Chapter 4:
Evaluating Interface Design

Introduction

Evaluation

- tests usability and functionality of system
- occurs in laboratory, field and/or in collaboration with users
- evaluates both design and implementation

Evaluation should be considered at all stages in the design life cycle.

Goals of Evaluation

- assess extent of system functionality
- assess effect of interface on user
- identify specific problems

Evaluation Plan

- testing is a necessity
- The determinants of the evaluation plan include:
  - stage of design (early, middle, late)
  - novelty of project (well defined vs. exploratory)
  - number of expected users
  - criticality of the interface
  - costs
  - time available
  - experience of the design and evaluation team

- Usability evaluators are open to non-empirical methods, such as user sketches, consideration of design alternatives, and ethnographic studies.
- Recommendations needs to be based on ***observational findings***
- Tools and techniques are evolving
- Evaluation plans range from 3 days to 2 years
- Costs range from 5-20% of a project
- Usability testing has become an established and accepted part of the design process

Expert Reviews

- Informal demos to colleagues or customers can provide some useful feedback
- Formal expert reviews have proven to be more effective

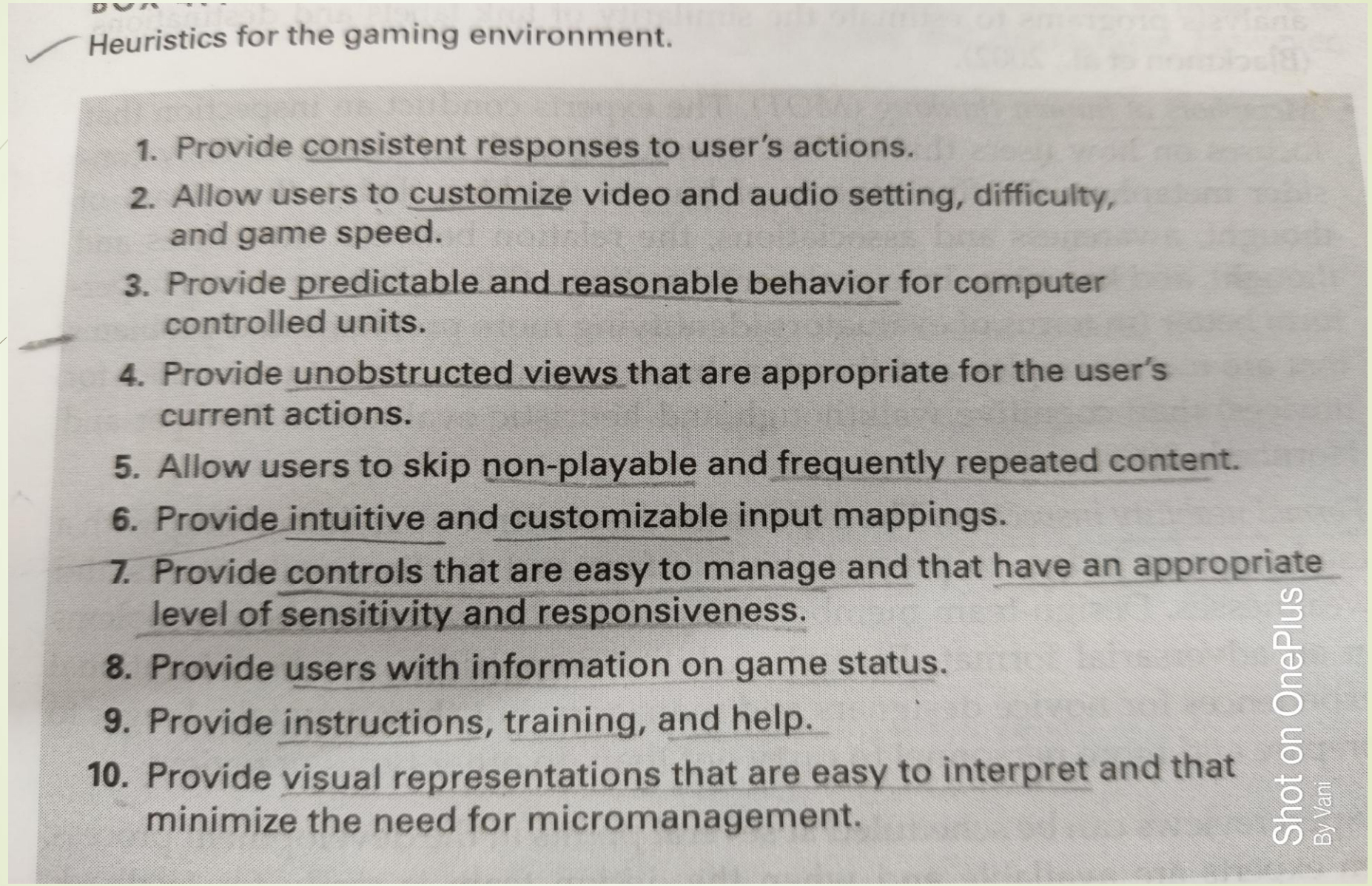
- Expert review methods:
  - Heuristic evaluation
  - Guidelines review
  - Consistency inspection
  - Cognitive walkthrough
  - Metaphors of human thinking
  - Formal usability inspection

