The Thesis committee for Travis Wayne Givens

Certifies that this is the approved version of the follow thesis:

Uses and Consequences of Data Visualization and Analytic Tools in Online Games

Approved By

Supervising Committee:

Supervisor: ________________________________
Karin Gwinn Wilkins

____
Wenhong Chen

Uses and Consequences of Data Visualization and Analytic Tools in Online Games

By

Travis Wayne Givens, B.A.

Thesis

Presented to the Faculty of the Graduate School

of the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2012

Uses and Consequences of Data Visualization and Analytic Tools in Online Games

By

Travis Wayne Givens, MA

The University of Texas at Austin, 2012

Supervisor: Karin Gwinn Wilkins

This thesis examines the usage of and attitudes toward data visualization and analytic tools in three genres of online games. Using an online survey, this research analyzes responses from participants regarding their play habits and attitudes online. Several scales are generated identifying different player demographics such as emotional attitudes, competitive attitudes, technological attitudes, spectator involvement, and overall attitudes toward information customization. In addition, several genre specific scales are created for massive multiplayer online games (MMO), real time strategy (RTS) and first person shooting (FPS) games. This research concludes that competitive attitudes are moderately correlated with information customization and implementation of data visualization tools. Additionally, the relationship between the usage of data visualization tools are strongest with the MMO genre compared to the RTS or FPS genres. In addition, my research shows a strong preference between the responses for the usage of data visualization tools amongst those who report higher levels of spectator involvement with online games. Finally, my research concludes that there is a strong relationship between the amount of time players spend playing online games and the attitudes toward and usage of data visualization tools.

Table of Contents

List of Tables

List of Figures

Chapter One: Defining the Need

In October of 2011, over 3,000 industry professionals and researchers from many of the game industries leading and emerging corporations met in Austin at the game developer's conference online. One of the key discussion topics throughout the conference was the proper use of analytic tools in game development (Cummings 2011). The use of these tools in the past several years has grown in response to an increasingly demanding and fractured player base, combined with an explosion of small game development companies. Analytic tools offer a unique solution to the need for information generation, filtering, and analysis that accompanies the massive amounts of data that game developers are accumulating. Analytic tools can take many forms such as algorithms, machine learning techniques, and data visualization. In this research I discuss the specific usage of data visualization as a necessary component of visual analytics that players and developers are using within the context of online games.

Despite the increased use of metrics, formulated measurements of player habits and attributes, and analytics, the process of analyzing the massive amount of information available, for informing game design, there are still many unanswered questions and challenges facing game developers. Many of these questions focus around the types of metrics used in analysis. Other problems involve implementation and resistance against the use of analytics in design teams. Misused, analytics can serve to limit the variance in play styles in online games. Used effectively, analytics can be a tool that allows developers to cater to future player demands through innovative design.

The use of analytic tools is not limited to the design process for online games. Players have actively engaged in the use of analytic tools for their games for decades. Recent games, such as *World of Warcraft*, have brought the implementation of player created analytic tools to the forefront of player attention. These tools are created by the players and distributed via websites designed for their distribution to the players. Access to certain activities in these virtual worlds is dependent on having a working knowledge and functioning copies of many of these analytic tools.

Therefore, if designers are not the only ones using analytics in games then I am inspired to ask several questions about the use of analytics that affect both players and designers. Of primary importance is to figure out who is using these analytic tools and at what level. My first question asks:

R1. Who uses analytic tools?

This question is important to ask because it allows us to create categories of use that explore who is using what type of analytics at which level. Additionally, this information guides the uses of these tools. The myriad number of uses for analytic tools in games is partially informed by the people who are using them. Analytic tools vary in frequency and process of use. Therefore, as I examine who uses analytic tools I must also ask:

R2. How are the analytic tools used?

By asking how analytic tools are used I can identify patterns of use and begin to cluster patterns together to form narratives of use and understanding. In addition, this allows us to examine situations where analytics is not used and under what conditions the use of analytic tools is increased or decreased. There are several genres of games, each with specific interface conventions. Also, attitudes related to the technology may play an important role in determining frequency and process of use of analytic tools in games. My next research questions look to what conditions influence how analytic tools are used.

R3. How does the genre of a game influence how the analytic tools are used?

Finally, I investigate the consequences of the use of analytic tools in games. These measures include revenue metrics, such as retention rates and time spent playing games. Additionally, they include non-revenue metrics, such as happiness and diffusion of the game across a player's social network. This results in my final research question:

R4. What are the consequences of the use of analytic tools in online games?

This last question is especially useful for both players and developers seeking to understand how analytic tools can impact play in games. Hopefully, this will lead to closer relationships between players and developers, and to informed game development. For players, this research has the potential benefit of increasing awareness about the use of analytic tools in games and during the development process. Developers may use the data regarding the consequences to design games with the various factors of use in mind, and ultimately increase both revenue and non-revenue metrics for their games.

What follows is a literature review that positions this research by defining online games and visual analytics. I then explore previous research that examines analytic tools and their uses and consequences from a player and developer perspective. I show how this research fills a gap in the literature created by two distinct styles of methodologies.

Chapter Two: Review of Literature

Language of Online Games

For the purposes of this research I use the term "online game" to represent the class of games that can and are primarily played online through a computer. The interaction that happens online in these games offers a rich and complex social environment of study, which is often as meaningful or real as everyday interaction (Taylor, 2006, p. 19). I should also note that the term online game is intended to include the interaction present in many virtual worlds. Bartle and others argue that virtual worlds should be considered separate from games (Bartle, 2004, p. 475; Boellstorrf, 2008, p. 32-36). For Bartle, virtual worlds exist as spaces for interaction and should not be confused with the interaction itself much in the way that "the Rose Bowl is a stadium and not a football game"(Bartle, 2004, p. 475).

Following this logic does not necessitate that I disregard virtual worlds entirely. Instead, I should note that when I speak of online games, I include those actions within virtual worlds where the primary mode of interaction is game-like. Thus, I will not speak of Azeroth or Norrath, just as Boellstorrf makes clear to separate *Second Life* from games (Boellstorrf, 2008, p. 33). Boellstorrf concludes his section on the separation between virtual worlds and games by stating that despite the important difference, there are many aspects that are still relevant and can provide insight into online games (Boellstorrf, 2008, p. 36). In this vein, I consider events such as raiding, harvesting, crafting, exploring, and selling within these virtual worlds. My goal is not to focus solely on character progression and related activities as the primary interaction, but also on the many ways that analytics is utilized in the other types of emergent game play. This research does an excellent job of providing opportunities to explore the consequences of data visualization in online games by giving examples for some of the events that these tools can be used for. I will discuss more about this when I discuss Consalvo's work regarding other types of game play.

The activities within virtual worlds form a large portion of the focus of this research. In the last few years, researchers have followed the wave of people who regularly go into these virtual environments using several different types of methodologies and have come away with strong findings about their economics (Castronova, 2002), *theory-crafting* (Nardi, 2009; Glas, 2010), and social interactions (Moberly, 2010; Taylor, 2009). The number of participants in these virtual worlds is staggering compared to the player bases of many other types of online games, such as real time strategy (RTS) and the sub-genre called multiplayer online battle arenas (MOBA), or first-person shooting (FPS) games. All of these communities utilize their methods of collecting analytics about the games they play to come up with optimum strategies. The number of people who play these games is large and researched. However, much of the research performed only considers traits within games or concrete measures such as age or gender. Along these research lines using my survey I wish to inspect different hypotheses to understand who the people are that are playing these online games. These hypotheses attempt to quantify and measure my research questions into measurable outcomes that examine various aspects of data visualization tools in online games. The first of these is the hypothesis that:

H1. Those who are technologically and mathematically proficient are more likely to use and be aware of data visualization and analytic tools available in online games.

Often there is overlap with the types of games people play. Playing *World of Warcraft* does not preclude someone from playing *StarCraft II* or *Modern Warfare*. Exploring the differences in how people use analytics within these types of online games is not well understood. Whereas research has examined the uses of different types of data visualization already in the genre of MMOs, by observing these other genres I can compare the uses of these same data visualization tools and see if there are differences between the genres.

Just as there is a long and invested debate among the definitions of game and online game, the path to defining analytics is equally murky. Traditionally, analytics has been used in fields such as economics or business, which lend themselves readily to mathematical modeling and computation (Bator, 1957; Hirshleifer, 1988). In this view, analytics is very closely aligned with mathematical computation of data. In fields such as data visualization analytics has a slightly different definition:

> Visual analytics combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets. The goal of visual analytics is the creation of tools and techniques to enable people to … synthesize information and derive insight from massive, dynamic, ambiguous, and conflicting data… (Keim, et al., 2007).

This definition has several points that require emphasis. This definition of visual analytics has three primary components: (1) a data set that can be massive and have multiple conflicting data points; (2) tools and techniques that form the automated analysis combined with data visualization; and (3) an understanding of the data and a proscribed plan of action.

Data visualization is an important and necessary component of visual analytics. Proper data visualization is the key necessary to provide successful visual analytics communicated to players. Therefore, it is necessary for us to explore data visualization and its role in providing visual analytics that are used by players. First, I explore the definition for data visualization. Following my definition, I examine various examples of how data visualization exists within many popular online games. Following this, I will look at why data visualization is so useful within the context of online games by incorporating elements of data visualization and cognition theory.

The practice of data visualization has existed for some time primarily through data maps, which combine cartographic skill and statistical representation (Tufte, 1997, p. 20). These first appeared in western civilizations during the 17^{th} century with examples such as Edmond Halley’s map showing trade winds and monsoons on a world map (Tufte, 1997, p. 23). Perhaps the most famous example of this type of data visualization is the *Carte Figurative* by Minard, which shows the fate of Napolean’s army as he traveled to and back from Moscow (Tufte, 1997, p.40).

Data visualization in the context of analysis is a process in which the inputs are information in various forms such as issues, data gathered through exploration, or hypothesis, scenarios and models, and the outputs are a discourse between the analyst and their information (Thomas and Cook, 2005, p. 35). In the last few decades the importance of effective data visualization has increased dramatically not only for presenting, but also for filtering and understanding the vast amounts of data that has become available since the creation of the computer. The problems of data overload have been examined by several studies ranging from pilot cockpits (Woods, Patterson, & Roth, 1998) to documents related to homeland security (Thomas and Cook, 2005). The common thread behind these studies is the use of data visualization to increase the ability to make quick decisions. The ability to make quick and accurate information is tantamount to any player attempting to win a game against other players. If data visualization is a useful channel for providing information then I expect there to be some relationship between the use of data visualization and modern competitive play.

Our hypothesis specifically examines the first of these possibilities:

H2. Competitive players are more likely to use data visualization and analytic tools.

Data visualization is a major component of nearly any modern digital game. One of the earliest examples of data visualization that appeared in online games is the scoreboard (See *Figure 1*). The scoreboard generally contains information about the current game to the players and includes a list of players and information about their performance such as who is alive, the number of kills versus deaths they have achieved, connection information, and other general statistics.

Figure 1 (Above): A typical scoreboard from Team Fortress 2

Figure 2 (Below): A mini-map from World of Warcraft

Another visualization tool just as ubiquitous in online games as the scoreboard is the mini-map (See *Figure 2*). The popularity of the mini-map in online games fills the need for examining data on multiple levels within online games. In his examination of 'fish-eye' visualization, Furnas (1986) explores the cognitive tendency of people to represent their immediate surroundings in great detail but to only reference landmarks further away from their immediate neighborhoods. *Figure 2* displays the mini-map that is used to display information about the general surroundings where the yellow dots represent useful resources to be harvested and exclamation marks display a potential quest of importance to the player. These landmarks act as signals that allow players to make decisions with data that would otherwise not be readily available to them.

This last point touches on how data visualization and analytics are used in online games. As Thomas and Cook (2005) suggest, the end product of data visualization is a discourse between the analyst and their information (p. 38). From a functional standpoint, in online games, this means that data visualization is used to help a player make decisions about what they should or need to be doing. From a cognitive perspective every decision players make has an embedded cost and cost structure (Thomas and Cook, 2005, p. 45). In online games this takes the form of time required to see an input such as an enemy unit, comprehend the information, decide on a course of action, and execute the planned response. Data visualization can reduce the cost structure of this process in two primary ways: by offloading some of the cognitive process to easier mechanisms; and by allowing software to perform the filtering, reasoning, and representation (Thomas and Cook, 2005, p. 46). The mini-map is an example of a software tool that performs the latter of these two functions very well. The tool takes in all of the information regarding the larger surroundings of the player and filters down to only those vitally important details that are represented to the player.

In online games, players sometimes compete against computers but are more likely to be playing against one another. The purpose of this data visualization in online games likewise is to help players make decisions toward the ultimate goal of winning the game. Players are willing to seek out tools that allow them to beat games and this is one reason that data visualization pairs so well with online games. As players become more competitive and capable in online games, data visualization tools such as guides or walkthroughs with higher cognitive costs give way to other quicker tools such as mini-maps. Similarly, once players have achieved mastery in a game, the need for cognitive tools such as data visualization is smaller as they react and base their decisions more on experiences and instinct, a concept widely popularized by Csikszentmihalyi (1989) as flow.

