Loading… The Journal of the Canadian Game Studies Association
Vol 6(10): 30-52
http://loading.gamestudies.ca

L2P NOOB: Examining Tutorials in Digital Games

Matthew M. White
Penn State Erie, The Behrend College
mmw22@psu.edu

Abstract

It has been well-noted that contemporary digital games tend to design for a relatively high skill threshold engineered to appease a well-entrenched and digitally literate audience (Hayes, 2005). Such design practices, however, serve to disenfranchise new and novice players wanting to learn to play. This novice-expert divide is a significant barrier to entry for individuals wanting to play digital games, and given that digital games are seeing increased use in pedagogical contexts (Akilli, 2007; Becker, 2007; Nieborg, 2011; Shelton, Satwicz, & Caswell, 2011; Ulicsak, 2010), such skill-based barriers further complicate the seamless incorporation of digital games into the classroom. In an effort to explore how we might bridge the gap between new and weathered players, I created three tutorials for *World of Warcraft* (2004) in an attempt to improve the existing tutorials for newer entrants to the game. These new tutorials offered different modalities of instruction, as well as instructional strategies in assisting newer players. Tutorials were designed using the Structured Sound Functions (SSF) model of instructional design, following the Attentional Control Theory of Multimedia Learning (ACTML). The tutorials were then analyzed for their effects on play outcomes, player engagement, and player motivations using the Dick and Carey (2011) three-stage model of formative evaluation. This work thus makes two important contributions. First, this research conducts a much-needed in-depth study of game tutorials, which is an area yet to be well-charted in the disciplines of either education or games studies. Secondly, by analyzing the results of the formative evaluation, I conclude that players react favorably to a faded or "just-in-time" instructional strategy—an approach to player scaffolding which showed significantly increased motivation for play, engagement, and play mastery among novice participants. Implications for game design and future research are discussed.

Author Keywords

Game design; tutorials; education; modality; audio

Introduction

This work illustrates a problem confronting novices new to digital games, and digital games researchers, namely: players are expected to experiment with games in order to gain play mastery (Thomson, 2009). Digital games form an important part of modern pedagogy and educational research (Akilli, 2007; Becker, 2007; Nieborg, 2011; Shelton, Satwicz, & Caswell, 2011; Ulicsak, 2010), game designers take for granted the skills possessed by players (Tosca,

2003), and in doing so, fail to teach them the requisite skills necessary for gameplay. Novice players are left instead to contend with minimal or absent help systems in games, or are left to solve problems on their own. In education, this has been called "minimally-guided instruction," an element of "discovery learning" (Sweller, Kirschner, & Clark, 2007, p. 117). Discovery learning, defined as learning that has learners solve problems with minimal instructional guidance, can pose challenges for entry-level learners (Sweller et al., 2007), such as novice game players. This paper illustrates the design and implementation of three new tutorials for the best selling Massively Multiplayer Online Role-Playing Game (MMORPG) currently available: *World of Warcraft* (2004). It then assesses the effectiveness of these tutorials using Dick, Carey, and Carey's (2011) model of formative evaluation with sixteen young adult females. The female sample was the result of opportunistic sampling procedures, which take advantage of emergent or unforeseen opportunities as they arise during the course of the research (Ritchie & Lewis, 2003, p. 81). Given the well-documented marginalization of women and girls in game culture (Jenson & de Castell, 2010), it was simply much more feasible to gather a participant group with no prior gaming history by recruiting women. It is important to note that this paper does not present a study of gender and gameplay, but rather a formative evaluation of three new tutorial systems for *World of Warcraft* with the goal of offering improvements and revisions to digital game tutorials generally.

Tutorials Present in Contemporary Digital Games

Problematically, many digital games rely on a tutorial system that fails to take into consideration the experience of the player. "Good tutorials are essential for new gamers" (Hayes 2005, p. 27) and many game developers have even recently begun to leave them out of titles, as is evident with *Star Ocean: The Last Hope* (2008). Some of this may stem from the recent trend of releasing very similar games, sequels, and so on (White, 2010), assuming that consumers have already played all of the other games in the series or genre. This relates to the body of experience (Kirschner, Sweller, & Clark, 2006) gamers have built around particular genres; often it is enough to refer to "RPG", "shooter", or other genre conventions in order to recall the relevant skills, expertise, and strategies long-term gamers have built over years of experience. Players are often expected to fully 'experiment' with the controls in order to figure out how to manipulate their characters. This form of absolute discovery learning can cause difficulty for beginning learners (Sweller, Kirschner, & Clark, 2007). Tutorials leave out foundational material such as camera controls, movement controls, basic RPG menu navigation, and other beginner skills necessary to play the game; this is based in the assumption that the players would not want to be burdened with this kind of basic information. The interface's programming does not seem to respond or provide feedback when players are having trouble with these basic concepts. Numerous games employ absolute discovery learning in their approach to novice players. For example, in *Bayonetta* (2010), players are quite literally 'dropped' into a fighting situation with no instructions regarding the control or movement of their player character.

In contrast to the minimally-guided tutorial design typically found within contemporary digital games, innovative and efficient tutorial systems, designed from the ground up, are beginning to emerge. A contemporary example is *Heavy Rain* (2010), which starts players with benign tasks and on screen prompts that do not interrupt game action. The actions very quickly become more

significant if players master them, but continue to allow players to experiment if they are having trouble. There are no major penalties for failing to grasp the controls immediately. Tutorial systems such as this, built using scaffolding and cognitive apprenticeship models of teaching (Dennen, 2004), that are not "too didactic" (Hayes, 2005, p. 27) might allow future researchers to cultivate game skills *in vivo* without requiring 'starting from scratch' by looking for old titles that had tutorials. *Half-Life* (1998), for example, contained a full basic tutorial system, as it was one of the first extremely popular three-dimensional shooters. Because of this, the anticipated core gaming audience all required instruction in the novel control scheme. The better-designed tutorials discard assumptions about players and introduce novel concepts such as movement in a way that is endogenous and non-threatening. In *Fable II* (2008), for example, a trail is presented on the ground that leads players to the next objective when they are having trouble navigating the space. Advanced players can disable the trail.

Instructional Strategies

Whether intentional or not, any tutorial system that appears on screen in a digital game employs some kind of instructional strategy. Many of these do not take learner mastery into account, ignoring the difference between novice and expert play (Jenson, Fisher & de Castell, 2011).

