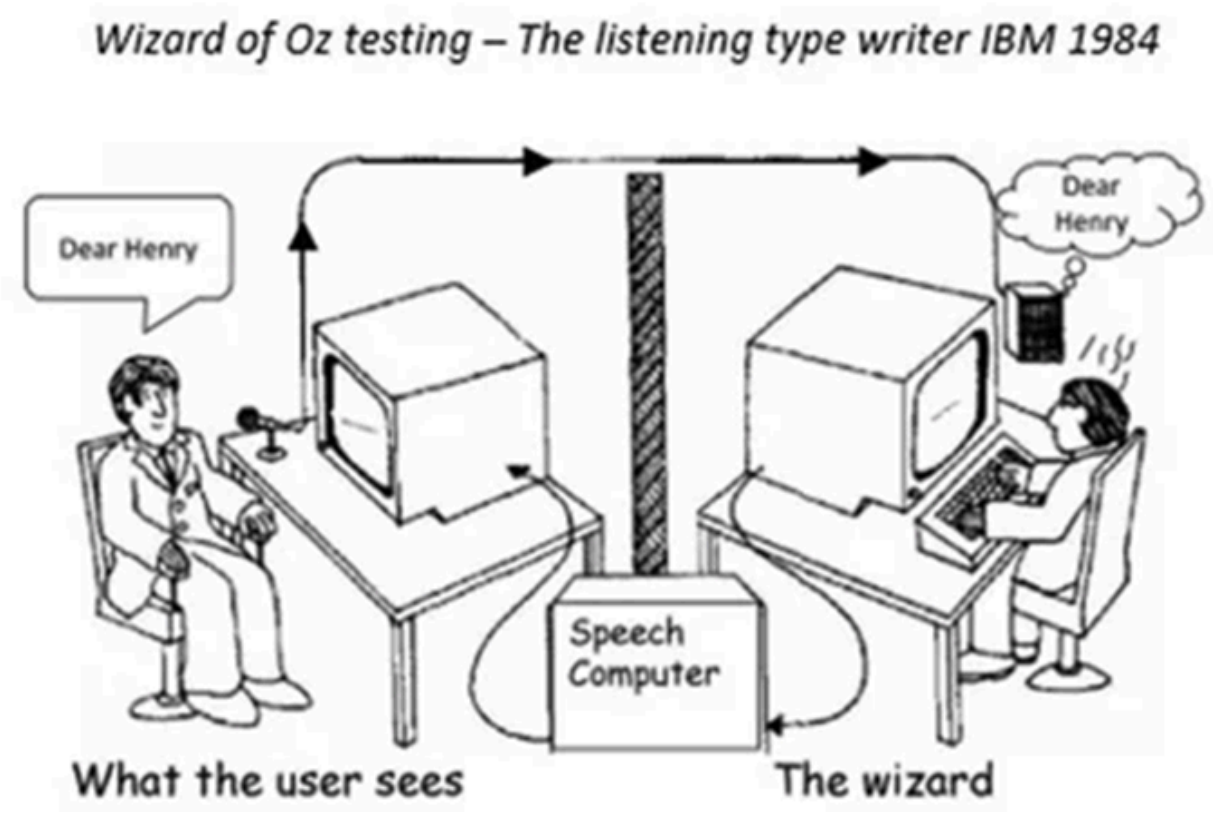
Chapter - 8

What is 'Wizard-of-Oz' Prototyping?

Wizard-of-Oz Prototyping is a way to test a new idea without actually building the technology. A human pretends to be the technology to see how people interact with it. Here's a basic example:

- **Example**: Imagine you're testing a new voice assistant (like Siri). Instead of programming it, a person listens to the user's questions and types back answers. The user thinks they're talking to a real assistant, but it's just a person pretending.

Why Use It: This lets you learn what users want and how they'll use the technology before you spend time and money actually building it.



What is Focus Group Discussion? How to perform this technique to collect data and validate it? Discuss interview techniques to collect data.

