Comparing Human and Synthetic Group Behaviors: A Model Based on Social Psychology

Natalie Fridman, Gal A. Kaminka
The MAVERICK Group
Computer Science Department
Bar Ilan University, Israel
{fridman, galk}@cs.biu.ac.il

Abstract

Existing models of group behavior, in a variety of fields, leave many open challenges. In particular, existing models often focus only on a specific phenomenon (e.g. flocking, pedestrian movement), and thus must be switched depending on the goals of the simulation. In contrast, we investigate a general cognitive model of simulating group behaviors, based on Festinger's Social Comparison Theory (SCT), a prominent social psychology theory. In previous work, we have show SCT covers a variety of pedestrian movement phenomena. In this paper we present evidence for SCT's generality by describing the use of the SCT model (using the Soar cognitive architecture) in generation of imitational behavior in loosely-coupled groups. Since the imitational behavior does not have clear standards of evaluation, we propose a method for such evaluation. Based on experiments with human subjects, we show that SCT generates behavior more in-tune with human crowd behavior.

Introduction

Models of crowd behavior facilitate analysis and prediction of the behavior of groups of people, who are in close geographical or logical states, and are affected by each other's presence and actions. Existing models of crowd behavior are often simplistic, and typically not tied to specific cognitive science theories or data. Moreover, existing computer science models often focus only on a specific phenomenon (e.g. flocking, pedestrian movement), and thus must be switched depending on the goals of the simulation.

We propose a novel model of crowd behavior, based on Social Comparison Theory (*SCT*) (Festinger, 1954), a popular social psychology theory that has been continuously evolving since the 1950s. The key idea in this theory is that humans, lacking objective means to evaluate their state, compare themselves to others that are similar. Similarity in SCT is very loosely defined—indeed, much of the literature on SCT addresses the exploration of different ways in which humans judge similarity.

In this paper we describe the implementation and adaptation of SCT the model in the Soar cognitive architecture. SCT was implemented as a secondary parallel thread within Soar. Whereas normally, operators are proposed (and selected) by Soar based on their suitability for a current goal, in our agent, operators were also proposed based on their suitability for SCT. We also briefly discuss mechanisms in the architecture, necessary for enabling SCT: a memory mechanism and an exploration mechanism.

We evaluate the use of SCT in generation of imitational behavior and show that SCT generates behavior in-tune with human crowd behavior. As the imitational behavior does not have clear standards of evaluation, we propose a method for evaluation of imitational behavior. The SCT model was evaluated in studies with human subjects. The subjects ranked SCT to be a middle-ground between completely individual behavior, and perfect synchronized ("soldier-like") behavior. Independently, human subjects gave similar rankings to short clips showing human crowds.

Background and Motivation

Social psychology literature provides several views on the emergence of crowds and the mechanisms underlying its behaviors. These views can inspire computational models, but are unfortunately too abstract to be used algorithmically. In contrast, computational crowd models tend to be simplistic, and focus on specific crowd behaviors (e.g, flocking). A common theme in all of them is the generation of behavior from the aggregation of many local rules of interaction, e.g. (Reynolds, 1987; Yamashita & Umemura, 2003).

Social psychology. A phenomenon observed within crowds, and discovered early in crowd behavior research, is that people in the crowd act similar to one another, often acting in a coordinated fashion which is achieved with little or no verbal communications.

There are several psychological theories that explained this coordinated behavior. For example, Le Bon (Le Bon, 1895) emphasized a view of crowd behaviors as "Collective Mind" that transform an individual who becomes a part of the crowd into becoming identical with the others in the crowd. Le Bon explains the homogeneous behavior of the crowd by two processes: *Imitation* and *Contagion*. Allport, (Allport, 1924) states that crowd behavior is a product of the behavior of like-minded individuals. According to Allport's theory, individuals become a part of the crowd behavior when they have a "common stimulus" with people inside the crowd. Additional explanation of coordinated crowd behaviors (Tajfel & Turner, 1986; Reicher, 2001) suggest that this coordination emerges because people in the crowd share a common social identity. Unlike Allport's individualistic behavior of people in crowds, Social Identity theory combines together the society aspects with an individual aspects.

Computational models. Work on modeling crowd behavior has been carried out in other branches of science, in particular for modeling and simulation. Reynolds (Reynolds, 1987) simulated bird flocking using simple, individual-local rules, which interacted to create coherent collective movement. There are only three rules: avoid collision with neighbors, match velocity with neighbors and stay close to the center of gravity of all neighbors.

