# Real-time experiential geosimulation in virtual reality with immersion-emission

Paul M. Torrens<sup>†</sup>
Department of Computer Science and Engineering
New York University
New York, NY, USA
torrens@geosimulation.com

Simin Gu Integrated Design & Media New York University New York, NY, USA sg5292@nyu.edu



Figure 1: User participation in the VR experiment. The background projection shows the streetscape and user's view.

# **ABSTRACT**

The aim of this work is to develop closer connectivity between real, lived human encounters in everyday life and geosimulation that might be tasked and designed to experiment with synthetic variations of those experiences. In particular, we propose that geosimulation can be used in close connection with virtual geographic environments and virtual reality environments to build human-in-the-loop interactivity between real people and geosimulation of the geographies that they experience. We introduce a novel scheme based on immersion and emission by socio-visual gaze to facilitate connectivity between human users and geosimulation. To examine the utility of the approach, we present a worked demonstration for examining road-crossing behavior in downtown settings for Brooklyn, NY.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs International 4.0 License. GeoSim'21, November 2, 2021, Beijing, China © 2021 Copyright held by the owner/author(s). ACM ISBN 978-1-4503-9101-6/21/11. https://doi.org/10.1145/3486184.3491079

# **CCS CONCEPTS**

Computing methodologies ~ Modeling and simulation ~
 Simulation types and techniques ~ Agent / discrete models

# KEYWORDS

Geosimulation, computer-human interaction, virtual reality, human behavior, socio-visual gaze, human-in-the-loop

#### **ACM Reference format:**

Paul M. Torrens and Simin Gu. 2021. Real-time experiential geosimulation in virtual reality with immersion-emission. In 4th ACM SIGSPATIAL International Workshop on GeoSpatial Simulation (GeoSim'21), November 2, 2021, Beijing, China, 10 pages. https://doi.org/10.1145/3486184.3491079

#### 1 Introduction

Geosimulation has long been considered beneficial for driving the dynamic behavioral interactions of agents and other automata in computer models [1]. A usual instantiation of this approach involves deploying geosimulation atop theoretical models of human behavior, using the geosimulation as a vehicle for experimentation and knowledge discovery [2, 3]. In these cases,

geosimulation is matched in different ways to human behavior. These include using geosimulation to test theories of how behavior unfolds within systems context (particularly for complex adaptive systems with geographic scaling) [4, 5]; representation of the geographical agency of particular phenomena [6]; and building synthetic human process dynamics in support of studying other system factors [7]. Geosimulation is also employed to extend models for uses beyond existing data records, e.g., to animate static data through simulated time by forecast or hindcast, to scale-up the representation of a few agentrepresentatives to large populations of modeled people, or even to use geosimulation to produce synthetic data [8]. In this context, geosimulation is particularly useful for working in applications with sparse behavioral data [9], which is often illustrative of most situations in urban applications, where behavior data is routinely difficult to acquire empirically. Recent developments in expert-inthe-loop type geosimulation [10] and automated intervention in user-based geosimulations [11, 12] are particularly useful in these instances, as a scheme to encode available knowledge bases and on-the-fly user input into the geosimulation. Another usual proofof-usefulness for geosimulation in studies of human behavior is to use the framework to build representations of scenarios that would normally sit outside the reach of academic inquiry, e.g., for hazardous situations, for scenarios that sit far into the future as mere possibilities, or for normative planning. In these cases, one must often speculate about potential behaviors, because the applications being considered are rare, novel, or have yet to occur.

A conundrum, particularly in the latter two use cases of geosimulation (where behavior data are difficult to acquire, or for behavioral scenarios that are rare in occurrence) is that the geosimulation is reliant upon the fidelity of the underlying model of human behavior, which itself may be significantly opaque in understanding [2]. Chasing the "fidelity gap" between the realworld, the model, and the simulation involves significant work to build and calibrate automata states and rule-sets on the front-end of experiments, and then further subsequent effort to validate and verify the simulations on the back-end of the experiment. These gaps can be "massaged" into workable geosimulation applications, using mathematical adjustments such as equation-free [13], for example. But, for many facets of human life, the underlying human behavior may not even be knowable and the purpose of the simulation is to generate candidate threads for building that knowledge. This may place the geosimulation on unsteady footing, typically at construction. Consequently, traditional geosimulations proceed from automata that have been built from (usually highly) abstract conceptual models [14]. In essence, the model does claim much fidelity relative to individual behavioral agency, and instead focuses on phenomena at coarse scale [15]. such that didactic aims of the geosimulation are at arms-length from many real-world scenarios they support [16, 17].

In this paper, we examine how virtual reality (VR) environments for geosimulation of human behavior might be used atop virtual geographic environments (VGEs) [18-20] to build real-time individual interactivity between model users and agent-automata for urban scenarios in which (1) data are generally sparse because

of difficulties in observation, and (2) the scenarios involved are rare and novel due to low likelihood of occurrence relative to the broader dynamics in which they sit. In essence, we aim to provide an extensible scheme for introducing real human-in-the-loop connections to geosimulation. To accomplish this aim, we will show how immersive-emissive functions can be used within VR/VGE environments for geosimulation of busy urban scenes to facilitate real-time experiential connections between real-world model users and synthetic agent representations of road-crossing behavior among pedestrians, vehicles, and signal systems. We conclude by outlining how VR/VGE can be used as an environment for immersive geosimulation to support naturalistic forms of user interaction with simulated scenes of human behavior in everyday urban contexts. We believe that the approach we have outlined can be helpful in moving geosimulation closer to the reality of everyday lived experiences, thereby expanding the experimental reach of geosimulation, particularly for human behavior research.

# 2 Methods

To accomplish our objectives of building a human-in-the-loop framework for geosimulation, we developed a multi-agent geosimulation to characterize an everyday urban scene involving road-crossing behavior at a busy downtown intersection. The geosimulation follows existing approaches for pedestrian modeling in urban environments, but we enabled one significant pathway for the direct involvement of real human users in the model: we endowed the agents with interactive personality, which we achieved using socio-visual gaze. We allow real human users to participate directly in the geosimulation through immersion. This was accomplished by developing the geosimulation atop a realistic-appearing VGE representation of a real intersection in Brooklyn, NY. By physically walking around a real space, and looking within a virtual space, human users directly control their own ego-agent character in the geosimulation. From the perspectives of other agents in the simulation, the ego-agent is no different than any other agent, so model agents may interact with the human participant with parity of spatial interaction. However, the human user is immersed directly in the geosimulation through VR, allowing them to participate naturally with the simulation events as they unfold.



