given experiment trial, the target is delivered verbally (spoken by a text-to-speech synthesizer), and then a short time later, the Mac/Windows OS appears with an application open in a given state (a different state in each trial for that application and user). The subject is instructed to find and click on the target menu item as soon as he hears its name spoken, and to emphasize both speed and accuracy. If the target has not been acquired after 60 seconds has elapsed, the trial is ended and recorded as 60 seconds.

Experimental Conditions (including randomization):

To increase consistency of conditions between OS, the experiment uses one set of three applications (Microsoft Windows, Excel, Powerpoint) which have relatively consistent interfaces between Mac and Windows versions and which users of each are likely to have used. The n subjects for each OS do m trials on each application, where a good value for m might be 5-10, depending on overall length of experiment, any learning effect identified in pilot sessions, etc.

An experiment session begins with the subject being instructed in the experiment task.

Then the subject is presented with $m \times 3$ trials, where 3 is the number of applications tested. The trials are randomly ordered in terms of application.

Type of Statistical Analysis:

This data can probably be analyzed using a 2-tailed paired t-test, since it has a single independent variable and we expect its data to be normally distributed given the appropriate selection of the subject pool.