To give a short example of this type of transition, I provide an example of a player who is just trying out the Real Time Strategy (RTS) game *Starcraft II* for the first time. After toiling around and learning the basic functionally of the game the player goes online and reads some strategies about how to play the multiplayer and what sorts of techniques are most useful. This player starts utilizing these strategies and one strategy in particular is known as scouting. Scouting is the process of sending a unit to look at an opponent's base so that the scouting player can see what strategy their opponent is using and counter it. In my example, this player rigorously scouts every game and begins to identify trends in opponents' play. After several months or years the player no longer feels the need to scout, because he instinctively knows when he can be attacked and with what types of units based on his many games of experience.

Our player in this example is able to progress through a process of using data visualization tools that require less cost to make decisions as their experience grows. Certainly however, a large portion of people who play games do not have the time, patience, or ability to become a true master of a game to the point that they can rely on experience and instinct. Many of the people who play online games do so at a level below the expert tier. Thomas and Cook (2005) illustrate a key property of data visualization is that it is designed to accessible and useful to those below the expert tier (p. 47). This makes data visualization extremely useful for new players of a game and offers incentive for players to learn, play, and stay with a particular online game. In this way data visualization allows players to make decisions quickly given their information in order to win and be competitive. Before, I discussed that some positive relationship should exist between competitive play and the use of data visualization tools. In comparison, we should also expect some type of relationship for those who are in this top tier of competent players that might be different from the general population.

In addition to measuring the relationship between competitiveness and the use of data visualization I also want to compare success and competency in terms of online games and data visualization. Since data visualization is useful as a tool to help analyze player situations I expect that using data visualization in games would be positively correlated with winning games and feelings of competency. Therefore, my hypothesis is:

H3. Use of data visualization and analytic tools is positively correlated with feelings of competency in online games.

Another important component of consuming games is not only playing, but also watching other people play games. Most research in this area is couched in cultures of Asia, such as Chee's (2006) examination into *StarCraft* in South Korea or Nardi's (2009) ethnographic work on *World of Warcraft* in China. This research is ethnographic and involves trips to cybercafés and personal interviews with players and fans of *StarCraft II* in South Korea and elsewhere (Chee 2006). Cheung and Heung (2011) used textual analysis of over 100 online blog entries, forum postings, and news article to create nine different audience phenotypes. These studies interpret games within a larger cultural context but never examine the act of watching broadcasts as individuals within an economic or analytic perspective. However, this does provide a useful avenue for researching the 'who' of online games that can translate to this research.

Additionally, since these studies were produced the game broadcast industry has experienced a renaissance throughout the world. In 2007 the only place to find high quality broadcasts between competitive players of online games was to tune in or watch a recorded video of one of two South Korean broadcast stations twice a week streaming *StarCraft*. Now at any given moment there are hundreds of streams of players who are streaming high quality content of their games and tournaments to an audience that regularly numbers above thirty thousand. For some major events, the number of viewers can be even higher. At the conclusion of Riot Game's Season One Tournament for their titular game, *League of Legends*, over 215,000 concurrent connections were recorded for the finals matches (Riot Games, 2011). These large numbers of viewers allow broadcasters to make money by running advertisements between matches creating a legitimate source of revenue that can fund million dollar prize pools. For the teams that win, the real money prizes supplement player incomes and the organizations that sponsor the players by paying for plane rides and trips to other countries for more tournaments.

There are two major types of streaming content at the moment: tournaments, which include daily and non-daily tournaments, and streams from a first-person perspective, which allow the audience to see all of the moves that players make. Daily tournaments and non-daily tournaments are very similar except for their broadcast schedules. Daily tournaments are generally performed online only and stream three to five times a week. Non-daily tournaments are often larger in scope in terms of having a physical location where players compete on a stage in front of a live audience. First person streams can vary wildly in their schedule depending on how often a player is practicing. Many games include a non-tournament based ladder ranking system and many streams are these types of games where there is no prize pool, but offer an opportunity for players to practice skills against competent opponents.

Game developers are beginning to incorporate tools within games for the specific purpose of tapping in to these new ancillary markets for their games. For example, the game developer, Blizzard Entertainment, patched into their game, *StarCraft II*, tools and overviews that are only accessible to spectators of a match or through replay. These tools included various meters and overviews that allowed for a more thorough breakdown of each match to occur. Other developers have followed suit. Riot Games, creators of *League of Legends,* has similar tools available only to spectators that display a graph of how much gold each team has which is the most important indicator of who is currently winning a match. Valve has taken this concept one step further with their recent project, *Dota2*, by incorporating a spectate feature that allows players not only to play games against public players, but also to search for matches of other players and spectate them easily. Within this spectate mode, players have the ability to examine graphs of gold (see *Figure 3*) and experience that in many cases illustrate the flow of the game through these data visualization tools.

Figure 3: An example of a graph available to spectators in Dota 2

I have to consider the role of data visualizations for spectators watching matches of online games. As developers move in to additional markets such as e-sports their exposure and dependency on providing positive user experiences to spectators the need for providing tools and quick updates to broadcasters and players about the game is becoming more and more important. My hypothesis regarding the connection between broadcasting and data visualization is the following:

H4. Spectators are positively affected by charts and graphs to help explain situations.

In many ways, the rise of these data visualization tools is similar to the incorporation of technologies in other media, such as the telestrator on television. As opposed to the player who uses data visualization tools in games to inform decisions toward the goal of winning, these data visualization tools for spectators serve the purpose of quickly and accurately understanding the match they are watching. During a broadcast of a major tournament, spectators must observe the game through the lens of a dedicated commentary team. The job of these commenters is to direct the attention to the exciting parts of each match and generally this follows kills, deaths, and other exciting engagements between players. However, in all of these games there is an element of play that exists, which is not as flashy and therefore may not be highlighted by broadcasters. In situations like these, broadcasters may not allow viewers to see these less flashy moments but thanks to data visualization tools the impact of these other play elements can be seen and understood by the audience in a succinct manner that keeps the pace of the broadcast quick and exciting.

In addition to the data visualization tools available to players, broadcasters and spectators may have access to more visualization tools. *Figure 4* displays various analytic tools that take the form of real time metrics such as bar charts that display current actions per minute (apm) by players or production values, while *figure 3* displays the line graphs that can display the difference in total experience and gold between two teams, as well as scoreboards with more quantitative information on all players at once. Broadcasters have a wide array of visual tools at their disposal for helping spectators understand what is occurring in the game. While broadcasters have started using these tools, the industry takes it for granted that people use and appreciate these tools. To date, no extensive research has been performed that examines the attitudes of spectators toward the use of data visualization tools in broadcasted online games.

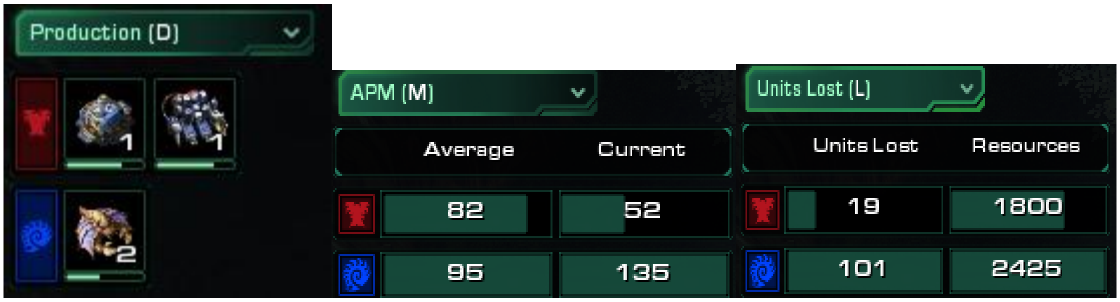

Figure 4: Production, APM, and Units Lost Tabs available to spectators in Starcraft II

While visual analytics has been around in some form for hundreds of years, with the advent and widespread use of computers, the field has experienced a renaissance in the number of tools and techniques and their focus. In game studies specifically there are two styles of research that are defined primarily by the populations of use and their available resources. The first of these styles is low technology in tools and technique but widespread in use. This includes what Consalvo (2007) considers paratexts: game guides, walkthroughs and FAQs, as well as many tools created by players in virtual worlds and techniques for recording data through theory-crafting (Nardi 2009). The second style of visual analytics used in game studies is truer to the original mathematical foundations of analytics. This style is heavy on the technology requirements and low in terms of populations of use. This style has been used by researchers in partnership with the game industry to analyze and create models for activities such as detecting play patterns (Weber et. al., 2011), cheaters (Ahmed et. al, 2009), and economic behaviors (Castranova et. al., 2009). I will discuss both of these styles and how they apply to online games more in depth in the next few sections.

Analytics of Players

As mentioned previously, the first style of analytics research uses methods that are low in technology to examine play habits and attitudes widely in use among players. These methodologies include ethnographies, participant observation, interview and textual analysis. In the next section I examine walkthroughs, game guides and FAQs for games, known as paratexts, as an example of textual analysis research on analytics used by players.

Walkthroughs, game guides and FAQs have been around for some time. All three act on the same basic principle, with one person who has played through a game sharing knowledge about how the game is played. A game guide is traditionally released by a third party publisher, commissioned by the original game developer to act as a supporting tool for the players of the game. The contents of these game guides often include sections on controls, characters, and a specific scenario by scenario how-to guide that goes through every detail in depth.

A walkthrough is similar to a game guide, but there are several distinctions between the two. First, though these are written by one author, walkthroughs are collaborative works derived from the efforts of a player community. Second, walkthroughs typically only engage the audience through a particular play through style. In some cases, several walkthroughs for the same game may be posted next to each other with each identifying a separate play style such as “quick run-through” or “maximum items.” Other walkthroughs may engage only in describing particular events in the game such as boss battles. Finally, walkthroughs are written in a separate style convention from game guides. This includes the use of text characters, fixed widths, and ASCII diagrams. Both game guides and walkthroughs can be quite lengthy, easily surpassing 100 pages for longer games.

FAQs exist as a separate convention from either the walkthrough or the game guide. FAQs are much shorter than their counterparts and utilize the format known traditionally as the question and answer. FAQs are still asked mostly in relation to specific events within games or to get around a particular encounter.

As varied as these three styles are in their production and consumption, they all share qualities that indicate they are a type of visual tool of analytics. The first element of conceptualizing analytics is met by considering the game as the conflicting data set. The data set becomes the decisions in the game that the player makes in response to specific stimuli. Often games will have multiple ways of completing objectives that may lie in conflict with each other. The second and third elements I can argue together.

In his 2010 book, *Players of Stake*, Rene Glas devotes several sections to the way that players in *World of Warcraft* behave outside of the intended design boundary using what he calls walkthroughs or leveling guides. In one example, Glas (2010) participates in a community of players who created guides to creating a "*twink"* character for PvP (player versus player) battlegrounds (p. 124). By following the guide, Glas was led through a series of decisions and shown combinations of items would present the best opportunities for victory (p. 128 - 133). In this example paratexts operate as tools that cull available data and produce a prescriptive data set. The possible items available to be worn act as the set representing all of the item possibilities, while the guide assesses the value of each of the items and chooses the items that work best together for the future. While this information is well understood in ethnographies of *World of Warcraft* and other MMORPGs the use of these paratexts is not as understood in the context of real time strategy and first person shooting games.

Glas (2010) and others (Consalvo, 2007; Nardi, 2009) have recognized the extensive usage of these guides for *World of Warcraft* and the impact they have on the social dynamics within the game. The widespread nature and practice of these game-play activities, such as twinking, eventually led Blizzard, the designers of *World of Warcraft*, to implement specific features that facilitated these game-play activities and the tools that supported them (Glas, 2010, p. 133-137). Often communities come together to create these tools and guides. For example, in her research into *World of Warcraft*, Nardi (2009) interviewed players who were participating in the practice known as theory-crafting.

Theory-crafting is the practice whereby the players of a game attempt to create, or piece together the underlying algorithms behind a game. For some players theory-crafting offers a chance to understand the rules behind the game that the developers do not wish to divulge (Nardi, 2009, p. 137-151). Theory-crafting in practice is typically the result of simple experiments with vast numbers of data points. These data points are then decoded and analyzed to discover patterns and quantifiable measures of improvement to determine the answers to various questions about the game. Many times these results are posted to forums and guides and are transmitted to players seeking specific information about a question (Nardi, 2009, p. 141; Glas, 2010, p. 120). More important however, is that the results behind this theory-crafting form the basis for many of the analytic tools used within online games. For example, players can perform pseudo-experiments to determine exactly how much damage it takes for players to obtain the attention of the monster (Chen, 2010) through repeated trials and scenarios.

Within a virtual world as old as *World of Warcraft*, the presence of these tools is ubiquitous and necessary to participate in many of the in-world activities (Taylor, 2009). Raiding in *World of Warcraft*, one of many game-play activities, often requires a dizzying array of player designed add-ons and real-time analytics tools in addition to the basic game. The tools record every event that happens within the raid and report to players vast amounts of data. Some of the examples that are reported on include how much damage or healing a player does, whether or not they are at their screen and ready to participate in the battle, or an avatar’s status (Taylor, 2009).

The advantages to this type of knowledge include an increased ability to coordinate with others and the ability to account for other players when interacting towards a common goal. Additionally, for many players who prefer alternative play activities within virtual worlds, these tools can allow such activities to occur. For example, a person may use a tool that analyzes the economic information within a virtual world to gauge whether or not certain items are undervalued in the market in order to profit. Researchers (Glas, 2010, p. 32; Nardi, 2009; Chen, 2010) have identified these activities through participant observation and interviews. While there has been extensive research into these activities in virtual environments there is a critical lack of research observing the extent of these analytic tools by players of other online games. My hypothesis highlights the difference between these genres and the incorporation and expectation of visual analytic tools within each genre.