"I know nothing about the subject,
but I'm happy to give you my expert opinion."

- **Expert review methods:**
- Heuristic evaluation – 8 golden rules followed or not, check whether heuristics match the application. Eg: 29 playable heuristics check.
  - 3 categories
    - Game usability
    - Mobility heuristics
    - Gameplay heuristics
- Guidelines review - conformation with organisational or other guidelines document , may take long time to observe and review the interface according to the guidelines doc
- Consistency inspection – check for terminology, font, color schemes, layout, formats for i/p and o/p in interfaces, documentation and online help.

- Cognitive walkthrough – experts simulate users walking through the interface to carry out the tasks like high frequency tasks and rare critical tasks(error recovery), may walkthrough privately or with group members
- Metaphors of human thinking – conduct inspection which focuses on how users think when interacting with the interface.
  - 5 aspects of human thinking – habit, stream of thought, awareness and associations, relationship between utterances and thought ,and knowing
- Formal usability inspection – hold courtroom style meeting – to present the interface and to discuss its merits and demerits. Formal usability inspections can be educational experiences for novice designers and managers

- Gaming environment Heuristics



Heuristics for the gaming environment.

1. Provide consistent responses to user's actions.
2. Allow users to customize video and audio setting, difficulty, and game speed.
3. Provide predictable and reasonable behavior for computer controlled units.
4. Provide unobstructed views that are appropriate for the user's current actions.
5. Allow users to skip non-playable and frequently repeated content.
6. Provide intuitive and customizable input mappings.
7. Provide controls that are easy to manage and that have an appropriate level of sensitivity and responsiveness.
8. Provide users with information on game status.
9. Provide instructions, training, and help.
10. Provide visual representations that are easy to interpret and that minimize the need for micromanagement.

