THE EFFECT OF CHARACTERS' LOCOMOTON ON AUDIENCE PERCEPTION OF CROWD ANIMATION

by

Wenyu Zhang

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Computer Graphics Technology

West Lafayette, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Prof. Nicoletta Adamo, Chair
Department of Computer Graphics Technology

Dr. Tim McGraw
Department of Computer Graphics Technology

Dr. Robert Proctor
Department of Psychology

Approved by:
Prof. Nicoletta Adamo

ACKNOWLEDGMENTS

I would like to thank Prof. Nicoletta for her patient and consistent support on this thesis from beginning to the end. I would also like to thank Dr. McGraw and Dr. Proctor for their professional suggestions regarding my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

A common practice in crowd animation is the use of human templates. A human template is a 3D character defined by its mesh, skeletal structure, materials, and textures. A crowd simulation is created by repeatedly instantiating a small set of human templates. For each instance, one texture is randomly chosen from the template's available texture set, and color and shape variety techniques are applied so that multiple instances of the same template appear different (Thalmann & Musse, 2013). When dealing with very large crowds, it is inevitable to end up with instances that are exactly identical to other instances, as the number of different textures and shape modifications is limited. This poses a problem for crowd animation, as the viewers' perception of identical characters could significantly decrease the believability of the crowd simulation. A variety of factors could affect viewers' perception of identical characters, including crowd size, distance of the characters from the camera, background, movement, lighting conditions, etc. The study reported in this paper examined the extent to which the type of locomotion of the crowd characters affects the viewer's ability to perceive identical instances within a medium size crowd (20 characters). The experiment involved 51 participants and compared the time the participants took to recognize two identical characters in three different locomotion scenarios (i.e. standing, walking, and running). Findings show that the type of locomotion did not have a statistically significant effect on the time subjects took to identify identical characters within the crowd. Hence, results suggest that audience perception of identical characters in a medium size crowd is not affected by the type of movement of the characters.

CHAPTER 1. INTRODUCTION

Traditionally, off-line crowd rendering has been commonly used in the animation and visual effects industries. However, nowadays real-time crowd rendering approaches are being rapidly developed. An essential criterion for real-time crowd simulation is that the simulation has to run at a rate higher than 30 frames per second (Thalmann et al. 2009). Various crowd simulation techniques have been developed and are widely applied in the visual effects, animation, and video game industries. However, factors that affect the quality of audience perception, such as the degree of realism and level of detail, still need to be considered for these techniques to have validity.

1.1 Significance

This research examined the effects of characters' locomotion on the viewer's perception of identical characters in medium-sized crowd simulations. Findings from this research can help those involved in animation production save time and resources by decreasing the number of individual characters that are necessary in certain crowd simulation scenarios.

1.2 Research Questions

To what extent does characters' locomotion affect viewer's perception of identical instances in a medium size crowd animation?

1.3 Scope

This research focuses on crowds which require a heterogeneous character appearance and motion. The goal is to determine whether it is possible to use a lower number of entity characters depending on the particular type of locomotion seen in the scenario. It seeks to determine the minimum quantity of entities and instances required to portray a realistic crowd. This can help reduce unnecessary production costs while still maintaining an acceptable level of production quality. Therefore, this research has implications for real-time simulations (e.g. computer games) as well as off-line rendering simulations (e.g. crowds in film). In some cases, heterogeneous characters might actually have a distracting effect on the viewers' perception of the virtual environment. For example, in a scene showing a large group of people running out of a burning building, a more uniform crowd simulation would make the crowd behavior and motion path more conspicuous (Ulicny et al., 2002). The ultimate goal is increasing the efficacy of animation production by maximizing the visual complexity of moving pictures with a minimal expenditure of technical resources.

1.4 Assumptions

Assumptions for this study include:

1. Participants have paid their full attention when viewing the testing video clips.
2. The number of participants used in the study is sufficient for statistical data analysis, according to previous related studies.

1.5 Limitations

Limitations for this study include:

1. The testing video clips do not have a fully realistic quality to them and do not rely on a physical-based rendering effect.
2. The study focuses on medium size crowds only.
3. The study considers only 3 types of locomotion as standing, walking, and running.

1.6 Delimitations

Delimitations for this study include:

1. This research only displays scenes in which the crowd characters are heterogeneous, such as city street pedestrians, stadium audience, etc., instead of more homogeneous groups (e.g. soldiers in uniforms).
2. This research does not consider the potential effects of character appearance on perception and identification.

3. Every video clip is shown from a bird's eye perspective, to ensure viewers can easily perceive the entire scene. It therefore does not consider the potential effects of visual angle on perception.
4. Facial expression does not vary across the different characters in each scenario.

1.7 Definitions

Crowd Animation: A method used to simulate the motion of real-world crowds. The crowds can either be homogeneous, like a group of uniformed soldiers, or heterogeneous, like a group of real-world street pedestrians.

Locomotion: The change in the position of a model during a certain period of time.

Entity: The original model in a crowd simulation program, which can be modified and duplicated to generate additional models.

Instance: Duplicated copies of the same character rendered in a scene. Each instance may have different appearances and poses.

1.8 Summary

This chapter provided an overview of the research topic. It discussed the research questions that the study attempts to answer and the scope of the study, along with its significance and limitations. The next chapter consists of a literature review of all the relevant topics covered in this study.

CHAPTER 2. LITERATURE REVIEW

Previous research has shown that there are numerous factors that could potentially affect audience perception of crowd animation. In this chapter, a literature review consisting of three sections is presented. It includes a description of the production process of 3D character animation, a discussion of some of the challenges and problems that may arise when creating crowd animations, and relevant previous perceptual studies in this field.

2.1 Character Animation

The production process of character animation can be broken down into five basic parts: modeling, texturing, rigging, animating and rendering. Modeling, texturing, and rendering all determine the physical appearance of characters, while rigging and animating control the characters' movements and facial expressions.

Modeling is a process whereby the creator defines the shape of the characters without giving consideration to their texture. It allows the creator to display several basic properties of a character, including height, gender, age, body shape, hair style, and muscle level. Zell et al. (2019) noted that "shape is the main descriptor for realism, and material increases realism only in case of realistic shapes." Lighting and shading are also important components in the creation of realistic as well as stylized character animations. In computer animation, some recent research (Wusessubg et al. 2016) suggests that the lighting method does not have a significant impact on the perception of characters' emotion. Also, it turns out that dark shadows are often ignored by viewers in certain animation scenarios, in contrast to their more salient role in film (Kardos, 1934).

Rigging and Animating are two key factors in crowd animation. Traditionally speaking, the character models in crowd animation are polygonal meshes rigged by bones. When the joints are rotated, the vertices cluster attached to the joints becomes deformed along a predetermined trajectory, which is how character animation is generated (Dong et al., 2019). In real world production, character motion can be created either by animators' key-framing and adding frame interpolation, by using physics-based animation generated by computer simulation tools, by capturing real-time motion data from devices on actors (motion capture), or via any combination of the three aforementioned techniques (Zell et al. 2015).

However, certain comprehensive methods need to be adopted in creating crowd animation because moving crowds involve complicated mechanics which require algorithmics (Lemercier et al. 2012). Such methods, which go beyond an individual character's locomotion, have been studied by previous researchers. For example, walking is a common mode of locomotion that can easily be produced for an individual character. However, in the case of a group of walking characters (e.g. pedestrians on the street), factors such as collision avoidance must also be considered (Reynolds, 1987). Additionally, many algorithms related to the motion trajectories of crowds have been developed in recent years. For instance, Yu & Terzopoulos (2007) developed a novel framework for pedestrian characters, including behavioral interaction in urban settings. Guy et al. (2011) presented a technique called Personality Trait Theory to create heterogeneous crowd motion. Sun et al. (2013) simulated realistic crowd trajectory in an urban scenario surrounded by traffic, vehicles, intersection, etc.

In a crowd simulation, characters' locomotion and behavior inevitably rely on the nature and quality of the algorithms operating behind the scenes.

2.2 Crowd Rendering

3D renderers are tools used to output final image sequences in animation production. They generate various results depending on the shading models and numerous physical parameters inputted by the animators. Traditionally, crowd animation was rendered through animation programs. However, with the development of 3D games, many programs tend to have the capability to simulate crowds in real time. In the earlier studies on crowd rendering, limited computational speed was the main problem in the creation of 3D scenes with populated characters. Many acceleration techniques for the rendering of large environments were subsequently invented.