Figure 2: 13 different types of agent-avatars were used in simulation, with varying characteristics.

We accounted for four classes of agents in the model: pedestrianagents, cars, signals, and the ego-agent. Pedestrian-agents were used to represent synthetic characters in the model, which form the basis for building dynamic human activity in the simulated scenes. We provide for 13 different pedestrian variations in the model, with diversity in demographics (sex, age, race), physical attributes (height, body shape), and appearance (hairstyle, clothing). The avatar-representations of the agents in simulation were designed to reflect these characteristics (Figure 2). Agents were also assigned a random field of view ranging between 140 and 190 degrees, within which they had (visual) access to the objects within the scene graph for that portion of the VGE (Figure 3). *Ego-agents* were used to instantiate the simulation user within the VGE. Pedestrian-agents and vehicle agents act and interact with ego-agents as if they were pedestrian-agents. However, the decision-making, personality, path-planning, way-finding, steering, locomotion, attention, gaze, collision detection, collision avoidance, signal adherence, and road-crossing behavior of the ego-agent was driven directly by the human user through immersion in the VRE.  ${\it Car-agents}$  were used to model individual vehicles for road-crossing scenarios, as well as to produce basic dynamic traffic patterns for the VRE streetscape scene. Signals were introduced to handle traffic control for car-agents, and crossing safety for pedestrian-agents and for ego-agents (Figure 4, at left).





Figure 4: (Left) Based on personality, agent-pedestrians in the VGE use socio-visual gaze to examine things that grab their attention. (Right) Agents may also gaze at each other [man at right of image] and at the ego-character [woman in center].

#### 2.1 Data structures

All agents in the model were encoded as geographic automata (G) [21], with variations in their typology (K), state arrays (S), state transition rules ( $T_S$ ), movement rules ( $M_L$ ), neighbor arrays (R), and neighborhood rules ( $T_S$ ). Geographic automata shared location conventions (L), with positions noted in both the VGE's navigation mesh and the VRE's scene graph (using slipstreaming [20]). Typology transition rules ( $T_S$ ) were not used as agent remained fixed within their initial typologies.

$$\begin{split} G &\sim (S_t, K, L, N(I_t), R_S, R_K, R_L, R_N) \\ R_S: S_t &\longrightarrow (S_t + n|I_t); \\ R_K: K_t &\longrightarrow K_{t+n}; \\ R_L: L_t &\longrightarrow L_{t+n}; \\ R_N: N(I_t) &\longrightarrow N(I_{t+n}) \end{split} \tag{1}$$

Path information was stored separately as a navigation graph [22]. The navigation graph was adapted from a GIS layer file of the intersection under study, based on TIGER line data (sidewalks and road lanes) and building footprints and geometry. All objects and automata in the model were indexed to the graph. Geographic automata accessed the graph using their neighborhood input  $N(I_t)$  lookups and neighborhood transition rules  $(R_N)$ , e.g., when caragents detect a nearby collision with another car-agent, they enter into a new  $N(I_t)$  lookup to calculate their deceleration or stopping rules. All objects in the VGE were held in a scene graph [23, 24] for rendering in the VRE, We relied on the built-in scene graph for Unity3D [25] as well as its materials, bump-mapping, lighting, and rendering engine.

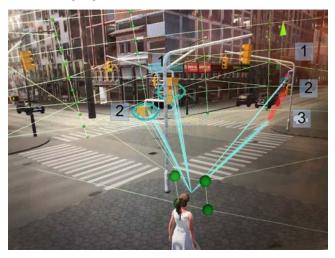


Figure 3: Ray-casting to assess users' ego-agent gaze upon VGE objects within the VRE. A user's field of vision is shown in pale green; gaze rays are shown in white; gazed-upon objects are illustrated as cyan eye icons. Above, as a user approaches the road crossing, she is gazing upon two traffic lights (blue boxes 1), two crossing signals (boxes 2), and the curb (box 3) on the other side of the crossing. (Red lines are ray-snappings to the nearest candidate scene object because the user gazes near—but not at—the opposing curb. The large three-dimensional grid corresponds to the geofence for the VRE experiment.)

#### 2.2 Designing agent behavior

We introduced three types of behavioral models. In each case, the models were used to produce immersive streetscape experiences in simulation and to elicit actions, interactions, and reactions from the (real, human) simulation user. The data recorded from these experiences then became the primary performance indicators and outputs of our experiments. Human behavior for pedestrianagents was designed to produce realistic movement, life-like locomotion, place-based behaviors for road-crossing, socio-visual gaze, and inter-personal interactions. Traffic behavior was designed to control the movement of car-agents to drive plausible dynamics of inner-city traffic, to determine collision avoidance rules between car-agents and between (other) car-agents, facilitate pedestrian-agents, and ego-agents, and to

communication between implied drivers and pedestrian-agents and ego-agents. *Signal behavior* was used to provide street control of foot traffic and car traffic at the four approaches to the modeled intersection. We developed *road-crossing behavior* as combinations of the other behaviors.

### 2.3 Agent-based human behavior

The aim of this work is to develop specific capabilities for geosimulation to support real-time experiential interaction between human users and synthetic agent equivalents within a realistic urban streetscape setting. Further, the simulations for this study were designed as VGEs to represent the urban geography of a real-world site in downtown Brooklyn, NY. We therefore emphasized spatial (and largely movement-associated) behavior in the design of the agent models. To build model human behavior for agent-pedestrians, we focused on high-level movement of path-planning [26]; medium-level movement of navigation, and wayfinding [20]; low-level movement of steering and collision avoidance [27]; locomotion [28]; and perception [29-31]. We also introduced specific road-crossing behavior for the area of sidewalk and pelican crossing. In addition, we introduced sociospatial behavior, through socio-visual gaze and personality.

The design of the models for this paper built on some of our previous work in high-fidelity behavioral modeling for geosimulation [32]. Elsewhere, we have explored building sociobehavioral models in geosimulation through data-mining and machine learning [33, 34], as well as anti-social behavior among crowds of agents in geosimulation [35]. However, the development of the socio-behavioral components through sociovisual gaze and personality are novel to this paper.