> **H5.** There exists a difference between genres of online games and the preference for the information displayed in the UI.

While there are benefits to the use of analytic tools within online games, there are also possible negative social ramifications to their use. These analytic tools act in a manner very similar to surveillance technologies and in doing so have the potential for being good and bad (Lyon, 2001). Citing the use of these analytic tools, Moberly (2010) observes players:

> Required to work together in teams ranging in size from two to forty, players routinely critique each other's performance, often using in-game tools and add-ons to measure damage and healing done, killing blows, deaths, and other variables such as bases and flags captured (p. 232).

This observation is echoed by Chen (2010), who using actor network theory through participant observation, argues that analytic tools such as raid meters transform the network of trust between players (p. 179). Prior to the widespread use of the tools skill as measured by success was inferred. While a player may not have been performing as well as another, it was hard to detect as long as the group was succeeding overall. However, Chen found following the implementation of the analytic tools, a player's skill was quantifiable and eventually the tools became the authority on player skill; effectively usurping the trust network (Chen, 2010, p.179).

Moberly (2010) argues further that analytics are used as tools of domination whereby players are reduced to a commodity and their self-value is measured in terms of abstract units of measure and minimizing deviation from the “play styles the community as a whole endorses as the most effective” (p. 232). Viewed in this paradigm, analytics act as a tool that also helps bounds the community of players by helping individuals rank, judge, and stratify each other. If you are part of the community, then you abide by the best practices of play styles that are indicated by the in-game tools and add-ons. This has ramifications in the use of relationship between the use of data visualization and emotional attitudes of online games. I will return to this relationship later in the paper.

The methods used by these researchers all focus on the individual as part of the larger system and utilize methods that emphasis the participant either through participant observation, interviews or through textual analysis. Several of these studies specifically mention that they were unable to offer a complete picture of these virtual environments, because the staffs of the businesses that design and maintain the online games were unwilling to discuss their work with the researchers (Glas 2010 see also: Nardi 2009, and Chen 2010). This limitation on the design of these studies makes it so only the player perspective and not that of the designers of a game is examined. Finally, many of these studies are performed in virtual worlds, specifically *World of Warcraft*. The research design of these studies in virtual worlds limits the applicability of the findings to only *World of Warcraft*. Certainly, there are other online games and while the interaction in other types of online games may not be as diverse, they still represent a serious time commitment as well as a site for social interaction by the people who play them.

However, research into these online games has been largely limited to virtual world environments. In addition, these ethnographies are limited in scope to the observations of a single researcher and not generalizable to larger populations.

Predicting Players

The problem of making decisions based on the attributes of entire populations of players is a recent focus of developer attention. Developers are using data visualization and as analytic tools to develop insight into the players of their games. In this context, analytics has begun to become synonymous with information gathering, modeling, and prediction performed primarily on computers or through cloud computing (Iosup, 2009). This type of analytic computing is performed primarily by the businesses that run the platforms themselves. This style of analytics in online games is considered high technology and is closely aligned with the fields of computer science and human computer interaction. Running this style of analytics in online games can have several goals from understanding play patterns (Tychson, 2008), detecting cheaters (Ahmad, 2009), to understanding emergent and economic behavior (Castranova, 2009).

High technology analytics has been made possible in the last decade through advances in computing power, storage, and accessibility. Games, and especially online games, have not always been able to record player data. Some of the earliest methods of record keeping in games were limited to high scores and initials. Eventually this practice evolved into concurrent, progressive and longitudinal measures (Medler, 2009). These measurements have been implemented in game design as replays, scoreboards, ghosts, and achievement systems. The ultimate goal of these features is to provide tools and rewards for the players and to keep them playing. The question of whether or not these data visualization tools actually increase the time that players spend in these games is never addressed. These questions manifest in my research in terms of hypotheses such as:

H6. The amount of money spent on online games is positively correlated with the use of data visualization tools in online games.

One of the fundamental challenges of collecting the data in online games that has to be overcome is the storage of all of the data that is generated on players. Depending on the type of game and the scope of recording, the amount of data generated each month can be as small as gigabytes or as large as terabytes of data (Williams, 2010). Research using virtual worlds often has data size issues on the higher end of the scale where often every action and movement is recorded. This is one reason that virtual worlds are frequently examined by researchers; the possibility of using a quantified measure for every interaction is a treasure horde if the data can be used properly. For example, quantitative measures were used to measure guild and group life cycle tendencies in *World of Warcraft* (Thurau, 2010).

Other drawbacks to this style of analytics focus on accessibility and computing power. Games can keep massive amounts of data on their players and in order to analyze this data properly requires supercomputing or cloud computing. Several techniques have been developed for generating responsive databases such as robust unique effect analysis (RUnEA) (Weber, 2011b), online analytic processing (OLAP) (Medler, 2011), as well as heat maps, social network analysis and machine learning. Many of these techniques require proficiency in languages such as MSSQL or Java to program, which creates a barrier to accessibility. Additionally, businesses are not often privy to give away game data regarding players for confidential and competitive reasons. This trend may be slowly changing as game development corporations are slowly integrating analytic analysis into their design framework. Finally, a major hurdle is the margin of error on these algorithms that attempt to predict human behaviors. Despite progress, detecting cheaters in a game is still a very difficult task (Ahmad, 2009). Combined, all of these factors combine to make this style of analytics extremely difficult for the typical player to use or to create.

Despite the disadvantages to performing this style of analytics in online games there are a number of advantages. The setting for this style of research is not limited to mostly virtual environments that are seen with the ethnographic research. This is due to the similarity of the goals of the game developers across genres of game type. The setting for this type of analysis can be more variable depending on the metrics the business is interested in using. For example, in a study examining players of the Madden NFL game, researchers quantified the impact of several different game modes and play options to understand how they impacted overall play time (Weber, 2011a). Other research into self-generating Mario platform also uses data to maximize the time the player spends in the game (Weber, 2011b). This goal of maximizing the time that players spend in games is a common thread for this style of research. Therefore, my hypothesis is presented as:

H7. The amount of time spent playing per week is positively correlated with the amount of data visualization tools players use.

Other positive effects of utilizing analytics that may influence game design include identifying and quantifying player motivations and goals in online games (Yee, 2006a). Originally, player motivations were viewed through the paradigm of Bartle's four player categories (Bartle, 1996). However, this number has grown to produce as many as ten currently accepted different motivations for players in an online game (Yee, 2006a). Analytics is used to parse through the different activities that players spend time on in order to discover what activities and how long they spend at each event. Additionally, by mapping player trends in events they can detect outliers or machine-like event repetition and make specific recommendations regarding those players, or bots, and their play styles (Ahmad, 2009).

One of the key advantages to this style of research is that it allows us to examine and say definitively what actions people are taking inside online games. Unfortunately, these methods do not consider the attitudes of the player toward the advances made in the name of increased tools. While they are effective at aggregating entire populations of players together they suffer at the level of the individual player. Additionally, this style of research cannot answer why people are doing the actions in these online games. This is the key distinction between the two styles of game analytics. Additionally, the benefit to these studies is largely for the designers of the online games themselves. The only benefit to the players is seen through the next iteration of the game design and not through any direct interaction with the high technology analytics.

Chapter Three: Methods

The current study focuses on the impact and use of data visualization and analytic tools in online games by players. Most of the previous research in this area occurs in two methodological categories. The first is ethnographic review, often incorporating on-site interviews with players in China, South Korea, and the United States (Nardi, 2009; Chen, 2010). The second category utilizes survey methodology to reach large numbers of participants and make assessments on the players based on self-reporting responses.

Ethnographic studies tend to incorporate vast amounts of play time by the authors to create a contextual basis for their work (Nardi, 2009; Glas, 2010). In these studies, all conversations are recorded and recoded into categories and used to examine various aspects of these games. The online game that received the most attention and study from these studies is *World of Warcraft*.

Perhaps the best known research involving surveys and online games has been performed by Yee (2006b; 2006a) and others (Williams, 2008). The number of responses they received in each round of their studies generally ranged from 2,000 – 4,000 participants and for one study they received roughly 7,000 participants (Williams 2008). This study in *EverQuest 2* was unique because it was a stratified sample due to the partnership with Sony, the developers behind the game. Most other studies represent convenience samples drawn from populations of visitors to online gaming sites that cater to specific games.

Our research methodology utilizes survey methods to quantify the impact data visualization and analytic tools have in online games. My research is focused around four primary research questions: i) what types of players use these data visualization and analytic tools; ii) how are these tools used by these players; iii) how does the genre of game and attitude impact how the tools are used; iv) what consequences for game preference exist because of these data visualization and analytic tools.

The first hypothesis looks at the interaction between knowledge outside of the context of online games and the interaction within those online games in order to understand who is using data visualization tools. Mathematical and technological proficiency are key skills in proper use of data visualization tools such as models, graphs, and scenarios that are prevalent in current online games. By looking at the correlation between these skills and the use of data visualization I can begin to understand some of the reasons why data visualization is not used or recognized in online games.

A second factor that impacts gameplay is the level of competitiveness in the player. My second hypothesis connects the level of competitiveness with the use of analytic tools. One possibility is that competitiveness and the desire to win correlates with other measures and results in the use of data visualization and analytic tools as one of many to help a competitive player succeed in online games. A second possibility is that as players become more competitive and gain experience with games they tend to rely less on analytic tools and more on experience and instinct to make decisions in online matches.

In order to measure competency for the third hypothesis I use the self-assessment standard of placement in games and how confident players feel they are ranked in their games. While I cannot guarantee that players are giving an accurate representation, I feel that I can get a sense of general competency from self-reporting. Similarly, the fourth hypothesis uses self-reporting to understand the relationship between those who watch broadcasts of online games and the role that data visualization plays in their attitudes towards spectating these matches.

The first four hypotheses exist within the research question examining who uses data visualization tools. With these four hypotheses I can see the impact of different player attitudes on the use of data visualization tools. Additionally, these four hypotheses give insight into my second research question that asks how people are using these tools. If for example, only the competitive players are using the data visualization tools then I can understand some connection between competitive play and these tools. Likewise, if a play style of spectating is also heavily correlated with the use of these tools then I can examine use of these analytic tools in that context.

Finally, I need to consider the difference between genres of online games. Not all online games are developed and play the same or include the same data visualization and analytic tools. To take these differences into account I ask similar questions about play style relative to each genre of game. Specifically, for questions about MMOs such as *World of Warcraft* and *Star Wars: The Old Republic* I included questions about experience gain, gear maximization, and the use of tools such as mini-maps, map overlays, and user interfaces. For questions related to real time strategy games such as *Starcraft II* and others I incorporated elements most closely associated with real time strategy and similar games such as mini-maps, scoreboards and post-game metrics.

Finally, for first person shooting games I consider the data visualization tools most prevalent in first person games such as scoreboards and mini-maps.

Each section about a genre also includes a question for comparison between genres. This question asks each respondent whether they are influenced in their game preference based on how much information is displayed in the user interface of each game. Throughout measuring all of these hypotheses two separate analysis are presented: a combined measure that observes the tendency to utilize data visualization tools in online games in general and a unique measure for each genre measured by the survey. This is used to test my third research questions, which examines the relationship between genre and the use of data visualization tools.

The fourth research question asks what consequences exist with the use of data visualization tools. To test this research question I incorporate questions about the various time, financial, and emotional commitments that players make with the game.

The idea behind this hypothesis is that players who use data visualization tools are likely to enjoy their games more and would be tempted to buy more games or spend money on games they already play. However, consequence is interpreted to represent a time commitment to a game as well. From a game design position, player retention is important to maximize when so many options exist for games to play. If effective data visualization and analytic tools provide value to players then all things equal they play games more that provide more data visualization tools and to play more in general.

Instrument

This research relies on survey tools to record data about data visualization and analytic tools in online games. The survey was generated online through the use of a third party free survey site, ZipSurvey. ZipSurvey offered a one month free trial to host a survey and allowed me to input textual questions as well as instructions for the survey and a link to disseminate the survey to others. The survey implement consists of sixty-four questions designed to ascertain user profiles and the extent to which players rely and make decisions based on data visualization in various types of online games. Approximately ten minutes is needed to answer every single question; however, not all respondents had to answer all questions. Prior to release the survey was pretested and changes were made to the survey to make it more robust and to remove as much potential bias in the questions as possible.

The survey consists of three clusters of questions. As mentioned in the research questions the goal of the first section is to generate player profiles for different play types. Determining different player types and profiles has been a major effort in game studies since Bartle (1996) originally came up with the four player tropes of achievers, explorers, socializers, and killers. Since his original look at MUDs in 1996, these player types have becoming increasingly complex and robust. In Yee's (2006b) more recent work, they identified nearly four subgroups within each original category. However, these models all included information collected solely about activity that occurred within the context of game play and did not include information about a player's outside proficiencies.