A technique called "instance" has been frequently used in crowd simulation. In a shading API (Application Programming Interface) such as OpenGL or DirectX, a geometry shader can be used to deform the vertices and the triangle mesh of a crowd with only one call in GPU (Graphics Processing Unit) (Carucci, 2005). Ashraf and Zhou (2007) applied a hardware-accelerated method through programmable shaders to animated crowds. In Peng et al. (2011), developments have been made to utilize GPU in their parallel architecture to improve the performance of graphics computation. They invented a mesh simplification algorithm which can render a real-time crowd system on the GPU. Klein et al. (2014) created an innovative method which allows instances of

3D characters with controllable parameters to be rendered on the web. In the work of Dong & Peng (2019), a novel crowd rendering system that simultaneously runs real-time on GPU and decreases the computation load on the graphics card was developed. The scene includes 30,000 instances in real-time motion.

Shading is also a significant component in rendering real-time images. Maciel & Shirley (1995) implemented the LOD (Level-of-Detail) technique to create impostors to reduce the complexity of rendering. This method later evolved to an IBR (Image-based Rendering) technique which was adopted by many researchers. In a study by Tecchia et al. (2000), the Image-based Rendering (IBR) method was adopted, and the characters were pre-computed and animated. A multi-pass algorithm was first used to retouch different parts on the character, followed by the addition of efficient shading and shadow. In a study by Tecchia & Loscos (2002), they used an approach whereby each character was transformed into an image-based impostor which possesses an adaptive resolution depending on one's viewing angle. Ciechomski et al. (2005) presented a customized hardware rendering pipeline which created texture variety from a single texture in HSB color space. Millan & Rudomin (2006) combined the imposter and instancing techniques and created a program which is more efficient in rendering large crowds. Ciechomski (2006) used another IBR approach with human characters, which involves taking photos of humans. For example, in the movie The Matrix, hundreds of cameras aligned in a spherical array take photos at a very high shutter speed. This allows the animator to generate a reconstructed 3D model from the spherical image data. After that, all the snapshots can be converted into texture map images, which is known as the Billboarding technique.

2.3 Perception of Virtual Characters

Ciechomski et al. (2005) has stated that "For a human crowd, variation can come from the following aspects: gender, age, morphology, head, kind of clothes, color of clothes and behaviors." In other words, the perception of human crowds in animation mainly depends on two aspects – appearance and behavior.

All the virtual CG characters can be classified into two categories: photo-realistic and stylized. In a study by Zell et al. (2015), it was found that factors such the shape of a character's body and its material (especially the albedo texture) can significantly affect audience perception. These two factors have a strong influence on how realistic the characters are perceived to be.

Another factor that affects the believability of perception is the facial proportion of characters. Green et al. (2008) concluded that facial height, jaw width, and eye separation are all considered to be important factors which can increase the appeal of animated characters.

Besides their exterior appearance, the behavior or motion of characters likewise plays an important role in creating realistic perceptions. Based on a study by Johansson (1973), when the characters are in motion (e.g. walking or running) as opposed to staying still, viewers can appreciate that the virtual characters resemble real world human beings, instead of perceiving them as a group of static dot-shape objects. Research by McDonnell et al. (2008) compared the reaction times in spotting appearance-based duplicated characters versus motion-based duplicated characters. They concluded that characters cloned by appearance are more conspicuous than characters cloned by motion. Also, they discovered that the position layout of characters affected the viewers' perception – horizontal layout makes it easier for the audience to spot cloned characters compared to a vertical or diagonal layout. One limitation of their experiment is that all the testing characters were positioned facing forward, which is not considered typical in crowd animations. Prazak & O'Sullivan (2011) studied the locomotion variety in crowd animation perception. They adopted motion capture techniques to capture 83 actors' real-world motion data (including both males and females) and created a virtual scene to perform the experiment. They claimed that at least three different locomotion types are needed to be displayed for each gender to achieve a realistic level of behavioral variety in a pedestrian scene. However, their character set was relatively small, with only 24 characters being shown at a time in each scene. Moreover, they did not examine the effects of the various types of motion in the experiment.

Eye tracking has become quite popular in perception studies in recent years. Using an eye-tracking device, McDonnell et al. (2009) found that head and upper body are the first part viewers tend to notice, regardless of the character's position, motion, gender, size, etc. They also found that creating more kinds of head accessories and variable top textures is more effective at increasing variety than alternating the facial geometry of characters.

When it comes to facial close-ups, the eyes tend to catch viewers' attention more than other body parts. A recent study (Schwind & Jäger, 2016) confirmed that viewers primarily maintain their glance at the virtual characters' eyes and mouth. On average, it was found that participants spend around 35% of the time looking at the eyes, while spending no more than 10% of the time focusing on other parts of the body.

Figure 2.1 is a screenshot of an animated commercial short for WestField Shopping Mall. Some of the CG characters are walking randomly in the mall; while some are standing still. They all have different appearance and slightly difference behavior which increases the perception fidelity of crowd animation.

Figure 2.1 A Screenshot of WestField Visualization Commercial

2.4 Summary

This literature review examined computer animation, crowd simulation, and perception of virtual characters. It addressed state of the art techniques, as well as some of the challenges and problems related to crowd animation and perception. It is worth to conduct research in audience perception of crowd animation which can greatly optimize production efficiency depending on the research the result.

CHAPTER 3. METHODOLOGY

3.1 Introduction

This section discusses the methodology of this research study. The goal of this study was to determine whether different types of locomotion would affect viewers' perception of the crowd. The participants watched randomized video clips representing three scenarios and were then instructed to complete a related online survey. The study adopted a quantitative research that compared the length of time that participants spent on each scenario. A customized Bayesian Linear Mixed Model was employed to analyze the collected data.

The independent variable in this research was the type of locomotion (standing, walking, running) of 3D characters as well as the gender of participants The dependent variable was the length of time that the subjects spent to identify two identical characters in the crowd.

3.2 Hypotheses

H_{01}: Participants will spend the same amount of time to identify identical characters in all the three locomotion scenarios.

H_{a1}: Participants will spend different amount of time to identify identical characters in each of the three scenarios. Specifically, participants will spend more time to identify identical characters in the Running Scenario than in the Walking and Standing Scenario, respectively. The rationale for the hypothesis is based on the results of a prior study by Prazak & O'Sullivan (2011) that showed that participants had lower accuracy in identifying identical characters when watching a crowd animation with fast movement, and medium accuracy when watching a crowd animation with slow or no movement.

H_{02}: Participants will spend the same amount of time to identify identical characters regardless of participants' genders.

H_{a2}: The time participants will spend to identify identical characters will vary depending on participants' genders.

3.3 Subjects

A total of 83 participants took part in this study. Thirty-three participants were students from Computer Graphics Technology department at Purdue University. The additional 50 participants were selected via a survey posted on Amazon Turk. The participants were recruited without regard to gender and resulting in 46 males and 37 females in the pool. Participants' age ranged from 18 to 64 years old. Participants' familiarity with computer animation ranged from zero experience to very familiar with computer animation. All the participants could see the computer screen clearly, with or without corrective lenses.

3.4 Stimuli

The stimuli used in this study consisted of three online videos demonstrating different types of character locomotion within crowd animation, along with an online survey. The crowd animation video clips were created using Maya 2016 with Golaem plugin and were rendered using Mental Ray renderer. The rendered picture contains both highlight and shadow in order to simulate realistic lighting. However, the materials on the characters do not include any other channels besides diffuse textures. All the characters' exterior, such as garment texture, is from the preset package of Golaem plugin. The characters' locomotion (e.g. walking, running) was also created using Golaem presets. We customized the characters' moving trajectories to allow them to have specific paths without moving out of the frame. Also, to assure all the other parameters stayed uniform, the camera angle, lighting, shadow, contrast, were set up completely identical in each video clip. The camera was positioned at one side of the scene with a tilting angle of 30 degree towards the ground. The lens has a view angle of 35 degree to capture the full scene.

In each scene, there are 18 characters with heterogeneous appearance and only two characters with homogeneous appearance, which includes skin color, hair color, color of shirt, pants and shoes. In the standing scenario, characters stand still on the ground surface and exhibit casual turning-in-place movements. In the walking scenario, characters walk in random trajectories on the ground surface. In the running scenario, characters run around in random trajectories.

After the animation images were exported from the animation package. Image sequences for each scenario were processed in video editing programs and output as three 98-second video clips. Each video clip had a 10-second opener with instruction reminding viewers to be prepared for the experiment. Each video clip looped 10 times itself. All the experiment videos were in sRGB color space without any post-processing or visual effects.