High-level movement for agent-pedestrians was introduce as follows. For path-planning, agent-pedestrians were initially seeded with origin-destination (OD) pairs at randomized locations on the VGE sidewalk within a close distance to the intersection being studied. Agent-pedestrians used an A\* best-first traversal heuristic [36] to plan a path between the OD pairs. Upon reaching their destination, the agent-pedestrians would call for a new set of OD-pairs with their terminal node on the previous path as the origin for the subsequent OD-pair. These paths were stored in run-time using a navigation mesh [22], with links to the simulation scene graph. We demarcated the mesh into four distinct movement zones: roads (for car-agents), sidewalks (for pedestrian-agents), pelican crossings (pedestrian light controlled; for pedestrian-agents and for car-agents), and the area covered by the VRE geo-fence (for ego-agents). The use of a navigation graph structure has some basis in theory and reality, as pedestrians are understood to make use of environmental cues that are encoded in (and read from) streetscapes in determining their high-level movement, i.e., movement decisions that are revisited and updated relatively infrequently [37, 38]. Within the navigation mesh, portions of the VGE that were traversable by agentpedestrians were distinguished from those corresponding to roads, such that agent-pedestrians were not permitted to step into the road (this was not the case for human users, who had no such limitation). The portion of the roadway that corresponded to

pelican crossings was also demarcated in the navigation mesh, and agent-pedestrians engaged in specific movement behavior (see "road-crossing behavior") when entering those spaces.

We facilitated medium-level movement for agent-pedestrians in the following approach. Agents were programmed to follow their settled (shortest) navigation path through the navigation mesh. The A\* heuristic promotes heavy-tailed distributions of turns [39, 40], such that agent-pedestrians will generally travel in a straight line between O-D pairs, steering (by turning) to make adjustment relatively infrequently [41]. Elsewhere, we have shown that the patterns of coarse movement that this produces match observational data in downtown settings [26]. However, if agentpedestrians detect a potential collision (with other agentpedestrians or with an ego-pedestrian), they may decouple from their high-level movement routine to engage in "medium-level" movement. (We term this as medium-level as it takes place over spaces of the VGE that sit between paths and step-by-step locomotion; it also involves neighborhood scanning that is localized to small bands of space-time around the agent.) Mediumlevel movement was invoked using basic navigation and steering. Navigation was handled straightforwardly: an agent was tasked to proceed from origin to destination; when the destination is reached, it then calculated a new destination, pulled the shortest path from the navigation mesh, and continued moving along the new path. Steering was handled using Reynold's [42] well-known rule-set for tiered autonomous movement. For most collisions, steering provides a collision-free path for agent-pedestrians. The real human ego-agent, however, may move relatively erratically and suddenly depending on how the user experiences the VRE. To accommodate realistic collision behavior between ego-agents and pedestrian-agents, we developed an attention system. Pedestrian-agents are assigned a random set of attention priorities, including for agents of a given typology. If pedestrianagents detect an immediate collision that they cannot steer around, lower priority agent-pedestrians will calculate a sub-path to yield the right of way to higher-priority agents. As we will discuss, gaze is also tied to the same attention system, which results in situations in which agent-pedestrians may gaze upon an ego-pedestrian, then decide to slow down and steer or speed-up without yielding the right of way, all the while locking eye contact with the ego-agent (see Figure 4 at right). The agent-pedestrians were designed to never collide with cars. This is not the case for the ego-agents controlled by the human user in simulation.

We modeled low-level movement for agent-pedestrians in the following fashion. We used motion blending [43] to produce realistic-appearing locomotion (stride and matching ambulation, as well as transitions between motion states such as standing  $\rightarrow$  idling  $\rightarrow$  walking  $\rightarrow$  jogging  $\rightarrow$  running, as well as whole-body turning maneuvers, and head-turning maneuvers. Elsewhere, we have shown how motion blending can be tied to movement rules for geographic automata within GIS [28]. For the experiments represented in this paper, we relied on existing motion capture libraries with *Unity 3D*. ("Idle", "quarter turn", "half turn" (with rotational directionality for both turn volumes); "stand"; "walk", "sharp walk", "run" and "run sharp".) In addition, we made use of

a state-transition blend tree to generate a number of blends and combinations between these base libraries using motion warping to resolve state-to-motion transitions [44]. This was handled in run-time for the VRE using the "third person controller" and "animator controller" in *Unity 3D* [25].

Personality. Agents were also specifically programmed with different *personalities*, which we based on rank-attraction to other agent-objects in the scene. The primary responsibility for the personality model was to drive pedestrian-agent gaze. At each step in a simulation run, as objects fall within the visual range of agent-pedestrians, they will gaze at the object that has the highest rank-attraction (for example, they will turn their head to look away from a traffic signal and toward a user as she walks past them). A secondary aspect of pedestrian-agents' personality controlled their verbal interactions with other agents. If a pedestrian-agent comes into potential collision with another pedestrian-agent, it will gaze at that agent and vocalize a warning by playing recorded audio announcing, "Hey, watch out!". As highlighted when discussing collision detection above, agentpedestrians also deploy an attention priority system that yields different collision personalities (tendency to yield by steering to adjust course when encountering another agent-pedestrian or ego-pedestrian, or not).

#### 2.4 Car behavior and agent-based traffic



Figure 5: Simulated traffic patterns flowing into and out of the experimental intersection.

Cars followed a roadway mesh in the VGE, which limited them to the road spaces in simulation. Their velocity and maximum speed were straightforwardly set with a user-defined parameter. Cars were programmed to obey traffic signaling rules. If a car neared an intersection as the signal shifted from red to green, they would slow and come to a stop before the pedestrian crossing area. Cars were also programmed to avoid collisions with *other cars* by slowing and stopping. To generate simple traffic, we gave cars shifting destinations around the VGE (Figure 5). Car-agents were specified to maintain a safe driving distance behind other cars, by velocity-matching. This matching, coupled with stopping behavior was anecdotally successful in producing realistic traffic dynamics of congestion, for example. We specifically introduced car-pedestrian interactions in the model by enabling cars to stop (abruptly if necessary) if a pedestrian crossed their path. (Cars

would also play an audible honking sound to alert pedestrians that they were in danger of colliding with them.)