Blue and Adler (Blue & Adler, 2000) used Cellular Automata (CA) in order to simulate collective behaviors, in particular pedestrian movement. The focus is again on local interactions: each simulated pedestrian is controlled by an automaton, which decides on its next action or behavior, based on its local neighborhoods.

Helbing et al. (Helbing & Molnar, 1997; Helbing, Molnar, Farkas, & Bolay, 2001) also focused on simulating pedestrian movement. Each entity moves according to forces of attraction and repulsion. Pedestrians react both to obstacles and to other pedestrians.

Yamashita and Umemura (Yamashita & Umemura, 2003) take a different approach in simulating group panic behavior. While inspired by Reynolds' model, they propose a model where each simulated person moves using three instincts: An escape instinct, a group instinct and an imitational instinct. According to Yamashita and Umemura, when a person is in panic, she acts based on these instincts, simplifying the decision making process.

Our work differs from those described above in that we aim to develop a general cognitive model of simulating group behaviors, one based on psychology. We have already shown that our model covers pedestrian movement phenomena as was presented in our previous work (Fridman & Kaminka, 2007), together with initial results on imitational behavior. Here, we present additional evidence for such generality by describing implementation in Soar, and evaluation of SCT model on imitational behavior in loosely-coupled groups. We discuss the full set of results, and the evaluation methodology, in detail.

A Model of Social Comparison

Our research question deals with the development of a computerized cognitive model which, when executed individually by many agents, will cause them to behave as humans do in groups and crowds.

We took Festinger's social comparison theory (Festinger, 1954) as inspiration for the social skills necessary for our agent. According to social comparison theory, people tend to compare their behavior with others that are most like them. To be more specific, when lacking objective means for appraisal of their opinions and capabilities, people compare their opinions and capabilities to those of others that are similar to them. They then attempt to correct any differences found.

We believe that social comparison theory may account for some characteristics of crowd behavior:

Imitation. Using social comparison, people may adopt others' behaviors. Festinger notes (Festinger, 1954): "The drive for self evaluation is a force acting on persons to belong to groups, to associate with others. People, then, tend to move into groups which, in their judgment, hold opinions which agree with their own".

Contagion. One implication of SCT is the formation of homogeneous groups. Festinger writes (Festinger, 1954): "The existence of a discrepancy in a group with respect to opinions or abilities will lead to action on the part of members of that group to reduce the discrepancy".

To be usable by computerized models, social comparison theory must be transformed into a set of algorithms that, when executed by an agent, will proscribe social comparison behavior. A first step towards this goal has been take by Newell, who examined the axioms of social comparison (Newell, 1990), a subset of which appears here:

1. When lacking objective means for evaluation, agents compare their state features to those of others.
2. Agents compare themselves to those who are more similar; comparison increases with similarity.
3. Agents take steps to reduce differences to the objects of comparison.

Newell argued that these axioms are not social, in the sense of requiring active interaction between the agents. Rather, they utilize uni-directional observations and actions by the comparing agents.

We turn these abstract axioms into a concrete algorithm. The algorithm is described in (Fridman & Kaminka, 2007), and we provide only a brief description here. Each observed agent is assumed to be modeled by a set of features and their associated values. For each such agent, we calculate a similarity value $s(x)$, which measures the similarity between the observed agent and the agent carrying out the comparison process. The agent with the highest such value is selected. If its similarity is between given maximum and minimum values, then this triggers actions by the comparing agent to reduce the discrepancy:

1. For each known agent x calculate similarity $s(x)$
2. $c \leftarrow \texttt{argmax}\, s(x)$, such that $S_{min} < s(c) < S_{max}$
3. $D \leftarrow$ differences between me and agent c
4. Apply actions to minimize differences in D.

SCT Implementation in Soar

We implemented SCT in the Soar cognitive architecture (Newell, 1990). Soar was connected to the GameBots virtual environment (Kaminka et al., 2002). Here, multiple agents, each controlled by a separate Soar process (each executing SCT) can interact with each other in a dynamic, complex, 3D virtual world (see Figure 1).

A detailed discussion of Soar's role as a cognitive architecture is beyond the scope of this paper. We provide a very brief overview here, and refer the interested reader to (Newell, 1990) for additional details.

