Evolution of Player Skill in the America's Army Game

Il-Chul Moon
Mike Schneider
Kathleen M. Carley
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213, USA
imoon@andrew.cmu.edu
mike4@cmu.edu
carley@cs.cmu.edu

Many disciplines utilize computer games as interactive training simulations. However, their use is often limited to training mechanical skills, and they are not viewed as a sophisticated training tool with which to teach human interactions within organizations and social/organizational skills. Therefore, in this paper we examine how the players of the game America's Army changed their performance, play styles and social positions after one year of game play experience. For the initial investigation, we compare performance measures and play style at the beginning and the end of our survey period. Also, we calculate social network measures, such as centrality measures and clustering measures, to see how the social positions of the players change. After the comparison, we observe that players with one year of experience are more sensitive to communication and have tighter and denser communication networks around them.

Keywords: Interactive training simulation, evolution of player behavior, military training with simulations, social network analysis

1. Introduction

Computer games are increasingly being used for interactive training simulations. Numerous games have been developed for training purposes, and some existing games have been repurposed as training simulations. For example, a computer game Hazmat:Hotzone (see http://simopsstudios.com/), which was originally developed by the Entertainment Technology Center at Carnegie Mellon University in collaboration with the New York Fire Department, was developed to train fire fighters. Also, DOOM II, one of the most well-known first-person-shooting (FPS) games, was transformed by the United States (U.S.) Marine Corps Modeling and Simulation Management Office (MCMSMO) for use as a training tool for marines (see http://www.tec.army.mil/TD/tvd/survey/Marine_Doom.html).

The training goals of these simulations are so limited that they are only used for conveying mechanical knowledge or teaching skills that are easily demonstrated and repetitive. Also, to the best of our knowledge, it is commonly accepted that training simulations are not the best way for trainees to experience real person-to-person interaction. However, they can be very beneficial in giving some depth of knowledge about teamwork, group dynamics and ways to interact with team members. In addition, the increasing number of multiplayer online games suggests that it might be interesting to investigate how players develop social skills and teamwork through their game play.

Unlike the games intentionally developed for training purposes, the computer game America's Army (see http://www.americasarmy.com) was developed for recruitment and communication purposes, but it has also been extensively researched for its potential and effectiveness as a training tool [1]. In 2005, America's Army had more than 5.8 million registered players. Developed by the MOVES Institute for the U.S. Army, the game falls into the FPS genre, and was designed to portray a realistic portrait of squad-level combat in the U.S. Army [2, 3]. Each game consists of two teams, an offensive team and a defensive team, consisting of between one and 14 players each. A team can win the game by killing all of the opposing

SIMULATION, Vol. 82, Issue 11, November 2006 703-717

DOI: 10.1177/0037549706073689

players, or by accomplishing the particular goal for that mission, such as securing an oil pipeline, crossing a bridge, etc. The original role of America's Army was about the recruitment of young adults. In spite of its original purpose, and because of its very realistic features, Moon et al. [1] assessed and found its potential as a training tool, and it has been used effectively as a training tool for soldiers [3].

Although many aspects of America's Army have been researched, its influence on players has not been investigated thoroughly. Particularly, the interaction among team members may be an interesting aspect of how the game shapes the behavior of the players. For instance, if America's Army can teach good social skills and leadership to its players, it will demonstrate the value of America's Army not as a simple knowledge/skill training simulation, but as a more realistic combat training simulation. Therefore, in this paper, we investigate how players evolve their performance, play styles and social positions with various social network analysis measures.

2. Previous Research

2.1 Research on Interactive Training Simulation

Martens and Himmelspach [4] have pointed out that the value of combining intelligent tutoring systems and simulation has been realized for a long time. They have categorized the combination into three groups: interactive training simulations, demonstrative training simulations and character simulations. Interactive games fall into the category of interactive training simulations because players decide, act and learn in the context of the game. Also, there have been many attempts to use games as training simulations, although these trials have not been supported by any theory.

Among training simulation systems, there are some systems that implement detailed evaluation systems of their players. For instance, Auzende et al. [5] have introduced the Pedagogical Platoon Training System (PPTS), which aims to implement an evaluation environment for strategic and tactical skills during simulation training. The PPTS evaluates the simulation exercise while it is running and immediately generates the After Action Review (AAR) report when the exercise concludes. The system also shows expert knowledge related to the assessment result. However, the research does not show team interaction mechanisms and evaluation related to the players.

Freeman et al. [6] have reported on computer simulation training for Airborne Warning and Control System (AWACS) air weapons officers. In particular, they looked at the communication behavior of the trainees before and after the simulation training and found out that the training helped them to improve their communication skills. This research paper demonstrates a good strategy for evaluating the effectiveness of a training simulation. Therefore, we have attempted to use a similar approach to show how players of America's Army change their behavior.

The emphasis on the communication and organization of the players was, in part, introduced to the training simulation community by Salas and Cannon-Bowers [7]. They surveyed trends in training research and proposed the important tasks in the field. They have claimed that the progress in training research with regards to cognitive and organizational concepts is a newly emerging field. They have also stated that the new development promises to change how we conceptualize, design and institutionalize learning and training in organizations, and they have concluded that this training research field will require a deeper understanding of the cognitive and organizational aspects of training. Thus, it seems worthwhile to develop a framework that can examine how trainees interact with others and find their roles in the context.

2.2 Research on Military Training with Simulation

Training military personnel is a critical task for the U.S. Army, and the importance of training is becoming ever more critical because of the increasing technological sophistication of today's military. To tackle the increasing demands of training, the U.S. Army has developed several new methodologies. The most well-known approach is distance learning. According to Wisher et al. [8], distance learning has a number of merits: life-long education, minimum changes in daily life and tasks, and high cost-effectiveness. However, one area where distance training is currently not very effective is combat training and flight training.

First, flight training requires intensive mechanical knowledge learning and repetitive maneuver practices. Thus, some computer games are used instead of real flight training to reduce training cost and to increase the number of training sessions. For example, Herz and Macedonia [9] have reported that Microsoft® Flight Simulator is one of the most successful commercial-off-the-shelf (COTS) training simulations. According to Herz and Macedonia [9], the U.S. Navy distributed a customized Flight Simulator to all student pilots and undergraduates in Naval Reserve Officer Training Courses. They have also stated that there is research on the training value of Flight Simulator, and this research has confirmed that students who used Flight Simulator during early flight training tended to perform better in tests. Additionally, Flight Simulator has been used as a core component of other training simulations. For example, the U.S. Navy utilized the Operator Machine Interface Assistant (OMIA) system to teach operators about the new common-cockpit of MH-60R and MH-60S helicopters (see http://www.baseops.net/flightsimulators/), and Flight Simulator was integrated with the OMIA system for flight displays and other functionalities of these helicopters. Moreover, Flight Simulator was also used as a standalone training simulation during the project.