Our survey attempts to test for player types related to both events within the play of online games and external factors. Most of these questions consisted of either five-point likert-scale questions from strongly disagree to strongly agree with a neutral possibility as well as a not applicable answer. In order to test the hypothesis 1 that a higher level of comfort with technology applications and mathematics is a factor in the use of data visualization tools questions I included questions to create a scale around mathematical and computer proficiency. To measure whether the level of competitiveness affects the use of data visualization tools I included questions related to player experiences and attitudes related to competing against others in online games. Similarly, I incorporated questions that self-assess player competency in online games and whether that impacts the use of data visualization tools. In order to accommodate players whose primary interaction with online games is through the spectator role I also included several questions related to gauge the level of spectating and involvement in watching broadcasts that occurs with players. Finally, in an attempt to provide external validity to my survey I also included questions from Yee's 2006 survey that allow us to compare responses related to questions about mental health in games.

The second cluster of questions uses open-ended formats asking how often players play games, how much they spend on them, and to name the last three online games they played. Due to the limitations of the online survey site, this last question was only available as a generic text box allowing respondents to fill in the response with more than three answers. Only one respondent included more than three games and in this case I only included the first three games they listed. These responses were separated into three different variables representing the order they were listed from most recent going backward in time. These responses were then coded by which genre of game they belong to in order to allow for analysis based on games within genre.

The final cluster of questions are associated with the research questions associated with the genre of online game that respondents answered they play. This allows us to compare responses within genre and discover trends related to data visualization and tools between genres of games. Specifically this research focuses on three genres of online games: virtual worlds, first person shooter, and real time strategy games. These three genres represent 20 percent, 10 percent, and 34 percent of the total units sold in 2010 respectively (ESA, 2011). This allows us to see whether players use these visual analytic tools different by genre. These questions were all five-point likert-scale questions with a not applicable option available. Not applicable was coded as a missing variable during analysis.

Recruitment

Recruitment for this survey was performed online. I used a non-random convenience sample method. Due to this, I will not report on statistical significance numbers for my correlations because they are not appropriate measures given my sample. This introduces a limitation on my research that limits the results of my survey to the population of respondents specifically. The initial call for participants occurred on a Friday afternoon in the hope that respondents would be more likely to take a survey during downtime either at work or after school but before prime gaming hours in the evening and weekend. Several sites were targeted specifically for their large influence in the e-sports scene. The first of these sites is TeamLiquid.net, a site devoted to StarCraft, StarCraft II, and several MOBA games where users keep each other updated on tournaments related to their games.   This site has approximately 10,000 active members. The second site I recruited participants from is a bulletin board like website known as Reddit. Users on Reddit generate content for each other in several sub-forums devoted to specific topics. I posted the call for recruitment in a sub-forum titled, games, whose missions statement declares, "the goal of /r/Games is to provide a place for informative and interesting gaming content and discussions" (Reddit, 2012). This community site has over 80,000 subscribed readers and boasts a history of topics related to games and development of online games.

A second call for participants was made at the beginning of the next week in order to garner a larger pool. Additionally, several responses were generated when the survey was circulated amongst a game company located in Austin, Texas. A copy of the call for participation can be found in the appendix. The survey closed at the end of the following week for a total of 10 days of responses at the beginning of March 2012.

In each call for participation there is a link for the survey from the third party site. This site records each visit to the survey to gauge the number of respondents who viewed the survey in order to receive an accurate completion rating. There were a total of 115 views of the survey cover page. The cover page contained standard IRB information as well as information regarding the four main research questions. In both the recruitment posting and the IRB cover page respondents were guaranteed that no identifying information would be collected other than age. Of the 115 views, there were

71 recorded attempts to start the survey. Of these, 11 did not fill out questions beyond the first cluster. This led to 60 complete survey responses for a 52 percent response rate to the survey.

All responses were first cleaned in excel. This included changing responses from string character type to numeric. Additionally, two questions were offered as select all that apply. Due to the unique design of the third party hosting site, each possibility was given a variable and not checking was represented as a missing data point. If respondents checked any responses for that question, the remaining missing ones were recoded as 0 instead of missing. At this stage 11 incomplete responses were culled from the study. After cleaning the data in Excel, the survey responses were imported in to SPSS for analysis.

Creating Scales

In addition to the basic demographic questions I designed the survey with the intention of being able to create scales that combine questions in the survey that fall in to thematic categories. To create these scales I took the questions from the first cluster and performed a traditional factor analysis on the results of the surveys. The first iteration of this process yielded a group of questions whose eigenvalue was over four and accounted for 19 percent of the variance. Seven questions were identified from this group as thematically compatible and part of the same component according to the factor analysis. These seven questions are listed in *table 1*. A reliability analysis of these seven questions yielded an alpha value of .741, which is an acceptable value for combining these responses in order to create the scale based around customizing the information available in game.

These questions share the common thread of the user seeking to find more information about what they are watching and processing that information. One question asks respondents to consider their comfort level when understanding charts and another questions attitude when broadcasters use charts or graphs to explain what happens in a game. Additionally, the other questions directly question the amount of information preferred by respondents in their user interfaces. Another important component of this scale is the attitude toward manipulation of the information. Three of the questions examine attitudes related to the ability to manipulate the data available in the game: (i) games that do not feature a customizable user interface frustrate me; (ii) I like online games that allow add-ons; and (iii) I customize my user interface. All together these questions form a scale around attitudes towards understanding, having more, and having the ability to manipulate information in their games.

After creating this scale I re-ran the factor analysis excluding the seven questions that existed in the previous scale. This new analysis yielded a component with an eigenvalue above 3 that represented 22 percent of the variance. This component consisted of five questions aimed at measuring the level of competitiveness in the respondents. These questions asked the respondent to rate their attitude towards playing against other players, how competent their performance was, and how often they played against other people. The reliability analysis of these five questions resulted in an alpha value of .727, which places it within acceptable levels.

Repeating this process again I found two further components with eigenvalues of 2.5 and 2 whose components accounted for 21 percent and 18 percent of the variance respectively. The first of these two a scale composed of three questions related to respondents' self-reported comfort with technology. These questions ask if the respondents do well with computer technology, are literate in computer programs, and whether they report being interested in learning new computer technologies. These three questions have a reliability alpha of .710.

Scale	Questions	alpha	N
Spectator	1. I have watched a stream for a competitive tournament 2. I have attended a competitive tournament for an online game as a spectator 3. I have watched other people stream games of them playing online 4. I have watched broadcasts of people commenting during online games	0.799	57
Information	1. I prefer to have as much information as possible in the user interface 2. I am comfortable understanding charts 3. I prefer to have less information in my user interface 4. Games that do not feature a customizable user interface frustrate me 5. I like it when broadcasters use a chart or graph to help explain what is happening in the game 6. I like online games that allow add-ons 7. I customize my user interface	0.741	49
Competition	1. I enjoy competing against other people in-game 2. I perform in the top 20% of people for an online game 3. I am one of the more competent players in an online game I play 4. I spend a majority of my time playing against other people 5. I devote a lot of time to playing games	0.727	59
Technology	1. I do well with computer technologies 2. I am interested in learning new computer technologies 3. I am literate with computer programs	0.710	57
Emotional	1. Playing online lets me relieve the stress from the day 2. I like the escapism aspect of online games 3. Playing the game lets me forget some of the real life problems I have	0.773	57
MMO	1. It is very important to me to get the best gear possible 2. I try to optimize my experience (XP) gain 3. I am influenced by which MMORPGs I play based on how much information is displayed in the user interface 4. I am influenced by which MMORPGs I play based on whether they allow me to modify the user interface	0.741	31
RTS/MOBA	1. I primary use the mini-map to navigate around in-game 2. I use the in-game metrics to gauge my abilities 3. I keep track of my in-game metrics while in a game 4. I use guides to stay informed on the latest strategies	0.739	41
FPS	1. I often look at the scoreboard 2. I use guides to stay informed on the latest strategies 3. The scoreboard is not a good source of information 4. I am influenced by which FPS games I play based on how much information they provide in the user interface	0.760	34

Table 1: Alpha reliability scores by scale

Another component is composed of questions from Yee's 2009 survey designed to establish external validity between this research and previous survey research. The questions used from that survey involve the self-reporting of the mental health of the respondents. These questions asked whether online games help participants to relieve stress, forget some of the real life problems they have, or appreciate the escapism aspect of online games. A reliability analysis on these questions yielded a result of .773, which is the highest of these clusters.

In the first cluster of questions i asked participants several questions to help identify the type of content that participants watched as spectators. I was able to combine four binary questions into a scale that measured the different amount of content respondents watched as spectators. These four questions are listed in *table 1* above. These four questions (n=52) have an alpha reliability of 0.799, which confirms my ability to combine them in to a single scale.

In addition to the competition, technology, emotional, spectator, and information scales, I also generated a scale for each genre of game my survey examines. Each of these scales is a compilation of four questions, though the exact questions do vary with each genre depending on the different types of data visualization best practices currently exist in these games.

The MMO scale consists of four questions that ask the respondent how important it is to them to optimize experience gain and to obtain the best gear possible. In addition to these two questions is another pair of questions that ask whether decisions about which MMO respondents play is dependent on how much information is in the user interface as well as how well they allow this interface to be customized. The questions in this scale have an alpha reliability value of .741, indicating an acceptable level of agreement between these questions.

The real time strategy and multiplayer online battle arena scale also consists of four questions, which all relate to the use of visual analytics while playing. One of the questions relates to mini-map usage whereas two others measure the use of in-game metrics by respondents and the final question asks about the usage of guides to stay informed on the latest strategies. These four questions have an alpha reliability value of .739 and thematically fit together. For this reason I created the scale using these four questions.

The final scale is composed of responses from the questions participants were asked to answer regarding first person shooting games. These four questions have players report whether they often look at the scoreboard; or if they do not find it a useful source of information. In this case, the response on the scale was inverted to positively correlate with the rest of the questions. The other two questions asked if players use guides to stay informed on the latest strategies and whether they are influenced about which first person shooting games they play based on the information available in the user interface. The questions that made up this final scale had an alpha reliability of 0.760 indicating a strong level of agreement between the questions.

Chapter Four: Results

Who uses analytic tools?

We were able to analyze sixty responses to my survey. In order to create a basic demographic profile of my respondents I asked appropriate questions such as age, gender, and country of origin. The country of origin question was added for several reasons. First, online games have a strong presence in Europe and Asia and this study aims to go beyond the North American boundary. Additionally, the websites used for recruiting have a strong cadre of participation from those outside of the United States.

The mean age of my respondents was 27.8 years old. However, the median was 25.50 and the largest number of participants reported being 25. The value at the 75th quartile was 33 and the range of age was 18 thru 49. Out of my respondents, 83 percent are male and 17 percent are female respectively, with two respondents choosing not to answer. Of the sixty responses, over a third, (35%) were from participants outside of the United States. After the United States, most respondents were from Canada (7%) and Europe (22%).

The first scale I examine is the information scale. This scale consists of seven likert-scale questions with value ranges from 0 to 4. Answers to questions were summed to produce a single numerical value. The theoretical range for this scale is 29 with a minimum of 0 and a maximum of 28. Of the 60 participants in the survey, 49 answered every question in the scale. The median score on the scale is 18 with a mean of 18.96. The scale for information customization in game is bimodal with seven respondents scoring 15 and 17 on the scale. The standard deviation for the scale is 4.2 indicating roughly 2/3 of respondents received a score between 15 and 23.

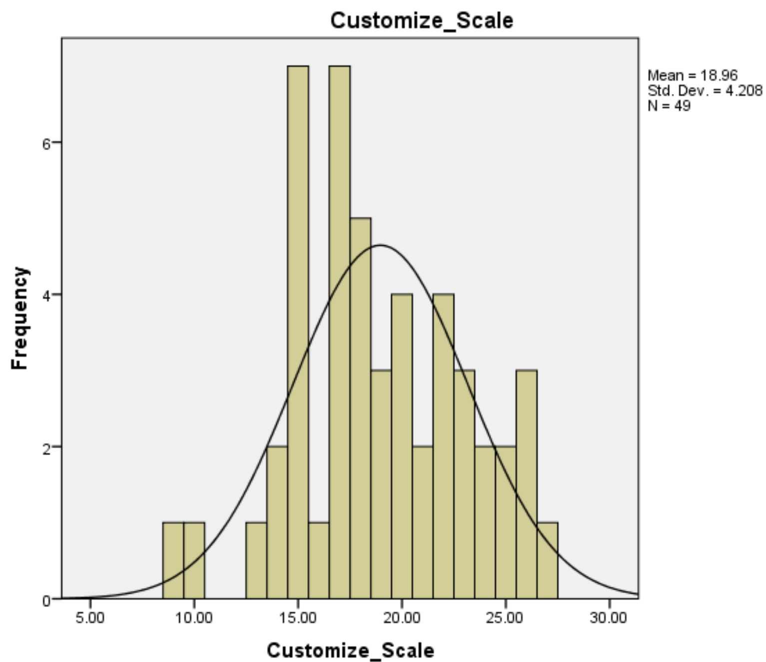

Figure 5: Histogram of customize information scale (N=49)

The customized information scale used throughout this research measures the extent to which participants in the survey are aware of, utilize, and modify the information tools available to them in online games. The alpha reliability for this scale was above 0.7, which indicates for the purposes of this research that the scale does an accurate job of identifying these attitudes in my survey participants. Most of the survey respondents appeared in the middle range of the scale indicating that awareness and customization of information takes place for many but only a few in my population went further and heavily customize the information in their games or base decisions on this information availability.