#### 2.5 Signal behavior

Signals in the traffic intersection are represented in the model as agents, but really just run as events that other agents interact with as cues for their movement (i.e., agents do not have an influence on the signals by pressing a crossing button, for example). We included two types of signal agents: pedestrian signals and traffic signals. Both are designed to work together. Traffic signals regulate the flow of car traffic in the model, and pedestrian signals notify pedestrians when it is safe to cross the intersection (using a walk or don't walk signal; see Figure 4 at left). We programmed the timing of the signals to match the commensurate timing in the corresponding real-world traffic intersection at our study site.

#### 2.6 Road intersection and crossing behavior

Pedestrian-agents and car-agents were programmed with dedicated movement rules when in the portions of the navigation graph marked for road-crossing. Both car-agents and pedestrian-agents were made aware of the state of the traffic signal and the pedestrian signal respectively. Car-agents adhered to traffic rules: if a traffic light changed state from green to amber to red, they would slow to a stop at the cusp of the pelican crossing. Additionally, if car-agents detected pedestrian-agents or ego-agents in the crossing while the signal indicated green, they would wait for the crossing to completely clear of agents before moving. If the traffic light changed state to red while car-agents were in the crossing, they were programmed to continue their movement.

Pedestrian-agents, upon entering the pelican crossing, can actually see two pedestrian signals: one on the sidewalk at which they are standing, as well as the opposing signal on the other side of the road. We therefore had to be careful to build rules for pedestrian-agents to follow. This is a situation that is common in the real-world, where a pedestrian may judge how much time they have to cross based on signals at the other side of the crossing.

#### 2.7 Building a Virtual Geographic Environment (VGE).

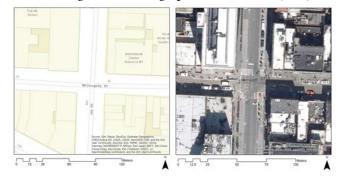


Figure 6: Our study site in downtown Brooklyn (map at left; aerial image at right).

We developed a virtual geographic environment (VGE) for simulation scenarios [20]. We modeled the VGE on a very busy

intersection in downtown Brooklyn, NY, at the confluence of Jay Street and Willoughby Street (Figures 6 and 7). Jay Street carries a lot of vehicle traffic to and from the Brooklyn Bridge, as well as large flows of commuters to nearby offices, schools, and universities via the transit hub at Jay Street/MetroTech. Willoughby Street typically caters to smaller volumes of local residents and shoppers patronizing its small community shops; vehicle traffic volumes are usually light. The mixture of the two flows produces some interesting rhythms and motifs for streetscape dynamics, with frequent potential for pedestrianvehicle interactions at its two signalized pedestrian crossings.



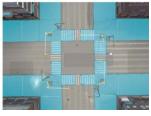


Figure 7: Our VGE model of the study site (left), with walkable area (sidewalk and crossing) in blue (right).

# 2.8 Running VGEs as a Virtual Reality Environment (VRE).

The VGE was rendered and presented to the users in experimental scenarios as a virtual reality environment (VRE) (Figure 8). The key design factor in developing the VRE was to present the VGE with significant fidelity to appearance, dynamics, and particularly to endow it with experiential *immersion*.

We rendered our VGE model via *Unity 3D* for participant use in the *HTC Vive Pro* virtual reality headset. Several city blocks around our study site were modeled with detail. This provided both background sense of place and foreground recognition for collision detection, use of waypoints from street and building features, sense of movement, motion flow, sense of proximity, and projection of users' personal space into the VRE (Figure 8).





Figure 8: Geographically-faithful VGE background (left) and foreground (right) urban environments rendered in the VRE.

The VGE was rendered in the VRE in appreciably "real-time". We ran the entire VRE at 60 frames per second; users reported no problems with jittering or lag during the experiments. The *Vive Pro* provided tracking between the headset (and the user's head position and head movement) and a set of referenced base stations in our laboratory. The *Vive Pro* allows the user to *physically walk around* while engaged in the VRE. This allowed us to use a roughly

100 m<sup>2</sup> (~1,000 sq. ft.) tangible space for participants to walk around in our laboratory. We geo-fenced the area within the corresponding VRE so that users were made visually aware when they neared the limits of the physical space (see Figure 3). To map the physical space of the laboratory to the VGE, we applied a slight acceleration function to the user's walking speed, so that space covered in the real world corresponded to a larger space in the VRE. Some users reported that this was somewhat strange, but they quickly got used to it. Our colleague Xi Sun recently introduced a technique for mitigating these accelerations, using a folding technique that warps the VR space, alongside redirected physical walking to yield the illusion of traversing large distances [45]. The technique is not suitable for our approach here because pedestrian movement is heavy-tailed. We also adjusted the height of the agent representations in the VRE per experiment participant to adjust for taller or shorter participants, so that the user was always close to eye-level with the tallest agentpedestrians in the VRE.

A number of studies in psychology have reported that VR environments can be useful proxies for real-world spatial experiments, particularly those involving spatial behavior [46, 47]. Movement and vision within the *Vive Pro* produces rather faithful representations of real-world perception when rendered through *Unity 3D*. Users reported that the experience appeared and felt realistic, particularly because they received tactile response to walking, as well as their own physical feelings of balance and acceleration while moving in the lab. Indeed, one of the first things that users did upon entering the VRE was begin to look around, followed quickly by quite confident walking. Similarly, when departing from curbed areas in the VRE, users would lift and release their feet as if negotiating a real three-dimensional infrastructure.

We made limited use of audio to establish an aural sense of place in the VRE. Specifically, the audio interactions between pedestrian-agents (and between pedestrian-agents and egoagents) were scaled based on the ego-agent's distance to colliders. So, for example, if an ego-agent turns abruptly into a collision, the audio will play loudly, but if there are collisions among other agents far behind them, they will hear the audio more softly and spatially displaced (using the *HTC Vive* built-in spatial audio capabilities). Car-agents were also designed to play a honking alert if pedestrians entered the roadway space and neared a collision with them. This was also built with spatial audio, with the result that users would receive a sudden alert in their near proximity near a collision.

