CPSC 444: User Interface Design
2006-07/Term 2: Final Exam Review Questions (with solutions)

Note: these review questions concentrate mostly on material from lectures, but the exam will cover tutorials and the labs (project) as well.

Question 1: Experimental Design and Statistical Analysis

Recall that Hick's Law defines a log-linear decision-time relationship for choosing one item from *n* equally-likely alternatives. Suppose you wanted to verify Hick's Law as part of developing general design guidelines for a menu-based GUI. Design a simple user study task that verifies Hick's Law and explain what kind of statistical analyses would be appropriate for this study. You do **not** have to provide a description of your subject pool and you do not need to provide a complete description of your equipment/apparatus.

A user task that verifies Hick's Law might proceed in the following manner:

- *Have users sit in front of a computer display that is capable of simultaneously displaying 1, 3, 5, or 7 graphical icons of different colours. Ensure that these graphical icons can be randomized so that they are unlikely to be displayed in the same spatial order in successive trials. Ensure the number of graphical icons displayed in a given trial is also randomized so that users cannot predict how many icons will be presented in the next trial.*

- *The icons displayed in a given trial will be numbered (e.g. in a display of five icons, "1"-"5" left to right).*

- *Inform the users they will be completing a series of trials that will require them to select a single graphical icon from a list of one or more icons. To select a graphical icon, they will verbally indicate the number that is displayed underneath the appropriate icon.*

- *At the beginning of each trial, the display will be blank. One of the colours in the graphical icons to be presented will be randomly chosen. Each colour is equally-likely to be chosen. This colour will be announced or indicated to the user.*

- *Once the announcement has been made, the display will change to present the list of 1, 3, 5, or 7 graphical icons for that particular trial. A timer will be started.*

- *The timer will stop once the user has chosen one of the five icons. The amount of time taken to choose an icon will be recorded, as well as whether the user chose the correct icon or not.*

- *These trials will be repeated several times to ensure consistency of user responses.*

- *Once these data have been collected, a linear regression analysis should be used to determine whether the equation given by Hick's Law is appropriate. A reliable log-linear fit to the data can be used to verify Hick's Law. Other "goodness-of-fit" tests could also be used to independently confirm the presence of a linear fit.*

Question 2: 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 hope that their interface will be better for both novices and experts, but suspect it may only be the case for novices. 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 variables are: (1) software package (Buggy's interface and its competitor), and (2) user expertise (novice and expert)

Dependent variables are both related to performance: (1) time to locate syntactic code errors, and (2) 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) Describe the most powerful statistical test to compare how quickly users can find syntactic errors in their code and whether there is an impact of expertise? What are the possible outcomes from the statistical test? 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 for novices and experts. A two-by-two ANOVA is needed here (interface design x expertise), where interface design should be within subjects to increase statistical power and expertise will have to be between subjects.

The test will tell you about 3 possible effects:
1. main effect of interface – one interface is faster overall (regardless of expertise)
2. main effect of expertise – one group of users is faster overall (regardless of interface)
3. interaction effect – that performance on the given interfaces is dependent on the expertise of the subjects (e.g., novices are faster using Buggy, but experts perform equally well on both.

The ANOVA 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, nuisance variables could be the variability in coding the task performance, time of day, subject energy level, or the difficulty / nature of the bugs used in the experiment task.

Question 3: Field methods

The following questions refer to the CHI 2002 paper: "An Evaluation of a Multiple Interface Design Solution for Bloated Software" by McGrenere et al.

(a) What commercial software application was used in the research?

MSWord

(b) Briefly describe the setting where the experimental study took place.

It took place in the field. The prototype software was installed on participant's own work machines, wherever they normally did their word processing.

(c) There were two independent variables in the experimental study. What were they and what were their specific values?

Interface: MSWord Personal (adaptable UI) and MSWord 2000 (adaptive UI)

Type of user: feature-keen and feature-shy

(d) Name two limitations to this study as it was conducted **and** briefly provide alternative study designs that would address each of the limitations.

1. *Main measures discussed in the paper are all self reported so could be impacted by subject bias. To address this, the study could have been run in the lab instead of the field and some of these measures could have been measured.*
2. *The study was only 6 weeks long. To really understand the impact of personalization, more time is probably needed. To address this, the same basic study design could be used, but it should be extended, perhaps to 4 months or more.*

Question 4: Video

(a) In 444 you used video to capture data in a lab experiment of a hi-fi prototype. Video can also be a powerful tool in the early stages of design, as demonstrated by Wendy Mackay. Briefly explain what those early stages are, how video is used in those design stages, and why it is a useful tool.

Preliminary interviews and observation in context – video is used to capture richer data than what could be recorded by pen & paper, audio recorder, and digital camera alone.

Brainstorming – video is used to visually capture a design idea brainstorming session. This is richer than simply recording it in text format and sketches.

Prototyping – video is used to create low-fidelity prototypes that really bring to life the system in action. This is richer than a static paper prototype.

User feedback on prototype – video prototype better enables users to understand how a system will work and so they can provide more detailed feedback.

(b) Name and explain one primary drawback to using video in these early stages of design.

Using video is very time consuming – capturing it takes time and so does editing it.

Question 5: History of HCI

List 5 innovations by Doug Englebart as described and shown in Alan Kay's lecture and explain to what extent HCI has evolved today with respect to those innovations.

the first mouse – we are largely using the exact same mouse design today

notion of ergonomics (comfortable seating and use of keyboard) – exact setup by Englebart is not seen today, but we now have ergonomic keyboards, wrist guards, and a focus on safe seating

chord keyboard – this has not become mainstream. It (or a variant) is used in isolated circumstances, like court stenography, but not by the general population

shared screen control (ability for remote users to access a single desktop) – have quite a few tools today that provide this support, e.g., NetMeeting

video mediated communication – this is not yet mainstream, but we are seeing more and more users using video links with cheap webcams

many more…

Question 6: Cognitive Engineering

One heuristic in Heuristic Evaluation is to "Help users recognize, diagnose, and recover from errors." Which of the two "gulfs" identified in Norman's 7-stage model of human-computer interaction does that heuristic best attempt to address. Explain your answer.

Gulf of evaluation

This heuristic refers to provide users with helpful system feedback (for example, helpful error messages). Helpful feedback is what allows users to evaluate whether or not they performed the correct action with respect to their goal.

Question 7: HCI research and industry

You have been asked to design a technology supported small workroom to support face-to-face collaboration for small to medium size groups (2-6 people). Identify and briefly explain five factors that you need to consider in terms of how to outfit that room with displays.

Size – wall size vs desktop size vs PDA size

Number of displays – perhaps one big one or multiple smaller ones

Angle – tabletop display, wall displays, something in between (like a drafting board)

Superimposition of input space – touch displays vs. indirect input (like a mouse)

Arrangement of users to the display – sitting around display, standing in front of display

Other answers possible as well…