Gamification of Crowd-Driven Environment Design

Brandon Haworth, *Member, IEEE,* Muhammad Usman, Davide Schaumann, Nilay Chakraborty, Glen Berseth, *Member, IEEE,* Petros Faloutsos, and Mubbasir Kapadia, *Member, IEEE*

***Abstract*—This paper describes using human creativity within a gamified collaborative design framework to address the complexity of predictive environment design. This framework is predicated on gamifying crowd objectives and presenting environment design problems as puzzles. A usability study reveals that the framework is considered usable for the task. Participants were asked to configure an environment puzzle to reduce an important crowd metric, the total egress time. The design task was constructed to be straightforward and uses a simplified environment as a probe for understanding the utility of gamification and the performance of collaboration. Single-player and multiplayer designs outperformed both optimization and expert-sourced designs of the same environment and multiplayer designs further outperformed the single-player designs. Single-player and multiplayer iterations followed linear and exponential decrease trends in total egress time respectively. Our experiments provide strong evidence towards an interesting novel approach of crowdsourcing collaborative environment design.**

***Index Terms*—Architectural Design, Crowd Simulation, Gamification, Crowd Sourcing, Co-Design.**

I. INTRODUCTION

Architectural design, from the layout of interior elements to the structural spaces, is a complex problem with a vast solution space that is difficult to navigate. Including crowd-awareness in the design process guarantees that the search for solutions is further complicated. Additionally, scenarios in which design choices impact the people using them or are critical, such as egress and evacuations, are difficult to predictively explore. To address this problem space, synthetic crowds are used as dynamic analytical resources [1], [2]. Synthetic crowds afford low-cost, large scale, and flexible testing of designs, which enables optimization-based approaches that can automatically produce solutions to said design problems. However, such approaches are highly dependent on their fitness functions, may miss or misinterpret useful solutions, and may distance the designer and the end-user from the process. Work in this field has sought user-in-the-loop automation processes to alleviate the shortcomings of solely automated approaches. These approaches typically utilize optimization or machine learning algorithms to generate solutions based on varying modes of user input [2], [3]. However, these approaches do not directly handle multi-user design and do not facilitate co-design or afford crowd-sourcing of designs.

B. Haworth, M. Usman, and P. Faloutsos are with the Department of Electrical Engineering and Computer Science, York University, Toronto, ON, M3J 1P3 Canada.

N. Chakraborty, D. Schaumann, and M. Kapadia are with the Department of Computer Science Rutgers University, Piscataway Township, NJ, 08854 USA

P. Faloutsos is also an Affiliate Scientist with the University Health Network: Toronto Rehabilitation Institute, Toronto, ON, M5G 2A2 Canada.

Manuscript received XXXXX XX, XXXX; revised XXXXX XX, XXXX.

This paper presents an approach to gamify the process of design exploration for crowd-aware environments in a way that opens the process to a broader community. Our approach replaces the designer-as-user in the aforementioned area of research with the community and the designer as co-operative sources of information. The focus of this paper is the gamification of the process of crowdsourced design, but the described system also affords co-design as an inherent feature of this approach.

This game's scoring mechanism is based on measures derived from simulations of crowd movement in the environment. Instead of utilizing automated optimization processes the players compete to produce better designs for a given crowd simulation-based measure. The game also provides feedback in the form of dynamic crowd simulation, agent path traces, and statistical heatmaps. Players are provided tools to edit designs within the constraints defined by the architect or designer–making the core game mechanic a multi-solution puzzle solving exercise. Additionally, players are given the means to parametrize the crowds and explore their dynamics in any given design.

To evaluate this gamified approach, this paper focuses on the usefulness and usability of the system in this domain. Participants are given a straightforward but difficult environment design task as a puzzle, then they are asked to select a puzzle-type (in the form of a crowd configuration) and play in either single or multiplayer modes. In the single player mode, participants work alone to produce solutions to the puzzle that minimize the total egress time for the given puzzle-type while aiming to attain higher positions on the leader board (a public high score list for the puzzle-type). In the multiplayer mode, players collaborate by exchanging ideas in the form of designs and analyses, and compete to attain high positions on the leader board. Results from the single player mode show that gamifying the process is useful in that participants successfully decreased total egress time. Results from the multiplayer mode show that competition and collaboration is even more beneficial in that participants produced higher value designs than single players. Comparative regression analyses on design iterations and total egress times in both modes reveals a linear relationship in the single player mode and an exponential relationship in the multiplayer mode–further solidifying the value of multiplayer play. Finally, the results from a usability survey reveal a high degree of usability even though participants had non-designer backgrounds.

II. RELATED WORK

In this section, we explore the intersection of architectural design, crowd simulation, and the gamification of difficult

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2020.2965069, IEEE Computer Graphics and Applications

problems. In particular, the issue of moving from manual architectural design to automated and user-in-the-loop methods for iteratively producing and evaluating design candidates is the most relevant work in this area. The space of crowd simulation and gamifying complex problems is explored in the context of simulation, problem-solving, and training.

Commercial CAD solutions, such as Autodesk Revit and Rhino3D, are de facto standards in the design of complex 3D structures and environments. More recently, interactive commercial suites for exploring the behaviours of crowds, groups, customers, and end-users in environments have been deployed, such as Pedestrian Dynamics®and MassMotion.

Research efforts have focused on the automatic and user-in-the-loop generation of optimal architectural design solutions with respect to design criteria [4], [5], [6], [7]. Furthermore, an active area of research has been utilizing crowd simulation as a dynamic source for human-centric design criteria and objective analysis [1], [2]. Work has also shown that the user-in-the-loop approach can be made real-time and highly interactive by training neural networks to statically estimate crowd flow in environments [3]. Our work is inherently multi-designer, multi-user, multi-player and directly supports collaboration. Our work uses a real-time dynamic crowd simulation that is highly parametrizable. A user may wish to test a large variety of crowded scenarios, and our system allows for that directly. This includes the ability to find and explore design issues such as bottlenecks, laminar flow, corner bottlenecks, group crossing, vortexes, etc which are not possible in a statically estimated approach.

