

# Assessing the perceived realism of agent grouping dynamics for adaptation and simulation

O'Connor, S, Shuttleworth, J, Colreavy-Donnelly, S & Liarokapis, F

**Author post-print (accepted) deposited by Coventry University's Repository**

O'Connor, S, Shuttleworth, J, Colreavy-Donnelly, S & Liarokapis, F 2019, 'Assessing the perceived realism of agent grouping dynamics for adaptation and simulation', *Entertainment Computing*, vol. 32, 100323.

<https://dx.doi.org/10.1016/j.entcom.2019.100323>

DOI 10.1016/j.entcom.2019.100323

ISSN 1875-9521

Publisher: Elsevier

**NOTICE: this is the author's version of a work that was accepted for publication in *Entertainment Computing*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Entertainment Computing*, 32, (2019) DOI: 10.1016/j.entcom.2019.100323**

© 2019, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

# Assessing the Perceived Realism of Agent Grouping Dynamics for Adaptation and Simulation

Stuart O'Connor<sup>a</sup>, James Shuttleworth<sup>b</sup>, Simon Colreavy-Donnelly<sup>a</sup>, Fotis Liarokapis<sup>c</sup>

<sup>a</sup>*De Montfort University, The Gateway, Leicester, LE1 9BH, United Kingdom*

<sup>b</sup>*Coventry University, Priory St, Coventry, CV1 5FB, United Kingdom*

<sup>c</sup>*Masaryk University, erotnovo nm. 617/9, 601 77, Brno, Czechia*

---

## Abstract

Virtual crowds are a prominent feature for a range of applications; from simulations for cultural heritage, to interactive elements in video games. A body of existing research seeks to develop and improve algorithms for crowd simulation, typically with a goal of achieving more realistic behaviours. For applications targeting human interaction however, what is judged as realistic crowd behaviour can be subjective, leading to situations where actual crowd data is not always perceived to be more real than simulation, making it difficult to identify a ground truth. We present a novel method using psychophysics to assess the perceived realism of behavioural features with respect to virtual crowds. In this instance, a focus is given to the grouping dynamics feature, whereby crowd composition in terms of group frequency and density is evaluated through thirty-six conditions based on crowd data captured from three pedestrianised real-world locations. The study, conducted with seventy-eight healthy participants, allowed for the calculation of perceptual thresholds, with configurations identified that appear most real to human viewers. The majority of these configurations correlate with the values extracted from the crowd data, with results suggesting that viewers have more perceptual flexibility when group frequency and density are increased, rather than decreased.

*Keywords:* Adaptation, Crowd Simulation, Intelligent Agents, Perceptual

---

*Email addresses:* [stuart.oconnor@dmu.ac.uk](mailto:stuart.oconnor@dmu.ac.uk) (Stuart O'Connor), [csx239@coventry.ac.uk](mailto:csx239@coventry.ac.uk) (James Shuttleworth), [simon.colreavy-donnelly@dmu.ac.uk](mailto:simon.colreavy-donnelly@dmu.ac.uk) (Simon Colreavy-Donnelly), [liarokap@fi.muni.cz](mailto:liarokap@fi.muni.cz) (Fotis Liarokapis)

## 1. Introduction

Realistic crowd simulation is complex and a broad topic for research, due in part to the many elements that comprise a system for producing emergent crowd behaviour (Lerner et al., 2007; Thalmann and Musse, 2012). Similarly contributing, is the fact that virtual crowds are utilised for a wide range of applications (Azahar et al., 2008). We make a distinction between serious applications and those for entertainment, as both are trying to achieve different end results and thus typically utilise different design approaches. Examples of serious applications include simulations for evacuation procedures (Almeida et al., 2013), virtual crowds for specific scenarios such as bioterrorism (Song et al., 2013), and simulations for the reconstruction of cultural heritage (Ulicny and Thalmann, 2002). These types of applications are often modelled using crowd data (Zhao et al., 2018) and have a limited amount of human interaction, being built towards specialised purposes. In contrast, virtual crowds utilised for entertainment are typically geared towards a user-centered focus, wherein user interaction and perception are highly important considerations (Pagulayan et al., 2002). A clear example can be seen looking at crowds applied as interactive elements in video games, with developers stressing the importance of believability or perceived realism (Bernard et al., 2008). Serious games on the other hand, offer a middle ground between a serious focus and entertainment, often borrowing elements from both in their design approach. This differentiation of purpose is crucial, both for research and implementation, especially for crowd simulations designed for user experience where the development process is often highly constrained in terms of time and computational resources (Cournoyer and Fortier, 2015). The choices developers make greatly impact the resultant crowd behaviour and can be the difference between achieving plausible crowds or not.

In this context, crowd behaviour can be defined as the overall outwardly visible results of the underlyingWhen viewing a crowd, some of the most noticeable elements are directly related to the group dynamics. “How many individual groups make up the crowd?” “How many people are in each of these groups?” These are important questions when analysing a crowd at large and are an equally important consideration when implementing a virtual crowd. The number of groups and their density has the potential to alter how the

crowd is perceived. In some urban locations, a crowd consisting of a lot of highly dense groups can be considered more perceptually realistic than a lot of low density groups for example. In others a few highly-dense groups might be considered more perceptually realistic than a lot of highly dense groups. Both frequency and density of groups have the potential to alter the overall composition of the virtual crowd and thus change how it is perceived, building upon previous studies that highlight crowd composition as an important factor (Peters and Ennis, 2009). As such there is room for the exploration of these two key variables to highlight certain perceptual thresholds regarding which values are considered perceptually realistic. ing artificial intelligence algorithms employed in a simulations design (Duives et al., 2013; Anderson, 2003). This emergent crowd behaviour thus incorporates the combined outcome of agent internal motivations and reactions between crowd members, as well as the virtual environment, and other potential stimuli based on the configuration. This crowd behaviour can be broken down further into the concept of behavioural features, whereby specific elements of the resultant behaviour can be identified as focal points (O'Connor et al., 2015).

Typically, a crowd system will include a form of decision making (Luo et al., 2009), pathfinding (Hajibabai et al., 2007; Cui and Shi, 2011) and steering (Reynolds, 1999), which allow agents to perceive, think, and act, to a limited extent. Sometimes, however, these base algorithms do not produce crowd behaviour realistic enough for a simulation to realise its purpose (Schrom-Feiertag et al., 2016). In these instances, additional algorithms can be implemented to impose different effects upon the resultant crowd behaviour (Ijaz et al., 2015; Becker-Asano et al., 2014; Duives et al., 2013). For example, by implementing a Social Forces Model (Helbing and Molnar, 1995) that seeks to replicate social forces between agents, certain dynamics are created. Agents can be attracted or repelled from one another depending upon the specific forces modelled and the weight factors involved, all of which results in different crowd behaviour. As such, the choices made when implementing a crowd system greatly determines the resultant crowd behaviour and particularly in cases involving user interaction, it can be difficult to know how the results will be perceived (Durupinar et al., 2011; Charalambous et al., 2014).

There is a body of existing research for developing and improving artificial intelligence algorithms for crowd based simulation (Malinowski et al., 2017; Banerjee and Kraemer, 2011; Leggett, 2004). Some research aims to utilise new techniques or models to simulate crowd behaviour (Beltaief et al.,

2011; Heïgeas et al., 2010; Kim et al., 2012), whilst other works extend existing behaviours to make them more realistic (Lemercier et al., 2012; Mehran et al., 2009). These lines of research typically result in new and interesting behaviours or refinements, but often do not consider effectiveness from a user perspective (Zhou, 2012). Industry developers of AI for commercial purposes such as video games, particularly note that “AI is always a minefield (Grossman, 2013, p.154), with various examples of ineffective agents (Grossman, 2013). A recent project, for example, that has explored the value of authentic crowd simulation AI in a cultural heritage context, was the Inside Joycean Dublin Project and later the I-Ulysses: Poetry in Motion project, with which the project is associated (Colreavy-Donnelly et al., 2019).

Realism in terms of simulation and virtual environments is an important metric (Loomis et al., 1999; Kort et al., 2003; Kuliga et al., 2015; Bařak et al., 2018), particularly with efforts in research and development seeking to produce realistic virtual crowds (Kang and Han, 2017; Fuchsberger et al., 2017). Additionally, realism has been shown to an influential factor in a number of arenas, including immersion, engagement, and education (Tashiro and Dunlap, 2007; Chalmers and Debattista, 2009). As noted however, depending on the application of the virtual crowd, the requirements in terms of realism can vary (Guy et al., 2012). For serious applications, simulated behaviour needs to be as close to that of a real crowd as possible; this ensures results are accurate for reasoning about real-world environments, for example in areas such as construction and safety (Hajibabai et al., 2007; Aschwanden et al., 2008; Moussaïd et al., 2016). Crowd behaviour exhibited by an application for entertainment however, needs to be perceptually plausible to the users to aid interaction and immersion (McMahan, 2003; Cheng and Cairns, 2005), which is not necessarily the case for current high-fidelity reproductions of real-world behaviours. In these cases, realism is subjective and influenced by perception, rather than being true to what is actually realistic (Sommerseth, 2007). It has been shown that simulation correctness based on actual crowd data is not guaranteed to be the most perceptually realistic configuration for human viewers and that it is possible for parametrised and synthesised behaviour to produce equal or better results (Peters and Ennis, 2009).

In the case of virtual crowds intended for user interaction, since neither simulation correctness based on crowd data, or algorithmic complexity, are universally applicable as metrics for perceived realism, perceptual evaluation can instead be utilised to assess the effectiveness of simulated crowd behaviour. Such methods are more generally used to evaluate graphical ap-

plications, and research exists showing the benefit provided by probing human perception (Ennis et al., 2011; McDonnell et al., 2007a; Peters et al., 2008). The use of psychology, in particular, is a fundamental component under paradigms whereby the end-goal of a visualisation is the human response, rather than a rendered image. One such psychological method is psychophysics (Baird and Noma, 1978).

