Dashboard

Usability Test Plan

version: UT-RD_extended(32019-1)

Cheryl Lewman
March 18, 2019

Table of Contents

Document Overview

This document describes a test plan for conducting a usability test during the development [   ] e goals of usability testing 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, productivity, and end-user satisfaction. Our specific goals for these set of two tests was for icon recognition, to follow click paths, and to see how users interact with the chatbot.

The usability test objectives are:

1. To determine design inconsistencies and usability problem areas within the user interface and content areas. Potential sources of error may include:
   a. Navigation errors – failure to locate functions, excessive keystrokes to complete a function, failure to follow recommended screen flow.
   b. Presentation errors – failure to locate and properly act upon desired information in sc   due to labeling ambiguities.
   c. Control usage pr   u or entry field usage.
2. Exercise the application   olled test conditions with representative users. Da   ss whether usability goals regarding an effective, efficient, a   terface have been achieved.
3. Establish baseline user p   atisfaction levels of the user interface for future usability evalu

Testing included two ses   , a remote User Testing service. Maze is online and users are not mor   sers are given a specific set of instructions and asked to comp

Executive Summary

We're lo   e with specific parts of the app: 1. Are users accessin   rd or through the menu? 2. Icon recognition- specifica   es, calendar, settings, and menu. 3. How are users interacti   users able to locate various links with few misclick

Upon re   an, including the draft task scenarios and usability goals fo   nted acceptance of the plan is expected.

Methodo

The usa   edesign will be conducted on March 14 and 18, 2019. All tests   sign and are unmonitored.

Participants

For this round of testing, we recruited remote testers through Maze.design. We will perform two testing sessions with 25 users in each session, for a total of 50 users. Users are all English speakers, men and women, between the ages of 18-70 years.

Training

Maze doesn't require any kind of special training.

Procedure

Users are provided by Maze.design and are given their requirements through Maze. Maze has an interface that allows users to test directly from the ~~InvisionApp~~ screens in their web browser. Maze states the task (what Maze calls a Mission), the task is then displayed on the left side of the browser. The right side of the browser contains a mock iPhone X, which is what they use to do click-thrus. They are provided with comprehensive click-thrus that allow the user to explore the app. After the testing session is over, users are asked the question, "What did you think of this prototype?" Their answers are recorded.

Users are required to use a web browser during the testing session. The web browser simulates an iPhone X where they are able to click-through the app. Users have no previous known knowledge of any previous Gov2Go products.

The participants' responsibilities are 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 are directed to provide honest opinions regarding the usability of the application, and to participate in post-session feedback.

The participant will be seated at their workstation in their work environment. They will login to Maze.design and be greeted with a message explaining the test. They will use their personal computers and go through each task, or mission. On the left side of the screen, they will see their mission. On the right side of the screen, they will see the Invisionapp prototype in Maze's environment. The environment provided for in this test is iPhone X. Users' time-on-task measure will begin and all click thrus will be recorded, with hotspots created and recorded for each click. Each mission has three possible outcomes: 1. *Direct Success*. This means the user followed the exact path we expected. 2. *Indirect Success*. The user was able to complete the task, but found it using an unexpected path. 3. *Fail*. The user wasn't able to complete the task. After all tasks have been attempted, the participant will answer a post-test satisfaction question.

Roles

Trainer

·   Users are vetted through Maze.design and are given information on how the process works.

Facilitator

·   There is no official facilitator for this test.

Data Logger

·   Maze.design records participant's actions and comments.

Test Observers

·   There are no observers for this test.

Test Participants

·   Maze.design provides overview of study to participants

Usability Tasks

The usability tasks were derived from test scenarios developed from use cases. Due to the range and extent of functionality provided in the application or Web site, and the short time for which each participant will be available, the tasks are the most common and relatively complex of available functions. The tasks are identical for all participants of a given user role in the study.