The usefulness of games, play, and gamification for educational purposes is well-established [8]. Furthermore, game-based collaboration has been successfully deployed to crowd-source solutions to design problems - from protein structures to narratives [9], [10], [11], [12]. Most related to our approach are serious games utilizing crowd simulations in training, planning, and evaluating emergency evacuation plans [13]. These games differ from our approach, in that, they are egocentric (first-person) evacuation training and evaluation games, while ours is an omniscient (knowing of all persons in the scenario) crowd-sourced and collaborative environment design model.

This work is a significant extension of the previously accepted short paper publication [14]. We propose harnessing the power of crowd-sourcing in environment design as a means to address the complexity of predictively designing for dynamic scenarios such as egress. We show the value of gamifying crowd simulation-based objectives and of multiplayer play as a high-value design modality. This is accomplished through a large crowd-sourced user study with hundreds of participants, an extensive comparative regression analyses on single-player and multiplayer performance, discussion of differences in findings between single-player and multiplayer modes, and a usability study showing a high degree of usability for this approach.

III. Proposed Framework

Our framework aims to provide an interactive collaborative platform for architectural and urban design in the form of an online, multiplayer game. By providing the game via the internet, design solutions may be crowdsourced. By utilizing state-of-the-art crowd simulation the solutions may be crowd-aware, or predictive of particular types of crowds or events. The rest of this section describes the proposed framework in detail.

Our framework is predicated on facilitating the following functions: co-design between architects and players (community members, colleagues, or the public); multiplayer collaboration and competition; and the crowdsourcing of environment designs. This approach is driven by two distinct cycles of collaboration and design, as seen in Figure 1.

A high-frequency cycle of collaboration and competition between players generates new high-value designs as seen expanded on the right side of Figure 1. Crowd simulation is utilized in a gamified design tool for environments. Parametrized crowd simulations are a criterion for objective quantitative analysis of scenarios.

A low-frequency cycle of co-design facilitates collaboration between the player community and the designer or stakeholders. This facilitates the uptake of new designs provided by the community, as seen on the left side of Figure 1. This portion of the framework affords the designer or stakeholders the ability to generate problem starting points, new puzzles, or constraints.

The following subsections delineate the modules which make up the framework of the game system.

A. Crowd Simulation

The crowd simulation system utilized is based on a modified version of Optimal Reciprocal Collision Avoidance (ORCA) [15], a successor of reciprocal velocity obstacles. Global navigation is handled using navigation meshes (NavMesh) for the walkable area representation with A* search for global path planning [16]. Navigation meshes are recomputed after every environment change during play. The Unity® Mecanim system, in tandem with a modified version of the ADAPT system [17], handles animation of each bipedal humanoid character model. Additionally, ORCA has been shown to produce generalizable results in crowd-based environment optimization and reproduce "human-like" fundamental diagrams in common egress scenarios [1].

The agents' initial positions in the simulations are distributed either randomly or based on parameters or regions set by the designer. The destinations of the agents are predicated on the scenario set by the designer (such as choosing egress points) or they may be generated randomly. A simulation ends successfully once all agents have reached their goals. If an agent does not complete its goal, the scenario is considered incomplete (the player has not provided a viable solution to the puzzle). Thus, design solutions must be traversable by all agents to be submitted and scored.

Crowd behaviours may be modified by choosing aggression, heterogeneity, and density quality levels. These three traits modify the instantiation parameters of crowd agents–desired velocity/acceleration, size, and density. In particular, these values are important in heterogeneous crowd perception and

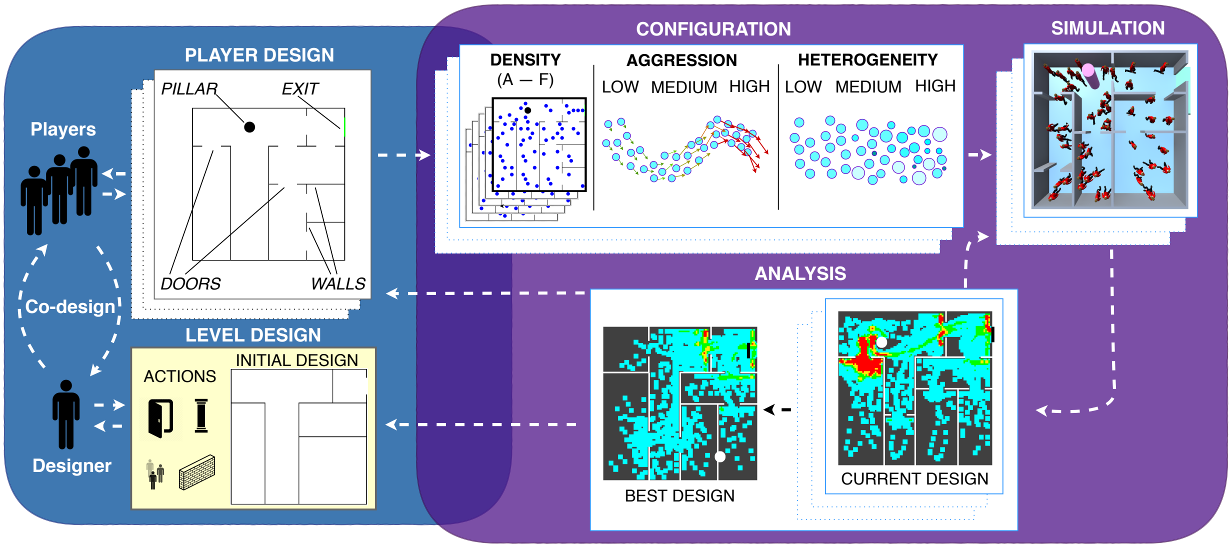


Fig. 1: Framework Overview. To the left, tools facilitate the design and review of environments both on the designer end and the player end. This section facilitates a lower frequency cycle of co-design. To the right, a crowd analysis and animation suite afford the configuration, simulation, and analysis of egress scenarios core to the game's objective. This section expands on the high-frequency cycle the players participate in to produce designs, and highlights the iterations they make after joining a particular puzzle-type..

