

Chapter 9

- Evaluation time is when system is working the way we desired. We have baseline which must be available. Baseline is taking manual system to computer and check if it gives benefit or not. Baseline can be old product which we are modernizing.

- Usertesting is expensive, we test it after making design so changes would be difficult. We want that we do it in every stage.

Prototype is made & checked by user. Short prototype is made & checked by user.

- Methods are proposed for the evaluation of system such that a disaster change is not made. Expert analysis is system is check by the one who have expertise in usability. Involve expert evaluator and designer other than users. Non experts can't tell about how to solve it, it only tell about issues.

- Symbol and metaphors are two different thing. E.g. this symbol is not the metaphor of music. Symbol is not real world thing we just assign something to objects. Metaphors are real world thing.

2

Evaluation is not made of anything like login screen unless you made any changes in it. You made evaluation where difficulty can be caused. Expert user gives more fruitful result.

no people
Training, less training required so, expert evaluation is cheaper. It's like mechanic vs engg. However expert user are not users. Users are still required. Some things are only evaluated by users.

Design Critique: User can be used to solve minor issues.
• Finding flaws that should be redesign.
Not whole system is redesign.
• When you know there is problem but you need evidence so expert can be used here. They can give idea that it's because this type problem.

• Difference b/w user & expert test.
can't articulate problem → can tell prob with solution.

(کس طرح پر)
Qualitative
research

3

(آپ اور کیسے پر)
Quantitative
research

- gain insight
- why something happens

- Next time if we use prior knowledge, it will apply or not.

- Think about protocol

- get to know about problem. we take about summary here. It just tell system is good but not why.

- If system is better so it is better or not

- No think about protocol

① Cognitive Walkthrough

what user is thinking.

- How system is helpful to user. System support user learning. Without user we come to know about user's cognition, this is the goal here.
- Expert evaluation technique.
- Domain expert can do cognitive walk for smaller usability testing.
- We user can do it by metaphors, by checking recall & recognition.
- It is walkthrough interface by expert. Each steps have question and answer is given by interface.
- Sequence of actions that an interface have is

4

goal → cognitive load decrease

evaluated by user.

Every experts follow steps and evaluate if normal user pass through these steps what problem will they face.

For each task walkthrough:

- 1) What impact interaction have on user?
- 2) What ^{cognitive} process? ← ^{recognition} ^{recall} ^{emotion}
- 3) What learning problem? ← metaphor problems
slip & mistake problem
system not consistent

When we see design, our goal is one point.

What is required for it?

- 1) layout must be brainstormed. It should be detailed, not complete.
- 2) Full system is not tested, an important task is tested.
- 3) Designers tell about list of actions to achieve some goal.
- 4) who and what user are.

Evaluate:

- 1) User goal is same as user action.
- 2) Will user see, action is available.
- 3) It will recognize by seeing it, when see the available action.
- 4) Will user get the feedback. (Perception/Intercept)

Eg VCR was complicated.
- Firstly we break it into task/action sequence / user response.

- Second evaluation technique which is based on heuristic evaluation.

Task adequacy is your task complete but your task remain due to some reason. So you can now make plans.

② Heuristic Evaluation:

Heuristic improve usability. Structuring the critique of system is important. If you don't have structure to evaluate, insight/info is not useful. If 10 persons are evaluating so they work together.

Such heuristics can be used for making design & also ^{can} measure that it's good or not.

- Common sense method or experiential method are heuristic. They are based on experience. They are steps which we know that it will lead to good result.

- If heuristic is violated means our system is not working properly.

6

- It can be used for design evaluation / specification
We want to show always a good design for
evaluation. "Discount Usability Technique" is its
another name.

- It helps in finding usability problem in UI.
Different evaluator evaluates different problem
& then communicate afterwards.
(Independent first, gather afterwards)

- We do heuristic evaluation by following
principles.

- Phases of heuristic evaluation:

- 1) We have to give domain knowledge to
evaluators.
- 2) Step / scenario to follow to build
consistency that each evaluator follows same
thing.
- 3) Severity rating → defined scaling for
each problem.