The task descriptions below are required to be reviewed by the UX Manager to ensure that the content, format, and presentation are representative of real use and substantially evaluate the total application. Their **acceptance is to be documented** prior to usability test.
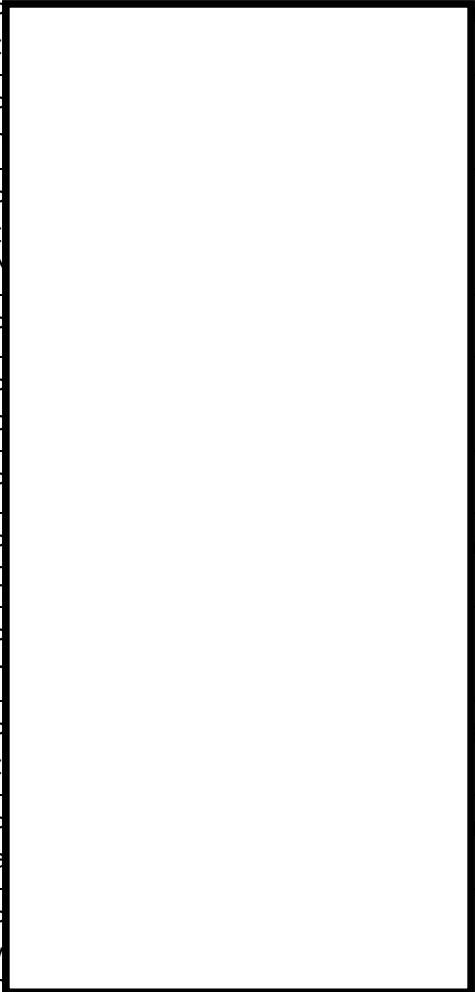
Test 1

Mission (Task)	Heading	Detailed Instructions
1	LOGIN	P
2	MENU DISCOVERY	P
3	SETTINGS DISCOVERY	P
4	NOTIFICATIONS DISCOVERY	P N
5	ACCESSIBILITY STATEMENT	P S
6	CAR TAG RENEWAL SERVICES	P n the CAR TAG RENEWAL SERVICE.

7	MESSAGES DISCOVERY	P...M...
8	P... DISCOVERY	P...on to...dia N...
9	CALENDAR DISCOVERY	P...dar.
10	CALENDAR DATE VIEW(1)	P...ary 1...
11	CALENDAR DATE VIEW(2)	P...AILS v...
12	CHATBOT DISCOVERY	P...op.
13	CHATBOT TEXT	P...LD. S...
14	CHATBOT SEND	P...top.
15	CHATBOT CONFIRM	P...d...ce.
16	CHATBOT CLOSE	P...s...
17	LANGUAGE DISCOVERY	P...c...S...

Test 2

Mission (Task)	Heading	Detailed Instructions
1	LOGIN	P...
2	NOTIFICATIONS DISCOVERY	P...n...
3	ACCESSIBILITY STATEMENT	A...o...a...e...d...e t...S...

4	SIGNUP FOR A SERVICE	P...n t...
5	M̶̶̶̶̶̶̶̶̶̶DISCOVERY	P... n...
6	P̶̶̶̶̶̶̶̶̶̶COVERY	P... t... N...
7	CHATBOT DISCOVERY	P...
8	CHATBOT SEND	P... c...
9	EVENTS DISCOVERY	P...
10	ALERTS DISCOVERY	P...s C...
11	INVOICE DISCOVERY	P... n...
12	DISCOVERY	P...t t...
13	CHANGE PASSWORD	P...r a...
14	C̶̶̶̶O SUPPORT D̶̶̶̶RY	P... w... h...

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. We will record if they completed the task using the route we prescribed, an indirect route, or if they failed to complete the task. We will also record the hotspots to see where users are clicking, as well as what pages they click through.

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.

Critical Errors (Fail)

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.

Non-critical Errors (Indirect Success)

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.

Subjective Evaluations

Subjective evaluations regarding ease of use and satisfaction will be collected at the conclusion of the session.

Scenario Completion Time (time on task)

The time to complete each scenario, not including subjective evaluation durations, will be recorded.

Usability Goals

The next section describes the usability goals for Gov2Go-Dashboard.

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. .

A completion rate of 90% is the goal for each task in this usability test.

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.

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.

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. We will also examine all remote users to verify the validity of their testing session.

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.

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)

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: 11% - 29% of participants experience the problem
- · Low: 10% or fewer of the participants experience the problem

Problem Severity Classification

The identified severity for each problem implies a general reward for resolving it, 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. They occur in varying frequency and are characteristic of calls to the Help Desk. Reward for resolution is typically exhibited in fewer Help Desk calls and reduced redevelopment costs.

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. Reward for resolution is typically exhibited in reduced time on task and decreased training costs.

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. Reward for resolution is typically exhibited in reduced time on task and increased data integrity.

Severity 4 - Low impact problems faced by few participants; there is low risk to not resolving these problems. Reward for resolution is typically exhibited in increased user satisfaction.

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 implementation strategy.