



1-1-2012

Egress Online: Towards Leveraging Massively, Multiplayer Environments for Evacuation Studies

Aline Normoyle

University of Pennsylvania, alinen@seas.upenn.edu

John Drake

University of Pennsylvania, drake@seas.upenn.edu

Alla Safonova

University of Pennsylvania, alla@cis.upenn.edu

Follow this and additional works at: https://repository.upenn.edu/cis_reports

Recommended Citation

Aline Normoyle, John Drake, and Alla Safonova, "Egress Online: Towards Leveraging Massively, Multiplayer Environments for Evacuation Studies", . January 2012.

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-12-15.

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/cis_reports/977
For more information, please contact repository@pobox.upenn.edu.

Egress Online: Towards Leveraging Massively, Multiplayer Environments for Evacuation Studies

Abstract

Large datasets of real human behaviors are of huge benefit across numerous domains, including evacuation safety, urban planning, marketing, and ergonomics. However, because large-scale experiments involving real human subjects are expensive and prohibitively difficult to organize, such datasets are scarce. Thus in this paper, we propose the use of massively multiplayer online (MMO) communities as an inexpensive and innovative way to capture datasets of large numbers of people under different conditions. We describe our implementation of an online data collection system, based on games, inside the popular massively multiplayer, online environment of Second Life. We evaluate the use of this system for performing evacuation experiments using a mix of Second Life residents and players recruited on campus. Our system was able to draw online participants, support data collection needs, and provide potential insights into high-level evacuation behaviors such as the choices of exit, effects of building debris, and the use-patterns of a building. Through experiments performed using our system, we found that Second Life residents found the game controls and environment to be significantly more compelling than lab participants; that players unfamiliar with our office building tended to evacuate primarily via the front entrance; and that in-game debris significantly increased the numbers of participants who failed to exit a building safely.

Keywords

Serious Games, User studies and evaluation, Behavior capture, Building evacuation.

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-12-15.

Egress Online: Towards leveraging massively, multiplayer environments for evacuation studies

Aline Normoyle¹, John Drake², and Alla Safonova

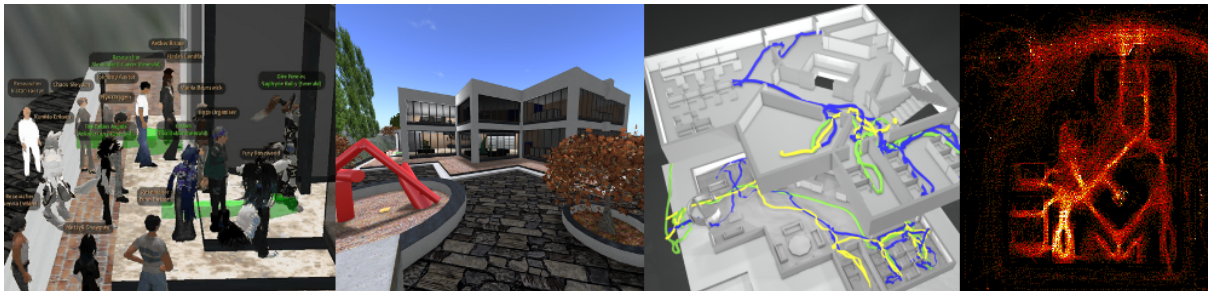


Figure 1: From Left to Right: Participants from Second Life wait to start an experiment. The fictional office building used for fire evacuation experiments. Playback of captured trajectories in Maya. A frequency map showing high usage areas around the building, collected over many sessions.

Abstract

Large datasets of real human behaviors are of huge benefit across numerous domains, including evacuation safety, urban planning, marketing, and ergonomics. However, because large-scale experiments involving real human subjects are expensive and prohibitively difficult to organize, such datasets are scarce. Thus in this paper, we propose the use of massively multiplayer online (MMO) communities as an inexpensive and innovative way to capture datasets of large numbers of people under different conditions. We describe our implementation of an online data collection system, based on games, inside the popular massively multiplayer, online environment of Second Life. We evaluate the use of this system for performing evacuation experiments using a mix of Second Life residents and players recruited on campus. Our system was able to draw online participants, support data collection needs, and provide potential insights into high-level evacuation behaviors such as the choices of exit, effects of building debris, and the use-patterns of a building. Through experiments performed using our system, we found that Second Life residents found the game controls and environment to be significantly more compelling than lab participants; that players unfamiliar with our office building tended to evacuate primarily via the front entrance; and that in-game debris significantly increased the numbers of participants who failed to exit a building safely.

¹alinen@seas.upenn.edu

²drake@seas.upenn.edu

Keywords: Serious Games, User studies and evaluation, Behavior capture, Building evacuation.

Note: This work was supported by a University Research Foundation grant from the University of Pennsylvania.

1 Introduction

Large datasets of real human behaviors would greatly assist in solving problems across numerous domains, such as the evaluation of emergency procedures, the prediction of consumer activities using different marketing strategies, the evaluation of use-patterns of urban structures, and the simulation of more realistic crowds.

Unfortunately, datasets of real human behaviors are scarce due to the difficulties – legal, ethical, logistical and financial – of capturing human behaviors under realistic, controlled conditions. For example, although video recording devices for capturing data are ubiquitous, extracting human behaviors from raw video remains time consuming and impractical. In particular, tracking human subjects in crowded places and under different lighting conditions can be problematic, especially across multiple cameras. Additionally, data from video lacks annotations describing the intentions of recorded subjects, which would help with data analysis. Beyond the problems of recording people, we also require the ability to run controlled experiments, which in addition to requiring that a building already be built, often can only be performed infrequently and at small scales due to the cost and logistics of their setup. Finally, for some situations, such as evacuation under emergency conditions, obtaining large datasets is nearly impossible, a point we motivate in the following section.