Figure 1: **Soar agents in the GameBots environment. Each agent has limited field of view and range, and may move about and turn.**

Soar has two components: A graph-structured working memory, and a set of user-defined production rules that test and modify this memory. Efficient algorithms maintain the working memory by executing rules that match existing contents. All the agent's knowledge, sensor readings, and decisions are recorded in the working memory. Soar operates in a classic sense-think-act cycle, which includes a decision phase in which all relevant knowledge is brought to bear to propose, and then select, an *operator*, that will then carry out deliberate mental (and sometimes physical) actions. Once the operator finishes its actions, it is automatically de-selected (terminated), and the cycle repeats. Unlike simple production rules, whose effects on working memory are temporary, operator-induced the actions of rule firings on working memory (and in turn, on physical actions) are persistent, even after the operator has been de-selected. Overall, a Soar agent's behavior is the result of the sequential selection of operators, each performing an action on the environment and/or internal memory.

For our experiments, several basic task-oriented operators were implemented, to allow the agents to move about, turn towards each other, measure distances to others, etc. Thus one thread of control, always running, is in control of the agent's actions towards whatever tasks it was given.

SCT was implemented as a secondary parallel thread within Soar (Figure 2). Whereas normally, operators are proposed (and selected) by Soar based on their suitability for a current goal (e.g., through means-end analysis), in our agent operators were also proposed based on their suitability for SCT. In other words, at every cycle, a Soar agent would consider operators that advance it towards its goal. In our implementation, it would also consider operators that seek to minimize perceived differences to other agents.

Thus SCT-proposed operators compete with the task-oriented operators for control of the agent. This may appear to contradict Festinger's theorizing that social comparison comes into play only when people are at an impasse. However, this is not the case. By setting Soar's decision preferences to prefer SCT-proposed operators only when no task-oriented operators are available, one gets the behavior predicted by Festinger's theory. Further exploration of this issue is beyond the scope of this paper.

The SCT thread proposed operators by following the algorithm described previously, though in a way that is adopted

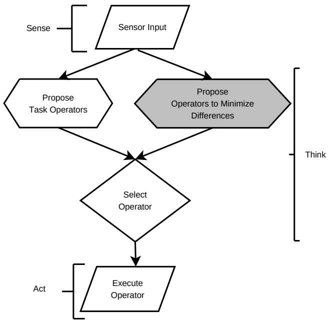


Figure 2: **The Soar sense-think-act decision cycle, SCT process highlighted.**

for Soar's decision cycle: At every cycle, for each observed agent and for each difference, the SCT process would propose an operator that would minimize the difference. Then, a set of preference rules is triggered that ranks the proposals based on feature weight. Additional rules prefer the most similar agent (that is still not sufficiently similar). Thus at the end, only one SCT operator is supported.

Here additional cognitive components became necessary. Suppose an agent X decided to turn towards the same angle as an agent Y that is next to it. Due to the limited field-of-view of X, it would lose track of Y once it makes the turn. From that point on, it could no longer keep track of Y, to minimize additional differences. This would cause it to become overly reactive, turning about immediately to seek Y again, or to select a different operator altogether (now that Y could no longer be imitated).

We thus found it necessary to utilize two mechanisms: (i) a memory mechanism that keeps track of the whereabouts of agents, once seen; and (ii) an exploration mechanism that occasionally would turn towards remembered agents, to provide an update on their state (for the purpose of comparison). Both of these mechanisms (memory and exploration) are of course present in many cognitive architectures, and are not necessarily linked to SCT. We thus leave discussion of such mechanisms outside of this paper.

Modeling Imitational Behavior

An attractive feature of social comparison is its hypothesized prevalence in human group behavior, i.e., its generality across different behaviors. Indeed, we believe that the SCT model we present in this paper is sufficiently general to account for a wide variety of group behaviors. This is in contrast to many existing computational models, that typically focus on specific tasks.

In previous work (Fridman & Kaminka, 2007) we evaluated the use of the SCT model in generation of pedestrian movement phenomena like bidirectional movement and movement in groups with and without obstacles. The SCT model accounts for group formation in pedestrians that are inter-related, a phenomenon not addressed by previous models. And where previous techniques apply, SCT shows improved results.

Here, we discuss in detail the implementation in Soar, and the evaluation methodology, providing additional evidence for such generality by describing the application of the SCT model to the problem of generating imitational behaviors in loosely-coupled groups. Unlike individual imitation, where one agent imitates a role model, crowd imitational behavior spreads across a group of individuals who dynamically select role models for imitation, from the level of observable actions to the level of unobservable internal mental attitudes (e.g., goals). Here, imitation occurs more loosely, as the role models do not necessarily intend to play their role, and indeed may not even know that they are being imitated. Also, the imitators potentially switch their role-model targets from one moment to the next. Psychology literature describes such imitational behavior as one of the keystones of crowd behaviors (Le Bon, 1895).