“Psychophysics is commonly defined as the quantitative branch of the study of perception, examining the relations between observed stimuli and responses and the reasons for those relations (Baird and Noma, 1978, p.1). This method has the key benefit of allowing for the calculation of perceptual thresholds through examination of the psychometric function. These thresholds being fundamental to psychophysics (Fechner, 1966) are the reason for its application in auditory and visual stimuli identification experiments (Melo et al., 2014). However, its application to virtual crowds, and in particular crowd behaviour, has been limited to non-existent, with most studies applying more traditional perceptual evaluation techniques (Peters and Ennis, 2009). However, by combining the use of crowd data with psychophysics through what we refer to here as comparative psychophysics, it adds a potential new dimension to assess and refine resultant crowd behaviour. This comes from identification of the absolute thresholds and being able to compare human perception to the reality of the situation through the use of this crowd data. As we have noted, the choices made when implementing a virtual crowd greatly impact the resultant crowd behaviour and thus it is of benefit to inform developers through these ‘perceptual metrics, particularly when considering user-centered applications.

Realism is shown to be an important aspect of crowd behaviour and different applications necessitate differing qualities and definitions of realism. It is clear that the efficacy of simulated crowd behaviour for user-centered applications must be evaluated against something more sophisticated than just adherence to crowd data and simulation correctness. This work is a proof of concept for the novel methodological approach of applying aspects of psychophysics for threshold calculation towards elements of simulated crowd behaviour. To this end, we examine grouping dynamics, consisting of two variables, group frequency and group density. This selection is due to a number of factors, including the features prominence and wide-applicability to most all crowd simulation scenarios. Thus, given this objective and our focus on grouping dynamics for crowd simulation, we present four research questions:

- RQ1: What are the thresholds for group frequency and group density to be considered perceptually plausible?
- RQ2: What composition of group frequency and group density provides the highest levels of perceived realism?
- RQ3: What effect does location have on the perceived realism of group frequency and group density?
- RQ4: What are the differences for crowd compositions, if any, between crowd data and perceived realism?

By considering these four questions, an outcome of this work is to also provide perceptual metrics in the form of thresholds and optimum configurations to inform and support crowd simulation development, in which human interaction and acceptance is of intrinsic value.

Towards this goal, in this paper, we present here a novel framework for employing a comparative psychophysical method to evaluate the perceived realism of behavioural features. Through a three step-methodology, the grouping dynamic feature is identified, implemented into the urban crowd simulation (Figure 1) and subjected to a quantitative psychophysical experiment using crowd data from three real-world locations. The perceptual thresholds are calculated, the optimum configurations examined and the differences between crowd data and the analysed perceptions are explored. The method and novel application of comparative psychophysics towards behavioural features, is to inform algorithm choice and allow for development flexibility while remaining within the bounds of plausibility through threshold identification, ensuring user requirements are kept the forefront but giving implementation options when resources are constrained.

The remainder of this paper is structured as follows: Section 2 considers related work. Section 3 highlights the methodology and framework. The analysis of crowd footage is detailed in Section 4. Section 5 covers implementation of the crowd simulation through synthesis. In Section 6 perception is outlined with the psychophysics experiment and its results. Finally, Section 7 provides conclusions and areas for further research.

## 2. Related Work

Existing research highlights that psychological methods can be successfully applied to certain aspects of crowd simulation. McDonnell et al. has



Figure 1: The urban crowd simulation developed for the purpose of recreating real-world locations and populating them with virtual crowds for the perceptual evaluation of behavioural features.

published a series of papers using psychophysical methodologies to assess certain graphical elements and animations of agents, the first of which focused on evaluating the level of detail (LoD) effects regarding the clothing of virtual humans (McDonnell et al., 2006). With current consumer graphics hardware, displaying large crowds of agents with fully deformable clothing is typically not possible in real-time, and thus LoD techniques are commonly employed to switch between high and low-quality models. Using psychophysical methods, high resolution models were tested against their lower resolution counterparts and an imposter-based technique. In particular both high resolution models and imposters could reproduce the fluidity levels of cloth more effectively than low resolution alternatives. Using a forced choice method, participants judged the low-resolution models to have the stiffest cloth, whereas imposters were less stiff and high-resolution models were the most fluid. Results indicated that the imposters clothing was closer to the ideal and thus more effective when replacing the high-resolution models, keeping the virtual crowds as realistic as possible. Further studies identified a perceptual metric for the most appropriate update frequency for the different imposter types. Finally,



the effectiveness of these insights was perceptually evaluated in a full crowd-based system, with results complementing earlier studies showing imposters to be perceived as the most realistic substitute. An unexpected result led to the insight that crowd size did not have an effect on the perception of differences between the character types. This study shows in its application that psychophysics can be utilised successfully to judge perceived realism and also highlights the utility of psychophysics for identifying these useful perceptual metrics that can inform development choices in an often-constrained environment.

A further study sought to assess the perceptual thresholds for smooth animation in the form of a pose update rate (McDonnell et al., 2007b). Measured as poses per second (pps), this is defined as the frequency of updates that display when animating a character. This is an important factor to be assessed, since despite technological advances, the utilisation of resources is still an important consideration, and the perceptual significance of certain aspects of animation have not yet been explored. From the results of several psychophysical experiments, acceptance thresholds were discovered that provide guidance for developers wishing to implement smooth animated characters and crowds. From a movement based study considering the poses with linear velocity, it was determined that for the animation to be perceived as smooth the rate at which the pose is updated needs to be around 40 pps. This rate was selected as it is when the probability of acceptance curves become linear, implying that 40pps is the rate at which participants considered all the motion to be smooth 100% of the time. As such, increasing the pps rate above 40 would not yield any higher results. Utilisation of these findings in simulation development has the potential to avoid wasting computational resources on a higher pose update rate than can be perceived by the user. While this study has a focus towards the animation aspects of crowds rather than looking at behavioural elements, it again highlights the potential of applying psychophysics in this realm and the impact these perceptual metrics can have on real-time crowd simulation choice.

Peters and Ennis conducted an analysis of pedestrians using a corpus of static images and video data picturing some of the areas around Trinity College, in Dublin (Peters and Ennis, 2009). This research towards modelling groups of plausible virtual pedestrians provides some insights into using perceptual methods to assess crowds. Using MetroPed, a specifically developed tool, these analysed scenes could be recreated. However, it was hypothesised that some of the new modelling techniques may affect viewer perceptions

of scene plausibility and thus these factors should be explored. As the primary variable, the group ratios were considered. These consisted of seven different ratios, showing different numbers of individual pedestrians, duos, and trios. An additional “no group” condition was also used. These scenes were recorded as short 2-3 second video clips to be utilised during the trials. Participants judged the various clips on their plausibility. The researchers hypothesised that the ratios with more individuals and pairs of pedestrians would be more plausible than the other ratios. Results from the perceptual studies indicated that adding groups to crowd-based simulations does increase the plausibility of the scene. This comes with a stipulation however, as the ratios are an important consideration. It was found the ratios that favoured either more individuals or groups of two were the most plausible for viewers. This provides practical insights into the development of virtual crowds, highlighting that perhaps more attention should be given to the behaviour of individuals or pairs, rather than groups of three or more pedestrians. This study particularly highlights the benefits that can be gained from using crowd data to inform synthesis of the scenario and crowd variables for perceptual evaluation. Furthermore, we again see that tangible insights can be gathered from analysing perception and that crowd composition forms an intrinsic element of whether a virtual crowd will be perceived as realistic or not. Thus, this study informs our methodology and we build upon the crowd composition factor with our focus on grouping dynamics.

As mentioned before, the Inside Joycean Dublin Project, which was associated with the Metropolis project at TCD, explored the value of crowd simulation in a cultural heritage context. Later, the I-Ulysses: Poetry in Motion project, which used assets and development expertise from this project, continued this research, but with a renewed focus on a mobile application using the Unity engine. In this context the user of the I-Ulysses: Poetry in Motion game would encounter puzzles and characters from James Joyces novel Ulysses in a realistic 3-D rendering of Victorian Dublin. An augmented reality experience was intended for the prospective user, whereby they consult real-time locations and context-sensitive events on their phone, as they walked around modern Dublin. As part of this, realistic crowd simulation and an authentic virtual recreation of the historical city was a key element of the experience (Colreavy-Donnelly et al., 2019).

Concluding, there is enough evidence to show the value in applying perceptual and psychophysical methods to evaluate specific aspects of virtual crowds. It has become apparent that by assessing the human perceptions,

clear insights into how the crowds are perceived can be identified. Often, these perceptual studies will result in additional unexpected discoveries, which indicates how a certain factor might be of more significance than was initially thought. This can lead to finding key configurations and specific aspects that are important for the perceptual realism of crowds. This knowledge aids developers as it means they can focus their time and resources on the more important aspects, which influence how the virtual crowd is perceived. As demonstrated using psychophysics has been successful in probing human perception, with regards to specific aspects of virtual crowd simulation. However, these studies have mainly focused on visual aspects, such as graphics and animation. It is also possible to test a number of different variables or a whole simulated scene, however while this can provide a useful indication of how a combination of factors can affect the perceptual realism, it does not necessarily judge the merits of individual components. This is where testing a specific behavioural feature is important, as it not only highlights the importance of that specific factor in comparison to others for modelling purposes, but perhaps the best configurations when implementing said feature in a number of scenarios.

In a psychophysical experiment, a stimulus is selected and altered in intensity and the responses from viewers is measured in order to determine certain perceptual thresholds through the psychometric function. There are several different psychophysical methods that can be employed, ranging from classical to adaptive methods, such as constant stimuli or staircase procedures, with or without forced-choice or comparative elements, to outline a few. Each has distinct advantages depending on the type of stimuli being presented and the experimental setup (Ehrenstein and Ehrenstein, 1999). For our experiments, a comparative constant stimuli approach was selected. Constant stimuli is a classical method, which in the case of this experiment, offers the benefit of exploring the entire psychometric function rather than focusing data around only the threshold point, as typically occurs during adaptive procedures. Additionally, the element of randomness inherent in the method of constant stimuli is invaluable, especially considering the multiple locations and large number of individual trials. Of the classical methods, constant stimuli has been found to be effective (Simpson, 1988) and allows for the psychometric function to be plotted and the absolute thresholds identified. The forced-choice paradigm is not applied here as it is known to produce higher seen responses than is typical, which for discrimination of motion perception can be useful (Gold and Shadlen, 2000),

however in a previous constant stimuli experiment (O'Connor et al., 2015) focusing on crowd behaviour with social forces, this potentially led to issues with threshold identification due to high responses. The comparative method is employed however, which borrows an element of forced choice in terms of stimuli comparison between the simulated and real footage to ground participant's responses. This type of comparison has shown success within similar applications that compare reality and generated content (Peters et al., 2008).