Not only flight training, but also combat skill training requires more real-world interaction than many other skills, and existing distance-training methods are not good at reproducing real-world interaction. It is increasingly recognized that FPS games are an alternative way to train combat soldiers [10, 11] (see also http://www.tec.army.mil/TD/tvd/survey/Marine_Doom.html). FPS games can simulate the combat area, weapon attributes and actions of the opposing soldiers with a high level of detail. These advantages of FPS games have long been recognized by the U.S. Army and U.S. Marine Corps (USMC), and, before the development of America's Army, two attempts were made to use an FPS game as a training tool for combat soldiers. The first FPS game used by the USMC was Marine Doom, a modification of Doom II, which was an early and popular FPS game. According to the MCMSMO (see http://www.tec.army.mil/TD/tvd/survey/Marine_Doom.html), Marine Doom has been used to train four-man fire teams. It teaches concepts such as mutual fire-team support, the protection of the automatic rifleman, the proper sequence of an attack, and so on. However, to the best of our knowledge, the effectiveness of Marine Doom for training has not been examined. The second FPS game used by the U.S. Army was Tom Clancy's Rainbow Six Rogue Spear (Rogue Spear) [11]. The U.S. Army used the game to train soldiers how to conduct a military operation in urban terrain. However, the game was not used for weapons training, but rather to help hone decision-making skills at the small-unit level. Specifically, the game was expected to teach soldiers three small-unit tasks: (i) how to prepare for a mission; (ii) how to work as a team during mission execution; (iii) how to conduct AARs. Although Rogue Spear does enable this type of training, we were not able to find any research papers evaluating the tool's effectiveness or realism.

Even though these two games were predecessors of the combat training game, their effectiveness and outcomes have not been well researched, and we conjecture that the reason for this is the relatively unrealistic features of the game play and limited usage of these games as training tools in military communities. However, their successor, America's Army, addresses these limitations. The most well-known case study of America's Army as a combat training tool is the research carried out by Farrell et al. [12]. Farrell et al. used America's Army as a land navigation simulator for training cadets who were taking a class in ground maneuvering. They have demonstrated the ability of America's Army to be a land navigation simulator but not its ability to be a training tool, because the training used other course materials and their research did not evaluate the realism of the game.

Other researchers have looked at America's Army as part of a framework for training soldiers, not merely a land navigation tool. Moon et al. [1] examined the realism of America's Army to see whether it has enough features to be used as a training tool for real-world soldiers. They compared America's Army to real-world research on squad-level activities and concluded that it has enough realism to match real-world combat training. However, they did not investigate the behavior of the players and its evolution.

The most interesting report about America's Army as a training tool has come from real-world use, not from a research center. Zyda [13] introduced the experimental use of America's Army at Fort Benning, GA. He has reported that a staff sergeant used America's Army to train new recruits who were having trouble with the rifle range or the obstacle course; these recruits passed the range test after playing America's Army. Although Zyda [13] has introduced an interesting episode in the real world, he has concentrated on the production of serious games, and not on the use of America's Army as a training tool, and there is no significant validation of the training given by America's Army. Additionally, it is discouraging that the reported use was limited to training of weapon characteristics and physical training. However, we suspect that America's Army may support a greater level of training. Therefore, in this paper, we evaluate the evolution of the players in terms of performance, play styles and social positions, in order to verify whether America's Army can really teach such sophisticated concepts to players.

3. Data Description

The log record datasets [14, 15] were collected twice in a one-year interval. The first dataset was recorded from 200 America's Army game servers over the course of 14 days in June 2004, and the second dataset from 138 servers over the course of 23 days in April 2005. Each line of the log files represents one event recorded by the servers.

There are always two teams per game, playing against each other. A team can have up to 14 players. While we can capture the initial performance measures and play styles with the dataset, we cannot identify the social position of the players in the teams directly from the dataset. Therefore, we reconstruct social networks with the dataset and with the "who-talked-after-whom" heuristic. This heuristic assumes that someone who talks after another person is likely responding to the previous speaker. With this assumption, the heuristic can capture the response network among team members (see Figures 1 and 2). We use the reconstructed social networks to evaluate the social position of a player in terms of centrality and other measures from the social network analysis field.

From the two datasets, we extracted the play records of 1829 players and analyzed them. The 1829 players are those who played one or more games in 10-men teams and

Table 1. Descriptive statistics for the two datasets

	Period 1	Period 2
Number of teams	502,796	484,544
Number of teams (size $\geqslant$ 10)	99,993	184,433
Number of players	73,497	91,322
Number of analyzed players	11,648	63,998
Number of overlapping players	1829	

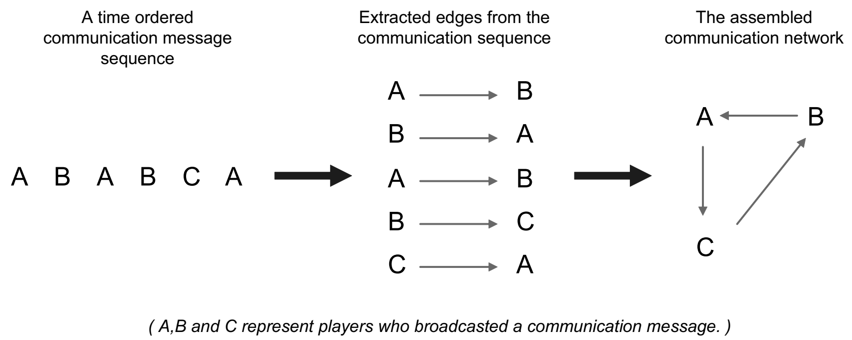

Figure 1. Schematic diagram showing how to create a communication network with a “who-talked-after-whom” heuristic

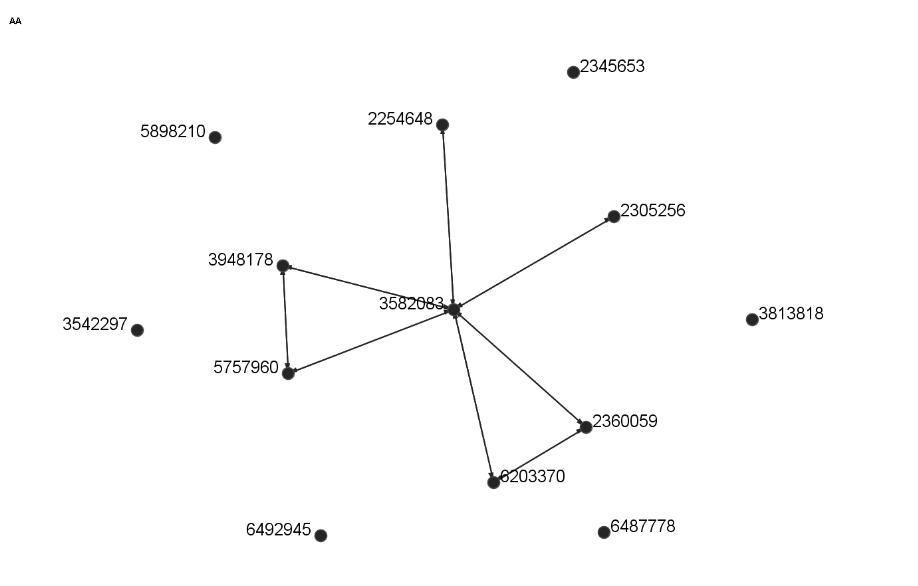

Figure 2. Diagram of the social network of a sample America’s Army team. Nodes represent the team members (numbers are unique IDs of players), and edges denote responses with Report-Ins

communicated through Report-Ins. We applied this regulation to make sure that the behaviors of all the analyzed players came from plays with many others. Also, some players may not know how to send or read the Report-In communication messages, so we analyzed only players who used that communication at least once. Additionally, because of the nature of our research, we examined players whose player IDs were in the two datasets at the same time. Finally, we assumed that all the records from the same player ID were made by only one player in the real world.