Figure 3.1 Locomotion Scenarios Illustrating Standing, Walking, Running, Respectively

3.5 Evaluation Instrument

This experiment required participants to view a series of animation video clips. Therefore, a laptop or personal computer with proper display and fast Internet access was required. Mobile devices were not allowed in this study given that the screen resolution on such devices and the various nuances of person-device interaction might affect the perception results.

Data collection was performed via an online survey created using the Qualtrics survey platform. The videos were embedded into the survey platform and all user interaction controls were disabled. The survey included the IRB consent form, detailed experimental instructions, a demo video, three formal testing videos, a demographics questionnaire, and optional feedback.

Since this study required quick responses from the participants, detailed instructions along with a video tutorial were displayed to participants at the beginning of the experiment to ensure they thoroughly understood the experimental procedure. In addition, a demo video was presented to allow participants to familiarize themselves with the procedure and promote reliable results. Participants were expected to adjust page zoom to a suitable resolution in the browser to allow them to watch the entire frame.

Formal video clips for testing began to play automatically as soon as the participant jumped to the page. Each video clip had a text reminder displaying “Please move the cursor on the blue button. (Do not click until you have found two identical characters).” Along with each formal experiment video, there were required questions on the next page letting participants select the identical character they found, as a method to detect the accuracy. Each question had only one correct answer out of three choices. The answer did not contain any text but only a pair of screenshots of the characters (full body front and back) appeared in the video. Thus, viewers might have had a more intuitive impression to select the character they believed they have found. Participants were forced to select an answer before they could jump to the next page.

In order to decrease potential confounds stemming from the learning effect (whereby participants' performance improves over time as they are exposed to the same stimulus), the order of the three video scenarios was randomized. We randomized the video groups into three different combinations to make sure each scenario would not always appear at the first. This greatly reduced the audience's learning effect. The order combinations were Standing-Walking-Running, Walking- Running-Standing and Running-Standing-Walking.

3.6 Procedure

For each scenario, the video clip started to play automatically and looped for 15 times. All the interaction controls were disabled on the videos. Thus, participants were not able to pause, adjust speed, download, or loop the video by themselves. Participants were asked to click on the blue button showing "CLICK ME" at the bottom right corner of each scenario page as soon as they spotted the two identical characters. The system recorded the exact response time for each participant. Figure 3.2 shows a screenshot of the Walking Scenario stimuli.

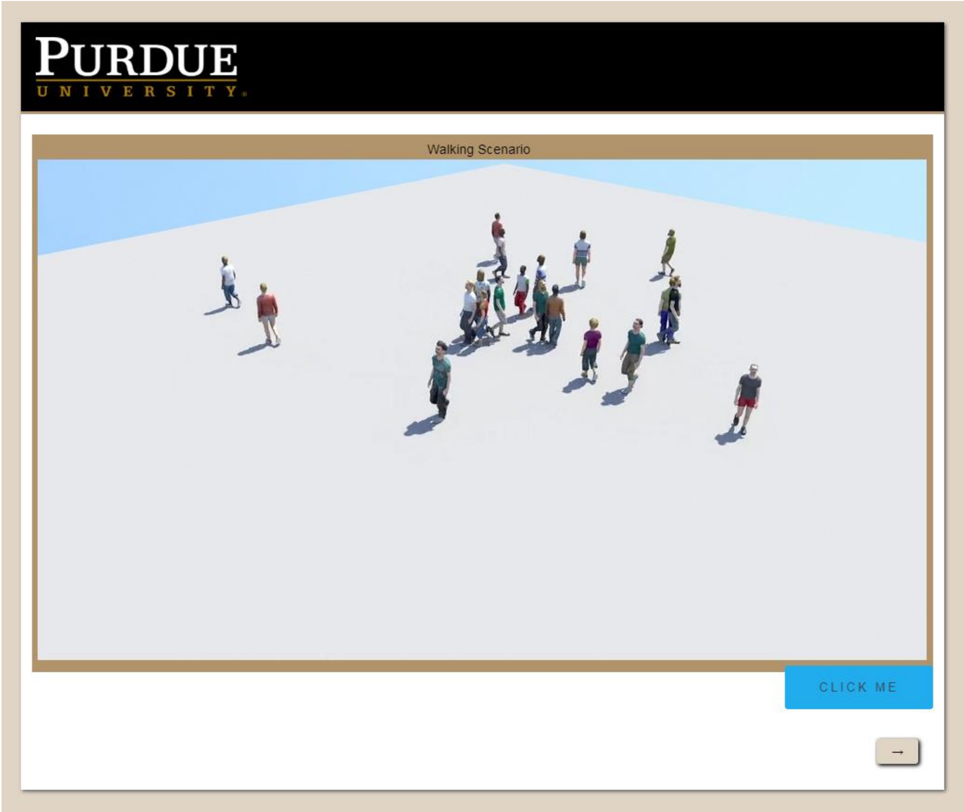


Figure 3.2 A Screenshot of Formal Testing Video Clip (Walking Scenario)

Next, participants were asked to select which of the three types of characters were identical in the video clip. After a selection was made, the page would progress to the next video. After viewing all the video clips and answering the pertaining questions, the participants were asked to fill out a brief demographic questionnaire. It collected participants gender, age and their familiarity of computer animation. Finally, they were given the option to share any feedback or comments they may have had regarding their experience before concluding the study.

3.7 Summary

This section is an overview of the framework and methodology of this research. It explained the variables, hypothesis, subjects pool, stimuli, and experiment instrument and procedure. Hypotheses can be analyzed by the procedure mentioned below.

CHAPTER 4. DATA ANALYSIS

After the experiment was conducted, participant response times (i.e. the amount of time each participant spent to identify identical characters in each video) were collected. A Bayesian Linear Mixed Model was used to determine whether the response times varied significantly across the three locomotion scenarios (standing, walking, and running).

4.1 Data Pre-Processing

The dependent variable in this study was the response time, or the length of time that participants spent on each scenario before clicking the mouse (indicating that they identified two identical characters). First, an accuracy check was performed to clean up the data result. Since there was only one correct answer after each question, participants who selected incorrect answers were subsequently removed from data set. Standing Scenario had an accuracy of 75%; Walking Scenario had lowest accuracy of 62%; Running Scenario had highest accuracy of 87%. Figure 4.1 is a bar graph to visualize the accuracy result. Since as stated in Chapter 3.6, the actual video would not play until the 10^{th} second and would terminate at the 98^{th} second. Participants' who had spent less than 10 seconds and greater than 98 seconds in watching each video clip were considered outliers and were likewise removed from data set. After the clean-up, there were 51 available subjects in the data set, 28 males and 23 females. The reaction times across the three video types were then analyzed using Bayesian Linear Mixed Model.

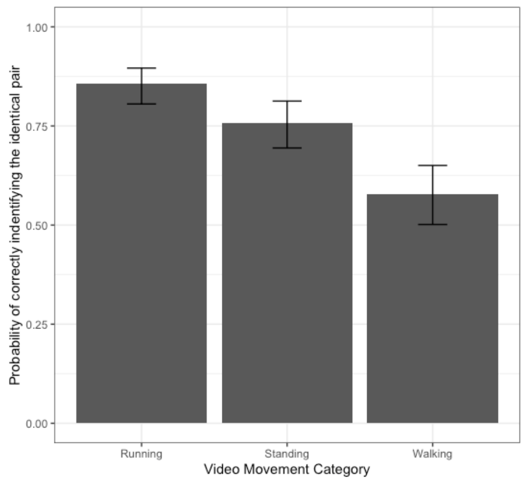


Figure 4.1 Response Accuracy of Each Locomotion Scenario

4.2 Data Model

Each participant in our study was exposed to all three video categories. Only the subjects who identified every pair correctly and responded within the acceptable range of response times (as explained above) were included in the response time analysis.

Combining all the factors which might affect the result of this study, we attempted to fit the model as below:

$$\boldsymbol{Time_{ijk} = \mu + Video_i + Subject_k + Period_j + Sequence + Gender + \varepsilon_{ijk}}$$

where:

1. $Time_{ijk}$ is the actual response time for subject k watching video i in time period j.
2. μ is the overall mean expected response time.
3. $Video_i$ is the effect of the i^{th} video category (Running, Walking, Standing) on the expected response time.
4. $Subject_k \sim N(0, \sigma^2_{subj})$ is the random effect of subject k on expected response time.
5. $Period_j$ is the effect of the j^{th} time period on the expected response time.
6. *Sequence* is the effect of video display order on the expected response time.
7. *Gender* is effect of different gender on the expected response time.
8. $\varepsilon_{ijk} \sim N(0, \sigma^2)$ is the error between expected and actual response time.

