Chapter 3
An Adventure in Usability:
Discovering Usability Where it was not Expected

Holly Blasko-Drabik
University of Central Florida, USA

Tim Smoker
University of Central Florida, USA

Carrie E. Murphy
University of Central Florida, USA

ABSTRACT
Usability testing can be used as an effective tool throughout the design and development of a serious game. User responses to the usability measures can help designers meet the challenges of creating a game that is educational, playable, and entertaining. However, there are a multitude of measures to choose from, some empirically created and tested and some created by developers to test a single game. This chapter will provide a broad overview of usability, two major methods of conducting usability testing and how usability can be applied to testing, and more importantly, improving games. Two possible usability measures are discussed from a game evaluation standpoint and compared highlighting on which aspects of usability heuristics they focus.

INTRODUCTION
Imagine you are playing a simple driving game on your computer. The goal is simple: drive as long as you can without striking any obstacles in the roadway. To do this, you hit the space bar to shift from one lane to another. However, the game itself is drab and simplistic. The car is represented by a colored oval, obstacles are represented with black and white rectangles, and the background is grey. While not exactly exciting or engaging, this game is perfect from the standpoint of usability: all you need to know is how to press the spacebar and you can play.

Now imagine a second, similar game with the same goal of avoiding obstacles in the roadway. However, in this game, the colors are vivid and realistic; the car is futuristic and portrayed in stunning and imaginative detail and large trees and meticulously rendered landscapes fill the screen.
This game has a more intensive interface as well. Where in the first game a single key press controlled the car’s movement, this game requires two buttons to play the game: a move left button and a move right button. Which game would the average user consider more “usable?” Would they judge it on which game is more “real” or more “enjoyable?”

To address these questions, usability analysts turn to the user’s behavior as well as their stated preferences. Surprisingly, these are not always the same. For example, the user’s behavior in game one would largely be a measure of reaction time. The game, boiled down to its essence, tests how quickly the user can respond. In the second game, though, the user must also make a decision for which button to press to avoid the obstacle. Because of this additional decision component, the user’s performance would suffer in the second game. This reduced performance would normally be a symptom of poorer usability. However, when asked, users would rate the second game as being preferable to the first. The truth being that the visual aesthetics of the game would outweigh any perceived loss of performance (if the user even noticed a decrease in their performance level). So what are usability analysts actually testing if not simple performance or preference? If the customer, in this case the user, is not always right and better performance is only “good enough”, how can usability consider itself a scientific approach and not just marketing with trendy vocabulary?

When it comes to flashy new technology and applications, the gaming industry is usually at the fore. However, while usability has become a buzzword for much technology, gaming has remained more cautious in its adoption of explicit standards. In fact, in the gaming industry it is an oft-debated topic, usually concerning if it should be considered at all. Usability is often seen by serious researchers and developers as the stepbrother who only shows up for money, the metaphorical black-sheep of the development cycle that nobody really wants around but it looks good for the family pictures. This arises from a single problem; there is no universal standard for usability. Usability is itself a ubiquitous term but it is usually coupled with multiple standards, variations, and definitions that are often not only industry-specific, but company-specific as well. One company may use validated measures like the Systems Usability Scale (SUS) or the Questionnaire of User Interface Satisfaction (QUIS) but many companies choose to develop their own standard for usability measurement, and sometimes their own definition of just what a “usable” product entails.

So, why is there so much disagreement? One reason is because there are no single, simple answers to just what the best measure of usability is. Two of the more popular measures, the SUS and the QUIS, have several advantages and disadvantages. Because these measures have been validated they give an accurate scale, but they may not have the ability to test what the developers care about. More specific tests that have been individually developed preserve the developer’s rights and may test specific aspects of interest but the accuracy of these measures is often dubious at best. When attempting to apply a usability test to a new product, these costs and benefits must be weighed against further constraints of resources and the time available to conduct the analysis. In the 24-hour development rush to create games, this is especially the case. Usability is often a last minute, rubber stamp process, more concerned with accessibility than usability professionals would prefer. Ideally, usability would be based off of an industry standard that can be applied universally to software development at all levels. For example, tests that that can measure the effectiveness of a webpage while still having the capabilities of measuring user satisfaction in gaming could increase the effectiveness of usability immensely by allowing comparisons currently impossible.

As a result of this, developers will spend thousands of hours developing a new game that may be visually pleasing, then turn around and ignore