2.8.1 Immersive behavior of real human users within the VRE. The behavior of the ego-agent was not modeled; it is directly supplied by the perception, cognition, decision-making, movement, locomotion, and social interactions of the human user that participates in the experiment. The users are dropped into the VRE and must rely on their perception and cognition of the VGE to rapidly make sense of the simulated scene's geography and dynamics. Users reported no problems in building a sense of space and place within the VRE. Most looked around briefly, commented

that it felt realistic, and began engaging in the experiment tasks. Most picked out the street corner at the intersection, and the pelican (pedestrian light controlled) crossing consisting of a pedestrian signal and a marked crossing area. All participants adopted a shortest path to complete their traversal tasks in experiments. They also engaged in collision avoidance with agent-pedestrians and street objects (traffic signal poles). All participants stayed within the bounds of the sidewalk. Most participants also negotiated the curb at the pelican crossing carefully. Behavior at crossing opportunities was mixed, with some users waiting for walk signals and others jaywalking. This mixture of behaviors is somewhat typical in New York. Participants were all surprised when agent-characters would gaze at them, and quite taken aback when they would talk to them. Similarly, most participants would stop their movement abruptly when a car honked at them as they were crossing. In essence, our experiment allowed us to use real human users to drive character movement in simulation, and our observations of the behaviors of the human users in our physical setting enabled the discovery of new behaviors that were not represented in the simulation, particularly emotional behaviors (surprise) and rapid-onset perception and cognition (being caught unawares, startled) that correspond to spatial behaviors (abrupt movements, rapid scanning of hyper-local surroundings to gather visual information and update sense of space, proximity scanning around personal space). Our focus in this paper is on determining what users were gazing at during the simulation, but these anecdotal observations from our time studying users in the VRE suggests that much more could be done to examine users' verbal and non-verbal communications while in simulation. We propose that behavior coding and motion capture could be useful to capture structured and empirical data for these purposes.

2.8.2 In-VRE data collection using data listeners. Because the experiment is conducted in a hybrid real/virtual environment (the material space of our research laboratory and the virtual space of our VGE city), we were able to generate a huge volume of highresolution empirical data during the experiment. We classified these data using a three-fold information typology based on Tuan's theory of spatial information coding and processing [48]. Ego-centric data refers to the (behavioral) geography formed between the human subject and objects that they perceive in the space of the VGE/VRE. These data are sourced primarily through the immersion-emission framework and constitute (real) visual connections (by socio-visual gaze) between the user's egocharacter and VGE objects. Allo-centric data are attributed to the geography that forms between objects in the localized spaces of the VGE/VRE. Allo-centric data are distinct from ego-centric data in their lack of ego-connection to the user. While the user may use allo-centric geography for spatial cues, the user does not usually form a self-object relationship with allo-centric data. For example, a user might identify open space between two pedestrians on the simulated sidewalk and determine that she has latitude to walk through that space. This is a different form of spatial relationship than would be experienced when an agentcharacter looks at the user and begins to walk toward them, for

example. One key distinguishing factor is that for ego-centric data, the user has a direct involvement in creating the geography; for allo-centric data, the user is placed in more of an observer relationship. We added a third class of data for events, which we consider as key spatial interactions in the experiments through collision. Specifically, ego-centric and allo-centric data are tagged using event data, which allow us to focus our data queries on important substantive incidents in the experiments. In particular, we are therefore able to build data bundles for collision events that include the event location and timing, as well as the ego-centric and allo-centric data that correspond to that event. We refer to these as context bundles. Each context bundle is structured using a time stamp as well as a location stamp. We streamed context bundle data to a dashboard for data collection and visualization (Figure 9); this enabled us to fine-tune settings between experiments if necessary.

#### 2.9 The immersion-emission model.

The aim for this paper is to explore whether closer-couplings might be built between real human user experiences and geosimulation, and we have proposed VGEs and VREs at experimental media for building these connections. We propose a specific method to facilitate tangible experiments and their corresponding representation in the VGE/VRE, which we term as, "immersion-emission". The immersion component corresponds to the virtual setting for the geosimulation, which in the case of this paper is a realistic-appearing and lifelike-operating streetscape with built fabric, logical geography, space-specific events, and dynamic agent characters in the form of synthetic pedestrians and cars with perceptive and interactive capabilities that allow the VRE user to build spatial (and in some ways also social) awareness of the geosimulation, using their natural spatial behaviors and skills.

The emission component of our framework comes from the techniques that we use to collect data in the model. Here, we follow terminology from since-debunked science of the past, which supposed that the human eye worked by generating rays of light that would be cast toward objects [49]. Our computational procedure follows a similar method, by ray-casting through a focused window of the user's field of view [50, 51]. Specifically, we use ray-casting through the user's field of vision to connect their gaze with objects in the VGE scene graph (as shown graphically in Figure 3). (Work on space syntax and visibility graphs [52, 53] proposes that a similar scheme is used by pedestrians in the real world to build natural movement responses in built environments.) In our models, agents' personality provides the social components of this gaze. Immersion-emission allows us to empirically and tirelessly build data about what the user is looking at, for how long, and what other things were going on in the model for those bundles of space and time. This produces a huge amount of data as the users continually scan their surroundings and shift their gaze as part of their normal stereo vision and locomotion [54-56]. We may analyze these data for any of the 60 frames per second that a user spends in the experiment, or for any feature of interest in the VGE (see Figure 3, where we

analyzed and visualized which components of a streetscape a user gazed upon).

Gaze is just one modality in which we might examine how people map their perception and cognition to their surroundings, and clearly there remains many open questions about how other aspects of their activity or non-verbal communications might be studied while they participate in the geosimulation. In prior work, we spent considerable time examining connections between affect and spatial behavior, albeit through coded observation and trajectory sampling over two-dimensional spaces [33, 34, 57]. Much of the same methodology could be applied within VGEs/VREs, with the benefit that they VGE would return empirical data to match observed codes. In our data analysis that follows, we focused only on gaze at objects in the few seconds before and after ego-agent collisions with car-agents. (This also enables us to differentiate between normal saccade-type looks from our human participants, and meaningful gaze fixation on objects.)

# 3 REAL HUMAN USER EXPERIMENTS

To evaluate the utility of our approach, we conducted a series of experiments with the system. We recruited ten participants to individually spend between three and five minutes in the VRE. (Admittedly this is a relatively small number of users, but our intent was to demonstrate the viability of the platform. In future work, we will demonstrate experiments with the system using larger user cohorts that we have recruited.) Each participant was tasked with performing a circuit of the modeled intersection (four crossings) and they were instructed to take care to avoid walking into the road and to avoid collisions with other pedestrians that they might encounter, with the same care that they might afford in real life. After the experiment, participants were given a questionnaire survey designed to elicit qualitative information about their experiences with road crossing in the real world, and their experiences in the geosimulation.

