# Shop Aware

# Usability Test Plan

**Protoype**

**Theresa DeRosa**

**June 24, 2015**

## Table of Contents

[Document Overview 3](#_Toc359583313)

[Executive Summary 3](#_Toc359583314)

[Methodology 3](#_Toc359583315)

[Participants 4](#_Toc359583316)

[Training 4](#_Toc359583317)

[Procedure 4](#_Toc359583318)

[Roles 5](#_Toc359583319)

[Trainer 5](#_Toc359583320)

[Facilitator 5](#_Toc359583321)

[Data Logger 6](#_Toc359583322)

[Test Observers 6](#_Toc359583323)

[Ethics 6](#_Toc359583324)

[Usability Tasks 6](#_Toc359583325)

[Usability Metrics 7](#_Toc359583326)

[Scenario Completion 7](#_Toc359583327)

[Critical Errors 7](#_Toc359583328)

[Non-critical Errors 7](#_Toc359583329)

[Subjective Evaluations 7](#_Toc359583330)

[Scenario Completion Time (time on task) 8](#_Toc359583331)

[Usability Goals 8](#_Toc359583332)

[Completion Rate 8](#_Toc359583333)

[Error-free rate 8](#_Toc359583334)

[Time on Task (TOT) 8](#_Toc359583335)

[Subjective Measures 8](#_Toc359583336)

[Problem Severity 8](#_Toc359583337)

[Impact 9](#_Toc359583340)

[Frequency 9](#_Toc359583341)

[Problem Severity Classification 9](#_Toc359583342)

[Reporting Results 9](#_Toc359583343)

## Document Overview

This document describes a test plan for conducting a usability test during the development of ShopAware.gov The goals of usability testing include establishing a baseline of user performance, identifying potential design concerns to be addressed in order to improve the efficiency and end-user satisfaction of data results returned to protect their health.

The usability test objectives are:

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

We are encompassing a group of people to represent the Public that consists of thirteen people. The testing will be done on June 30th in the conference room by some users using portable devices. The other users will be observed at their local workstations.

## Executive Summary

Upon review of this usability test plan, including the draft task scenarios and survey questions for ShopAware.gov, documented acceptance of the plan is expected.

## Methodology

We will have thirteen participants from all areas of the company. Each participant with be observed either seated at their own workstation or using a portable device in the conference room. We will be assessing the sites ease of use and return of useful information to the public.

### Participants

The thirteen 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 website. The participants will be directed to provide honest opinions regarding the website, and to participate in the post-session survey.

The participants were selected based on a variation of age and vision concerns.

### Procedure

Multiple computers and portable devices 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 or conference room.

The facilitator will brief the participants on the Web site and instruct the participant that they are evaluating the website, rather than the facilitator evaluating the participant. The facilitator will ask the participant if they have any questions.

After all task scenarios are attempted, the participant will complete the post-task survey.

## Roles

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

### Facilitator

* Provides overview of study to participants
* Defines usability and purpose of usability testing to participants
  + Assists in conduct of participant
  + Responds to participant’s requests for assistance
* Records participant’s actions and comments
* Silent observer
* Identifies problems, concerns, coding bugs, and procedural errors
* Serves as note taker

### Ethics

All persons involved with the usability 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.

## 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 Web site, and the short time for which each participant will be available, the tasks are the most common of the 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 application owner, business-process owner, development owner, and/or deployment 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.

The scenarios are to create a shopping list that encompasses food, drugs, and medical devices that will return a recall, or adverse action. This is giving the consumer the ability to see this risk first hand when they are making out their shopping list. We will provide the ability to have your shopping list save for when you return to the site. You can have a login if you want to receive alerts if an item you buy has a recall on it.

## 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. [include or delete any metrics not used in the planned test]

### 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

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

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 via surveys, and during debriefing at the conclusion of the session. The survey will utilize a rating scale.

### 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 ShopAware.gov

### 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%/enter completion rate 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%/enter error-free rate 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. Combined with the interview/debriefing session, these data are used to assess attitudes of the participants.

## 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. The report is anticipated to be delivered to the Project UCD Contact by 7/1/2015.