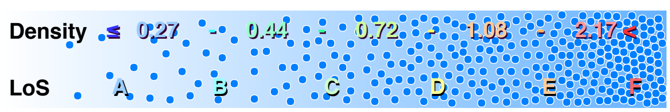


Fig. 2: Level of Service (LoS) categories and their respective crowd density range mappings, where density is measured in agents/m^2. Using LoS grades instead of continuous values, like the levels for homogeneity and aggression, is intended to reduce the gulf of understanding for non-expert players.

outcomes. Three levels of aggression (LoA) are available: low, medium, and high. This quality controls the distribution of speed and acceleration of agents. Three levels of heterogeneity (LoH) are available: low, medium, and high. This quality controls the distribution of the radii, heights, and masses of the agents. This allows the simulation to represent the diversity of body types in a crowd. There are six crowd densities available to the player. LoA and LoH are based on common crowd heterogeneity and aggression proxies across the particle-based synthetic crowds and crowds analysis literature (desired velocity & particle radii) [18]. These are presented in terms of Levels of Service (LoS), since continuous density may not be intuitive to the average player. LoS map a label (A-F) to a range of crowd density capacities in terms of flow rate (0.27 - 2.17 agents/m^2), as seen in Figure 2. Here we invert the original definition of LoS to be a function of density in the crowd rather than the capacity of the environment because we sample the initial conditions over the entire scenario (assuming higher density produces lower quality egress) [19].

The aforementioned categorical qualities of agent behaviour map to underlying parameters for the steering model. For each agent in the crowd, each one of these underlying parameters is sampled from a normal distribution based on the selected categorical level for each quality. That is, each category determines the value of μ and σ for the distribution of each parameter's explicit values amongst the agents. For example, LoA controls the underlying desired speed (v_d) and desired acceleration (a_d) parameter values for each agent. With the categorical setting LoA: Low, the parameters are set at approximate normative human walking values, so $v_d = 1.33m/s \pm 0.5m/s$ and $a_d = 0.68m/s^2 \pm 0.3m/s^2$ sampled for each agent. Together, these categorical levels are LoS: A-F, LoA: Low, Medium, High, and LoH: Low, Medium, High–making for a total of 54 combinations. Each combination is considered a puzzle-type for a given environment in the game, since unnormalized crowd objective metrics are not comparable across densities and low densities often have broader solution spaces. Effectively, each of the possible crowd configurations has a leader board in the game.

B. Game Mechanics

Players (community members, colleagues, or the public) drive an iterative refinement process of a given design problem. This module provides tools for modifying elements of an environment design such as walls, pillars, and doorways, subject to the constraints imposed by the designer. These tools are provided in a puzzle move like format, in which, players may select a particular puzzle move and apply it to the environment. Puzzle moves include add, delete, move, or rotate for pieces such as doors, walls, or pillars.

The primary scoring mechanism of the game is crowd objective metrics. In this paper, total egress time has been used as a proxy for score. However, any computable crowd objective metric is possible, such as flow rate, effort, collisions, path length, etc. These scores are presented amongst players as leader boards (high scores) for different scenarios (crowd parametrizations and environments). After a simulation has completed successfully, the player may additionally review the outcomes using several tools including: the crowd animation itself, agent path traces, and heat maps of aggregate crowd density.

The game is playable in two modes, single and multiplayer. In both modes, a puzzle is a given environment design problem. The players first decide a puzzle-type, or crowd configuration, which relates to the difficulty of the puzzle (for example, it is difficult to find good solutions for higher density scenarios). This puzzle and puzzle-type combination has an associated public leaderboard, or high scores list.

Single Player. In single player mode, a player will iterate on the design by making puzzle moves and running analyses while attempting to improve the design. When their design is submitted the final score is registered and the player is entered into the respective position on the leader board. In the multiplayer mode, the player is provided the ability to share and iteratively revise their own and other players' designs and analyses.

Multiplayer. In the multiplayer mode, a player iterates on the current leading solution for the puzzle-type. When a player submits their design for scoring, the score, design, and analyses are made publicly accessible for that puzzle-type. In this way players are always working on the best current design and may draw from the pool of designs. A player can reinforce their design decision by learning from their analyses, other players' analyses, and other players' designs.

IV. Materials and Methods

To facilitate the evaluation of the usefulness and usability of the proposed gamification of crowd-aware environment design and the game system, participants were asked to solve a common but complex problem - emergency egress. In this way, reducing the score is considered better and players compete (in either single or multiplayer modes) to produce progressively better designs for egress scenarios. Two studies are conducted to explore player performance in the available modes of play: single and multiplayer. Finally, the usability of the system is evaluated using the well-established System Usability Scale (SUS). This research was approved by the University's Research Ethics Board.

The game is provided as an online web-based Unity® application. Participants were recruited and asked to play the game by emailing various mailing lists. The mailing lists included both undergraduate and graduate students. Instructions for tasks and controls were provided in the game menus accessible before and during play. After the participant completed their session, they were asked to complete a survey on system usability and redirected to an online form. Informed consent was provided by all participants and data was collected anonymously and stored securely.

Participants were provided with a basic layout, as seen in the frame "Initial Design" of Figure 1. There were several constraints associated with each element in the layout for this study. First, the static elements of the layout, provided when the game starts, may not be removed or moved - particularly the bounding and structural walls. There were a fixed number of available additional elements such as walls (2), pillars (1), and doorways (6) that must be placed by the participant. Walls could only be placed if they formed an enclosed space. For these experiments, partial dividers were considered invalid, but participants could place a wall and then place a door on that wall. A wall may not be made to pass through an outside, or bounding, wall. All walls had a fixed width of 0.1m and a height of 3m and must be equal to or greater than 2m in length. Finally, all agents in the crowd simulation were required to evacuate the environment - there must be paths to the exit from everywhere in the environment.

