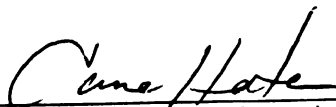This is to certify that the
dissertation entitled

THE EFFECT OF RPG NEWNESS, RATING, AND CHARACTER

EVILNESS ON THE NPC BELIEVABILITY

presented by

SANGYEOB LEE

has been accepted towards fulfillment
of the requirements for the

_____Doctoral_____   degree in   _Media and Information Studies_

_____
Major Professor's Signature

_____December 9, 2009_____

Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

THE EFFECT OF RPG NEWNESS, RATING, AND CHARACTER EVILNESS ON
THE NPC BELIEVABILITY

By

Sangyeob Lee

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILSOPHY

Media and Information Studies

2009

ABSTRACT

THE EFFECT OF RPG NEWNESS, RATING, AND CHARACTER EVILNESS ON
THE NPC BELIEVABILITY

By

Sangyeob Lee

Characters in most role-playing games often are not at all plausible. Non Player

Characters (NPCs) in role-playing games may seem at first glance to be complex, robust

beings. However, it doesn't take long to notice that their behaviors are extremely limited.

Computer scientists working on Artificial Intelligence have recognized the importance of

believable characters. Studies on the definition of character believability have been

relatively rare, and the definition of believability has different meaning in different

disciplines. In this paper, a new definition of character believability is proposed, based

on several theoretical frames. This definition is compared with other definitions of

believability in diverse areas. A multi-disciplinary literature review explores various

qualities of computer characters different scholars and practitioners have described that

are expected to contribute to those characters being believable. These attributes are

categorized into five believability qualities. Overall NPC believability and specific

believability within those five qualities is investigated.

Eight NPCs were chosen for the study based on game quality (high and low Metacritic

scores), modernity (games released before and after 2006) and game function (good guy

or bad guy). The believability of each character was surveyed by research subjects

recruited from a sophomore level introductory digital media course at a large Midwestern

university with an online survey. Within the survey participants were shown a 2 to

3 minute video of each NPC interacting with a player, and then answered believability questions about the NPC they had just seen. Role play gaming experience and demographics were also measured. ANOVA analysis of the results suggested that NPCs from newer and/or high rated games were perceived to be more believable than characters from older or low rated games. NPCs from newer and/or high rated games were also found to be more enjoyable than characters from older or low rated games.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Characters in games are often not plausible at all. For example, the owner of a pizza house in Grand Theft Auto waits to take an order when the player character enters the restaurant. However, the owner shows no personality or emotion when he waits for the player to pay for pizza. He doesn't even try to run away or ask for help if the player beats him to steal his cash. At first glance, Non Player Characters (NPCs) seem to do what they are supposed to do. However, it doesn't take long to notice that their behaviors are extremely limited. A simplistic NPC may actually prevent gamers from feeling fully immersed in game experience. For example, NPCs whose role in the game is to give a hint or direction often don't do anything other than wait for players to approach to them. Once a player enters an NPC's activation boundary, the NPC approaches the player gives the hints, often regardless of whether the player requests the information or not. Then the NPC returns to the exactly the same position and resumes their waiting-posture. If the same player approaches the NPC again, the NPC merely repeats the same hint. Although NPCs may look like a player avatar, their behavior clearly marks NPCs as artificial and limited. Because they look like player characters but act like machines, NPCs are usually not believable participants in the game world.

Artificial Intelligence researchers working on interactive drama systems have recognized the importance of believable characters (Aylett, 1999; Mateas, 1997; Szilas, 2003). In an interactive drama, players can participate in a story in a virtual world that is populated with computer-controlled characters from a first person perspective. In both the interactive dramas and games, appealing and interactive parallel the importance of characters in fiction. Making game characters believable matters because their perceived

believability is expected to increase players' feeling of immersion and their enjoyment of the game.

Computer scientists have proposed a relationship between the believability of characters and the level of immersion players experienced (Bhatt, 2004; van Doom & de Vries, 2006; Watson, 2002). Believability has also been linked to enjoyment (Brown & Cairns, 2004). Whether they are player characters (controlled by players) or NPCs (controlled by the game engine), creating more believable characters are likely to result in better game experiences.

LITERATURE REVIEW

What is Believability?

What does it mean that something is believable? This is the question that researchers have tried to answer in many contexts. For example, in a study of workplace gossip, Berkos (2003) defined believability as "the perception of truth and the absence of lies" (p. 13). Beltramini (1982) measured the believability of cigarette warning labels using scales such as trustworthiness and credibility. Believability has been studied in relating to advertising. In studies of alcohol warning labels, Andrews, Netemeyer and Durvasula found that the user characteristics (frequent users or occasional/non users) and the alcohol warning label message (labels warning about birth defects, driving impairment, hypertension, drug combination, and addictive nature) have significant impact on the believability and the attitude on labels (1991). Source credibility is a related topic from communication and journalism research. When people decide whether they will accept or reject the information they are presented with, they make their judgment based on the "credibility" of the messenger. Hovland, Janis and Kelly (1953) theorized that source credibility is a function of both the perceived expertise and trustworthiness of messenger. Ewing (1940)'s foundational study of source credibility looked at how the opinion-change occurs when the message is consistent or inconsistent with the prejudice of the audience as well as the characteristic of the message. He found the greater the difference of opinion between a messenger and audiences, the greater the change of the opinion might be. The Ewing's findings about opinion change and the discrepancy between messenger and audience has been confirmed by later studies (Brehm & Lipsher, 1959; Pastore & Horowitz, 1955). In general, research shows that the

credibility of a *messenger* is necessary for acceptance of the *message* the messenger delivers.

However, the believability of characters in a game is not necessarily related to truthfulness or credibility. Mateas (1997) points out that a believable agent may NOT be trustworthy and may possibly not tell the truth. Characters are *"artistic abstractions* of people, whose behavior, motivations, and internal life have been simplified and exaggerated in just such a way as to engage the audience in the artist's vision" (Mateas, 2002, p. 8).

Believability has been studied in relation to the arts (literature, theater, film, radio drama, etc). Traditional animators have described how they achieved the believability in the characters they drew. Chuck Jones, master of animation who created many Warner Brother's animation characters such as Bugs Bunny, Daffy Duck, Elmer Fudd and Porky Pig stressed the importance of believability in bringing a character to life (1989). Animators at Disney have tried to illustrate their characters as entities that can think and act by their own volition to create the illusion of life (Thomas & Johnson, 1981). The goal of these early animators was to create the illusion of life, facilitating the viewers' willing suspension of disbelief. Modern computer agent designers borrowed the practices of these early animators to make their character believable. However, interactive computer characters pose additional challenges. Unlike watching TV cartoon animations, the interaction between users and computer agents is bidirectional.

Believability is a key goal in creating embodied agents (Lester & Stone, 1997; Lester, Voerman, Towns, & Callaway, 1997; Nijholt; Ortony, 2002). It has been, Especially, the main quality of interactive entertainment applications such as computer

games (Bhatt, 2004; Mac Namee & Cunningham, 2001) and interactive drama (Mateas, 1997). Namee and Cunningham (2001) addressed the problem of static NPCs by creating autonomous, proactive, persistent NPCs, which have desires of their own. The problem remains in choosing what kinds of desires and behaviors might make an NPC more believable. The Oz Project at Carnegie Mellon University identified believability as an important AI goal. They defined a set of "requirements for believability" that includes personality, emotion, self-motivation, change, social relationships, and the illusion of life (Loyall, 1997, p. 15).

Communication and psychology researchers have studied presence, a concept related to believability. Creating a sense of presence has a core goal of mediated communication systems, telepresence, virtual reality and games, and linear mass media such as television, movies, and radio. Presence researchers have developed definitions and measurement constructs (Barfield, Zeltzer, Sheridan, & Slater, 1995; Lombard, Ditton, Crane, & Davis, 2000). According to Schroeder (2002), presence describes a medium's capability of evoking the feeling that a human user is physically present in a virtual environment. Heeter defined three dimensions of presence as a feeling of being there in "her" study of virtual reality; personal presence (I am there), social presence (you are here), and environmental presence (the virtual environment exists) (1992). Naturalness (believability) is one of the four sub concepts of presence along with spatial presence (being in a physical space) and engagement (user's interest) which contribute to a sense of presence, and negative effects (headache) that interfere with presence (Lessiter, Freeman, Keogh, & Davidoff, 2001). The main difference between believability and the other presence sub concepts lies in the fact that believability depends upon intentional

cognitive involvement while other dimensions are perceptual rather than judgmental. It seems that the believability makes the feeling of presence smooth by taking care of the cognitive part of it. Believability is associated with *willing suspension of disbelief* which is not automatic while belief, for example, is automatic (Reeves & Nass, 1996).

Bates considered believability in "his" study on the role of emotion. He describes believability as providing "the illusion of life, thus permitting the audience's suspension of disbelief" (1994, p. 122). The origin of the expression, *suspension of the disbelief* can be found in a Samuel T. Coleridge's autobiography, Biographia Literaria.[1] Coleridge described the relation between reader and literature by expressing it as *willing suspension of disbelief.*

The phenomenon of willing suspension of disbelief arises when fictional or virtual characters, events, or worlds contradict what the audience knows to be real. For example, when people read a story such as Beauty and the Beast, for example, we do not doubt whether a beast actually could wake up from death when the beauty confesses her love toward him. Even though we know that human beings cannot return to life by the tear of a true love, we don't focus on doubting it when we are watching. We willingly suspend this disbelief that the dead can't return to life for the pleasure of the story. However, suspending disbelief does not mean that viewers believe everything they see or read; rather it means that they don't reject the story because they disbelieve what they perceive. Bhatt explained suspension of disbelief as a quality of a fabricated story enabling people to believe unrealistic facts for the enjoyment of reading (2004).

---

[1] http://www.bartleby.com/66/78/12878.html

Then, what is believability? Especially, what does it mean by believability when it comes to the study NPCs and computer agents?

Character believability is the size and nature of the cognitive gap between the character players experience and the character they expect. When the player's expectations exactly match their experience, a character is fully believable. The larger the gap, the more likely it is to interfere with suspension of disbelief.

Theoretical frameworks

Schemata is a kind of cognitive network of similar thoughts (Slavin, 1988). Similarly, Mandler described a schema as a unique and harmonious representation (1984). Each individual has many unique schemata depending on his or her experience and cognitive ability. The concept of schema has been used by cognitive scientists and psychologists to study how humans interpret, remember information they encounter in daily life in relation to their previously developed schemata (Armbruster, 1986; Ausubel, 1967; Bartlett, 1932; Duis, 1996).

The theory of schema explains well how the human brain perceives, interprets and constructs knowledge of the world. The basic idea is that people remember new information by relating this information to the pre-existing information (schemata). Thus, they use schemata to make sense of the world. Schemata make it possible to make prediction about their next behavior. Schemata are not conscious. How we store and process information occurs at a subconscious level. However, hints at the schemata people accumulated through their daily life can be revealed in various attitudes and behaviors such as stereotypes, social roles, etc. When new information doesn't fit into

7

the existing schemata, the new information may not be comprehended correctly or cannot be comprehended at all.

Psychologists, Rumelhart and Norman, identified three possible effects of new information on people's existing knowledge structure: accretion, restructuring and tuning (1978). Accretion occurs when new information fits well into the existing schema. The schema structure remains unchanged as the new information is added. Restructuring happens when a pre-existing schema can't explain new information, and people have to change their schema to accommodate the new information. Tuning effect describes a situation when people use new, somewhat contradictory information in order to tune or modify an existing schema. Nowak and Biocca investigated differences in how people respond to different entities (agents, controlled by a computer and avatars, controlled by person) as well as different levels of anthropomorphic visual representation (high-anthropomorphic, low-anthropomorphic, and no image) with three different forms of presence (Telepresence or a sense of being there, Copresence, or a sense of being with and Social presence, a sense of interpersonal interaction (2003). They found that whether the entity was an agent or avatar didn't influence the feeling of presence in general. An entity represented by an image resulted in a stronger experience of presence than entities with no image. One interesting finding of Nowak and Biocca's study is that the subjects in low-anthropomorphic condition experienced higher telepresence than those in other conditions (high-anthropomorphic and no image condition). Also, there was no difference in perceived telepresence between entities with high anthropomorphic image and those with no image. These results support the notion that by default people tend to assume any entity is human when sparse information is provided. There is an underlying

assumption that someone else is "like me" unless information contradicts that assumption.

When people form impressions of others, they construct mental models of new people based on people they already know (Gordon, 1986). Users of a low bandwidth communication system tended to think the distant others they were communicating with were like them because of a lack of cues to contradict that assumption, whereas users of a higher bandwidth system could perceive more detail and recognized more differences between themselves and the distant others. High bandwidth revealed differences, whereas low bandwidth carried little information resulting in greater reliance on default schemas (Walther, 1996). People "filled in the blanks" and assumed similarly to themselves. People first try to apply one of their existing schemata to interpret a character or situation they encounter, and modify or create a new one only if familiar schemata are a poor fit.

Uncertainty Reduction theory (URT) explains how humans interact with strangers based on their schemata. Berger and Calabrese (1975) developed URT in order to explain a role of communication forming new interpersonal relationship. They borrowed the concept of uncertainty from information science where it was used to describe data transmission between machines (Shannon & Weaver, 1963). According to the URT, uncertainty is unpleasant, and people try to avoid or reduce it when they encounter strangers. Berger and Calabrese proposed three stages of interactions when people encounter strangers: an entry phase, a personal phase, and an exit phase. During the entry phase, people share very basic information such as sex, age, socioeconomic status, etc. In the personal phase, more personal information will be shared including attitudes,

beliefs, values, etc. During the exit phase, people decide if they want to continue to have a relationship with the stranger. The three interaction steps may apply to the interaction with a strange computer generated character as well. The entry phase is especially important because the believability of character will be judged by the character's qualities shown in entry phase.