Focus Group Discussion (FGD)

- **What it is**: A focus group discussion gathers a small group of users to discuss their thoughts, experiences, and needs about a specific product or system. The discussion is guided by a facilitator who asks questions and steers conversation to get insights.
- **How to Conduct It**:
    - **Gather a Group**: Typically 6-10 users who represent the target audience.
    - **Facilitator's Role**: Ask open-ended questions about specific aspects of the product or interface.
    - **Data Collection**: Record the discussion, noting key feedback and common themes.
    - **Validation**: Look for recurring themes or opinions that indicate strong trends or issues.
- **Example**: Imagine gathering a group of students to discuss their experience using an e-learning platform. Their feedback could show usability issues or suggest new features.

2. Interview Techniques for Data Collection

- **What it is**: Interviews involve speaking one-on-one with users to collect detailed insights about their interactions with a system.
- **How to Perform It**:

    - **Types of Questions**: Simple question-and-answer or structured questions about specific product features.
    - **Setting Up Scenarios**: If the product isn't fully developed, a mock-up or simulated experience can be used to get user input.
    - **Variations**:
        - **Cognitive Walkthrough**: Ask users to "speak aloud" their thought process while interacting with the system to identify gaps.
        - **"Can You Break This?" Testing**: Encourage users to find flaws, bugs, or inconsistencies in the design.
- **Example**: In developing a new app, you might ask users to perform tasks on a mock-up while describing their steps and challenges.

3. Interviewing Techniques to Collect Data

- **Key Points**:
    - **Observing Behavior**: Direct observation during interaction reveals issues users might not express verbally.
    - **Variety in Interviews**: Some users may represent different expertise levels or demographics, helping identify if the design works across a broad audience.
    - **Feedback Types**: Collecting feedback about layout, language, ease of navigation, and aesthetic appeal.
- **Example**: For a hospital management system, a designer could interview doctors, nurses, and administrative staff, each providing insights unique to their roles, which help design a more effective system.

What are the data gathering techniques? What are the pros and cons?

1. Interviews

- Description: Talking one-on-one with users to understand their needs, challenges, and preferences. Can be structured (with set questions) or unstructured.
- Pros:
    - Allows in-depth insights into user experiences.
    - Flexible; can adapt questions based on user responses.
- Cons:
    - Time-consuming.
    - Requires skilled interviewers to gather quality data.
- Example: Interviewing students about their experience using an educational app to identify specific learning challenges.

---

2. Focus Groups

- Description: Conducting a guided discussion with a small group of users to gather opinions, preferences, and attitudes.
- Pros:
    - Encourages dynamic conversation and diverse viewpoints.
    - Faster than one-on-one interviews for broad feedback.
- Cons:
    - May lead to groupthink, where people conform to others' opinions.
    - Dominant voices can overshadow quieter participants.

- Example: Conducting a focus group with parents to discuss features for a children's app.

---

3. Surveys and Questionnaires

- Description: Using structured questions to collect information from a larger group, usually via paper or online forms.
- Pros:
    - Can gather data from many people quickly.
    - Cost-effective and easy to distribute.
- Cons:
    - Limited depth; doesn't capture detailed user emotions.
    - Risk of low response rates.
- Example: Sending a questionnaire to app users about their satisfaction with its interface.

---

4. Observations

- Description: Watching users interact with a product or service to see how they use it in real-time.
- Pros:
    - Provides real insight into user behavior and interaction.
    - Identifies unspoken usability issues.
- Cons:
    - Requires time and presence with users.
    - Users may behave differently if they know they are being watched.

- Example: Observing a user navigate a website to see where they get confused or stuck.

---

5. Prototyping (Wizard of Oz Testing)

- Description: Creating a basic, interactive version of a product to test ideas before building the actual system. A person (the "wizard") simulates system responses.
- Pros:
    - Allows testing of features early in the design process.
    - Users provide feedback on real interactions.
- Cons:
    - Time-consuming to set up.
    - Might not fully simulate the real experience.
- Example: Testing a voice assistant by having a person type responses to user questions, rather than programming the assistant.

