

Training with Automated Agents Improves People's Behavior in Negotiation and Coordination Tasks[☆]

Raz Lin^a, Ya'akov (Kobi) Gal^{b,d}, Sarit Kraus^{a,c}, Yaniv Mazliah^b

^a*Department of Computer Science, Bar-Ilan University, Ramat-Gan 52900, Israel*

^b*Department of Information Systems Engineering, Ben-Gurion University of the Negev, Be'er-Sheva 85104, Israel*

^c*Institute for Advanced Computer Studies, University of Maryland, College Park MD 20742, USA*

^d*School of Engineering and Applied Sciences, Harvard University, Cambridge MA 02138, USA*

Abstract

There is inconclusive evidence whether practicing tasks with computer agents improves people's performance on these tasks. This paper studies this question empirically using extensive experiments involving bilateral negotiation and three-player coordination tasks played by hundreds of human subjects. We used different training methods for subjects, including practice interactions with other human participants, interacting with agents from the literature, and asking participants to design an automated agent to serve as their proxy in the task. Following training, we compared the performance of subjects when playing state-of-the-art agents from the literature. The results revealed that in the negotiation settings, in most cases, training with computer agents increased people's performance as compared to interacting with people. In the three player coordination game, training with computer agents increased people's performance when matched with the state-of-the-art agent. These results demonstrate the efficacy of using computer agents as tools for improving people's skills when interacting in strategic settings, saving considerable effort and providing better performance than when interacting with human counterparts.

Keywords: training, automated agents, automated negotiation, coordination.

[☆]Preliminary results of this research were published in the proceedings of AAMAS 2009.

1. Introduction

Settings in which people and computers make decisions together arise in a wide variety of application domains (e.g, hospital care-delivery systems, systems administration applications) as well as in virtual reality and simulation systems (e.g, disaster relief, military training). The automated computer agents in these settings are designed for the purpose of supporting people, acting as proxies for individuals or organizations, or working autonomously. However, there is scant work on the influence of autonomous agents on people's behavior.

The evidence on the use of computer agents to change people's behavior in strategic settings is inconclusive. On the one hand, autonomous agents designed by researchers and students commonly use opponent modeling, game theoretic reasoning and machine learning, approaches that allow them to perform successfully in their respective setting [13]. On the other hand, when deciding whether to cooperate, people prefer to cooperate with other people rather than with computer agents. In particular, people have been shown to offer less to computer agents when making agreements than to people [24].

To address this gap, we study the question of whether using automated agents to train people can improve people's performance in two representative settings involving negotiation and coordination among multiple participants. We propose two methods for training people in these settings that are evaluated empirically in extensive experiments. The first training method involves people practicing a given task with other participants (whether other people, or computer agents that are designed by researchers and students). The second method involves people designing an automated agent to serve as their proxy in the given task. We compared the efficacy of these approaches by measuring people's behavior during training with that of their performance during a separate testing phase conducted on the same task. A challenge to evaluating people's performance in these multi-participant task is that their behavior depends in part on the strategies of the other participants. We therefore used a standardized agent to interact with people when comparing between their performance in the testing phase. This agent was chosen from the state-of-the-art in each of the respective settings, meaning that its proficiency was already demonstrated when interacting with other computer agents (or people) in separate studies. The use of the standardized agent provided an objective metric with which to evaluate people's performance.

Our empirical methodology consists of three settings. The first two consisted of different types of strategic multi-attribute bilateral negotiation tasks of imperfect information. The first simulated a job interview between an employer and

candidate, while the second simulated diplomatic negotiations (preliminary results on the first setting were published by Lin et al. [16]). In both cases, an agreement consisted of an assignment of possible values for each of the attributes, and the negotiation was conducted using an alternating offer protocol. The third setting was purely competitive and consisted of a three-player multi-round coordination game commonly used in the literature to evaluate computer agents [25]. We compared people's performance in these settings under some or all of the following training conditions:

- classical role playing (training) with another human counterpart;
- training with an automated agent;
- designing and coding an automated agent to act as a proxy.