Social Response to Communication Technology perspective (also known as the "media equation") posits that people react very similarly to social cues from humans and social cues from media acting like a human (Nass, Fogg, & Moon, 1996; Nass & Moon, 2000; Nass, Moon, & Carney, 1998; Reeves & Nass, 1996). For example, people are even polite to computers. When asked to evaluate performance of a computer, people tended to give more positive feedback about the computer if the computer they were evaluating was the same computer asking the evaluation questions. If instead they answered the evaluation on a different computer, people gave less polite responses, perhaps because they were not as concerned about hurting the original computer's feelings.

Humans seem to construct expectation about an entity they are interacting with by evaluating the level of perceived realism. Japanese robotist Mori, introduced the concept of an "uncanny valley" in human reactions to anthropomorphic robots (1970). According to Mori, as the realistic human likeness of a robot is increased, human attraction to and familiarity with the robot will increase, but only up to a certain point. After this point, attraction is replaced by fear, unease, or revulsion created by a robot that appears to be, but is not quite, human-like. Eventually, in theory, as the human likeness level keeps increasing, the human perception of the robot will rebound and response to the robot will

approach the level of attraction to a healthy person. The area where attraction plummets

and is replaced by unease is called the "uncanny valley."
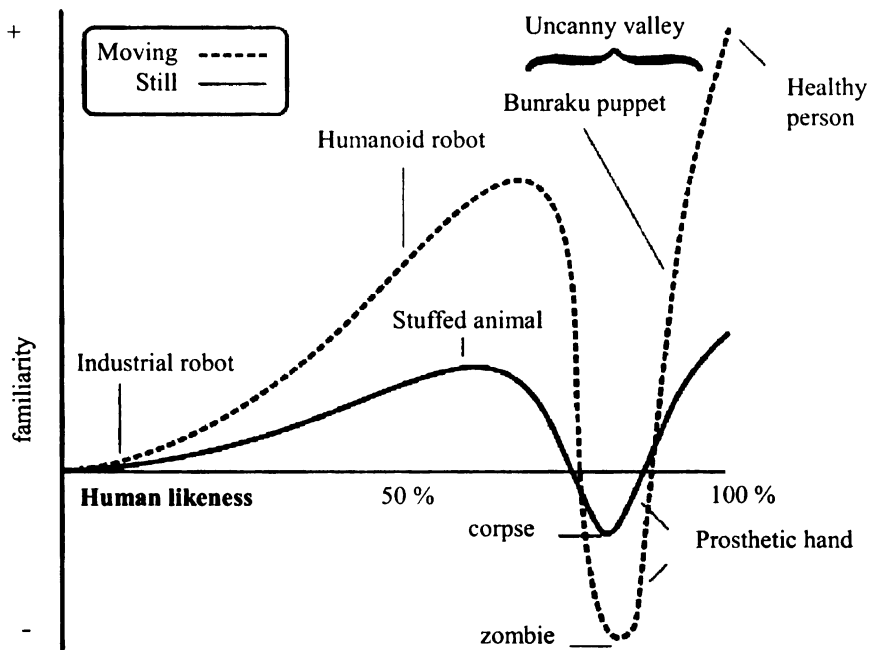
Figure 1. Uncanny Valley

Game designer and journalist Clive Thompson (2004) suggests that the "Uncanny

Valley can make games less engrossing." Mori suggested the peak of the hill right

before the uncanny valley should be the target for robot design.

BELIEVABILITY QUALITIES

Many studies in computer science address the goal of creating believable agents. The qualities as well as meaning of character believability varies by field of study. A comprehensive literature survey can help to combine and categorize diverse believability paradigms.

In this paper, many character believability qualities were investigated in many related fields such as computer science and communication. These qualities are grouped into several categories that have similar concepts. The analysis of believability qualities produced 5 key believability categories: appearance, personality, goals, emotions, and social relations. Each of these qualities is expected to contribute to overall general believability. Certain qualities may be more important to general perceived believability than others.

1. Appearance

Table 1. Characters Qualities Related to Appearance

Quality	Author	Context
Behavior	Hayes-Roth & Doyle 1998	7 behavioral requirements of believability in animate characters
Appearance, Identity, Manner of gesture, Manner of speak, Content of speech	Hayes-Roth, Maldonado & Moraes 2002	10 key qualities in animate characters
Behavioral cues	Sengers 1999	Comprehensive agents (comprehensibility)
Situated Liveness, Controlled visual impact, Complex behavior pattern	Lester & Stone 1997	Three believability criteria in Animated pedagogical agents
Consistency in expression, appearance of goals	Loyall 1997	Requirements of believable agents
Language	Loyall and Bates 1997	Requirements of believable agents
Lack of ambiguity in expression	Letster et al 1997	Deictic believability in animated pedagogical agent

Computer scientist Hayes-Roth and colleagues described the appearance of an animated character as "the encoding of each characters' identifying demographic information – age, weight, gender, socioeconomic background and culture in the chosen embodiment of the character, as well as the representation of this embodiment" (Hayes-Roth, Maldonado, & Moraes, 2002, p. 7). According to the definition, character appearance not only includes static, visible attributes, but also animated gestures and nonverbal behaviors (Hayes-Roth & Doyle, 1998). Believability qualities that can be perceived visually are included in the character appearance category.

According to Hayes-Roth, behavior animation is the core quality of believability of an animated character (Hayes-Roth & Doyle, 1998). In defining seven requirements of believable behaviors in animated characters, Hayes-Roth and Doyle specify that the behaviors have to be diverse to offer many possible scenarios but normative to appear life-like, ambiguous enough to invite different interpretations, and unique enough to distinguish one animated character from all others. Hayes-Roth and colleagues propose that manner of gesture (for example, facial expression and hands), manner of speech (for example, intonation and pronunciation), content of speech (for example, idiomatic expression), identity and appearance are key qualities of a believable animated character (Hayes-Roth, et al., 2002). In the study of constructing comprehensibility as an essential factor of a believable agent, Sengers said "The agent's comprehensibility comes from thinking out the connections between behaviors and displaying them to the users" (1999, p. 2). Similarly situated liveness, controlled visual impact, and complex behavior pattern are suggested as techniques for increasing the believability in animated pedagogical agents (Lester & Stone, 1997). Situated liveness means that the behaviors of computer

13

agents should be ongoing to show the alertness. Controlled visual impact calls for a carefully chosen magnitude of character behavior, not to distract users' attention. For example, moving from one location to another is visually bigger and has more impact than blinking an eye. Also, the behavior pattern of a computer character should be complex enough to avoid creating an impression of simplicity. The techniques for increasing believable behaviors suggest that the representations of computer agent behaviors mimic life. In the same context, consistency in expression and appearance of goals (Loyall, 1997) and lack of ambiguity in expression (Lester, et al., 1997) have been proposed as techniques for increasing the believability in agents.

The character appearance category describes all the qualities that are exposed to the human sensory input including visual hints that communicate information such as gender, age, ethnicity, height, socioeconomic status, etc. When people look at a character, they make guesses on not only what kind of character it is but also how they can interact with it based on context and the information perceived by sensory input.

However, none of the qualities in the character appearance category implies that having a human appearance is necessary for believability (Loyall, 1997; Mates, 1997). According to Mateas, non-human computer agents also could be useful because of lower expectation compared to anthropomorphic agents. However, the lowering of expectation doesn't come entirely from the non-humanness of agents. Lower expectations are also related to lower level of fidelity and realism of the agent regardless of whether it is human-like or not. Comic artist Scott McCloud (1994) notes that the more visually realistic a comic character is, the harder it is for people to think of that character as being like themselves. A smiley face could be anyone, including the player. But a photorealistic

Arnold Schwarzenegger is clearly not the player. Echoing McCloud's sense that seeing oneself in a character is different than seeing someone else in a character, Bailenson et al. (2001) studied user reactions to an agent that had the users' own photorealistic face. They found that people treated agents embodied with their own face (a virtual self) fundamentally differently than agents embodied with a stranger's face in regards to both measured nonverbal behaviors and questionnaire ratings. Users were more intimate with their virtual self, standing closer and expressing less reluctance to engage in embarrassing behaviors in the presence of their virtual self than in the presence of a virtual other.

As research suggests, qualities in the character appearance category can be representations of other believability qualities. For example, the identity quality in the character appearance category can reflect another believability quality, personality. From this perspective, all visual and audio information human users can perceive are closely related to the other key qualities of believability. The appearance of computer characters should be carefully designed with consideration of other related believability qualities. In this sense, it is very important to design the appearance of NPCs based on their role in specific settings. Appearance is usually perceived first, and players establish a kind of expectation based on information from an NPC's appearance. Even in a case when the information from appearance is very limited, users construct some kind of image of the character based on the appearance. Hayes-Roth, Maldonado, and Moraes stressed the importance of the appearance by saying

"Appearance affects the character's effectiveness and credibility at performing their assigned role, and directs the patterns of interaction. Even before the character speaks a single word, even before the page is completely loaded, the

visitor has already processed the subliminal cues embedded in the characters' representation, such as the relative status and occupation of the interactors, and formed a model of what pattern the ensuing interaction will follow" (2002, p. 8).

In order to create a believable agent, the character appearance must do something more than simply reveal demographic information. Rather, character appearance should suggest other believability qualities such as emotion and personality. This leads to the following hypothesis:

H1: A character's appearance will be closely related to that character's general perceived believability.

2. Personality

Table 2. Character Qualities Related to Personality

Quality	Author	Context
Personality	Drennan 2004	Believable conversational NPC
	Rousseau & Hayes-Roth 1997	Believable synthetic actors
	Mateas 1997	Requirement of believability
	Loyall 1997	Requirement of character believability
	Bates, Loyall, Reilly, Castelfranchi, Werner 1994	Believable agent
	Romano,Wong 2004	Believability of virtual character
	Reilly, Bates 1995	Believable social agents
	Reilly 1997	Believable social agents
Persona	Hayes-Roth & Doyle 1998	Believable animate character
Identity	Hayes-Roth, Maldonado and Moraes 2002	10 key Believable qualities in animate characters
Back story	Hayes-Roth, Maldonado and Moraes 2002	10 key Believable qualities in animate characters
Goal based personality	Rizzo et al 1999	Personality and social behavior in believable agent

Personality has been one of the most critical aspects of believability for cartoon animators. Thomas and Johnson (1981) describe the importance of the personality. "For a character to be that real, he must have a personality, and, preferably, an interesting one" (p. 19). More recently personality has been suggested as one of the most critical factors in

creating believable agents (Allbeck & Badler, 2002; Bates, Loyall, Reilly, Castelfranchi, & Wemer, 1994; Reilly, 1997; Romano & Wong, 2004).

The Oz project at Carnegie Mellon University was the first attempt to develop a believable agent in an interactive story environment. In the Oz project, computer scientist Loyall (1997) defined personality as "all of the particular details – especially details of behavior, thought and emotion – that together define the individual" (p. 16). Similarly Hayes-Roth et al. defined personality as a group of psychological characteristics that differentiate one entity from others (Hayes-Roth, van Gent, & Huber, 1997).

Rousseau and Hayes-Roth proposed a computer agent model whose actions are driven by its personality and mood rather than its goal to produce a more dramatically believable and interesting character (1997).

Mateas, another researcher of the Oz project, defined personality as the thing that inspires every single behavior of a character from very simple behavior, such as talking to very cognitive activity, such as reasoning (1997). To him, the personality is "something unique and specific"(1997, p. 6). Mateas discusses other qualities of believability such as emotion and change but stresses they must be consistent with the personality of the character.

Hayes-Roth, Maldonado and Moraes used the term backstory to refer to any kind of (fictional) past individual experience that contributes to personality while the term identity indicates a slightly different quality which exists in the overlap between character appearance and personality (2002).

Personality defines the uniqueness and peculiar qualities of computer characters that distinguish them from other computer characters. Some qualities of personality are closely related to the psychological traits. Goldberg proposed the "big five" personality traits (agreeableness, extraversion, neuroticism, conscientiousness and openness to experience) through empirical study of the human personality (1993). Some computer scientists have tried to construct the personality on synthetic characters by giving a unique combination of the parameters that constitute basic personality traits as Goldberg suggested (Bates, 1992; Rousseau & Hayes-Roth, 1997). Rizzo et al. designed goal-based personalities (Rizzo, Veloso, Miceli, & Cesta, 1999). It seems that the personality cannot be conveyed by itself and can only be only revealed by some other mechanics such as appearance, emotion, etc. This leads to the following hypothesis:

H2: A character's personality will be closely related to that character's general perceived believability.

3. Goals

Table 3. Character Qualities Related to Goals

Quality	Author	Context
Goals	Bates 1994	Believable agents
Self-motivation	Loyall 1997	Requirement of character believability
	Mates 1997	Requirement of believability
Role, Role dynamics	Hayes-Roth, Maldonado and Moraes 2002	10 key Believable qualities in animate characters
Goal preference	Rizzo et al 1999	Personality and social behavior in believable agent
Intentionality	Riedl and Young 2005	Character believability

Hinting at thought processes through visual, nonverbal means was not easy to implement for early animators. They knew expressing the thought process of cartoon characters was important but didn't know how to reveal it. According to Thomas and

Johnston (1981), it was the animation of a dog that looked into the camera and snorted that gave them the idea of making cartoon characters appear to think.

