My Art My Access

User Testing Plan

Nicholas Mitchell
04/25/2013

1 Table of Contents

2 Document Overview

This document describes a test plan for conducting a user test during of the My Art My Access user-interface The goals of user test include establishing a baseline of user performance, establishing and validating user performance measures, and identifying potential design concerns to be addressed in order to improve the efficiency, effectiveness, and end-user satisfaction.

The user test objectives are:

- To determine design inconsistencies and usability problem areas within the user interface and content areas. Potential sources of error may include:
  - Navigation errors – failure to locate functions, excessive keystrokes to complete a function, failure to follow a recommended screen flow.
  - Presentation errors – failure to locate and properly act upon desired information in screens, selection errors due to labeling ambiguities.
  - Concept errors – failure to understand website purpose or content value.
- Exercise the web site under controlled test conditions with representative users. Data will be used to access whether user testing goals regarding an effective, efficient, and well-received user interface have been achieved.
- Establish baseline user performance and user-satisfaction levels of the user interface for future user-testing evaluations.

User test participants should be representative of user groups and demographics that My Art My Access will serve. This will include faculty, staff, and students currently being underserved by the current artwork database interface as well as faculty, staff and graduate students that are well-served in disciplines, such as art history. For both the well-served and the underserved groups, user samples sizes of 3-4 will be most appropriate. It is also recommended that user-testing take place in the School of Information usability lab to ensure consistent and accurate results.

3 Executive Summary

This user test is designed to meet two important goals for the development of the user interface for My Art My Access.

[Summarize specific details of the usability test for the given application or Web site; describe specific functions to be evaluated. Summarize the usability goals.]

Upon review of this usability test plan, including the draft task scenarios and usability goals for My Art My Access documented acceptance of the plan is expected.

4 Methodology

User Testing Group

For the user test, it is recommended that individual responsible for conducting and administering the test locate 6-8 participants total. There should be anywhere from 3-4 users in a user-group that is representative of faculty, staff, or students from disciplines and demographics that are under-served (for example: French History, Architecture) by the existing artwork interface. The second user group should consist of users that are well-served, or at least targeted by the existing DASE interface in the field of art history

User Test Setting

To ensure the most accurate and consistent results are collected, the user test should be held in the usability testing lab. The lab environment will allow testers to both observe and record actions that may otherwise go un-noticed

Tools and Data Collection

The user test should be conducted in two parts. Part one will be the task portion of the user test. Users should work through the tasks on the web site using the latest version of *Mozilla Firefox*. Task actions and data should be gathered and recorded using the usability software *Morae.* Test administrators should also record user tasks by taking observational notes at the time of the test. In part two, participants will be asked to take part in a follow-up interview where they will be answer questions for gathering some qualitative and some quantitative data.

4.1 Participants

The user test should include a minimum of six to a maximum of eight users. These 6-8 users should be representative of two distinct user groups. A user group of 3-4 users should be from under-served groups from diverse disciplines. These users will be a university faculty, staff, or graduate student that is an intended audience for My Art My Access. Possible options in this user group might be individuals with discipline knowledge in Architecture, French History, Philosophy, Studio Art or Materials Science. The second user group should be a representative group that has been the targeted-audience for the existing DASE interface. Possible options in this category might be faculty or graduate students in art history, museum staff, or staff in a field related to art history. Each participant should meet the following criteria:

- At least some interest in artwork.
- Experience researching or studying artwork
- Experience with traditional digital art collections.
- A plus but not necessary, experience with tagging

The participants' responsibilities will be to attempt to complete a set of representative task scenarios presented to them in as efficient and timely a manner as possible, and to provide feedback regarding the usability and

acceptability of the user interface. The participants will be directed to provide honest opinions regarding the usability of the application, and to participate in post-session interviewing.

Test administrators will reach out to potential participants by connecting through email and in-person conversations. They will also work with Dr. Unmil Karadkar to work on building and acquiring participants.

4.2 Training

The participants will receive and overview of the usability test procedure, equipment and software.

4.3 Procedure

Participants will take part in the user test at Usability Lab in the School of Information at the University of Texas at Austin. A Windows computer with the Web site/Web application and supporting software will be used in a typical office environment. The participant's interaction with the Web site will be monitored by the facilitator seated in the same office. Note taking and data logging will done during sessions in observation room, connected by video camera feed and/or from a one-way mirror. The test sessions will be videotaped.