Participants were recruited as volunteers and experiments were conducted under an experimental design approved by the New York University Institutional Review Board (IRB). The age range of participants was 21–45 years (average age of 26.6 years) and five female and five male participants participated in the experiments. None of the participants reported physical disability. Of the ten participants, six had no prior experience or little experience with VR. As participants immersed themselves in the VRE, one member of our team watched their physical movement and checked-in with them verbally at regular intervals. The other team member studied the incoming data as well as the participant's view within the VRE, using a dual projection to a screen (Figure 1).

The survey instrument asked for participant background information (age, sex, education, profession, and whether the participant felt that they had a locomotive disability). Participants were also asked to rate their experience with VR ("I have a very good experience of interacting with the virtual reality

environment"), as well as their potential susceptibility to motion sickness ("I do not feel dizzy or tired or sick while interacting with the virtual reality environment"). A second set of questions were posed to elicit experiential feedback. Users were asked to rate their locomotive skills, their driving skills, and their usual adherence to traffic signals and pedestrian crossing. This latter question was designed to evaluate participants' sense of risk-taking in road-crossing. A third set of questions asked participants to recall their behavior within the VRE geosimulation. They were asked a free-form question, "Please specify all events and objects that you remember from your experience". They were also asked to rate the immersion and quality of the virtual experience. Finally, participants were asked a free-form question about whether they had faced any difficulty with the equipment or technology used to run the experiment.



Figure 9: A frequency visualization of objects that all users gazed upon immediately before and during collisions with cars in the geosimulation experiments.

# 4 RESULTS

#### 4.1 Survey results.

All participants in our experiments responded with high self-rating of their locomotion and driving skills (average rating of 8.1 out of 9, where 9 corresponds to "strongly agree"). They similarly self-reported strong adherence to the rules of the road as pedestrians and drivers (average 7.3 out of 9). All participants also responded that they had enjoyed the experience of walking through the geosimulation (average 8 out of 9). The only issues reported with the technology were the limitations of using a wired tether to the *Vive Pro* headset. (We have since developed a version for untethered headsets, which we have tested on the *Oculus Quest.*) The free-form responses to the recall question of objects

and events that were encountered illustrated that participants recalled looking at pedestrians, noticed cars passing them by as they walked, were aware of traffic and pedestrian signals, and noticed the built surroundings. These free-form reports match the empirical data that we gathered from the geosimulation, which suggests at least anecdotally that the immersion–emission model is useful in collecting information that matches to real user perceptions and experiences.

#### 4.2 Immersion-emission data

As mentioned, we focused on the gaze-to-object connections (fixation) in the few seconds preceding and during participants' collision with cars in the geosimulation. The aggregate results are shown visually in Figure 9, which scales the frequency of gaze to the size of the circle in the visualization. The most frequent objects of participants' attention during collisions were other pedestrians and cars. The next most-frequent set of objects were the ground (participants were looking down when they collided). The least frequent objects were individual buildings and traffic signals.

#### 4 CONCLUSIONS AND FUTURE WORK

This work aims to expand connectivity between real human behavior and geosimulation representations of that behavior. To accomplish this, we developed a system for immersing human participants directly in a geosimulation through a combination of VGE and VRE. We coupled immersion with an emission-based framework designed to collect data on human gaze fixation within the environment. To establish tight-coupled links between immersion and emission, we developed basic personality rules for agent-characters, designed to elicit corresponding reactions and interactions from the human participants. We ran a set of experiments to test the approach. Our results show that the system is effective in building human-in-the-loop connections to geosimulation; in particular, it supports a range of natural interactions between real human users and synthetic characters designed to mimic human behavior in a complex urban setting.

The approach has a number of limitations. The VGE could be developed with significantly more sophistication, including a larger number of agent pedestrians and cars that would provide a more faithful representation of a real urban scene. Similarly, the buildings in our geosimulation were inert; this could be adapted to produce more realistic interactions between agents and the urban geography of the scene, to produce rhythms and motifs of commuting, shopping, tourism, etc. Critically, the behavioral rulesets for the geosimulation could be expanded to produce high-fidelity streetscape dynamics. (For example, cars turning right on red lights is a large factor in road crossing safety in New York.)

The VRE could also be adapted to handle some technical and immersion issues. Our use of the *HTC Vive* limited us to a fixed space. In recent work we have begun to use wireless headsets that free us from these limitations, but the graphical resolution of those devices has proven problematic in supporting busy geosimulations on VGEs. Our VRE experiments were limited to single participants; this could be expanded to include multiple

human users at once, which would allow for more realism and would also expand the reach of the experiments into issues of group behavior (e.g., peer influence). The *HTC Vive* is also limited to head movements; in future work we hope to include full-body representation using inertial measurement devices.

Data collection could be expanded. Indeed, the data production capabilities of our system is one of its main advantages. We focused narrowly on critical data during collision events, but this could be expanded to generate data on all interactions within the geosimulation, e.g., the spatial behavior of participant interactions with agents (bias, leader-follower behavior, social gaze), and as VGE objects corresponding to urban design factors.

The experiments that we ran were limited and could be expanded in the future. In particular, we have an interest in exploring the influence of different times of day (which can easily be represented in the VRE). Similarly, weather effects (particularly rain and its influence on visibility and urgency in movement) could be accommodated. We could also develop a wider range of experimental scenarios, as the geosimulation would support perhaps limitless configurations. In particular, we have an interest in examining the influence of gap acceptance (how much space and time a crosser uses to move through an intersection) on road crossing, as a canonical space-time concern that maps directly to human perception, cognition, behavior, and spatial skill. We are also interested in using the environment to assess design options for urban infrastructure. For example, different signal timing, curb geometries, traffic islands, and signage could be examined in simulation before being considered for real-world applications.

We have recently run a series of experiments with a larger cohort of participants (35 people). However, more users produce (significantly) more data and we are currently exploring reliable data-base schemes that can work in parity with the system (at 60 fps), as well as dashboard tools that can give us real-time feedback about key performance indicators (KPIs) within the simulation.

Despite tlimitations, in conclusion, we hope to express that all of these "fixes" are possible in the system that we developed. Indeed, this is one of the advantages of connection geosimulation with human-in-the-loop interactivity. The message that we have tried to demonstrate here is that the intricacy of geosimulation and its abilities to support huge swaths of what-if experimentation can be brought to parity with the intricacies of human behavior, in a way that supports empirical knowledge discovery.