During the testing and training phase, subjects were not told that they were interacting with an agent. Thus, any difference in their behavior can be attributed to the history of their prior interaction in the training phase.

Results showed that training with state-of-the-art agents helped people improve their performance for all role contingencies in the job candidate and the coordination setting and all role contingencies but one in the diplomatic negotiation setting. Training with agents designed by the subjects themselves improved their performance for all role contingencies in the job candidate and the coordination setting, but had a negligible effect in the diplomatic negotiation setting. Further analysis revealed that in the coordination game, training with people improved the performance of those people that coordinated more often with the standardized agent.

These results have insight for agent designers for human-computer decision-making as well as social scientists. They suggest that in settings requiring coordination and agreements, people can learn to be more skillful by learning to play from computer agents. These agents can be used as tools for training people in such tasks. This can result in considerable savings in cost and effort as compared to using people for training purposes.

The remainder of the paper is organized as follows. In Section 2 we review related work focused on the evaluation of training methods and the use of simulation and role-playing for training. Sections 3 and 4 present experiments and results for the negotiation and coordination settings in our study. Finally, we conclude the paper with open questions and future directions for research.

2. Related Work

We first discuss related work relating to training people to perform negotiation tasks. The use of simulations and role-playing is common for training people in negotiations (e.g., the Interactive Computer-Assisted Negotiation Support system (ICANS) [22], the InterNeg Support Program for Intercultural REsearch (INSPIRE) [9] and virtual humans for training [8]). Surprisingly, little research has been conducted that measures the effect of simulations and role-playing directly on people's negotiation skills, despite underlying assumptions that role-playing improves people's negotiation skills [21, 5]. Specifically, several works have evaluated the role of simulation in training students' skills as diplomatic negotiators using questionnaires and subjective reporting [20, 4]. Susskind and Corburn [21] study the usefulness of negotiation simulations by questioning leading practitioners in the field about why and how they use simulations to teach negotiation. Kenny et al. [8] and Traum et al. [23] have used virtual humans to facilitate people's negotiation, leadership and interviewing skills. These virtual humans were tested in several negotiation scenarios in social and military contexts in which culture plays a crucial role. Lennon et al. [11] have studied the extent to which training improves people's negotiation skills across cultures, as measured by their performance in a post-training negotiation task. There is no prior work that uses automated agents for the purpose of improving human performance in negotiation.

Another strand of research has studied the role of media, GUIs and decision support tools on people's negotiation behavior. Ross et al. [19] and Butler [2] studied whether watching negotiation simulations on video helped students increase their learning of negotiation concepts, as measured by students' reaction to the video and their ability to recognize pivotal points in the negotiation process. Other works have studied the role of web-based GUIs for facilitating negotiation [12, 9]. None of these methods have measured the effect of these support tools on people's performance in real time.

The use of automated agents in human-computer negotiations is a burgeoning field in Artificial Intelligence. For a comprehensive summary, see the survey by Lin and Kraus [13]. Most work in this field has focused on the design of agents that can reach more beneficial agreements than do people [3, 6, 10, 15]. Notable exceptions include Kamar et al. [7] who designed a computer agent that used collaborative decision-making strategies to interact with people in a cooperative game, and Bachrach et al. [1] who showed that agents playing strategies that implement solution concepts from cooperative game theory can play well with

people in a weighted voting game. None of these works have studied the effect of prior play in coordination games on people’s performance.

3. Training Methods in Bilateral Negotiation

In this section we study whether role-playing with people or training with automated agents can enhance the negotiation experience by improving the negotiation skills of human negotiators.

3.1. The Bilateral Negotiation Settings

Following Lin et al. [15] we consider a bilateral negotiation settings in which two agents, either automated negotiators or people, negotiate to reach an agreement on conflicting issues with uncertainty, expressed by the fact that the exact score function of the rival is private information.