Thus, we explore an alternative method for obtaining large datasets of real human behaviors. This method leverages the millions and millions of hours spent by real people online in 3D virtual communities such as Second Life, BlueMars, Kaneva and others. Millions of virtual residents inhabit these online communities: socializing, throwing parties, playing games, trading and conducting business. In addition to being large-scale, these environments are configurable and support the labeling of subjects (e.g., with usernames or anonymous ids). Thus, the conditions of experiments can be controlled, and tracking and disambiguating human subjects is trivial.

Online communities therefore provide us with a unique opportunity to collect large datasets of human behaviors. However, several challenges must be addressed before such data collection can be effective. First, how can researchers motivate participants to act as required (for instance, when participating in an evacuation experiment)? Second, in an environment where people have countless choices for activities, how can researchers attract people to their experiments? Third, how can researchers analyze the fidelity of the experimental setup to ensure the captured human behaviors contain meaningful insights?

1.1 Motivating Example: Building Evacuation

The design of any structure – shopping malls, airports, theaters, stadiums, or entire city blocks – must consider the structure’s suitability for evacuation. People’s lives depend on the ability to exit a building quickly in cases such as fires, earthquakes, and terrorist attacks. Clearly, evacuation procedures must be evaluated as part of the design process; yet, robust and accurate design tools for evacuation and disaster response do not exist, particularly because such tools would rely on the ability to robustly predict crowds and human interactions. Organizing live experiments is clearly impractical: experiments would require thousands of people evacuating all possible structures under all possible emergency conditions.

Validation of crowd simulators relies predominately on tracked data based on real crowd movements, such as during evacuation drills, which are then compared against the simulation. However, an evacuation drill does not mimic a realistic emergency situation and the data obtained from real recorded scenarios and tracked crowd movements do not generalize easily to different structures, which may have their own complex spatial characteristics. Research into tracking crowds from video, such as from security or surveillance cameras, could benefit building evaluation and design (Lee et al., 2007). However, the problem of collecting large amounts of data to study evacuation is still difficult: emergencies do not happen often; cameras may not be installed; and even when cameras are installed, emergencies such as fires often damage both cameras and footage.

In this paper, we show how we address these challenges using games during which participants must evacuate a building to escape a fire, implemented inside Second Life. The benefits of using 3D online communities, such as Second Life, is that we can leverage the thousands of hours people already spend online to test arbitrary building designs with large numbers of people. A system such as ours could be a low cost complement to the simulation tools available to architects and safety engineers for designing safe, comfortable structures. Because participants are *real people*, they show the variety of personalities, preferences, and styles of real people. Thus, our online system can provide insights and help generate hypotheses into crowd reactions and movements – some of which may be more realistic than the reactions produced by simulation.

In the following sections, we describe our approach to the game design, whose goals are to attract sufficient numbers of online residents; to support data collection requirements; and to provide an environment which addresses the needs of the experiment. We then evaluate the system using several independent criteria: a domain-independent evaluation of the ability of the system to attract and engage participants and the extents to which the collected behaviors displayed common social norms, and a domain-dependent evaluation of the abil-

ity of the system to provide insights into fire evacuation. Particularly, we conducted three experiments related to egress behaviors identified by fire safety researchers: the tendency for people to prefer known exits for egress, the effects of debris and blocked exits for hindering egress, and the phenomenon of task fixation, where people delay egress until after they finish an activity.

2 Background and Related Work

Simulation-based evaluation has long been an invaluable alternative to real-life experiments. In fact, numerous crowd simulators have been developed (see surveys by Pelechano et al. (2008) and Thompson & Marchant (1995)). Unfortunately, because no real data is available, it is difficult to evaluate the fidelity and robustness of a simulator, especially with respect to evacuation scenarios. For example, most current crowd simulators are based on cellular automata. Such simulators rely on discrete grids which can produce fixed densities and unrealistic flow rates through portals. Agent speeds are typically derived from ideal conditions while agent paths can include distorted lengths and travel times due to the underlying grid (Andersen et al., 2005). These factors clearly affect the realism of the simulators, but to what extent is unclear.

Early evacuation research focused on architectural features affecting the ease and speeds with which people can move through corridors, stairs, and doorways. This work has led to the development of calculators for computing structure egress times based on occupant density which are important for the legislation and development of building regulations (Kobes & Post, 2010). For example, the International Building Code (IBC, 2008) dictates minimum staircase sizes and minimum numbers of exits based on these findings. Current research has extended its focus to human factors regarding how people use structures, choose exits, and make decisions regarding when and how to evacuate. Because of the difficulty of performing behavioral experiments, current research is predominately based on reports and survivor interviews collected after disasters occur (Kobes & Post, 2010; Johnson, 2005). Such interview data is extremely valuable, but we wish to enhance tools so that we might generate hypotheses and insights before such disasters occur.

Historically, virtual reality and simulation have long been a means for reproducing dangerous and expensive scenarios for the purposes of training, design, and manufacturing. In the case of architecture and evacuation, realistic models already exist to accurately describe the spread of fire and the corresponding deterioration of building materials (Bukowski & Squin, 1997). However, these simulators make use of existing crowd simulation systems and therefore suffer from the same set of limitations described above.

The emergence of realistic, massively multiplayer virtual environments offers several unique and exciting opportunities for research (Blascovich et al., 2002; Castronova, 2006; Hendaoui et al., 2008; Bainbridge, 2007). Blascovich et al. (2002) discuss the potential of online worlds to alleviate problems of laboratory social research, thanks to the ability to create controlled experiments, their potential for realistic content, and the potential to draw participants in greater numbers who are more representative of the population than college students.