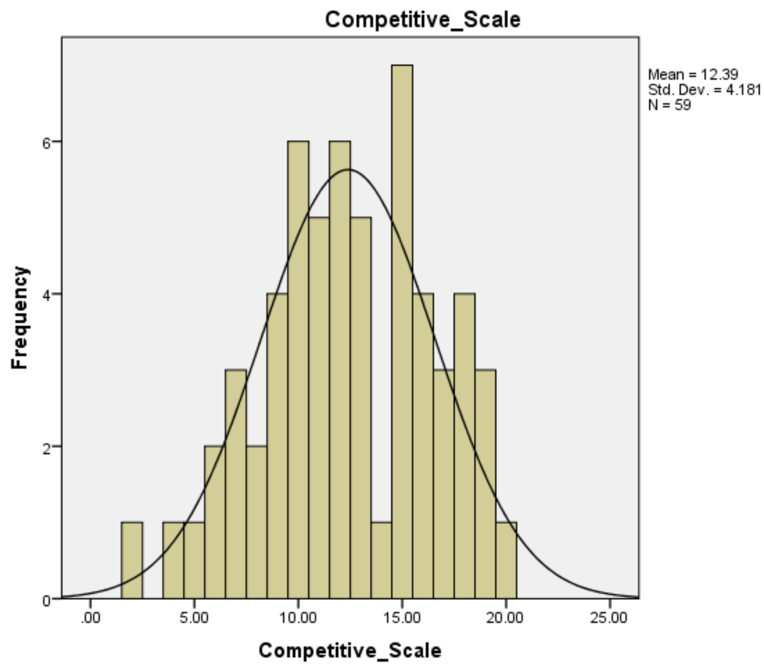

Figure 6: Histogram of competitive attitude scale (N=59)

The second scale I examine is the competition scale. This scale measures the level of competitiveness reported by respondents and includes questions about a participant's attitude toward playing against other people, how competent they felt they were, and how often they played against other people. This scale consists of five, five-point likert-scale questions resulting in a theoretical range of 21 with values possible from 0 to 20 inclusively. The competition scale had the highest N of the scales with 59 of 60 participants answering all of the questions in the scale. The median score on this scale is 12 with a mean of 12.39. The mode of the scale is 15 and the competition scale has a standard deviation of 4.18 indicating a relatively high level of variance in the scale.

Next I look at the technology scale, which measures the level of comfort that respondents have with technology and indicate an interest in computer technologies. This scale is composed of three questions resulting in a theoretical range of 13 with values from 0 to 12 available. Nearly every participant in the survey responded to all three questions for a total of 57 respondents with a defined value for the scale. Contrary to the other scales, which demonstrated relatively evenly distributed values, 27 (47%) of the 57 respondents scored the maximum 12 of 12 on the scale. Not surprisingly, the median score was 11 with a mean of 10.61 and standard deviation of 1.8.

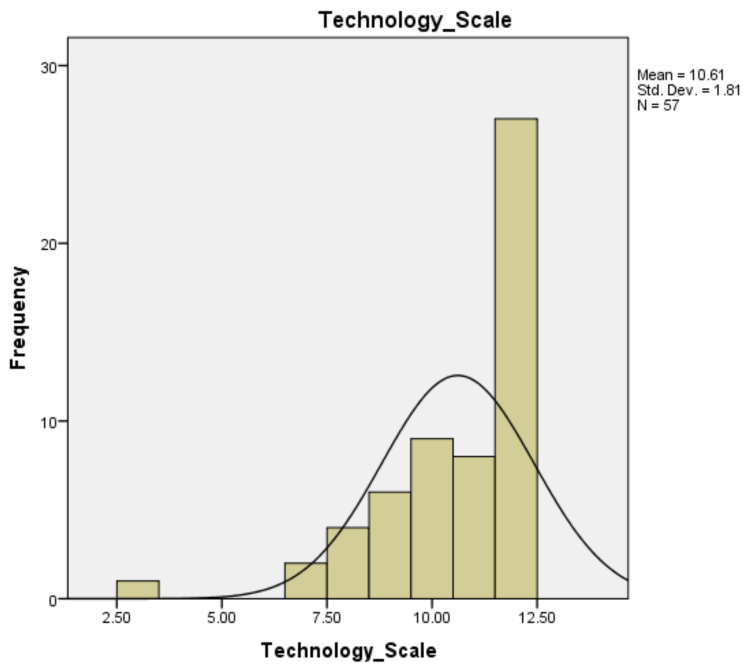

Figure 7: Histogram of technology scale (N=57)

The emotional scale measures the level of emotional attachment respondents have with online games and how participants use online games to relieve stress, escape, and forget about any real life problems the respondent may have. This scale is also composed of three questions with a possible range of values from 0 to 12. 57 out of the 60 participants answered every question in the scale to define a score. The mean value in the scale is 8.28 with a median of 9 and mode of 10. The standard deviation of the scale was 2.1.

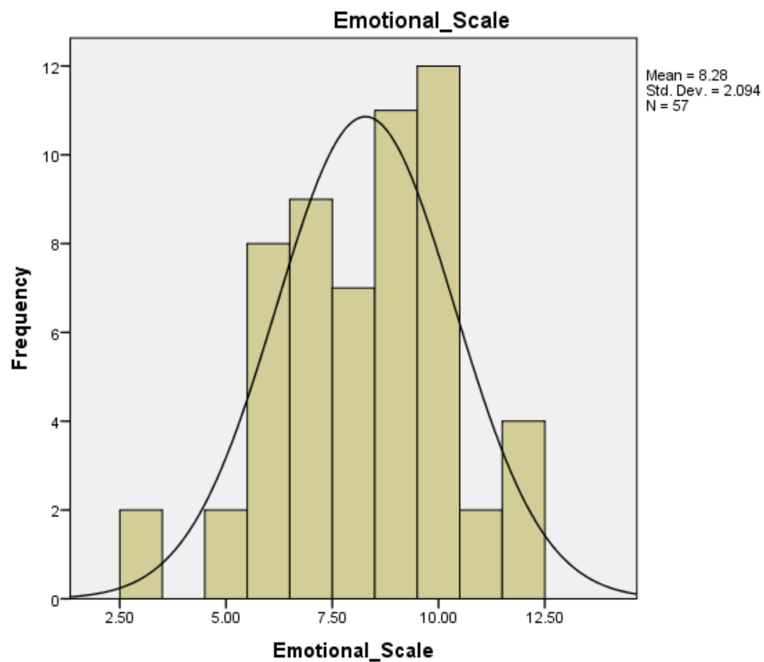

Figure 8: Histogram of emotional scale (N=57)

The spectator involvement scale incorporates four binary variables to generate a single scale that represents the involvement of the respondent with regard to spectating online matches. These questions were all near the start of the survey, yielding a sizeable N=57. The median and mode for the scale is 3 and the mean score on this scale for respondents is 2.2, indicating tendency toward agreement. Interestingly, for all respondents who scored a three on the spectator involvement scale, they all answered negatively to the same question, which asked if they had physically visited a tournament site as a spectator.

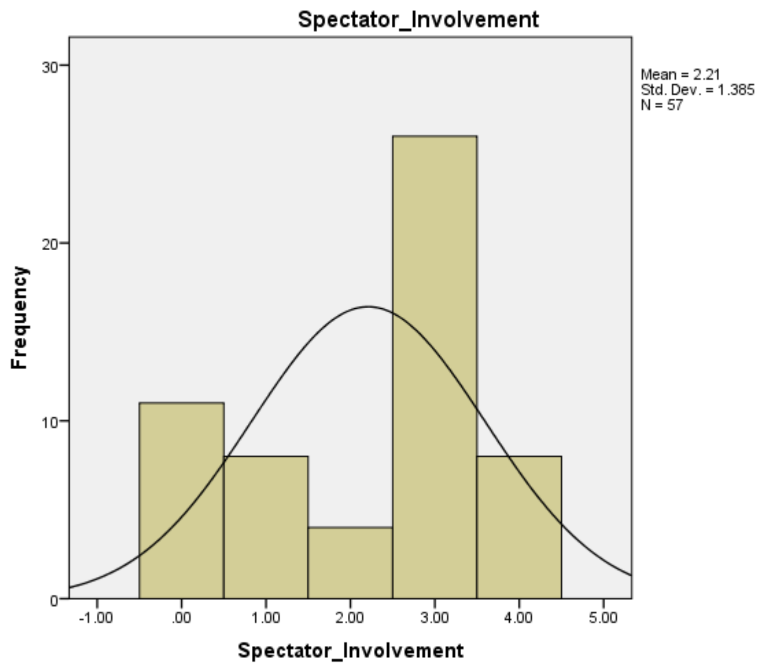

Figure 9: Histogram of spectator involvement (N=57)

A correlation check on the data was performed order to see if there were any interesting relations amongst the scales and the collected demographic data. For the table of discussed correlations see *table 2*. The strongest correlations between age and gender are the moderate negative correlations with spectator involvement and competitive attitudes. This would indicate that as age increases, the level of spectator involvement decreases. Likewise, as females are coded as one with males as zero, spectator involvement is moderately negatively correlated with the female participants in my survey. Similarly, competitive attitudes are both moderately negatively correlated with age and gender. This indicates that older people tend to be less competitive and females also tend to be less competitive than their male counterparts. Conversely, young males are the most likely to be involved in spectating and to have competitive attitudes toward online games.

Age and gender are weakly negatively correlated to the technology scale. The emotional scale is weakly positively correlated with age and gender and information customization has almost no measurable correlation with age and a weak positive correlation with gender.

Scale		Age	N	Gender	N
Spectator Involvement	Correlation	-0.36	57	-0.50	55
Technology	Correlation	-0.24	57	-0.16	55
Emotional	Correlation	0.19	57	0.23	55
Competitive Attitudes	Correlation	-0.26	59	-0.30	57
Information Customization	Correlation	-0.01	49	0.08	47

Table 2: Age and Gender correlations for different scales

Our first hypothesis is that those who are technologically and mathematically proficient are more likely to use data visualization and analytic tools available in online games. *Table 3* shows the relevant correlation between technology and the information customization scale and *table 4* shows the cross-tabulation between the technology and information customization scale. Both figures show the weak relationship between technology and the customize information scale which measures the overall attitude and use of data visualization and tools.

Scale		Information	N
Technology	Correlation	0.010	46

Table 3: Correlation coefficient between technology scale and information customization scale

Information Scale	Technology Scale Low (0-4) (N = 1)	Medium (5 – 8) (N = 5)	High (9-12) (N = 40)	Gamma
Low (0 – 9)	0.0%	0.0%	2.4%	
Medium (10 – 18)	100%	60%	46.4%	
High (19 – 27)	0.0%	40%	51.2%	-0.01

Table 4: Cross-tabulation between technology and information customization scales

In *table 4* I see that nearly all of the data is contained in the high end of the technology scale, which corresponds to the extremely high value of the respondents on the technology scale. Respondents who scored in the high in the technology scale were approximately split evenly between the medium score in the customize information scale, 46.4 percent, and the high score, 51.2 percent. I should note that the technology scale is comprised of scores where 47 percent of the respondents scored the maximum possible value, which would adversely affect any attempt to create a meaningful relationship between other scales.

Table 5 displays similar results, focusing instead on the relationship between the attitude toward competition scale and the use of data visualization and analytic tools. My second hypothesis is that competitive attitudes are positively correlated with the use of data visualization and attitudes toward their use. The information in *table 5* compares the correlation coefficient between the competition scale and the information scale. The relationship between the competitive attitudes scale and information scale is moderate with a positive correlation coefficient of 0.3 (N=48).

Scale		Information Customization	N
Competition	Correlation	0.300	48

Table 5: Correlation between competition and information customization scale

More specifically, I can observe the relationship between the two levels of reported competency in online games with the use of these analytic tools. The two questions in the survey related to player competency ask participants to respond on a standard five-point likert-scale with the statements: I perform in the top 20 percent of people in a game I play; and I am one of the more competent players in online games I play. Both of the questions are combined into a single competency scale with an alpha reliability score of 0.80. *Table 6* displays the correlation coefficients between this competency scale and the information customization scale. The relationship between the competency and information customization is moderate at 0.30.

Scale		Information Customization	N
Competency	Correlation	0.30	49

Table 6: Correlation between competency and information customization scales

Our fourth hypothesis tasked us with examining the relationship that spectators have with the analytic tools and data visualization while observing online games. My survey implement did not include questions related to spectating for specific genres. However, a separate question from any of my scales asks participants to agree or disagree with the statement, “I like it when broadcasters use a chart or graph to help explain what is happening in the game”. The use of these charts and graphs specifically reference analytic tools that spectators and broadcasters utilize when watching games. Responses to this question were recoded into a binary variable. The cross-tabulation between this question and the spectator scale can be seen in *table 7* as well as the correlation between the two measures in *table 8*.