Flashcard Instructional Strategy

The Flashcard Instructional Strategy is common in digital games. This instructional strategy provides the player with 'pop ups' which are similar to flashcards. Flashcards are often used to study memory-intensive subjects within a defined curriculum such as languages, math, science, medicine, law, and so on. Typically individuals will use flashcards for independent practice, or pairs of students will use them to practice with one another. In the context of digital games, flashcards take the form of on-screen 'pop-ups' that display a piece of information about a game mechanic (how to navigate, for example) that the designers felt would be relevant to players based on intuition (Mann, 2009). These pop-ups vary from game to game, but universally take the place of an instructional flashcard. They vary in application, but are mostly rather didactic in their approach to conveying information. Flashcards show players tidbits of context-relevant, but not necessarily temporally-relevant, information via an on-screen pop-up. These popups often interrupt gameplay, literally removing control from the player while the popup is on screen. Still others require the player to use her long-term memory in performing gaming to simultaneously attend to the task of reading and understanding the tutorial pop-up whilst avoiding in-game consequences such as death by enemy attack, loss of timed quest bonuses, disorientation in three-dimensional space, or removal from a player group due to idling.

Just-In-Time Scaffolded Instructional Strategy

Any instructional design strategy that provides decreasing instruction as learners' proficiency and mastery grows is defined as being partially or wholly scaffolded (Lajoie, 2005), while the removing of instructional material as learners gain mastery is known as instructional fading

(Burton, Moore, & Magliaro, 2005). Providing quick, relevant information as learners become frustrated or approach disengagement is known as just-in-time information (Prion, 2007). Scaffolding is often employed in cognitivist teaching methods, such as cognitive apprenticeship, which is a method of allowing learners to effectively borrow instructor cognition to achieve difficult tasks (Dennen, 2004). Any digital game that employs some or all of these methods is employing a just-in-time, scaffolded instructional strategy. While this strategy is less common in games, instructions that fade as players improve in skill, and that provide quick information at key moments when players are becoming frustrated or repeatedly making errors comprise the Just-In-Time (JIT) instructional strategy.

In a JIT instructional strategy, players can expect instructions to occur in direct relation to their skill in playing the game, which the game itself detects through programming hooks. Players who make more errors are greeted with more frequent and in-depth instruction, while players who make few or no errors are not burdened with superfluous over-instruction.

Needs Assessment

The implication from the description of tutorial strategy inefficiency outlined in earlier sections is that there is a discrepancy between the way things are and they way they ought to be. In the field of education this is known as a "need" (Seels, Fullerton, Berry & Horn, 2004). After establishing a problem and a need, the next step is to determine if one or more of these needs can be met with an educational intervention. This process is known in education as a "needs assessment" (Kaufman, 1993; Stufflebeam, 1985; Witkin, 1984). In this paper a contemporary model of assessing need referred to as Delivery, Environment, Content and Learner (D.E.C.L.) was used to determine whether one or more of these needs could be met by tutorial instruction (Mann, 2008). The D.E.C.L. was chosen over other models because of its ability to break the interrogation of need down into specific sections, or factors.
The four factors of which D.E.C.L. is composed are applied to the default *World Of Warcraft* (2004) tutorial strategy and to contemporary digital game tutorials studied as part of this work. The D.E.C.L. factors were considered in a stepwise fashion.

Delivery (D)

The delivery factor or "D" in D.E.C.L comprises four variables: presentation, strategies, scope and sequencing.

Presentation

Presentation includes the speed of the display, the graphics, audio, colour, movement, text sizes/fonts, and the quality of the interactions, including any help or tutorial accessible by the player. Before 1988, the tutorial or "tutor mode of computer operation" required content to be developed in a specific discipline with substantial coding by expert computer programmers (Mann, 2009). Computer tutors could accommodate a wide range of user differences. Today,

many users are developing their own tutorials for education or training (Mann, 2009). The tutorials in most digital Role Playing Games
(RPGs) typically employ a visual-only "flashcard tips pop-up" instructional strategy, as previously stated, that does little to accommodate novice players wanting to learn to play the game, and frequently assumes prior knowledge of relevant gameplay. There may well be deficiencies in the presentation of digital RPG tutorials that could be improved to accommodate novice players.

Strategy

Strategy in this context concerns the rules and regulations of digital RPGs. Rule systems in digital RPGs are fairly similar between games, and specifically refer to how users interact with the player characters, or "avatars" found within these games. While there are numerous types of RPGs, this study focused specifically on MMORPGs. These are fairly standard in their strategies for completion. For example, all MMORPGs rely on RPG conventions such as quests, slaying monsters, inventory management, and gaining experience points. *Runescape* (2001)*, Ragnarok Online* (2002), *Final Fantasy XI* (2004)*,Star Trek Online* (2010)*,* and *Lord Of The Rings Online* (2007), all MMORPGs, employ extremely similar gameplay systems and strategies.

Scope

Scope refers to the size of the game. Size is a difficult factor to gauge in an RPG that does not occupy physical space. Size can, however, be measured by a factor of time taken to completion. Some digital role-playing games require relatively little time to traverse, with games like *The Legend Of Zelda: Spirit Tracks* (2009) spanning 10-15 hours. MMORPGs such as *World Of Warcraft* (2004) often demand 2000 hour commitments, with players often spending 20 or more hours per week in-game (Yee, Ducheneaut, Nelson, & Likarish, 2011). Such a significant time investment could potentially alienate or intimidate novice players.

Sequencing

Sequencing concerns the ordering of goals, events, and challenges within the digital RPG. While structure and sequence vary from game to game, *World Of Warcraft* follows a very linear, well defined set of goals that the player must accomplish in order to progress. All players of the game will therefore encounter a similar sequence of events regardless of differences in their individual play styles. In *World Of Warcraft*, for example, all players can expect to encounter a new-player area where monsters are passive and quests are fairly easy. The game becomes more difficult after they have passed a certain point in gameplay. As in the previous section, MMORPGs employ similar sequences for advancing the game, regardless of the game's title.

Environment (E)

The environment factor or "E" in D.E.C.L. comprises two variables: climate and setting.

Climate

Climate refers to the level of organizational involvement. This involvement can be in-game, such as in a guild (Ang & Zaphiris, 2010), or external, including schools, colleges, or society at large. While investigation of digital games in education is being undertaken, and suggestions for the use of games in education exist (Akilli, 2007; Becker, 2007; Miyamoto, 2010; Nieborg, 2011; Shelton et al., 2011;
Ulicsak, 2010), it is not yet common to see digital games appearing in curriculum documents, or in lesson plans at a classroom level. Organizational adoption of video games is still relatively nascent, the vast majority of games being used in computer labs still being "edutainment" software (Egenfeldt-Nielsen, 2011). This lack of widespread organizational adoption eliminates a venue through which novice players might be exposed to games, and thus contributes to the need illustrated in this assessment.

Setting

Setting, according to the D.E.C.L. model, refers to the physical locations in which digital games are typically played. The setting for interacting in a digital game could be anywhere, anytime on a desktop computer or handheld device. In the current study, the setting was controlled to accurately illustrate how the digital game itself, and more specifically the instructional design of the tutorial system within the game, affected learner reception of the game.

Content (C)

The content factor or "C" in D.E.C.L., comprises three variables: domain, tasks, and the mental operations required. Digital roleplaying game content varies, but generally follows themes such as fantasy – *Lord of the Rings Online, Final Fantasy II* (1997) – or science fiction – *Fallout 3* (2008)*, Mass Effect 2* (2010). The content poses a potential source of player alienation, as not all content is appealing or suitable to all players.