LoS	**LoA**	**LoH**	**E**	σ	$\mathbf{E_p}$	$\sigma_\mathbf{p}$
A	Low	Low	9.29	0.12	11.29	1.63
C	Low	Low	10.53	0.39	12.49	0.79
C	High	Low	10.19	0.28	13.00	0.59
D	Low	Low	12.29	0.61	14.50	0.65
E	High	High	13.64	0.73	15.90	0.65
F	Low	Low	15.73	0.63	17.18	0.71

TABLE I: The mean of total egress times E, in seconds, with standard deviation σ in the expert-designed environment, and the mean total egress time $\mathbf{E_p}$ with standard deviation $\sigma_\mathbf{p}$ in the CMA-ES optimized environments, for all crowd configurations reported in the study. Each configuration is simulated 10 times with different initial conditions. Note that there are 54 total possible combinations, these 6 are those which participants chose to play.

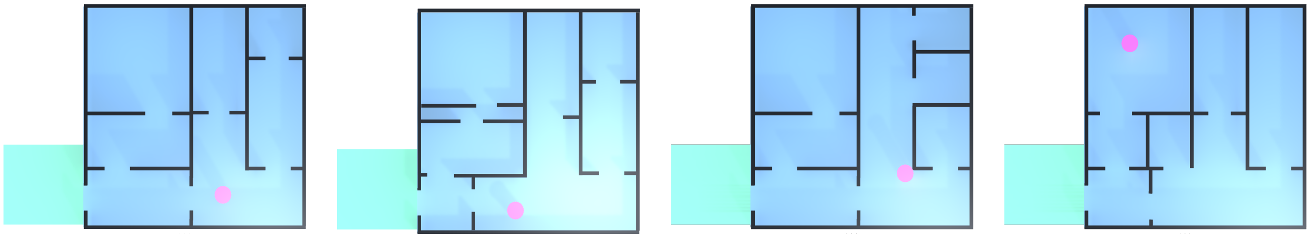
Each participant was identified by a unique ID assigned to them at the start of play. The data gathered included Player ID, architectural element positions, total egress time of the simulation, as well as the participants' choices of crowd configuration parameters (LoS, LoA, LoH), of viewing the heat map, and of viewing the Top Scorer's heat map.

A human design baseline, within the constrained parameters for the required elements in the experimental environment, was generated by an expert designer comfortable with the system, as in Figure 1 *Initial Design*. Additionally, automated optimization baselines were generated using Covariance Matrix Adaptation-Evolutionary Strategy (CMA-ES) on the parameter space of the puzzle [20]. For each crowd configuration available in the study, the CMA-ES optimal designs and the expert design were simulated ten times with randomized initial agent configurations. Table I shows the mean and standard deviation of total egress times for the baselines. Examples of these environments can be seen in Figure 3.

V. Single Player Study

This study hypothesizes that gamification of such a difficult problem in environment design provides a non-expert player with a means to improve their design and total egress time. The total egress times of iterations completed by the players were stored for analysis. This data serves as a comparison to multiplayer and expert performance.

In the single-player mode, participants work on the puzzle, or environment, under the given constraints for the puzzle-type, or crowd parametrization, they choose at the beginning of the session. In this mode, a participant may view their



(a) (b) (c) (d)

Fig. 3: Examples of environment layout designs from each of the experimental design source (a) the expert design (b) the CMA-ES design for the LoS: E, LoA: High, LoH: High, scenario (c) a single player design for the LoS: A, LoA: Low, LoH: Low scenario and (d) a multiplayer design for the LoS: C, LoA: Low, LoH: Low scenario. Note how (a-c) are somewhat similar, likely an artefact of exploration by a singular source, and (b) has a non-useful room structure, an artefact of completely automated design.

analyses, as described in Section III-B, between each design iteration. When the participant is satisfied, they may submit their final design for scoring and placement on the public leaderboards. This mode represents a typical single-player mode common in many puzzle-like games, wherein players work on solutions alone.

A. Analysis

A Kruskal-Wallis one-way analysis of variance was conducted to compare the total egress times of the CMA-ES optimal, expert scenario, and the final single-player authored designs in each crowd configuration challenge. This was followed by a Conover's pairwise multiple comparison test with both Bonferoni-Holm and False Discovery Rate (FDR) corrections for ties. Comparative regression analysis helps to further understand how participants performed when using the game. The ordered set of iterations (independent variable) and egress time (dependent variable) is regressed for three models: linear ($y = ax + b$), exponential ($y = ae^{bx}$), and logarithmic ($y = a\ln(x) + b$). The slope, or curvature, sign (increasing[+], decreasing [-]) parameter a (b for the exponential model) is of particular interest. This parameter controls the rate and sign of improvement over iterations (large values are faster rates and negative values are decreasing total egress time).

B. Results

The summary statistics overall single-player iterations and over final designs are reported in Table II. A box plot comparison of the median, IQR, and extrema of the CMAE-ES optimal, expert design, and final multiplayer design mean total egress times are shown in Figure 4. There was a significant difference in the total egress times amongst the three conditions for each crowd parametrization ($p < 0.001$). The post-hoc tests revealed that for both crowd parametrizations the expert and CMAE-ES optimal design was not significantly different, while the single-player final designs are significantly different from both other conditions ($p < 0.01$). The distributions of relevant parameters, Pearson r^2 values, and Mean Squared Error (MSE) values, for all three models in the regression analysis, are shown in Figure 5.

LoS	LoA	LoH	N	I	$\mathbf{E^t}$	$\sigma^\mathbf{t}$	$\mathbf{E^f}$	$\sigma^\mathbf{f}$
A	Low	Low	60	330	8.51	1.47	7.20	1.10
C	Low	Low	60	297	10.66	1.10	9.55	0.92

TABLE II: Summary statistics for the overall data in the single player gamification dataset by crowd configuration. N is the total number of participants for a given configuration; I is the total number of iterations this group made; E^t and σ^t are mean of total egress times, in seconds, and standard deviation of the total set; E^f and σ^f are the mean of total egress times, in seconds, and standard deviation of the final designs for each player.