In this research then, we apply these psychophysical methods to specifically assess the behaviour of agents within an urban scene. With crowd behaviour being one of the most important factors for how a simulation is perceived (O'Connor et al., 2013), psychophysical studies may prove to be a useful method for enhancing the perception of crowd simulation.

### 3. General Methodology

For the perceptual evaluation of crowds, an adapted iterative three-stage methodology is employed, covering analysis, synthesis, and perception, with the results being added to a corpus of perceptual data which then informs the next iteration of synthesis (Figure 2). These stages, in more detail, are: (1) analysis of real-world instances of crowd behaviour; (2) synthesis of crowd behaviour into a simulation; and (3) perceptual evaluation of the resultant crowd behaviour. Related research has shown that conducting these three tasks is a necessity for producing a robust study that contains a model, which is 'grounded in reality' (Peters and Ennis, 2009). By consulting crowd data and the corpus to build the scenario and parameters it ensures real-world grounding exists for the model, as often variables are easily neglected if this step is not taken. This method is intended for multiple iterations, each focusing on a different behavioural feature so a corpus of perceptual data is built up that can be utilised for various applications within the scope of crowd simulation and beyond. This corpus consists of the optimal and threshold values for various behavioural features along with the insights gained, so that over time more comprehensive crowd simulations can be constructed based on prior results. The intention is for contribution towards user-centered design principles for virtual crowds, with the corpus informing the development process in terms of algorithm and parameter implementation, as well as factors such as crowd composition.

Algorithm construction and modelling are informed by analysis of video data that shows real-world crowd behaviour within the location of interest.

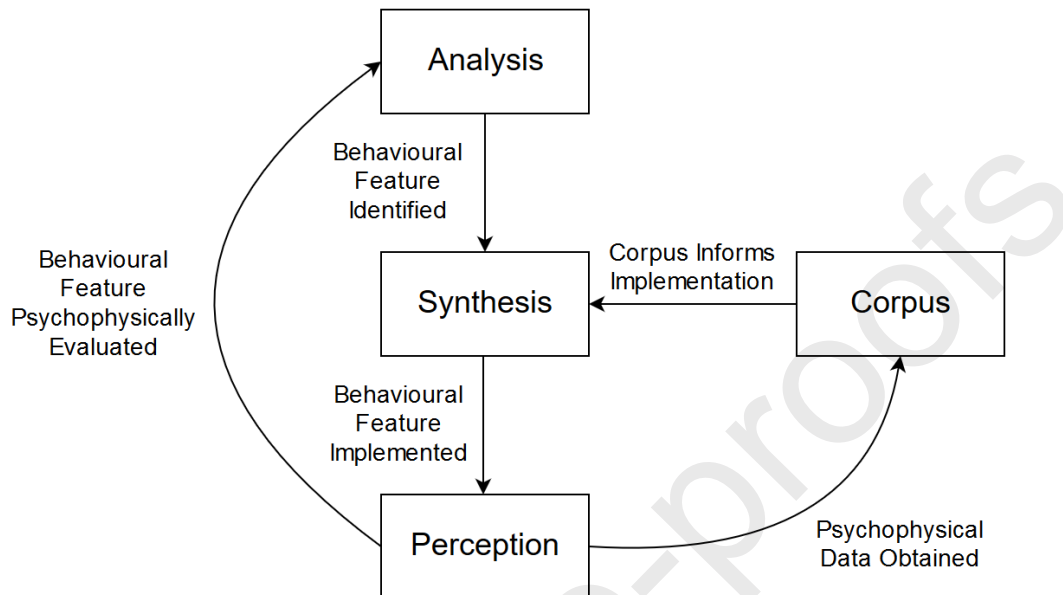


Figure 2: The adapted form of the analysis, synthesis and perception methodology, with the perceptual data being built-up into corpus over multiple iterations.

On occasion, variables that are not predicted by the developers are discovered through these visual analyses. This can lead to a more robust simulation for perceptual evaluation. Modelling a simulation based on crowd data is not an uncommon process, with a growing number of examples of these types of methodologies in the literature (Guy et al., 2012). In addition to informing the modelling process for the simulation, the analysed crowd data can also highlight possible behavioural features that can be perceptually evaluated, and their initial values can then be altered to produce the required stimuli. Specifics of this analysis stage are discussed in Section 4.

Using the analysed data from the previous stage, it is then possible to model algorithm construction and the virtual environment. In a process of synthesis, behaviours are created with the parameter spaces left open for perceptual evaluation to determine the optimum and threshold values. Using values drawn from the real-world instances, behaviour is initially grounded in reality before being subtly altered for psychophysical stimuli. Additionally, the environment is modelled based upon the real world location, taking into account the most important features that have an impact on crowd behaviour. Details on synthesis are discussed in Section 5.

The final stage is to perceptually evaluate the synthesised behaviour, using a psychophysical based study. This adapted comparative experiment shows participants both simulated footage, and the original footage the simulation was based upon. This provides a point of reference for evaluation and in this manner the perceived realism of the specific behavioural feature can be assessed. It allows for the calculation of perceptual thresholds through the psychometric function and identification of the most ideal intensities for a given feature. These results can then inform the synthesis stage on future iterations of method or other related crowd simulations. The psychophysical experiment and the results are discussed in Section 6.

#### 4. Analysis

Crowd behaviour in general may appear to be fluid in nature, and though some virtual crowds have been modelled using fluid dynamics, it is more common that crowd behaviour is modelled using agent-based systems (Kapadia et al., 2012), since this allows for control over individuals and flexibility in implementing unexpected behaviours present in reality. When developing behaviour, it is possible for some behaviours to be overlooked, e.g. emergent behaviours such as abrupt changes in direction or spontaneous ceasing of motion.

In the work presented here, in the analysis stage, real world crowd behaviour is considered in order to highlight these potentially unexpected behaviours and inform the development of virtual crowds and environments. Over time, and with multiple iterations of the general methodology, a corpus of analysed perceptual data can be built-up.

In this paper, the main medium considered for analysing crowd behaviour is video footage, showing actual pedestrians in a variety of urban locations. It is possible to use other mediums such as photographs, however video gives a more accurate representation of the motion, grouping and behaviour of the pedestrians. A number of high profile urban areas were considered for this study, however three main locations were selected for the behaviour shown and their implementation viability. These three locations are Bourbon Street in New Orleans, 41st Street in New York and Temple Bar in Dublin (Figure 3).

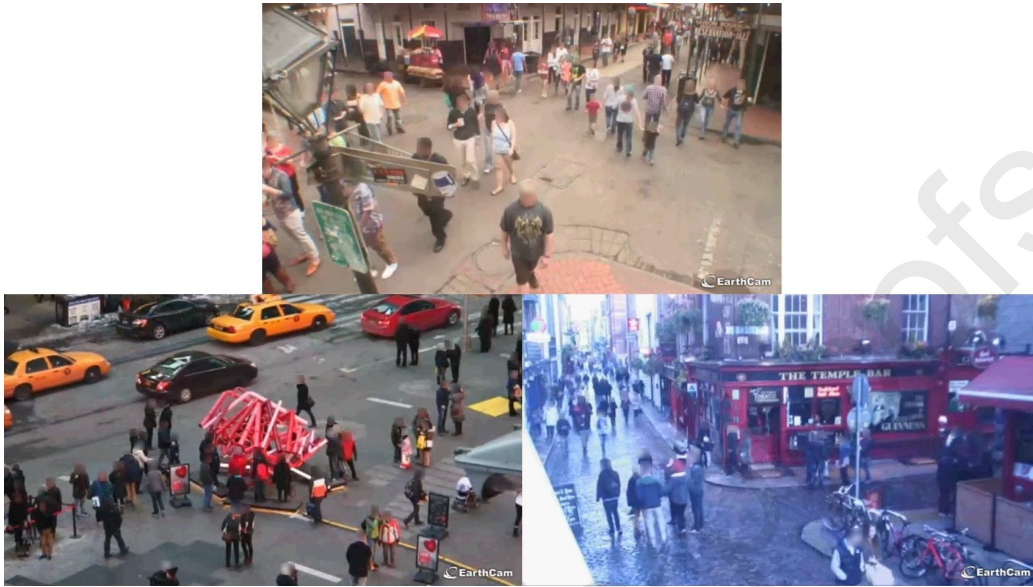


Figure 3: Three stills from the video footage that shows crowd behaviour in three high profile locations for the purposes of analysis.

#### 4.1. Behavioural Feature

The identification of a behavioural feature is an important aspect for this research, as it provides a focal point for the analysis and subsequent perceptual evaluation.

A behavioural feature is a specific element of the overall emergent crowd behaviour. Over multiple iterations of this methodology, an increasing range of behavioural features are identified, offering key ideas about how to implement virtual crowds. The most apparent behavioural feature that has not already been examined and is shown prominently within the crowd footage is selected. In this manner the most common features are evaluated first, as it can be argued they have the most initial value when producing crowd behaviour. Over time, however, more specific and less prominent behavioural features would also be analysed. Past features of this research include varying velocity (O'Connor et al., 2013) and social forces (O'Connor et al., 2015), both of which encapsulated multiple variables.

The steps for behavioural feature selection are as follows:

1. Consult crowd footage based on scenario and composition.

2. Identify instances of emergent behaviour, which have not been previously evaluated, through annotation.
3. Compare identified instances of emergent behaviour for prominence across footage.
4. Select the most prominent behavioural feature and any identified variables therein based on annotation results.

For this iteration, the identified behavioural feature is grouping dynamics, with two key variables for group frequency and group density.