The model includes fixed-effects stemming from our independent variable video category (standing, walking, running) and factors corresponding to video order, period, gender as explained above. In addition, we included random subject effect, in order to control for heterogeneity of each subject.

In this study, there are basically three categories under $Video_i$, 51 different individuals under $Subject_k$, three categories under $Period_j$, three different orders under *Sequence* and two categories under *Gender*. $Video_i$ includes VideoS, VideoW, VideoR; $Period_j$ includes Period1, Period2, Period3; *Sequence* includes Sequence1, Sequence2, Sequence3; *Gender* includes GenderMale and GenderFemale.

4.3 Data Analysis

As the graph suggests, participant response accuracy was highest in the running category, followed by standing and walking scenarios, respectively.

Using Bayesian Linear Mixed Model, such result was yielded from the data model with a 95% credible interval ranging from 2.5% to 97.5%.

Table 4.1 Multiple-factor Credible Interval

	.lower	.upper	.width	.point	.interval
VideoS	-18.0	13.9	0.95	median	qi
VideoW	-2.75	28.1	0.95	median	qi
Period2	-22.3	9.85	0.95	median	qi
Period3	-24.3	7.85	0.95	median	qi
Sequence2	-23.6	12.7	0.95	median	qi
Sequence3	-22.5	15.3	0.95	median	qi
GenderMale	-31.1	-0.44	0.95	median	qi

In this case, VideoR, Period1, Sequence1 and GenderFemale are used as baselines. The most plausible values with higher probability of representing the true estimate indicate that the mean of the intervention group VideoS and VideoW should be either lower or higher compared to the comparison group VideoR. As 0 lies within the interval, we do not have statistically significant evidence to claim that there is difference between VideoS, VideoW, and VideoR.

Credible interval for Period2 and Period3 contains 0. This indicates we do not have statistically significant evidence to claim that there is difference between Period1, Period2 and

Period3. Accordingly, credible interval for Sequence2 and Sequence3 contains 0. This indicates we do not have statistically significant evidence to claim that there is difference between Sequence1, Sequence2, and Sequence3, either.

However, GenderMale has both negative lower bound and upper bound which does not contain 0. Thus, Gender turned out to be significant factor in this data model. Two interaction plots regarding Video Types and Gender were generated after this interesting finding. In the first plot, it shows that male participants always had shorter response time than female participants across all the three video types, especially in Standing and Running Scenario. In the second plot, female participants tended to have lower variance while male participants had a higher variance. However, both genders performed worst in Walking Scenario.

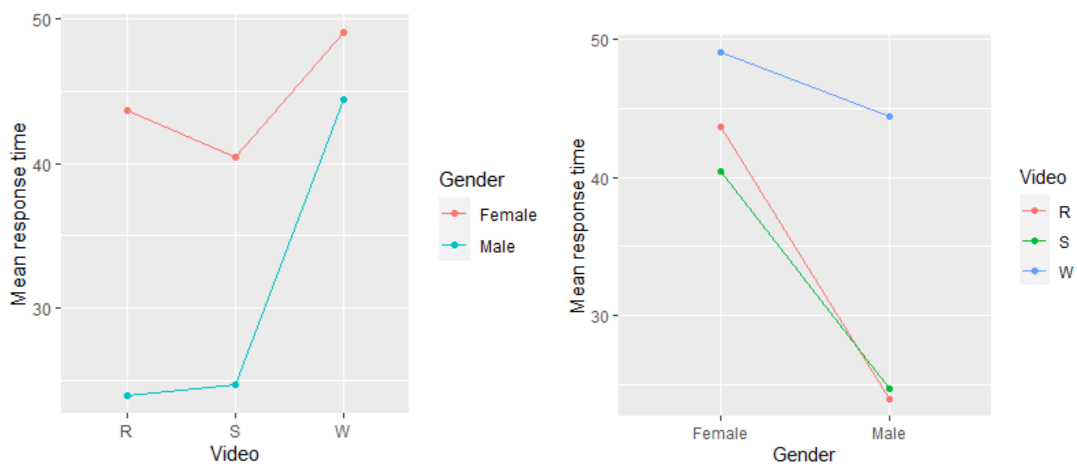


Figure 4.2 Visualization of Response Time per Gender

4.4 Results

Results from the data analysis showed that the time participants took to identify two identical characters in the crowd were not significantly affected by different locomotion categories. Hence, we failed to reject the null hypothesis. There was no significant difference in reaction time across the three different crowd animation scenarios.

However, gender had a significant effect on participants' perception of identical characters within the crowd. Male viewers tended to be able to spot identical characters quicker than female viewers. In the three types of scenario, male and female viewers had smaller difference in Walking Scenario while they had major difference in Standing and Running Scenario.

4.5 Summary

Between-subject repeated measurements method was applied to data analysis. The result is not significant compared with all our null hypotheses. The experimental results suggest that response times did not vary significantly across the three locomotion categories. In conclusion, different types of locomotion did not affect audience perception in crowd animation. Different gender could affect the perception in crowd animation. Female participants tended to spend more time to identify identical characters in the video.

CHAPTER 5. CONCLUSION AND DISCUSSION

5.1 Conclusion

In conclusion, the results of this experiment indicated that the type of locomotion used in an animation (i.e. standing, walking, or running) has no significant effect on the audience's perception of the scene. In particular, the type of locomotion exhibited by characters in a crowd scenario did not significantly impact the time it takes to spot identical characters, since the time participants spent on the video did not differ significantly across all three locomotion scenarios. Gender did have impact on perception of crowd animation.

5.2 Discussion

This study has several limitations and potential confounds. First, a power analysis was not performed prior to the actual experiment. The researcher used as much of the subject's background characteristics and demographics in the design and analysis of the study to obtain as much power as possible under the circumstances. The pool of subjects included different ages ranging from 18 to 54. The median age lied in the range of 25 to 34. It also included people with zero experience of computer animation as well as people who have some experience in computer animation. Nineteen participants have no experience with animation while 32 participants stated that they were familiar with computer animation. The pool was representative of the target population, however, in the future studies with larger pools of participants might yield more generalizable results.

Second, the position of the characters in the crowd at any given moment of time might have had an effect on participants' perception. For example, identifying two identical characters that happened to be running close to each other may have been easier than if the characters were far apart. Thus, distance between two identical characters could have been a significant factor that affected the perception in such scenarios.

Third, all the shots were static without any camera movement, which is not always true in real world films. In a case with camera movement (e.g. a top-down view with a dolly shot), the audience might not be able to focus on a specific area. Hence. the probability that viewers spot identical characters may be lower.

Fourth, the videos used in this experiment are quite rudimentary and are considerably lower in quality compared to real-world commercial film productions. Visual fidelity is relatively low due to quality of character texture assets and lack of surrounding environment. The videos also lack elements used in compositing such as smoke, fog, haze, dust, and flares - all of which are inevitably present in the real world. Further, all the testing scenarios do not include any 3D objects which might become blockers (e.g. buildings, poles, signs), but only an open space on a flat ground. As a result, the audience might be able to perceive identical characters more quickly and easily in our study as compared to real-world animated films. Therefore, the results of our experiment may be limited in their generalizability.

Fifth, a phenomenon known as the learning effect might have also played a role in this experiment. Participants might be able to achieve better result with more and more familiarity with the testing procedure in a short period of time. The researcher used randomization tricks to mitigate this effect. A demo video was given at the beginning of the study, so participants could become familiar with spotting identical characters before conducting the actual experiment.

Finally, viewers' perception of the characters might have been affected by the intrinsic design features of the characters, in addition to our variable of interest (locomotion). For example, it is known that human eyes are more sensitive to certain colors of the visible spectrum (e.g. solid red and yellow) than to others, and so participants' response times might have been affected by the different colors of the characters.

5.3 Future work

In future experiments, characters' motion paths could be varied to exhibit different trajectories. For example, all the characters could be running towards the same target, or all of them could be running around in a loop. It would be interesting to see whether the moving path of the crowd as a whole would affect viewers' perception of identical characters.

In addition, certain camera angles, such as the absolute top view, could make it very difficult to spot identical characters. The difficulty of perception would also depend on the distance between the render camera and the characters. It will be also worth conducting research on the perception under moving cameras. This will match numerous shots in real world films since moving shots are more common than static shots. Future experiments can explore the effects of all these variables.

Future experiments could also diversify characters' appearance, so that differences in skin color, gender, body shape, and other variables can be included and their effects on audience perception could be analyzed. Characters could also be made to wear glasses, hats, and other accessories to investigate their effects on viewers' perceptions.