In particular, virtual communities provide unique opportunities for collaboration – especially among strangers. Harnessing the Internet to coordinate massive human efforts has its roots in human computing, which strives to take advantage of the great numbers of people currently online to perform tasks either too difficult for computers or too large and too tedious for a single person to do by hand. Human computing systems can pay people, such as Amazon’s Mechanical Turk; can rely on hobbyists and enthusiastic volunteers, such as the Hubble telescope Galaxy Zoo (2010) galaxy classifier or bird watchers’ logging site, eBird (2010); or can rely on games for recruiting participation. Games have several potential advantages. They can disguise tedious tasks such as image labeling and text translation as entertaining and often addictive activities (von Ahn, 2006); can provide mechanisms for checking participant answers; and can motivate players through prizes and competition. Additionally, Cooper et al. (2010) have demonstrated how a game can facilitate collaboration between non-experts to help solve real scientific problems. However, designing such games is non-trivial and requires much iteration and play-testing to both appeal to players and provide useful data (Cooper et al., 2010). Whereas previous work has looked at how dedicated online sites and games can be used for collaboration, this work investigates how existing virtual communities might be leveraged for human collaboration.

Because of the ease with which users can build original content, Second Life has proved to be useful in the domains of simulation, education, and training. The Children’s Memorial Hospital in Chicago used Second Life (Linden Labs, 2009) to plan and rehearse evacuation protocols. These protocols cannot be practiced otherwise, because they entail clearing wings of the hospital which would disrupt patient care. Stott (2007) looked at the use of Second Life to train medical professionals, including its use in treating people with post traumatic stress disorder and phobias. The University of Arkansas’s “Modeling Healthcare Logistics” project (University of Arkansas, 2010) uses Second Life to prototype and test logistics models. Harvard University’s River City Project developed games in Second Life to teach middle school kids about disease transmission and the scientific method (The River City Project, 2007).

Our use of virtual online communities has been motivated by previous research which has found that a surprising number of human behaviors in virtual worlds mimic the real world. A 2007 survey found that people’s avatars look

similar to themselves 65% of the time and have the same gender 77% of the time (de Gortari, 2007). In particular, Second Life residents often maintain a consistent online persona as well as maintain online friendships. Failing to adhere to rules of conduct can have long term consequences and evoke real emotions in members (de Gortari, 2007). Yee & Bailenson (2007) have shown that avatar attractiveness affects extroversion and height effects aggressiveness in negotiation similarly to trends in real life. Concerning nonverbal communication, Yee et al. (2007) studied interpersonal distance and eye contact and found correspondences with real life. In particular, Yee et al. (2007) found that interpersonal distances between males were typically larger than between females, that eye contact occurred less frequently between males than between females; and that decreases in interpersonal distance were compensated with gaze avoidance. Friedman et al. (2007) also studied proxemics and found similar results even when one of the interacting avatars was controlled by a Bot rather than a real person.

3 Design of MMO experiments

Second Life has millions of members, approximately 50,000 of which are online throughout the day (Second Life Statistics, 2012). Second Life members often maintain a consistent online persona, contribute to a virtual economy, and spend the majority of their time exploring, socializing, and customizing their belongings and appearance (de Gortari, 2007). In the following sections, we describe our experimental setup for running controlled evacuation studies in a massively multiplayer online environment.

3.1 System Design and Data Collection

Our experimental setup consists of a two-story office building and grounds, constructed using the built-in modeling capabilities of the Second Life viewer (Figure 2). Additional game mechanics are implemented with the Linden Scripting Language (LSL). We collect experiment data using custom Second Life avatars, or bots, which we built using the libOpenMetaverse API (libOpenMetaverse, 2012) – similar to the implementation in La & Michiardi (2008). Bots were invisible to players and collected chat, trajectories as (x,y,z) coordinates, and experiment events to a MySQL database. The trajectory bot recorded avatar trajectories within a given radius at a resolution of 0.5 meters. Since user chat is received within a limited radius, chat detectors were scattered throughout the experiment site and forwarded messages to a bot for database entry.

All participants were given access to a tutorial area where they could practice

controlling their characters, opening doors, and picking up objects. Once participants were ready, an experiment coordinator (in Second Life) would ask participants to gather at the building's front entrance. There, participants would be asked to sign an electronic IRB form, told the rules for the upcoming game, and given a heads up display which showed their current health level. Participants were told to keep their health as high as possible. The Second Life experimenter would then initiate the experiment and tell everyone to start.

After the experiment, players filled out a questionnaire and received prizes and Linden dollars based on their performance.

3.2 Participants

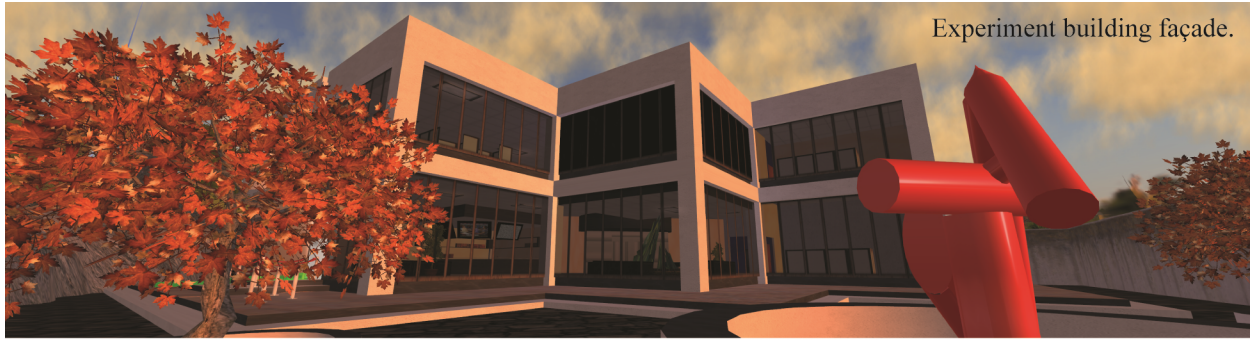
Our participants were a mix of Second Life residents and students recruited to play Second Life on campus. To advertise experiments to Second Life residents, we built an announcement and sign-up board on our land, installed a similar board in a popular public location, posted events on the Second Life event calendar, and hosted a Second Life group where fans could receive notifications about experiments.