When viewing a crowd, some of the most noticeable elements are directly related to the group dynamics. “How many individual groups make up the crowd?” “How many people are in each of these groups?” These are important questions when analysing a crowd at large and are an equally important consideration when implementing a virtual crowd. The number of groups and their density has the potential to alter how the crowd is perceived. In some urban locations, a crowd consisting of a lot of highly dense groups can be considered more perceptually realistic than a lot of low density groups for example. In others a few highly-dense groups might be considered more perceptually realistic than a lot of highly dense groups. Both frequency and density of groups have the potential to alter the overall composition of the virtual crowd and thus change how it is perceived, building upon previous studies that highlight crowd composition as an important factor (Peters and Ennis, 2009). As such there is room for the exploration of these two key variables to highlight certain perceptual thresholds regarding which values are considered perceptually realistic.

While these were the two variables that were identified with respect to this behavioural feature, there are other elements of crowd composition that are potentially relevant grouping dynamics. This includes attributes such as the velocity magnitude, ratio between groups, and physical qualities of the individuals. These do not specifically form a part of this behavioural feature however and so are not considered key variables, however they could pose as potential limitations and thus are noted in order to remain constant at simulation runtime to reduce potential impact in terms of perceptual evaluation.

#### *4.2. Video Analysis*

As discussed in Section 4, three high profile locations were selected for analysis (Figure 3). These locations were selected due to two main factors;



firstly, they each consisted of highly trafficked urban areas with a comparable high density of pedestrians present, and secondly, they were similar in terms of their behavioural features as identified through feature selection in Section 4.1. High density crowds with corresponding behavioural features are important for this psychophysical experiment, as one of the main areas of investigation is the absolute thresholds. By providing a comparable baseline over the three locations, the influence of environmental configuration on these thresholds can be assessed. In addition, by selecting typical pedestrianised streets, which are a prominent type of location utilised in both video games and serious applications, results can be applicable to a wide range of future endeavours. Other important factors for the selection of these three locations include; all being central locations in major cities, all having viable a viable perspective on their pedestrian lanes, all having limited boundaries comparable to the size of the space, and all being viable for reconstruction as part of synthesis.

The crowd footage was acquired at peak times during clear weather from live web camera feeds that were a part of the EarthCam network (EarthCam, 1996). Each of the locations featured high numbers of pedestrians which allows for the collection of useful behavioural data, to later influence simulation development. By selecting three locations with similar crowd behaviour, perceptual thresholds for each area can be compared more accurately in the perception stage as detailed in Section 6.

With the identification of crowd grouping dynamics as the behavioural feature for analysis Section 4.1, it means that only this specific behaviour now needs to be taken into account in order to arrive at initial values for main variables and inform synthesis. Previously viewing the overall footage has highlighted different and unexpected behaviours that can be used to inform simulation design, but for the final part of the analysis stage a manual annotation is tailored towards the behavioural feature. In this process an annotator manually adds semantically relevant descriptors. Generally, the annotation should be as simple and efficient as possible, only considering those factors relevant to the study (Peters and Ennis, 2009). This is of particular importance when considering video footage, where different frames need to be analysed. In this case, we adopted use of a process using simple graphical mark-ups where you can detect the important feature throughout a scene at a quick glance (Figure 4).

A short video clip of 30 seconds for each location is cut from the captured footage, showing what is determined to be average behaviour. These three



Figure 4: A keyframe that highlights the process of annotation whereby the present groups are tagged using specific colours to denote density.

clips are then used for the analysis and later presented to participants during the perceptual study outlined in Section 6. Through visual identification at key frames from the beginning of each clip and each 10 second interval thereafter, groups are manually annotated with their configuration recorded. The behavioural feature consists of two key variables, frequency and density. Group frequency refers to the number of different groups that are present within the given scene and group density refers to the number of individuals in these groups. By annotating the groups using the mark-up method, values for these two variables are tagged and can then be calculated at each key frame. These are then combined and averaged to give values for that specific video clip. These values are important for the synthesis stage, in order to inform algorithm parameters and modelling efforts, as outlined in Section 5.

## 5. Synthesis

The synthesis stage sees the utilisation of data gathered during analysis to inform development of a tailored crowd simulation for eventual perceptual analysis. This consists of two key areas of implementation. Firstly, in Section 5.1 the development of the virtual environment to match the scene that was depicted in the analysed video footage is outlined. Secondly, in Section 5.2 the implementation of the behavioural feature through algorithm design and parametrisation for intensity alteration is discussed. The chosen development environment for this crowd simulation is the Unity Engine. Using game-based technologies is not uncommon for these types of simulations (Szymanczyk

et al., 2012), as they offer a flexible platform to build from with support of a number of useful technologies. By using Unity we were able to render the urban locations and populate them with virtual crowds. In addition, it allowed for the creation of varied scenarios and the different configurations required for the psychophysical evaluation as noted in Section 6.1.

### 5.1. *Virtual Environments*

In crowd simulation the virtual environment is a crucial aspect for the interaction as it provides the navigation platform for the agents. In certain cases, the environmental setup determines some of the behaviours exhibited by the virtual crowds. For example, a narrow area can cause groups to use more extreme measures to avoid collisions. This being considered, for this research three similar real urban locations were chosen and recreated in virtual space. This consisted of three distinct steps that allowed for the creation of virtual environments that were ready to be populated with agents. Firstly, modelling the 3D geometry and structuring the environment as presented in the crowd footage. Secondly, setting the walkable areas and informing collision detection. Thirdly, setting the different agent paths through the environment for high level navigation. It should be noted that a bespoke implementation was employed without utilisation of Unity’s NavMesh functionality.

As highlighted in Section 4, three urban locations were selected from a variety of crowd footage. For the purposes of the psychophysical experiment, the virtual environment was modelled to resemble the original footage (Figure 5), with focus given to environmental configuration rather than identical graphical representation. Unity utilises a system of prefabs, whereby game objects with their defined properties can be stored in this manner, to act as a template for future use within a scene. Multiple instances of the same object can be added to the scene, inheriting their initial properties from their prefab. By producing prefabs for common urban objects and architecture, the three scenes needed could be efficiently built from these templates. The main consideration was ensuring that the areas available for navigation by pedestrians closely matched those presented in the footage. By using stills from the footage as a visual guide, the camera perspective was defined based on the position and orientation from the real-world scene. From there the prefabs were positioned and transformed manually using a zone-based approach from this perspective. In that manner, the virtual scenes were built from the ground up using these prefabs.



Figure 5: A comparison between the three real-world locations and the modelled virtual locations.

The virtual environments provide the three-dimensional geometry in which the agents can be generated, however other properties had to first be implemented to ensure that the agent systems behaved appropriately. The walkable areas and inaccessible areas needed to be communicated to the agents so that they kept to the locations required and did not collide with elements of the geometry. By incorporating vector fields of varying magnitude into the environment that agents would adhere to, it was possible to keep the agents within the walkable areas whilst still allowing them the freedom of movement required by their steering mechanisms. These vector fields remained the same over all runs of the simulation.

The paths that agents followed were manually defined within the virtual

environment based on the lanes utilised within the crowd footage. Using a series of connected nodes, common to pathfinding algorithms, the main paths were defined so that when initialised groups of agents would have a starting point and a destination, with a path of nodes to work through. Pre-defined paths were created to match the maximum number of groups used across all instances, meaning that while all paths may not be in use in due to the number of groups in an instance, the paths themselves remain constant throughout runtime.

By combining these three elements, namely the virtual environment, the vector fields and the paths, agents could be initialised for the purposes intended. The virtual environment is modelled to be navigated, with vector fields denoting the key areas where the agents should walk and avoid. Agents will follow the paths set within the virtual environment to reach their respective destinations. A total of three scenes, modelled after the real-world locations, were created for the purpose of stimuli creation when coupled together with the agent systems as outlined in Section 5.2.

### *5.2. Agent Systems*

The agent systems are responsible for giving the virtual crowds their form of intelligence. The types of algorithms implemented affect the kind of behaviours exhibited by the agents. A typical implementation of an agent system will often contain a decision-making system, high-level pathfinding navigation, and local steering mechanics. The purpose of simulations conducted in this research are to psychophysically test the crowd grouping behavioural feature, which consists of the two variables crowd frequency and density. It was important to limit the behavioural variability of the agents for the purpose of perceptually evaluating the behavioural feature, as implementing more complex algorithms that consider more factors such as social forces would potentially lead to emergent behaviour influencing results. As such, the sophistication of the agent systems was kept suitably limited with two core components: steering mechanics and the group controller.

Steering mechanics were pioneered by Reynolds initially as Boids (Reynolds, 1987). It is a type of flocking model typically shown with three main steering behaviours, separation, alignment and cohesion. Separation controls the steering force responsible for avoiding other local flock members, alignment controls the steering force for the average heading of other local flock members, and cohesion controls the steering force to move towards the average position of other local flock members. Each member has a form of simple

radial perception with a blind spot at the rear, in order to access the local neighbourhood around itself and enforce the steering behaviours. The outcome is a simplistic behaviour that serves to capture the group dynamics for a flock.

To limit emergent behaviour, steering is utilised with an adapted Boids model consisting of multiple forces calculated based on local neighbourhoods and weighted into determining an agent's final position during an update cycle. Several key steering behaviours are initialised for each agent in a defined group:

- Cohesion: Determines how closely each member of the group sticks together.
- Alignment: Determines how closely the group follows in formation.
- Separation: Determines how far agents in the group can move apart from each other.
- Separation from other groups: Determines how the group as a whole move to avoid colliding with other groups.
- Path-following: Determines the path groups follow within the virtual environment to reach their intended destination.