Loyall (1997) insisted that a self-motivated character should not only appear to think, but also have to show the emotion of its own volition. He wrote, "Pluto snorting was not what was powerful; it was that he was doing it of his own accord, instead of in reaction to some external stimulus" (Loyall, 1997, p. 20).

Character intentionality was suggested as an important aspect of character believability. According to Riedl and Young, character intentionality is "...the way in which the choice of actions and behaviors that a character makes appears natural (and possibly rational) to external observers" (p. 2).

Goals are one of the most important fundamentals for building believable agents because goals directly affect all other qualities of believability. For instance, the outfit of an agent should be designed carefully with a consideration of its role. If it is a kind of tutoring character that teaches cooking, for example, it may be more natural for a player to expect the agent to wear a high cook's hat and white colored chef's coat. Like personality, goals are a precursor to other believability traits - they help to define appearance, emotion, and behavior (Bates, 1994; Hayes-Roth, et al., 2002; Rizzo, et al., 1999). This leads to the following hypothesis:

H3: The goals of a character will be closely related to that character's general perceived believability.

4. Emotions

Emotion is one of the two most important qualities of believable agents (Bates, et al., 1994; Hayes-Roth & Doyle, 1998; Hayes-Roth, et al., 2002; Loyall, 1997; Reilly & Bates, 1995; Romano & Wong, 2004).

Table 4. Character Qualities Related to Emotions

Quality	Author	Context
Emotion(s)	Drennan 2004	Believable conversational NPC
	Ortony 2002	Believable emotional agent
	Loyall 1997	Requirement of character believability
	Mateas 1997	Requirement of believability
	Bates 1994	Emotions in believable agents
	Wooldridge and Jennings 1995	Believable agents in the context of intelligent agents
	Bates, Loyall, Reilly, Castelfranchi, Wemer 1994	Believable agent
	Romano,Wong 2004	Believability of virtual character
	Reilly, Bates 1995	Believable social agents
Empathy	Hayes-Roth & Doyle, 1998	Believable animate character
Emotional dynamics	Hayes-Roth, Maldonado and Moraes, 2002	10 key Believable qualities in animate characters
Emotional reaction	Rizzo et al, 1999	Personality and social behavior in believable agent
Emotional reactivity	Bates 1994	Suggested as believability demands on an interactive character

The early animators, Thomas and Johnson (1981) said, "From the earliest days, it has been the portrayal of emotions that has given the Disney characters the illusion of life" (p. 505). They especially stressed several considerations to convey the emotional state of the character: clear definition of an agent's emotional state, revealing its thought process through the emotions, and use of time to emphasize the emotion.

Bates (1994) first adapted Thomas and Johnson's principles of revealing cartoon character emotions to apply to the design of emotions of "Woggles" (self-animating creatures in a simulated world that was a small portion of the Oz project). In order to

20

define their emotional states clearly, Woggles only have one or two primary emotions with appropriate intensities. Also Woggles' emotions were mapped to specific behaviors of a character with a specific personality to reveal its thought process. In other words, one emotion could be mapped to two different behaviors based on the personality of characters. However, Bates didn't use animation techniques (for example, exaggerating) to emphasize Woggles' emotions. Bates described the value of character emotion "...it helps us know that characters really care about what happens in the world, that they truly have desires" (1994, p. 3). Emotions signal an agent's aliveness, creating the illusion of life. Bates (1994) considered the emotional reactivity essential to believability in interactive character design.

Ortony re-categorized the conditions in which emotions can be generated and suggested five positive and negative conditions for believable agents (2002). Positive conditions *(the first entry is the undifferentiated reaction):*

- because something good happened (joy, happiness etc.)

- about the possibility of something good happening (hope)

- because a feared bad thing didn't happen (relief)

- about a self-initiated praiseworthy act (pride, gratification)

- about an other-initiated praiseworthy act (gratitude, admiration)

Negative conditions *(the first entry is the undifferentiated reaction):*

- about the possibility of something bad happening (fear, etc)

- because a hoped-for good thing didn't happen (disappointment)

- about a self-initiated blame worthy act (remorse, self-anger, shame, etc)

- about an other-initiated blameworthy act (anger, reproach, etc)

- because one finds someone/thing unappealing or unattractive (hate, dislike, etc.)

He pointed out the importance of consistency between internal responses (emotions) and external responses (behaviors) in believable emotional agents. Mateas (1997) also suggested that the emotions of a believable character should be expressed in unique ways (akin to and supporting the uniqueness of personality). Wooldridge and Jennings also cite emotion as a key component of believable agents that makes them appear to respond to human emotions in appropriate and consistent ways (1995).

The emotion category doesn't prescribe what kinds of emotions believable characters have to possess. Rather all of the research literature about emotions suggests that it is important for believable agents to clearly exhibit emotions of their own. The emotion category requires believable agents to reveal their emotions as an outcome of unseen internal processes. At the same time, believable agents should respond to players' emotions correctly in a given context.

However, it is not easy to show a right expression of a right emotional state in right time. Unlike cartoon characters who can express their emotional states based on pre-scripted, linear narratives, computer agents' expressions of emotional states are hard to plan because their emotional responses should be impacted by interactions with players. NPC character designers must design a range of possible emotional states anticipating specific circumstances. Each actual emotional state of a computer agent should be appropriate for the circumstance under which it is enacted so that players do not experience cognitive dissonance.

The explicitness of emotion needs to be clear enough that players can distinguish each emotion at a glance. The expression of an agent's emotional state should be neither

ambiguous in meanings nor weak in strength. Emotional states should be revealed through multiple channels such as facial expression, gestures, etc. In some cases, for example, it may be hard to notice emotions only by looking at the facial expression of a computer agent either because the face is too small or the face may be looking away. Body language and motion can help convey character emotion. This leads to the following hypothesis:

H4: A character's emotions will be closely related to that character's general perceived believability.

5. Social relations

Table 5. Character Qualities Related to Social Relations

Quality	Author	Context
Social-relationship	Loyall 1997	Requirement of character believability
	Mateas 1997	Requirement of believability
Situated Social competence	Bates 1994	Believable agents
Social relation	Hayes-Roth & Doyle 1998	Believable animate character
Social interaction patterns	Hayes-Roth, Maldonado and Moraes 2002	10 key Believable qualities in animate characters

The early animators didn't recognize characters' social relationships at first because different animators drew each character in a cartoon. When one animator drew all the characters in a scene, the importance of character relations was recognized. Thomas and Johnston (1981) wrote:

...the Bambi and Thumper sequence had something that the Pluto and Donald

sections did not have. That was a character relationship with strong beginnings in

the story department. ...With this as a springboard, the animator continued

developing this relationship, which only could have been done by one person

handling both characters and completely controlling every single bit of action, timing, and cutting. …This new way of working with character relationships encompassed the whole range of relations between two or more characters–from the broadest to the most delicate. It involved expression scenes that often registered the most secret thoughts and inner emotions of the characters, which as they became more subtle were also more revealing (p. 164).

Many studies on believable agents described social relationships among computer characters that influence an interaction and are influenced by the interaction in turn among characters (Bates, 1994; Mateas, 1997; Thomas & Johnson, 1981). Some studies described a social aspect to the interaction between computer characters and players (Hayes-Roth & Doyle, 1998; Hayes-Roth, et al., 2002).

Loyall especially insisted that social relationships among characters should be designed carefully with detailed behaviors and interactions to reveal the relationships among characters (1997). Hayes-Roth, Maldonado and Moraes found that the context of the interaction is even more important than the actual content itself (2002). According to them, cultural differences, the tempo of turn taking in conversation, initiatives of the conversation, etc. can be more important to create believable agents than social relations between characters, especially in a one to one interaction environment with only the player and a single character.

A character that appears to have a social relationship with other characters helps players willingly suspend their disbelief and conceive of the character as real. Of course, social relationships between NPCs cannot exist in environments limited to one-to-one

interaction between the player and an isolated NPC. Most interactions with pedagogical agents, for example, only occur directly between the agent and human.

The social relationship between a computer agent and the player also can affect the perception of believability. Nass et al. found that people feel more attraction and trust with a computer character that has the same ethnic background as the user than with one that has a different ethnic background (2000). The more users feel a social connection with the computer characters, the more believable the character will seem.

In role-playing game environments that have multiple NPCs, the social relation among NPCs should enhance yet not overwhelm or otherwise interfere with the interaction between players and NPCs. Both NPC-NPC and NPC-player social relations should seem natural. This leads to the following hypothesis:

H5: A character's social relations will be closely related to that character's general perceived believability.

Summing up all the character believability qualities above, Table 6 shows the complete list of character believability qualities derived from the literature.

Table 6. Five Believability Qualities

	Quality	Author	Context
Character Appearance	Behavior	Hayes-Roth & Doyle (1998)	7 behavioral requirements of believability in animate characters
	Appearance, Identity, Manner of gesture, Manner of speak, Content of speech	Hayes-Roth, et al (2002)	10 key qualities in animate characters
	Appearance of reactivity	Bates (1994)	Suggested as believability demands on an interactive character
	Behavioral cues	Sengers (1999)	Comprehensive agents (comprehensibility)
	Situated Liveness, Controlled visual impact, Complex behavior pattern	Lessiter & Stone (1997)	Three believability criteria in Animated pedagogical agents
	Consistency in expression Appearance of goals	Loyall (1997)	Requirements of believable agents
	Language	Loyall & Bates (1997)	Requirements of believable agents
	Lack of ambiguity in expression	Letster et al (1997)	Deictic believability in animated pedagogical agent
Personality	Personality	Drennan (2004)	Believable conversational NPC
		Rousseau & Hayes-Roth (1997)	Believable synthetic actors
		Mateas (1997)	Requirement of believability
		Loyall (1997)	Requirement of character believability
		Bates, et al (1994)	Believable agent
		Romano,Wong (2004)	Believability of virtual character
		Reilly, Bates (1995)	Believable social agents
		Reilly (1997)	Believable social agents
	Persona	Hayes-Roth & Doyle (1998)	Believable animate character
	Identity	Hayes-Roth et al (2002)	10 key Believable qualities in animate characters
	Back story	Hayes-Roth et al (2002)	10 key Believable qualities in animate characters
	Goal based personality	Rizzo et al (1999)	Personality and social behavior in believable agent
Goals	Goals	Bates (1994)	Believable agents
	Self-motivation	Loyall (1997)	Requirement of character believability
		Mates (1997)	Requirement of believability
	Role, Role dynamics	Hayes-Roth et al (2002)	10 key Believable qualities in animate characters
	Goal preference	Rizzo et al (1999)	Personality and social behavior in believable agent
	Intentionality	Riedl & Young (2005)	Character believability

Table 6 cont'd

Emotions	Emotion(s)	Drennan (2004)	Believable conversational NPC
		Ortony (2002)	Believable emotional agent
		Loyall (1997)	Requirement of character believability
		Mateas (1997)	Requirement of believability
		Bates (1994)	Emotions in believable agents
		Wooldridge & Jennings (1995)	Believable agents in the context of intelligent agents
		Bates et al (1994)	Believable agent
		Romano,Wong (2004)	Believability of virtual character
		Reilly & Bates (1995)	Believable social agents
	Empathy	Hayes-Roth & Doyle (1998)	Believable animate character
	Emotional dynamics	Hayes-Roth et al (2002)	10 key Believable qualities in animate characters
	Emotional reaction	Rizzo et al, (1999)	Personality and social behavior in believable agent
Social relations	Social-relationship	Loyall (1997)	Requirement of character believability
		Mateas (1997)	Requirement of believability
	Situated Social competence	Bates (1994)	Believable agents
	Social relation	Hayes-Roth & Doyle (1998)	Believable animate character
	Social interaction patterns	Hayes-Roth et al (2002)	10 key Believable qualities in animate characters

RESEARCH QUESTIONS AND HYPOTHESES

The qualities that have been insisted to create believable agents were examined and categorized into five different categories: appearance, personality, goals, emotions, and social relations. However, these believability qualities are not necessary conditions for character believability. Humans require very few cues in order to react socially to a computer (Reeves & Nass, 1996). Also researchers found that the social presence could be triggered by only minimum intelligence (Biocca, Harms, & Burgoon, 2001). One difficulty in applying the believability qualities to NPCs is that the individual believability qualities are not independent: they are inter-related to each other. This is why two different believability qualities in one computer character don't guarantee doubled-believability. Often, character believability can be established mainly by one or two distinct individual believability qualities. Perceived character believability may be less than the sum of each individual believability quality if the believability qualities are not designed with a combing principle such as goals and/or personality. In other words, an individual believability quality may have a negative effect on the total character believability if it is in conflict with another main principle. This leads to the following research question:

> **RQ1: How will the general believability level relate to the five believability qualities? Will the general believability be established by one or few believability qualities?**

> **RQ2: If so, how will other believability qualities affect the general believability? Will lower perceived believability of specific believability qualities be associated with lower general believability?**

28

Gaming for entertainment is a voluntary activity. Players tend to have preferred genres they play often and other genres they avoid (for example, see 2008 Pew Foundation tables comparing the frequency of genre play among female and male teenagers). In a study of strategic difference in a computer game with 76 elementary students, Hong and Liu found that students in the expert group used more analogical approaches while the novice group repeated a pattern of trials and errors (2003). Similarly, the difference in visual attention to a first person shooter game was studied. Researchers found that experts who played more than eight hours a week showed faster and more accurate responses to the game than the novice players who played less than 30 minutes a week (Smith, Tsai, Wong, Brooks, & Peterson, 2008). Higher sport skills and better understanding of game situations (Blomqvist, Luhtanen, & Laakso, 2000) and better problem solving skill (DeVane & Durga, 2008) were found in the expert's group. Considering the above performance difference, it is expected that the perception of believability is different between experts and novices. Returning to believability, those who often play Role Play Games are experienced with the genre. Through experience, they have developed expectations about what to do and the role of NPCs within a game. Others who never play RPGs do not have those expectations. Experts' more developed schemas about RPG NPCs may be associated with higher overall believability ratings. This leads to the following hypothesis:

> **H6: RPG players will rate NPC believability higher than those who are not experienced with RPGs.**

Regarding human perception on character believability, time seems be an important factor. Certain believability qualities can be perceived faster and with less ambiguity than other qualities. For example, the appearance of a character is easy to perceive and doesn't require much time or cognitive processing compared to perceiving personality or experiencing social relations. It is not clear how much each individual category contributes to the total character believability. If an interaction between players and NPCs is very short, appearance probably has a bigger impact on overall believability. Other categories such as personality can make more contribution to believability as the interaction unfolds over time.