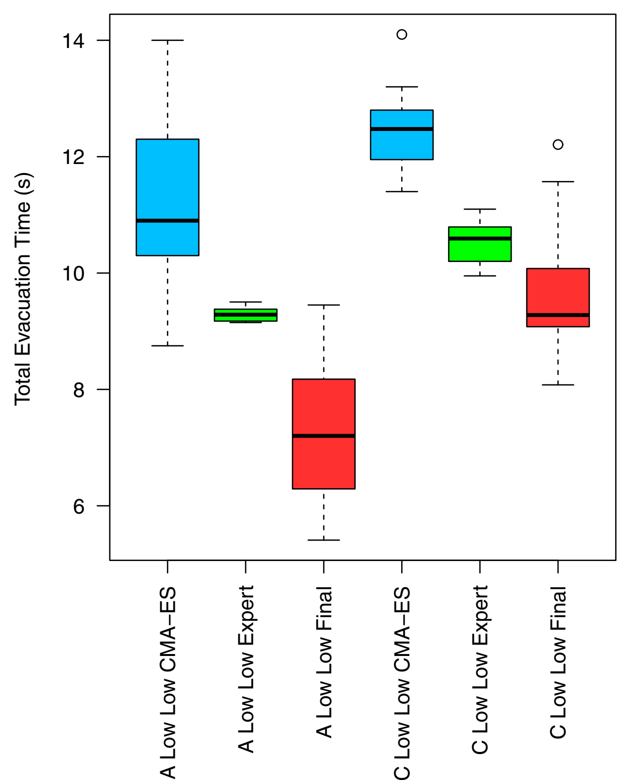


Fig. 4: Box plot comparison of CMA-ES optimal designs versus expert versus final mean total egress times for the single player mode, showing the median, IQR, and extrema. The player created designs show greater variance, but the total egress time is reduced significantly.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2020.2965069, IEEE Computer Graphics and Applications

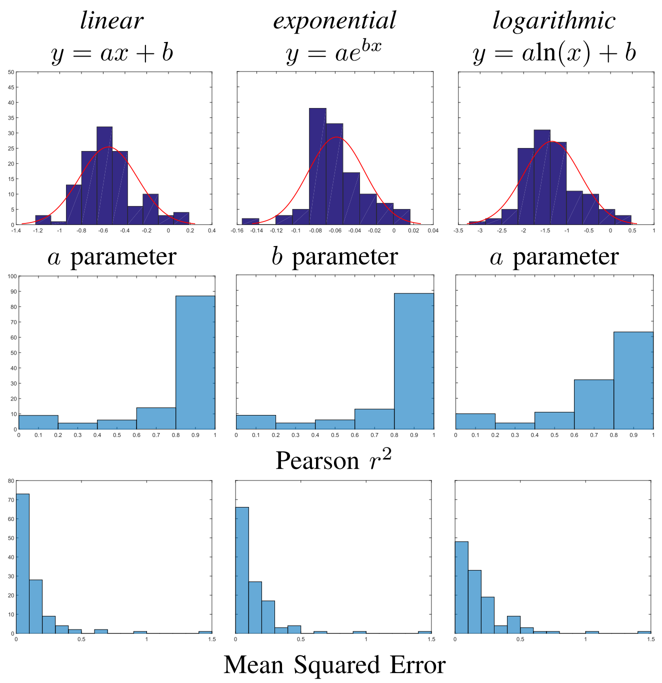


Fig. 5: Regression results over all single player design sessions. Each participant's ordered set of iterations (independent variable) and total egress times (dependent variable) is fitted for each model. The complete set of results is shown as Top row: the distribution of fitted parameters, b for the exponential model, a otherwise; Middle row: the distribution of Pearson r^2 values; and Bottom row: the distribution of Mean Squared Error (MSE) values. The results favour the linear model in terms of lower MSE with r^2 values similar to the exponential model.

C. Discussion

Regression analysis shows that individual participants follow either an exponential or linear improvement in their efforts to reduce total egress time. The distribution of Pearson r^2 values favours an exponential curve. However, while the MSE values show a normal distribution about zero for both linear and exponential models, the distribution of MSE favours the linear model. Most importantly, all regression models show a strong negative slope. This highlights the fact that all participants are decreasing the egress time in their design iterations.

VI. Multiplayer Study

This study hypothesizes that the competitive and collaborative processes found in multiplayer modes [12] afford a more comprehensive search of the solution space resulting in even shorter total egress times (in comparison to both the expert designs and single player designs). The total egress times of crowd configurations for the groups of multiplayer participants were stored for analysis.

In the multiplayer mode, participants work on the puzzle, or environment, under the given constraints for the puzzle-type, or crowd parametrization, they choose at the beginning of their session. That is, each puzzle-type has a group of fellow players all collaborating and competing for the top position on that puzzle-type leaderboard. In this mode, a participant may view both their analyses and those of the current leaders for the particular puzzle-type, as described in Section III-B, between each design iteration. The participants have access to both the leading analyses and designs and may use the current leading design as a starting point for their own. In this way, the multiplayer mode is both collaborative, players work on designs together, and competitive, players are explicitly attempting to submit the highest value designs. When participants are satisfied, they may submit their designs and analyses for scoring, distribution, and placement on the public leaderboards.

LoS	LoA	LoH	N	I	$\mathbf{E^t}$	$\sigma^\mathbf{t}$	$\mathbf{E^f}$	$\mathbf{E_{min}}$
A	Low	Low	9	115	7.17	2.14	5.72	4.02
C	High	Low	7	107	7.88	1.85	6.52	5.06
D	Low	Low	7	107	9.79	2.37	6.84	6.02
E	High	High	7	107	12.68	2.98	9.90	8.50
F	Low	Low	6	99	14.04	3.09	12.04	10.09

TABLE III: Summary statistics for the overall data in the multiplayer gamification dataset by crowd configuration. N is the total number of participants for a given configuration; I is the total number of iterations this group made; E^t and σ^t are mean total egress time, in seconds, and standard deviation for the total set; E^f is the total egress time, in seconds, of the final design; and E_{min} is the minimum, or fastest, total egress time, in seconds.