Expert Reviews

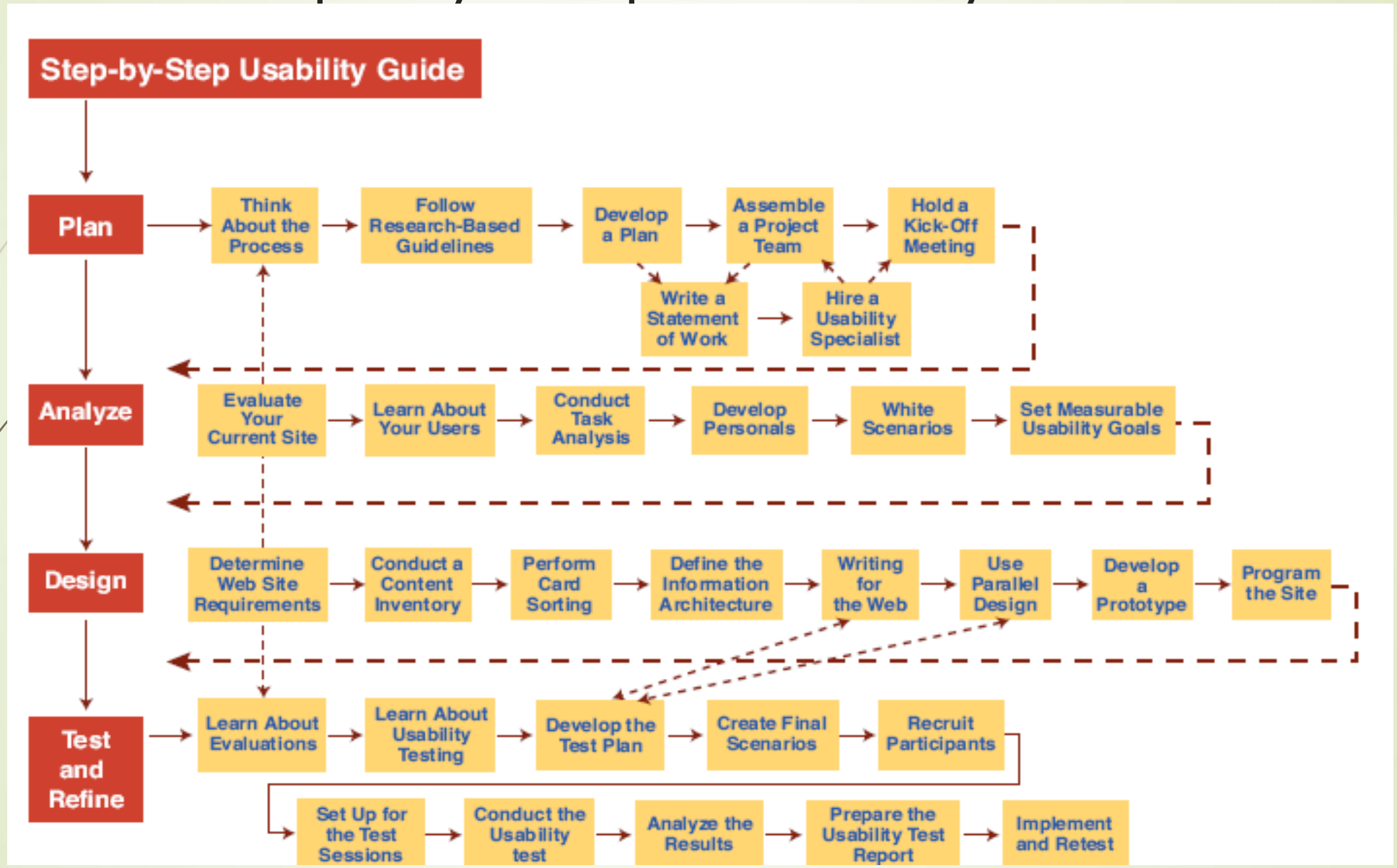
- Expert reviews can be scheduled at several points in the development process
- Different experts tend to find different problems in an interface, so 3-5 expert reviewers recommended
- The dangers with expert reviews are that the experts may not have an adequate understanding of the task domain or user communities.
- Even experienced expert reviewers have great difficulty knowing how typical users, especially first-time users will really behave.

Making usability recommendations useful and usable.

- 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.

From Molich et al., 2007.