Over a period of two months, we performed two series of data collection. In the first series, we performed 7 experiments over 3 days, and drew 63 participants (39 Second Life, 24 lab). For these experiments, participants were only allowed to play once and thus were unfamiliar with the building. In the second series, we performed 15 experiments over 4 days and drew over 100 participants (59 Second Life, 50 lab). In the second series, we allowed players to replay.

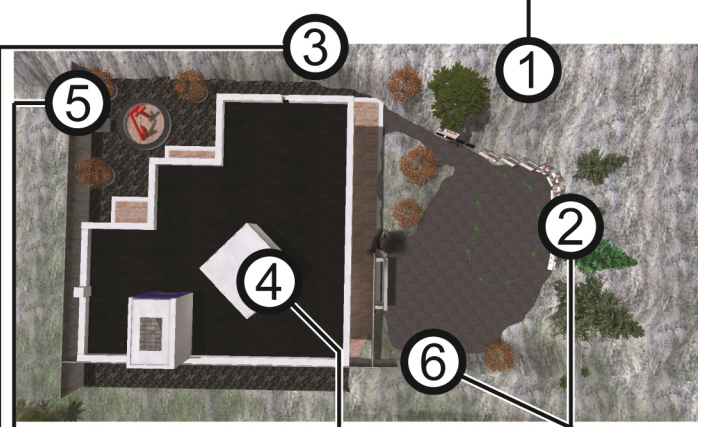
3.3 Game Design

Framing experiments as games strives to engage participants while performing the serious matter of testing hypotheses. Importantly, game rules should encourage people to provide meaningful data and should provide sufficient challenge and competition to encourage people to replay. Additionally, the type and style of the game needs to appeal to Second Life users (to attract participants), support our data collection needs, and provide insights into experimental questions.

Environment and Rules: For this study, we developed two very simple games. In both, players are invited to an office building (Figures 3(a) and 3(c)) where they are given a health bar and warned that it will drain if they become injured. If the health bar drains to zero, the player is immediately transported outside the building. In the simplest game, which we name "The Waiting Game", we ask players to find a specific room in the building and gather there. In the

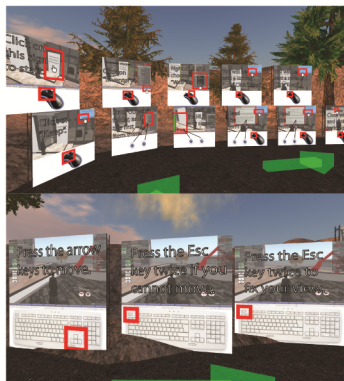


1 This sign advertises the experiment and allows people to register for reminders.



Overhead view.

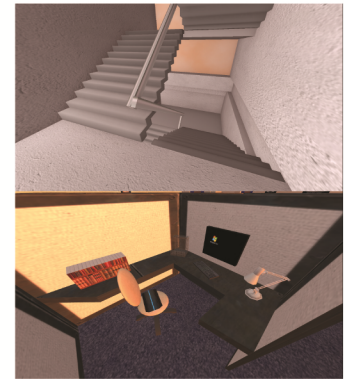
2 Participants first visit this tutorial space with clear visual instructions.



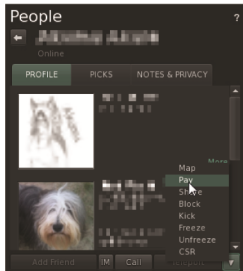
3 Each participant must attach the “HUD” object and accept the terms of a consent form before starting the experiment.



4 Participants enter the experiment to achieve the set goals.



5 At the end of an experiment, each participant receives their reward, a debriefing explaining the research, and a link to a questionnaire.



6 The recording “bots” are hidden, but nearby.



Figure 2: Overview of our system.

second game, named "The Collection Game", the players are encouraged to explore the building and gather parts for a virtual jet pack, which they could assemble and keep after the experiment.

Asking players to find a specific room allows us to capture the effects of people leaving a room together when the evacuation occurs. Conversely, the collection game encourages participants to explore the building and learn its layout.

To discourage cheating, participants are not allowed to fly while in the building and are disqualified for teleporting.

Fire and Evacuation: In both games, we trigger a fire throughout the building approximately 10 minutes after the game starts. To simulate emergency conditions, we implemented the simulation of fire and smoke within Second Life (Figure 3(d)) including how fire propagates within the building over time.

Our simulated fire spreads rapidly and is highly damaging, giving participants approximately 100 seconds to evacuate. This is not typical but also not unrealistic: fire tragedies, such as the Cocoanut Grove fire (Boston, 1942) and the Station Nightclub fire (Rhode Island, 2003) resulted when fires spread very quickly after detection. In particular, video footage of the Station Nightclub fire showed the building engulfed in flames within a mere 90 seconds after ignition. Researchers later reproduced the disaster in a mock-up of the building and empirically measured the building reaching deadly levels of heat flux within 61 seconds, deadly temperatures within 76 seconds, and insufficient oxygen within 87 seconds (Grosshandler et al., 2005).

After the fire is triggered, alarms sound and players must escape before their health drains to zero (Figure 3(b)). Fire drains the health of the players proportionally to how close they are to it.