	Do you like broadcasters using charts?		
Spectator Scale	No (N = 24)	Yes (N = 28)	Gamma
0 – 2	50%	21%	
3 – 4	50%	79%	0.57

Table 7: Cross-tabulation of spectator involvement scale and question asking whether respondents agree or disagree with the statement, "I like it when broadcasters use charts and graphs to explain what is happening in game"

		Do you like it when broadcasters use charts?
Spectator Scale	Correlation	0.34
	N	52

Table 8: Correlation between spectator involvement scale and whether respondents like it when broadcasters use charts

In the cross-tabulation I see that the split between those who prefer it when broadcasters use charts is roughly even with 46 percent of responses in the negative with 54 percent of participants agreeing with the statement. The largest percentage of respondents scored a 3 or higher on the spectator involvement scale. In every case where respondents scored a 3 on the spectator involvement scale, the question asking if survey respondents had attended a tournament as a spectator was the only question they would respond negatively to. When respondents agree they like broadcasters using charts, they score a three or higher on the spectator involvement scale 79 percent of the time. The gamma strength statistic for this cross-tabulation is strong at 0.57. Similarly, spectator involvement is moderately correlated with the sentiment that respondents like it when broadcasters use data visualization tools such as charts or graphs to help explain what is occurring in the game.

How are the analytic tools used?

The first research question focused on the different types of players. These questions focused on survey responses about their use of technology, emotional, competitive attitudes, and level of spectator involvement. Combined, this helps us examine who uses analytic tools.

Our second research question focuses on how these analytic and data visualization tools are utilized and to what purpose. *Table 9* below displays the frequency distributions for several questions in the survey related to what purpose these tools have in online games. For example, there are three questions (3, 4, 5) related to using the mini-map to find information in game, one for each genre. Likewise, three questions (6, 7, 8) relate to using paratexts such as game guides to help players stay up on the latest strategies or to defeat encounters. Question one involves the use of data visualization tools to increase efficiency in game. The second question asks users whether they customize their user interface to display more information about the game. Whereas questions three, four, and five all relate to using the mini-map to find information, question nine asks respondents whether they use the mini-map for navigation rather than information. Similarly, the last question asks respondents who play real time strategy or multiplayer online battle arena games whether they use analytic tools for analysis rather than information generation or navigation. These questions are not part of any scales used in the analysis in the rest of this research, but each of these questions offers insight and will be considered as separate items in future analysis.

Questions (Relevant Genre)	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	N
1. I primarily use hotkeys when I play games	0%	13.5%	21.7%	31.6%	31.7%	59
2. I customize my user interface	5.2%	13.8%	31.0%	37.9%	12.1%	58
3. I use the mini-map often to help me find information (MMO)	0%	0%	3.2%	38.7%	58.1%	31
4. I use the mini-map to alert me during gameplay (RTS/MOBA)	0%	0%	4.9%	56.1%	39.0%	41
5. I use the radar to find information in-game (FPS)	0%	3.0%	0%	51.5%	45.5%	33
6. I use game guides to help me complete encounters (MMO)	10.0%	16.7%	26.7%	36.7%	10%	30
7. I use guides to stay informed of the latest strategies (RTS/MOBA)	12.2%	24.4%	14.6%	39.0%	9.8%	41
8. I use guides to stay informed on the latest strategies (FPS)	20.6%	26.5%	29.4%	23.5%	0%	34
9. I primarily use the mini-map to navigate around in-game (RTS/MOBA)	0%	14.6%	31.7%	41.5%	12.2%	41
10. I use metrics to analyze games when I play them (RTS/MOBA)	7.7%	7.7%	28.2%	30.8%	25.6%	39

Table 9: Frequency distributions for questions related to how many respondents use analytic tools

Of particular interest is the positive agreement to use the mini-map for information retrieval across all three genres. The frequency for the use of paratexts to inform decisions is more evenly distributed, indicating that not all respondents are using these types of data visualization tools when they play online games. Finally, over 50 percent of respondents who play real time strategy games (N=39) either agreed or strongly agreed with the statement they use data visualization tools for analysis of game situations indicating this type of usage for data visualization is widespread among respondents.

How is use influenced by Genre?

Our third research question seeks to understand the difference between genres of online games. By understanding the differences between genres I can examine how different conventions in online games can influence the type and attitude toward data visualization. To start, I examine the number of responses for each genre that participants affirmed they played. These numbers are critical for this survey, because these questions act as a filter for each section related to that genre.

Of the 60 respondents 52 percent indicated they play MMORPGs and 57 percent of the participants said they play FPS games online. However, the largest percentage of people, 68, reported they play RTS or MOBA games online. In addition to these distinct categories there are several instances of overlap between genres respondents report playing. Out of the 60 respondents, 20 percent reported playing all three genres, 35 percent reported playing both MMO and RTS/MOBA games, 27 percent play both MMO and FPS games, and 43 percent play both RTS/MOBA and FPS games. For a complete description please see *figure 10* below.

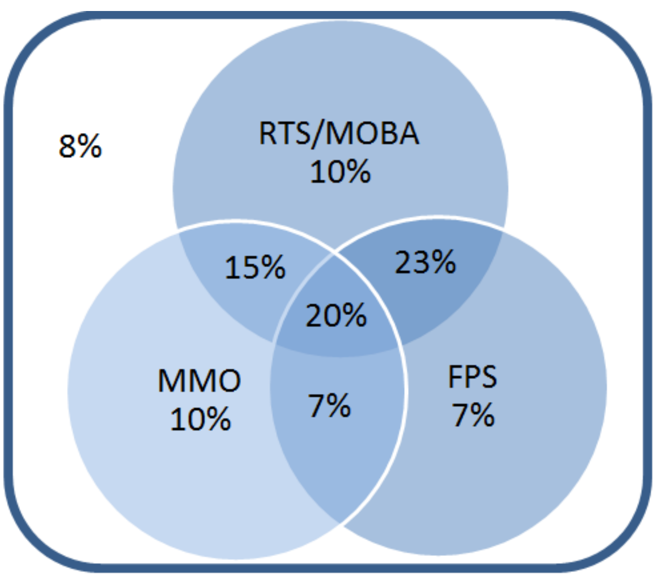

Figure 10: Venn diagram showing responses by genre (N=60)

Within each section, several questions related to the usage and preference of analytic tools in online games is combined into a single scale for each genre. The following three scales only consider responses from participants who played their respective genres. The massive multiplayer online scale measures participant responses to playing MMOs and consists of four questions that ask respondents about their habits and preferences playing these online games. This scale has an alpha of 0.741. The scale has theoretical range of 17 from a minimum of 0 to a maximum of 16. The scale is defined for the 31 respondents who answered all four questions. The median and mode value on the scale is nine and the mean score on the scale is 9.68 with a standard deviation of 3.3. Of interest is the 5 (16%) of the 31 respondents who ended in the negative half of the scale with a score less than eight.

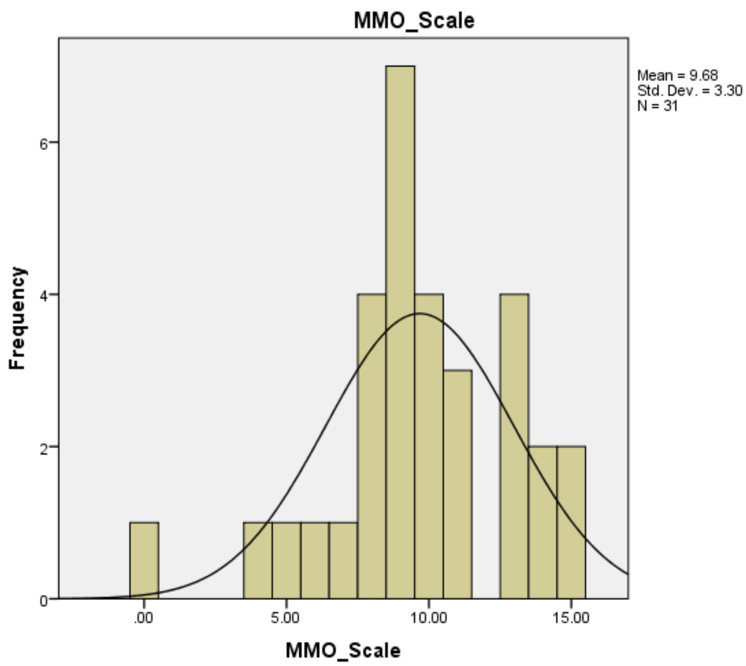

Figure 11: Histogram of MMO Scale (N=31)

The number of survey participants with a defined real time strategy scale is much higher than the number with a defined MMO scale. In fact, 68 percent of respondents affirmed playing real time strategy or multiplayer online battle arena games. This scale consisted of four questions primarily concerned with the presentation of information through various metrics while participants are playing online games. The four questions comprising this scale have an alpha score of 0.739. Similar to the MMO scale, the RTS scale spans a theoretical range from 0 to 16. Also of interest is that 31 percent of the 41 respondents were in the negative half of the scale with a score of less than eight. The median score on the RTS scale is 10, with a mean of 9.54 and a mode of 11. The standard deviation for this scale is 3.25.

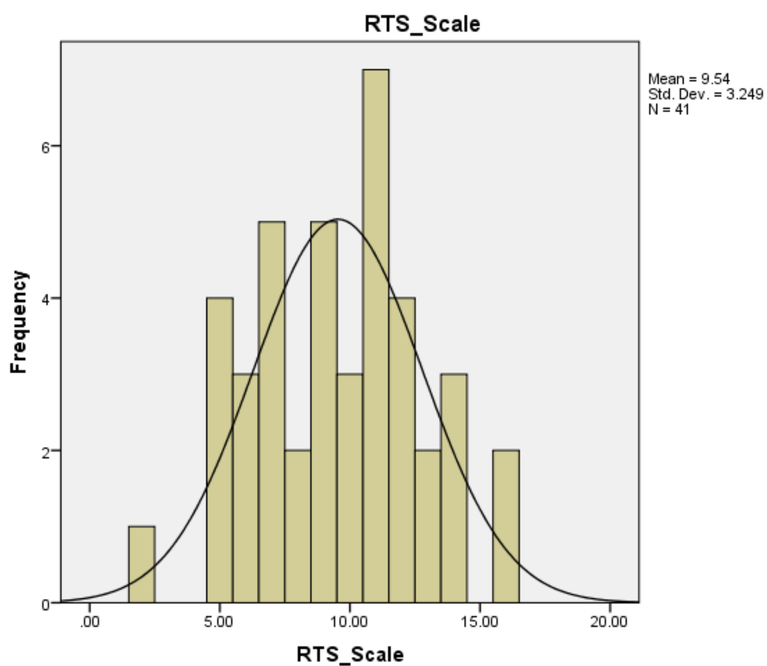

Figure 12: Histogram of RTS and MOBA scale (N=41)

The final scale is based on the 34 responses to the survey participants that play first person shooting games. The four questions in this scale have an alpha value of 0.760. This four-question scale also has a theoretical score that range from 0 to 16 with a median score of 9, a mean of 8.6 and a mode of 10. The standard deviation for this scale is 3.4. The scale recorded overall disagreement with the scale from 32 percent of the 34 respondents who play FPS games and one participant responded with the minimum possible score in the scale.

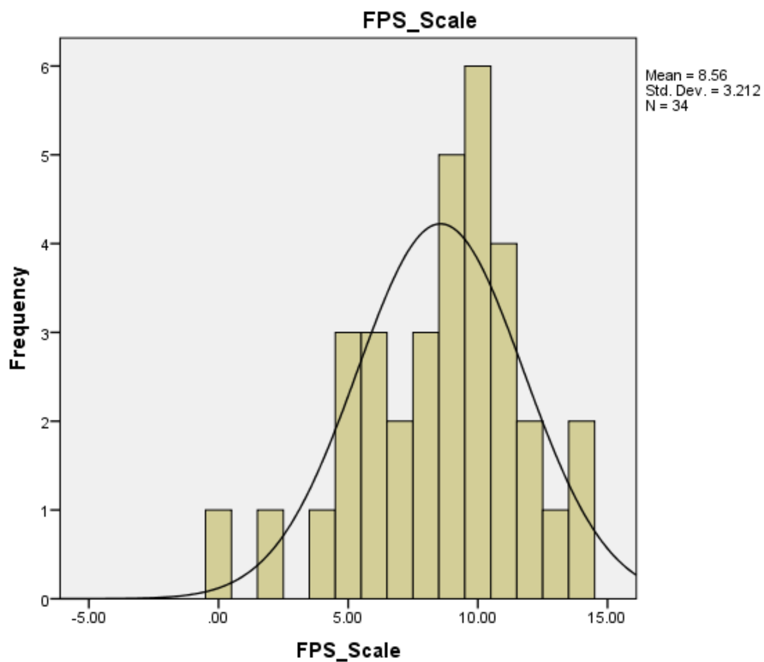

Figure 13: Histogram of scores on FPS scale (N=34)

Scale	MMO	N	RTS	N	FPS	N
Information Customization	0.52	29	0.41	35	0.28	28
Competitive Attitudes	0.70	30	0.17	40	0.43	32
Spectator Involvement	0.23	29	0.29	41	0.17	31
Technology	0.09	29	0.08	40	0.09	33
Emotional	-0.18	31	0.48	40	0.43	32

Table 10: Correlations between genre specific scales and player scales

The table above displays the correlations between my genre specific scales and the scales I have previously mentioned that measure different aspects of player demographics. Comparing these correlations allows us to examine differences between genres of online games with regard to my player demographics.