These behaviours consist of two key variables, a weight to determine how much impact a specific behaviour has in relation to the other behaviours, and a radius to determine the size of the local neighbourhood considered for that behaviour. The key difference to a standard Boids model is the weighting factor and the two additional steering behaviours, separation from other groups and path-following. Separation from other groups functions the same as separation, except that it is calculated at group level instead of agent level. Path-following was similarly outlined by Reynolds in his work for autonomous characters (Reynolds, 1999), producing a corrective force pointing agents down the specific paths and allowing them to respond dynamically to the environment. A path-following steering behaviour was more effective for the purposes of this simulation than a specific pathfinding algorithm, due mainly to the short paths that were required for stimuli production which removed the need for a decision-making algorithm to select new destinations. As paths needed to be kept the same over different runtimes, the dynamic nature

of an agent system with decision making and pathfinding would not have been beneficial to the psychophysical evaluation. When taken with the other steering behaviours in our adapted model, it prevents agents from exhibiting the flocking behaviour and instead they become autonomous pedestrians in their select groups. As it is the behavioural feature that is being altered, the values for these steering behaviours are kept constant for all agents within the simulation, across all runtimes.

The group controller is overhead of the steering mechanics and conducts the specification of the group frequency and group density based on the selected configuration file at runtime (Table 1). Firstly, the number of groups are initialised, each of which holds a variable for the number of agents, which are then subsequently populated in a designated area calculated and visualised through a connected spline. The configuration can be altered as needed for the group frequency and density simply by changing the integer values of the corresponding variables, allowing for the capture of varied stimuli as required. Individual agents are selected from a large quantity of agent prefabs previously modelled and animated. In this manner, the majority of agents seen within the simulation will be unique so as to not influence results by using only a few different agent models that are repeated obviously within the scene. In addition, each agent encapsulates its own variables, including a minimum and maximum speed, a size represented as a radius, a maximum force, and a mass, which are initialised with values from the configuration. As the behavioural feature being tested is specifically related to the group frequency and density, these values are kept constant for all agents within the simulation.

Taking both components in to account then, the crowd behaviour is defined as  $F_R = F_1w_1 + F_2w_2 + F_3w_3 + F_4w_4 + F_5w_5$ , where  $F_R$  is the resultant force,  $F_1...F_5$  are the five calculated steering forces, and  $w_1...w_5$  are the corresponding weight factors. This is then applied through the physics with Newton's 2nd law  $a = \frac{F_R}{m}$ , where  $a$  is acceleration and  $m$  is mass. With an agent's acceleration at a specific update cycle, a new velocity is then calculated and from that its updated position determined by this velocity and its previous position. In total, taken with the virtual environment detailed in Section 5.1, this has produced a simulation whereby the crowd grouping behavioural feature can be altered for both group frequency and density, with the crowd behaviour itself kept standardised, in order to record effective stimuli for psychophysical experimentation (Figure 6). The experiment setup is discussed in Section 6.1, outlining the stimuli recorded through the

configurations of group frequency and density based on analysis.



Figure 6: Output stimuli from the urban crowd simulation showing both low and high intensities for each location. Groups follow paths defined within the environment and can be seen walking on the roads due to the pedestrianisation of these locations and is representative of the paths found in the footage.

## 6. Perception

The previous sections detail how a behavioural feature can be identified and then incorporated into a crowd simulation. The final step in the methodology is to perceptually evaluate this behavioural feature. This provides feedback regarding the perceived realism of the feature and can identify key intensities or parameter values that appear most realistic to viewers.



### 6.1. Method

A total of seventy-eight participants (69 males; 9 females) (4 aged 13-18; 57 aged 19-25; 12 aged 26-35; 5 aged 36-50) from various education backgrounds but all with basic computing skills took part in this experiment, registering responses for each of the experiment's 36 trials. In order to reach large numbers of participants, a psychophysics platform was developed and deployed online to improve accessibility (Figure 7). Upon navigating to the webpage and beginning the test, a participant is presented with instructions and must agree to be part of the experiment to continue. Next, basic demographic data is taken and the experiment begins. The group frequency trials are presented first, showing one location after the other. Finally the group density trials are then presented in the same manner. On each trial a participant is shown the real video clip on the left and the simulated video clip on the right for a duration of 10-seconds. They are asked how realistic do you find the groups in the simulated video and must select a percentage from a drop down box, in intervals of 5%, before they can submit and move on. All participations were informed of the scale prior to starting the experiment as part of the instructions. The reason this scale was selected was due the resolution required for accurate probing of the psychometric function and thus identification of the perceptual thresholds. This was further informed through our previous studies, in particular our work on social forces (O'Connor et al., 2015).

For the method of constant stimuli without any forced choice element, the absolute thresholds are taken at the 50% seen level and a seen result is classified as a perceived realism response of 50% or higher. As noted previously, three virtual scenes are utilised based on three real-life locations. Each of these locations had 6 trials or intensities for each of the two variables being examined. This adds up to 12 trails per location, for a total of 36 trials in all. The numbers in each group are altered and the simulations conducted to record 36 short video clips that last no more than 10 seconds each. 6 key intensities values are used, 10%, 30%, 50%, 70%, 90% and 100%, however the actual group frequency or density these translated into were slightly different depending on the location (Table 1). The 50% value was extracted from the crowd data, so it is the value that most accurately represents reality. This value is then altered on a scale to give the other intensity values. By using 6 different intensities for each variable and at each location, a distribution of data was generated and used to plot the psychometric function to identify the key perceptual thresholds.

Table 1: Intensity value metrics for each variable at each location.

Intensity (%)	Bourbon Street (Frequency)	Bourbon Street (Density)	41st Street (Frequency)	41st Street (Density)	Temple Bar (Frequency)	Temple Bar (Density)
10	4	1.0	4	1.0	3	1.0
30	12	1.4	13	1.4	10	1.4
50	20	2.4	22	2.3	16	2.3
70	28	3.4	31	3.2	22	3.2
90	36	4.3	40	4.1	29	4.0
100	40	4.8	44	4.6	32	4.5

### 6.2. Results

For the each of the three locations and for both variables, group frequency and group density, psychometric functions are plotted to highlight responses (Figure 8). In addition, all three graphs show that the range of stimuli intensities were broad enough to have allowed for the absolute thresholds to be identified, due to the fact that the point at which the psychometric function crosses the 50% observed mark on the ordinate is visible for all functions for both the upper and lower thresholds.

The absolute thresholds for group frequency in Bourbon Street are identified as 35% intensity and 90% intensity, which links to a frequency of 14 and 36. In addition, the absolute thresholds for group density in Bourbon Street are identified as 35% intensity and 81% intensity, which links to a density of 1.6 and 3.9. The absolute thresholds for group frequency in 41st Street are identified as 41% intensity and 94% intensity, which links to a frequency of 18 and 42. In addition, the absolute thresholds for group density in 41st Street are identified as 39% intensity and 84% intensity, which links to a density of 1.8 and 3.8. The absolute thresholds for group frequency in Temple Bar are identified as 35% intensity and 95% intensity, which links to a frequency of 12 and 31. In addition, the absolute thresholds for group density in Temple Bar are identified as 33% intensity and 81% intensity, which links to a density of 1.5 and 3.7.

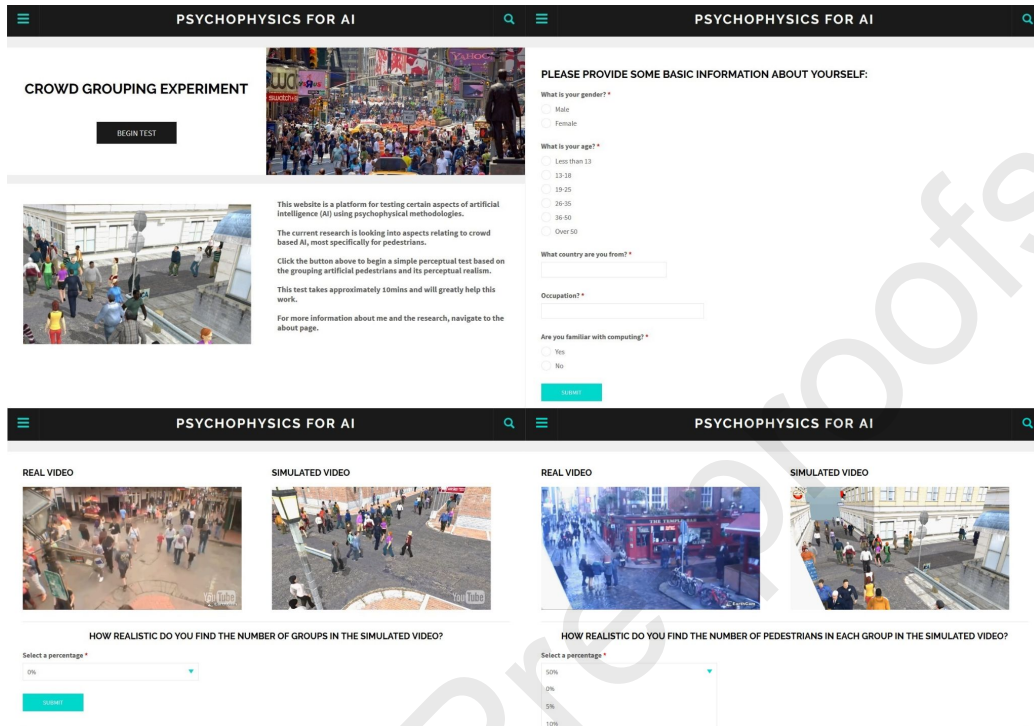


Figure 7: Different pages of the online platform; top left image is the home screen, top right is demographic data collection, bottom left is a Bourbon Street frequency trial and bottom right is a Temple Bar density trial.