Technology has been one of the important factors in making a game realistic. As technology developed, many factors in games such as character design and sound and video quality became more and more realistic/believable. Researchers found that many technological developments in game design such as sound, graphic, CPU speed, etc. has made games more realistic (Frauenfelder, 2001; Kramer, 1995; Loftus & Loftus, 1983; Newman, 2002). Moreover, studies on game environment such as sounds and graphics revealed that gamers liked a more realistic game environment, and male gamers liked these realistic settings more than female gamers did (Wood, Griffiths, Chappell, & Davies, 2004). Shapiro and Chock studied the relation between typicality and reality using video taped television programs. They found that the drama that contains more realism evokes more enjoyment (2003). Also, believability was insisted as one of the prerequisites of media enjoyment (Vorderer, Klimmt, & Ritterfeld, 2004). Usually new games adapt more advanced technologies than relatively old games. Concurrent with

technological advances, game designers continue to innovate and attempt to make better

games. This leads to the following hypothesis:

> **H7a: Characters in Newer games will be more believable than characters in old games.**
>
> **H7b: Characters in Newer games will be more enjoyable than characters in old games.**

The quality of games varies widely, whether the criterion is game sales or critical

acclaim. Many rating websites have published the quantified scores of all various media

forms such as movie, video games, TV, music, etc. based on various categories.

Technological improvement can also be seen in many good-rated games in general. The

ingredients for good video games have revealed various aspects of the video games

ranging from character to interface. Especially for RPG, deep world history design,

evolution of MOBs (evil NPCs) and believability principles of MOB behavior were

suggested as factors for making enhancing Massively Multiplayer Online Games

(Tychsen). Also, character was reported as one of the important elements with other

qualities in creating a believable world (Dormans, 2006). This leads to the following

hypothesis:

> **H8a: Characters in games with good rating will be more believable than characters in lower rating games.**
>
> **H8b: Characters in games with good rating will be more enjoyable than characters in lower rating games.**

Except several recent games, RPG generally put game players in a good guy

position completing a main quest by fighting against villain characters, or going through

adventures. Related to the game players' role in games, NPCs usually have two different roles: helpers or antagonists. All NPCs in First-Person-Shooter games are all villains trying to kill game players while NPCs in RPG are populated good guys, bad guys, and neutral characters. Good characters in RPG are characters that help game players by providing some hints or information that can be critical in completing quests while evil characters try to threat/kill hindering game players in finishing quests. It is assumed that evil characters appeared to be someone liver posting tangible threats than good characters because it might be easy for designers to build few attributes more vivid for evil characters rather than putting various important attributes to good characters. On the other hand, good characters are given relatively important roles than evil characters in general. It is assumed that character designers put more hours in creating good characters than evil characters because good characters perform more important roles than evil characters. This leads to the following research question:

RQ3a: How will the general believability level relate to the character good/evilness? Will the evil characters be more believable or less believable?

RQ3b: How will the enjoyment level relate to the character good/evilness? Will the evil characters be more enjoyable or less believable?

METHODS

To test the hypotheses, four RPG games were selected (two new and two old; two with high Metacritic scores and 2 with low Metacritic scores). Within each game, one antagonist NPC and one protagonist NPC were selected. 30 second digital videos of the NPC's interaction with a player were recorded. Study participants viewed each video, answered a series of believability questions about that character, and repeated this process for all eight characters.

NPC Selection

Different game genres include particular NPC roles and NPC-player interactions. For the character believability study with NPCs a game genre was needed that met some basic requirements of the study. First, the game should guarantee that game players' interaction time with NPCs is long enough to experience believability qualities that go beyond physical appearance, such as emotion and personality. A computer character's internal state can have three different layers (emotion, mood, and personality), and the effect duration for experiencing these layers were reported to be short, medium, and long respectively (Schaap & Bidarra, 2009). In some game genres such as first person shooter games, the interaction between game players and NPCs is unidimensional (shoot and be shot at) compared to other genres such as Role Play Games (RPGs) in which the player and the NPCs have more diverse roles and interactions. Similarly, the role of NPCs in racing games is very limited, having little effects on game directions, and the behavior patterns of the NPCs in the game is very simple such as just sitting next to your character in a car or driving other cars competing with you. On the other hand, RPGs and Massively Mutliplayer Online Games (MMOs) deeply integrate NPCs into gameplay.

Finally, it is important for the game to be a single player game. In a multiplayer game, live humans represented by avatars would co-exist in the game with NPCs. In those games the believability of the NPCs may suffer by comparison to real humans.

For the actual screening of game genres, the list of computer game genres in Metacritic was used. The Metacritic website (www.metacritic.com) uses a unique scoring system called Metascore to evaluate the quality of various electronic medium such as games, television, movie, music, etc. According to the website, the Metascore is "a weighted average of all of the scores assigned by individual critics to that movie, game, book or album" (citation: http://www.metacritic.com/about/scoring.shtml). All the game genres in Metacritic were scrutinized with the above screening rules. Many genres such as racing, first person shooter, etc. were eliminated by the first rule: guaranteeing enough interaction time with NPCs. Some genres such as simulation were rejected by the second rule above: revealing various behavior patterns. MMOs were eliminated because by definition they are multiplayer. Also, the interactions with NPCs in MMORPG games often remain minor and supplementary compared to the interactions between avatars. RPGs satisfy all of the above basic requirements of this study. Specifically, RPGs generally provide a game environment in which game players are able to interact with NPCs without any time limitation. Also, most RPGs have rich and diverse narrative structures in which NPCs take diverse roles ranging from primary (main quest) to minor (side quest).

A first step in recruiting RPG characters to test the research questions/hypothesis of this study was to determine an objective sampling method for selecting RPG games. A starting point was a website listing RPGs, Metacritic scores, and a year of publish. RPGs

with third person perspective interface (especially the game with top-down interface) were eliminated due to the fact that many believability qualities would not be visible in the top-down interface. For example, the emotion could be found in a character's facial expression which was impossible to reveal in top-down interface. Games for this study were selected based on two criteria -- published year and game quality score. Four games would be selected – one newer game with a high Metacritic score, one newer game with a low Metacritic score, one older game with a high Metacritic score, and one older game with a low Metacrtici score. Within each of the chosen games, two NPCs would be selected based on the NPC's role (antagonist or protagonist).

RPGs published earlier than year 2006 were grouped in "old games" while games published later than or in the year 2006 were grouped in "new games." For the published year category, year 2006 was selected as the determining year considering the duration of game developments (one to three years)[2]. The Metacritic web site listed 63 RPGs initially. The Metascore categorization was decided by the distribution of all RPGs in the Metacritic website considering the lowest score (44) and the highest (95). The cutoff score for the high group was (84) and (71) for the lowest group. Games in middle Metascore were eliminated from the study in order to maximize the difference between good game and bad game. In a similar reason, the cutoff Metascore for the high group was set to 90 while the score for the lowest group was set to 70. The group with higher than 90 Metascore was labeled as "good games" while games with Metascore lower than 70 were labeled as "bad games" for the game quality score. The original list of RPGs with their Metasocre is listed in the Appendix.

---

[2] http://en.wikipedia.org/wiki/Game_development

With the two between-game criteria, all games in the Metacritic were grouped in four different groups: high Metascore-new games (n=1), high Metascore-old games (n=7), low Metascore-new games (n= 6), and low Metascore-old games(n=13),. (Medium Metacritic score games were omitted from consideration to maximize the good game/bad game differences.) The four games for the study were randomly selected from those four game groups.

The actual NPC recruitment was conducted within the four games. One protagonist or helpful NPC and one antagonist NPC was recruited from each game. The final NPC recruitment was made based on convenience (characters who appeared early in game play) due to the fact that game design companies don't provide the exhaustive list of the NPCs in their games. The final NPCs were recruited along with the storyline of the game based on first-timeness and importance of character's role. For example, the final good NPC was the first good and important character that game players may encounter along with main quest, and the bad NPC was the first evil and important character that game players may encounter in a storyline. For deciding the importance of a character's role, the decision was made by whether the character was involved in main quest.

RPGs challenge game players with various goals (they are called 'quests'). There are two different types of quests in RPGs: main quests and side quests. In most RPGs, a main quest is the primary mission for game players and must be completed to advance through the game. In order to achieve the main quest, the game player must go through many pre-defined events populated with primary and minor NPCs. The success of a RPG game usually depends on whether game players accomplish the main quest or not. The

completion of side quests seldom affects the success of a game, and only exists for additional exploratory enjoyment for game players.

All final NPCs were important characters staged around events that are related to a main quest. For example, one of the final NPCs is a monk whose job is giving important information that is a critical piece of information in terms of completing a main quest to game players. Without contacting and getting information from him, it is impossible for game players to complete the main quest. The final eight NPCs were the first good or evil characters encountered along with a main quest. For this, many cheating websites that provided step-by-step information along with main quest were used. The final RPG NPCs are recruited in shown in Table 7.

Table 7. Recruited NPCs

			Character Name	Game Title	Metascore	Year
New	High	Good	Jauffre	Oblivion	94	2006
		Bad	Dremora	Oblivion	94	2006
	Low	Good	Vesit	Two Worlds	65	2007
		Bad	Young Grom	Two Worlds	65	2007
Old	High	Good	Paul	DeusEx	90	2000
		Bad	Terrorist	DeusEx	90	2000
	Low	Good	Yrsa	Might and Magic IX	55	2002
		Bad	Skeleton	Might and Magic IX	55	2002

Procedures

The purpose of this study is to measure the believability on various NPCs and compare them in order to retrieve some valuable information for game character designers. For this, the five believability qualities (personality, emotion, appearance and behavior, goal, and social relation) were revealed through extensive literature reviews. The believability of each NPC was measured in general believability questions asking about subjects' general believability perception, as well as in specific believability

questions asking about subjects' believability perception of each five believability quality.

The game players' subjective perception on character believability needs to be measured in a way that guarantees the maximum exposure of those qualities without any distractions or intervention. A research method in which participants are not distracted by the interaction with game characters was needed. For the above reasons, video clips were selected as test materials for this study rather than letting participants experience the actual NPCs in the very game environment. Eight video clips were recorded from the very game environment of the eight recruited NPCs. The video clips included almost all behavior patterns of each individual NPC in their game environments. The NPC videos varied in length from one minute thirty seconds to two minutes depending on the behavior patterns.

Subjects in the study were required to access a website which contains all the materials and questionnaires. The order in which the eight NPCs were presented was varied randomly in order to avoid the practice effect in repeated measure. Subjects were able to control the video clips in terms of play, pause, stop, rewind, and fast-forward so that they could replay the video clips if they wanted. After each NPC video clip, subjects were asked to answer questions regarding the believability of the NPC they just watched.

The survey consisted of two parts. The first set of questions measured game players' perception on believability of each character they just watched. Four general believability questions asked about general believability of the NPC and six believability questions asked about their perception on five different NPC qualities (two for appearance, and one each for behavior, emotion, goal, social relation, and personality). A

final question asked about subjects' general enjoyment of the NPC. All of the

operationalizations of these concepts were developed for this study due to the lack of

previous research on believability. The response categories used a five level Likert Scale

with 5 representing strong agreement and 1 representing strong disagreement. In the last

part of the survey, subjects were asked about demographic information such as age,

gender, college year, etc.

RESULTS

Subjects in this study were recruited from a sophomore level introductory digital media course at a large Midwestern university. A total of 161 subjects participated. Participants were given extra credit in exchange for their participation. Thirty-eight were female, 119 were male, and 4 left gender blank on the survey. There were 38 subjects in their freshman year, 51 sophomores, 39 juniors, and 28 seniors. The subjects' average age was roughly 21 (20.97) years old. Age of 19 and 20 occupied 49 % of the whole participant population. Among various game genres, 72 subjects (44.7%) reported action games as their favorite along with other genres: 66 participants (41%) for sports and first person shooter games, 59 (36.6) for adventures, 57 (35.4%) for RPG, 56 (34.8) for strategy, etc. Participants reported an average of 3.42 hours of RPG playing in a week ranging from 0 hour to 52 hours and an average 2.97 hours of playing in a session. Regarding the game player types, 59 participants showed preference on explorer type, 35 for socializer, 31 for achievers, and 12 killers.

General Believability

Four general believability questions measured the participants' general perception on general believability including "In general, this character is believable within the game (Q1)," "The character in the game is what I would expect it to be (Q2)," "I think this character is realistic within the game (Q3)," and "I think I could guess how this character might respond under different circumstances (Q4)." Table 8 shows individual character average believability across the eight NPCs. The four general believability questions were combined into a single additive scale. Cronbach's alpha for the GB scale was 0.897.