4. Research Method

To compare the behavioral changes of the players within our dataset, we set up three stages for comparison. First, we compared their behavior before and after, based on the initial performance measures. Secondly, for each player we compared the two play styles with regards to the amount of weapon fires, the amount of damage received or inflicted, and the frequency of different communication message types. Finally, we compared how the players positioned themselves in team communication networks using centrality and clustering measures. To verify the comparisons, we used paired t-tests, and we created bar charts to visualize the data. The paired t-test is used for the following reasons: (i) the performance measures of periods 1 and 2 are paired; (ii) the measure differences between the two periods are approximately normally distributed; (iii) the number of observed players (above 1800) is large enough. The t-tests can verify the average difference between the two groups with a significance number.

In addition, to see how the variance among the players can be explained differently, we used principal component analysis on the player style measures and social network measures [16–18]. With the comparison between the coefficients of the first principal components from the two periods, we see which factors are common or different among players. The measures related to social network analysis, such as centrality measures, clustering coefficients, Simmelian ties, etc., are calculated using the Organization Risk Analyzer [18].

5. Result

5.1 Evolution in Performance

To see the evolution of player performance, we calculated the survival rates, the average number of opponents killed and the winning rates (see Figure 3). The most distinct result of this examination is the changes in the survival rate and the average number of opponents killed. The survival rate decreased after one year of play, and the kill number increased. It seems strange that the analysis shows that the players became worse at surviving through the game after one year of play. However, when we consider the increased kill number, we can suggest that the players are more aggressive than previously. On the other hand, the increased winning rate demonstrates that with more experience they are better at winning a game, although they are more aggressive and more likely to die in the games. To verify this idea, we carried out paired t-tests with three hypotheses: that the players will have a lower survival rate, a higher kill number and a higher winning rate after a year. The results verified these hypotheses with high confidence (see Table 3).

The aggressiveness related to experience might reflect the nature of America's Army as a computer game. We can imagine that a soldier in the real world would be more willing to survive rather than die. However, in America's Army, the players tend to act more boldly because there is no serious penalty even though their avatars in the virtual world are dead. Therefore, this aspect of America's Army makes it less successful as a training tool for real military training.

5.2 Evolution in Play Styles

We examined the evolution of play styles with regards to the amount of inflicted/received damage, the volume of weapon fire and the frequencies of the three different communication types (Normal communication, Commo and Report-In). The first two factors, amount of damage and volume of weapon fire, present the damage management skills of the players, and the frequencies of communication show how players use communication differently as their experience grows.

First, the amounts of inflicted and received damage show a similar trend as the performance measures, as both amounts of damage increase. The other trend is the reduced volume of weapon fire after one year. The players shoot less frequently than previously, which is surprising because previous research has claimed that overwhelming weapon fire is a key to success [14, 15]. Thus, we conjecture that strategies for the experienced players and for others might be different. In the previous analyses, there is no discrimination between players in terms of experience, so the recommendations in the paper might be for inexperienced players who form the majority of the population under study. As we can see in Figure 4, the experienced players shoot less frequently, but are more deadly than previously. If we regard the players as trainees, this result gives us an insight into the evaluation of trainees of interactive training simulation. We have to treat the experienced and inexperienced trainees differently. Additionally, good strategies for the inexperienced players would not always be good for experienced players. We have confirmed these results with paired t-tests (see Table 3).

To understand communication usage, we have observed that players are more actively engaging with some communication types after one year's experience (see Figure 5 and Table 4). The players communicate with Commo (predefined messages that are mapped to a hot key) and Report-In more often than previously. The common feature of the two message types is the shortness of the messages and the ease of transmission. They are very short messages such as "roger", "negative", "move out", "roof top", "pipeline entrance", etc. We think that the experienced players like to communicate with others in short messages. Also, the ease of the transmission of these two communications may be a great advantage for the players. They can map the messages to one to three key input sequences, so they do not have to spend much time typing the messages. In contrast, the frequency of the Normal communication (which requires typing the message directly) decreased. When we consider the evolved communication usages of the players, we can see that they are learning to use communication styles which are fast and efficient.

5.3 Evolution in Social Positions

5.3.1 Social Positions Based on Centrality Measures

The above analyses show how the players evolve in their play styles and performance. Whilst we can perform these analyses without making any assumptions, we cannot

Table 2. Measures calculated with a one-year gap

Measure	Category	Meaning
Winning rate	Performance measure	Percentage of winning out of game participation
Survival rate	Performance measure	Percentage of survival out of game participation
Number of opponents killed	Performance measure	Average number of killed opponents during one game
Inflicted damage	Performance measure	Average amount of inflicted damage during one game
Received damage	Performance measure	Average amount of received damage during one game
Number of weapon fires	Basic play style	Average number of weapon fires during one game
Number of normal communications	Basic play style	Average number of Normal communications during one game
Number of report-ins	Basic play style	Average number of Report-Ins during one game
Number of commo communications	Basic play style	Average number of Commo communications during one game
In-degree centrality	Agent measure	Player's normalized in-flow communication edge
Out-degree centrality	Agent measure	Player's normalized out-flow communication edge
Eigenvector centrality	Agent measure	Degree of connected players who are themselves connected to many players
Betweenness centrality	Agent measure	Across all node pairs that have the shortest path containing the player, the percentage that pass through the player
Clustering coefficient	Agent measure	Density of the player's ego network, which is subgraph induced by its immediate neighbors
Triad count	Agent measure	Number of triads centered at the player
Simmelian ties	Agent measure	Number of ties with strongly, reciprocally connected players when there are one or more third-party players who commonly have strong and reciprocal edges to themselves and the connected player