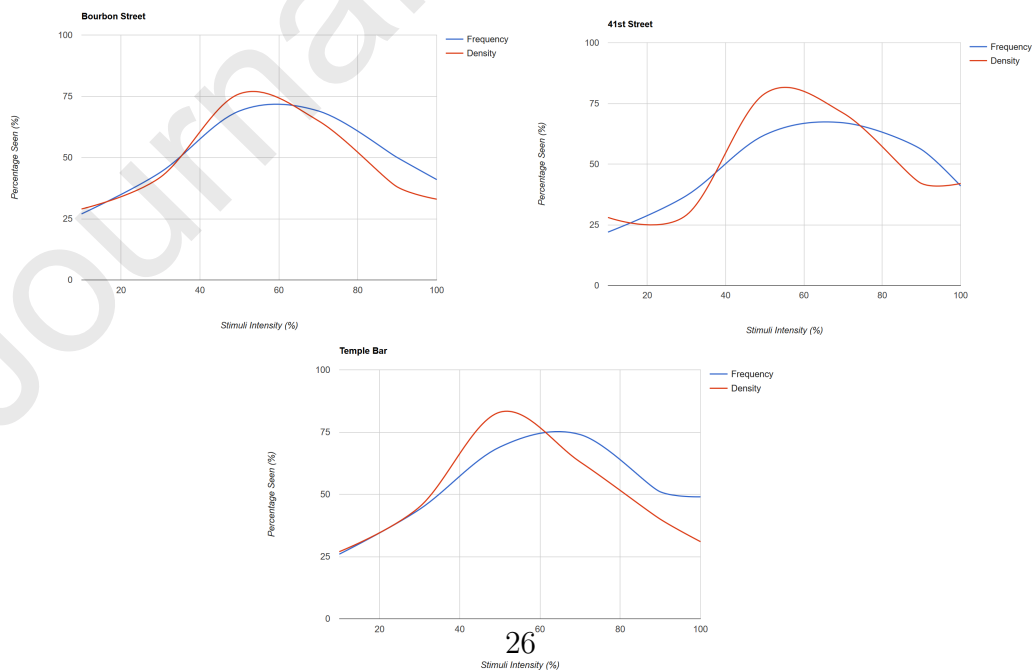


Figure 8: Graphs showing the psychometric functions for group frequency and group density at the Bourbon Street, 41st Street, and Temple Bar locations, with percentage seen responses on the y-axis and stimuli intensity on the x-axis.

The optimum configuration is identified using the perceived realism values collected through the psychometric data (Figure 9). For group frequency in Bourbon Street, the highest perceived realism value of 0.7 is at 50% intensity, equating to a value of 20 groups. For group density in Bourbon Street, the highest perceived realism value of 0.81 is at 50% intensity, equating to a value of 2.4 agent per group. For group frequency in 41st Street, the highest perceived realism value of 0.68 is at 70% intensity, equating to a value of 31 groups. For group density in 41st Street, the highest perceived realism value of 0.83 is at 50% intensity, equating to a value of 2.3 agents per group. For group frequency in Temple Bar, the highest perceived realism value of 0.71 is at 70% intensity, equating to a value of 22 groups. For group density in Temple Bar, the highest perceived realism value of 0.88 is at 50% intensity, equating to a value of 2.3 agents per group.

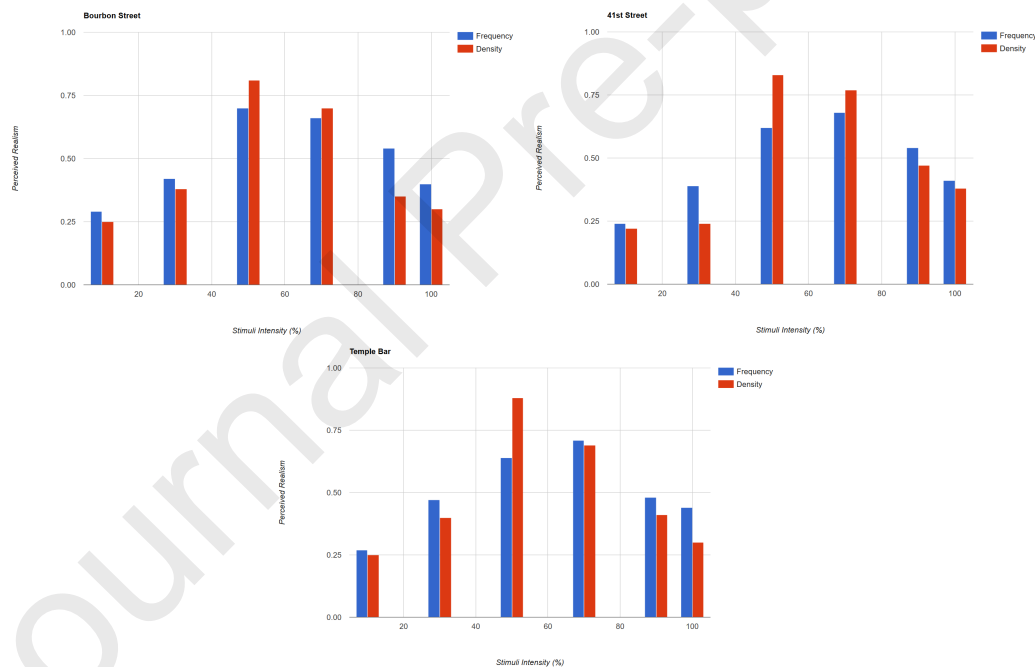


Figure 9: Graphs showing the perceived realism values for group frequency and group density at the Bourbon Street, 41st Street, and Temple Bar locations, with perceived realism on the y-axis and stimuli intensity on the x-axis.

Two one-way repeated measures analysis of variance (ANOVA) were conducted on the data to test the null hypothesis that there is no change in par-

participant's responses when measured at the different locations of 41st Street, Bourbon Street and Temple Bar (N=6), for both group frequency and group density taken separately, where the dependant variable is the perceived realism. The results of the group frequency ANOVA indicated no significant effect, Wilks' Lambda = .456,  $F(2, 4) = 2.386$ ,  $p = .208$ . The results of the group density ANOVA similarly indicated no significant effect, Wilks' Lambda = .849,  $F(2, 4) = .356$ ,  $p = .721$ . Thus, there is no significant evidence to reject the null hypothesis.

The average absolute thresholds for group frequency are calculated at 15 groups and 36 groups. The average absolute thresholds for group density are calculated at 1.6 agents per group and 3.8 agents per group. The average optimum configuration for group frequency is calculated at 24 groups and the average optimum configuration for group density is calculated at 2.3 agents per group (Table 2).

Table 2: A summary of the absolute thresholds and optimum configuration for the grouping dynamics behavioural feature.

Grouping Dynamics	Lower Threshold	Upper Threshold	Optimum
Frequency (Bourbon)	14 Groups	36 Groups	20 Groups
Density (Bourbon)	1.6 Per Group	3.9 Per Group	2.4 Per Group
Frequency (41st)	18 Groups	42 Groups	31 Groups
Density (41st)	1.8 Per Group	3.8 Per Group	2.3 Per Group
Frequency (Temple)	12 Groups	31 Groups	22 Groups
Density (Temple)	1.5 Per Group	3.7 Per Group	2.3 Per Group
Frequency (Average)	15 Groups	36 Groups	24 Groups
Density (Average)	1.6 Per Group	3.8 Per Group	2.3 Per Group

### 6.3. Discussion

For discussion of results, we refer back to our research questions:

- RQ1: What are the thresholds for group frequency and group density to be considered perceptually plausible?
- RQ2: What composition of group frequency and group density provides the highest levels of perceived realism?
- RQ3: What effect does location have on the perceived realism of group frequency and group density?

Table 3: A summary of the absolute thresholds and optimum configuration with respect to stimuli intensity.

Grouping Dynamics	Lower Threshold	Upper Threshold	Optimum
Frequency (Bourbon)	35% Intensity	90% Intensity	50% Intensity
Frequency (41st)	41% Intensity	94% Intensity	70% Intensity
Frequency (Temple)	35% Intensity	95% Intensity	70% Intensity
Density (Bourbon)	35% Intensity	81% Intensity	50% Intensity
Density (41st)	39% Intensity	84% Intensity	50% Intensity
Density (Temple)	33% Intensity	81% Intensity	50% Intensity

- RQ4: What are the differences for crowd compositions, if any, between crowd data and perceived realism?

To summarise our findings for the grouping dynamics experiment, in Table 2 we show both the perceptual thresholds for group frequency and density, as well as their identified optimum values (i.e., the range in which these variables are perceived as realistic and the value at which they are perceived to be the most realistic). The corresponding intensity values can also be seen in Table 3 for reference. This has answered RQ1 and RQ2, highlighting that the methodological approach presented in this work can be applied to behavioural features for crowd simulation in order to calculate perceptual thresholds and optimum for achieving perceptual realism with respect to virtual crowds. The overall shape of the psychometric functions shown in Figure 8 presents with a peak in terms of the percentage seen responses that drops off as the stimuli intensity is either increased or decreased away from this optimum level. This can be attributed to the fact that too few groups or too many in the scene evidently cause the perceptual plausibility to fall below acceptable levels. The in turn applies to having too few individuals or too many individuals in each group.

If we consider that the ANOVA results do not allow us to reject the null hypothesis that there is no change in participants responses when measured at the different location, it suggests that for grouping dynamics the environment does not play a significant role. It should be noted that these results are based on the intensity rather than the specific values for the group frequency and density variables at those intensities. Thus, we found for example that 50% intensity at one location gives a similar result to 50% intensity at another, which can map to different values in practice. This further shows the

utility of our methodological approach and the effect of grounding results with the crowd data in the analysis stage. In this case, we specifically looked at three similar heavily pedestrianised urban areas, so the result is limited in that respect, however it further highlights the applicability of the perceptual metrics for different scenarios. Furthermore, it answers RQ3 and provides an interesting insight for further exploration.

In terms of RQ4, the majority of the optimum configurations for both group frequency and group density at the three different locations, correlate with the initial 50% intensity based on reality as seen in Table 3. The only two that varied are the intensities for 41st Street group frequency and Temple Bar group frequency, with viewers favouring an increased intensity of 70% for both. In addition, the upper absolute thresholds possessed a bigger difference from 50% intensity, when compared with the lower thresholds. This suggests, along with the two optimum configurations at 70%, viewers had more perceptual flexibility in terms of intensity when it is increased from the initial 50% rather than decreased. It can also be seen that all of the group frequency upper thresholds are higher than their density counterparts, suggesting more viewer tolerance for additional groups rather than increased group size. Additionally, it is worth noting that the 50% intensity value falls within the absolute thresholds for all of the locations, showing virtual realism was always within the realms of acceptable perceived realism within this experiment. This is an important consideration, as if this was found to not be the case, then it would suggest a disparity between what is actually realistic and what the viewers considers realistic.