#### **ACKNOWLEDGMENTS**

This material is based upon work supported by the National Science Foundation under Grant No. 2027652 and 1729815.

#### REFERENCES

- Benenson, I. and Torrens, P. M. 2004. Geosimulation: Automata-Based Modeling of Urban Phenomena. John Wiley & Sons, London.
- [2] Ållen, P. M. and Torrens, P. M. 2005. Knowledge and complexity. Futures, 37, 7, 581-584.
- [3] Batty, M. and Torrens, P. M. 2005. Modeling and prediction in a complex world. Futures, 37, 7, 745-766.

- [4] Batty, M. and Torrens, P. M. 2001. Modeling complexity: the limits to prediction. CyberGeo, 201, 1, Online 1035
- [5] Anderson, T. and Dragićević, S. 2020. Representing complex evolving spatial networks: geographic network automata. ISPRS Int. J. Geo-Inf., 9, 4, 270
- [6] Pesavento, J., Chen, A., Yu, R., Kim, J.-S., Kavak, H., Anderson, T. and Züfle, A. 2020. Data-driven mobility models for COVID-19 simulation. In Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Advances in Resilient and Intelligent Cities, Seattle, WA, November 3-6, 2020, eds. C.-T. Lu, F. Wang and G. Trajcevski, 29-38. New York, NY: Association for Computing Machinery
- [7] Shashidharan, A., van Berkel, D. B., Vatsavai, R. R. and Meentemeyer, R. K. 2016. pFUTURES: A parallel framework for cellular automaton based urban growth models. In Proceedings of 9th International Conference, GIScience 2016, Montreal, QC, Canada, September 27-30, 2016, eds. J. A. Miller, D. O'Sullivan and N. Wiegand, 163-177. Cham: Springer International Publishing
- [8] Bush, B. W. 2001. Portland Synthetic Population. Los Alamos National Laboratory Report LA-UR-00-5972, Los Alamos National Laboratory, Los Alamos
- [9] Pires, B. and Crooks, A. T. 2017. Modeling the emergence of riots: a geosimulation approach. Comput. Environ. Urban Syst., 61, 66-80.
- [10] Kim, J.-S., Jin, H. and Züfle, A. 2020. Expert-in-the-loop prescriptive analytics using mobility intervention for epidemics. In Proceedings of the First Workshop on Prescriptive Analytics for the Physical World (PAPW 2020) (KDD2020), eds. Z. Li, S. Chawla, Y. Li, N. Ramakrishnan, C. Shahabi, W. Zhang, D. Gunopulos, H. Wei, G. Zheng, S. Song and L. Nocera, Online at https://prescriptiveanalytics.github.io/file/4-hjin.pdf. New York, NY: ACM
- [11] Li, X., Cai, W. and Turner, S. J. 2017. Cloning agent-based simulation. ACM Trans. Model. Comput., 27, 2, Article 15.
- [12] Kim, J.-S., Kavak, H., Manzoor, U. and Züfle, A. 2019. Advancing simulation experimentation capabilities with runtime interventions. In Proceedings of the IEEE 2020 Spring Simulation Conference (SpringSim), (virtual), May 18-21, 2020, eds. F. J. Barros, X. Hu, H. Kavak and A. A. Del Barrio, 2019.8732869. Los Alamitos, CA: IEEE
- [13] Zou, Y., Torrens, P. M., Ghanem, R. and Kevrekidis, I. G. 2012. Accelerating agent-based computation of complex urban systems. Int. J. Geogr. Inf Sci., 26, 10, 1917-1937
- $[14]\,$  Couclelis, H. 2002. Why I no longer work with agents: a challenge for ABMs of human-environment interactions. In Agent-based Models of Land Use and Land Cover Change, eds. D. Parker, T. Berger and S. M. Manson, 3-5. Bloomington, IN:
- [15] Faith, J. 1998. Why gliders don't exist: anti-reductionism and emergence. In Artificial Life VI: Proceedings of the Sixth International Conference on Artificial Life, ed. C. Adami, 389-392. Cambridge, MA: MIT Press
- [16] Kavak, H., Padilla, J., Diallo, S. and Barraco, A. 2020. Modeling the modeler: an empirical study on how modelers learn to create simulations. In Proceedings of the IEEE 2020 Spring Simulation Conference (SpringSim), (virtual), May 18-21, 2020, eds. F. J. Barros, X. Hu, H. Kavak and A. A. Del Barrio, 1-12. Los Alamitos, CA: IEEE
- [17] Torrens, P. M. 2002. Cellular automata and multi-agent systems as planning support tools. In Planning Support Systems in Practice, eds. S. Geertman and J. Stillwell, 205-222. London: Springer-Verlag
- [18] Lin, H. and Batty, M. 2011. Virtual Geographic Environments: A Primer. ESRI Press, Redlands, CA.
- [19] Lin, H., Batty, M., Jørgensen, S. E., Fu, B., Konecny, M., Voinov, A., Torrens, P. M., Lu, G., Zhu, A.-X., Wilson, J. P., Gong, J., Kolditz, O., Bandrova, T. and Chen, M. 2015. Virtual environments begin to embrace process-based geographic analysis. Trans. GIS, 19, 4, 1-5.
- [20] Torrens, P. M. 2015. Slipstreaming human geosimulation in virtual geographic environments. Ann. GIS, 21, 4, 325-344.
- [21] Torrens, P. M. and Benenson, I. 2005. Geographic Automata Systems. Int. J. Geogr. Inf Sci., 19, 4, 385-412.
- [22] Gayle, R., Sud, A., Andersen, E., Guy, S. J., Lin, M. C. and Manocha, D. 2009. Interactive navigation of heterogeneous agents using adaptive roadmaps. IEEE Trans. Vis. Comput. Graph., 15, 1, 34-48. Osfield, R. and Burns, D.
- [23] Osfield, R. 2019. OpenSceneGraph http://www.openscenegraph.org/.
- [24] Sowizral, H. 2000. Scene graphs in the new millennium. IEEE Comput. Graph. Appl, 20, 1, 56-57.
- [25] Unity Technologies. 2021. Unity 3D 2021.1.20, San Francisco, CA.
- [26] Torrens, P. M., Nara, A., Li, X., Zhu, H., Griffin, W. A. and Brown, S. B. 2012. An extensible simulation environment and movement metrics for testing walking behavior in agent-based models. Comput. Environ. Urban Syst., 36, 1, 1-17.
- [27] Torrens, P. M. 2014. High-resolution space-time processes for agents at the built-human interface of urban earthquakes. Int. J. Geogr. Inf. Sci., 28, 5, 964-
- [28] Torrens, P. M. 2012. Moving agent pedestrians through space and time. Ann. Am. Assoc. Geogr., 102, 1, 35-66.
- [29] Torrens, P. M. 2018. A computational sandbox with human automata for exploring perceived egress safety in urban damage scenarios. Int. J. Digit. Earth, 11, 4, 369-396.

