Chapter 4 Questions

1. What is expert review? Describe the benefits and limitations of expert review. Describe the six different types of expert review methods.

Expert reviews provide knowledgeable, objective, comprehensive insight for user interface designers. A risk is getting off track if the expert review is not structured in advance or the expert is not particularly familiar with the task domain.

* **Heuristic evaluation:** The expert reviewers critique an interface to determine conformance with a short list of design heuristics, such as the Eight Golden Rules
* **Guidelines review:** Interface is checked for conformance with organizational or other guidelines document
* **Consistency inspection:** The experts verify consistency across a family of interfaces, checking the terminology, fonts, color schemes, layout, input and output formats, and so on within the interfaces as well as in the documentation and online help.
* **Cognitive walkthrough:** The experts simulate users walking through the interface to carry out typical tasks.
* **Metaphors of human thinking (MOT):** The experts conduct an inspection that focuses on how users think when interacting with an interface. They consider metaphors for five aspects of human thinking: habit, the stream of thought, awareness and associations, the relation between utterances and thought, and knowing.
* **Formal usability inspection:** The experts hold a courtroom-style meeting, with a moderator or judge, to present the interface and to discuss its merits and weaknesses. Design-team members may rebut the evidence about problems in an adversarial format.

1. Summarize the guidelines expert reviewers should follow in order to write effective usability recommendations.

* Communicate each recommendation clearly at the conceptual level.
* Ensure that the recommendation improves the overall usability of the application.
* Be aware of the business or technical constraints.
* Show respect for the product team’s constraints.
* Solve the whole problem, not just a special case.
* Make recommendations specific and clear.
* Avoid vagueness by including specific examples in your recommendations.

1. State the considerations of selecting and handling participants for a usability experiment.

Selecting the appropriate subjects is important. The subjects should be representative. Make sure sampling techniques for a random subject selection are appropriate for your application. Consider demographics, experience, confidence level, motivation, convenience, and subject groups.

Participants should always be treated with respect and should be informed that it is not they who are being tested; rather, it is the software and user interface that are under study. They should be told about what they will be doing and how long they will be expected to stay. Participation should always be voluntary, and informed consent should be obtained. In most places, an Institutional Review Board (IRB) governs any research performed with human subjects

1. Describe and compare types of usability testing.

**Paper mockups and prototyping.** Early usability studies can be conducted using paper mockups of screen displays to assess user reactions to wording, layout, and sequencing. A test administrator plays the role of the computer by flipping the pages while asking a participant user to carry out typical tasks. This informal testing is inexpensive, rapid, and usually productive.

**Discount usability testing.** Quick-and-dirty approach to task analysis, prototype development, and testing. A controversial aspect is the recommendation to use only three to six test participants. Advocates point out that most serious problems are found with only a few participants, enabling prompt revision and repeated testing, while critics hold that a broader subject pool is required to thoroughly test more complex systems. Good for formative evaluation.

**Competitive usability testing.** Compares a new interface to previous versions or to similar products from competitors. Close to a controlled experimental study, must be careful to construct parallel sets of tasks and to counterbalance the order of presentation of the interfaces. Within-subjects designs seem the most powerful, because participants can make comparisons between the competing interfaces—fewer participants are needed, although each is needed for a longer time period.

**Universal usability testing.** Tests interfaces with highly diverse users, hardware, software platforms, and networks. Trials with small and large displays, slow and fast networks, and a range of operating systems or Internet browsers will do much to raise the rate of customer success. Being aware of any perceptual or physical limitations of the users (e.g., vision impairments, hearing difficulties, motor or mobility impairments) and modifying the testing to accommodate these limitations will result in the creation of products that can be used by a wider variety of users.

**Field tests and portable labs.** Puts new interfaces to work in realistic environments or in a more naturalistic environment in the field for a fixed trial period. Often logging software is used to capture error, command, and help frequencies, as well as productivity measures. Portable usability laboratories with recording and logging facilities have been developed to support more thorough field-testing . A different kind of field testing involves supplying users with test versions of new software or consumer products; tens or even thousands of users might receive beta versions and be asked to comment. ]

**Remote usability testing.** Usability tests done online, avoiding the complexity and cost of bringing participants to a lab. Possible to have larger numbers of participants with more diverse backgrounds, and it may add to the realism, since participants do their tests in their own environments and use their own equipment. The downside is that there is less control over user behavior and diminished ability to observe their reactions, although usage logs and phone interviews are useful supplements. Tests can be performed both synchronously (users do tasks at the same time while the evaluator observes) and asynchronously (users perform tasks independently and the evaluator looks at the results later). Some studies have shown remote usability testing to find more problems than traditional usability testing. Synchronous remote usability testing is a valid evaluation technique.