Test Administrators will brief the participants on the Web site and instruct the participant that they are evaluating the application, rather than the facilitator evaluating the participant. Participants will sign an informed consent that acknowledges: the participation is voluntary, that participation can cease at any time, and that the session will be videotaped but their privacy of identification will be safeguarded. The facilitator will ask the participant if they have any questions.

Participants will complete a pretest demographic and background information questionnaire. The facilitator will explain that the amount of time taken to complete the test task will be measured and that exploratory behavior outside the task flow should not occur until after task completion. At the start of each task, the participant will read aloud the task description from the printed copy and begin the task. Time-on-task measurement begins when the participant starts the task.

The test administrators will instruct the participant to 'think aloud' so that a verbal record exists of their interaction with the Web site. The test administrator will observe and enter user behavior, user comments, and system actions in the data logging application.

After the test administrators complete the user task portion of the test, the participant will begin the second part interview to elaborate on the task session and gather additional information regarding system usefulness.

5 Roles

The roles involved in a user test are as follows. An individual may play multiple roles and tests may not require all roles.

Trainer

- Provide training overview prior to user testing.

Facilitator

- Provides overview of study to participants
- Defines usability and purpose of user testing to participants
- Assists in conduct of participant and observer debriefing sessions
- Responds to participant's requests for assistance

Data Logger

- Records participant’s actions and comments

Test Observers

- Silent observer
- Assists the data logger in identifying problems, concerns, coding bugs, and procedural errors
- Serve as note takers.

Test Participants

- Provides overview of study to participants
- Defines usability and purpose of usability testing to participants
- Assists in conduct of participant and observer debriefing sessions
- Responds to participant's requests for assistance

5.1 Ethics

All persons involved with the user test are required to adhere to the following ethical guidelines:

- The performance of any test participant must not be individually attributable. Individual participant's name should not be used in reference outside the testing session.
- A description of the participant's performance should not be reported to his or her manager.

6 User Testing Tasks & Interview

Listed below are the tasks that should be completed during the Usability portion of the User Test and the Follow-up interview questions. The usability tasks seek to establish a baseline understanding user interaction with the features and functions available in My Art My Access. During all usability tasks, users should be encouraged to talk-aloud about their processes. Section 7 and Section 8 discuss Usability Metrics and Usability Goals respectively. The interview component seeks to establish a proof-of-concept feedback analysis as well as call attention to website usefulness.

Part 1: User Tasks (Usability Test)

1. For the first task, users will be asked to find five images using the website search. The artwork will be defined prior to the user test. During this task,

test administrators should give users an artist, title, or other artwork with defined criteria to the user one at a time.

2. The second task will be similar to the first task. This time users will be asked to find five images using the website browse navigation. The artwork will be defined prior to the user test. During this task, test administrators should give users an artist, title, or other artwork with defined criteria to the user one at a time.
3. The third user task will ask that the user add one field of artwork description. The user will be asked to do this to five different images. The artwork will be defined prior to the user test. During this task, test administrators should give users an artist, title, or other artwork with defined criteria to the user one at a time..
4. The final task will be for users to add 3-5 metadata description fields to a single artwork. The artwork will be defined prior to the user test. During this task, test administrators should give users an artist, title, or other artwork with defined criteria.

Part II: Follow-up Interview

- At any point was there something missing, or something that you expected to see that was not there?
- Was there ever any unnecessary information that you felt was distracting or could be removed?
- Using a Likert scale, how useful would a website like this (one that allows expert description of artwork) useful to you or your institution?
- If you answered that it would, in what ways would it benefit? If it would not be very useful, why would it not be useful to you?
- Using a Likert scale, do you believe site layout is sufficient for allowing users to find and describe artwork?
- Were items organized in a way that was clear to you?
- On a Likert scale, do you believe that a website like this could meet or help to meet educational objectives?
- What were three things that you liked about the website?
- What were three things you did not like about the website?
- On a Likert scale, was browse navigation sufficient with Artist, Title, Work Type, and Nationality for finding artwork?
- If you disagree, that it is not sufficient, what do you wish would have been added?
- If you could change one thing about the site (it may be a big change or small change) , what would that be?
- On a Likert scale, do you believe it is useful to you have a domain specific (French History, Architecture, etc.) labels attached to custom description fields?
- If you disagree that this would not be useful, why?

7 Usability Metrics

Usability metrics refers to user performance measured against specific performance goals necessary to satisfy usability requirements. Scenario completion success rates, adherence to dialog scripts, error rates, and

subjective evaluations will be used. Time-to-completion of scenarios will also be collected.