Table 8. Individual Character Average Believability Across the Eight NPCs

		Jauffre	Dremora	Vesit	Young Grom	Paul	Terrorist	Yrsa	Skeleton	Avg.
Q1	Avg.	3.85	3.72	3.58	3.59	3.73	3.36	3.23	3.47	3.57
	SD	0.87	0.87	0.99	0.91	0.92	1.01	1.05	0.98	0.95
Q2	Avg.	3.80	3.54	3.47	3.59	3.59	3.41	3.13	3.61	3.52
	SD	0.85	0.93	0.97	0.91	0.96	1.01	1.01	0.85	0.94
Q3	Avg.	3.82	3.55	3.60	3.54	3.66	3.23	3.28	3.51	3.52
	SD	0.87	0.96	1.01	0.97	0.99	1.03	1.00	0.93	0.97
Q4	Avg.	3.50	3.29	3.46	3.29	3.54	3.33	2.75	3.34	3.31
	SD	0.93	0.99	0.99	1.07	0.96	1.04	1.07	1.07	1.02
General Believability	Avg.	3.74	3.53	3.53	3.50	3.63	3.33	3.10	3.48	3.48
	SD	0.88	0.94	0.99	0.96	0.96	1.02	1.03	0.96	0.97

Average GB ranged from a low of 3.10 (SD = 1.03) for Yrsa to a high of 3.74 (SD = .88) for Jauffre, both of whom were protagonist NPCs. Jauffre was rated the highest on general believable (M = 3.74, SD = .88) among the NPCs with the lowest standard deviation (0.88) indicating relatively high agreement on the score. The least general believable NPC was Yrsa (3.10) with the highest standard deviation (1.03). Other NPCs showed general believability ranging from 3.3 to 3.6. Yrsa and Terrorist were NPCs that showed the lowest agreement on the believability (a standard deviation of 1.02 for Terrorist and 1.03 for Yrsa). However, the agreement on believability score was about the same across the NPCs ranging from 0.88 to 1.03. In general, all NPC general believability ratings had consistent levels of standard deviation across the four questionnaires.

Specific Believability Qualities

The specific believability (SB) qualities measured the participants' perception on five specific believability qualities that were derived from the believability literature review. The specific believability qualities included "The personality of this character is convincing to me (personality)," "The appearance of this character makes sense to me

(appearance 1)," "The way this character behaves makes sense to me (appearance 2),"

"The way this character responds to me resembles how humans respond to each other

(social relation)," "I can clearly understand this character's motivations (goal)," and "The

emotional expression of this character is very clear and convincing to me (emotion)." The

five specific believability questions were combined into a single additive scale.

Cronbach's alpha for the SB scale was 0.917. Table 9 shows detailed NPC specific

believability ratings by character.

Table 9. NPC Specific Believability Quality Ratings

		Jauffre	Dremora	Vesit	Young Grom	Paul	Terrorist	Yrsa	Skeleton	Avg.
Personality	Avg.	3.73	3.43	3.25	3.38	3.52	3.14	2.82	3.27	3.32
	SD	0.88	0.94	1.12	1.02	0.97	1.14	1.06	1.04	1.02
Appearance	Avg.	3.91	3.58	3.56	3.56	3.59	3.28	3.11	3.51	3.51
	SD	0.87	0.94	1.03	0.94	1.00	1.12	1.06	0.95	0.99
Behavior	Avg.	3.79	3.51	3.39	3.32	3.69	3.24	2.81	3.40	3.39
	SD	0.86	0.88	1.08	1.03	0.89	1.04	1.15	0.97	0.99
Social relation	Avg.	3.56	2.59	3.21	2.50	3.50	2.62	2.38	2.53	2.86
	SD	0.91	1.11	1.14	1.04	1.01	1.11	1.01	1.10	1.05
Goal	Avg.	3.74	3.34	3.31	3.22	3.69	3.22	2.53	3.34	3.30
	SD	0.85	1.09	1.05	1.13	0.88	1.16	1.09	1.14	1.05
Emotion	Avg.	3.41	3.24	3.11	3.27	3.19	2.81	2.59	3.22	3.11
	SD	1.03	1.11	1.20	1.14	1.09	1.12	1.09	0.97	1.09

Paralleling the general believability results, Jauffre was rated highest on the

specific believable qualities (3.69) across all NPCs while Yrsa was the least believable in

terms of the specific believability qualities (2.71). The specific believability qualities for

all other NPCs were in between 3.05 and 3.53. Again the result was very similar to that of

general believability. In general, the appearance of NPCs was rated more believable

(3.32 on average) than other believability qualities while social relation believability

seemed to be the hardest to achieve (2.86 on average). Again all NPCs showed roughly

the same standard deviation ranging from 0.90 (Jauffre) to 1.12 (Terrorist). The standard

deviation across believability qualities showed almost the same level of agreement on each score ranging from 0.99 (appearance 1 & 2) to 1.09 (emotion).

Enjoyment

The enjoyment question measured the participants' general perception on how much they expected to enjoy playing with the NPC: "It would be enjoyable to play with this character in a game (enjoyment)." Table 10 shows average responses by individual NPC.

Table 10. Individual NPC Average Enjoyment

		Jauffre	Dremora	Vesit	Young Grom	Paul	Terrorist	Yrsa	Skeleton	Avg.
Enjoyment	Avg.	3.17	3.16	3.06	3.00	3.42	2.98	2.41	3.09	3.32
	SD	1.06	1.10	1.03	1.02	0.96	1.11	1.10	1.03	1.02

Jauffre and Dremora had almost the same expected enjoyability, while Yrsa was rated the least enjoyable. The standard deviation showed roughly the same level across all NPCs.

H1 to H5: General and specific Believability

H1: A character's appearance will be closely related to that character's general perceived believability.

H2: A character's personality will be closely related to that character's general perceived believability.

H3: The goals of a character will be closely related to that character's general perceived believability.

H4: A character's emotions will be closely related to that character's general perceived believability.

43

H5: A character's social relations will be closely related to that character's general

perceived believability.

For these hypotheses, the general believability index (combining all four GB

questions) and five believability qualities (six questions) were the average scores across

eight different NPCs. The general believability score was calculated by averaging four

general believability questions across eight NPCs measuring general perception of

believability. In order to study the relation between the general believability and the

specific believability qualities, the general believability score was compared to each

specific believability qualities. These scores were subjected to a Pearson correlation. For

an alpha level of .01, all correlations were found to be statistically significant. Table 11

shows the result of correlation of believability qualities among NPCs.

In general, personality and appearance (both appearance and behavior) showed

high correlation with general believability. The highest correlation with the general

believability among believability qualities was personality ($r(159)= 0.805, p < 0.01$)

while the second and third highest was behavior ($r(159)=0.789, p < 0.01$) and appearance

($r(159) = 0.769, p < 0.01$) in turn. Each NPC showed the highest correlation only with

personality (Terrorist, Skeleton, Vesit, and Yrsa), appearance (Dremora and Grom) and

behavior (Jauffre and Paul). Table 11 shows the result of correlation between each

believability quality and general believability.

Among five believability qualities, appearance 1, appearance 2, goals, emotion,

and personality were strongly correlated with general believability while the correlation

between social relation and general believability was significant but weak. **Hypothesis 1,**

2, 3, 4 were supported by the data while hypothesis 5 (general believability and

personal relations) was not supported.

Table 11. Correlation Between Each Believability Quality and General Believability

		Personality	Social relation	Goal	Emotion	Appearance1	Appearance2
Dremora	Pearson Correlation	0.638	0.167	0.560	0.489	0.693[a]	0.621
	Significance	0.000	0.000	0.000	0.000	0.000	0.000
	N	161	161	161	161	161	161
Jauffre	Pearson Correlation	0.719	0.630	0.570	0.651	0.715	0.754[a]
	Significance	0.000	0.000	0.000	0.000	0.000	0.000
	N	161	161	161	161	161	161
Terrorist	Pearson Correlation	0.656[a]	0.425	0.545	0.555	0.572	0.646
	Significance	0.000	0.000	0.000	0.000	0.000	0.000
	N	161	161	161	161	161	161
Paul	Pearson Correlation	0.706	0.543	0.642	0.490	0.697	0.717[a]
	Significance	0.000	0.000	0.000	0.000	0.000	0.000
	N	161	161	161	161	161	161
Skeleton	Pearson Correlation	0.723[a]	0.260	0.571	0.543	0.719	0.676
	Significance	0.000	0.000	0.000	0.000	0.000	0.000
	N	161	161	161	161	161	161
Vesit	Pearson Correlation	0.727[a]	0.590	0.589	0.552	0.541	0.652
	Significance	0.000	0.000	0.000	0.000	0.000	0.000
	N	161	161	161	161	161	161
Grom	Pearson Correlation	0.675	0.295	0.380	0.557	0.716[a]	0.621
	Significance	0.000	0.000	0.000	0.000	0.000	0.000
	N	161	161	161	161	161	161
Yrsa	Pearson Correlation	0.697[a]	0.458	0.596	0.542	0.611	0.645
	Significance	0.000	0.000	0.000	0.000	0.000	0.000
	N	161	161	161	161	161	161
Average		R=0.805	R =0.473	R=0.599	R=0.595	R=0.769	R=0.789

< a: the highest correlation within NPC>

The five believability qualities were subjected to a Pearson correlation. For an alpha level of .01, all correlations were found to be statistically significant. The strongest correlation was found between personality and appearance ($r(159) = 0.826, p < 0.01$) while the weakest correlation was between goal and social relation ($r(159) = 0.498, p <$

0.01). The table 12 shows the result of correlation among believability qualities. All believability items were significantly correlated, but social relations were the most distinct.

Table 12. The Result of Correlation Among Believability Qualities

	Personality	Social relation	Goal	Emotion
Social relation	.566			
Goal	.719	.498		
Emotion	.785	.594	.677	
Appearance	.826	.481	.714	.680

RQ1 and RQ2: Which specific believability qualities are most related to general believability

RQ1: How will the general believability level be constructed in relation to the five believability qualities? Will the general believability be established by one or few believability qualities?

RQ2: If so, how will other believability qualities affect the general believability? Will they lower the general believability?

Regression was used to analyze the relationship between the general believability and the specific believability qualities. Significant variables are shown in Table 13. Consistent with the result from correlation analysis, the regression result showed higher coefficients in personality, appearance, and behavior with statistical significance. In all NPCs, six believability qualities explained 60 to 78 percent of the whole variation in general.

46

Table 13. Regression between General Believability and Specific Believability Qualities

	Coefficients	b	p	R^2	F/sig
Dremora	Personality	0.158	0.005	0.628	43.368 / 0.000
	Appearance	0.302	0.000		
	Behavior	0.066	0.276		
	Social relation	-0.20	0.547		
	Goal	0.178	0.000		
	Emotion	0.032	0.441		
Jauffre	Personality	0.205	0.000	0.778	90.068 / 0.000
	Appearance	0.198	0.000		
	Behavior	0.246	0.000		
	Social relation	0.094	0.023		
	Goal	-0.17	0.703		
	Emotion	0.153	0.000		
Terrorist	Personality	0.136	0.014	0.607	39.613 / 0.000
	Appearance	0.174	0.000		
	Behavior	0.164	0.004		
	Social relation	0.003	0.940		
	Goal	0.128	0.003		
	Emotion	0.138	0.004		
Paul	Personality	0.138	0.011	0.715	64.545 / 0.000
	Appearance	0.213	0.000		
	Behavior	0.215	0.000		
	Social relation	0.092	0.025		
	Goal	0.181	0.000		
	Emotion	0.028	0.437		
Skeleton	Personality	0.215	0.000	0.719	65.718 / 0.000
	Appearance	0.229	0.000		
	Behavior	0.193	0.000		
	Social relation	0.009	0.746		
	Goal	0.023	0.522		
	Emotion	0.078	0.048		
Vesit	Personality	0.303	0.000	0.631	43.972 / 0.000
	Appearance	0.152	0.001		
	Behavior	0.122	0.035		
	Social relation	-0.004	0.943		
	Goal	0.155	0.001		
	Emotion	-0.012	0.800		
Grom	Personality	0.191	0.000	0.661	50.100 / 0.000
	Appearance	0.345	0.000		
	Behavior	0.131	0.008		
	Social relation	0.010	0.795		
	Goal	0.005	0.897		
	Emotion	0.067	0.111		
Yrsa	Personality	0.293	0.000	0.693	57.890 / 0.000
	Appearance	0.259	0.000		
	Behavior	0.159	0.000		
	Social relation	-0.049	0.331		
	Goal	0.120	0.019		
	Emotion	0.011	0.814		

47

Personality, appearance 1 (outfit), and appearance 2 (behavior) were significant predictors for general believability with similar level. General believability seemed to reflect surface level of believability. Table 14 shows the five believability quality scores of eight different NPCs along with their general believability scores as well as the average of five believability quality scores.