In order to simulate imitational behavior we used position and direction as the agents' feature set. For each observed agent and for every difference found, the SCT process proposes a corrective operator to be performed in order to minimize the difference in the selected feature. In this task, the corrective operators were 'move-to' (minimizing distance to the observed agent, correcting position differences) and 'turn-to' (imitating angle of the observed agent).

In addition to the proposed SCT operators, Soar also proposes operators based on their suitability for the current goal, and based on an exploration mechanism which proposes operators seeking new information. In this task, goal operators were 'turn-to' (a random angle); the exploration mechanism operators turned towards previously seen agents.

We used Soar preference rules to rank the feature weights such that the position feature gets higher priority than direction. This means that a closest agent is considered to be more similar, however the chosen feature for correction is direction. The S_{max} value was unbounded, which means that there is no such thing as too similar. In our case Soar can propose corrective operator with value equal to zero if there is no correction to make with respect to the observed agent. We used additional Soar preference rules to give higher priority to exploration mechanism operators than to goal operators. Thus, each agent prefers the SCT operators ('turn to') and in the case when there are no seen agents (i.e. there is no proposed SCT turn-to operator) an agent will prefer the exploration mechanism operators, and only afterwards the goal operators. The resulting simulated behavior has the agents standing in their initial locations, turning to some direction or doing nothing.

Evaluation of imitational behavior

We conducted experiments to evaluate whether SCT can indeed generalize to account for imitational behavior in groups. Unlike the pedestrian movement domain, where clear measures are available for objective measurement of the success of a model (e.g., flow, lane changes), imitational behavior does not have clear standards of evaluation.

We propose a method for evaluation of imitational behavior. We propose a questionnaire composed of general questions and specific tasks related questions. The general questions can be used as a common method for evaluation of all kinds of imitational behaviors. We rely on experiments with human subjects, which judged the human crowd behavior and the resulting SCT behavior in comparison to completely individual behavior (i.e., arbitrary decisions by each agent, independent of its peers), and to completely synchronized behavior (i.e., all agents act in complete unison).

The first hypothesis underlying the experiments was that groups controlled by SCT would generate behavior that would be ranked somewhere in-between the individual and perfect-coordination models, i.e., that SCT would generate behavior that would be perceived as coordinated, but not perfectly so. Another hypothesis is that human crowd behavior would also be ranked somewhere in-between the individual and perfect-coordinated behaviors.

To examine the first hypothesis, we created three screen-capture movies of 11 Soar agents in action. All movies were shot from the same point of view, and showed the agents in the same environment. In all screen-capture movies there is one blue agent that stands in front and turns up to 90° left or right. All others are red agents that act according to one of the models.

In one movie (*individual*), the red agents act completely independently of each other, randomly choosing an angle and turning to it. In another (*unison*), the red agents act in almost perfect coordination, turning towards the same angle as the blue agent almost instantaneously (small timing differences result from asynchronous responses of the simulated environment). Finally, in the *SCT* movie, the red agents act according to our model as described above.

These experiments were carried out using 12 subjects (ages: 18–40, mean: 28; male: 6; additional 4 subjects dropped due to technical reasons). Each subject was given a brief description of the appearance of the environment and agents, sometimes aided by a snapshot from a movie (e.g., as in Figure 1). The subjects were told that the purpose of the experiment was to evaluate the use of perception models embedded in the agents; that there was a red dot—visible to the agents but not to the subjects—that moves about on the walls surrounding the group. The agents' goal is to individually locate this dot, and then track it in place by turning around. The purpose of the cover story was to focus the attention of the subjects away from group behavior and imitation, so as to not bias the results. After the description, the movies were shown to the subject.