A. Analysis

The multiplayer mode culminates in a single final design produced by the players for their puzzle-type. The best performing expert scenario and the final multiplayer design total egress times for each puzzle-type are compared. To further understand how participants performed when using the game, a comparative regression analysis study is carried out on the multiplayer participant design iterations for each puzzle-type. This parameter controls the rate and sign of improvement over iterations (large values are faster rates and negative values are decreasing total egress time).

B. Results

A scatter plot comparison of the CMA-ES optimal, expert, and final mean total egress times are shown in Figure 6. The summary statistics over all multiplayer iterations and final designs are reported in Table III. The results for the comparative regression analysis are reported in Table IV. Examples of interesting regression results from two multiplayer design challenges are shown in Figure 7.

C. Discussion

Regression analysis shows that participants iteratively decrease the egress time when collaboratively modifying their environment configurations. The best-fitting, in terms of high Pearson r^2 and low MSE, are the exponential model across all crowd configurations, except for the configuration LoS:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2020.2965069, IEEE Computer Graphics and Applications

Crowd Configurations			*Linear Regression Summary*			*Exponential Regression Summary*			*Logarithmic Regression Summary*		
LoS	**LoA**	**LoH**	**a**	$\mathbf{r^2}$	**MSE**	**b**	$\mathbf{r^2}$	**MSE**	**a**	$\mathbf{r^2}$	**MSE**
A	Low	Low	-0.059	0.86	0.63	-0.0085	0.88	0.55	-1.93	0.71	1.30
C	High	Low	-0.056	0.87	0.45	-0.0071	0.88	0.42	-1.72	0.75	0.85
D	Low	Low	-0.072	0.88	0.65	-0.0073	0.88	0.67	-2.08	0.67	1.85
E	High	High	-0.088	0.84	1.38	-0.0071	0.87	1.19	-2.72	0.72	2.44
F	Low	Low	-0.093	0.74	2.48	-0.0070	0.79	2.02	-3.06	0.84	1.52

TABLE IV: Results of the regression analysis for the collaborative data set for all linear, exponential, and logarithmic regressions: the parameters, b for the exponential model, a otherwise; the Pearson r^2 values; and the Mean Squared Error (MSE) are reported.

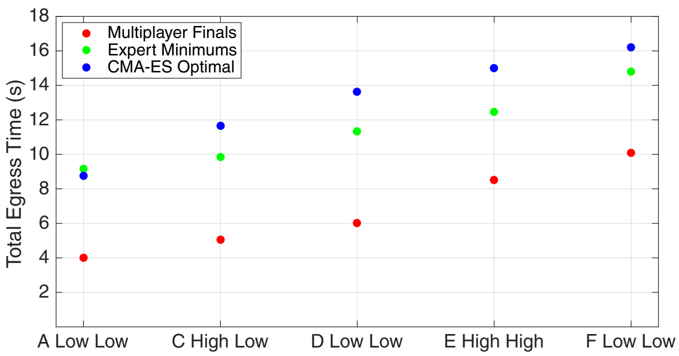
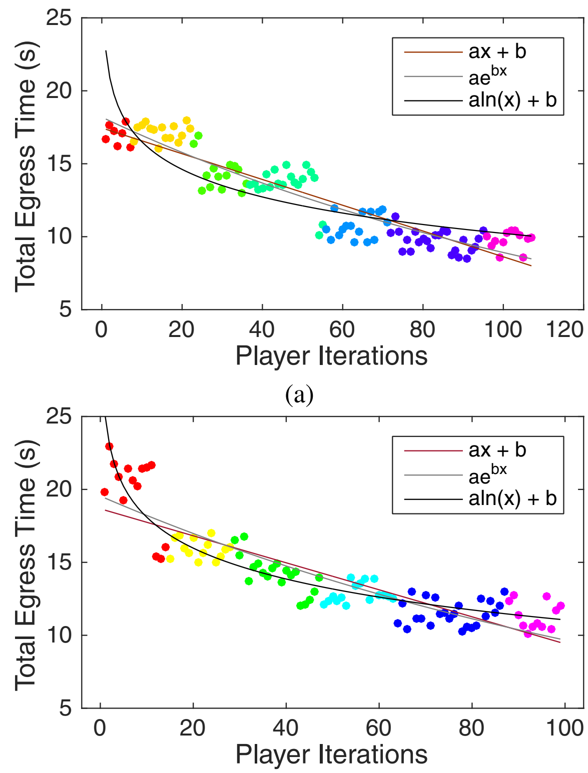


Fig. 6: Comparison of expert and CMA-ES optimal design minimum total egress time versus multiplayer final total egress times for the multiplayer mode. These data are for the shortest expert and optimal design total egress time out of ten simulations and the final collaborative design produced by the multiplayer participants for each of the crowd configurations.

Count	Mean	Median	Min	Max	σ	*IQR*
61	78.69	80	70	92.5	8.84	7.5

TABLE V: Summary statistics for SUS results, where the score range is from 0 to 100.



(b)

Fig. 7: Example regression analysis results (lines & curves) for all functions over all design iterations for the a) LoS: E, LoA: High, LoH: High and b) LoS: F, LoA: Low, LoH: Low; multiplayer mode design iterations. Each data point is colour coded by player for that particular puzzle-type. These two puzzle types exemplify the extremes of the experiment outcomes, a) shows an exemplary case where the iterations plateau in an exponential fit, while b) shows the only crowd configuration where the logarithmic fit is best because of the difficulty in initially searching the solution space of high-density scenarios.