Table 14. The Five Believability Quality Scores Among 8 NPCs

	New				Old			
	High Score		Low Score		High Score		Low Score	
	Evil	Good	Evil	Good	Evil	Good	Evil	Good
	Dremora	Jauffre	Terrorist	Paul	Skeleton	Vesit	Grom	Yrsa
Personality	3.43	3.73	3.14	3.52	3.27	3.25	3.38	2.82
Emotion	3.24	3.41[b]	2.81	3.19[b]	3.22	3.11[b]	3.27	2.59
Social relation	2.59[b]	3.56	2.62[b]	3.5	2.53[b]	3.21	2.50[b]	2.387[b]
Goals	3.34	3.74	3.22	3.69[a]	3.34	3.31	3.22	2.53
Appearance	3.58[a]	3.91[a]	3.28[a]	3.59	3.51[a]	3.56[a]	3.56[a]	3.10[a]
Behavior	3.51	3.79	3.24	3.69[a]	3.40	3.39	3.32	2.80
General B	3.52	3.74	3.33	3.63	3.48	3.53	3.50	· 3.09
Correlation with General B	R=0.731 P=0.000 N=161	R=0.866 P=0.000 N=161.5	R=0.764 P=0.000 N=161	R=0.828 P=0.000 N=161	R=0.803 P=0.000 N=161	R=0.768 P=0.000 N=161	R=0.733 P=0.000 N=161	R=0.779 P=0.000 N=161

< a=the highest believability among five believability quality>
< b=the lowest believability among five believability quality>

Each highest believability score among five believability qualities was slightly higher than the general believability score except Terrorist. The general believability score was slightly higher (3.33) than the highest score among five believability qualities (3.28). In general, the general believability score was almost the same with the highest believability score among five believability qualities. It seemed that the general believability perception was constructed mainly by one or few well-designed believability quality without being hassled by other believability qualities much. The schema game players created from each character may be big and vague enough to accommodate all the combination the five believability qualities created. In other worlds,

each five believability qualities worked together in order to build a schema for the character not necessarily conflicting each other. The clearest quality may boost the general perception of believability while other qualities added up some uniqueness.

The highest believability quality scores were from appearance in most cases (Dremora: 3.54 from its appearance, Jauffre: 3.85 from its appearance, Terrorist: 3.26 from its appearance, Paul: 3.69 from its goals, Skeleton: 3.46 from its appearance, Vesit: 3.47 from its appearance, Grom: 3.44 from its appearance, and Yrsa: 2.96 from its appearance). The highest score for Paul was from Goals, but the appearance of Paul was about almost the same. It seemed that game players perceived the appearance of character with the highest degree of believability.

Individual Differences and Believability

H6: RPG gaming hours will result in higher general believability.

Regression was used to analyze the relationship between hours of RPG playing and general believability. The hours of RPG playing did not significantly predict the general believability (b = .011, $t(151)$ = 1.663, p = 0.098). Hypothesis H6 was not supported.

Believability Factors

Factor analysis was conducted on the four general believability and six specific believability quality questions which were averaged with eight NPCs to look for underlying dimensions within each character. Principal components analysis with Varimax rotation was used. Table 15 shows the result of factor analysis.

Table 15. The Result of Factor Analysis

Rotated Component Matrix[a]		
	Component	
	1	2
General Believability 1	**.913**	.278
General Believability 2	**.887**	.275
General Believability 3	**.900**	.282
General Believability 4	.330	**.707**
Social relation	.087	**.861**
Emotion	.449	**.737**
Appearance 1	**.848**	.306
Appearance 2	**.724**	**.563**
Personality	**.697**	**.606**
Goal	*.504*	*.663*
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.		

Two factors emerged accounting for 79.68% of the variance. Three of the general

believability questions (general believability question 1, 2, and 3) and three of the

specific believability qualities (personality, appearance1 and appearance2) loaded .6 or

higher on factor 1. General believability question 4 and four of the specific believability

qualities (personality, social relation, goal and emotion) loaded .6 or higher on factor 2.

Three questions -- "the personality of this character is convincing to me (personality),"

"the way this character behaves makes sense to me" (appearance2), and "I can clearly

understand this character's motivations (goal)," loaded higher than .5 on both factors.

The items loading high on both factors have something in conceptual common -- they

relate to how well the respondent feels they could predict how the NPC might behave.

Dimension 1, Surface Believability, combines four questions that had higher loadings

than 0.8 in only component 1. Surface Believability relates to judgments about the NPC

which could be made quickly, at first glance. Surface Believability combined "In general, this character is believable within the game," "The character in the game is what I would expect it to be," "I think this character is realistic within the game," and "The appearance of this character makes sense to me." Average responses (3.53) combining all characters were higher than neutral (3) on the five-point scale of Surface Believability ranging from 1.59 to 5.

Surface believability included three of the general believability questions but excluded the fourth general believability question: "I think I could guess how this character might respond under different circumstances." The excluded question seemed to be different with the other three questions in requiring more cognitive capacity to process the information while the other three questions asked cognitively shallow attributes of believability relatively. Also, the Specific Believability appearance 1 question ("The appearance of this character makes sense to me.") asked about static appearance of NPC that was easy to perceive and didn't need much cognitive capability to process. In general, Surface Believability dealt with snap judgment attributes of believability that didn't need much cognitive power to process and were faster to perceive.

Dimension 2, Personal Believability, combined three believability qualities (the appearance 2 behavior patterns, personality and goal): "The way this character behaves makes sense to me," "The personality of this character is convincing to me," and "I can clearly understand this character's motivations." Average responses combining all characters were 3.33 on a five-point scale of Personal Believability ranging from 1.71 to 4.75. Personality Believability measured a more intermediate type of believability

attributes that need a little more cognitive capability to process than Surface Believability. However, these believability attributes didn't need heavy cognitive processing compared to the Predictable Believability.

Predictable Believability, the special third scale combined two believability qualities (emotion and social relation) and one general believability: " The emotional expression of this character is very clear and convincing to me," "The way this character responds to me resembles how humans respond to each other," and "I think I could guess how this character might respond under different circumstances." Average responses were slightly higher than neutral (3) on the five-point scale of Predictable Believability ranging from 1.58 to 4.75. Predictable Believability consisted of attributes that need heavy cognitive capability to process and time interacting with the character. Predictable Believability is probably the hardest for designers to incorporate and might be expected to evidence lower believability ratings. It may be the highest threshold for believability. An NPC could perform well on Surface and Personal Believability, yet fail on the Predictability dimension. Considering these three believability dimensions over time, Surface Believability is an immediate, snap judgment. Personal Believability judgments follow quickly, and Predictable Believability qualities take time and character interaction to form.

Because the goal of this research is to explore relationships among different facets of NPC believability, the three items that loaded onto both factors were treated as their own unique scale. The three believability scales were constructed by summing items that loaded .6 or higher on the factor with their relative contribution on the other factor. The summed scales were then divided by the number of items so that the resulting means

could be interpreted as roughly corresponding to the 5 point Likert scale used for the individual items.

The Surface Believability has the highest mean value while Personal Believability is lower than the Surface Believability but higher than the Predictable Believability. Although they emerged as two separate factors, Surface Believability and Personal Believability are highly correlated. Table 16 shows the correlations between three new believability dimensions. All the correlations were statistically significant at the 0.01 level.

Table 16. The Result of Correlations between Three New Believability Dimensions

	Surface Believability	Personal Believability
Personal Believability	.854	
Predictable Believability	.619	.827

Table 17 summarizes the mean values, min/max, standard deviation, and Cronbach's alpha of those three scales. Cronbach's Alpha was calculated to check the reliability of these scales. All reliabilities were above .78, with Surface Believability at .954, Personal Believability at .9, and Predictable Believability at 0.781.

Table 17. The Mean, Min/Max, Standard Deviation, and Cronbach's Alpha of Three New Believability Dimensions

	Minimum	Maximum	Mean	Std. Deviation	Alpha
Surface Believability	1.59	5.00	3.5306	.58248	0.954
Personal Believability	1.71	4.75	3.3362	.56194	0.900
Predictable believability	1.58	4.75	3.0932	.50532	0.781

<Table ###>

Comparing Believability of High and Low Rated Games, New and Old Games, and Good and Evil Characters

For these comparisons, the three believability factors were used. The variation of the three believability scales was found between new and old games, good and bad characters, and high and low rated games. The Surface Believability of a good character in a high rated new game showed the highest mean where the Predictable Believability of a good character in a low rated old game was the lowest. In all three believability scales, the good character in a high rated new game was the highest (3.84 for Surface Believability, 3.75 for Personal Believability, and 3.49 for Predictable Believability) while the good character showed the lowest mean in a low rated old game (3.19 for Surface Believability, 2.72 for Personal Believability, and 2.57 for Predictable Believability). Table 18 summarizes the mean values of three scales across the three recruitment categories (character goodness, game rating, and game newness).

Table 18. The Mean Values of Three Believability Dimensions Across the Three Recruitment Categories

	Good				Bad			
	High rating		Low rating		High rating		Low rating	
	New	Old	New	Old	New	Old	New	Old
Surface Believability	3.84	3.64	3.55	3.19	3.60	3.32	3.57	3.53
Personal believability	3.75	3.63	3.32	2.72	3.43	3.20	3.31	3.34
Predictable believability	3.49	3.41	3.26	2.57	3.04	2.92	3.02	3.03

The Surface Believability, the Personal Believability, and the Predictable Believability scores were subjected to a four-way ANOVA with three levels of NPC

evilness (good, bad), game newness (new, old), game rate (high, low), and gender (male, female).

H7a: Characters in newer games will be more believable than characters in old games.

There was a significant main effect for game newness, $F(1, 155) = 32.296$, $p <$ 0.001 indicating that the mean **Surface Believability** score was significantly greater for new games ($M = 3.66$, $SD = 0.058$) than for old games ($M = 3.42$, $SD = 0.060$). Participants showed higher Surface Believability on NPCs recruited from newer games than NPCs recruited from older games.

There was a significant main effect for game newness, $F(1, 155) = 23.722$, $p <$ 0.001 indicating that the mean **Personal Believability** score was significantly greater for new games ($M = 3.48$, $SD = 0.056$) than for old games ($M = 3.27$, $SD = 0.058$). Participants showed higher Personal Believability on NPCs recruited from newer games than NPCs recruited from older games.

Also, there was a significant main effect for game newness, $F(1, 155) = 22.531$, $p < 0.001$ indicating that the mean **Predicable Believability** score was significantly greater for new games ($M = 3.21$, $SD = 0.052$) than for old games ($M = 2.98$, $SD = 0.054$). Participants showed higher Predictable Believability on NPCs recruited from newer games than NPCs recruited from older games. The hypothesis H7a was supported.

H8a: Characters in games with good rating will be more believable than characters in lower rating games.

There was a significant main effect for game rate, $F(1, 155) = 4.477, p = 0.035$, indicating that the mean **Surface Believability** score was significantly greater for high Metascore games ($M = 3.59, SD = 0.06$) than for low Metascore games ($M = 3.49, SD = 0.058$). Participants showed higher Surface Believability on NPCs recruited from games with a higher rating than NPCs recruited from games with a lower rating.

There was a significant main effect for game rate, $F(1, 155) = 47.236, p < 0.000$, indicating that the mean **Personal Believability** score was significantly greater for high Metascore games ($M = 3.53, SD = 0.058$) than for low Metascore games ($M = 3.22, SD = 0.057$). Participants showed higher Personal Believability on NPCs recruited from games with a higher rating than NPCs recruited from games with a lower rating. The hypothesis 3 was supported.

There was a significant main effect for game rate, $F(1, 155) = 19.115, p = 0.000$, indicating that the mean **Predicable Interaction** score was significantly greater for high Metascore games ($M = 3.19, SD = 0.053$) than for low Metascore games ($M = 3.00, SD = 0.051$). Participants showed higher Predictable Believability on NPCs recruited from games with a higher rating than NPCs recruited from games with a lower rating. The hypothesis H8a was supported.

RQ3a: How will the general believability level relate to the character good/evilness? Will the evil characters be more believable or less believable?

There was no significant main effect for character evilness in Surface Believability, $F(1,155) = 1.181, p=0.279$. The **Surface Believability** was not significantly different between good characters (3.57) and evil characters (3.52).

56

Participants showed no Surface Believability difference between good characters and evil characters.

There was no significant main effect for character evilness in Personal Believability, $F(1,155)= 0.606$, $p=0.437$. The **Personal Believability** was not significantly different between a good character (3.4) and evil character (3.35). Participants showed no Personal Believability difference between good characters and evil characters.

There was a significant main effect for character evilness in **Predictable Interaction**, $F(1,155)= 16.122$, $p=0.000$. The Predicable Interaction was significantly different between good characters and evil characters. Good characters showed significantly higher Predictable Believability (M = 3.4, SD =.054) than that of evil characters (M = 2.98, SD = .057). Participants showed higher Predictable Believability on good characters than evil characters.

Comparing Enjoyment of High and Low Rated Games, New and Old Games, and Good and Evil Characters

ANOVAs were used to compare the enjoyment of each character among good/bad NPC, high/low Metascore, new/old game and between gender. The item for enjoyment was "It would be enjoyable to play with this character in a game." Average responses (3.04) were about neutral (3) on the five-point scale of enjoyment. The Surface Believability scores were subjected to a four-way ANOVA with three levels of NPC evilness (good, bad), game newness (new, old), game rate (high, low), and gender (male, female).

H7b: Characters in newer games will be more enjoyable than characters in old games.

There was a significant main effect for game newness, $F(1, 155) = 4.844$, $p = 0.029$, indicating that the mean enjoyment score was significantly greater for new games ($M = 3.10$, $SD = 0.067$) than for old games ($M = 2.97$, $SD = 0.067$). Participants showed higher enjoyment on NPCs recruited from newer games than NPCs recruited from older games. The hypothesis 7b was supported.

H8b: Characters in games with good rating will be more enjoyable than characters in lower rating games.

There was a significant main effect for game rate, $F(1, 155) = 12.824$, $p = 0.000$, indicating that the **enjoyment** was significantly greater for high Metascore games ($M = 3.14$, $SD = 0.065$) than for low Metascore games ($M = 2.93$, $SD = 0.067$). Participants showed higher enjoyment on NPCs recruited from games with a higher rating than NPCs recruited from games with a lower rating. The hypothesis 8b was supported.

