**Formative Usability Test**

Formative evaluation is a type of usability evaluation that helps to "form" the design for a product or service. Formative evaluations involve evaluating a product or service during development, often iteratively, with the goal of detecting and eliminating usability problems.

One important aspect of formative evaluation is that the audience for the observations and recommendations is the project team itself, used to immediately improve the design of the product or service and refine the development specifications. Results can be less formal than in summative evaluation, as suits the needs of designers, developers, project managers, and other project participants. As in all user-focused methods, it is important to recruit participants that are very similar to the expected end-users of the system. The wrong participants can lead to unreliable or misleading results that can put projects at risk. HFE guidance on [**participant recruiting and use**](https://vaww.portal2.va.gov/sites/humanfactors/HFBoK/SitePages/Working%20with%20Study%20Participants.aspx) can help avoid problems in this area.

Although some sources group usability inspection methods under formative evaluation, using the term to describe many methods that are suitable early in the design process, HFE primarily uses the formative moniker in relation to testing with users and representative tasks. Several types of formative usability tests used by HFE include:

**A/B Test** – The Comparison (Contrast) Test is carried out by having a group of representative users perform representative tasks on two or more alternative system designs for the purpose of comparing user performance between the designs. Using within-group (comparing the same users performance across all interfaces) or between-group (each participant is only exposed to one interface, and their performance is compared to similar participants using another interface) comparisons will provide project teams with empirical data for design decisions. Usability.gov and others use this moniker for unmoderated tests done on a live site, but HFE uses the more traditional meaning.

**Assessment Test** – A test to compare the usability of a design against specific measureable performance and/or satisfaction goal, much like a summative evaluation, but with a lower level of effort and in an earlier development phase than a summative evaluation.

**Exploratory Test** – Exploratory Testing is a user-based evaluation of unimplemented technology or software concepts. The technique is used to uncover interaction design issues during an exploratory phase of a project; it often involves having the project team simulate some or all the responses of the system. This approach can be used to explore design and usability of emerging technologies such as natural language applications or pervasive computing applications. It is also well suited to testing paper prototypes early in the design process.

**Benefits**

* Identifies major usability problems that frequently aren't caught by inspection methods. It can uncover problems that are driven by user-specific expectations, needs, and abilities.
* Obtains reliable measures of users' effectiveness, efficiency and satisfaction.
* Provides feedback that reveals possible design flaws and other issues.
* Experienced test facilitators can elicit feedback from users to help understand why they had problems.
* Low and medium-fidelity prototypes are cost-effective to test.
* It is easy to have business office representatives, UI Designers and developers as observers.
* In some cases you can produce video clips, screen shots, and quote from test sessions to illustrate problems.

**Limitations**

* Not all problems will be found with small samples of users. (Nielsen, 2012)
* You may not have access to users that match the user profile.
* Lab testing takes users away from their natural work environment.
* Technical setup may be complex and require domain experts and additional time for setup and debugging.
* An inexperienced facilitator can influence the results by using too many hints, asking biased questions, or providing nonverbal cues about the tasks.

**Study Execution**

|  |  |
| --- | --- |
| **​Milestone** | **​Owner** |
| ​Initiate Kick-off Call | HFE |
| Develop Proposal | ​                                  HFE |
| ​Attend Objectives Finalization Meeting | HFE, ​Business Office |
| ​Develop Tasks | ​Business Office |
| ​Verify that tasks meet objectives | ​                                  HFE |
| Provide system access or a representation of the system | ​                         Business Office​ |
| ​Create test data | ​Business Office |
| ​Verify Environment | ​HFE |
| ​Recruit Participants | ​Business Office​ |
| ​Schedule Participants | HFE |
| ​Pilot test the study materials and proceedures | ​HFE |
| ​Conduct usability assessment | HFE |
| Attend results briefing | HFE​, Business Office, Developer |
| ​Analyze findings, write report | HFE |
| ​[**Conduct After Action Review**](https://vaww.portal2.va.gov/sites/humanfactors/SitePages/AAR%20Process.aspx) | HFE |

**Outcomes**

* A results briefing to discuss what solutions are possible for usability issues within the given constraints.
* Summarize the results of the satisfaction questionnaire, task time and effectiveness (accuracy and completeness) measures if used.
* Produces a list of usability problems, categorized by severity and frequency, and an overview of the types of problems encountered.

The type of objective and subjective data collected during testing may include:

* Ability and time to complete a task.
* Sequence and number of steps to complete a task.
* Types and numbers of errors.
* Number of repeated errors.
* Number of design issues that lead to user confusion.
* Ratings of ease of performing a task.

Severity rankings are assigned to each problem. These rankings can be determined by the impact of the problem and the persistence of the problem.

|  |  |
| --- | --- |
| **Phase of Development​** | |
| ​ | ​Planning, Scoping & Definition |
| ​ | ​Requirements Gathering |
| **✔** | **​Early Design** |
| **​✔** | **​Detailed Design & Development** |
| **​✔** | **​Field Testing** |
| ​ | ​Deployment |
| **✔** | **​Post-Deployment** |

|  |  |
| --- | --- |
| **​Study Characteristics ​** | |
| ​**Timeframe** | ​5-6 weeks |
| **​Level of Effort** | High |
| ​**Data Collection** | Behavioral, Attitudinal |
| ​**Data Reporting** | ​Quantitative, Qualitative |

|  |
| --- |
| **Referenced on this page...** |
| Nielsen, J. (2012). How Many Test Users in a Usability Study?. In Alertbox. Retrieved February 1, 2014, from [**http://www.nngroup.com/articles/how-many-test-users/**](http://www.nngroup.com/articles/how-many-test-users/).  Nigel Bevan, Narendra Singhal, Ben Werner, Duane Degler, Chauncey Wilson. (2009). Formative Evaluation. In Usability Body of Knowledge. Retrieved February 1, 2014, from [**http://www.usabilitybok.org/formative-evaluation**](http://www.usabilitybok.org/formative-evaluation). |