After each movie, the subjects were asked to fill a short questionnaire (described below) based on what they saw. Each movie was shown only once. The order of presentation of movies was randomly selected for each subject, to control for learning and order effects. The questionnaire included the following questions:

1. If there is only one red dot in the room, to what degree did

all agents see it? (1 - nobody saw the red dot; 6 - all agents saw it)

2. To what degree were the movements of the agents random? (1 - not random at all; 6 - very random)

3. To what degree was there cooperation between the agents? (1 - no cooperation at all; 6 - full cooperation)

4. To what degree was there agreement between the agents? (1 - no agreement at all; 6 - full agreement)

5. To what degree were the agents coordinated in terms of the direction of their movements? (1 -no coordination at all; 6 - fully coordinated)

6. How quickly did the agents find the red dot? (1 - dot not found at all; 6 - immediately found)

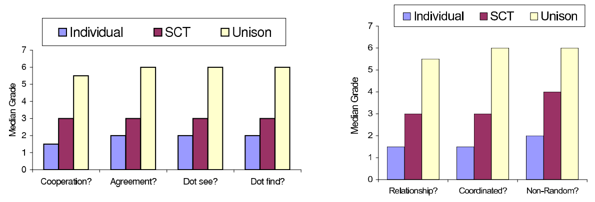
7. To what degree were the agents related to each other? (1 - no relation at all; 6 - tight relation)

8. Do you see any leaders? If so, how many? (1-11) (1- one leader; 11 - all agents are leaders, i.e., no leader).

In this experiment, the subjects were asked to grade the movies on an ordinal scale of 1–6, with 1 being a low score (typically associated with more individual behavior), and 6 being a high score (typically associated with perfect unison). In order to keep consistency in presentation of results, the scale of the second question (Non-Random) was reversed. The results of the last question (Number of leaders) are presented separately due to inconsistency in scale with other questions.

Agents results

In general, the responses to the questions in this experiment have placed SCT between the individual and unison models. Results are summarized in Figure 3(a) and 3(b). The questions in Figure 3(a) are associated with agents' performance on a given task. In the presented questionnaire the number of questions are 1, 3, 4 and 6. Figure 3(b) refers to more general questions (i.e. the same questions that were used in human crowd movie). In questionnaire the relevant numbers of questions are 2, 5, and 7. The categories in the X-axis correspond to questions given to the subjects. The Y-axis measures the median result. Each bar correspond to compared model and as explained above we compare SCT model to Individual and Unison models.



(a) (b)

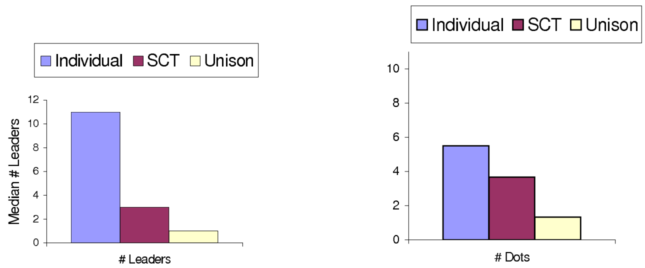
Figure 3: Results of questionnaire on agents performance.

The results clearly demonstrate that the SCT model lies in between the individual and perfect-unison model. While in some questions it appears to be somewhat closer to the individual model, it is significantly different from it at the $\alpha = 0.05$ significance level (t-test, one-tailed).

Figure 4(a) shows the results for the question on the number of leaders. The median result for the individual was 11 (i.e., every agent is a leader, or in other words, no leader). For the unison model, the median result was 1. For the SCT model, the median result was 3. In this question the SCT model result is very close to the Unison model. According to t-test (one-tailed) the SCT model significantly differs from the Individual model ($p = 0.02$). However, in comparison to Unison model there is no significance found ($p = 0.3$).

We conducted an additional experiment, in which static images—snapshots from the movies—were shown to subjects who were then asked how many red dots were present, based on the number of different directions in which agents were watching. The results of this experiment are summarized in Figure 4(b). Again the categories in the X-axis correspond to question given to the subjects. The Y-axis measures the average of median results that belong to each model.

Again the results demonstrate that the SCT model lies in between the individual and perfect-unison model and it significantly differs from the individual model ($p = 0.011$, t-test, one-tailed) and from perfect-unison model ($p = 0.012$, t-test, one-tailed).



(a) Number of leaders in screen-capture movies.

(b) Screen snapshot results.

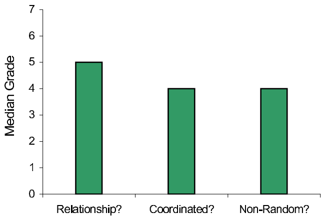
Figure 4: Additional results for the simulated agents.

Human crowd experiment

Another hypothesis underlying the experiments is that human crowd behavior would also be ranked somewhere in-between the individual and unison models. To examine this, we search for a human crowd movie where individuals perform the same action as in simulated agents movies. We used a news clip movie which shows people, grouped together, standing and waiting for some event to occur. The only action they perform in the movie is to turn occasionally.