RQ3b: How will the enjoyment level relate to the character good/evilness? Will the evil characters be more enjoyable or less enjoyable?

There was no significant main effect for game rate, $F(1,155)= 2.205$, $p=0.14$. The enjoyment was not significantly different between good characters (3.09) and evil characters (2.98). Participants showed no enjoyment difference between good characters and evil characters.

58

Interaction Effects

The effect of two between-game criteria (game newness and game rate) was found to be even stronger when they are related to each other. The effect of game newness was found only with low rated games in Predictable Believability and enjoyment. The high rated games didn't show any significant believability difference between new and old games.

There was a significant main effect between game rate and game newness, $F(1,155)=7.865$, $p=0.006$). The Predictable Believability difference between new games and old games was higher in low Metascore games (0.23) than high Metascore games (0.11). Table 19 summarizes the interaction effect between game rate and game newness.

Table 19. The Interaction Effect Between Game Rate and Game Newness in Predictable Believability

	New	Old	Total
High	3.24	3.13	3.19
Low	3.17	2.83	3.00
Total	3.21	2.98	

There was a significant main effect between game rate and game newness, $F(1,155)= 5.867$, $p= 0.017$. The possible-enjoyment of characters across game newness and game rate was about the same except for characters from low rated and old games. Table 20 summarizes the interaction effect between game rate and game newness.

Table 20. The Interaction Effect Between Game Rate and Game Newness in Enjoyment

	New	Old	Total
High	3.14	3.13	3.14
Low	3.06	2.80	2.93

Total	3.10	2.97	

A Pearson correlation addressed the relationship among enjoyment (M = 3.04, SD = 0.64), Surface Believability (M = 3.53, SD = 0.58), Personal Believability (M = 3.34, SD = 0.56), and Predictable Believability (M = 3.09, SD = 0.51). For an alpha level of .01, all correlations were found to be statistically significant. The possible enjoyment was related a little more with Personal Believability than Surface Believability or Predictable Believability. Table 21 shows the result of correlation analysis between enjoyment and the believability scales.

Table 21. The Result of Correlation Analysis Between Enjoyment and Believability Scales

		Enjoy	Surface Believability	Personal Believability	Predictable Believability
Enjoy	Pearson Correlation	1.000	.653**	.715**	.655**
	Sig. (2-tailed)		.000	.000	.000
**. Correlation is significant at the 0.01 level (2-tailed).					

DISCUSSION

General believability was found to be closely related to specific believability qualities. Four of five specific believability qualities (personality, emotion, goal, and appearance) were significantly contributed to build the general believability while social relation was not related to the general believability. It may be due to the fact that many NPCs in RPGs appear by themselves without having any social relation with other NPCs. Also the NPC's social relationship with the game player can be revealed only by the interaction with game players. Perhaps having research participants watch video of the NPC in the study instead of actually playing with the characters limited their experience of interaction. Furthermore, NPCs don't engage in social relations with game players until the player approaches the boundary which triggers the NPC's interaction.

The believability perception game players construct about a character through game playing seem to be driven by one or a few qualities. In this study, the general believability perception of the characters was driven by the appearance of each character regardless of games' newness, game ratings, or characters' good/bad role. The general believability score was almost the same as the highest believability score among five believability qualities. It seemed that the general believability perception was constructed mainly by one or few well-designed believability qualities.

The results clearly show that characters from high rated games are more believable than characters from low rated games. Also, characters from newer games were found to be more believable than characters from older games. Expected NPC enjoyment was very consistent with NPC believability. Characters from newer games or high rated games were expected to be more enjoyable. The hypotheses regarding game

rating and game newness were strongly supported. Newer games are able to draw upon newly developed technologies such as better 3D graphics and sound as well as newer AI and other game algorithms. Also, highly rated games are probably highly rated in part because they incorporate well designed, believable NPCs. Highly rated new games were particularly strong in NPC Predicable Believability.

Protagonist and antagonist NPCs were most different in terms of Predictable Believability. Good character showed significantly higher Predictable Believability than evil characters. No significant difference was found between good character and bad character design for Surface Believability or Personal Believability, or for expected NPC enjoyment. The result may be due to the fact that Surface and Personal Believability are mostly based on visual cues, whereas Predictable Believability requires inferring intent and extrapolating about behavior beyond the immediate gaming situation. Perhaps we expect good people to be consistent, whereas bad guys are more of a mystery. Or perhaps antagonist NPCs are not as carefully or fully developed in a game. People seem to perceive the good character as more believable than the evil character only after a certain amount of interaction with it because the character intention such as goodness and evilness of character can only be revealed in the game storyline. In other words, people can feel a good character or an evil character more believable when they can make some prediction on that character.

The effect of improvement in good characters was found within game newness. The improvement of newer games seems to be found more on good character design than evil characters. This is where improvements were found for NPC Personal Believability, Predictable Believability and enjoyment. Good characters were more believable than bad

characters from newer games in the Personal Believability and Predictable Believability which need some interaction time to be revealed. Also good characters were more possibly enjoyable than evil characters in newer games. The improvement from new game design seems to have more impact on the believability which needs more interaction time compared to immediate, snap judgment Surface Believability.

The improvement of Predictable Believability could be found in characters from either new games or high rated games or both. Characters from old and low rated games both showed significantly lower Predictable Believability.

Surprisingly, RPG experience, represented by RPG playing time, was not a significant predictor of any of the believability scales. There was no difference in how believable the NPCs were for study participants who frequently played RPGs and those who never played. Those two groups of subjects would be expected to have vastly different schemas for RPG NPCs. Yet both types of respondents made similar judgments of the NPCs.

Several believability qualities have been considered by computer scientists as a pre-requisite for evoking willing suspension of disbelief in characters. In this study, those character believability qualities were categorized into five different groups of personality, emotion, social relation, goals and appearance. Those five categories were regrouped into three related items based on participants' data. The finding of three believability scales (Surface Believability, Personal Believability, and Predictable Believability) is consistent with the previous research on interpersonal communication.

It can be argued that a two-factor solution emerged and this is what should be used for the analysis. However, this is an exploratory study, and the three-factor solution

fits the theory better and was deemed more informative at this phase of research. Uncertainty Reduction Theory (URT) predicted three different phases (entry, personal, and exit) of human interaction with a stranger (Berger & Calabrese, 1975).

The items, which constructed the Surface Believability, are very similar and consistent with the entry phase in Uncertainty Reduction Theory. In the entry phase of URT, people exchange very surface level information such as sex, age, race, etc. The four items of the Surface Believability are "In general, this character is believable within the game," "I think this character is realistic within the game," "The character in the game is what I would expect it to be," and "The appearance of this character makes sense to me." Three of the Surface Believability items measured participants' general perceptions on each character. The other item was used to measure the appearance of each character.

During the personal phase of URT, people tend to share or seek to know more personal information such as attitudes, beliefs, values, etc. The items used to construct the Personal Believability were "The way this character behaves makes sense to me," "The personality of this character is convincing to me," and "I can clearly understand this character's motivations." All three items used for Personal Believability measured peoples' personalized perception on each character.

During the exit phase of URT, people decide if they want to continue to have a relationship with the stranger. The Predictable Believability items include "The emotional expression of this character is very clear and convincing to me," "I think I could guess how this character might respond under different circumstances," and "The way this character responds to me resembles how humans respond to each other." The

Predictable Believability measured peoples' perception of predicted interaction on each character.

As Uncertainty Reduction Theory described, the three believability subscales are dependent on time. Surface Believability is a kind of perception people construct in the first place when they encounter strange characters in a game setting. Surface Believability is the first perception people build based on character. Generally, Surface Believability is built on the visual information such as appearance. Once people build an initial perception on character, they tend to seek further information which can be interpreted personally to them. It will take a longer time for game players to construct Personal Believability than Surface believability. In Surface Believability phase, game players construct more perception on a character's personality side on top of Surface Believability. That's why Personal Believability is slower than Surface Believability in terms of the speed of recognition. Personal Believability is situated in the middle between Surface Believability and Predictable Believability in terms of its source of information. On one hand, Personal Believability is constructed on the very low level perception and believability quality such as appearance. On the other hand, Personal Believability shares a lot of common cognitive grounds with Predictable believability. Predictable Believability will be the last perception game players construct on a character due to its cognitive nature. In this phase, game players try to create a schema for a character and make a prediction on it.

The Uncertainty Reduction Theory, however, cannot explain the whole interaction between game players and game character. Different with Uncertainty Reduction Theory, game players are more active and generous in receiving and

interpreting the information they get in the course of interaction with a game character. In a game environment, game players much more willingly suspend their disbelief than people in the interaction with strangers. Just because of the definition of believability: a willing suspension of disbelief, the process of building believability from Surface Believability, through Personal Believability, to Predictable Believability is much easier for game players than people who come across strangers in a real life setting.

The mean value of the three believability scales showed a snap shot of time variation among the scales. The materials for this study are around two minutes long, and it may not be enough for the participants to construct perceptions down to Predictable Believability level. For the two minutes, participants showed the highest mean value on Surface Believability while Predictable Believability is the lowest among the scales. Personal Believability is in the middle. The variation of believability values among the three scales shows that Surface Believability contributed the most in terms of building general believability on a character while Predictable Believability contributed the least in the first two minutes of interaction. It cannot directly apply to the real game setting because this study used a game video rather than a real game. However, it seems that game players depend more on Surface Believability quality than the qualities that need more cognitive power in order to interpret a character as believable.

The theoretical communication frameworks add new insights and reasoning behind the design and study of believable NPCs. Applying the frameworks one could posit:

1) Players' first reaction to an NPC will be to apply an existing schema to understand what to expect and how to interact with the NPC (schema theory).

2) If no existing schema fits the situation, players will be forced to construct a new schema. Even so, the new schema will probably be constructed from existing schemas.

3) Players are likely to assume the NPC is like them, unless available information contradicts that assumption. When no information is available, we fill in the blanks. When contradictory information is available, we adapt our schema. Doing so takes cognitive effort and can draw someone out of suspension of disbelief.

4) Players will react to NPCs who look like player avatars and to NPCs who look like signs. We tend to be very forgiving of forms, assigning the benefit of anthropomorphism even to rocks and disembodied text.

5) More detail is not always better. Low detail lets the player fill in the blanks. The more an NPC looks but does not act like a player avatar, the more cognitive friction will result.

6) People are uncomfortable with ambiguity and uncertainty. They are not sure which schema to call upon. Drawing upon familiar schemas requires less brainpower to understand.

Dawson et al. (2007) posit that games involve either less or different suspension of disbelief than watching a movie. "Gamers appear to forget that they are playing a game less readily than filmgoers forget that they are watching a film"(Dawson, et al., 2007, p. 11). Research on what breaks suspension of disbelief in games and in movies could inform understanding of the differences between these entertainment media. The relationship between character believability and player types should also be considered. As discussed above, perceptions of character believability are by definition subjective

and in part depend on the player's motivation for playing. Gaming experience, both overall and within a game is likely to influence character believability. Players may experience a drop in believability as experience increases, if characters behave the same way each time. Or believability may increase due to familiarity, as players construct schemata for each NPC.

FUTURE STUDY

Character believability includes three different stages with different response time. Due to the fact that each believability stage has a distinct mode of interaction, investigating each individual stage in detail will be necessary in order to answer some questions: which stage is more important than others in terms of making character more believable? How does each stage affect some important concepts such as enjoyment, like, satisfaction, etc? How are these stages different with those in Uncertainty Reduction Theory explained? It is especially important to define what kind of information is delivered to game players in each different stage. Fundamentally, human beings' interaction with a computer (game characters in this case) is expected to be different than that with other human beings (as URT expected) in the precedent assumption about their interaction partner. Research on how this precedent assumption on the interaction partner can change the mode and attitude toward the partner will be needed.

In a bigger picture, the causal relationship between character believability and the quality of game will be interesting even though it was not studied in this research. Good role playing games include many other factors including character believability. This study found empirical evidence that high rated and relatively new games are populated with more believable characters than low rated and older games in general. However, it is not clear how the character believability is related to the general game assessment. If there are other factors affecting the game assessment, what are they? Especially in a Role Playing Game, the character believability may be more important than other factors due to the nature of the game. Studies on factors that affect the Role Playing Game evaluation

not only from experts' perception but also causal game players' perception will be important in terms of providing some tips to character designers.

A study of character believability within a game will be valuable to measure and determine general character believability. In this study, two characters were recruited from each game based on the assumption that characters that are critical characters for game players to complete their important missions (main quests) are more well and believably designed than characters who are just populated in a game environment waiting for game players. The big variation between important/major characters and unimportant/minor characters may break gamers' immersion in games because side quests (secondary goals of the game) can be as important as the main quests in many cases. Good games are expected to have little variation in character believability among characters maintaining a similar level of character believability regardless of character roles. One can assume that game players can immerse themselves in a very similar level of willing suspension of disbelief throughout all the quests in the game.

On the other hand, this study found some empirical evidence that there was some believability variation among characters from different games. Each character showed a different level of character believability based on its gender, goodness/evilness, occupation, race, etc. The importance of character believability can be different by the game genres. The character believability could be more/less important depending on its genres as well as its game characteristics. In some genres such as Role Playing Games and first person shooter games, Non Player Characters tend to do more important roles in game mechanics while Non-Player Characters in racing games, for example, affect less

game evaluation in general. Studies of character believability in various genres will reveal the effect of character believability on game evaluation.