7.1 Scenario Completion

Each scenario will require, or request, that the participant obtains or inputs specific data that would be used in course of a typical task. The scenario is completed when the participant indicates the scenario's goal has been obtained (whether successfully or unsuccessfully) or the participant requests and receives sufficient guidance as to warrant scoring the scenario as a critical error.

7.2 Critical Errors

Critical errors are deviations at completion from the targets of the scenario. Obtaining or otherwise reporting of the wrong data value due to participant workflow is a critical error. Participants may or may not be aware that the task goal is incorrect or incomplete.

Independent completion of the scenario is a universal goal; help obtained from the other usability test roles is cause to score the scenario a critical error. Critical errors can also be assigned when the participant initiates (or attempts to initiate) and action that will result in the goal state becoming unobtainable. In general, critical errors are unresolved errors during the process of completing the task or errors that produce an incorrect outcome.

7.3 Non-critical Errors

Non-critical errors are errors that are recovered from by the participant or, if not detected, do not result in processing problems or unexpected results. Although non-critical errors can be undetected by the participant, when they are detected they are generally frustrating to the participant.

These errors may be procedural, in which the participant does not complete a scenario in the most optimal means (e.g., excessive steps and keystrokes). These errors may also be errors of confusion (ex., initially selecting the wrong function, using a user-interface control incorrectly such as attempting to edit an un-editable field).

Noncritical errors can always be recovered from during the process of completing the scenario. Exploratory behavior, such as opening the wrong menu while searching for a function, will be coded as a non-critical error.

7.4 Subjective Evaluations

Subjective evaluations regarding ease of use and satisfaction will be collected via post-task interview. The interview will utilize free-form responses and rating scales.

8 Usability Goals

The next section describes the usability goals for My Art My Access.

8.1 Completion Rate

Completion rate is the percentage of test participants who successfully complete the task without critical errors. A critical error is defined as an error that results in an incorrect or incomplete outcome. In other words, the completion rate represents the percentage of participants who, when they are finished with the specified task, have an "output" that is correct. Note: If a participant requires assistance in order to achieve a correct output then the task will be scored as a critical error and the overall completion rate for the task will be affected.

A completion rate of 100% is the goal for each task in this usability test.

8.2 Error-free rate

Error-free rate is the percentage of test participants who complete the task without any errors (critical **or** non-critical errors). A non-critical error is an error that would not have an impact on the final output of the task but would result in the task being completed less efficiently.

An error-free rate of 80% is the goal for each task in this usability test.

8.3 Time on Task (TOT)

The time to complete a scenario is referred to as "time on task". It is measured from the time the person begins the scenario to the time he/she signals completion.

8.4 Subjective Measures

Subjective opinions about specific tasks, time to perform each task, features, and functionality will be surveyed. At the end of the test, participants will rate their satisfaction with the overall system. Combined with the interview/debriefing session, these data are used to assess attitudes of the participants.

9 Problem Severity

To prioritize recommendations, a method of problem severity classification will be used in the analysis of the data collected during evaluation activities. The approach treats problem severity as a combination of two factors - the impact of the problem and the frequency of users experiencing the problem during the evaluation.

9.1 Impact

Impact is the ranking of the consequences of the problem by defining the level of impact that the problem has on successful task completion. There are three levels of impact:

- High - prevents the user from completing the task (critical error)
- Moderate - causes user difficulty but the task can be completed (non-critical error)
- Low - minor problems that do not significantly affect the task completion (non-critical error)

9.2 Frequency

Frequency is the percentage of participants who experience the problem when working on a task.

- High: 30% or more of the participants experience the problem
- Moderate: 12.5% - 29% of participants experience the problem
- Low: 12.5% or fewer of the participants experience the problem

9.3 Problem Severity Classification

The identified severity for each problem implies a general reward of an improved user experience, and a general risk for not addressing it, in the current release.

Severity 1 - High impact problems that often prevent a user from correctly completing a task.

Severity 2 - Moderate to high frequency problems with moderate to low impact are typical of erroneous actions that the participant recognizes needs to be undone.

Severity 3 - Either moderate problems with low frequency or low problems with moderate frequency; these are minor annoyance problems faced by a number of participants.

Severity 4 - Low impact problems faced by few participants; there is low risk to not resolving these problems.

10 Reporting Results

The Usability Test Report will be provided at the conclusion of the usability test. It will consist of a report and/or a presentation of the results; evaluate the usability metrics against the pre-approved goals, subjective evaluations, and specific usability problems and recommendations for resolution. The recommendations will be categorically sized by development to aid in further implementation strategy.