The information customization scale has a moderate correlation with the MMO scale and a RTS/MOBA scale and a weak correlation with the FPS scale. The relationship between the competitive attitude scale and the MMO scale is very strong at 0.7 whereas the relationship between the RTS scale is weak at 0.17 and moderate with the FPS scale indicating different attitudes between all three genres. Conversely, both the spectator involvement scale and the technology scale are weakly related to all three genres with little difference between the correlations. Finally, the emotional scale is weakly negatively correlated with the MMO scale and moderately correlated with the RTS and FPS scales.

Consequences of analytic tools?

In addition to traditional measures of demographics, I included specific demographic questions related to playing online games. When asked about their typical day, respondents replied they played an average of 2.36 hours, or 2 hours and 22 minutes, a day. Of the 58 respondents who answered the question, five, or 8.6 percent, replied they typically spent no time playing games. A similar number of respondents indicated they generally are not able to play online games during the week. For this question, 4 of 54 (7.4%) respondents replied they did not play online games in a typical week with a median response of 5 days a week. Combining these two questions gives us a workable range of time spent playing games in a typical week from 0 hours up to a possible maximum of 35 hours a week.

Table 11 displays the relationship between time spent playing games and whether respondents rated in the low, medium, or high range of the customize information scale. The only respondent to score in the low end of the customize information scale also reported playing between zero and five hours a week. Above the 10 hour per week mark, respondents scored in the high range of the customize information scale 66.7 percent of the time and 33.3 percent of the time they scored in the medium range of the scale. This is in contrast to respondents who reported playing between six and 10 hours every week. In that instance, respondents score in the high end of the customize information scale 33.3 percent of the time and score the medium range 66.7 percent of the time. The gamma strength statistic for this cross-tabulation is 0.58 indicating a strong strength relationship between customize information and time spent playing per week.

	Time Per Week			
Customize Information	0-5 Hrs (N = 11)	6-10 Hrs (N = 12)	11+ Hrs (N = 21)	Gamma
Low (0 – 9)	9.1%	0.0%	0.0%	
Medium (10 - 18)	63.6%	66.7%	33.3%	
High (19 - 27)	27.3%	33.3%	66.7%	0.58

Table 11: Cross-tabulation between customize information scale and time spent playing online games per week

One other relevant question related to playing online games asked participants to indicate approximately how much money they had invested in to online games in the last three months. For the 56 respondents who answered this question the median response was $50, which approximates the cost of one new game. The mode of the responses was zero, indicating for nearly 30 percent of the group they had not purchased anything in the last three months but instead were either playing or watching games they already owned.

Table 12 displays the cross-tabulation between the customize information scale and the amount of money spent on online games in the last three months. The cross-tabulation is broken into two categories for amount of money spent on online games over the last three months. Respondents who report spending more than $60 over the course of the last three months scored in the high range of the customize information scale 58.8 percent of the time compared to the 35.3 percent who scored in the medium range and the minimal 5.9 percent who scored in the low range of the scale. Comparatively, those who spent $60 or less scored 55.2 percent in the medium range and only 44.8 percent in the high range of the customize information scale. Overall, respondents are about equally likely to score in the medium range on the customize information scale and in the high range. The gamma statistic for this cross-tabulation is 0.20 indicating a weak relationship between money spent and the information customization scale.

Information Customize	Cash Spent $0 - $60 (N = 29)	$61+ (N = 17)	Gamma
Low (0 – 9)	0.0%	5.9%	
Medium (10 – 18)	55.2%	35.3%	
High (19 – 27)	44.8%	58.8%	0.2

Table 12: Information customization scale by money spent in the last three months

In addition to asking participants how much they had spent on online games in the last three months I asked participants to name the last three games they had played. Responses were re-coded in to genres to break down the play habits of respondents by genre. *Table 13* below displays the results of these questions. Of interest is the decline in the number of people who reported playing RTS/MOBA games as they move further from their last game played. Meanwhile, the relative number of people who report playing both FPS and MMORPG games remains fairly consistent throughout all three past histories of games. I should note that all 12 respondents previously mentioned who work for a game developer indicated they had played the MMO they worked on as one of the last three games they played which would skew the percentages in favor of MMOs.

	MMO (N = 56)	RTS/MOBA (N = 45)	FPS (N = 28)	Other (N = 31)
Last Game Played	32.1%	48.9%	39.3%	19.4%
Second Last Game Played	35.7%	31.1%	25.0%	45.2%
Third Last Game Played	32.2%	20.0%	35.7%	35.4%

Table 13: Frequencies of last games played reported by respondents by genre

Chapter Five: Discussion

Our survey generated 60 responses due to recruitment from various online community gaming sites. The makeup of these responses is fairly similar to what Yee (2006a) found in his study of 7,000 players of various MMOs. The male to female ratio in my survey is nearly identical to the response rate that Yee cites in his research of 84 percent male and 16 percent female (2006b). In addition, many of those surveyed were in their 20s, representing the relatively younger age group that has grown up with online games and data visualization tools. This can be understood in the context of the wide range of years of experience that survey participants reported; between 1 and 20. While these numbers are not surprising, extra care should be taken to consider that these survey responses represent the opinions and self-assessment of a small community of players out of a much more robust and diverse broad grouping of those who play online games.

At the same time, these relatively small communities of players are also vocal, dedicated, and extremely loyal to products and games they deem worthy. Specifically, this is the type of audience that developers are seeking when they design games because the respondents to this survey make up many of the consumers for ancillary game markets. Indeed, there were five respondents who reported playing less than one hour on a typical day, yet they rated above 0 on the spectator scale, indicating their primary engagement with the game is through these secondary markets. Additionally, 30 percent of the respondents reported they had not spent any money in the last three months on online games indicating they were either watching or playing games they already owned.

Indeed, almost all of the titles mentioned in the last played category were released in 2011 or prior.

The distribution between online game genres was much closer than expected. With the reported widespread popularity of games such as *World of Warcraft* and others in the MMO genre, to find that the least number of players played MMOs compared to real time strategy or first person shooting games was surprising given the relative lack of attention paid attention to these other genres in scholarly work. The bias toward real time strategy and multiplayer online battle arenas is understandable and should be noted. The community websites used for recruitment originally started out as information portals for *Starcraft* professional matches, prior to the explosion of massive multiplayer games or recent generations of first person shooters such as counter-strike: source or halo.

While the players on these sites play other games as well, the community resources are geared toward these genres specifically. The growth of e-sports to be synonymous with real time strategy and multiplayer online battle arena games also explains part of the popularity of these types of games in my survey. I should also note that 12 survey participants are on staff at a game development company that creates and distributes content for a massive multiplayer game, which introduces additional bias in terms of the last games played statistics.

Survey respondents were highly positive in their attitudes towards technology, comfort with existing, and learning new computer technologies. Unfortunately, this causes issues with my attempts to quantify the relationship between technology and the use of data visualization tools. This scale attempted to measure the level of computer expertise that players have with computers to see if there was some relationship between technology literacy and the use of data visualization and analytic tools. *Table 4* demonstrates the extremely weak relationship between these two factors. My data therefore suggests that I should reject my first hypothesis.

We can identify at least two possibilities for the limitation of my data. The first possibility is the scale suffers from measurement validity. Nearly half of all respondents scored the maximum value in the scale. While a general level of expertise and comfort is expected future iterations of this research will need to revisit this scale and determine a more appropriate measure of technology literacy and its effects on data visualization utilization. Future scales could offer more gradations or more difficult measure of what it means to be truly technologically literate.

A second possibility is the population of my respondents is especially technologically savvy and the measurement is accurate. Evidence for this possibility is the inherent connection between online computer games, computer science, and technology. Many children use computer games as a gateway to learning various computer technology skills including programming. In addition, community websites tend to foster and attract many users who are interacting and constantly learning about new technologies through each other. Whichever possibility i attribute the results of the technology scale to, I am forced to conclude this first hypothesis is not supported.

The competitive attitude scale measures the devotion and sense of competition that survey participants report feeling when they play online games. The community sites that the survey were posted to are frequently traversed by highly competitive players whose primary goal is winning tournaments or enticing viewers to watch their streams to procure advertising revenue.   I defined a competitive attitude as having competency in an online game to a high degree, a devotion to spending time to playing the game, and a desire to play and compete against other players specifically in online games. This scale accrued the highest level of variance amongst the survey respondents, which serves to indicate the relatively diverse competitive interests of the respondents. For example, professional players who visit the site would rate highly on the competitive attitudes whereas spectators who primarily observe other players would rate low on the competitive attitudes scale.

Table 5 displays the correlations for competitive attitudes and scales representing general and genre specific uses of data visualization and analytic tools. My second hypothesis is confirmed by moderate correlation coefficient of 0.3 between competitive attitudes and information customization.

Our third hypothesis relating competency and use of data visualization tools in game follows a similar pattern with my second hypothesis. *Table 6* displays the results of the correlation between the two questions combined in to a single competency scale related to the customized information scale. These weak correlations indicate that those with increased competency tend to utilize information customization more. However, the relationship is not absolute indicating a large possibility of non-experts utilizing data visualization tools. This corroborates with the uses that Thomas and Cook (2005) mention as one of the key advantages to data visualization.

One of the main uses for the community sites that I used for recruitment of my survey is to provide members with a list of currently live broadcasts and tournaments. This way, at any given time, those who wish to spectate professional matches between players have the ability to watch other players perform in matches for online games. As part of my research questions, the fourth hypothesis posited that spectators are positively affected by the use of charts and graphs by broadcasters to help explain what is happening in the game. *Table 7 and table 8* show the heavy relationship between spectator involvement and appreciation for the use of data visualization tools by broadcasters confirming my fourth hypothesis. All of the categories except for the zero and two on the spectator involvement scale have a higher percentage of respondents who like broadcasters using charts.

The data for respondents who scored a zero on the spectator scale is skewed negatively, because neutral responses were recoded as zero when the variable was turned into a binary variable. Regardless of this issue, I can confirm the hypothesis that respondents positively react to broadcasters when they use charts or graphs to explain what is occurring in the games they are watching based on the bivariate relationship between spectator involvement and broadcaster approval.

Our fifth hypothesis proposed that differences exist between genres in terms of the uses for data visualization. Examining *table 10* confirms this hypothesis. There is a trend between genres where the strength of the relationship is strong, medium, and weak with MMOs, RTS and FPS games respectively. I propose a few possibilities to help explain this observed relationship between these genre factors and why they are strongest for MMOs and only medium or weak for RTS and FPS genres.

One possibility is that online games in the MMO genre allow for more player-types, which may be influenced by competitive attitudes. This would be due to the relatively undefined victory conditions that are possible in MMOs. For explorers, who may be less competitive, this may involve discovering a new zone. At the same time, those players who are killers or achievers, a more inherently competitive play-style, can be driven by a victory condition that involves killing a particular boss or player versus player instance. This is in contrast to most real time strategy and first person shooting games, which necessarily restrict the types of victory conditions that exist which narrow the possibilities of play style. With less play styles there may be less of competitive attitudes because the games are designed where everyone has to play competitively.

Along these same lines, my questions may suffer from measurement validity in MMOs because it is harder to define what competency is in these types of games. Competency in this sense can refer to one of several different player types that may exist in MMOs. An explorer might consider themselves competent if they have mastered a particular wall climbing technique not many others would know. Alternatively, beating a specific boss or a ranking in a player versus player match could also define competency depending on the player type.

Real time strategy games generally offer a ladder based system that allows players to rank themselves after every match against others. Here, the top 20 percent is rather well defined as either a position in the ladder system, or a win-loss ratio that can be calculated after every match. Competency can be likewise measured easily as mastery over game mechanics, a win-loss ratio greater than 1, or proper execution of certain roles within a team setting. Likewise, victories, kills and deaths are all easily calculated in many first person shooting games in real time allowing for an easy measurement of ranking or competency using those statistics.

Of particular interest is the single negative relationship between the emotional and MMO scales. This indicates that as players who utilize more analytic tools in an MMO are also less likely to use the MMO as an escape or to relieve stress from the day. Nardi (2009) and Chen (2010) examine this phenomenon in their research into raiding in *World of Warcraft*. Chen observed the often long list of tasks that have to be completed in order to successfully raid in the *World of Warcraft* environment. Nardi observed these often repetitive tasks can create a sense of work, often the antithesis of the feeling of escape or venting through online games that the emotional scale measures.

Another possibility is the practice of information customization and the use of data visualization tools are much more developed and regularly used in MMOs. Indeed, several ethnographic studies have exclaimed upon the value and necessity of incorporating add-ons if a player was serious about performing in the highest levels of the online games (Nardi, 2009; Chen, 2010; Glas 2010). This phenomenon could lead to such ubiquitous use of these tools that players who consider themselves competent in the game are the ones who are able to navigate these additional add-ons and data analytic tools successfully.   While data visualization tools are also ubiquitous in real time strategy and first person shooting games they serve a different role. Whereas in MMOs, these tools can inform and determine every decision, in RTS and FPS games they offer a role more akin to advisory, offering information when wanted and ignored just as easily.

The consequences of effective data visualization can be observed from *table 11* through *table 13*. My hypothesis that the amount of money spent on online games in the last three months is positively related to the use of data visualization tools is confirmed with the weak relationship observed by the 0.24 gamma from the cross-tabulation between these two variables. The weakness of this relationship is not surprising. Whether a player has bought a game is of tantamount importance to a developer but difficult to quantify across a broad scope of online games through survey methodology.