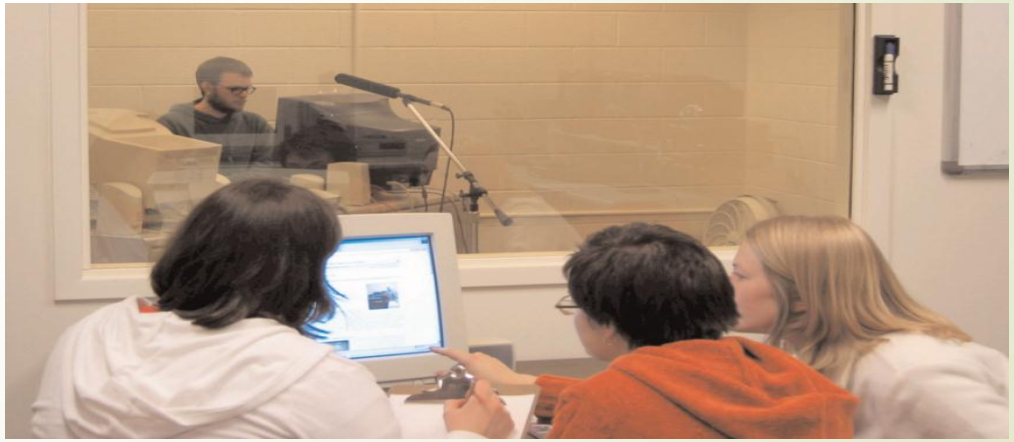
Step-by-Step Usability Guide



Usability Testing and Laboratories


Participant observer in Workstation


eye-tracking setup



Usability Testing and Laboratories

- Emergence of usability testing and laboratories since early 1980s
- Sped up many projects and produced dramatic cost savings.
- A typical modest usability lab would have two 10 by 10 foot areas
  - one for the participants to do their work
  - and another, separated by a half-silvered mirror, for the testers and observers
- Participants should represent the intended user communities, with attention to
  - background in computing, experience with the task, motivation, education, and ability with the natural language used in the interface.
- Participation should be voluntary, and informed consent should be obtained

- Practice is to ask all individuals to read and sign a statement like this:
  - I have freely volunteered to participate in this experiment.
  - I have been informed in advance what my task(s) will be and what procedures will be followed.
  - I have been given the opportunity to ask questions, and have had my questions answered to my satisfaction.
  - I am aware that I have the right to withdraw consent and to discontinue participation at any time, without prejudice to my future treatment.
  - My signature below may be taken as affirmation of all the above statements; it was given prior to my participation in this study.
- Videotaping participants valuable for later review and for showing designers or managers the problems that users encounter.
  - Use caution in order to not interfere with participants
  - Invite users to think aloud as they are performing the task.
- Think aloud brings a Informal atmosphere
- Retrospective think aloud
- Institutional Review Boards (IRB) often governs human subject test process

Think Aloud

- I predict that ...
- I can picture ...
- A question I have is ...
- This reminds me of ...
- This is like ...
- I am confused about ...
- The big idea here is ...
- I believe ...

- **Variant forms of usability testing** – can use a combination of:
  - **Paper mockups and prototyping** – mockup of screen displays to assess user reactions to working, layout and sequencing. Informal testing is Inexpensive, rapid and productive
  - **Discount usability testing** - quick and dirty approach to task analysis, prototype development and testing. Formative evaluation – discount usability testing – identifies problems that guide redesign and summative evaluation – provides evidence for product announcements. 3-6 participants.
  - **Competitive usability testing** – compares new interface to previous versions or to similar products from competitors. Few participants are needed for longer duration.
  - **Universal usability testing –** tests interfaces with diverse users, hardware, s/w platforms , networks. For consumer electronics products, e governance, web based services – tests to clean up problems and help ensure success for wider variety of users.
  - **Field test and portable labs –** puts new interface to work in realistic environment. By including logging s/w to capture errors, command & help frequencies, productivity measures.
  - **Remote usability testing -** conduct tests online- reduce cost bringing participants to lab. More participants , diverse background recruted by email or thr online communities. Synchronously or asynchronously. s/w support –synchronous WebEx. NetMeeting.
  - **Can-you-break-this tests** –teenagers challenged to try to beat new games. Destructive approach.

Survey Instruments

- Written user surveys are
  - informal
  - subjective
  - familiar
  - cheap
  - good complement to usability tests and expert reviews.

- Keys to successful surveys
  - Clear goals in advance
  - Development of focused items that help attain the goals.