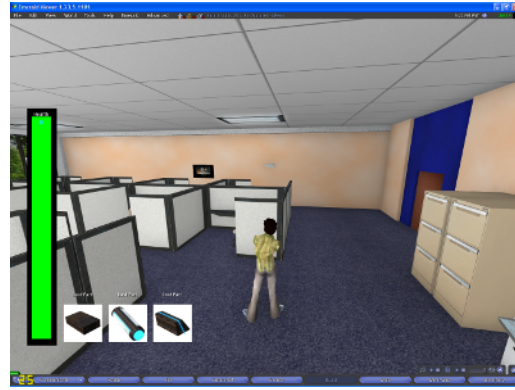
Incentives: Both games provide several incentives to encourage participants to exit as quickly as possible. Foremost, the scenario encourages people to evacuate quickly due to the shrill alarm, the presence of other evacuating participants, and the easily seen health level. Some users, however, might ignore these factors out of curiosity. Thus, we also award small prizes to those participants who maintain the most health (and collect the most jet pack parts if applicable). Prizes are also awarded to a daily winner.

4 Evaluation

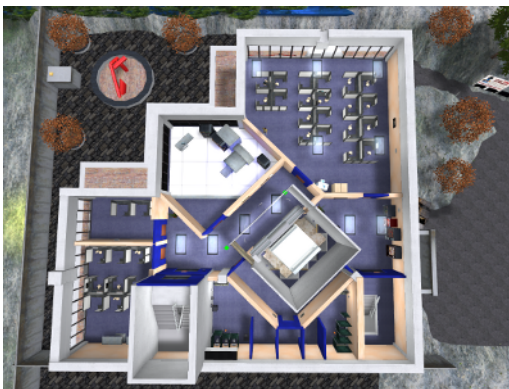
We wish to assess the viability of a Second Life setup for attracting participants and providing useful insights with such a system. For example, what types of behaviors can such a system capture? What types of questions can such a



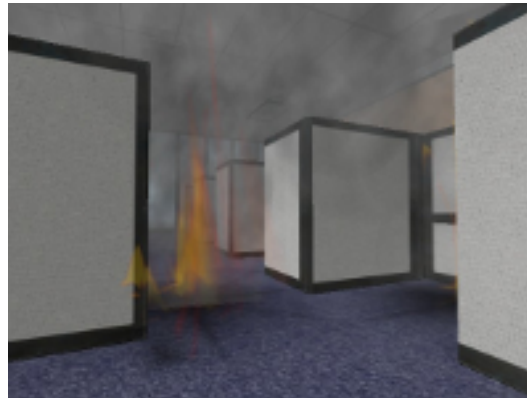
(a) Layout of first floor



(b) A player with the collector heads up display



(c) Layout of second floor



(d) Fire and Smoke

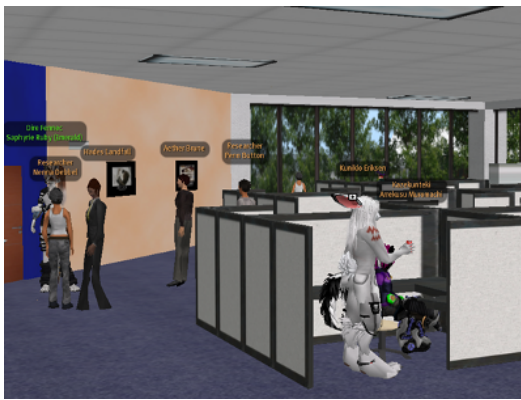
Figure 3: Screenshots from our game.

system answer?

Although Second Life avatars are controlled by real people, we do not expect that people will behave in the virtual environment as they would in the real world. In fact, the ability to live larger than life is part of the appeal of such environments. Avatars do not feel pain, have greater strength, and limitless stamina. Additionally, interactions inside virtual worlds (between people and with objects) are limited due to the restrictions of canned gestures and pre-defined object affordances. For example, visitors to our building open doors by left-clicking on them and can only pick up pre-tagged objects. Lastly, Second Life avatars, regardless of size, travel at the same speed. Additionally, the default speed is much faster than the average speeds of real people (3 m/s for walking and 5 m/s for running). Future experiments involving evacuation times, fatigue, and injury would need to address these factors.

4.1 Domain Independent Evaluation

Previous research has suggested a surprising number of ways in which human beings behave similarly in virtual environments as they do in real life – particularly regarding social interactions and norms (section 2). In our system, direct observation confirms the tendencies described by Yee et al. (2007) for avatars to stand near and face each other while in conversation (Figure 4(a)). After each experiment, most participants congregated outside of the building to talk, much like a real life fire drill (Figure 5(b)). Before an experiment, participants would crowd around the entrance, often respecting the personal space around other avatars (Figure 4(c)). After adding queue ropes to the entrance, avatars will stand in a line to enter, despite the fact that it is easy to jump over the rope (Figure 4(d)). Additionally, unlike many MMORPGs such as World of Warcraft, avatars cannot pass through each other.



(a) In conversation, standing in cliques



(b) Congregating after an experiment



(c) Gathering at the front entrance



(d) Queuing at the front entrance

Figure 4: Examples of social behaviors captured during our experiments.

4.1.1 Presence Evaluation

A section of our questionnaire was devoted to the level of presence, or the feeling of “being there in the scenario”, felt by participants. We wished to determine whether recruited lab participants experienced our Second Life scenarios equivalently to Second Life users. They did not.

Presence is a useful means for estimating the immersiveness of a scenario (Lombard & Ditton, 1997; Sanchez-Vives & Slater, 2005) and for our purposes provided a useful tool for identifying problems. Though imperfect, questionnaires were the most practical way to determine presence and obtain feedback from our online participants. We modeled the questions after early presence research (Slater et al., 1995; Witmer & Singer, 1998; Usoh et al., 2000): 4 questions asked about the experiment experience, e.g. how compelling was the office building, fire alarm, urgency to leave, and spread of fire; 6 questions determined “reported presence”, e.g. how well could the person focus on the scenario rather than the control mechanisms; 3 questions determined “behavioral presence”, e.g. how consistent outcomes were with expectations; and 3 measured “ease of locomotion”, e.g. ability to control the avatar including the level of involvement. Each question is rated on a 1-7 scale. 32 Second Life residents and 50 Lab participants filled out our questionnaire.