Domain of the game

Domain refers to type of game content. Since the nature of game content varies from one digital role-playing game to the next, it is difficult to assess which types of game content are least exclusionary to novice players. A thorough literature review in both game studies and education yielded no studies investigating various types of digital role-playing games across players sorted by experience level. Furthermore, classifications of digital games typically default to industry-developed genre conventions such as First Person Shooter (FPS) games such as *Call of Duty: Black Ops* (2010), or Real Time Strategy (RTS) games such as *Starcraft II* (2010). Within these limitations, however, academic studies have shown that RPG gameplay is linked to self-directed, novelty-seeking behavior (Kim et al., 2010). RPGs, as opposed to other genres, also tend to move players from a state of low to high lucidity (Conway, 2010).

Tasks in the game

The tasks variable refers to the goals set for the players by the game developers. Tasks can be remapped to game design elements, sometimes called mechanics. Game mechanics are elements of digital games, such as jumping, gravity, and so on. Notably, the degree of sophistication in tasks could increase cognitive load in novices through high element interactivity (Sweller, 2010). An option may be to decrease the level of sophistication of the RPG for novices. Content and delivery are inextricably linked in digital RPGs, as content in games is delivered to players as they progress through the tasks laid out for them by game designers (Conway, 2010). *World Of Warcraft* is representative of current trends in MMORPGs, and indeed digital gaming generally, as many games employ similar tasks.

Mental operations required

Mental operations required refers to the generic thinking skills required to interact with the game – the kind and level of motor skill required, intellectual abilities (discrimination, problem identification,
and solution) strategies (devising, predicting), reaction time, and other discrete motor skills. The difference between novice and expert play has been conflated with a male-female divide in games research (Jenson, Fisher, & de Castell, 2011). Digital RPGs typically require basic use of human interface devices such as keyboards and mice, some computer proficiency, and basic problem solving and spatial navigation.

Learner (L)

The learner factor or "L" in D.E.C.L. comprises four variables: competence, capacity, demographics, and attitude. Much of the data investigating novice game play habits can be explained by the learner component of the D.E.C.L. It is tempting with novice players or learners to levy research questions *ad hominem*, while failing to account for the instructional technology within the game.

Competence

The competence variable refers to general, non-game abilities that make learning easier or more difficult for inexperienced players. These are learning strategies, general intelligence, and other attributes of the learner.

Capacity

Capacity refers to the abilities of the player relative to the requirements of the RPG. While the needs assessment did not specifically address learner capacity for gameplay, as it was primarily focused on investigating the tutorials present in *World Of Warcraft*, the random sampling method employed in the pilot test and field test would account for variances in player capacity.

Demographics

Demographics refers to age, socio-economic status, and other measurable factors of players of digital RPGs. Socioeconomic status has been found to have a negative correlation with game playership (Tobias, 2011). Age-range is also an important factor controlled in this study (Greenberg, Sherry, Lachlan, Lucas, & Holmstrom, 2008). This research sought to involve 18-28 year olds, as this age group had ample opportunity to be exposed to digital games from a young age and are among the most frequent players of games (Greenberg et al., 2008). It was very difficult to find a large pool of novice gamers, resulting in the low sample size. Furthermore, due to opportunistic sampling practices (Ritchie & Lewis, 2003), the gender demographic was unintentionally controlled. Women are still less likely to play games, including *World Of Warcraft*, whose players are 80% male (onlineschools.org, 2012), which made it much easier to find a large group of female participants who had absolutely no prior experience with MMORPGs.

Attitude

Attitude in this model refers to motivation to play the game, both intrinsic and extrinsic, including peer pressure, to play or not play games. The novice players recruited for the research undertaken in this paper did not possess intrinsic motivation to play MMORPGs, having no particular experience with them. This factor was not explored as part of this study, however player motivations after having played through different tutorials were investigated.

The D.E.C.L. model is introduced here to explore the problem faced by inexperienced players. The D.E.C.L. factors and variables offers a more sophisticated method of assessing need from that of evaluating the gap between current and ideal situations. In this study, the D.E.C.L. model was used to identify the degree to which one or more of these factors affected play mastery in a digital RPG. The D.E.C.L. model may be used to conduct educational research on the Internet (Mann, 2006), to create digital educational materials (Mann, 2005), and to evaluate those materials for revision (Mann, 2006).

Formative Evaluation Plan

A formative evaluation of four tutorials in *World Of Warcraft* was conducted in accordance with an established formative evaluation model (Dick, Carey & Carey, 2011). The Dick and Carey model is a three-stage model that employs review and retesting of instructional materials for use in education. The model provides opportunities for instructional designers to revise prototypes before they are employed in final educative situations. The four tutorials evaluated were: SSF-JIT, SSF-Flashcard, WoW-JIT, and WoW-Flashcard. The formative evaluation also included an ongoing evaluation in the form of long-term delayed post-tests. The development and evaluation of an entirely new system from the ground up, as well as evaluating new materials for existing curricula are all part of the capabilities of formative evaluation. The evaluation contained herein consisted of three distinct phases:

- Phase 1 consisted of a quality review by an instructional designer and a subject matter expert. A quality review invites expert third parties to examine the prototype instructional materials, in this case, the three new tutorials for *World Of Warcraft*.

- Phase 2 was a pilot test of the prototype tutorial strategies and modalities with learners. In a formative evaluation, the pilot test seeks to eliminate correctable errors before evaluating the effectiveness of the instruction in the field test (Dick et al., 2011).

- Phase 3 was a validation and comparative study of the tutorials. This took the form of a field test designed to evaluate the efficacy of the new tutorials in effecting changes in players' skill, motivation for play, and game engagement.

Formative Evaluation Overview

Initially, a review of all three new instructional prototypes was conducted by both a subject matter expert and an instructional designer. The role of this review was to eliminate errors that the initial designer may have overlooked. While these suggestions were not necessarily implemented in full, they were intended to sensitize the designer to potential problems before the remainder of the formative evaluation was undertaken with learners. After implementing the changes suggested by the expert reviewers, a pilot test was conducted whose purpose was to identify and remove the most obvious errors in the instruction and to obtain initial performance indications and reactions to the content by learners (Dick, Carey, & Carey, 2011). Four learners were selected and randomly assigned to the tutorial groups, and allowed to play for ten hours across three play sessions. Finally, a small group evaluation was conducted. This stage is also known as the "field test." The field test had two primary purposes: to determine the effectiveness of changes made following the pilot test and to identify any remaining learning problems that learners may have had, as well as to determine if the learners were able to use the instruction without the instructor (Dick, Carey, & Carey, 2011). This stage involved an additional twelve learners, again randomly assigned to one of the four tutorials, and allowed to play for ten hours across three sessions.