By using these techniques thoughtfully, teams can understand user needs, identify usability issues, and make informed decisions to enhance the user experience in future product designs.

Mention pros and cons of observation.

- **Pros and Cons of Observation**:
    - **Pros**:

1. **Natural Behavior Insight**: Observation captures users' natural interactions, revealing unspoken problems and behaviors.
2. **Unfiltered Feedback**: Allows the designer to directly see where users struggle or succeed.

- **Cons**:
  1. **Observer Bias**: The presence of an observer can influence user behavior, leading to unnatural interactions.
  2. **Limited Insight**: Observation may not provide an understanding of why users make certain choices without follow-up questions.

What is Hypothesis? What is the purpose of the Null and Alternate hypothesis?

Hypothesis: There is no difference in user performance (time and error rate) and preference (5 point likert scale) when typing on two sizes (i.e., small, large) of an alphabetic, qwerty and random on-screen keyboard using a touch-based large screen, a mouse-based monitor, or a stylus-based PDA.

What are the independent variables here? What are the levels? What are the dependent variables here? What is the purpose of the control variable?

A hypothesis is an idea or assumption that you can test. It's like a guess that says, "I think this will happen if we do this." For example, if you think two types of keyboards will be equally fast for typing, your hypothesis could be, "There's no difference in typing speed between keyboard types."

Purpose of Null and Alternative Hypotheses

- **Null Hypothesis (H0)**: This hypothesis states there is no effect or difference. In this case, it's saying the different keyboard types won't make a difference in typing speed or accuracy.
- **Alternative Hypothesis (H1)**: This is the opposite of the null hypothesis, suggesting there *is* a difference. Here, it might say that one keyboard type is faster or has fewer errors than the other.

Using both helps researchers see if their results are meaningful. If the data rejects the null hypothesis, it supports the alternative hypothesis.

Independent Variables

Independent variables are the things that are changed in an experiment to see if they affect the outcome. In this case:

- **Independent Variables**: Keyboard type, device type, and screen size.
- **Levels**:
    - For keyboard type: alphabetic, QWERTY, and random layouts.
    - For device type: touch screen, mouse, and stylus.
    - For screen size: small and large.

Dependent Variables

Dependent variables are the outcomes being measured. Here, they include:

- Typing speed (time taken)
- Error rate (how many mistakes are made)
- User preference (on a rating scale)

Purpose of Control Variables

Control variables are kept constant so they don't affect the experiment. They ensure that only the independent variables impact the results, making the test fair.

What is a Dependent and Independent Variable in User Experiments?

Independent Variable

- **Definition**: This is the variable researchers change or manipulate to observe effects.
- **Examples in HCI**: Interface type (e.g., old vs. new design), input device (e.g., keyboard vs. touch screen), or interaction style (e.g., swipe vs. button press).
- **Purpose**: Helps determine how specific changes impact user behavior or performance.

Dependent Variable

- **Definition**: This is the outcome measured in response to changes in the independent variable.
- **Examples in HCI**: Task completion time, accuracy (number of errors), or user satisfaction scores.
- **Purpose**: Measures the effectiveness of a design based on user performance and feedback.

Example Experiment

For instance, in testing a new game controller:

- **Independent Variable**: Input device (keyboard vs. new controller).
- **Dependent Variable**: Time taken to complete a game level, number of errors made.

Importance in User Experiments

- **Comparison**: Allows comparison between a new design (test case) and a standard design (control).
- **Objective Evaluation**: Provides quantitative data (like time, error rates) and qualitative feedback (like user satisfaction).
- **Insight into Usability**: Helps quantify how changes impact usability, guiding improvements.

What are the typical Steps to Carry Out Controlled Experiments in HCI?