Careful scrutiny of animation in experiments is different from real-world animated films. In real-world cases, audience does not purposely focus on spotting identical characters among either foreground or background characters while watching films. Therefore, a scenario in which the audience is asked to spot identical characters by watching a looping video does not typically occur in the real world. The looping video in our experiment is thus somewhat artificial and represents a worst-case scenario.

5.4 Summary

As discussed earlier in this section, there are still many unexplored variables and scenarios worth examining in the future. This research focused on a practical perception problem in the 3D animation and video game industries. Different types of locomotion in crowd animation did not significantly affect the audience's perception of identical characters in our customized scenarios. The results of this study are important as they can help animators determine the minimum number of different character entities necessary to create believable crowds with various types of locomotion. Hence the findings of this study could help reduce the budget of production of real-world 3D animations and video games.

REFERENCES

Ashraf, G., & Zhou, J. (2007). Hardware Accelerated Skin Deformation for Animated Crowds. *International Conference on Multimedia Modeling 2007: Advances in Multimedia Modeling, pp. 226-237.*

Carucci, F. (2005). Inside Geometry Instancing. In M. Pharr (Ed.), GPU Gems 2 (pp. 47-67). Addison-Wesley.

Ciechomski, P. H., Schertenleib, S., Maïm, J., Maupu, D., & Thalmann, D. (2005). Real-time Shader Rendering for Crowds in Virtual Heritage. The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST.

Dong, Y., & Peng, C. (2019). Real-Time Large Crowd Rendering with Efficient Character and Instance Management on GPU. International Journal of Computer Games Technology, Volume 2019, Article ID 1792304. https://doi.org/10.1155/2019/1792304

Guy, S. J., Kim, S., Lin, M. C., & Manocha, D. (2011). Simulating Heterogeneous Crowd Behaviors Using Personality Trait Theory. ACM SIGGRAPH Symposium on Computer Animation.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. Perception & Psychophysics, Vol. 14, No. 2, pp. 201-211.

Klein, F., Spieldenner, T., Sons, K., & Slusallek, P. (2014). Configurable Instances of 3D Models for Declarative 3D in the Web. Proceedings of the 19th International ACM Conference on 3D Web Technologies, pp. 71–79. https://doi-org.ezproxy.lib.purdue.edu/10.1145/2628588.2628594

Lemercier, S., Jelic, A., Kulpa, R., Hua, J., Fehrenbach, J., Degond, P., Appert-Rolland, C., Donikian, S., & Pettré, S. (2012). Realistic following behaviors for crowd simulation. EUROGRAPHICS 2012, Vol. 31, No. 2. DOI: 10.1111/j.1467-8659.2012.03028.x

Maciel, P., & Shirley, P. (1995). Visual Navigation of Large Environment Using Textured Clusters. 1995 Symposium on Interactive 3D Graphics.

McDonnell, R., Larkin, M., Dobbyn, S., Collins, S., & O'Sullivan, C. (2008). Clone Attack! Perception of Crowd Variety. ACM Transactions on Graphics, Vol. 27, No. 3. https://doi.org/10.1145/1360612.1360625

McDonnell, R., Larkin, M., Hernández, B., Rudomin, I., & O'Sullivan, C. (2009). Eye-catching Crowds: Saliency based Selective Variation. ACM Transactions on Graphics, Article No.55. https://doi.org/10.1145/1531326.1531361

Millan, E., & Rudomin, I. (2006). Impostors and pseudo-instancing for GPU crowd rendering. GRAPHITE '06: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia, pp. 49–55. https://doi.org/10.1145/1174429.1174436

Peng, C., Park, S. I., Cao, Y., & Tian, J. (2011). A Real-Time System for Crowd Rendering: Parallel LOD and Texture-Preserving Approach on GPU. The Fourth International Conference on Motion in Games, pp. 27-38.

Pettré, J., Ciechomski, P. H., Maim, J., Yersin, B., Laumond, J., & Thalmann, D. (2006). Real-time navigating crowds: scalable simulation and rendering. COMPUTER ANIMATION AND VIRTUAL WORLDS, 17:445–455. DOI: 10.1002/cav.147

Prazak, M., & O'Sullivan, C. (2011). Perceiving human motion variety. Proceedings of the 8th Symposium on Applied Perception in Graphics and Visualization. DOI: 10.1145/2077451.2077468

Reynolds, C. W. (1987). Flocks, Herds, and Schools: A Distributed Behavioral Model. Computer Graphics, Vol. 21, pp. 25-34.

Schwind, V., & Jäger, S. (2015). The Uncanny Valley and the Importance of Eye Contact. Mensch und Computer - Proceedings. https://doi.org/10.1515/icom-2016-0001

Sun, L., Li, X., & Qin, W. (2013). Simulating realistic crowd based on agent trajectories. COMPUTER ANIMATION AND VIRTUAL WORLDS, 24:165-172. DOI: 10.1002/cav.1507

Tecchia, F., & Chrysanthou, Y. (2000). Real-Time Rendering of Densely Populated Urban Environments. Rendering Techniques, pp. 83-88.

Tecchia, F., Loscos, C., & Chrysanthou, Y. (2002). Image-Based Crowd Rendering. IEEE Computer Graphics and Applications, 22(2):36-43. DOI: 10.1109/38.988745

Thalmann, D., Grillon, H., Maim, J., & Yersin, B. (2009). Challenges in Crowd Simulation. International Conference on CyberWorlds. DOI: 10.1109/CW.2009.23

Thalmann, D., & Musse, S. R. (2013). Crowd Simulation (2nd ed.). Springer, London. https://doi.org/10.1007/978-1-4471-4450-2

Ulicny, B., & Thalmann, D. (2002). Towards Interactive Real-Time Crowd Behavior Simulation Computer Graphics Forum,December, Vol. 21(4), pp. 767-775. https://doi-org.ezproxy.lib.purdue.edu/10.1111/1467-8659.00634

Yu, Q., & Terzopoulos, D. (2007). A Decision Network Framework for the Behavioral Animation of Virtual Humans. Eurographics/ ACM SIGGRAPH Symposium on Computer Animation (2007).

Zell, E., Aliaga, C., Jarabo, A., Zibrek, K. Gutierrez, D., McDonnell, R., & Botsch, M. (2015). To Stylize or not to Stylize? The Effect of Shape and Material Stylization on the Perception of Computer-Generated Faces. ACM Transactions on Graphics, Article No. 184. https://doi.org/10.1145/2816795.2818126

Zell, E., Zibrek, K., & McDonnell, R. (2019). Perception of Virtual Characters. In Proceedings of ACM Conference (SIGGRAPH Courses'19). ACM, New York, NY. https://doi.org/10.1145

APPENDIX A. IRB APPROVAL DOCUMENTS

PURDUE
UNIVERSITY

HUMAN RESEARCH PROTECTION PROGRAM
INSTITUTIONAL REVIEW BOARDS

To:	ADAMO, NICOLETTA
From:	DICLEMENTI, JEANNIE D, Chair Social Science IRB
Date:	05/30/2019
Committee Action:(P100)	Determined Exempt, Category **(P100)**
IRB Action Date:	05 / 24 / 2019
IRB Protocol #:	1905022179
Study Title:	The Effect of Different Types of Character Locomotion on Audience Perception of Crowd Animation

The Institutional Review Board (IRB) has reviewed the above-referenced study application and has determined that it meets the criteria for exemption under 45 CFR 46.101(b).

Before making changes to the study procedures, please submit an Amendment to ensure that the regulatory status of the study has not changed. Changes in key research personnel should also be submitted to the IRB through an amendment.

General

- To recruit from Purdue University classrooms, the instructor and all others associated with conduct of the course (e.g., teaching assistants) must not be present during announcement of the research opportunity or any recruitment activity. This may be accomplished by announcing, in advance, that class will either start later than usual or end earlier than usual so this activity may occur. It should be emphasized that attendance at the announcement and recruitment are voluntary and the student's attendance and enrollment decision will not be shared with those administering the course.
- If students earn extra credit towards their course grade through participation in a research project conducted by someone other than the course instructor(s), such as in the example above, the students participation should only be shared with the course instructor(s) at the end of the semester. Additionally, instructors who allow extra credit to be earned through participation in research must also provide an opportunity for students to earn comparable extra credit through a non-research activity requiring an amount of time and effort comparable to the research option.
- When conducting human subjects research at a non-Purdue college/university, investigators are urged to contact that institution's IRB to determine requirements for conducting research at that institution.
- When human subjects research will be conducted in schools or places of business, investigators must obtain written permission from an appropriate authority within the organization. If the written permission was not submitted with the study application at the time of IRB review (e.g., the school would not issue the letter without proof of IRB approval, etc.), the investigator must submit the

written permission to the IRB prior to engaging in the research activities (e.g., recruitment, study procedures, etc.). Submit this documentation as an FYI through Coeus. This is an institutional requirement.