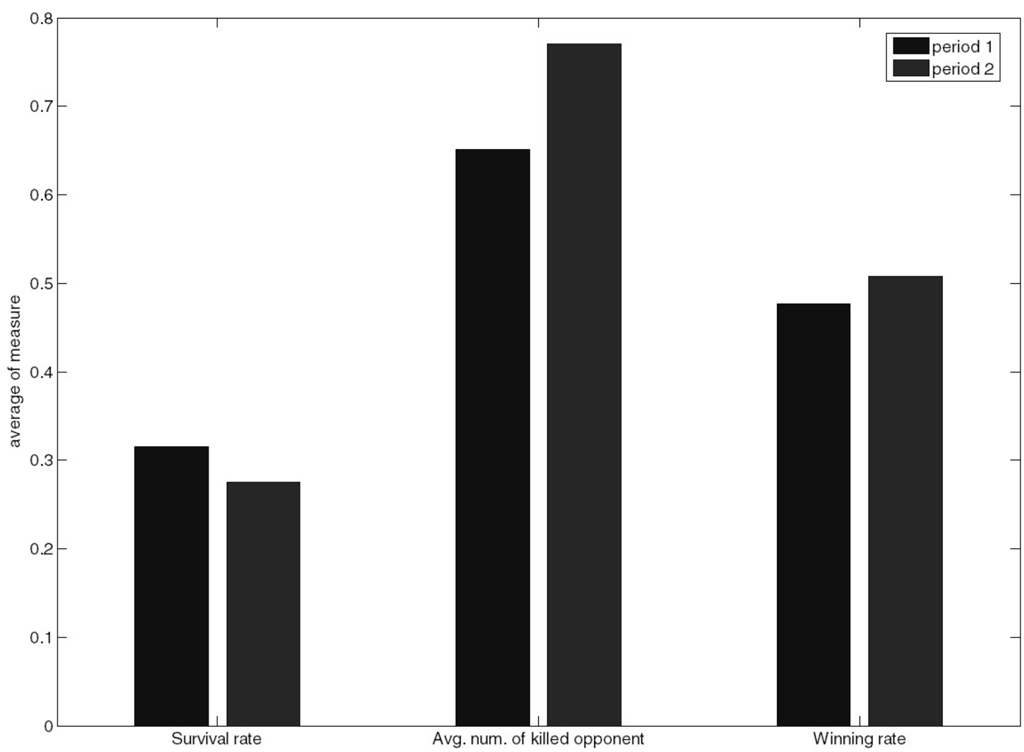

Figure 3. The performance measures of the players

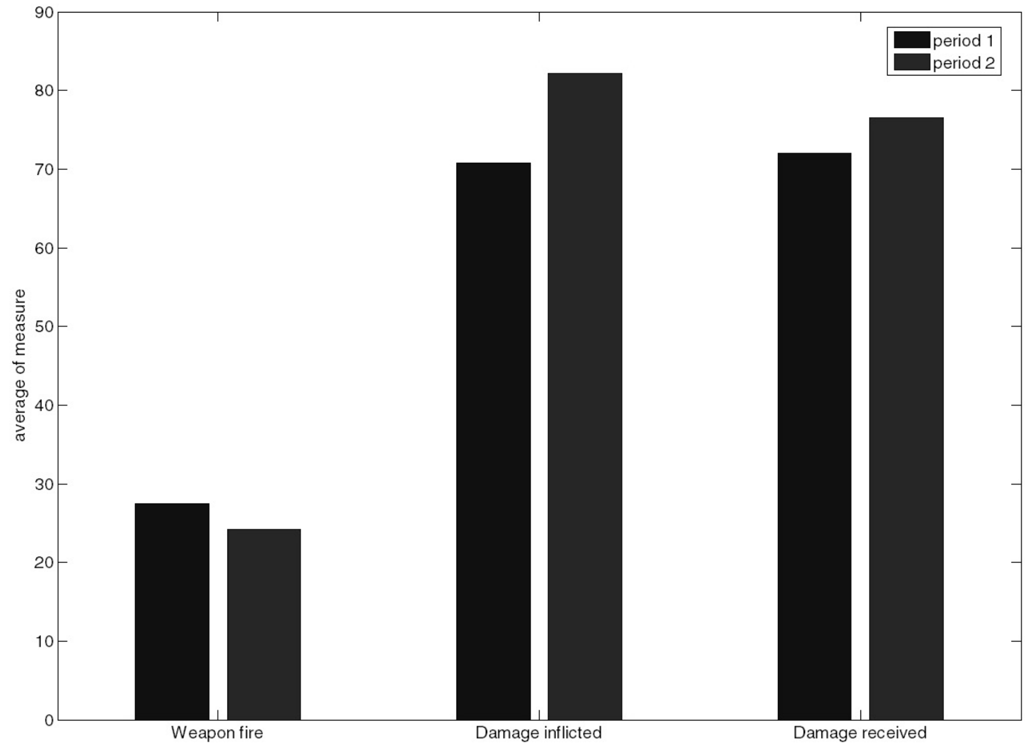

Figure 4. Play styles related to damage (average number of times a player fired a weapon, damage inflicted and damage received, in a game)

observe how the players communicate with team members, lead their teams, or take responsibility differently in a detailed manner. Therefore, we have reconstructed a communication social network for each team in our dataset, and we have calculated centrality measures that can tell us about the positions of the player in the networks.

The network centrality of most players increased with one year of experience, indicating that the players became more central to their teams. Also, the increment in in-degree and out-degree centrality means that they are more responsive to the communications of others. When we think of the edges on the social network as the information flow paths, we can argue that the players lead their team members by distributing and bridging critical information, such as where the team members are.

Particularly, it is noticeable that the increment in betweenness centrality is greater than any other centrality measure. It is possible that frequent communication usage may increase the centrality measures generally, but the extra increase of betweenness centrality shows that the players have tried to connect their team members as squad leaders do in the real world. Although the other centrality measures show clear evolution, the eigenvector centrality shows no difference. We conjecture that this result is caused by the other team members who are not experienced. The eigenvector centrality shows how well a player is connected to other well-connected players. Because the other team members may not be experienced, their centralities may remain unchanged. Therefore, the experienced member may have an unchanged eigenvector centrality because of the static centralities of others.

5.3.2 Social Positions Based on Clustering Measures

We can also observe the social positions of the players with clustering-related measures. If players have many clusters or cliques, they may be tightly connected to others and exchange communications frequently. The clustering coefficient for a node is the density of its ego network. In other words, if players have a high clustering coefficient, their neighbors are tightly connected to each other. This

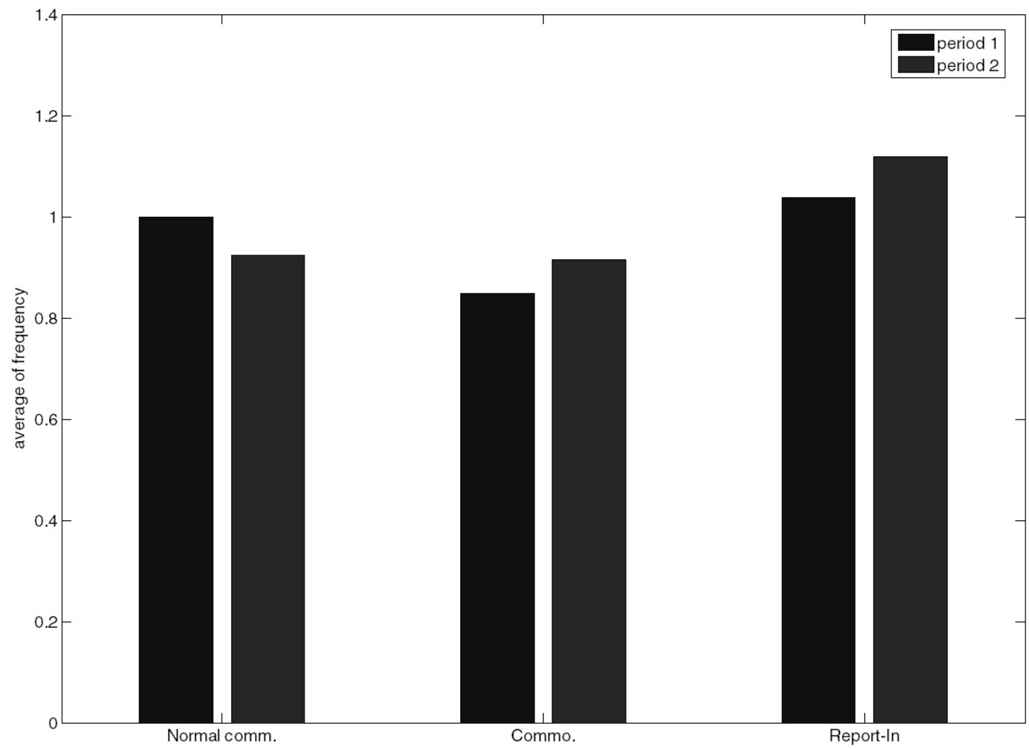

Figure 5. Play styles related to communication (average number of communications a player sent in a game)