1. Define Objective and Hypothesis

- Identify the goal of the experiment and formulate a hypothesis to test.

2. Select Participants

- Choose a sample of users relevant to the target audience (e.g., based on age, expertise).

3. Choose Independent and Dependent Variables

- **Independent Variable**: The factor you'll manipulate (e.g., interface design).
- **Dependent Variable**: The outcome you'll measure (e.g., task completion time).

4. Design Tasks and Scenarios

- Create realistic tasks for participants to complete, representing actual use cases.

5. Control Environment

- Minimize external factors by controlling the test environment (e.g., using a lab or quiet room).

6. Conduct the Experiment

- Guide participants through tasks while collecting data, ensuring consistent conditions.

7. Collect and Analyze Data

- Measure results like completion time, accuracy, or satisfaction using quantitative or qualitative methods.

8. Interpret Results and Draw Conclusions

- Analyze findings to see if they support or refute the hypothesis, using statistical tests if needed.

How to Conduct Study Design Decisions for Participants?

1. **Select Evaluation Method**:
    - Use methods like **interviews**, **cognitive walkthroughs**, or **simulations** to study participant interactions.
2. **Determine Participant Characteristics**:
    - Focus on relevant traits like **age**, **expertise**, or **specific needs** to ensure representative feedback.
3. **Choose Evaluation Focus**:
    - Decide whether the study will examine **interaction modeling**, **usability issues**, or **interface design flaws**.
4. **Set Environment**:
    - Conduct in a **controlled setting** or simulated space to keep results consistent.
5. **Administer Structured Tasks**:
    - Create clear tasks, or use techniques like **Wizard of Oz** (when simulating certain functions) to observe behavior effectively.
6. **Expert Heuristic Evaluation (if relevant)**:
    - Have HCI experts assess the design based on usability heuristics when direct user feedback isn't feasible.
7. **Analyze Results Based on Task Objectives**:

    - Look at task completion, errors, and usability to evaluate if the interface meets user needs.

What is evaluation? What are formative and summative evaluation?

- **Evaluation:** In HCI, evaluation refers to the process of assessing a system or product's usability, effectiveness, and user satisfaction. Evaluation helps identify design flaws and areas for improvement.
- **Formative Evaluation:** Conducted during the design and development phase, formative evaluation aims to improve the product iteratively. It helps designers make adjustments based on early user feedback.
    - **Example**: Testing an early prototype with a small group of users to refine features.
- **Summative Evaluation**: Conducted at the end of the development process, summative evaluation assesses the product's overall effectiveness and usability.
    - **Example**: Conducting usability tests on the final version of the software to ensure it meets user requirements and usability standards.

What are the evaluation criteria?

- **Usability:** Measures the ease of use and learnability of a user interface. It's assessed in two ways:

    - **Quantitative assessment:** Includes task performance measurements such as:
        - **Task completion time**
        - **Task success rate** (score or quantity achieved)
        - **Error rate**
    - **Qualitative assessment:** Gathers user feedback through surveys or questionnaires, like the NASA Task Load Index (TLX) and IBM Usability Questionnaire, focusing on factors like ease of use, effort, and user preference.
- **User Experience (UX):** Goes beyond usability, evaluating the user's emotional response and satisfaction. UX factors include:
    - **User emotions and perceptions**
    - **Context of use** (e.g., cultural, social, physical)
    - UX aims to measure satisfaction, focusing on how the product fits into the user's life.

What is the NASA TLX User Experience questionnaire?

1. **Mental Demand**
   "How mentally demanding was the task?"
   Measures cognitive load — whether the user found the task complex or mentally challenging.
    - Scale: **Very Low** to **Very High**
2. **Physical Demand**
   "How physically demanding was the task?"
   Assesses the physical effort required for task completion, such as repetitive movements or strain.