Categories 2 and 3

- Surveys and questionnaires should indicate
  - only participants 18 years of age and over are eligible to participate in the research; and
  - that participation is voluntary; and
  - that any questions may be skipped; and
  - include the investigator's name and contact information.
- Investigators should explain to participants the amount of time required to participate. Additionally, they should explain to participants how confidentiality will be maintained or if it will not be maintained.
- When conducting focus group research, investigators cannot guarantee that all participants in the focus group will maintain the confidentiality of other group participants. The investigator should make participants aware of this potential for breach of confidentiality.

Category 6

- Surveys and data collection instruments should note that participation is voluntary.
- Surveys and data collection instruments should note that participants may skip any questions.
- When taste testing foods which are highly allergenic (e.g., peanuts, milk, etc.) investigators should disclose the possibility of a reaction to potential subjects.

You are required to retain a copy of this letter for your records. We appreciate your commitment towards ensuring the ethical conduct of human subjects research and wish you luck with your study.

Ernest C. Young Hall, 10th Floor - 155 S. Grant St. - West Lafayette, IN 47907-2114 - (765) 494-5942 - Fax: (765) 494-9911

RESEARCH PARTICIPANT CONSENT FORM
The Effect of Different Types of Character Locomotion on Audience Perception of Crowd Animation
Nicoletta Adamo-Villani
Computer Graphics Technology
Purdue University

Key Information

Please take time to review this information carefully. This is a research study. Your participation in this study is voluntary which means that you may choose not to participate at any time without penalty or loss of benefits to which you are otherwise entitled. You may ask questions to the researchers about the study whenever you would like. If you decide to take part in the study, you will be asked to sign this form, be sure you understand what you will do and any possible risks or benefits.
This study is about 3D animation perception. The experiments are simply watching video clips on a computer screen, answer questions and filling out a questionnaire. Everything will be done online using computer. Usually, this study will take no longer than 10 minutes.

What is the purpose of this study?

The purpose is to check whether different characters' movement will affect audience's perception result or not. It can be a useful reference for real world animation production. We would like to explore if duplicated characters in a crowd will be identified by audience within the same time on different types of movement. Based on the result of this study, animation production can reduce the workload properly.
We plan to enroll about 100 participants in this study.

What will I do if I choose to be in this study?

If you would like to participate in this study, you will be invited to open an online link directing to Qualtrics questionnaire platform. Then you will watch three video clips (depend on your perception ability) and fill out questions for each video. After you have watched all the videos, you can also enter your email address if you would like to receive a gift card (optional). Finally, you can close the web page and the data will be recorded. No identifiable data for each individual will be collected.

How long will I be in the study?

This study generally takes 5 minutes to finish. Since participants will not be allowed to watch the video more than once or at their own pace, it takes less than one minute for each scenario. Once participants finish watching the video clips, it takes about one minute to fill out the extra questionnaire.

What are the possible risks or discomforts?

The risks in this experiment are minimal and no greater than daily life. Breach of confidentiality is always a risk with data, but we will take precautions to minimize this risk as described in the confidentiality section. Also, no personal identifiable information (e.g. IP address) will be collected by our survey.

Are there any potential benefits?

There are no direct benefits from this study. For indirect benefits, this study might provide some ideas and techniques from the aspect of perception for crowd animation production.

Will I receive payment or other incentive?

For all participants finished the study, you can choose to enter your email address at the end of the questionnaire. We will randomly draw five persons from all the participants and offer him/her a $25 Amazon gift card as reward.

Will information about me and my participation be kept confidential?

The record and result of this study may be reviewed by departments at Purdue University responsible for regulatory and research oversight. Only two investigators, Prof. Nicoletta Adamo-Villani and Wenyu Zhang will have access to the data. All the research data, including optional participants email address will be deleted after December 31st, 2019.

What are my rights if I take part in this study?

Your participation in this research is completely voluntary. If you agree to participate, you may withdraw your participation at any time without penalty before submitting the survey. If you want to withdraw the study after filling out the questionnaire, you can contact Wenyu Zhang to erase the data. Contact information is listed below.

Who can I contact if have questions about the study?

If you have questions, comments or concerns about this research project, you can talk to anyone of the researchers. Please contact
principle investigator Nicoletta Adamo-Villani, email: nadamovi@purdue.edu
or investigator Wenyu Zhang, email: zhan2586@purdue.edu.

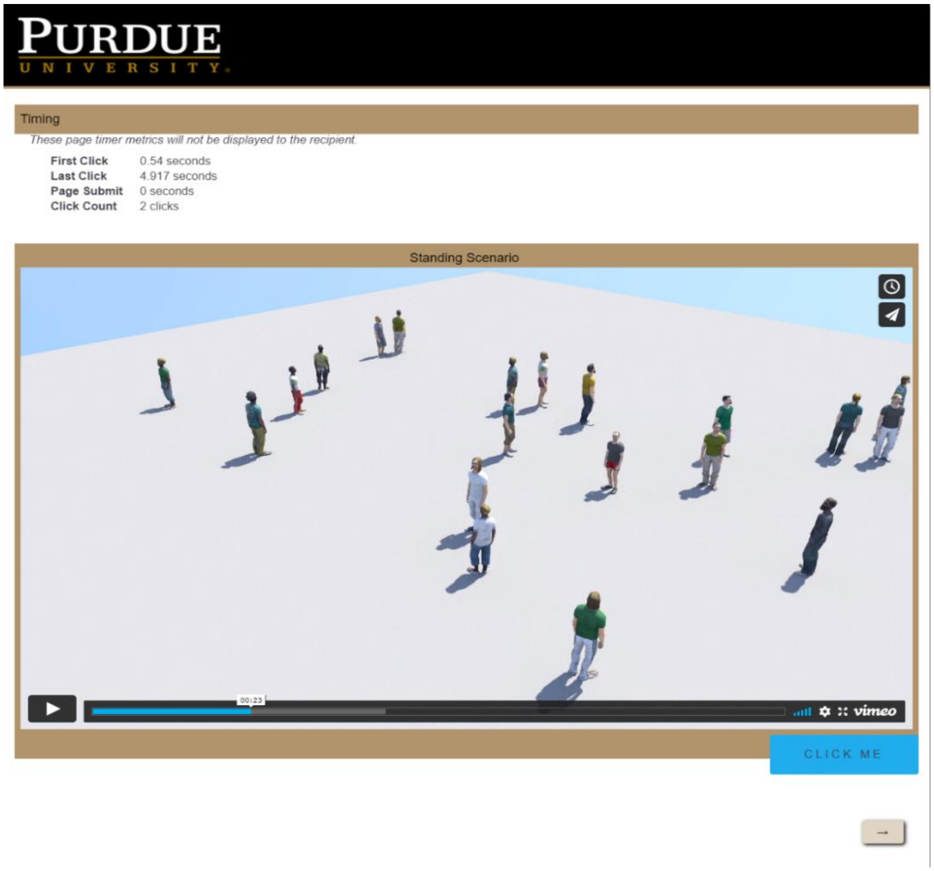
To report anonymously via Purdue's Hotline see www.purdue.edu/hotline.