Assessment Instruments

Assessment instruments are data collection tools used to evaluate the level of success learners had in reaching instructional objectives. This section describes the data collection instruments drafted for the purpose of assessment.

Entry Skills Assessment

As stipulated in the formative evaluation plan, and consistent with Mayer (2003), only low knowledge learners were sought for the purposes of the formative evaluation. In approaching potential participants, those individuals who had previously played an MMORPG such as *World*

Of Warcraft were excluded from the study. This assessment was conducted verbally by the researcher.

Play Observation

As stated previously, during the pilot test players were closely observed and asked to comment on the instruction, the game, the difficulty, and how the program was working. This took the form of the researchers sitting in close proximity to the participants and verbally asking for feedback. This form of observation is common in the pilot stage of formative evaluations (Dick et al., 2011). This kind of feedback was invaluable in improving the tutorials for the final stage of the formative evaluation. Being asked to comment on the game is the only assessment instrument employed in the pilot test that would not later be employed in the field study.

Protocol Analysis

Protocol analysis is a method of eliciting verbal reports from participants' 'internal monologues' (Ericsson, 2006). When players illustrate behaviors that are of interest – in this case, falling idle from the game, a behavior evident of disengagement – the researcher asks the participants to briefly describe what they are thinking or doing. According to Ericsson (2006), "[...] the closest connection between actual thoughts and verbal reports is found when people verbalize thoughts that are spontaneously attended during task completion" (p. 227). In the context of this research, protocol analysis was used in this manner to "[elicit] non-reactive verbal reports of thinking" (p. 227). Participant utterances in response to researcher-driven probing were recorded and used in later data analysis. This data collection tool was employed both for the pilot test, and for the later field study.

Idle Time Collection

Idle times are the number of times the player stops playing the game entirely. These are evidenced by either physically disengaging (removing one's hands from the keyboard/mouse), or by not performing input: hands in place, but not pressing keys. Disengagement was noted by the researcher visually and recorded. At the end of a play-session, the numbers of times players idled were added to collect a total idles score.

Death Collection

Death is an undesirable outcome, even in games. In *World Of Warcraft*, death results from failing to avoid enemy attacks until sufficient damage is sustained to bring about one's demise. Death punishes the player by making her run, as a ghost, from a nearby graveyard to her corpse. In terms of play metrics, repeated and frequent deaths illustrate failure to grasp tutorial concepts, or simply failure to be motivated to perform well in the game. *World Of Warcraft* records the number of deaths experienced by a player, so this number is easily accessible by the researcher.

Experience Point Collection

Experience points are a type of progress indicator in many RPGs, and particularly in *World Of Warcraft*. Players gain experience points until they advance in level, a measure of strength and power within the game. Experience points are positive, begin at 0 and increase perpetually until a level of 85 is reached. The number of experience points a player gathers in a limited time – in this case the ten hour play window allotted to participants in both the pilot test and field study – is a measure of the degree to which they are learning from the tutorials, as players will gather experience faster if they illustrate goal behaviors, such as completing quests. Players who frequently die, wander off path, and ignore the tutorials will, however, receive fewer experience points in a limited time frame.

Gearscore Collection

Although the original plan was not to measure the improvised statistic known as gearscore, it became apparent during the pilot test that those players who were spending time managing their in-game inventory - a behavior indicative of high mastery of the game - would have less time to gain experience within the limited ten-hour play session. Because of this, it was decided that experience alone was not a sufficient measure of efficient use of play time. Idle times, deaths, and experience when taken together provided insight into the variations in play habits and play mastery between groups. Item level, as it is known in *World Of Warcraft*, is a measure of the relative power of each of the players' items (such as a dagger or piece of armor). A rusty dagger would have a low item level (perhaps 3), whereas a sword spoken of only in legend would have a considerably higher item level (perhaps 50). Gearscore is measured by adding the item level of all of the players' items together, and gives a picture of the overall power of their character, and thus, mastery.

Game Engagement Questionnaire

A peer-reviewed, established game engagement questionnaire was taken from Brockmeyer, et al. (2009) for use in the pilot test and field study stages of the formative evaluation. This is a highly reliable instrument, with item reliability of .96, and a Cronbach's alpha score of .85 for the entire questionnaire. A full description of the reliability is provided in the Brockmeyer et al. (2009) work. The study of engagement in games is common, and is consistent with literature on game acceptance (Wang & Wang, 2008). Measuring player engagement during play sessions attempted to determine the difference, if any, that may have been present between tutorial groups.

Game Motivations Inventory

A game motivations inventory – a type of test probing for different types of motivations to continue to play games – was taken from Yee (2006). This peer-reviewed instrument was tested on over 7,000 subjects in a longitudinal study of player types. According to Yee, motivation is broken down into several subtypes, such as domination and achievement (2006). Yee's definition is founded on work by Bartle (1996) on players of Multi-User Dungeons (MUDs), a precursor to MMORPGs like *World Of Warcraft*. The definition of motivation to play as articulated by Yee

differs from the traditional educational definition of motivation for learning. Other works on digital games and learning have operationalized motivation for learning when a game is introduced (Cheng, 2009), however, this was not the intention of this study.

Results

Based both on prior and current research (Jenson & de Castell, 2010; Taylor, 2007), it was expected that novice players participating in the field study would report increased mastery playing the game, and that players in the just-in-time and SSF categories would report higher skill than those players in the flashcard groups. These assumptions were expected to become evident in play skill, motivation, and engagement data. Overall, results were similar between the pilot and field tests. The subject matter expert reviews revealed no substantial errors in programming or major changes. The following section will demonstrate the results of the implementation of the four tutorials on play metrics, player motivation, and player engagement using the instruments described above. This data is from the field test and is consistent with the pilot test data.