The results have been successful in answering the research questions, however there are some limiting factor to address. One in particular, is the limitation of these perceptual metrics for generalisation with other crowd simulations is apparent from cultural implications. When considering human perception, cultural bias is of great concern, not only for this research but many others that aim to create metrics or divine specific solutions to widely applicable problems in different arenas. For example, cultural differences between Western Europe and the Middle East are drastic, so it is likely that perceptual metrics regarding crowds would possess visible differences. Indeed, this backed up by the observation that factors such as typical pedestrian walking speed varies between countries and cultures (Levine and Norenzayan, 1999), meaning in addition to culture, the specific country or location can also play a great role on these types of metrics. In these types of situations, the metrics identified here would not be applicable without a

degree of inaccuracy. However, considerations were put in to place when choosing the behavioural feature and the locations of the crowd footage used to support the perceptual experimentation due to this limitation. In addition, the use of a comparative psychophysical element acted as a mechanism for firmly grounding the experiment and subsequent results (Peters and Ennis, 2009). All crowd footage utilised was based in either the UK or the US, which are heavily urbanised and pedestrianised countries that are often used as a basis for these types of simulations. The US features as the main location in a massive number of video games, with other fictional locations often based upon the architecture and people found therein. As such, while the application of the metrics to different locations and cultures is certainly a limitation, this was known, and care was taken to ensure the perceptual metrics would be as widely useful as possible. In terms of cultural and locational differences for crowd behaviour and its perceptions, it is a potentially rewarding route for future investigation.

An additional limitation is the visual differences between the real-world and virtual environments, which are evident from the comparison between the crowd footage and simulated videos. This is due to our reuse of assets through the Unity prefab system, rather than utilising a bespoke modelling solution for all three locations. Instead of focusing on graphical similarity, we instead focused on ensuring the environmental setup itself was as close to crowd footage as possible within the scope of study. In this way, we sort to mitigate potential differences in group behaviour and movement, rather than the appearance of the virtual environment.

Furthermore, conducting the psychophysical evaluation using an online platform can potentially introduce some bias to the results due to the types of responders and other factors such as being unable to control the setup in terms of the device the participants utilised. This can be seen in the fact that the majority of participations fall within a certain age range and are all considered computer literate. In addition, experimental design considerations had to be made to reduce impact on the results, including shortening video length and reducing the total number of trials. In this way we used design principles and usability to mitigate some of the potential disadvantages of online data collection.

The major advantage of utilising the methodological approach is that it allows for the identification of perceptual thresholds and optimum parameter values that are grounded in real-world data. This has the most potential in particular for game developers and those looking to implement crowd simu-



lation in applications that are user-centered. In this respect, the perceptual metrics, in particular the thresholds, have clear applications within the development environment, especially considering the constraints of real-time crowd simulation. Directly relating to the results of the grouping dynamics experiment, by identifying the lower thresholds for group frequency and density, the number of simultaneous agents within a scene can potentially be reduced to free-up resources while it is ensured that perceived realism is kept within sufficient levels to engage users. This is particularly important when developing for mobile platforms, which have become an important industry medium in recent years.

Additionally, we found that thresholds have more tolerance when increasing intensity rather than decreasing it. That insight has a direct application for developers who may want to tweak crowd composition for a specific scenario in applications where immersion and engagement are key, such as games, serious games, and educational simulations. An example can be seen with the I-Ulysses: Poetry in Motion virtual reality game (Colreavy-Donnelly et al., 2019), which focuses on the preservation of cultural heritage while needing to keep its users immersed within virtual environment to digest the educational content through its core mechanics. The virtual environments in this case are modelled around central Dublin, including the highly trafficked Temple Bar location. The benefit of applying our methodological approach in such an example is two-fold, preservation of authenticity and the flexibility to make changes while remaining within acceptable levels of plausibility. Thresholds when combined with optimum values can support configuration and algorithm selection, which highlights the potential versatility for perceptual metrics in user-centered design and the adaptability of the methodological approach for potential application outside of just crowd simulation.

## 7. Conclusions

We have presented here a set of perceptual metrics for evaluating and optimising simulated crowd behavior as part of a proof-of-concept application of a set of methods that can be used to determine appropriate metrics in other scenarios. The work is supported by experimental results and built upon a new crowd simulation platform created for this research. For the grouping dynamics behavioural feature, the optimum group frequency and group density for virtual crowds are identified for three different locations. Addi-

tionally, perceptual thresholds are calculated showing the limits that both group frequency and group density can be altered by, while still achieving perceptual plausibility. The results provide new insights into the links between reality and human perception, adding to the knowledge of perceptual plausibility in general, and as a consideration for virtual crowd behaviour in other similar lines of research.

The general methodology of analysis, synthesis, and perception, is continued through iteration to further build upon the corpus of perceptual data. As this corpus is built-up, it will allow for the comparison of behavioural features to show which specific configurations can potentially provide the best results within the design constraints present for a specific application. This provides the main avenue for future research with respect to the identification of behavioural features, their synthesis into simulation, and subsequent perceptual evaluation through psychophysics. In this manner, a comprehensive framework to achieve perceptually plausible crowd behaviour focused toward user requirements can be developed over time to include the most utilised features intrinsic to crowd simulation. This is of potential benefit to researchers and developers who can find it a challenging affair to implement effective AI within user focused applications, as it will allow design efforts to be informed by a robust corpus populated through user perception.

A prominent method for future assessment is behavioural annotation, whereby crowd behaviour can be influenced towards certain features within the virtual environment by embedding information therein (Anderson, 2003). This method of annotation for influencing the resultant behaviour is prominent in both academia and industry (Peters and Ennis, 2009; Shao and Terzopoulos, 2005), leading to emergent behaviours that can be quantified as behavioural features and furthermore tied to specific environmental objects or configurations shown within crowd data. This provides an opportunity for the methodological approach presented in this work to build upon behavioural annotations and add an additional dimension for environmental consideration when examining perceived realism for user-centered applications.

The perceptual metrics and insights discovered regarding grouping dynamics is an area for further investigation. Taking into account the limitations we have highlighted regarding other potential configurations for group composition, it is of benefit to examine these variables in order to further support the results of grouping dynamics feature. In particular, the consideration of physical attributes of individuals can have an effect upon group

compositions and resultant behaviour. Physical attributes are a consideration in various algorithms for crowd simulation, such as social forces, and can be pointed to within work on emergent crowd behaviour, such as with the HiDAC system (Pelechano and Badler, 2006; Pelechano et al., 2007). By applying the methodological approach presented here to examine perception in relation to this, insights can be found between these attributes and what constitutes plausible behaviour. These can then be addressed back to identified crowd data in order to highlight differences present and inform future development for both user-focused and serious crowd applications.

- Almeida, J. E., Rosseti, R. J., Coelho, A. L., 2013. Crowd simulation modeling applied to emergency and evacuation simulations using multi-agent systems. arXiv preprint arXiv:1303.4692.
- Anderson, E. F., 2003. Playing smart-artificial intelligence in computer games. In: In: Proceedings of zfxCON03 conference on game development. Bournemouth University, Fern Barrow, Poole, Dorset, BH12 5BB, UK.
- Aschwanden, G., Halatsch, J., Schmitt, G., 2008. Crowd simulation for urban planning. In: Proceedings of eCAADe. Vol. 2008.
- Azahar, M. A. B. M., Sunar, M. S., Daman, D., Bade, A., 2008. Survey on real-time crowds simulation. In: International Conference on Technologies for E-Learning and Digital Entertainment. Springer, pp. 573–580.
- Baird, J. C., Noma, E. J., 1978. Fundamentals of scaling and psychophysics. John Wiley & Sons.
- Banerjee, B., Kraemer, L., 2011. Evaluation and comparison of multi-agent based crowd simulation systems. In: Agents for games and simulations II. Springer, pp. 53–66.
- Başak, A. E., Güdükbay, U., Durupınar, F., 2018. Using real life incidents for creating realistic virtual crowds with data-driven emotion contagion. *Computers & Graphics* 72, 70–81.
- Becker-Asano, C., Ruzzoli, F., Hölscher, C., Nebel, B., 2014. A multi-agent system based on unity 4 for virtual perception and wayfinding. *Transportation Research Procedia* 2, 452–455.

- Beltaief, O., El Hadouaj, S., Ghedira, K., 2011. Multi-agent simulation model of pedestrians crowd based on psychological theories. In: Logistics (LOGISTIQUA), 2011 4th International Conference on. IEEE, pp. 150–156.
- Bernard, S., Therien, J., Malone, C., Beeson, S., Gubman, A., Pardo, R., 2008. Taming the mob: Creating believable crowds in assassins creed. In: Game Developers Conference (San Francisco, CA, Feb 18–22).
- Chalmers, A., Debattista, K., 2009. Level of realism for serious games. In: 2009 Conference in Games and Virtual Worlds for Serious Applications. IEEE, pp. 225–232.
- Charalambous, P., Karamouzas, I., Guy, S. J., Chrysanthou, Y., 2014. A data-driven framework for visual crowd analysis. In: Computer Graphics Forum. Vol. 33. Wiley Online Library, pp. 41–50.
- Cheng, K., Cairns, P. A., 2005. Behaviour, realism and immersion in games. In: CHI'05 extended abstracts on Human factors in computing systems. ACM, pp. 1272–1275.
- Colreavy-Donnelly, S., O'Connor, S., Homapour, E., 2019. I-ulysses: A technical report. Entertainment Computing: Special Issue on Adaptations.
- Cournoyer, F., Fortier, A., 2015. Massive crowd on assassin's creed unity: Ai recycling. In: Game Developers Conference (San Francisco, CA, Mar 2–6).
- Cui, X., Shi, H., 2011. A\*-based pathfinding in modern computer games. International Journal of Computer Science and Network Security 11 (1), 125–130.
- Duives, D. C., Daamen, W., Hoogendoorn, S. P., 2013. State-of-the-art crowd motion simulation models. Transportation research part C: emerging technologies 37, 193–209.
- Durupinar, F., Pelechano, N., Allbeck, J., Gudukbay, U., Badler, N. I., 2011. How the ocean personality model affects the perception of crowds. IEEE Computer Graphics and Applications 31 (3), 22–31.
- EarthCam, I., 1996. Earthcam network. Accessed on: 2017-04-25.  
URL <http://www.earthcam.com/>