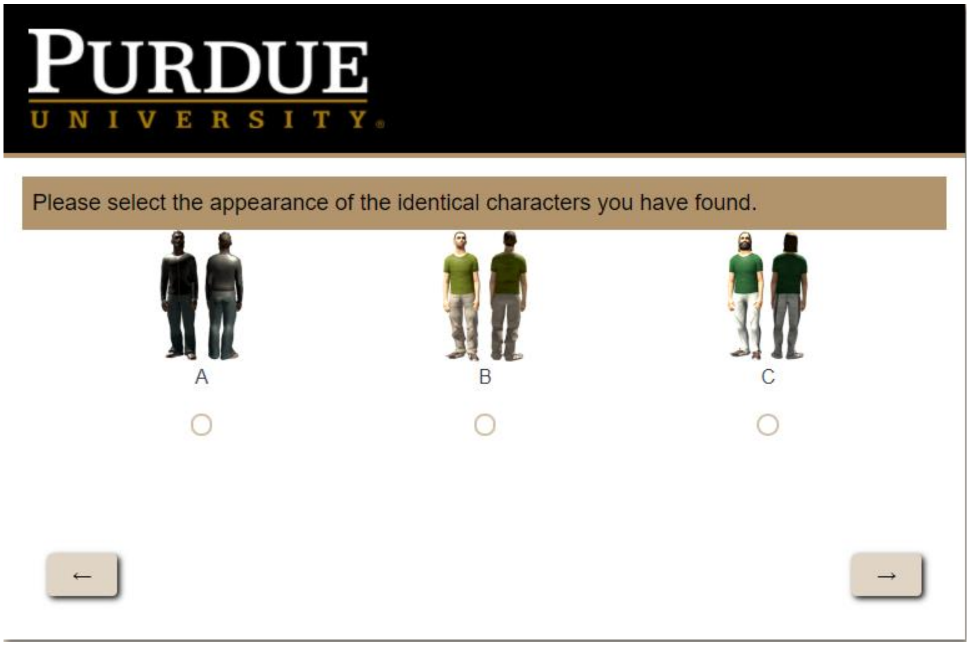
If you have questions about your rights while taking part in the study or have concerns about the treatment of research participants, please call the Human Research Protection Program at (765) 494-5942, email (irb@purdue.edu) or write to:

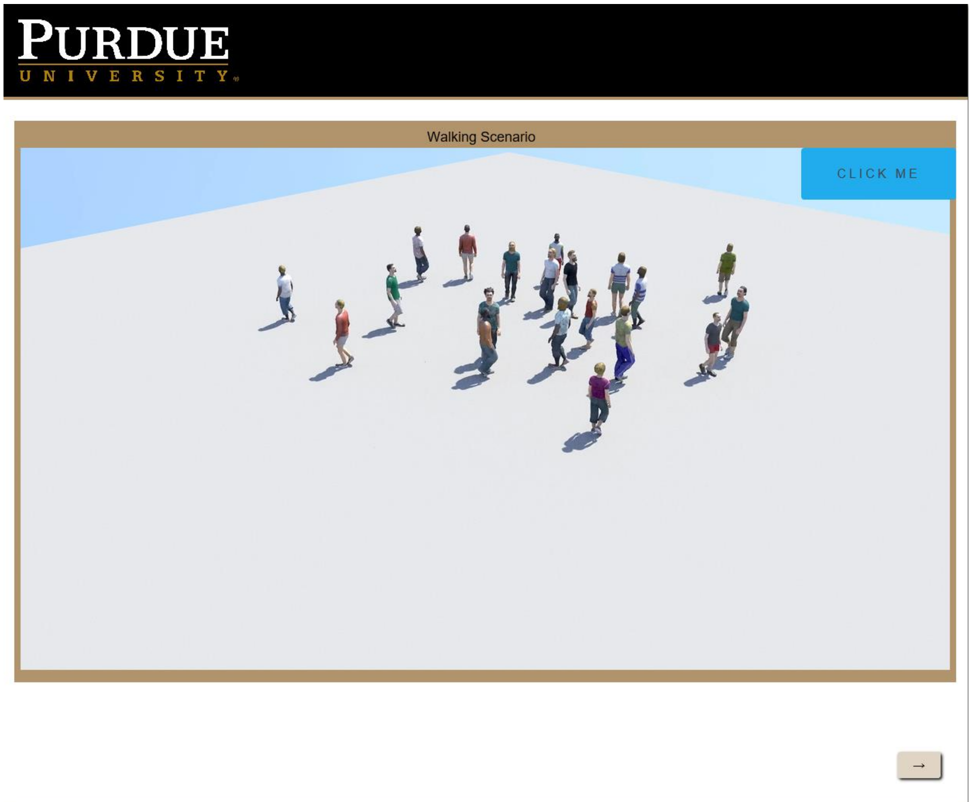
Human Research Protection Program - Purdue University
Ernest C. Young Hall, Room 1032
155 S. Grant St.
West Lafayette, IN 47907-2114

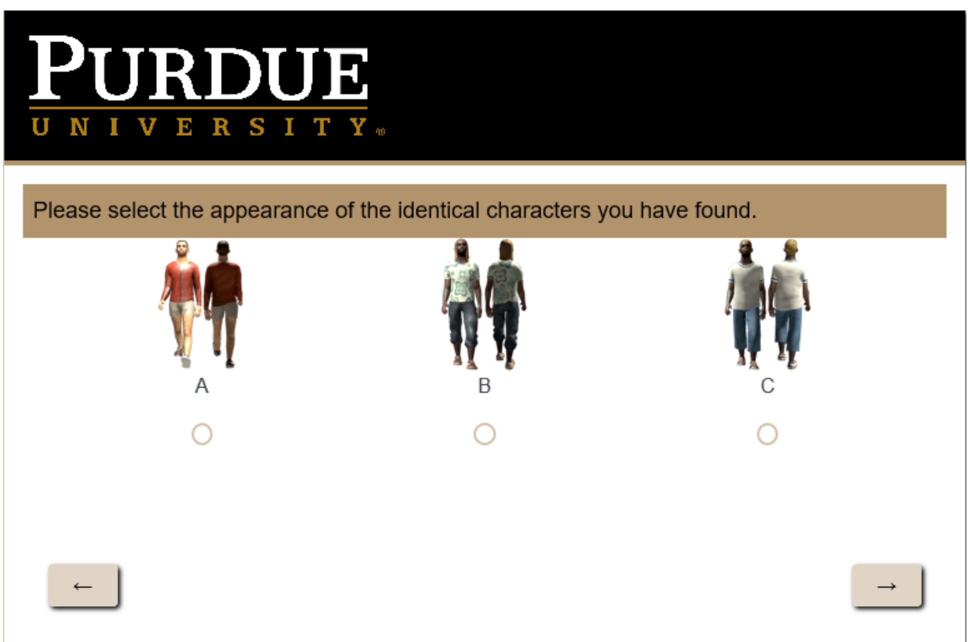
You can click here and proceed to the survey if you agree to participate in this study.