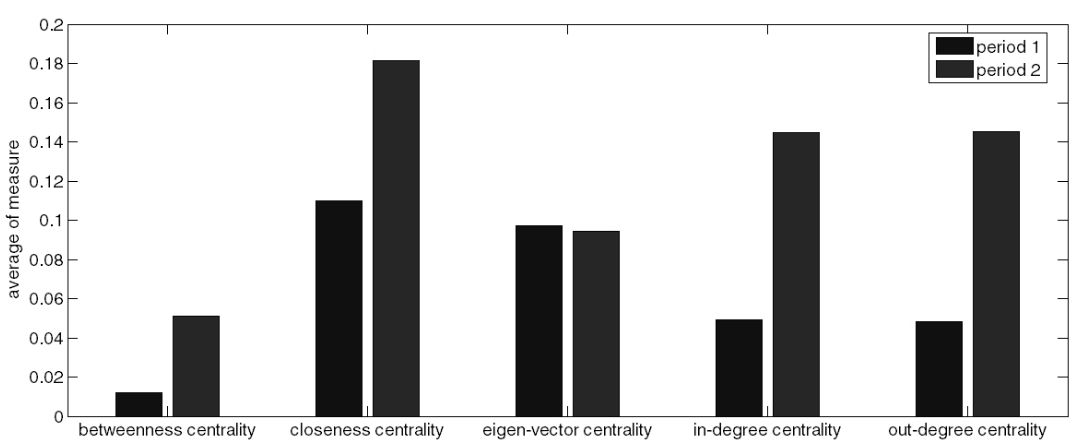

Figure 6. Centrality measures

Table 3. Paired t-test results for performance measures

	Period 1	Period 2
Survival rate		
Mean	0.315	0.275
Variance	0.090	0.051
t stat	4.710	
$P(T \leqslant t)$ one-tail	0.000	
Average number of killed opponents		
Mean	0.651	0.770
Variance	0.402	0.361
t stat	–6.231	
$P(T \leqslant t)$ one-tail	0.000	
Winning rate		
Mean	0.477	0.508
Variance	0.103	0.069
t stat	–3.231	
$P(T \leqslant t)$ one-tail	0.001	

Table 4. Paired t-test results for damage-related measures

	Period 1	Period 2
Weapon fire		
Mean	27.477	24.219
Variance	1744.796	439.114
t stat	3.195	
$P(T \leqslant t)$ one-tail	0.001	
Inflicted damage		
Mean	70.754	82.199
Variance	3794.059	3203.681
t stat	–6.313	
$P(T \leqslant t)$ one-tail	0.000	
Received damage		
Mean	72.065	76.478
Variance	782.560	435.725
$t$stat	–5.622	
$P(T \leqslant$ t) one-tail	0.000	

Table 5. Paired t-test results for communication-related measures

	Period 1	Period 2
Normal communication		
Mean	0.999	0.924
Variance	2.650	1.770
t stat	1.777	
$P(T \leqslant t)$ one-tail	0.038	
Commo		
Mean	0.848	0.915
Variance	2.110	1.711
$t$stat	–1.592	
$P(T \leqslant t)$ one-tail	0.056	
Report-In		
Mean	1.037	1.119
Variance	3.060	3.268
t stat	–1.551	
$P(T \leqslant t)$ one-tail	0.061	

means that not only does the player himself transmit many messages, but also the players around him exchange messages often.

Figure 7 shows how the clustering coefficients of the players increased. This result suggests that experienced players stimulate communications among their team members.

Furthermore, triad counts and Simmelian tie counts suggest that players with one year of experience make small and tightly connected groups more than they used to. For example, the triad count is the number of fully connected three-player groups, and the analysis result shows that the number of triads of players increased sharply after one year. This also demonstrates the increasing leadership of experienced players.

Simmelian ties are the links made when the two players are reciprocally, strongly connected, and to make the link of a Simmelian tie, there should be one or more third-party players who are reciprocally, strongly connected to the two players. This is a very difficult condition that can only be satisfied when there are three or more players connected very closely. The players with one year of play have almost eight times more Simmelian ties than they had previously. This clearly illustrates how the players are increasingly well connected to other team members.

These results consistently suggest that the players are learning the importance of communication and they are trying to establish a tight communication network. More importantly, the clustering coefficient suggests that the players' efforts make other players around them more connected. With this social network analysis, we can see how the interactive training changes the communication behavior of the players.

5.4 Overall Evolution

So far, we have analyzed measures from three perspectives: performance, play style and social position. Most of the measures show differences across the dataset, and we can interpret the changes based on the meaning of the measures. In this section, we look at the amount of change for each measure. We calculate the change rates by dividing the mean values of the first period by the mean values of the second period.

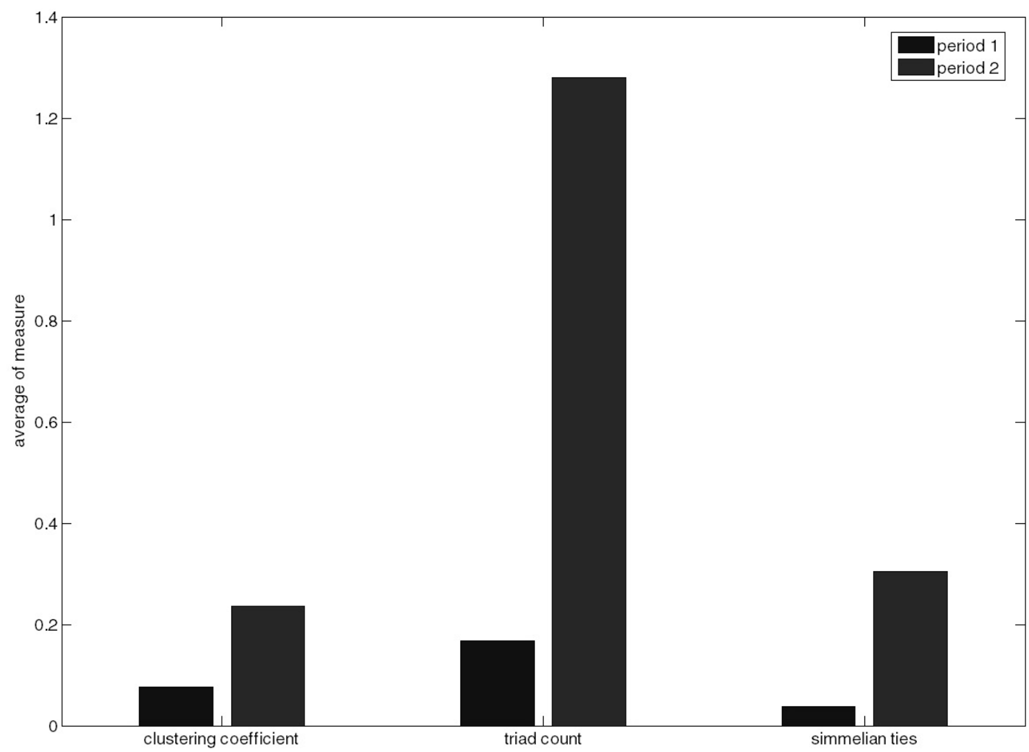

Figure 7. Clustering measures

According to Figure 8, the measures with the highest and second highest change rates are Simmelian tie count and triad count, respectively. This tells us that the experienced players act most differently in the area of subgroup organization. Additionally, the third largest change rate is in betweenness centrality. This shows that another area of change in player behavior occurred in communication behavior.

With this observation, we conjecture that America's Army teaches players not only the importance of communication, but also how they can organize it. Traditionally, interactive simulation training with FPS games focused on the training of weapon characteristics, mutual fire-team support, the protection of the automatic rifleman, the proper sequence of an attack and land navigation, etc. [5, 15, 17]. However, this result suggests that the players learned to use tight and frequent communication styles, which are not considered benefits of the interactive training simulation. We conjecture that players can learn these concepts because America's Army allows them to play with many unknown players with whom they must coordinate.