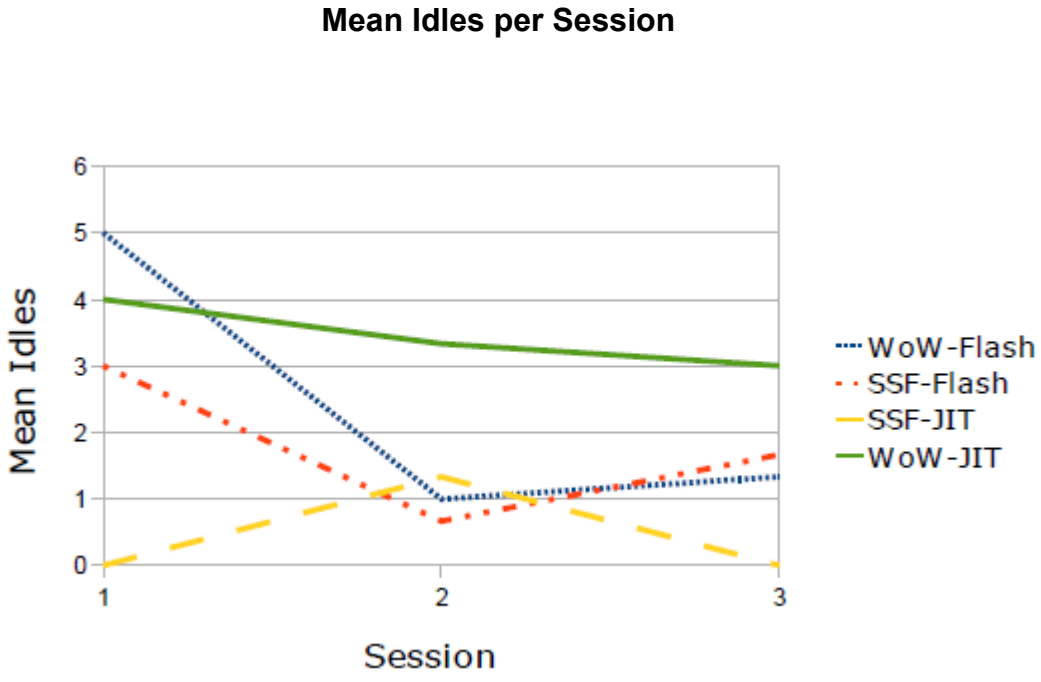

Figure 1: Mean Idles per Session

As gameplay progressed it became evident that SSF-JIT appeared to produce the fewest mean idles compared to the other groups. SSF modality played a greater role in this apparent change, by determining a greater gross number of idles across sessions than did instructional strategy. It can be seen that the *WoW* (visual) modality groups incurred many more idles on average than their SSF counterparts, a sign of lower overall game mastery.

Data analysis – Mean Idles	*t*	*df*	*p*	η^2
Modality	2.05	10	.07	.295
Strategy	-0.16	10	.88	.003

Table 1: Data Analysis for Mean Idles per Session

The independent samples t-tests of both modality and strategy were not statistically significant, though idles by modality did however approach significance at $p = .07$ with an η_2 of .295. It is therefore fair to assume that the *n* of 3 contributed to a lack of power in the statistics. The results however suggestive, cannot be generalized.

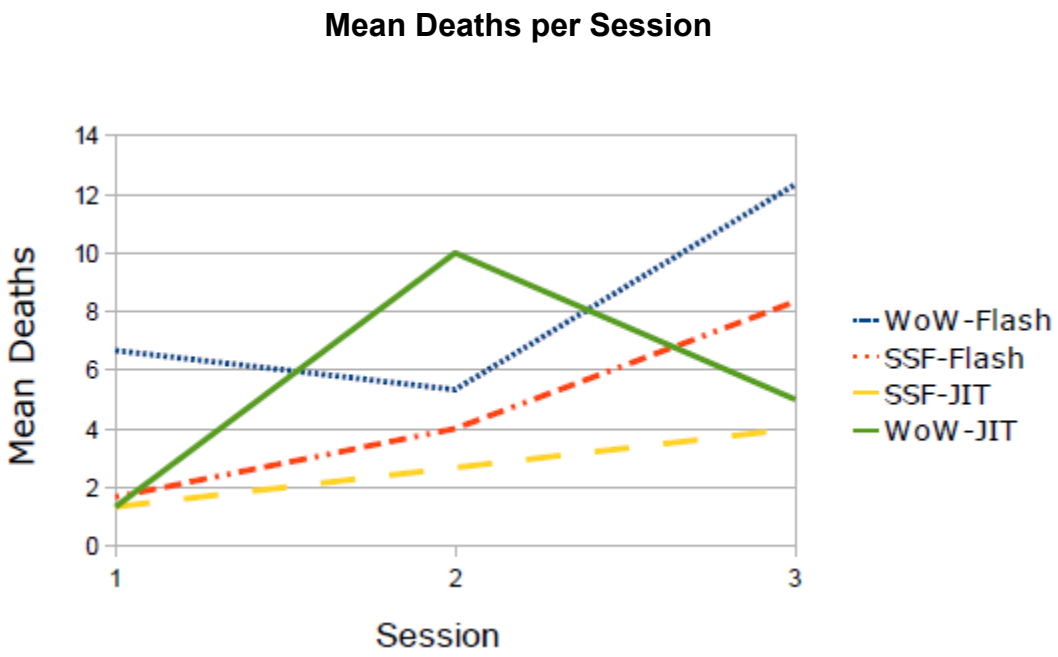

Figure 2: Mean Deaths per Session.

Differences between groups as measured by deaths were not statistically significant. *Figure 2* illustrates the mean deaths per session between groups, which show no apparent pattern. Oddly, deaths in the flashcard method groups seemed to increase, possibly suggesting that these players did not absorb enough content to move into mastery behaviours as the play sessions went on. The independent samples t-test contained in Table 2 examines mean deaths for significance:

Data analysis – Mean Deaths

	t	*df*	*p*	η^2
Modality	1.65	10	.13	.212
Methodology	-1.17	10	.27	.123

Table 2: Data Analysis for Mean Deaths per Session

Again no clear pattern emerges in deaths per session. It is unlikely that modality or instructional strategy played a part in determining the total number of deaths. These are tests of causality examining the differences between tutorials, not relationship between variables.

Gearscore per Session by Group

During the pilot test, it seemed that the modality of the instruction played some role in overall gearscore. In the final test data instructional strategy played some role as well, however small. It appears that the JIT groups emerged with higher overall mean gearscores than their flashcard counterparts, but failed to yield statistical significance.

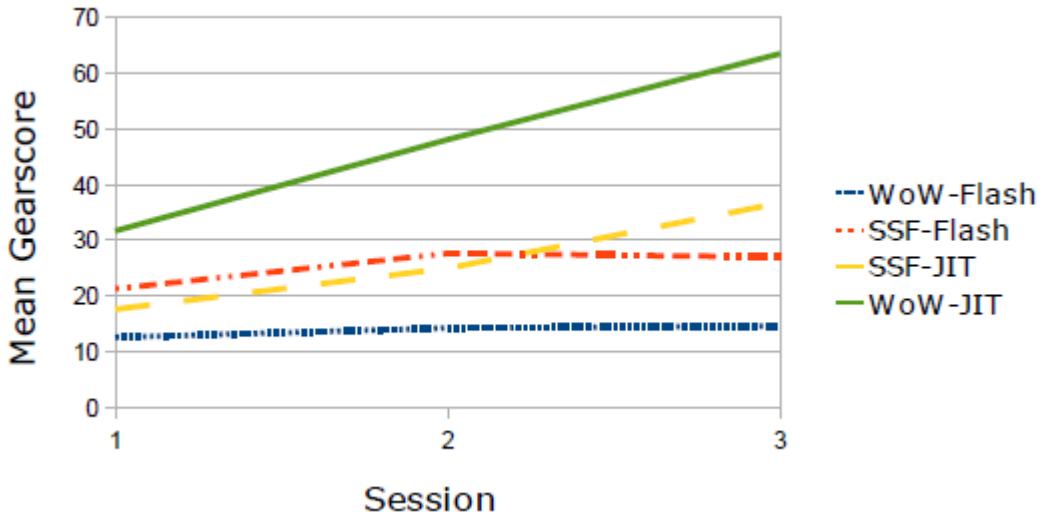

Figure 3: Gearscore per Session by Group

While the JIT groups appear to emerge with higher overall scores, the differences are not statistically significant.