- Ehrenstein, W. H., Ehrenstein, A., 1999. Psychophysical methods. In: *Modern techniques in neuroscience research*. Springer, pp. 1211–1241.
- Ennis, C., Peters, C., O’Sullivan, C., 2011. Perceptual effects of scene context and viewpoint for virtual pedestrian crowds. *ACM Transactions on Applied Perception (TAP)* 8 (2), 10.
- Fechner, G., 1966. *Elements of psychophysics*. vol. i.
- Fuchsberger, A., Tahmasbi, N., Ricks, B., 2017. A framework for achieving realism in agent-based pedestrian crowd simulations.
- Gold, J. I., Shadlen, M. N., 2000. Representation of a perceptual decision in developing oculomotor commands. *Nature* 404 (6776), 390–394.
- Grossman, A., 2013. *Postmortems from Game Developer: Insights from the Developers of Unreal Tournament, Black & White, Age of Empire, and Other Top-Selling Games*. Focal Press.
- Guy, S. J., Van Den Berg, J., Liu, W., Lau, R., Lin, M. C., Manocha, D., 2012. A statistical similarity measure for aggregate crowd dynamics. *ACM Transactions on Graphics (TOG)* 31 (6), 190.
- Hajibabai, L., Delavar, M., Malek, M., Frank, A., 2007. Agent-based simulation of spatial cognition and wayfinding in building fire emergency evacuation. In: *Geomatics solutions for disaster management*. Springer, pp. 255–270.
- Heïgeas, L., Luciani, A., Thollot, J., Castagné, N., 2010. A physically-based particle model of emergent crowd behaviors. *arXiv preprint arXiv:1005.4405*.
- Helbing, D., Molnar, P., 1995. Social force model for pedestrian dynamics. *Physical review E* 51 (5), 4282.
- Ijaz, K., Sohail, S., Hashish, S., 2015. A survey of latest approaches for crowd simulation and modeling using hybrid techniques. In: *17th UKSIMAMSS International Conference on Modelling and Simulation*. pp. 111–116.
- Kang, W., Han, Y., 2017. A simple and realistic pedestrian model for crowd simulation and application. *arXiv preprint arXiv:1708.03080*.

- Kapadia, M., Shoulson, A., Boatright, C. D., Huang, P., Durupinar, F., Badler, N. I., 2012. Whats next? the new era of autonomous virtual humans. In: *Motion in Games*. Springer, pp. 170–181.
- Kim, S., Guy, S. J., Manocha, D., Lin, M. C., 2012. Interactive simulation of dynamic crowd behaviors using general adaptation syndrome theory. In: *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. ACM, pp. 55–62.
- Kort, Y. A. d., Ijsselsteijn, W. A., Kooijman, J., Schuurmans, Y., 2003. Virtual laboratories: Comparability of real and virtual environments for environmental psychology. *Presence: Teleoperators & Virtual Environments* 12 (4), 360–373.
- Kuliga, S. F., Thrash, T., Dalton, R. C., Hölscher, C., 2015. Virtual reality as an empirical research tool exploring user experience in a real building and a corresponding virtual model. *Computers, Environment and Urban Systems* 54, 363–375.
- Leggett, R., 2004. Real-time crowd simulation : A review.
- Lemercier, S., Jelic, A., Kulpa, R., Hua, J., Fehrenbach, J., Degond, P., Appert-Rolland, C., Donikian, S., Pettré, J., 2012. Realistic following behaviors for crowd simulation. In: *Computer Graphics Forum*. Vol. 31. Wiley Online Library, pp. 489–498.
- Lerner, A., Chrysanthou, Y., Lischinski, D., 2007. Crowds by example. In: *Computer Graphics Forum*. Vol. 26. Wiley Online Library, pp. 655–664.
- Levine, R. V., Norenzayan, A., 1999. The pace of life in 31 countries. *Journal of cross-cultural psychology* 30 (2), 178–205.
- Loomis, J. M., Blascovich, J. J., Beall, A. C., 1999. Immersive virtual environment technology as a basic research tool in psychology. *Behavior research methods, instruments, & computers* 31 (4), 557–564.
- Luo, L., Zhou, S., Cai, W., Low, M. Y. H., Lees, M., 2009. Toward a generic framework for modeling human behaviors in crowd simulation. In: *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 02*. IEEE Computer Society, pp. 275–278.

- Malinowski, A., Czarnul, P., Czuryło, K., Maciejewski, M., Skowron, P., 2017. Multi-agent large-scale parallel crowd simulation. *Procedia Computer Science* 108, 917–926.
- McDonnell, R., Dobbyn, S., Collins, S., O’Sullivan, C., 2006. Perceptual evaluation of lod clothing for virtual humans. In: *Proceedings of the 2006 ACM SIGGRAPH/Eurographics symposium on Computer animation*. Eurographics Association, pp. 117–126.
- McDonnell, R., Jörg, S., Hodgins, J. K., Newell, F., O’Sullivan, C., 2007a. Virtual shapers & movers: form and motion affect sex perception. In: *Proceedings of the 4th symposium on Applied perception in graphics and visualization*. ACM, pp. 7–10.
- McDonnell, R., Newell, F., O’Sullivan, C., 2007b. Smooth movers: perceptually guided human motion simulation. In: *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*. Eurographics Association, pp. 259–269.
- McMahan, A., 2003. Immersion, engagement and presence. *The video game theory reader* 67, 86.
- Mehran, R., Oyama, A., Shah, M., 2009. Abnormal crowd behavior detection using social force model. In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, pp. 935–942.
- Melo, M., Bessa, M., Debattista, K., Chalmers, A., 2014. Evaluation of hdr video tone mapping for mobile devices. *Signal Processing: Image Communication* 29 (2), 247–256.
- Moussaïd, M., Kapadia, M., Thrash, T., Sumner, R. W., Gross, M., Helbing, D., Hölscher, C., 2016. Crowd behaviour during high-stress evacuations in an immersive virtual environment. *Journal of The Royal Society Interface* 13 (122), 20160414.
- O’Connor, S., Liarokapis, F., Jayne, C., 2015. Perceived realism of crowd behaviour with social forces. In: *Information Visualisation (iV), 2015 19th International Conference on*. IEEE, pp. 494–499.
- O’Connor, S., Liarokapis, F., Peters, C., 2013. A perceptual study into the behaviour of autonomous agents within a virtual urban environment. In:

- World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a. IEEE, pp. 1–6.
- Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., Fuller, T., 2002. User-centered design in games. CRC Press Boca Raton, FL.
- Pelechano, N., Allbeck, J. M., Badler, N. I., 2007. Controlling individual agents in high-density crowd simulation. In: Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation. Eurographics Association, pp. 99–108.
- Pelechano, N., Badler, N. I., 2006. Improving the realism of agent movement for high density crowd simulation. University of Pennsylvania, Center for Human Modeling and Simulation.
- Peters, C., Ennis, C., 2009. Modeling groups of plausible virtual pedestrians. IEEE Computer Graphics and Applications (4), 54–63.
- Peters, C., Ennis, C., McDonnell, R., OSullivan, C., 2008. Crowds in context: Evaluating the perceptual plausibility of pedestrian orientations. Proceedings of Eurographics Short Papers, 227–230.
- Reynolds, C. W., 1987. Flocks, herds and schools: A distributed behavioral model. In: ACM Siggraph Computer Graphics. Vol. 21. ACM, pp. 25–34.
- Reynolds, C. W., 1999. Steering behaviors for autonomous characters. In: Game developers conference. Vol. 1999. pp. 763–782.
- Schrom-Feiertag, H., Stubenschrott, M., Regal, G., Schrammel, J., Settgest, V., 2016. Using cognitive agent-based simulation for the evaluation of indoor wayfinding systems. arXiv preprint arXiv:1611.02459.
- Shao, W., Terzopoulos, D., 2005. Autonomous pedestrians. In: Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation. ACM, pp. 19–28.
- Simpson, W. A., 1988. The method of constant stimuli is efficient. Attention, Perception, & Psychophysics 44 (5), 433–436.
- Sommerseth, H., 2007. ”gamic realism”: Player, perception and action in video game play. In: DiGRA Conference.



- Song, Y., Gong, J., Li, Y., Cui, T., Fang, L., Cao, W., 2013. Crowd evacuation simulation for bioterrorism in micro-spatial environments based on virtual geographic environments. *Safety science* 53, 105–113.
- Szymanczyk, O., Duckett, T., Dickinson, P., 2012. Agent-based crowd simulation in airports using games technology. In: *Transactions on Computational Collective Intelligence VIII*. Springer, pp. 192–213.
- Tashiro, J. S., Dunlap, D., 2007. The impact of realism on learning engagement in educational games. In: *Proceedings of the 2007 conference on Future Play*. ACM, pp. 113–120.
- Thalmann, D., Musse, S. R., 2012. *Crowd Simulation*, 2nd Edition. Springer, Ch. State-of-the-Art, pp. 9–30.
- Ulicny, B., Thalmann, D., 2002. Crowd simulation for virtual heritage. In: *Proc. First International Workshop on 3D Virtual Heritage*. No. VRLAB-CONF-2007-042. pp. 28–32.
- Zhao, M., Cai, W., Turner, S. J., 2018. Clust: Simulating realistic crowd behaviour by mining pattern from crowd videos. In: *Computer Graphics Forum*. Vol. 37. Wiley Online Library, pp. 184–201.
- Zhou, B., 2012. Modeling collective crowd behaviors in video. Ph.D. thesis, Chinese University of Hong Kong.

- The average absolute thresholds for group frequency are calculated at 15 groups and 36 groups.
- The average absolute thresholds for group density are calculated at 1.6 agents per group and 3.8 agents per group.
- The average optimum configuration for group frequency is calculated at 24 groups.
- The average optimum configuration for group density is calculated at 2.3 agents per group.
- There is no change in participant's responses when measured at the different locations of 41st Street, Bourbon Street and Temple Bar, for both group frequency and group density taken separately, where the dependant variable is the perceived realism.