- Nielsen's ten point:

- 1) Always keep users inform about what is
going on.
- 2) Metaphors / natural group.
- 3) User control / freedom. User learn from
trial and error.

- 4) Consistency & standard should follow.
- 5) Error prevention. eg give pop ups on each stage.
- 6) Recognition rather than recall.
- 7) Flunbility / short cuts
- 8) Aesthetic & minimalist design → cognitive load should be low.
- 9) Recognize and Recovery of error.
- 10) Help & documentation.

Other Methods:

Model based evaluation

Design Rationale
↓

- Cognitive methods to filter design.
 - Mental model is used.
 - GOMS →
 - low leveling modeling technique is Keystroke-level (by pressing button or number of clicks) decide the evaluation.
 - GOMS → Goals are objective. Operators are all functions available to user. Methods are way to perform operations, there are many way to do one operation. Select is the way of choosing method.
- Process oriented have too many thinks.
 - Structure oriented have good simpler.

Keystroke model say that we want to calculate time for a single activity and we don't want to evaluate it through user, we want to do it by ourselves.

GOMS is used to make models and just apply the numbers of Keystroke to evaluate.

Review based evaluation: (mostly used)

Results from literature is used to support or deny parts of design.

Evaluating through user participation:

We can't ignore user evaluation. Here we have to deal with many problems. It is used for filtering / refining. Experts have done evaluation and then users. All the evaluation methods we studied, used here (combination of all)

- How to gather people
 - How to deploy system on their PC.
 - How to take responses from multiple user
- } Problem

Laboratory studies:

Environment is our own and user comes here to test. Facilities & equipment are of our own so everything will be according to our requirement.

- 9
- Facilities & environment are under our control
 - Content is lose because when user worked from home. The environment will be different.
eg: Election evaluation, flen logout after 5 min
 - Appropriate when system loc is dangerous
 - When we want to test isolate things.

Field Studies Deploy system to user's computer. User gets the real world env of their. Natural environment & interruptions are advantage. Distraction and noise are disadvantage. Its because req changes see to product.

Important Experimental factors

- 1) Subjects \swarrow Sufficient sample of each representative must be representation of real user
- 2) Variable \swarrow Dependent \rightarrow we change something
Independent \rightarrow affect of those change
- 3) Hypothesis \rightarrow Prediction of the outcome of exper
- 4) Experimental design.

3) \rightarrow Null hypothesis \rightarrow you can't prove anything
eg: It says DV don't have any affect on IV
If we reject it so it become false
and our claim is true.

Experimental design:

10

experimental condition Control

- We get familiar with activity between groups requires more user, no transfer of learning. It is bias if in some group there are people who are more genius so results will then bias.
- Within groups every one do the same experiment. Variation is not affect here. No of people required here is less. Its less costly and causes less biases. Transfer of learning is easy.

Analysis of data

Parametric	Non parametric	Contingency Table
<ul style="list-style-type: none">- Prediction making is easy, Powerful.- By looking normal distribution, prediction is made.	<ul style="list-style-type: none">- We have to see actual data point.- like recommendation system.	<ul style="list-style-type: none">- No of occurrence is noted here.

Subject groups:

Expensive, long time for settle down, difficult to timetable.

- Task must contain collaboration so it helps in testing. Creative task are one in which people do communication ^{to comp task} → Zoom (call or msg).
- Decision games is to select something after discussion. Control task is one in which single person can't do something, with collaboration you can do task.

- Data gathering, people use annotation tools in video.
- Variation can be in b/w or within gap
- Distribution cognition → on different loc or diff hardware is used.

Observational Methods:

- Think aloud → you observe user, but not disturb her. You give task & just observe.
- Cooperative → User is with designer & both can do questions with each other. User critique to system.

Protocol analysis → tools are given to
user to evaluate & observe. Record is
maintain by this.

- Post task walkthrough → You can user
about how to improve a task.

• You can analyse immediately or delayed.

↓
You just ask general
question but not
related to data and
user is fresh in mind
here.

↓
You get
time to analyse
data but
dis is that
user don't rem.
why he do that

There is no option which one is good, you
can use any acc to task or hybrid too.