**Can-you-break-this tests.** Users try to find fatal flaws in the system or otherwise destroy it. Game designers pioneered this approach to usability testing by providing energetic teenagers with the challenge of trying to beat new games.

1. Summarize the actions a designer should take to prepare an effective survey.

A survey form should be prepared, reviewed by colleagues, and tested with a small sample of users before a large-scale survey is conducted. Methods of statistical analysis (beyond means and standard deviations) and presentation (histograms, scatterplots, and so on) should also be developed before the final survey is distributed. Directed activities are more successful than unplanned statistics-gathering expeditions. Since biased samples of respondents can produce erroneous results, survey planners need to build in methods to verify that respondents represent the population in terms of age, gender, experience, and so on. It is important to pre-test or pilot-test any survey instrument prior to actual use. Additional attention may be needed when dealing with special populations. For example, questionnaires for children must be in age-appropriate language, questionnaires for international users may need to be translated, larger fonts may be needed for older adults, and special accommodations may need to be made for users with disabilities.

1. Write an example of an acceptance test benchmark task that uses measurable criteria for a user interface.

A good acceptance test benchmark will not use the vague and misleading criterion of “user friendliness.” It will establish measurable criteria for the user interface in these categories:

* Time for users to learn specific functions
* Speed of task performance
* Rate of errors by users
* User retention of commands over time
* Subjective user satisfaction

1. Describe the types of ongoing evaluation that take place after a product is released.

Interviews and focus group discussions, continuous user-performance data logging, customer service (online or telephone consultants, online suggestion boxes, e-mail), user discussion groups, wikis, and newsgroups. Automated software tools are also useful.

1. Describe the kinds of feedback a designer can get when using automated design tools.

* Can check spelling or concordance of terms
* Simple metrics: report numbers of displays, widgets, or links between displays
* More sophisticated: Assess whether a menu tree is too deep or contains redundancies, whether widget labels have been used consistently, whether all buttons have proper transitions associated with them
* Markup Validation Service
* Webpage Download speeds
* Run-time logging software, which captures the users’ patterns of activity
* Log-file-recording tool that captures clicks with associated timestamps and positions on the screen, keeping track of items selected and screen changes, capturing the screen shots, and recording when a user is finished

1. Describe how the scientific method is applied to human computer interface design evaluation.

* Deal with a practical problem and consider the theoretical framework.
* State a lucid and testable hypothesis.
* Identify a small number of independent variables that are to be manipulated.
* Carefully choose the dependent variables that will be measured.
* Judiciously select participants, and carefully or randomly assign participants to groups.
* Control for biasing factors (nonrepresentative sample of participants or selection of tasks, inconsistent testing procedures).
* Apply statistical methods to data analysis.
* Resolve the practical problem, refine the theory, and give advice to future researchers.

**Terms and Concepts to Know:**

1. Expert review
2. Heuristic evaluation
3. Guidelines review
4. Consistency inspection
5. Cognitive walkthrough
6. Metaphors of human thinking (MOT)
7. Formal usability inspection
8. Bird’s Eye View
9. Eye tracking Software
10. Think-aloud technique
11. Types of usability testing
    1. Paper mockups and prototyping
    2. Low-fidelity prototypes
    3. Discount usability testing
    4. Formative and Summative Evaluation
    5. Competitive usability testing
    6. Universal usability testing
    7. Field tests and portable labs
    8. Remote usability testing
    9. Can-you-break-this tests
12. Representative sample
13. Sample size and confidence level
14. Between-subjects versus within-subjects approach
15. Random sample
16. Dependent variable
17. Independent variable
18. Demographics
19. Likert Scale
20. Types of Usability questionnaires: QUIS, SUS, etc.

10. Acceptance test

11. Continuous user-performance data logging

12. Web analytics