Based on our questionnaire results, Second Life participants responded significantly more positively than our lab participants to our virtual environment and controls. In particular, Second Life participants reported feeling significantly more urgency to leave the building, found the spread of fire significantly more compelling, found the interactions with the environment and controlling the character significantly more natural, and felt significantly more involved in the virtual environment.

Many lab participants reported frustration with Second Life character controls and conventions, despite many being familiar with 3D video games. However, for Second Life participants, many who reported being frustrated with our games also reported significant lag. These findings agree with previous work establishing the importance of frame rate and controls on presence (Jörg et al. (2012); Lombard & Ditton (1997)). From these data, we hypothesize that Second Life residents are much more comfortable using the Second Life viewer and thus not encumbered with the interface, making them able to focus on and enjoy the scenario more.

4.2 Domain Dependent Evaluation: Evacuation Behaviors

In this section, we describe three evacuation experiments whose hypotheses are based on important behavioral and structural factors related to egress in

real life emergencies. The first looks at whether people in virtual environments share the tendency of people in real life to prefer familiar exits during egress. The second looks at whether virtual debris and inaccessible exits can affect egress in a virtual environment. The third looks at whether virtual participants would delay evacuation in order to finish a task.

4.2.1 Experiment 1: Choice of exit.

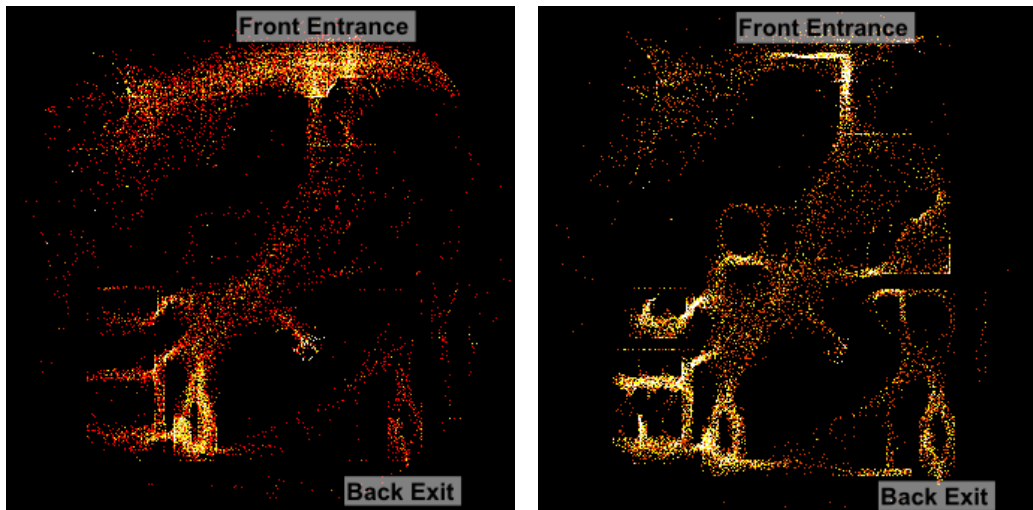
A common evacuation myth is that people will always choose the closest emergency exit. However, research has shown that people are influenced by an exit's proximity, its familiarity, and whether other people are using it (Sime, 1983, 1985). In practice, occupants will predominately use familiar exits, which are often the way in which they entered the building (Sime, 1984). Unfortunately, such a common sense, conservative strategy can sometimes have fatal consequences. In the Station Nightclub fire (Rhode Island, 2003), many fatalities and injuries occurred because people retraced their steps to the front entrance despite the existence of several alternative exit routes (Johnson, 2005). One route was via the kitchen and was only known to employees and regular customers. Another was next to the stage and could only be accessed via an inner door.

We performed an experiment to compare the choices of exit between participants who have never explored our building before versus those who had explored it at least twice. Figure 5 illustrates the differences between these two groups using heatmaps of the first floor of our building (lighter areas indicate higher foot traffic). Our building had both a front and back exit, where the front entrance was used to enter the building at the start of each experiment. Visitors unfamiliar with the building were less likely to explore the southeast corner, where our second exit was located.

We analyzed the choice of exit using a one-proportion z test against the null hypothesis that both exits are chosen equally often. Among unfamiliar participants, significantly more people, 28 out of 37 people (76%), chose the front entrance for egress ($z = 3.17$, $p < 0.01$). Among familiar participants, where players replayed at least twice, 20 out of 35 participants (57%) chose the front exit (n.s. $p > 0.05$).

4.2.2 Experiment 2: Maintenance of exits.

Lack of exits, due to either poor maintenance or structural deterioration from the fire, is a problem common to many fire disasters, such as the Triangle Shirtwaist Factory fire (NYC, March 25, 1911), Iroquois Theater Fire (Chicago, 1903) and the Cocoanut Grove fire (Boston, 1942). In these cases, exits were either blocked by debris or deliberately locked to prevent patrons from sneaking out



(a) Unfamiliar participants.

(b) Experienced participants.

Figure 5: A frequency map showing the usage of our first floor over multiple runs.

without paying. Blocked exits create bottlenecks and greatly increase the risk. For example, patrons of the Cocoanut Grove Nightclub were forced to exit via a revolving door that quickly jammed with bodies. Of course, current fire codes (IBC, 2008; OSHA, 2003) forbid these practices, specifying that exit routes be free of debris and that buildings must neither contain dead ends or long non-direct paths to the closest exit.

We replicated aspects of poor maintenance to test its effect in our virtual scenarios. We created a dead end at the upper stairwell leading to the roof; partially blocked the back exit with debris; and left only the slow revolving door unblocked at the front entrance, replicating the hazard of the Cocoanut Grove fire.

These modifications greatly frustrated participants whose feedback after these experiments included curses against the debris, blocked exits, and revolving door – “debris slowed me down”, “went to front door which was blocked”, “I wasn’t able to identify a clear path to the exit”, “the rotating door was too slow” and “the killer revolving door”.