5.5 Variance among the Observed Players

Although we have verified each difference with paired t-tests, it may be interesting to see how much diversity there was in the way the observed players developed their skills. For example, our analyses show that the players learned communication and organizational skills from the experience, and this is true when we compare their averages. However, the average score does not show the amount of variance among the players before and after. If communication-related measures show a large variance, this means that the players have diverse communication strategies. To carry out this analysis, we use principal component analysis. Using this method, we compute a linear sum that maximizes the variance among the players. Then, we observe the coefficients for the first principal components and detect which factor is most important for explaining the variance. It should be noted that the absolute values of the coefficients are important, not the exact value, because we examine how much a single factor can contribute to explaining the variance.

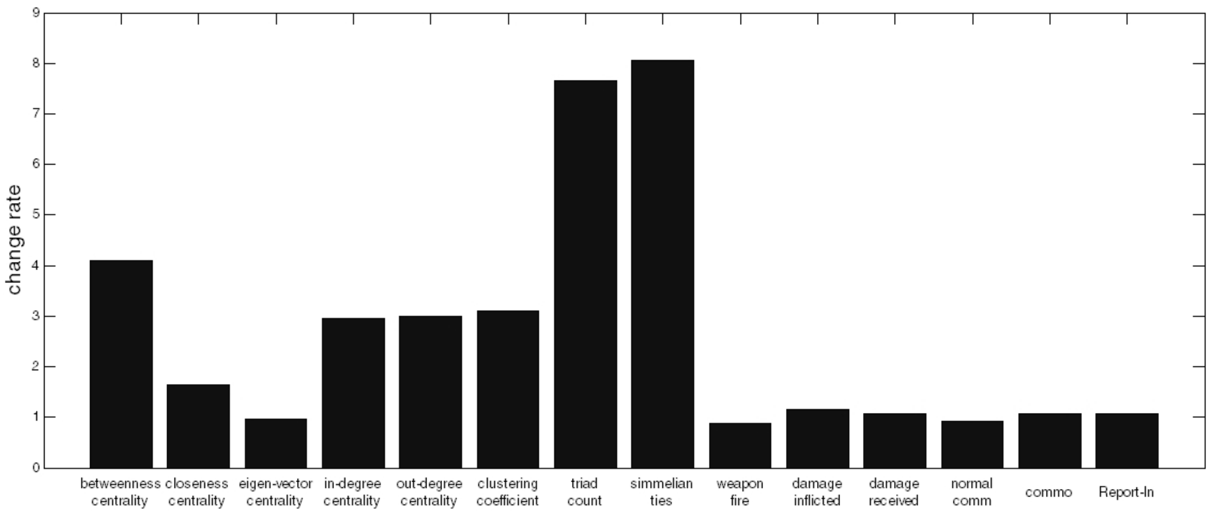

Figure 8. Change rate of measures

With the coefficients of the first principal component, we can see that some weights of the factors are changed. First, the largest and second largest amounts of change are from in-degree and out-degree centrality in both periods. This shows that the largest difference among the players is the tendency to respond and the frequency of this response to the communications of others, or their initiation of the communication.

However, the third largest coefficient comes from different factors. The third coefficient of the first period is from the number of Report-Ins, and that of the second period from the triad count. This illustrates that the players can be differentiated by the number of Report-Ins they transmitted in the first time period, but not during the second time period. In the second time period, we have to use a complex network measure, such as triad count, because the number of Report-Ins is too simple. We think that the players are experienced enough that most of them know the importance of Report-Ins and send these often to team members. However, it is still difficult to organize communication and send the Report-Ins at the right moment, so the players have different triad counts with a similar number of Report-Ins. To visualize this trend, we have drawn two scatter plots (see Figure 9). The first scatter plot shows that the triad counts of the players are consistently low, and the counts increase with many Report-Ins. However, the second scatter plot shows that the players can have high triad counts without sending more Report-Ins. Also, the triad counts of the players in the second scatter plot are more dispersed than in the first.

Moreover, the coefficients of the clustering-related measures are more highly ranked in the second period than the first in the principal component analysis. Thus, if we visualize the distribution of the players in terms of triad counts and Simmelian tie counts, the graphs will show to what extent the players evolved diversely after one year of game play. Figure 10 shows two scatter plots with clustering measures. During period 1, there are not many differences among the players. However, the second scatter plot clearly demonstrates that the players show various levels of Simmelian tie counts and triad counts.

This result suggests that the players become more diverse in terms of communication and organizational skills with more experience. At the same time, simple measures such as the number of Report-Ins become less effective for distinguishing players and evaluating them.

6. Conclusion and Discussion

Interactive training simulation is extensively used in military training. Squad-level training has been carried out with simulations that are developed or transformed versions of computer games. However, there has been no careful investigation of how trainees change their behavior or what types of knowledge and skills the simulations can teach trainees. Therefore, in this paper we have investigated these issues using a dataset from America's Army.

From the analyses, we can look at the evolution of players after one year in three perspectives: performance, play styles and social positions. In performance, we detected that the players were becoming more aggressive as they