Data analysis – Mean Gearscore

	t	*df*	*p*	η^2
Modality	0.37	10	.72	.014
Methodology	1.71	10	.12	.230

Table 3: Data Analysis for Mean Gearscore per Session by Group

The findings illustrated in Table 3 suggest that neither modality nor instructional strategy as operationalized in this study played a significant role in determining between-groups gearscore.

Experience Gathered per Session by Group

In the pilot test, it appeared that SSF-JIT earned higher experience points per session, and gathered greater total experience than all other groups. Figure 4 illustrates that not only did SSF-JIT continue to gather more experience points per session, but that WoWJIT also gathered a significantly increased number of experience points. This might suggest that the JIT method of instruction does not interrupt players and allows them to more readily gather experience points, a behavior evident of significant play mastery.

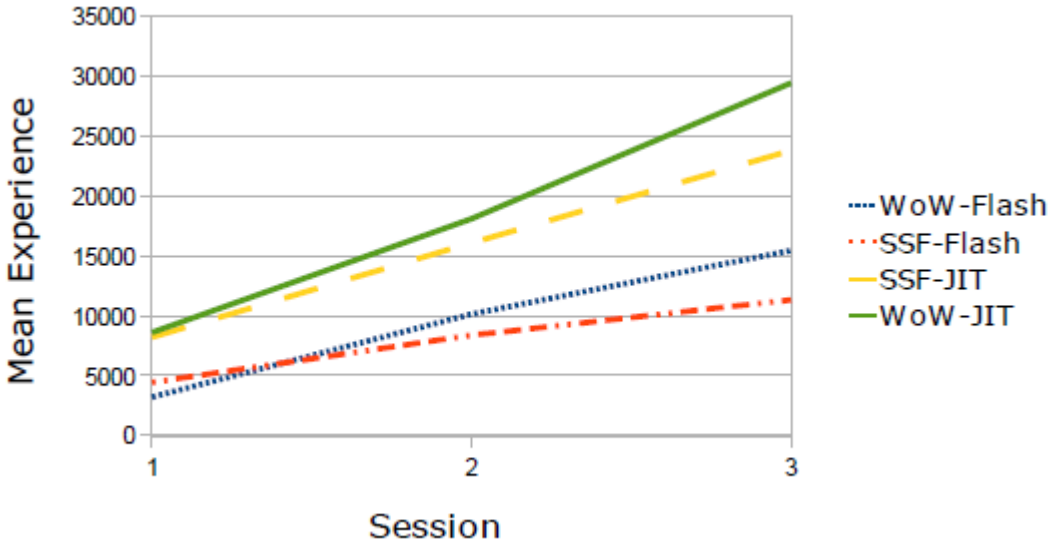

Figure 4: Experience Gathered per Session by Group

Data analysis – Mean Experience

	t	*df*	*p*	η^2
Modality	0.55	10	.59	.029
Strategy	1.67	10	.04	.221

Table 4: Data Analysis for Mean Experience per Session by Group

Table 4 illustrates that there is a significant correlation between strategy and experience gathered. Strategy is a variable transcoded for 1 = Flashcard, 2 = JIT. Modality did not play a significant role in altering the amount of experience gained across limited play sessions. This may indicate that strategy of instruction plays a greater role in play mastery, at least by this measure, than modality.

Engagement Data

Data on player engagement was also gathered as part of the study. This was accomplished by administering a Game Engagement Questionnaire that can be found in Appendix A. The following correlation analyses and independent samples t-tests explore the significance of the engagement findings between-groups. The data contained in Table 5 illustrates that strategy of instruction played a statistically significant role in player engagement, while modality did not. This raises interesting questions about the nature of game engagement as a measure, as significant correlations between player idle times and modality of instruction would seem to indicate a greater level of what educators typically refer to as "engagement", and the type of engagement measured by the questionnaire. Nevertheless, consistent with the development goals outlined for the field test of these new tutorials, it is clear that player engagement was heightened when a just-in-time instructional strategy was presented as an alternative to the WoW-Flashcard tutorial.

Data analysis – Mean Engagement

	t	*df*	*p*	η^2
Modality	0.32	10	.75	.001
Methodology	2.59	10	.03	.694

Table 5: Data Analysis for Mean Engagement by Group

Motivations for Play

Only instructional strategy seems to approach significance in altering player motivations. Motivations by strategy of instruction, however, approaches significance in affecting achievement play motivation (p = .057), and escapism motivations were found to be significant (p = .011). This finding may provide further evidence that instructional fading offered fewer interruptions by showing that escapism, a potential measure of immersion and engagement with the game system, as well as achievement oriented behaviors are increased when a minimally invasive strategy is employed. Table 6 provides an overview of the descriptive statistics for all of the quantitative analyses described in this chapter.

Descriptive statistics for quantiative analyses.

Variable	*n*	$\overline{X}$	*S*
Total Idles	12	6.1	5.3
Total Deaths	12	15.7	10.5
Final Gearscore	12	35.4	32.0
Final Experience / 1000	12	20.0	14.8
Social Motivations	12	17.3	5.6
Achievement Motivations	12	21.2	5.6
Exploration Motivations	12	22.3	4.6
Escapism Motivations	12	30.7	6.5
Grief Motivations	12	10.3	4.1
Engagement	12	101.8	27.8

Table 6: Overall Data Analysis

Discussion

A problem was identified in the ability of novice players to gain mastery in contemporary digital games. In response to this, a needs assessment known as the D.E.C.L. (Mann, 2006) was performed. This needs assessment interrogated the effectiveness of the tutorials within contemporary digital games by examining their instructional strategies, content, and modalities. It was decided that based on prior research on learning from media (Mann, 2009; Mayer, 2010), as well as considerable debate about novice play (Jenson & de Castell, 2010; Jenson, Fisher & de Castell, 2011), alterations to the tutorials might elicit change in player engagement and motivations toward the game, as well as player skill level as illustrated by play mastery

behaviors. Instruments were sought to measure these differences (Brockmyer et al., 2009; Yee, 2006), and Lua language modifications were made to *World of Warcraft*. Three new tutorials, SSF-Flashcard, SSF-JIT, and WoW-JIT, as well as the default WoW-Flashcard, were formatively evaluated using the Dick, Carey, and Carey (2011) model. The developmental goals for the field study assumed that a greater level of play skill, as evidenced through altered motivations for play and greater play mastery, would be present in the SSF-JIT group than any of the others. While this did not strictly hold true, it was found that the SSF-JIT group incurred fewer idles, gathered more experience, and had slightly altered motivations for play. Modality seemed to have a greater effect on play mastery, while instructional strategy had a greater effect on motivations for play.