We ran both the tile collection game and waiting game with and without traps and compared the numbers of participants who made it safely out of the building using two-proportion z tests (Figure 6). Significantly more people failed to escape when there were traps during the waiting game ($z=2.78$, $p < 0.01$) versus no significant difference for the collection game (n.s. $p > 0.05$). A possible explanation for this finding might be that people would follow others during egress. For the waiting game, players exited the same room at the same time whereas in the collection game players were spread throughout the building as they explored individually. However, when all results are considered together, significantly more people failed to exit safely when there were traps ($z=3.44$, p

< 0.001).

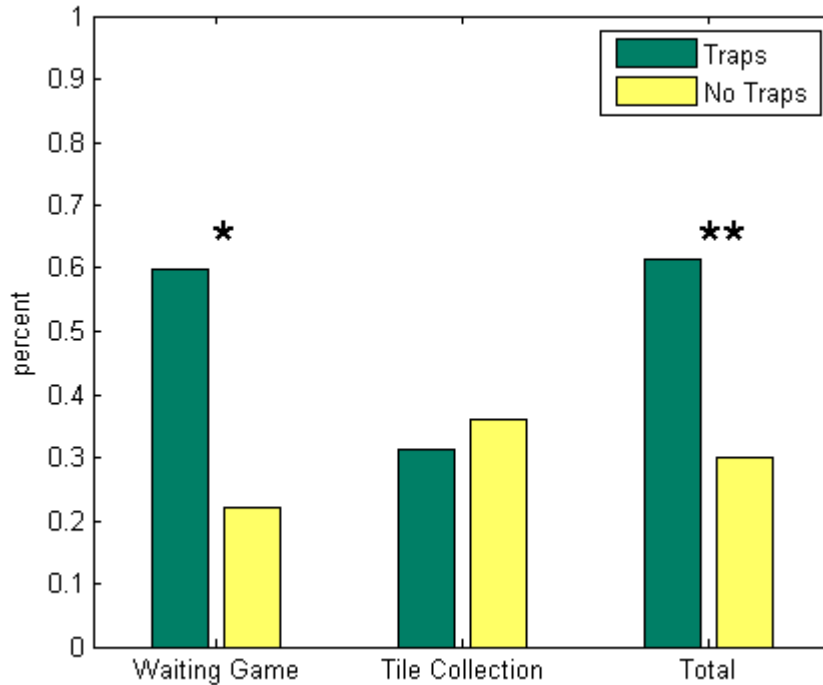


Figure 6: Trap results. The proportions of participants who did not safely evacuate in different scenarios with and without traps. A two-proportion z-test was used to test the significance between proportions of players who failed to safely exit the building. Asterisks indicate significant differences.

4.2.3 Experiment 3: Effect of game.

Another common evacuation myth is that people will start evacuating immediately after hearing an alarm. For example, in the 1979 Woolworth's fire in Manchester, eyewitnesses observed that patrons in the restaurant area were reluctant to leave without finishing their meals and paying first (Johnson, 2005). Nine out of the ten people who died in this incident were in the restaurant area. During the 1993 World Trade Center explosion, Fahy & Proulx (1997) tabulated the amounts of time between when people became aware of the fire and when

they began to evacuate and found that departure times ranged from as short as 1 second to more than four hours with a mean of 11.3 minutes and median value of 5 minutes. T.L. Graham (2000) defines this tendency of people to first finish a job before making an escape as *task fixation*.

For this experiment, we tested whether the presence of an activity affected the success rates of participants. Specifically, we compared the numbers of people who safely exited during our "waiting game" scenario (6/27) against our tile collection scenario (13/36) to test the hypothesis of whether significantly more people died during the tile collection scenario. A two-proportion z-test showed that there was no significant difference in the death rates between the waiting game and tile game scenarios (n.s. $p > 0.05$).

5 Conclusion and Discussion

Identifying and understanding the extent to which people behave realistically in online environments is of extreme importance for using such environments for experiments having real life significance.

With such an understanding, games can be designed to provide contexts for people to provide useful data. Data from these scenarios is trivially annotated and can be easily analyzed to answer experimental questions. For the domain of emergency evacuation, our participants, controlled by real people, choose when, where, and how to evacuate our virtual building, allowing us to run several behavioral experiments and obtain insights into potential weaknesses in the building design under different scenarios.

However, a desktop virtual game environment will likely never be completely realistic, so systems such as the one described in this paper are only proposed as a compliment to existing simulation techniques and real life drills.

Though our setup has potential to capture high level behaviors, it has problems at the low level. Particularly, obstacle avoidance is troublesome: people, especially those new to Second Life, are careless regarding collisions. Second Life residents are usually very careful to avoid other members, but may still bump into inanimate objects such as doors and desks. If collisions might adversely affect an experiment, one might add game elements to encourage careful avatar controls. Similarly, if timing information is needed, one would need to enforce realistic walking and running speeds.

Though our system has the potential to capture lots of data, experimenters currently need to be present during game sessions to mediate and give prizes. In the future, we would like to automate the system so that it may be left running for weeks at a time autonomously.

For the domain of evacuation, virtual evacuation drills could be used to cheaply test innovative safety features such as intelligent evacuation signs; to examine evacuation procedures before building construction; and to lessen the dependence on real-life drills which tend to be taken lightly or treated with apathy. The existence of a flexible and inexpensive test bed for evacuation could save many lives. VR systems are already widely used for training and our solution might motivate regular people to willingly participate in virtual evacuations of existing structures. Such games could educate people of available escape routes before an actual emergency occurs.

More broadly, a system such as ours could be of use in airline, school, and hospital safety and disaster training as well as for evaluating existing structures and could serve as a valuable, yet cost-effective, tool for architects, engineers, and government regulators.