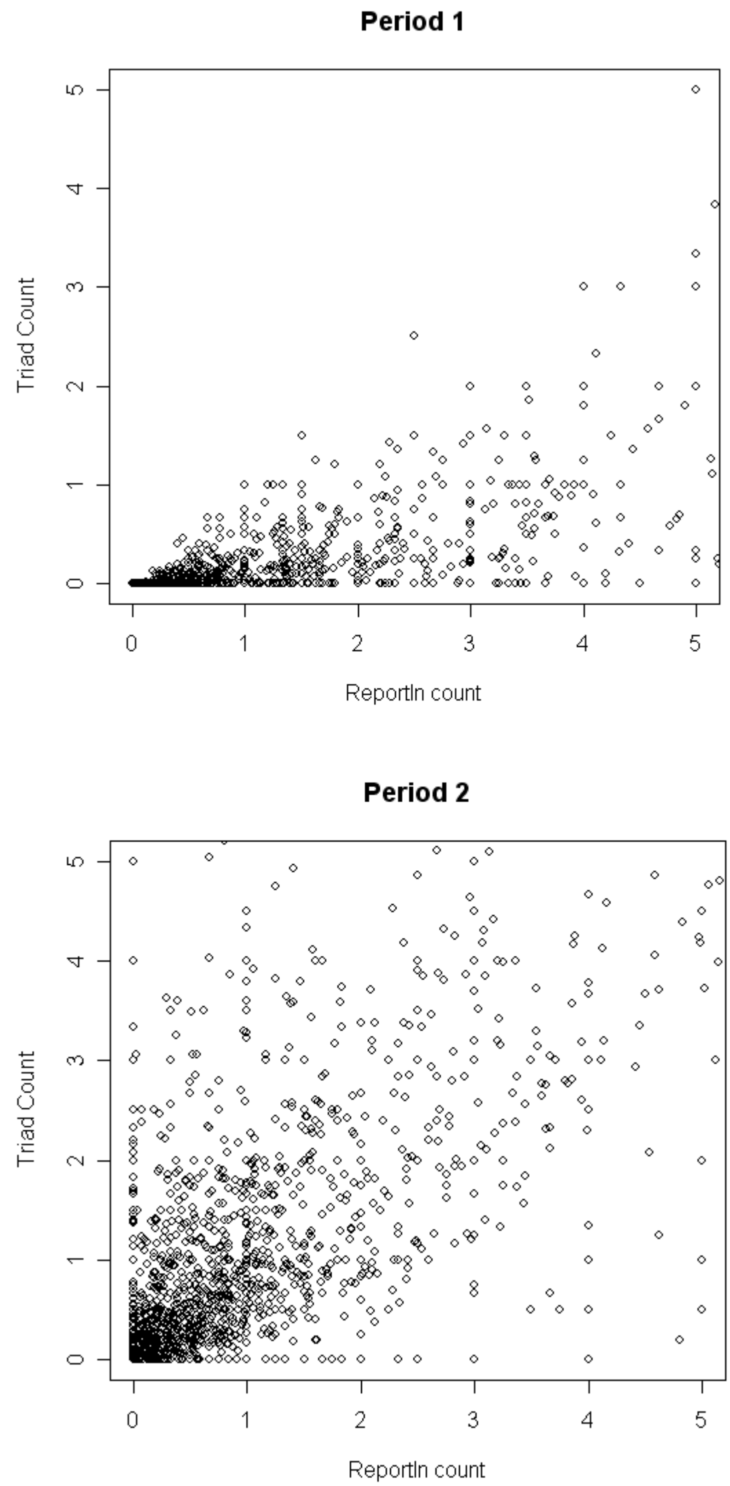

Figure 9. Scatter plots of average triad counts and average Report-In counts. Each dot represents one observed player. The plots show only the section where triad and Report-In counts are from 0 to 5

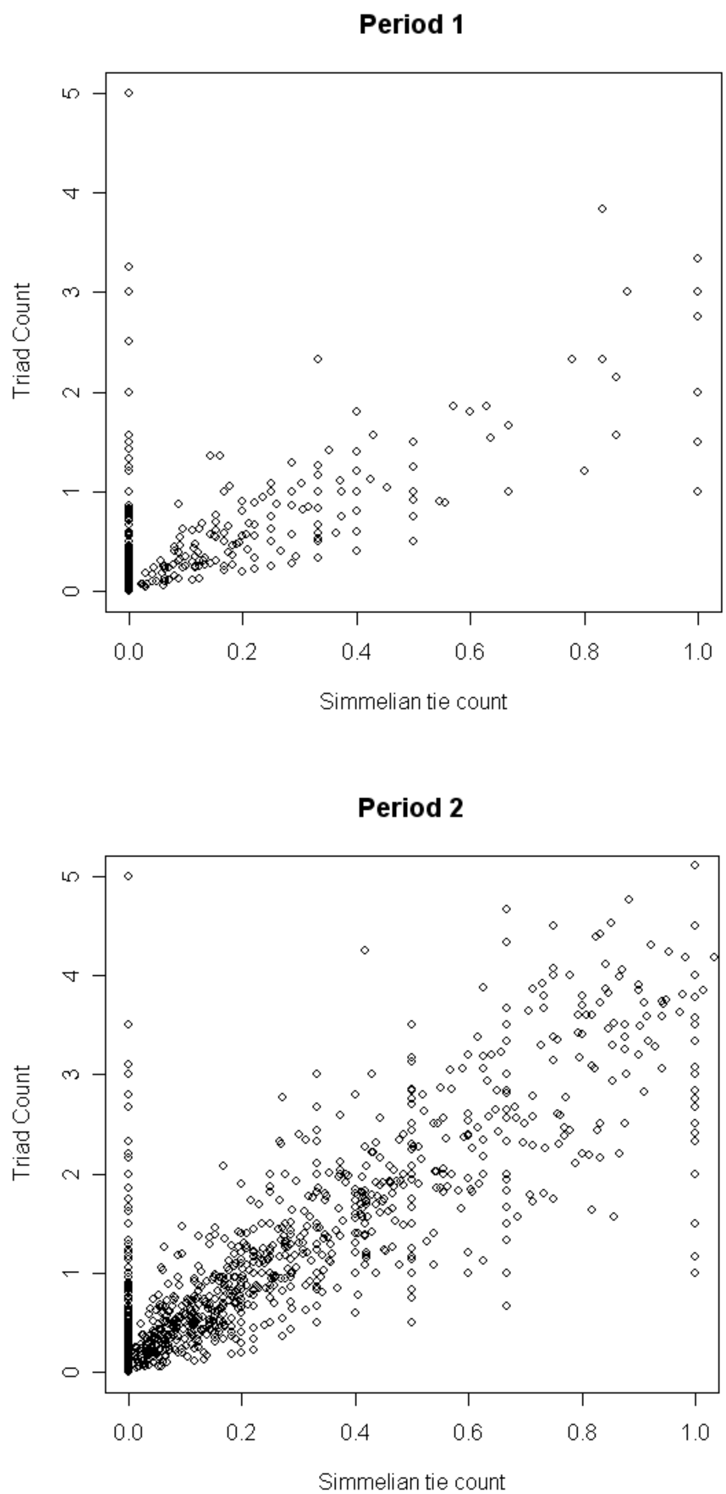

Figure 10. Scatter plots of average triad counts and average Simmelian tie counts. Dots represent players. The plots show only the section where triad count is from 0 to 5 and Simmelian tie count is from 0 to 1

Table 6. Paired t-test results for centrality measures

Betweenness centrality		
	Period 1	Period 2
Mean	0.012	0.051
Variance	0.001	0.003
t stat	–30.180	
$P(T \leqslant t)$ one-tail	0.000	
Closeness centrality		
	Period 1	Period 2
Mean	0.110	0.181
Variance	0.001	0.004
t stat	–43.853	
$P(T \leqslant t)$ one-tail	0.000	
Eigenvector centrality		
	Period 1	Period 2
Mean	0.097	0.094
Variance	0.014	0.003
t stat	1.118	
$P(T \leqslant t)$ one-tail	0.132	
In-degree centrality		
	Period 1	Period 2
Mean	0.049	0.145
Variance	0.005	0.011
t stat	–35.861	
$P(T \leqslant t)$ one-tail	0.000	
Out-degree centrality		
	Period 1	Period 2
Mean	0.048	0.145
Variance	0.004	0.011
Observations	1829.000	1829.000
$P(T \leqslant t)$ one-tail	0.000	

Table 7. Increase rate for centrality measures

	Increase rate
Betweenness centrality	4.096
Closeness centrality	1.648
Eigenvector centrality	0.967
In-degree centrality	2.957
Out-degree centrality	3.001

played. Although their winning rate increased, their received damage and inflicted damage also increased. This may reflect a limitation of America's Army as a training simulation because real-world soldiers will try to survive, rather than be aggressive and thus be more likely to be killed at the same time.

However, other analyses consistently show that after one year of experience players learned the importance of communication and new ways of organizing a communication network around themselves. For example, the number of desirable communication messages, such as Report-In and Commo, increased and the frequency of Normal communication, which is less efficient, decreased. This suggests that players learned which communication types are more effective in the game.

Table 8. Paired t-test results for clustering measures

Clustering coefficient		
	Period 1	Period 2
Mean	0.076	0.237
Variance	0.019	0.032
t stat	–32.229	
$P(T \leqslant t)$ one-tail	0.000	
Triad count		
	Period 1	Period 2
Mean	0.167	1.280
Variance	0.215	3.592
t stat	–24.894	
$P(T \leqslant t)$ one-tail	0.000	
Simmelian ties		
	Period 1	Period 2
Mean	0.038	0.304
Variance	0.031	0.265
$t$stat	–21.453	
$P(T \leqslant t)$ one-tail	0.000	

Furthermore, social network measures clearly demonstrate the evolution of players' communication behavior. Most centrality measures increased in our period of study, showing that experienced players displayed more leadership of their teams and that they are more responsive to the communications of others. When we only observe a simple count of communication messages, we cannot see whether the players are actively engaging the team in communication or simply sending out messages. However, using social network analysis, we can confirm that their evolution consists not just of sending of many messages, but deliberately bridging communications among team members.