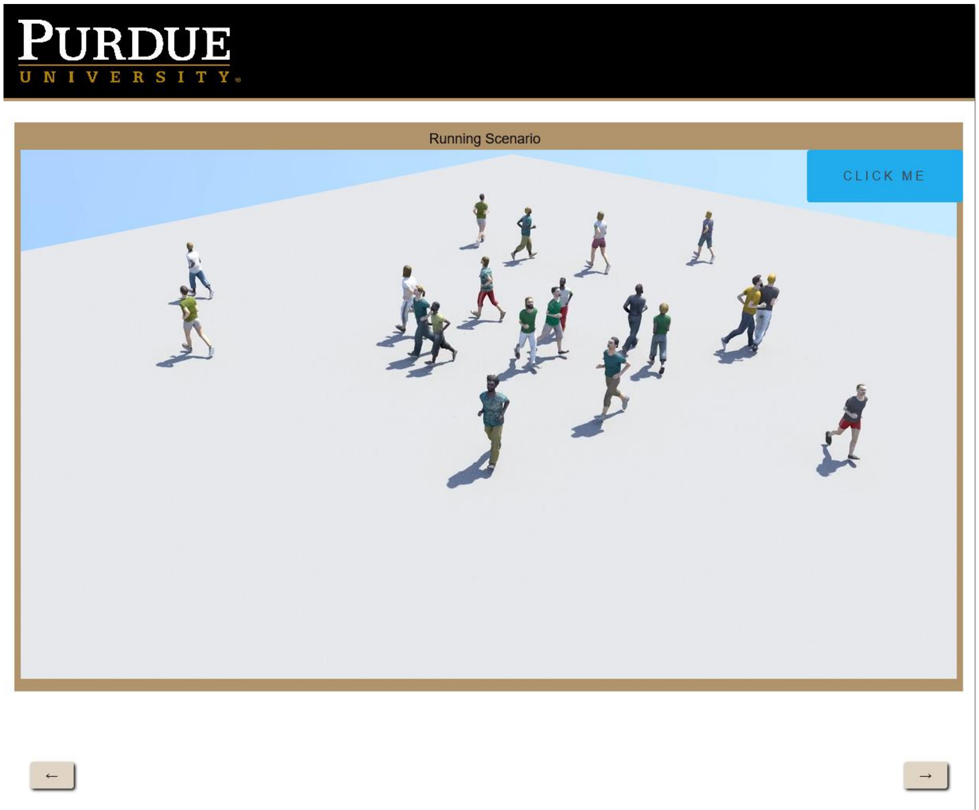
APPENDIX B. ONLINE SURVEY SCREENSHOTS

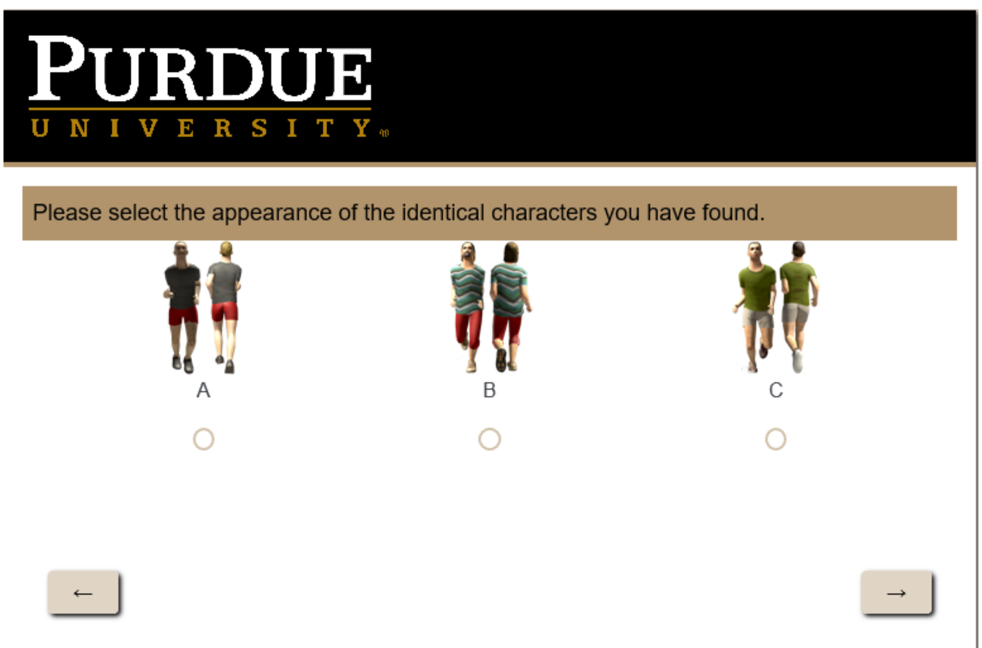





PURDUE
UNIVERSITY
Walking Scenario
CLICK ME
→


PURDUE
UNIVERSITY
Please select the appearance of the identical characters you have found.
A
B
C
←
→


PURDUE
UNIVERSITY
Running Scenario
CLICK ME
←
→


PURDUE
UNIVERSITY
Please select the appearance of the identical characters you have found.
A
B
C
←
→

APPENDIX C. DATA ANALYSIS MATERIALS

	Subject	Gender	Age	Familiarity	Time	Correct	Video	Period	Sequence
3	1	Male	3	2	17.32	1	R	1	2
4	2	Male	4	1	51.192	1	S	1	1
5	3	Male	4	1	24.772	1	W	2	1
6	4	Male	4	1	11.024	1	R	3	1
10	5	Female	2	1	23.954	1	S	3	3
14	6	Male	4	1	10.489	1	W	1	3
15	7	Male	4	1	13.797	1	R	2	3
17	8	Female	4	2	12.215	1	W	2	1
31	9	Female	2	2	26.804	1	S	2	2
32	10	Female	2	2	23.197	1	W	3	2
33	11	Female	2	2	47.275	1	R	1	2
58	12	Male	3	2	23.377	1	S	1	1
59	13	Male	3	2	74.844	1	W	2	1
60	14	Male	3	2	46.495	1	R	3	1
64	15	Female	2	2	86.439	1	S	3	3
65	16	Female	2	2	86.353	1	W	1	3
66	17	Female	2	2	21.952	1	R	2	3
79	18	Male	4	2	32.253	1	S	1	1
80	19	Male	4	2	57.138	1	W	2	1
81	20	Male	4	2	17.163	1	R	3	1
101	21	Female	2	1	84.219	1	W	1	3
102	22	Female	2	1	34.845	1	R	2	3
115	23	Female	4	2	38.457	1	S	1	1
116	24	Female	4	2	82.395	1	W	2	1
117	25	Female	4	2	69.641	1	R	3	1
124	26	Male	3	2	10.883	1	S	3	3
125	27	Male	3	2	81.163	1	W	1	3
146	28	Male	3	2	19.125	1	W	2	1
147	29	Male	3	2	23.626	1	R	3	1
160	30	Male	2	2	22.51	1	S	2	2
161	31	Male	2	2	34.897	1	W	3	2
178	32	Male	2	2	13.016	1	S	2	2
180	33	Male	2	2	65.53	1	R	1	2
187	34	Female	2	2	14.524	1	S	3	3
188	35	Female	2	2	10.218	1	W	1	3
189	36	Female	2	2	56.585	1	R	2	3
194	37	Female	1	2	80.14	1	W	3	2
199	38	Female	1	1	22.723	1	S	2	2
200	39	Female	1	1	28.139	1	W	3	2
201	40	Female	1	1	48.519	1	R	1	2
208	41	Female	1	2	70.515	1	S	2	2
209	42	Female	1	2	34.677	1	W	3	2
210	43	Female	1	2	27.078	1	R	1	2

219	44	Male	2	1	15.735	1	R	1	2
226	45	Male	1	1	32.241	1	S	1	1
227	46	Male	1	1	53.026	1	W	2	1
228	47	Male	1	1	15.915	1	R	3	1
229	48	Male	2	1	22.35	1	S	3	3
231	49	Male	2	1	15.329	1	R	2	3
241	50	Male	1	1	14.722	1	S	1	1
243	51	Male	1	1	21.598	1	R	3	1

Code used for data analysis in RStudio:

```
install.packages("lme4")
install.packages("lmertest")
install.packages("rstanarm")
install.packages("pbkrtest")
install.packages("tidybayes")
install.packages("ggplot2")
install.packages("cli")
install.packages("magrittr")
install.packages("ggplot")
library(magrittr)
library(lme4)
library(lmertest)
library(rstanarm)
library(matrix)
library(pbkrtest)
library(tidybayes)
library(ggplot)

getwd()
setwd("/users/wenyu/documents/purdue university/thesis/data")
dat <- as.data.frame(read.csv("19109_long_subset.csv"))
dat$subject <- as.factor(dat$subject)
dat$period <- as.factor(dat$period)
dat$sequence <- as.factor(dat$sequence)
dat$gender <- as.factor(dat$gender)

# fit the model
fit.all.bas <- stan_lmer(time ~ video + period + (1|subject) + sequence +
gender, data=dat)

# get a list of raw model variables names so that we know what variables we
can extract from the model
get_variables(fit.all.bas)

# spread_draws() lets us extract these indices as columns in the resulting
tidy data frame of draws from b
fit.all.bas %>%
  spread_draws(b[term,group]) %>% # term, group are user defined names of
columns
  head(10)
```

```
# in this particular model, there is only one term ((intercept)), thus we
could omit that index altogether to just get each group and the value of b
for the corresponding subject:
fit.all.bas %>%
  spread_draws('videos') %>% # term, group are user defined names of columns
  head(10)

# extract the draws corresponding to the posterior distributions of the
overall mean and standard deviation of observations
fit.all.bas %>%
  spread_draws(`(intercept)`, sigma) %>%
  head(10)

# if we want the median and 95% quantile interval of the variables, we can
apply median_qi()
fit.all.bas %>%
  spread_draws(`gendermale1`) %>%
  median_qi()

# long-format
fit.all.bas %>%
  gather_draws(`(intercept)`, sigma) %>%
  median_qi(`(intercept)`, sigma)

# plot
fit.all.bas %>%
  spread_draws(`(intercept)`, b[,group]) %>%
  median_qi(condition_mean = `(intercept)` + b) %>%
  ggplot(aes(y = group, x = condition_mean, xmin = .lower, xmax = .upper)) +
  geom_pointinterval()

summary(fit.all.bas)

# interaction plot
dat %>%
  group_by(gender, video) %>%
  summarise(time_groups = mean(time)) -> dat.time.group
  dat.time.group

dat.time.group %>%
  ggplot() +
  aes(x = video, y = time_groups, color = gender) +
  geom_line(aes(group = gender)) +
  geom_point() +
  labs(title = "interaction plot for locomotion type and gender",
       y = "mean response time")

dat.time.group %>%
  ggplot() +
  aes(x = gender, y = time_groups, color = video) +
  geom_line(aes(group = video)) +
  geom_point() +
  labs(title = "interaction plot for locomotion type and gender",
       y = "mean response time")
```