Consistent with research on contemporary digital games, participants in the study reported high levels of power and achievement motivations and were able to achieve significant play mastery by the end of the play sessions (Jenson & de Castell, 2008, 2010). In fact, as the players moved from the initial to final play sessions, mastery behaviors began to become evident, and in-game behaviors began to shift from exploration and disengagement, to fighting back and slaying monsters, which is consistent with prior research (Jenson & de Castell, 2010; Jenson et al., 2011). This can be seen in the curve of deaths by session illustrated in *Figure 2*. As players gained mastery, they began to learn to fight back when attacked instead of fleeing: a behavior that frequently leads to death in *World of Warcraft*. Participants in the SSF-JIT category had the fewest idles and deaths, and a similar amount of gearscore and experience (measures of play mastery) to the WoW-JIT group. Significantly, the WoW-Flashcard group, as expected, performed the worst in all measures of play mastery, while the JIT instructional strategy was found to have an effect on play motivation and reported "flow" responses (Csikszentmihayli, 2008). Interestingly, the altered escapism motivations and subjective reporting of "zoning out" did not correlate with increased engagement as measured by the game engagement questionnaire. This may suggest that engagement as measured by the game engagement questionnaire, and engagement as it is typically measured in education may not correlate meaningfully, or that the game engagement questionnaire needs to be re-evaluated.

Practical Implications of the Study

Statistically significant results were present in the reduction of idle-time through changing the modality of instruction from stochastic visual cues to auditory cues delivered via the SSF method of instructional design (Mann, 2008). These results may suggest that future game studies researchers, educators, and game designers might consider implementing in-game instructional systems and tutorials using an auditory modality delivered via such an instructional method. Furthermore, this may demonstrate continued evidence for the use of digital games for learning and teaching, as current multimedia learning literature already strongly advocates the use of sound in instruction (Mann, 2009; Mayer, 2010). This work provides additional evidence that auditory traces are more resistant to interference and less likely to be forgotten (Mann, 2006). A new finding is that these auditory cues also seem to cause less disengagement from an otherwise unfamiliar task in novice learners when compared to visual cues, which in this study resulted in increased mastery behaviors in-game.

Conclusion

This study found results that agree with previous researchers, and new results that have not yet been elucidated elsewhere. For example, it was clear in this study that auditory cues were more durable and resistant to misinterpretation and forgetting than visual cues (Mann, 2006). Furthermore, split attention caused poor results in the WoW-Flashcard group, which overloaded the visual channel, consistent with Mayer (2010) and others. The participants were found to have domination and mayhem motivations during final play sessions and post-play motivation interviews. This diverges from the work of previous researchers who have questioned the nature of female motivation and competitiveness (Graner Ray, 2004), when one considers the all female sample. These findings also provide evidence that expert status is a strong indicator of digital game play habits (Jenson, Fisher & de Castell, 2011). Novel to this research is the discovery that shifting from a visual to auditory modality increased play mastery significantly, and that instructional strategy plays a significant role in determining player motivations. These results present interesting implications for not only the design of tutorials for digital games, but also for educators who may now be able to meaningfully select digital games for learning and teaching based on their tutorial designs.

References

Adams, S., Mann, B., & Schulz, H. (2006). Can seventh-graders learn fractions from a web-based pedagogical agent? Using comparison groups three times over several weeks. In B. Mann (Ed.), *Selected Styles in web-based educational research* (pp. 332-346). Hershey, PA: Information Science Publishing.

Akilli, G. (2007). Games and simulations: A new approach in education. In C. Aldrich, M. Prensky, & G. Gibson (Eds.), *Games and simulations in online learning* (pp. 1-20). Hershey, PA: IGI Global.

Ang, C. S., & Zaphiris, P. (2010). Social roles of players in MMORPG guilds. *Information, Communication & Society*, *13*, 592–614.

Bartle, R. (1996). Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD Research, 1*(1). Retrieved from: http://www.mud.co.uk/richard/hcds.htm

Becker, K. (2007). Battle of the titans: Mario vs. MathBlaster. In C. Montgomerie & J. Seale (Eds.), *World conference on educational multimedia, hypermedia and telecommunications 2007* (pp. 2707–2716). Vancouver, Canada: AACE.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, *45*, 624–634.

Burton, J., Moore, D., & Magliaro, S. (1996). Behaviourism and instructional technology. In D. H. Jonassen (Ed.), *Handbook for research for educational communications and technology* (pp. 3- 36). New York, NY: Simon & Schuster MacMillan.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*, 293–332.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*, 233–246.

Cheng, G. (2009). Using game making pedagogy to facilitate student learning of interactive media. *Australasian Journal of Educational Technology*, *25*(2), 204–220.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*, 149–210.

Clark, R. (2007). Learning from serious games? Arguments, evidence, and research suggestions. *Educational Technology*, *47*(3), 56–59.

Conway, S. (2010). Hyper-lucidity, contra-lucidity, and the digital game. *Eludamos, Journal of Computer Game Culture.*, *4*(2), 135-147.

Csikszentmihalyi, M. (2008). *Flow : The psychology of optimal experience*. New York, NY: Harper Perennial.

Dennen, V. (2004). Cognitive apprenticeship in educational practice: Research on scaffolding, modeling, mentoring, and coaching as instructional strategies. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 813-828). New York, NY: Simon & Schuster McMillan.

Dick, W., Carey, L., & Carey, J. O. (2011). *The systematic design of instruction*. Upper Saddle River, NJ: Merrill/Pearson.

Egenfeldt-Nielsen, S. (2011, February). What makes a good learning game? Retrieved from eLearn Magazine: http://elearnmag.acm.org/archive.cfm?aid=1943210

Entertainment Software Association of Canada. (2011). *Essential facts about the Canadian video game industry 2011*. Retrieved from http://www.theesa.ca/wp-content/uploads/2011/10/Essential- Facts-2011.pdf

Ericsson, K. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. Ericsson (Ed.), *The Cambridge handbook of expertise and expert performance* (pp. 223-241). Cambridge, UK: Cambridge University Press.

Graner Ray, S. (2004). *Gender inclusive game design : expanding the market*. Hingham, MA: Charles River Media,.

Greenberg, B. S., Sherry, J., Lachlan, K., Lucas, K., & Holmstrom, A. (2008). Orientations to video games among gender and age groups. *Simulation & Gaming*, *41*, 238–259.
Harrigan, P. (2010). *SecondPerson : Role-playing and story in games and playable media*. Cambridge, MA: MIT Press.

Hayes, E. (2005). Women, video gaming and learning: Beyond stereotypes. *TechTrends*, *49*, 23–28.