Character believability studies aim to make a distinction between avatar and Non-Player Character in an environment in which both types of characters co-exist affecting each other. For human beings, it is very important to ask where the character we are interacting with are avatars (characters controlled by human beings) or Non Player Characters (character controlled by computer). Many virtual worlds such as Second Life or MMORPGs (Massively Multiplayer Online Role Playing Game) are populated with both avatars and NPCs. It is relatively easy for human beings to recognize NPCs because many NPCs are not believable. A kind of safe line between avatars and NPCs is clear by the less believable design of NPCs, not by any intentional intervention from the designer's side. It is critical to study how human beings recognize the existence of NPCs and what those factors are that make human beings guess who is what. Sooner or later, the NPCs will be more believable as the design and technology develops.

Regarding character believability, an interesting question raised from the literatures was about source credibility. A fundamental assumption of character believability is that users know that the computer characters they are interacting with are not real people. They willingly suspend their disbelief (ignoring the fact that characters are not real) for the joy of interacting with them. However, the credibility that players assign to the characters seems to have a different mechanism. People may not be as willing to suspend their disbelief when the scope of interaction moves outside of entertainment. For example, it may not be a good idea to have a clearly fictional computer agent offer medical advice or sports rehabilitation coaching. Humans may not

like the idea of being evaluated or advised by artificial computer characters. Credibility and believability are not synonymous. Believability contributes to enjoyment of games for entertainment. Credibility becomes important in "serious games." Research on how to create or increase the credibility of a computer character will be needed as patterns of human computer interaction become more diverse and serious.

Reference

Allbeck, J., & Badler, N. (2002). Toward Representing Agent Behaviors Modified by Personality and Emotion. *Proceedings of the First Annual Joint Conference on Autonomous Agents and Multiagent Systems*, 202.

Andrews, J. C., Netemeyer, R. G., & Durvasula, S. (1991). Effects of consumption frequency on believability and attitudes toward alcohol warning labels. *Journal of Consumer Affairs, 25*(2), 323.

Armbruster, B. B. (1986). Schema theory and the design of content-area textbooks. *Educational Psychologist, 21*(4), 253.

Ausubel, D. P. (1967). *Learning theory and classroom practice*: Ontario Institute for Studies in Education.

Aylett, R. (1999). Narrative in Virtual Environments-Towards Emergent Narrative. *Working notes of the Narrative Intelligence Symposium*.

Bailenson, J., Beall, A., Blascovich, J., Raimundo, M., & Weisbuch, M. (2001). Intelligent agents who wear your face: User's reactions to the virtual self. *Lecture Notes in Artificial Intelligence, 2190*, 86.

Barfield, W., Zeltzer, D., Sheridan, T., & Slater, M. (1995). Presence and performance within virtual environments. *Virtual environments and advanced interface design table of contents*, 473.

Bartlett, F. C. (1932). Remembering: An Experimental and Social Study. *Cambridge: C UP*.

Bates, J. (1992). *The Nature of Characters in Interactive Worlds and the Oz Project*: School of Computer Science, Carnegie Mellon University.

Bates, J. (1994). The Role of Emotion in Believable Agents. *Communications of the ACM, 37*(7), 122.

Bates, J., Loyall, A. B., Reilly, W. S., Castelfranchi, C., & Wemer, E. (1994). An architecture for action, emotion, and social behaviour. *Artificial Social Systems*

Selected Papers from the Fourth European Workshop on Modelling Autonomous Agents in a Multi-Agent World,MAAMAW-92 (LNAI, 830, 55.

Beltramini, R. F. (1982). Advertising perceived believability scale. *Proceedings of the Southwestern Marketing Association*, 1.

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research, 1*(2), 99.

Berkos, K. M. (2003). *The effects of message direction and sex differences on the interpretation of workplace gossip* Louisiana State University.

Bhatt, K. (2004). *Believability in Computer Games.* Paper presented at the Australian Workshop on Interactive Entertainment, Sydney, Australia.

Biocca, F., Harms, C., & Burgoon, J. K. (2001). Criteria for a theory and measure of social presence. *Retrieved February, 6*, 2003.

Blomqvist, M., Luhtanen, P., & Laakso, L. (2000). Expert-novice differences in game performance and game understanding of youth badminton players. *Physical Education & Sport Pedagogy, 5*(2), 208.

Brehm, J. W., & Lipsher, D. (1959). Communicator-Communicatee Discrepancy and Perceived Communicator Trustworthiness. *JOURNAL OF PERSONALITY, 27*, 325.

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. ACM Conf. on Human Factors in Computing: ACM Press.

Dawson, C. R., Cragg, A., Taylor, C., & Toombs, B. (2007). Video Games. *British Board of Film Classification.* Retrieved from http://www.bbfc.co.uk/downloads/pub/Policy%20and%20Research/BBFC%20Video%20Games%20Report.pdf

DeVane, B., & Durga, S. (2008). *Problem-solving in history: strategy games and schema.* Paper presented at the Proceedings of the 8th international conference on International conference for the learning sciences - Volume 3.

Dormans, J. (2006). Creating a believable gameworld. Retrieved from http://www.judithdormans.com/essays/Creating_a_believable_gameworld.pdf

Drennan, P. (2004). Conversational Agents: Creating Natural Dialogue between Players and Non-Player Characters. In Rabin, S. (Ed.), *AI Game Programming Wisdom 2*. Hingham, MA: Charles River Media, Inc.

Duis, M. (1996). Using Schema Theory to Teach American History. *Social Education, 60*(3), 144.

Ewing, T. N. (1940). *A study of certain factors involved in changes of opinion*. Duke University.

Frauenfelder, M. (2001). Smash Hits. *Wired, 9.08, August*.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *Am Psychol, 48*(1), 26.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language, 1*(2), 158.

Hayes-Roth, B., & Doyle, P. (1998). Animate Characters. *Autonomous Agents and Multi-Agent Systems, 1*(2), 195.

Hayes-Roth, B., Maldonado, H., & Moraes, M. (2002). *Designing for diversity: Multi-cultural characters for a multi-cultural world*. Paper presented at the IMAGINA 2002, Monte Carlo, Monaco.

Hayes-Roth, B., van Gent, R., & Huber, D. (1997). Acting in character. *Creating Personalities for Synthetic Actors*, 92.

Heeter, C. (1992). Being there: the subjective experience of presence. *Presence: Teleoperators and Virtual Environments, 1*(2), 262.

Hong, J., & Liu, M. (2003). A study on thinking strategy between experts and novices of computer games. *Computers in Human Behavior, 19*(2), 245.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and Persuasion: Psychological Studies of Opinion Change*: Yale University Press.

Jones, C. (1989). *Chuck Amuck: the Life and Times of an Animated Cartoonist.* New York: Farrar, Straus and Giroux.

Kramer, G. (1995). Sound and communication in virtual reality. *Lea'S Communication Series,* 259.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators & Virtual Environments, 10*(3), 282.

Lester, J. C., & Stone, B. A. (1997). *Increasing believability in animated pedagogical agents.* . Paper presented at the The First International Conference on Autonomous Agents, Marina del Rey, California.

Lester, J. C., Voerman, J. L., Towns, S. G., & Callaway, C. B. (1997). Cosmo: A Life-like Animated Pedagogical Agent with Deictic Believability. *Working Notes of the IJCAI, 97,* 61.

Loftus, G. R., & Loftus, E. F. (1983). *Mind at play: the psychology of video games*: Basic Books.

Lombard, M., Ditton, T. B., Crane, D., & Davis, B. (2000). Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. *Third International Workshop on Presence, Delft, The Netherlands.*

Loyall, A. B. (1997). *Believable Agents: Building interactive personalities.*, Carnegie Mellon University, Pittsburgh, PA.

Loyall, A. B., & Bates, J. (1997). Personality-Rich Believable Agents That Use Language. *Proceedings of the First International Conference on Autonomous Agents,* 106.

Mac Namee, B., & Cunningham, P. (2001). Proposal for an Agent Architecture for Proactive Persistent Non Player Characters. *Proceedings of the Twelfth Irish Conference on Artificial Intelligence and Cognitive Science,* 221.

Mandler, J. M. (1984). *Stories, scripts, and scenes: aspects of schema theory*: L. Erlbaum Associates.

Mateas, M. (1997). *An OZ-centric review of interactive drama and believable agents.* (No. CMU-CS-97-156): School of Computer Science, Carnegie Mellon University.

Mateas, M. (2002). *Interactive Drama, Art and Artificial Intelligence.* University of California.

Mates, M. (1997). *An OZ-centric review of interactive drama and believable agents.* (No. CMU-CS-97-156): School of Computer Science, Carnegie Mellon University.

McCloud, S. (1994). *Understanding Comics: The Invisible Art*: HarperPerennial.

Mori, M. (1970). The Uncanny Valley. *Energy, 7*(4), 33.

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies, 45*(6), 669.

Nass, C., Isbister, K., & Lee, E. (2000). Truth is beauty: Researching embodied conversational agents. In Cassell, J., Sullivan, J., Prevost, S. & E. Churchill (Eds.), *Embodied conversational agents* (pp. 374). Cambridge, MA: MIT Press. .

Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues, 56*(1), 81.

Nass, C., Moon, Y., & Carney, P. (1998). *Are People Polite to Computers?: Responses to Computer-based Interviewing Systems*: Division of Research, Harvard Business School.

Newman, J. (2002). The Myth of the Ergodic Videogame. *Game Studies, 2*(1), 1.

Nijholt, A. Agents, Believability and Embodiment in Advanced Learning Environments. *Proc. IEEE International Conf. on Advanced Learning Technologies*, 457.

Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 12*(5), 481.

Ortony, A. (2002). On making believable emotional agents believable. *Emotions in humans and artifacts*.

Pastore, N., & Horowitz, M. W. (1955). The influence of attributed motive on the acceptance of statement. *J Abnorm Psychol, 51*(2), 331.

Reeves, B., & Nass, C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*: Cambridge University Press New York, NY, USA.

Reilly, W. S., & Bates, J. (1995). Natural Negotiation for Believable Agents. *Technical Report CMU-CS-95-164, Carnegie Mellon University*.

Reilly, W. S. N. (1997). A methodology for building believable social agents. *Proceedings of the first international conference on Autonomous agents*, 114.

Riedl, M. O., & Young, R. M. (2005). An Objective Character Believability Evaluation Procedure for Multi-Agent Story Generation Systems. *Lecture Notes in Computer Science table of contents*, 278.

Rizzo, P., Veloso, M. M., Miceli, M., & Cesta, A. (1999). GOAL-BASED PERSONALITIES AND SOCIAL BEHAVIORS IN BELIEVABLE AGENTS. *Applied Artificial Intelligence, 13*(3), 239.

Romano, D. M., & Wong, A. K. L. (2004). *Personality Model of a Social Character*. Paper presented at the Workshop on Empathy in Human Computer Interaction, 7th of September at HCI 2004, the 18th British HCI Group Annual Conference.

Rousseau, D., & Hayes-Roth, B. (1997). Improvisational Synthetic Actors with Flexible Personalities. *Knowledge Systems Laboratory TR KSL97, 10*.

Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning and restructuring: Three modes of learning. *Semantic factors in cognition*, 37.

Schaap, R., & Bidarra, R. (2009). *Towards Emotional Characters in Computer Games*. Paper presented at the Proceedings of the 7th International Conference on Entertainment Computing.

Schroeder, R. (2002). Social interaction in virtual environments: Key issues, common themes, and a framework for research. *The social life of avatars: Presence and interaction in shared virtual environments*, 1.

Sengers, P. (1999). Designing Comprehensible Agents. *Proceedings of the Sixteenth Annual International Joint Conference of Artificial Intelligence*.

Shannon, C. E., & Weaver, W. (1963). *Mathematical Theory of Communication*: University of Illinois Press.

Shapiro, M. A., & Chock, T. M. (2003). Psychological Processes in Perceiving Reality. *Media Psychology, 5*(2), 163.

Slavin, R. E. (1988). *Educational Psychology: Theory Into Practice*: Prentice Hall.

Smith, C. F., Tsai, Y.-F. D., Wong, J. H., Brooks, D. T., & Peterson, M. S. (2008). More than meets the eye: Investigating expert and novice differences in action video games. *Journal of Vision, 8*(6), 105.

Szilas, N. (2003). IDtension: a narrative engine for Interactive Drama. *Technologies for Interactive Digital Storytelling and Entertainment Conference*, 187.

Thomas, F., & Johnson, O. (1981). *The Illusion of Life:Disney Animation*: Hyperion Books.

Thompson, C. (2004). The Undead Zone: Why Realistic Graphics Make Humans Look Creepy. *SLATE, June 9*. Retrieved from http://www.slate.com/id/2102086

Tychsen, A. (July 2 2007). Playing Stories - understanding role playing games and their players Retrieved Sep 23, 2009, from http://web.science.mq.edu.au/~atychsen/html2/believability.htm

van Doorn, M., & de Vries, A. P. (2006). Co-creation in Ambient Narratives: Ambient Intelligence for Everyday Life (AmI-Life'05), Lecture Notes in Computer Science.

Vorderer, P., Klimmt, C., & Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Communication Theory, 14*(4), 388.

Walther, J. B. (1996). Computer-Mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction. *Communication Research, 23*(1), 3.

Watson, M. (2002). Constructing Virtual Space.

Wood, R. T. A., Griffiths, M. D., Chappell, D., & Davies, M. N. O. (2004). The Structural Characteristics of Video Games: A Psycho-Structural Analysis. *CyberPsychology & Behavior, 7*(1), 1.

Wooldridge, M., & Jennings, N. R. (1995). Intelligent Agents: Theory and Practice. *The Knowledge Engineering Review, 10*(2), 115.