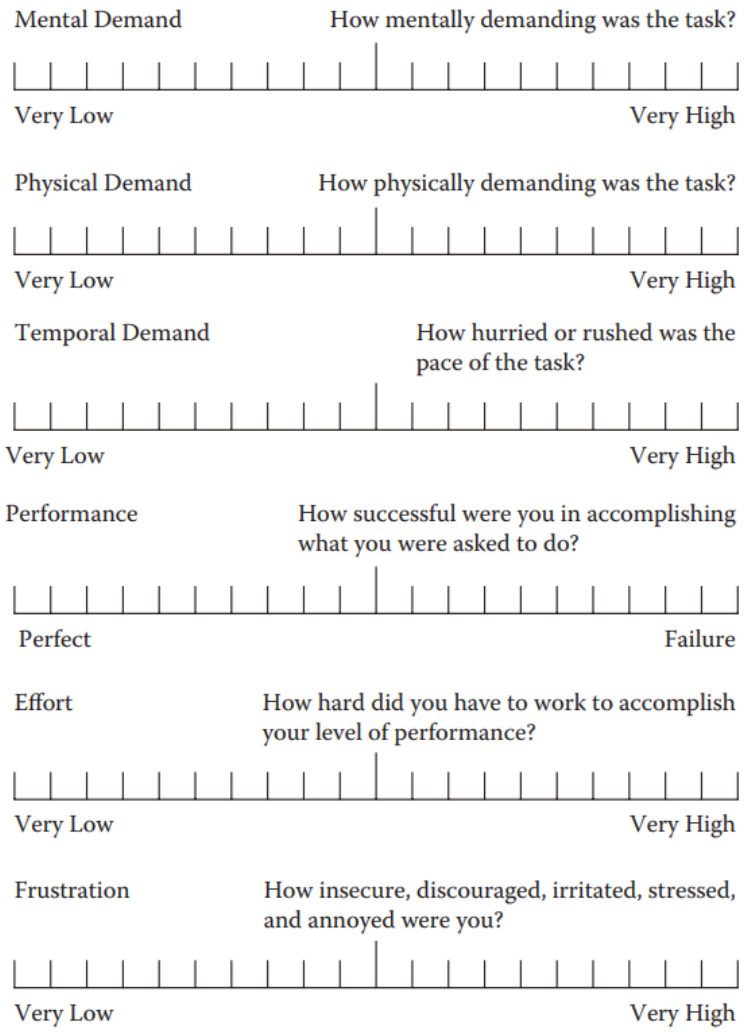
    - Scale: **Very Low** to **Very High**
3. **Temporal Demand**
   "How hurried or rushed was the pace of the task?"
   Look at time pressure, asking if the user felt that they had adequate time or were under a time crunch.
    - Scale: **Very Low** to **Very High**
4. **Performance**
   "How successful were you in accomplishing what you were asked to do?"
   Measures the user's self-assessment of task success and completion.
    - Scale: **Perfect** to **Failure**
5. **Effort**
   "How hard did you have to work to accomplish your level of performance?"
   Captures the perceived effort required, such as the physical and mental energy expended on the task.
    - Scale: **Very Low** to **Very High**
6. **Frustration**
   "How insecure, discouraged, irritated, stressed, and annoyed were you?"
   Assesses negative emotional responses and levels of frustration, often related to UX.
    - Scale: **Very Low** to **Very High**

How It Works

Each of these dimensions is rated on a 7-point scale by the user, with ratings ranging from very low to very high. After gathering

responses, the ratings can be analyzed to understand which aspects of the workload are impacting the user's experience the most.

NASA TLX is especially useful in identifying areas of a system that may need improvement to reduce user workload, making tasks more enjoyable and less frustrating, which are essential for enhancing UX.



What are the evaluation methods?

- Timing of analysis (e.g., throughout the application development stage: early, middle, late/after)
- Type and number of evaluators (e.g., several HCI experts vs. hundreds of domain users)
- Formality (e.g., controlled experiment or quick and informal assessment)
- Place of the evaluation (laboratory vs. in situ field testing)