- [30] Torrens, P. M. 2016. Exploring behavioral regions in agents' mental maps. Ann. Reg. Sci., 57, 2-3, 309-334.
- [31] Torrens, P. M. 2015. Intertwining agents and environments. Environ. Earth Sci., 74, 10, 7117-7131.
- [32] Torrens, P. M. 2014. High-fidelity behaviours for model people on model streetscapes. Ann. GIS, 20, 3, 139-157
- [33] Torrens, P. M. and Griffin, W. A. 2013. Exploring the micro-social geography of children's interactions in preschool: a long-term observational study and analysis using Geographic Information Technologies. Environ Behav, 45, 5, 584-614
- [34] Torrens, P. M., Li, X. and Griffin, W. A. 2011. Building agent-based walking models by machine-learning on diverse databases of space-time trajectory samples. Trans. GIS, 15, s1, 67-94.
- [35] Torrens, P. M. and McDaniel, A. 2013. Modeling geographic behavior in riotous crowds. Ann. Am. Assoc. Geogr., 103, 1, 20-46.
- [36] Hart, P. E., Nilsson, N. J. and Raphael, B. 1968. A formal basis for the heuristic determination of minimum cost paths. IEEE Trans. Syst. Man Cybern. Syst., 4, 2,
- [37] Kitazawa, K. and Fujiyama, T. 2010. Pedestrian vision and collision avoidance behavior: investigation of the information process space of pedestrians using an eye tracker. In Pedestrian and Evacuation Dynamics 2008, eds. W. W. F. Klingsch, C. Rogsch, A. Schadschneider and M. Schreckenberg, 95-108. Berlin: Springer
- [38] Cutting, J. E. 1993.Perceptual artifacts and phenomena: Gibson's role in the 20th century. In Advances in Psychology, ed. S. C. Masin, 231-260. New York: Elsevier Science Publishers
- [39] Brockman, D., Hufnagel, L. and Geisel, T. 2006. The scaling laws of human travel. Nature, 439, 462-465.
- [40] Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A. and Stanley, H. E. 1996. Lévy flight search patterns of wandering albatrosses. Nature, 381, 6581, 413-415.
- [41] Nams, V. O. 1996. The VFractal: a new estimator for fractal dimension of animal movement paths. Landsc. Ecol., 11, 5, 289-297.
- [42] Reynolds, C. W. 1999. Steering behaviors for autonomous characters. In Proceedings of Game Developers Conference, San Jose, CA, ed. Miller Freeman Game Group, 763-782. San Francisco, CA: Miller Freeman Game Group
- [43] Kovar, L. and Gleicher, M. 2003. Flexible automatic motion blending with registration curves. In Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation, San Diego, CA, July 26 - 27, 2003, eds. D. Breen and M. Lin, 214-224. San Diego, California: Eurographics Association
- [44] Witkin, A. and Popović, Z. 1995. Motion warping. In Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), Los Angeles, CA, August 6-11, ed. M. Cohen, 105-108. Los Angeles, CA: Association for Computing Machinery
- [45] Sun, Q., Patney, A., Wei, L.-Y., Shapira, O., Lu, J., Asente, P., Zhu, S., Mcguire, M., Luebke, D. and Kaufman, A. 2018. Towards virtual reality infinite walking: dynamic saccadic redirection. ACM Trans. Graph., 37, 4, Article 67.
- [46] Loomis, J., Blascovich, J. and Beall, A. 1999. Immersive virtual environment technology as a basic research tool in psychology. Behav. Res. Methods, 31, 4, 557-564.
- [47] Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M. and Beall, A. C. 2004. Does the quality of the computer graphics matter when judging distances in visually immersive environments? Presence: Teleoperators Virtual Environ., 13, 5, 560-571.
- [48] Tuan, Y.-F. 1975. Place: an experiential perspective. Geogr. Rev., 151-165.
- [49] Newton, I. 1730. Opticks, or, a treatise of the reflextions, refractions, inflections & colours of light. London.
- Appel, A. 1968. Some techniques for shading machine renderings of solids. In Proceedings of the AFIPS Spring Joint Computer Conference, 37-45. Atlantic City, New Jersey: Assocation of Computing Machinery
- [51] Roth, S. D. 1982. Ray casting for modeling solids. Comput. Graph. Image Process., 18, 2, 109-144.
- [52] Turner, A., Doxa, M., O'Sullivan, D. and Penn, A. 2001. From isovists to visibility graphs: a methodology for the analysis of architectural space. Environment and Planning B, 28, 1, 103-121.
- [53] Turner, A. and Penn, A. 2002. Encoding natural movement as an agent-based system: an investigation into human pedestrian behaviour in the built environment. Environ Plann B Plann Des., 29, 4, 473-490.
- [54] Cutting, J. E., Alliprandini, P. M. Z. and Wang, R. F. 2000. Seeking one's heading through eye movements. Psychon. Bull. Rev, 7, 3, 490-498.
- [55] Cutting, J. E., Springer, K., Braren, P. A. and Johnson, S. H. 1992. Wayfinding on foot from information in retinal, not optical, flow. J. Exp. Psychol. Gen., 121, 1, 41-72
- [56] Wang, R. F. and Cutting, J. E. 1999. Where we go with a little good information. Psychol. Sci., 10, 1, 71-75.
- [57] Griffin, W. A., Schmidt, S. K., Nara, A., Torrens, P. M. and Sechler, C. 2007. Integrating ABM and GIS to model typologies of playgroup dynamics in preschool children. In Agent 2007, eds. M. North, C. Macal and D. Sallach, 17-24. Evanston, IL: Argonne National Labs and the University of Chicago