F, LoA: Low, LoH: Low as shown in Figure 7(b). While all design collaborations plateau near some solution, (b) shows the most difficult challenge (the highest density egress) has a high starting point and the plateau is longer and flatter. The exponential regression captures the rapid improvement of the scenario over iterations and the plateau near the end of the collaborations as the players' near an optimal layout–at which point the scenario becomes difficult to improve significantly.

VII. Usability Study

Participants were asked to complete a System Usability Scale (SUS) survey following their play sessions. In total, 61 players provided feedback in an online SUS survey. For ease of use, SUS scores are scaled to a range between 0 and 100, with scores above 68 considered "above average" and "acceptable". Our average usability score is 78.69 out of a scale of 100. The summary statistics of the SUS scores are reported in Table V.

A mapping of scaled SUS scores to common adjectives provides an intuitive interpretation for each score range. The results show that the 61 participants mean and median scores fall within the adjective range of "good" and "excellent", which means that our game system is highly usable and "learnable" with some degree of confidence.

VIII. Conclusion

This paper describes an online game that re-imagines environment design problems as puzzle-solving games. To understand the effectiveness of gamification in this domain, comparative regression analyses show that single player and multiplayer modes produce high-value environment designs. The single player authored designs exceed the performance of the expert-designed environment, and the multiplayer designs exceed single player and optimization strategy performance. There is evidence that single players linearly improve total egress times over design iterations. Furthermore, collaborating players exponentially improve total egress times over design iterations.

The game system utilizes an iterative approach to optimizing the environment layout. Players drive this process with their design actions and analyses. Players become contributors to solutions for difficult problems. The results show that no matter what mode the players play in (single or multiplayer), they produce increasingly better designs.

Limitations The collaborative gaming approach is a technically greedy approach to optimization that may lead to local minima. This is highly dependent on the collaboration dynamics and the subjective value players give to solutions. The limitations in our studies included a free choice of puzzle types, or crowd parametrization, by players. Several of the 54 possible crowd configurations, or puzzle-types, were not selected by the players–limiting our analyses to those populated with data. However, this study provided ample evidence and motivation to pursue this line of research. Additionally, the study utilizes a simplified environment to reduce confounds and focus on the gamification of the process. This keeps the study and subsequent analysis tractable while easing the learning curve for participants who are not experts in this particular design space. Even seemingly simple crowd-driven design is a complex and non-convex solution space of an ill-conditioned problem, as small changes in parameters may drastically change outcomes.

Future Work It is hypothesized that the presented findings extend to more complex design problems. Future work will look into the usefulness of the approach in more difficult design problems. There are several relevant aspects of the player facing framework to explore. Extending the range of crowd parametrizations would be beneficial in an industry setting, but maybe less intuitive to non-expert gamers. We plan to explore the correct way to deliver both useful and usable tools to players. Additionally, the underlying steering model and its parametrizations are a largely simplified particle model of human movement and normative walking values respectively. Using this in a design platform assumes environments will be used by a specific subset of the population. This was useful for a proof-of-concept, but future work will address normativity as default in environment design especially where design interventions and accessibility necessarily intersect. The platform supports additional affordances which are left to future work to explore. Primarily, co-design, in that design stakeholders (architects, planners, government, institutions, etc.) may access the processes, platforms, and people of their community, potentially facilitating radically new ways of working, collaborating, or learning in the architectural domain. Finally, we plan for the exploration of the platform as an educational or training tool in the space of architectural design.

ACKNOWLEDGMENTS

This is research has been partly funded by ORF/ISSUM, NSF IIS-1703883, NSF S&AS-1723869, NSERC Discovery and Create Grants, and the Murray Postdoctoral Fellowship.

REFERENCES

[1] B. Haworth, M. Usman, G. Berseth, M. Kapadia, and P. Faloutsos, "On density–flow relationships during crowd evacuation," *Computer Animation and Virtual Worlds*, vol. 28, no. 3-4, 2017.

[2] B. Haworth, M. Usman, G. Berseth, M. Khayatkhoei, M. Kapadia, and P. Faloutsos, "Code: Crowd-optimized design of environments," *Computer Animation and Virtual Worlds*, vol. 28, no. 6, p. e1749, e1749 cav.1749.

[3] T. Feng, L.-F. Yu, S.-K. Yeung, K. Yin, and K. Zhou, "Crowd-driven mid-scale layout design." *ACM Trans. Graph.*, vol. 35, no. 4, pp. 132–1, 2016.

[4] H. Pottmann, M. Eigensatz, A. Vaxman, and J. Wallner, "Architectural geometry," *Computers and Graphics (Pergamon)*, vol. 47, pp. 145–164, 2015.

[5] X. Shi and W. Yang, "Performance-driven architectural design and optimization technique from a perspective of architects," *Automation in Construction*, vol. 32, pp. 125–135, 2013.

[6] M. Turrin, P. Von Buelow, and R. Stouffs, "Design explorations of performance driven geometry in architectural design using parametric modeling and genetic algorithms," *Advanced Engineering Informatics*, vol. 25, no. 4, pp. 656–675, 2011.

[7] J. Michalek and P. Papalambros, "Interactive design optimization of architectural layouts," *Engineering Optimization*, vol. 34, no. 5, pp. 485–501, 2002.

[8] D. R. Tobergte and S. Curtis, "A systematic literature review of empirical evidence on computer games and serious games," *Journal of Chemical Information and Modeling*, vol. 53, no. 9, pp. 1689–1699, 2013.

[9] D. Debkowski, A. Marrero, N. Yson, L. Yin, Y. Yue, S. Frey, and M. Kapadia, "Contained: Using Multiplayer Online Games to Quantify Success of Collaborative Group Behavior," in *Twelfth Artificial Intelligence and Interactive Digital Entertainment Conference*, 2016.

[10] S. Mueller, B. Solenthaler, M. Kapadia, S. Frey, S. Klingler, R. P. Mann, R. W. Sumner, and M. Gross, "Heapcraft: interactive data exploration and visualization tools for understanding and influencing player behavior in minecraft," in *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games*. ACM, 2015, pp. 237–241.