Jenson, J., & de Castell, S. (2008). Theorizing gender and digital gameplay: Oversights, accidents and surprises. *Eludamos: Journal for Computer Game Culture*, *2*(1), 15–25.

Jenson, J., & de Castell, S. (2010). Gender, simulation, and gaming: Research review and redirections. *Simulation & Gaming*, *41*, 51–71.

Jenson, J., Fisher, S., & de Castell, S. (2011). Disrupting the gender order: Leveling up and claiming space in an after-school video game club. *International Journal of Gender, Science, and Technology*, *3*(1). Retrieved from: http://genderandset.open.ac.uk/index.php/genderandset/article/viewFile/129/258

Kaufman, R. (1993). *Needs assessment : a user's guide*. Englewood Cliffs, NJ: Educational Technology Publications.

Kim, J. W., Han, D. H., Park, D. B., Min, K. J., Na, C., Won, S. K., & Park, G. N. (2010). The relationships between online game player biogenetic traits, playing time, and the genre of the game being played. *Psychiatry Investigation*, *7*, 17-23. Kinser, A. (2004). *Computer assisted remediation for at risk nursing students* (Unpublished doctoral dissertation). Pepperdine University, Malibu, CA.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75–86.

Mann, B. (1994). *Effects of temporal sound on computer-based learning* (Unpublished doctoral dissertation). Ontario Institute for Studies in Education, Toronto, Canada.

Mann, B. (2005). Making your own educational materials part I. *International journal of instructional technology & distance learning*, *10*(2). Retrieved from: http://itdl.org/journal/dec_05/article02.htm

Mann, B. (2006). Making your own educational materials part II: Multimedia design for learning. *International Journal of Instructional Technology & Distance Learning*, *10*(2). Retrieved from: http://www.itdl.org/Journal/Oct_06/article02.htm

Mann, B. (2009). Computer-aided instruction. In B. Wah (Ed.), *Wiley Encyclopedia of computer science & engineering* (pp. 583–592). Hoboken, NJ: John Wiley & Sons Inc.

Mann, B. L. (1997a). Shifting attention in multimedia: Stochastic roles, design principles and the SSF Model. *Innovations in Education & Training International*, *34*, 174–187.

Mann, B. L. (1997b). Evaluation of presentation modalities in a hypermedia system. *Computers & Education*, *28*(2), 133–143.

Mann, B. L. (2008). The evolution of multimedia sound. *Computers & Education*, *50*, 1157–1173.

Mann, B., Newhouse, P., Pagram, J., Campbell, A., & Schulz, H. (2002). A comparison of temporal speech and text cueing in educational multimedia. *Journal of Computer-Assisted Learning, 18*(3), 296-308.

Mann, B. L., Schulz, H., Cui, J., & Adams, S. (2012). Sustained learning in 4th and 5th graders but not 7th graders: Two experiments with a talking pedagogical agent. In J. Jia (Ed.), *Educational stages and interactive learning: From kindergarten to workplace training* (pp. 163-190). Hershey, PA: IGI Global.

Mayer, R., & Moreno, R. (1998). A cognitive theory of multimedia learning: Implications for design principles. Paper presented at the *15th ACM SIGCHI Conference on Human Factors in Computing Systems (CHI-98)* Los Angeles, CA, April 18-23. Retrieved from: http://unm.edu/~moreno/PDFS/chi.pdf

Mayer, R. (2003). *Multimedia learning*. Cambridge, UK: Cambridge University Press.

Mayer, R. (2010). Seeking a science of instruction. *Instructional Science*, *38*(2), 143–145.

Merrill, M. D., & Twitchell, D. (1994). *Instructional design theory*. Englewood Cliffs, NJ: Educational Technology Publications.

Miyamoto, S. (2010). Nintendo Guru Targets Education. *The Associated Press*, (CBC News). Retrieved from http://www.cbc.ca/technology/story/2010/03/19/nintendo-classroom.html?ref=rss

Nieborg, D. (2011). *Triple-A: The political economy of theblockbuster video game.* (Unpublished Doctoral dissertation). Universeteit van Amsterdam, Amsterdam, The Netherlands.

onlineschools.org. (2012). *The unbelievable World of Warcraft* [Infographic], Retrieved from http://www.onlineschools.org/blog/unbelievable-wow

Remo, C. (2009). ESA, Sony, Microsoft respond to Obama's call for STEM education. *Gamasutra*. Retrieved from http://www.gamasutra.com/view/news/26203/ESA_Sony_Microsoft_Respond_To_Obamas_Call_For_STEM_Education.php

Ritchie, J., & Lewis, J. (2003). *Qualitative research practice: A guide for social science students and researchers.* New York, NY: Sage Publications.

Seels, B., Fullerton, K., Berry, L., & Horn, L. J. (2004). Research on learning from television. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 249-334). Hillsdale, NJ: Lawrence Erlbaum Associates.

Shelton, B. E., Satwicz, T., & Caswell, T. (2011). Historical perspectives on games and education from the learning sciences. *International Journal of Game-Based Learning*, *1*, 83–106.

Stufflebeam, D. (1985). *Conducting educational needs assessment*. Boston, MA: Kluwer-Nijhoff.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. E*ducational Psychology Review*, *22*, 123–138.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, *119*, 176–192.

Sweller, J., Kirschner, P. A., & Clark, R. E. (2007). Why minimally guided teaching techniques do not work: A reply to commentaries. *Educational Psychologist*, *42*, 115–121.

Taylor, N. (2007). Mapping gendered play. *Loading...*, *1*(1), 24–31.

Thomson, G. (2009). *Orientation through playing: Improving the "new user experience" of virtual worlds through video game tutorial design.* (Unpublished Honour's Thesis). University of Huddersfield, West Yorkshire, England.

Tobias, S. (2011). *Computer games and instruction*. Albany, NY: State University of New York.

Tosca, S. (2003). The quest problem in computer games. In S. Goebel, N. Braun, U. Spierling, J. Dechau, & H. Diener (Eds.), *Proceedings of technologies for interactive digital storytelling and entertainment* (pp. 69-82). Stuttgart, Germany: Fraunhofer IRB Ferlag.

Wang, H. Y., & Wang, Y. S. (2008). Gender differences in the perception and acceptance of online games. *British Journal of Educational Technology, 39*(5), 787-806.

White, M. (2009). The senescence of creativity: How market forces are killing digital games. *Loading...*, *3*(4). Retrieved from: http://journals.sfu.ca/loading/index.php/loading/article/viewArticle/54

Witkin, B. (1984). *Assessing needs in educational and social programs*. San Francisco, CA: Jossey-Bass Publishers.

Yee, N. (2006). The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and Virtual Environments*, *15*, 309–329.

Yee, N., Ducheneaut, N., Nelson, L., & Likarish, P. (2011). Introverted elves & conscientious gnomes. Paper presented at *ACM CHI Conference on Human Factors in Computing Systems,* 2011 May 7 – 12, Vancouver, Canada.