Respondents need to represent population in terms of age, gender, experience

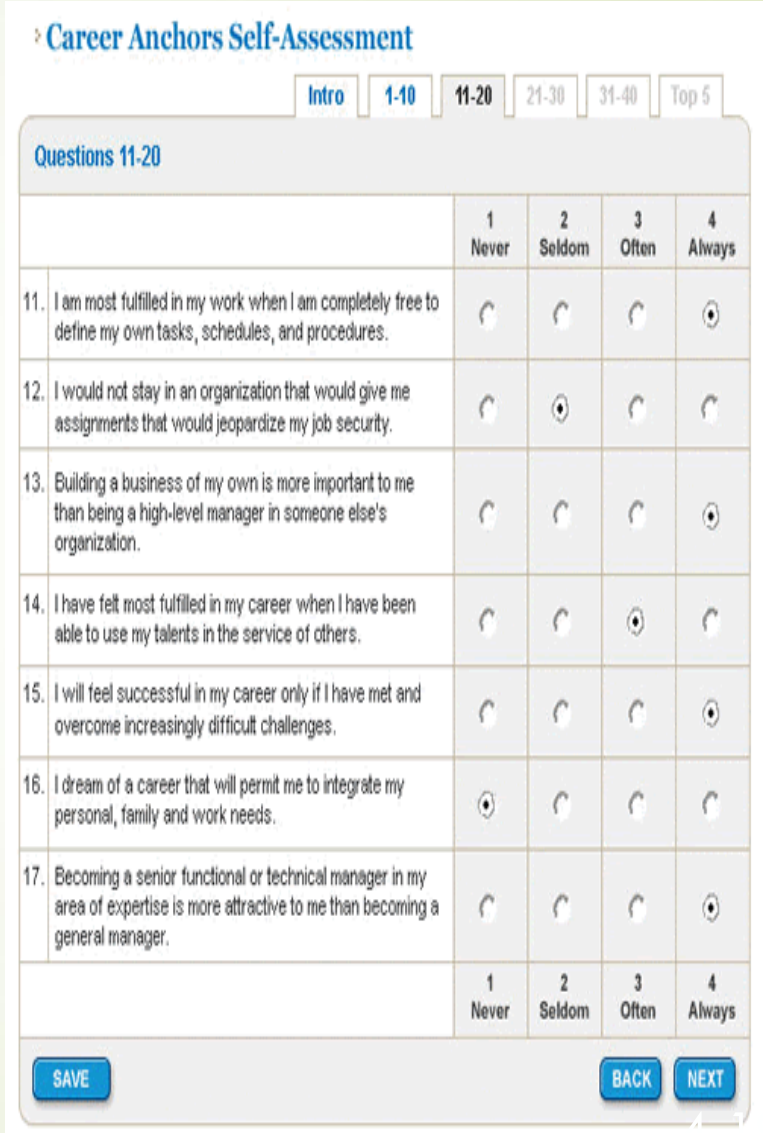
- Users asked for subjective impressions about aspects of the interface such as the representation of:
  - task domain objects and actions
  - Interface domain metaphors and action handles
  - syntax of inputs and design of displays.

Survey Instruments

- Other goals would be to ascertain characteristic about users including
  - users background (age, gender, origins, education, income)
  - experience with computers (specific applications or software packages, length of time, depth of knowledge)
  - job responsibilities (decision-making influence, managerial roles, motivation)
  - personality style (introvert vs. extrovert, risk taking vs. risk aversive, early vs. late adopter, systematic vs. opportunistic)
  - reasons for not using an interface (inadequate services, too complex, too slow)
  - familiarity with features (printing, macros, shortcuts, tutorials)
  - their feeling state after using an interface (confused vs. clear, frustrated vs. in-control, bored vs. excited).

Surveys

- Online surveys avoid cost of printing and the extra effort needed for distribution and collection of paper forms.
- Many prefer to answer a brief survey displayed on a screen, instead of filling in and returning a printed form
- A survey example is the Questionnaire for User Interaction Satisfaction (QUIS).