Finally, the clustering-related measures show that the communications around experienced players are tighter and denser than previously. The clustering coefficient, triad count and Simmelian tie count all increased. Therefore, we conclude that experienced players had an increased ability to stimulate communication among their teammates.

Interactive training simulations have concentrated on teaching some mechanical skills or specific knowledge. However, we propose that some training simulations will be improved by allowing multiple trainees to connect to the simulations in order to stimulate cooperation among them. We have shown that FPS games can not only teach procedural knowledge, such as weapon use and navigation, but also valuable "soft" skills, such as leadership, communication, organization and teamwork.

Table 9. Coefficients of the first principal components

Variable Name	Period 1	Rank 1	Period 2	Rank 2
Betweenness centrality	–0.330	5	0.278	6
Closeness centrality	–0.329	6	0.308	5
Eigenvector centrality	–0.174	9	0.245	9
In-degree centrality	–0.390	2	0.367	2
Out-degree centrality	–0.393	1	0.369	1
Clustering coefficient	–0.279	7	0.227	11
Triad count	–0.342	4	0.329	3
Simmelian ties	–0.268	8	0.319	4
Weapon fire	–0.066	13	0.130	13
Damage inflicted	–0.141	10	0.144	12
Damage received	0.000	14	–0.018	14
Normal communication	–0.130	11	0.267	7
Commo	–0.094	12	0.242	10
Report-In	–0.361	3	0.261	8

7. Acknowledgments

This work was supported in part by the Defense Advanced Research Projects Agency (DARPA) and the Office of Naval Research for research on massively parallel online games and dynamic network analysis. Additional support was provided by the Center for Computational Analysis of Social and Organizational Systems (CASOS) at Carnegie Mellon University. The views and conclusions contained in this paper are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of DARPA, the Office of Naval Research or the U.S. Government.

8. References

[1] Moon, I., M. Schneider, and K. M. Carley. 2005. Investigation of the realism of the on-line game – America's Army. Working Paper, CASOS, Carnegie Mellon University, Pittsburgh, PA.

[2] Davis, M., R. Shilling, A. Mayberry, J. McCree, P. Bossant, S. Dossett, C. Buhl, C. Chang, E. Champlin, T. Wiglesworth, and M. Zyda. 2003. Researching America's Army. *Design Research: Methods and Perspectives*. Cambridge, MA: MIT Press, pp. 268–75.

[3] Zyda, M., A. Mayberry, J. McCree, and M. Davis, M. 2005. From VizSim to VR to games: How we built a hit game-based simulation. In *Organizational Simulation: From Modeling and Simulation to Games and Entertainment*, W. B. Rouse and K. R. Boff, eds. New York: Wiley.

[4] Martens, A., and J. Himmelspach. 2005. Combining intelligent tutoring and simulation systems. In *Proceedings of the International Conference on Human–Computer Interface Advances for Modeling and Simulation (SIMCHI'05)*, New Orleans, LA, pp. 65–70.

[5] Auzende O., M. Joab, F. Riviere, P. Le Leydour, and M. Futtersack. 2004. A knowledge-based generator for after action review interactive documents. In *Proceedings of the International Conference on Computer Aided Learning in Engineering Education*, Grenoble, France, pp. 149–54.

[6] Freeman, J., F. J. Diedrich, C. Haimson, D. E. Diller, and B. Roberts. 2003. Behavioral representations for training tactical communication skills. In *Proceedings of the 12th Conference on Behavior Representation in Modeling and Simulation*, Scottsdale, AZ.

[7] Salas, E., and J. A. Cannon-Bowers. 2001. The science of training: A decade of progress. *Annual Review of Psychology* 52:471–99.

[8] Wisher, R. A., and T. M. Olson. 2003. The effectiveness of Web-based training. Research Report 1802. Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

[9] Herz, J. C., and M. R. Macedonia. April 2002. Computer games and the military: Two views. *Defense Horizons* 11. Washington, DC: Center for Technology and National Security Policy, National Defense University.

[10] Psotka, J., B. Black, and D. Horn. 2004. Symposium on PC-based simulations and gaming for military training. Research Product 2005-01. Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

[11] UBISOFT press release. 29 August 2001. UBISOFT licenses Tom Clancy's Rainbow Six Rogue Spear game engine to train U.S. soldiers. San Francisco, CA: UBISOFT. http://www.gamespot.com/dreamcast/action/rainbowsix/news.html?sid=2809208.

[12] Farrell, C. M., W. K. Klimack, and C. R. Jacquet. 2003. Employing interactive multimedia instruction in military science education at the U.S. Military Academy. In *Proceedings of the Interservice/Industry Training, Simulation, and Education (I/ITSEC) Conference*, Orlando, FL.

[13] Zyda, M. 2005. From visual simulation to virtual reality to games. *IEEE Computer* 38(9):25–32.

[14] Carley, K. M., I. Moon, M. Schneider, and O. Shigiltchoff. 2005. Detailed analysis of factors affecting team success and failure in the America's Army game. Technical Report, CMU-ISRI-05-120. Pittsburgh, PA: CASOS, Carnegie Mellon University.

[15] Moon, I., K. M. Carley, M. Schneider, and O. Shigiltchoff. 2005. Detailed analysis of team movement and communication affecting team performance in the America's Army game. Technical Report, CMU-ISRI-05-129. Pittsburgh, PA: CASOS, Carnegie Mellon University.

[16] Bonacich, P. 1987. Power and centrality: A family of measures. *American Journal of Sociology* 92(5):1170–82.

[17] Krackhardt, D. 1998. Simmelian tie: Super strong and sticky. In *Power and Influence in Organizations*, R. Kramer and M. Neale, eds. Thousand Oaks, CA: Sage, pp. 21–38.

[18] Reminga, J. and K. M. Carley. 2004. ORA: Organization Risk Analyzer. Technical Report, CMU-ISRI-04-106. Pittsburgh, PA: CASOS, Carnegie Mellon University. Available online at http://www.casos.cs.cmu.edu/projects/ora/index.html.

***Il-Chul Moon** is a PhD student at the Computational Analysis of Social and Organizational Systems (CASOS) laboratory,*

Carnegie Mellon University. His research interests include agent-based social simulation, dynamic network analysis, machine learning, artificial intelligence, etc.

Mike Schneider *is a research programmer and PhD student at Carnegie Mellon University. He specializes in social network analysis and user interface design and implementation.*

Kathleen M. Carley *(PhD, 1984, Harvard) is a Professor of Computer Science at the Institute of Software Research International and the director of the Center for Computational Analysis of Social and Organizational Systems (CASOS). Her research combines cognitive science, social networks and computer science to address complex social and organizational problems in areas such as dynamic network analysis, computational social and organization theory, adaptation and evolution, text mining, and the impact of telecommunication technologies and policy on communication, information diffusion, disease contagion and response within and among groups, particularly in disaster or crisis situations. She and her lab have developed infrastructure tools for extracting and analyzing large-scale dynamic networks (AutoMap and ORA) and various multi-agent simulation systems in areas such as counter-terrorism (DyNet), network evolution (Construct), BioDefense (BioWar), and organizational management (OrgAhead). Dr Carley is the founding co-editor with Al. Wallace of the journal Computational Organization Theory and has co-edited several books in the computational organizations and dynamic network area.*