The amount of money spent is a relative approximation for the purchase of an online game or the monthly subscription of a service to an MMO. This holds several assumptions. First, my measure assumes that games are not free to play. In some cases this is not the situation, as with a game in beta or a developer that utilizes a free to play model. Of more importance to these developers is the number of players who are actively playing their game, or who perform a particular action within the game such as clicking on an advertisement. Second, for the purpose of understanding, this measure assumes that all games cost the same, which is not the case. Additionally, only purchases that are made in the last three months are considered introducing external validity issues with regard to the timing of my survey.

Our last hypothesis concerning the relationship between time spent playing and the use of customization is confirmed by the moderate gamma displayed in *table 11*. This represents the amount of time that players spend in game as reported from the previous seven days from the survey. The high percentage of players who report on the high end of the customize information scale for respondents who also play more than 10 hours a week is expected. Of particular interest though is that even for players who are only able to play less than five hours a week, 27 percent still score high in the customize information scale and 91 percent overall report at least using some level of information customization in their online games. This displays the near ubiquitous nature of information customization across all levels of interest in online games.

A central limitation in the design of this research study is the issue of breadth versus depth. The topic of this research study is intentionally broad to allow for inclusiveness in terms of survey responses and response types. In order to create a broad survey that allowed for comparisons between genres and still be answerable within 10 minutes I was forced to intentionally leave some measurement questions vague and only include three or four questions per hypothesis in the survey. This design is appropriate when the sample size and recruitment are effectively large. In the case of this research, the sample size needs improvement in combination with finer questions that can adequately assess measures in the hypothesis such as level of comfort with technology.

Fortunately, my research does display characteristics of reliability both within the survey and with external sources such as Yee's work on profiles for MMO participants. As a measure of external reliability my survey performed an excellent job re-creating the escapism factor utilizing survey questions from Yee's 2006 paper on the demographics, emotions, and derived experiences of players in MMOs. Whereas in that work, the level of agreement between the questions in the escapism factor were only 0.62, my results demonstrated an alpha of 0.773 indicating a higher level of agreement between these questions for my sample. Players in all three genres utilize the paratexts confirming the use proposed by Consalvo (2007). Additionally, I observed many of the competitive and emotional aspects that Nardi (2009), Chen (2010), and Glas (2010) observe in players who raid in *World of Warcraft*.

Ultimately, this research shows that players are aware of and are utilizing these data visualization tools when they watch or play online games. Online games generate massive amounts of information for both players and developers. I expect this trend to continue and make this type of research focused on player perspectives to be more relevant and necessary in the coming years. Developers have been sifting through massive amounts of data for years with recent developments in data visualization and machine learning techniques that are finally allowing that data to be analyzed by corporations. However, players are also using these advances in innovative and novel ways to enhance their understanding of the games they play. The growth of both of these activities will continue as these tools appeal to players and spectators who want to learn more about what they are playing and watching. This will cause demand on developers to develop appropriate tools that players can use and want. Research such as this survey are useful tools for both developers and players seeking to understand who is currently using these tools, how they are using them, and what consequences that use has for the games they play.

Appendix

Below is a copy of the questions asked in the survey

Please check all that apply:

1. In the past 7 days I have:
    a. Used a social website such as Facebook
    b. Used an internet chat service
    c. Watched a video online
    d. Read news online
    e. Contributed to a blog or community website

For the following questions, please select the most appropriate choice

2. I have watched a stream for a competitive tournament for an online game
3. I have attended a competitive tournament for an online game as a spectator
4. I have attended a competitive tournament for an online game as a player
5. I own more than 1 computer
6. I have watched other people stream games of them playing online
7. I have watched broadcasts of people commenting during online games

For the following questions, please indicate whether you agree or disagree with each statement on a scale from 1 = Strongly Disagree to 5 = Strongly Agree. If the question is not applicable please select 6 = Not applicable.

8. I do well with computer technologies
9. I enjoy competing against other people in game
10. I prefer to have as much information as possible in the user interface of a game
11. Playing online games let me relieve the stress from the day
12. I am comfortable understanding charts
13. I am interested in learning new computer technologies
14. I prefer to have less information in my user interface
15. I primarily use hotkeys when I play games
16. Games that do not feature a customizable user interface frustrate me
17. I like it when broadcasters use a chart or graph to help explain what is happening in an online game
18. I am literate with computer programs
19. I like the escapism aspect of online games
20. I perform in the top 20% of people in an online game I play
21. Doing massive amounts of damage in games is satisfying
22. I like games that allow add-ons
23. I prefer to play against computer opponents in online games
24. I customize my user interface
25. I am one of the more competent players in an online game I play
26. Playing games let me forget some of the real life problems I have
27. I enjoy performing mathematical calculations

28. I spend a majority of my time playing against other people in games
29. I devote a lot of my time to playing games

For the following questions please provide the most appropriate answer

30. How much money have you spent on games in the last three months
31. How many years of experience do you have playing online games
32. Approximately how many hours a day do you spend playing online games
33. Approximately how many days a week do you spend playing online games
34. Please list up to the last three online games you have played
35. Do you play MMORPGs such as World of Warcraft (WoW), The Old Republic (TOR), etc.

If you answered 'Yes' to question 35 then please continue to question 36, otherwise skip to question 46.

All of these questions relate to playing MMORPGs such as WoW or TOR.

For the following questions, please indicate whether you agree or disagree with each statement on a scale from 1 = Strongly Disagree to 5 = Strongly Agree. If the question is not applicable please select 6 = Not applicable.

36. I use my mouse to click on spells or abilities in game
37. It is very important to me to get the best gear possible
38. I like to feel powerful in game
39. I try to optimize my XP gain as much as possible
40. I use the mini-map often to help me find information
41. I travel with the map overlay on the screen
42. I use game guides to help me complete encounters
43. I am influenced by which MMORPGs I play based on how much information is displayed in the user interface
44. I am influenced by which MMORPGs I play based on whether they allow me to modify the user interface

Please check all that apply

45. I have used the following in game
    a. Add-ons
    b. Map notes
    c. Game guides
    d. Game walkthroughs
    e. Scoreboards
    f. Auction house or similar feature
46. Do you play real time strategy (RTS) or multiplayer online battle arena (MOBA) games such as Starcraft II or League of Legends, etc.

If you answered 'Yes' to question 47 then please continue to question 36, otherwise skip to question 56. For the following questions, please indicate whether you agree or disagree with each statement on a scale from 1 = Strongly Disagree to 5 = Strongly Agree. If the question is not applicable please select 6 = Not applicable.

47. I primarily use the mouse to select or activate abilities
48. I primarily use the mini-map to navigate around in-game
49. I use the in-game metrics to gauge my abilities

50. I use the mini-map to alert me during game play
51. I keep track of my in-game metrics while in the game
52. Winning is my primary goal when I play
53. I use metrics to analyze games when I watch them
54. I use guides to stay informed of the latest strategies
55. I am influenced by which RTS or MOBA games I play based on how much information is displayed in the user interface
56. Do you play first person shooting (FPS) games online such as Counter-Strike, Team Fortress 2, etc.

If you answered 'Yes' to question 57 then please continue to question 36, otherwise skip to question 62. For the following questions, please indicate whether you agree or disagree with each statement on a scale from 1 = Strongly Disagree to 5 = Strongly Agree. If the question is not applicable please select 6 = Not applicable.

57. I use the radar to find information in-game
58. I often look at the scoreboard during my games
59. I use guides to stay informed on the latest strategies
60. The scoreboard is not a good source of information during my games
61. I am influenced by which FPS games I play based on how much information they provide in the user interface

For the following questions please provide the most appropriate answer

62. What is your current age
63. What is your gender
64. What country are you currently located in

References

Ahmad, M., Keegan, B., Srivastava, J., Williams, D., & Contractor, N. et. al. (2009). Mining for Gold Farmers: Automatic Detection of Deviant Players in MMOGs. *International Conference on Computational Science and Engineering*. (4) 340-345.

Bartle, R. (2004). *Designing Virtual Worlds*. Indianapolis, IN: New Riders.

Bartle, R. (1996). Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of Virtual Environments*. Available at http://www.mud.co.uk/richard/hcds.htm. Accessed Nov. 28, 2011.

Bator, F. (1957). The Simple Analytics of Welfare Maximization. *The American Economic Review*. (47.1) 22 – 59.

Boellstorrf, T. (2010). *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human*. Princeton, NJ: Princeton University Press.

Castranova, E., Williams, D., Huang, Y., Shen, C., Keegan, B., Ratan, R. (2009). As real as real? Macroeconomic behavior in a large-scale virtual world. *New Media & Society*. (11.5) 685-707.

Castranova, E. (2002). On Virtual Economies. *Game Studies.* (3.2).

Chen, M. (2010). Leet Noobs: Expertise and Collaboration in a World of Warcraft Player Group as Distributed Sociomaterial Practice (Doctoral dissertation). Available from University of Washington College of Education.

Csikszentmihalyi, M., LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*. (56.5) 815–822.

Consalvo, M. (2007). *Cheating: Gaining advantage in Video Games*. Cambridge, MA: MIT Press.

Cummings, J. (2011). GDC Online – Hunting Whales: Will Analytics Kill Off Diversity of Gameplay?. *Motivate. Play*. Retrieved from http://www.motivateplay.com/2011/10/gdc-online-%E2%80%93-hunting-whales-will-analytics-kill-off-diversity-of-gameplay/.

Keim, D., Andrienko, G., Fekete, J, Gørg, C., Kohlhammer, J., & Melancon, G. (2007).

Visual Analytics: Definition, Process, and Challenges. In *Dagstuhl group report*.

Entertainment Software Association. (2011). Essential Facts About the Computer and Video Game Industry: 2011 Sales, Demographics, and Usage Data. *ESA.* Retrieved from http://www.theesa.com/facts/gameplayer.asp.

Furnas, G. (1986). Generalized Fisheye Views. *Human Factors in Computing Systems CHI '86 Conference Proceedings*. 16-23.

Glas, R. (2010). Games of Stake: Control, Agency and Ownership in World of Warcraft (Doctoral dissertation). Available from the University of Amsterdam.

Hirshleifer, J. (1988). The Analytics of Continuing Conflict. *Synthese, Formal Analysis in International Relations*. (76.2) 201-233.

Iosup, A. (2009). CAMEO: Continuous Analytics for Massively Multiplayer Online Games on Cloud Resources. *EUROPAR*. Delft, Netherlands.

Lyon, D. (2001). *Surveillance Society: Monitoring Everyday Life*. Buckingham: Open University Press.

Medler, B. (2011). The Data Cracker: Building a DEAD SPACE 2 Visual Game Analytic Tool. *GDC Online.*

Medler, B. (2009). Generations of Game Analytics, Achievements, and High Scores. *Eludamos – Journal for Computer Game Culture.* (3.2) 177-194.

Moberly, K. (2010). Commodifying Scarcity: Society, Struggle, and Spectacle in *World of Warcraft. Eludamos - Journal for Computer Game Culture*. (4.2) 215-235.

Nardi, B. (2009). *My Life As A Night Elf Priest: An Anthropological Account of World of Warcraft*. Ann Arbor, MI: University of Michigan Press.

"Quality Gaming Content and Discussion -- /r/Games". Reddit.com/r/Games. Reddit. Web. 10 March 2012.

Taylor, T.L. (2009). The Assemblage of Play. *Games and Culture.* (4.4) 331- 339.

Taylor, T.L. (2006). *Play Between Worlds: Exploring Online Game Culture*. Cambridge, MA: MIT Press.

"The League of Legends Season One Championship Concludes: Over 1,690,000 Viewers Tune in to Watch FNATICMSI Take the Title." *Riot Games News*. Riot Games. Web. 27 June 2011.

Thomas, J. J. & Cook, K.A. (2005). *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. Washington, DC: IEEE Computer Society.

Thurau, C. & Bauckhage, C. (2010). Analyzing the Evolution of Social Groups in World of Warcraft. *IEEE Symposium on Computational Intelligence and Games*. Dublin, England

Tufte, E.R. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, Connecticut: Graphics Press.

Tychson, A. & Canossa, A. (2008). Defining Personas in Games Using Metrics. *FuturePlay*. Toronto, Canada.

Weber, B., John, M., Mateas M., & Jhala, A. (2011). Modeling Player Retention in Madden NFL 11. *Innovative Applications of Artificial Intelligence*. San Francisco, CA.

Weber, B., Mateas, M. & Jhala, A. (2011). Using Data Mining to Model Player Experience. *FDG Workshop on Evaluating Player Experience in Games*. Bordeux, France.

Williams, D., Poole, S., Contractor, N., & Srivastava, J. (2010). The Virtual World Exploratorium. *Communication Methods and Measures*.

Williams, D., Yee, N., & Caplan, S. (2008). Who plays, how much and why? Debunking the stereotypical gamer profile. *Journal of Computer-Mediated Communication*. (13) 993-1018.

Woods, D., Patterson, E., & Emilie, R. (1998). Can we ever escape from data overload? A cognitive systems diagnosis. United States Air Force Research Laboratory.

Yee, N. (2006a). Motivations for Play in Online Games. *CyberPsychology & Behavior*. (9.6) 772-775.

Yee, N. (2006b). The Demographics, Motivations and Derived Experiences of Users of Massively-Multiplayer Online Graphical Environments. *Presence: Teleoperators and Virtual Environments.* (15.3) 309-329.