6 Acknowledgments

The authors wish to thank Ben Sunshine-Hill, Catherine Stocker, and Sophie Joerg for their helpful comments; Kaitlin Pollock for her help running experiments and preparing demo materials; and all our experiment participants.

References

- Andersen, R., Berrou, J. L., & Gerodimos, A. (2005). On some limitations of grid-based (ca) pedestrian simulation models. In *First Intern. Conf. on Crowds Simulation*.
- Bainbridge, W. S. (2007). The scientific research potential of virtual worlds. *Science*, 317(5837), 472–476.
URL <http://www.jstor.org/stable/20037445>
- Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13(2), 103–124.
- Bukowski, R., & Squin, C. (1997). Interactive simulation of fire in virtual building environments. In *SIGGRAPH 97 Conference Proceedings*, (pp. 35–44).
- Castronova, E. (2006). On the research value of large games: Natural experiments in norrath and camelot. *CESifo Working Paper Series*, 1621.
- Cooper, S., Treuille, A., Barbero, J., Leaver-Fay, A., Tuite, K., Khatib, F., Snyder, A. C., Beenen, M., Salesin, D., Baker, D., & Popović, Z. (2010). The challenge of designing scientific discovery games. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games*, FDG '10, (pp. 40–47). New York, NY, USA: ACM.
- de Gortari, A. O. (2007). Second life survey: User profile for psychological engagement & gambling. The Virtual 2007 Conference: Interaction.
- eBird (2010). <http://ebird.org/content/ebird/>.

- Fahy, R., & Proulx, G. (1997). Human behavior in the world trade center evacuation. In Y. Hasemi (Ed.) *Fire Safety Science, Fifth (5th) International Symposium International Association for Fire Safety Science*, Proceedings. Fifth (5th) International Symposium International Association for Fire Safety Science.
- Friedman, D., Steed, A., & Slater, M. (2007). Spatial social behavior in second life. In *Intelligent Virtual Agents*.
- Galaxy Zoo (2010). <http://www.galaxyzoo.org/>.
- Grosshandler, W., Bryner, N., Madrzykowski, D., & Kuntz, K. (2005). *Report of the Technical Investigation of The Station Nightclub Fire*, vol. 1. US Government Printing Office, Washington, DC.
- Hendaoui, A., Limayem, M., & Thompson, C. W. (2008). 3d social virtual worlds: Research issues and challenges. *IEEE Internet Computing*, 12(1), 88–92.
- IBC (2008). International building code. International Code Council.
- Johnson, C. (2005). Lessons from the evacuation of the world trade center, sept 11th 2001 for the future development of computer simulations. *Cognition, Technology and Work*, 7, 214–240.
- Jörg, S., Normoyle, A., & Safonova, A. (2012). How responsiveness affects players' perception in digital games. In *Proceedings of the ACM Symposium on Applied Perception, SAP '12*, (pp. 33–38). New York, NY, USA: ACM.
- Kobes, H. I. d. V. B., M., & Post, J. (2010). Building safety and human behaviour in fire : a literature review. *Fire Safety Journal*, 45, 1–11.
- La, C.-A., & Michiardi, P. (2008). Characterizing user mobility in second life. In *WOSP '08: Proceedings of the first workshop on Online social networks*, (pp. 79–84). New York, NY, USA: ACM.
- Lee, K. H., Choi, M. G., Hong, Q., & Lee, J. (2007). Group behavior from video: A data-driven approach to crowd simulation. In *ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, (pp. 109–118). San Diego.
- libOpenMetaverse (2012). <http://www.openmetaverse.org/>.
- Linden Labs (2009). Linden labs case study: Children's memorial hospital chicago. URL <http://work.secondlife.com/en-US/successstories/case/chmc/>
- Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *J. Computer-Mediated Communication*, 3(2).
- OSHA (2003). Us occupational safety and health. Available on <http://www.osha.gov>.
- Pelechano, N., Allbeck, J., & Badler, N. (2008). *Virtual Crowds: Methods, Simulation, and Control (Synthesis Lectures on Computer Graphics and Animation)*. Morgan and Claypool Publishers.
- Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature reviews. Neuroscience*, 6(4), 332–339.
- Second Life Statistics (2012). <http://dwellonit.taterunino.net/sl-statistical-charts/>.
- Sime, J. (1983). Affiliative behaviour during escape to building exits. *Journal of Environmental Psychology*, 3, 21–41.

- Sime, J. (1984). Behavior in fire: panic or affiliation? *Department of Psychology, University of Surrey, UK.*
- Sime, J. (1985). Movement toward the familiar: person and place affiliation in a fire entrapment setting. *Environment and Behaviour, 17*, 697–724.
- Slater, M., Usoh, M., & Steed, A. (1995). Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Trans. Comput.-Hum. Interact., 2*(3), 201–219.
- Stott, D. (2007). Learning the second way. *British Medical Journal, 335* (7630), 1122–1123.
- The River City Project (2007). The river city project.
URL <http://muve.gse.harvard.edu/rivercityproject/>
- Thompson, P. A., & Marchant, E. W. (1995). A computer model for the evacuation of large building populations. *Fire Safety Journal, 24*(2), 131–148.
- T.L. Graham, D. R. (2000). Qualitative overview of some important factors affecting the egress of people in hotel fires. *Hospitality Management, 19*, 79–87.
- University of Arkansas (2010). Modeling healthcare logistics in a virtual world.
- Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using presence questionnaires in reality. *Presence: Teleoper. Virtual Environ., 9*(5), 497–503.
- von Ahn, L. (2006). Games with a purpose. *Computer, 39*, 92–94.
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*, 225–240.
- Yee, N., & Bailenson, J. (2007). The proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research, 33*, 271–290.
- Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & Merget, D. (2007). The unbearable likeness of being digital; the persistence of nonverbal social norms in online virtual environments. *Cyberpsychology and Behavior, 10*, 2007.