Styles of question

- general - establish background of user
- open-ended - 'Can you suggest improvements to interface?'
- scalar

Example: *How would you describe your experience navigating our website?*

	Very hard	Somewhat hard	Neither hard nor easy	Somewhat easy	Very easy	
Hard to Navigate	1	2	3	4	5	Easy to Navigate

- multi-choice
  - How do you most often get help with the system? Choose one.
    - online manual
    - contextual help
    - command prompt
    - ask a colleague
- ranked – place a list of items in order

Acceptance Test

- For large projects, the customer or manager usually sets objective and measurable goals for hardware and software performance. If the completed product fails to meet these acceptance criteria, the system must be reworked until success is demonstrated.
- **Central goal of acceptance test is not to find flaws but to verify adherence to requirements**

- Rather than the vague and misleading criterion of "user friendly," measurable criteria for the user interface can be established for the following:
  - Time to learn specific functions
  - Speed of task performance
  - Rate of errors by users
  - Human retention of commands over time
  - Subjective user satisfaction
- In a large system, there may be 8 or 10 such tests to carry out on different components of the interface and with different user communities.

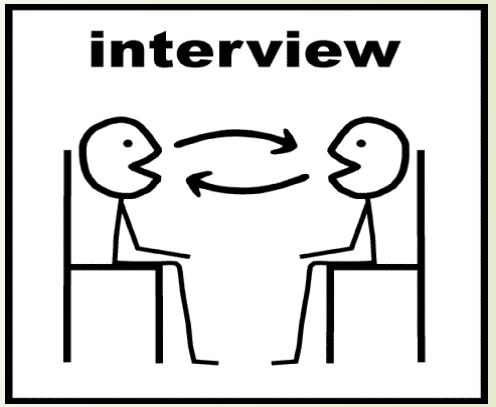
- After successful acceptance testing there may be a period of field testing before national or international distribution

- **Goal of**
  - Early expert review
  - Usability testing
  - Surveys
  - Acceptance testing
  - Field testing

is to force as much as possible of the evolutionary development into the pre release phase when change is relatively easy and inexpensive to accomplish

Evaluation During Active Use

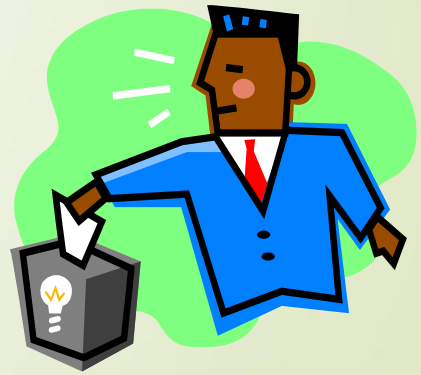
- Successful active use requires constant attention from dedicated managers, user-services personnel, and maintenance staff.
- Perfection is not attainable, but percentage improvements are possible.



- 1. **Interviews and focus group discussions**
  - Interviews productive because the interviewer can pursue specific issues of concern.
  - Group discussions ascertain the universality of comments.

Evaluation During Active Use

- 2. **Continuous user-performance data logging**
  - Collect data about
    - The patterns of system usage
    - Speed of user performance
    - Rate of errors
    - Frequency of request for online assistance
  - A major benefit is guidance to system maintainers in optimizing performance and reducing costs for all participants.
- 3. **Online or telephone consultants, e-mail, and online suggestion boxes**
  - Users feel reassured if human assistance available
  - On some systems, consultants can monitor the user's computer and see the same displays



Evaluation During Active Use

- 4. **Discussion groups, wiki's and newsgroups**
  - Permit postings of open messages and questions
  - Some are independent, e.g. America Online and Yahoo!
  - Topic list
  - Sometimes moderators
  - Social systems
  - Comments and suggestions should be encouraged.
  - Online suggestion box or e-mail trouble reporting