This experiment was carried out using 12 subjects different than in the screen-capture movies experiments. Each subject, after viewing a human crowd movie (Figure 5(a)) was asked to fill the same questionnaire as in previous experiments. However, since in the human crowd movie there was no cover story about red dot, there were some irrelevant questions that were dropped out. The remaining questions are more general and not tied to a specific task.

(a) A human crowd.



(b) Results: News clip.

Figure 5: **Human crowd.**

Results are summarized in Figure 5(b). As in previous results, the categories in the X-axis correspond to questions given to the subjects and the Y-axis measures the median result.

We compare the human crowd results to the individual and perfect-unison models results. It appears to be significantly different from the individual model in all questions ($p = 0.000016$, $p = 0.000033$, and $p = 0.04$, respectively; t-test, one-tailed). However, in comparison to the perfect-unison model, the results of the coordination and non-random questions are significantly different($p = 0.0034$, and $p = 0.0003$, respectively; t-test, one-tailed). The results of the relationship question shows no significant different between the perfect-unison and the news-clip movie ($p = 0.44$).

In response to the question "Do you see any leaders? If so, how many?", the median result in human crowd movie was 1.5. It appears to be significantly different from the individual model ($p = 0.001$, t-test, one-tailed) but not in comparison to the perfect-unison ($p = 0.374$). When the subjects were asked to qualitatively discuss their answer to this question, many subjects reported that they don't see any leader, however "one must be present outside of the view of the movie, since the crowd is waiting for something or someone". However, when they were asked to refer to only people seen in the movie, the answer was that there were several subgroups in the seen crowd. While this qualitative answer is similar to the answer we received in asking similar questions about the simulation movies, we do not believe that this necessarily suggests that the SCT model is completely accounting for realistic behavior. In the future, we will focus more explicitly on the issue of subgroups, by adding the following question to the questionnaire: "Are there any subgroups? If so, how many?".

Summary and Future Work

This paper presented a model describing crowd behavior, inspired by Festinger's social comparison theory (Festinger, 1954). The model intuitively matches many of the characteristic observations made of human crowd behavior. We presented an implementation of SCT model in Soar cognitive architecture, for experiments in imitational behavior. Though there is a lack of objective data against which the model can be evaluated, results of experiments with human test subjects are promising and seem to match intuitions as to observed behavior. The subjects ranked SCT to be a middle-ground between completely individual behavior, and perfect synchronized ("soldier-like") behavior. Independently, human subjects gave similar rankings to a short news clip showing human crowds.

Acknowledgements. This research was supported in part by ISF grant #1357/07, and by IMOD.

References

Allport, F. H. (1924). *Social psychology*. Boston: Houghton Mifflin.

Blue, V. J., & Adler, J. L. (2000). Cellular automata microsimulation of bidirectional pedestrian flows. *Transportation Research Record*, 135–141.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 117–140.

Fridman, N., & Kaminka, G. A. (2007). Towards a cognitive model of crowd behavior based on social comparison theory. In *AAAI*.

Helbing, D., & Molnar, P. (1997). Self-organization phenomena in pedestrian crowds. In F. Schweitzer (Ed.), *Self-organization of complex structures: From individual to collective dynamics* (pp. 569–577). London: Gordon and Breach.

Helbing, D., Molnar, P., Farkas, I. J., & Bolay, K. (2001). Self-organizing pedestrian movement. *Environment and Planning B*, *28*, 361–384.

Kaminka, G. A., Veloso, M. M., Schaffer, S., Sollitto, C., Adobbati, R., Marshall, A. N., et al. (2002, January). GameBots: A flexible test bed for multiagent team research. *Communications of the ACM*, *45*(1), 43–45.

Le Bon, G. (1895). *The crowd: A study of the popular mind.* (1968 ed.). Dunwoody, Ga., N.S. Berg.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, Massachusetts: Harvard University Press.

Reicher, S. (2001). The psychology of crowd dynamics. In *Blackwell handbook of social psychology: Group processes.* Oxford: Blackwell.

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th annual conference on computer graphics and interactive techniques (SIGGRAPH-87)* (pp. 25–34). New York, NY, USA: ACM Press.

Tajfel, H., & Turner, J. C. (1986). The social identity theory in intergroup behavior. In *Psychology of intergroup relations.* Chicago: Nelson-Hall.

Yamashita, K., & Umemura, A. (2003). Lattice gas simulation of crowd behavior. In *Proceedings of the international symposium on micromechatronics and human science* (pp. 343–348).