[11] B. Li, S. Lee-Urban, and M. O. Riedl, "Toward Autonomous Crowd-Powered Creation of Interactive Narratives," in *Intelligent Narrative Techologies V (INT5)*, vol. 8, 2012, pp. 20–25.

[12] S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović *et al.*, "Predicting protein structures with a multiplayer online game," *Nature*, vol. 466, no. 7307, pp. 756–760, 2010.

[13] J. F. Silva, J. E. Almeida, R. J. Rossetti, and A. L. Coelho, "A serious game for evacuation training," in *Serious Games and Applications for Health (SeGAH), 2013 IEEE 2nd International Conference on*. IEEE, 2013, pp. 1–6.

[14] N. Chakraborty, G. Berseth, B. Haworth, P. Faloutsos, M. Usman, and M. Kapadia, "Crowd sourced co-design of floor plans using simulation guided games," in *Proceedings of the Tenth International Conference on Motion in Games*. ACM, 2017, p. 1.

[15] J. Van Den Berg, S. J. Guy, M. Lin, and D. Manocha, "Reciprocal n-body collision avoidance," in *Springer Tracts in Advanced Robotics*, 2011, vol. 70, no. STAR, pp. 3–19.

[16] M. Kallmann and M. Kapadia, *Geometric and discrete path planning for interactive virtual worlds*. Morgan & Claypool Publishers, 2016.

[17] A. Shoulson, N. Marshak, M. Kapadia, and N. I. Badler, "Adapt: the agent development and prototyping testbed," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 7, pp. 1035–1047, 2014.

[18] N. Pelechano, J. M. Allbeck, M. Kapadia, and N. I. Badler, *Simulating heterogeneous crowds with interactive behaviors*. CRC Press, 2016.

[19] J. J. Fruin, *Pedestrian planning and design*. New York: Metropolitan Association of Urban Designers and Environmental Planners, 1971.

[20] G. Berseth, M. Usman, B. Haworth, M. Kapadia, and P. Faloutsos, "Environment optimization for crowd evacuation," *Computer Animation and Virtual Worlds*, vol. 26, no. 3-4, pp. 377–386, 2015.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2020.2965069, IEEE Computer Graphics and Applications

Brandon Haworth Brandon Haworth is a Post-Doctoral Fellow in the Department of Electrical Engineering and Computer Science at York University. He received the BSc degree in Computer Science from York University in 2013, MSc in 2016, and PhD in 2019. Brandon has also worked as a researcher and consultant with the University of Toronto, the University Health Network: Toronto Rehabilitation Institute, Teeple Architects, and other academic and industry partners. His research interests include crowd steering behaviours, artificial intelligence, digital game design, architectural optimization, assistive technologies, and rehabilitative technologies with a purpose to explore the intersections between visibility, representation, and decision-making in interactive technologies.

Muhammad Usman Muhammad Usman is a PhD candidate in the Department of Electrical Engineering and Computer Science at York University. He received his MSc degree in Computer Science from York University in 2016 and his BSc degree in Computer Science from University of the Punjab in 2013. His research interests include crowd steering behaviors, computer-aided architectural design, architectural design analysis and optimization, spatial visualizations, virtual reality, and assistive technologies.

Davide Schaumann Davide Schaumann is a Runway Startup Postdoc at the Jacobs Technion-Cornell Institute at Cornell Tech. His research focuses on developing computational tools to optimize the design and operations of complex buildings with a mission to create efficient, sustainable and human-centered environments. Schaumann previously held a post-doctoral position in the Department of Computer Science at Rutgers University. He received his Ph.D. in Architecture from the Technion - Israel Institute of Technology and he holds B.A. and MSc degrees in Architecture from the Politecnico di Milano.

Nilay Chakraborty Nilay Chakraborty is a Full-Stack Software Developer at WGACA LLC, NY. He received his M.S. degree in 2017 in computer sciences from Rutgers University, NJ. During His years at Rutgers University, he completed his Masters thesis on Multiplayer Game Frameworks for Crowd-Aware Co-design of Environments under the direction of Dr. Mubbasir Kapadia. He also worked at SIG Center for Computer Graphics, University of Pennsylvania under the supervision of Dr. Norman I. Badler on a project named Narrative Recreation using Agent's Autobiographical memory. His research interests are in crowd modeling, architectural design, and games.

Glen Berseth Glen Berseth is a PostDoc at UC Berkeley with Sergey Levine. He received his PhD at the Department of Computer Science at the University of British Columbia in 2019 where he worked on reinforcement learning, machine learning and motion planning with Michiel van de Panne. He received his BSc degree in Computer Science from York University in 2012 and his MSc from York University under the supervision of Petros Faloutsos in 2014 for optimization and authoring crowd simulations. He has published in a wide range of areas including computer animation, machine learning and robotics and is an NSERC scholarship award winner.

The focus of his research is applications of machine learning and optimization in motion planning for tasks with complex dynamics. His research interests include reinforcement learning, machine learning, character animation, robotics, crowd simulation and cognitive agents.

Petros Faloutsos Petros Faloutsos is a Professor at the Department of Electrical Engineering and Computer Science at York University, and an affiliate Scientist at the UHN-Toronto Rehabilitation Institute. Before joining York, he was a faculty member at the Computer Science Department at the University of California at Los Angeles, where in 2002 he founded the first computer graphics lab at UCLA. Faloutsos received his PhD degree (2002) and his MSc degree in Computer Science from the University of Toronto, Canada and his BEng degree in Electrical Engineering from the National Technical University of Athens, Greece.

Mubbasir Kapadia Mubbasir Kapadia is the Director of the Intelligent Visual Interfaces Lab and an Assistant Professor in the Computer Science Department at Rutgers University. Previously, he was an Associate Research Scientist at Disney Research Zurich. Kapadia's research lies at the intersection of artificial intelligence, visual computing, and human-computer interaction, with a mission to develop intelligent visual interfaces to empower content creation for human-aware architectural design, digital storytelling, and serious games.