The negotiation can end either when (a) the negotiators reach a full agreement, (b) one of the agents opts out, thus forcing the termination of the negotiation with an opt-out outcome (*OPT*), or (c) a predefined deadline is reached, whereby, if a partial agreement is reached it is implemented or, if no agreement is reached, a status quo outcome (*SQ*) is implemented. Let I denote the set of issues in the negotiation, O_i the finite set of values for each $i \in I$ and O a finite set of values for all issues ($O_1 \times O_2 \times \dots \times O_{|I|}$). We allow partial agreements, $\perp \in O_i$ for each $i \in I$. Therefore an offer is denoted as a vector $\vec{o} \in O$. Since no agreement is worse than any agreement, and a status quo is implemented if the deadline is reached, we assume that default values are assigned to each attribute. Thus, if both sides agree only on a subset of the issues and the deadline is reached, the unresolved issues are assigned their default “no agreement” value and thus a partial agreement can be implemented.

It is assumed that the agents can take actions during the negotiation process until it terminates. Let **Time** denote the set of time periods in the negotiation, that is **Time** = $\{0, 1, \dots, dl\}$. Time also has an impact on the agents’ scores. Each agent is assigned a time cost which influences its score as time passes. In each period $t \in \mathbf{Time}$ of the negotiation, if the negotiation has not terminated earlier, each agent can propose a possible agreement, and the other agent can either accept the offer, reject it or opt out. Each agent can either propose an agreement which consists of all the issues in the negotiation, or a partial agreement. We use an extension of the model of alternating offers (Osborne and Rubinstein [17], p. 118-121), in which each agent can perform as many interactions with its counterpart until the time period ends.

The negotiation problem involves incomplete information concerning the opponent's preferences. There is a finite set of agent types. These types are associated with different additive score functions. Formally, we denote the possible types of agents $\mathbf{Types} = \{1, \dots, k\}$. Given $t_i \in \mathbf{Types}$, $1 \leq t_i \leq k$, we refer to the score of an agent of type t_i as u_i , and $u_i : \{(O \cup \{SQ\} \cup \{OPT\}) \times \mathbf{Time}\} \rightarrow \mathbb{R}$. Each agent is given its exact score function. The negotiators are aware of the set of possible types of the opponent.

3.2. *Enhancing People's Negotiation Skills*

Three different approaches were employed to investigate their effect on people's negotiation skills. The first training method was the classical role playing of two people, that is, negotiating with another person. While role playing might be the simplest training method and used in classes, in the general case it is hard to find human negotiators with whom one can train.

The second approach that we evaluate is role playing with an automated negotiator. In this approach, the human negotiator is matched with the *KBAgent* [18], a concession-oriented agent that uses a general opponent modeling technique.

The third approach that we examine is the improvement of negotiation skills due to the actual design of an automated negotiator by the human subjects to play as their proxy. In this case, the human negotiators were given a task to implement an efficient automated agent. The students were provided skeleton classes to help them implement their agents. This also allowed them to focus on the strategy and the behavior of the agent, and eliminate the need to implement the communication protocol or the negotiation protocol. In addition, it provided them with a simulation environment in which they could test their agents and their strategies.

3.3. *Implementation using GENIUS*

The experiments were conducted using the GENIUS simulation environment [14] and a given multi-attribute multi-issue domains. We begin by describing the domains which was used in all the experiments and then continue to describe the experimental methodology and results.

We used existing domains from the literature [15]. The first domain is a Job Candidate domain, which is related to the subjects' experience, and thus they could better identify with it. In this domain, a negotiation takes place after a successful job interview between an employer and a job candidate. In the negotiation both the employer and the job candidate wish to formalize the hiring terms and conditions of the applicant. Below are the issues under negotiation:

1. **Salary.** This issue dictates the total net salary the applicant will receive per month. The possible values are (a) \$7,000, (b) \$12,000, or (c) \$20,000. Thus, a total of 3 possible values are allowed for this issue.
2. **Job description.** This issue describes the job description and responsibilities given to the job applicant. The job description has an effect on the advancement of the candidate in his/her work place and his/her prestige. The possible values are (a) QA, (b) programmer, (c) team manager, or (d) project manager. Thus, a total of 4 possible values are allowed for this issue.
3. **Social benefits.** The social benefits are an addition to the salary and thus impose an extra expense on the employer, yet they can be viewed as an incentive for the applicant. The social benefits are divided into two issues: **company car** and the percentage of the salary allocated, by the employer, to the candidate's **pension funds**. The possible values for a company car are (a) providing a leased company car, (b) no leased car, or (c) no agreement. The possible value for the percentage of the salary deposited in pension funds are (a) 0%, (b) 10%, (c) 20%, or (d) no agreement.
4. **Promotion possibilities.** This issue describes the commitment by the employer regarding the track for promotion for the job candidate. The possible values are (a) fast promotion track (2 years), (b) slow promotion track (4 years), or (c) no agreement. Thus, a total of 3 possible values are allowed for this issue.
5. **Working hours.** This issue describes the number of working hours required by the employee per day (not including over-time). This is an integral part of the contract. The possible values are (a) 8 hours, (b) 9 hours, or (c) 10 hours. Thus, a total of 3 possible values are allowed for this issue.

In this scenario, a total of 1,296 possible agreements exist ($3 \times 4 \times 12 \times 3 \times 3 = 1296$). Each turn in the scenario equates to two minutes of the negotiation, and the negotiation is limited to 28 minutes. If the parties do not reach an agreement by the end of the allocated time, the job interview ends with the candidate being hired with a standard contract, which cannot be renegotiated during the first year. This outcome is modeled for both agents as the status quo outcome.

Each side can also opt-out of the negotiation if it feels that the prospects of reaching an agreement with the opponent are slim and it is impossible to negotiate anymore. Time also has an impact on the negotiation. As time advances the candidate's score decreases, as the employer's good impression of the job candidate decreases. The employer's score also decreases as the candidate becomes less motivated to work for the company.

The score values range from 170 to 620 for the employer role and from 60 to 635 for the job candidate role. Both players had a fixed loss per time period – the employer of -6 points and the job candidate of -8 points per period.

Another domain was used to validate and bolster our confidence in the results, which involved reaching an agreement between Britain and Zimbabwe evolving from the World Health Organization’s Framework Convention on Tobacco Control, the world’s first public health treaty. The principal goal of the convention is “to protect present and future generations from the devastating health, social, environmental and economic consequences of tobacco consumption and exposure to tobacco smoke”. In this domain, 5 different attributes are under negotiation, resulting with a total of 576 possible agreements. The issues under negotiation are (a) the size of the fund, (b) the impact on other aid programs, (c) Zimbabwe’s trade policy, (d) Britain’s trade policy and (e) creation of a similar fund for other health issues.

Each turn in the scenario equals a week of negotiation, and the negotiation is limited to 14 turns (28 minutes, simulating 14 weeks). If the sides do not reach an agreement by the end of the allocated time, the Framework Convention will be seen as an empty document, which will cause Britain to lose political capital invested in the summit but save it money in the short term. For Zimbabwe such an event will cause financial hardship and deprive it of a precedent that can be used for future negotiations. This outcome is modeled for both agents as the status quo outcome. As time advances Zimbabwe’s score decreases since the aid measures discussed are not implemented. On the other hand, Britain gains score as time advances since it postpones the date it must transfer money to the fund.

The score values range from -575 to 895 for Britain’s role and from -680 to 830 for Zimbabwe’s role. Britain’s role has a fixed gain per time period of 12 points while Zimbabwe loses 16 points per period.

There are three possible types of agents for each role. These types are associated with different additive score functions. The different types are characterized as ones with short-term orientation regarding the final agreement, long-term orientation and a compromising orientation. Detailed score tables for the domain can be found in Appendix A.

3.4. Experimental Settings

We ran an extensive set of simulations, consisting of a total of 269 human negotiators. The human negotiators were mostly computer science undergraduate and graduate students, ages 18-30, while a few were former students who are currently working in the Hi-Tech industry. Table 1 summarizes the number of

Approach/Role	Employer	Job Candidate	Britain	Zimbabwe
Control Group	18	16	15	15
Training via Human Negotiation	18	18	20	20
Training via Automated Negotiator	20	20	18	18
Training via Agent Design	19	19	15	N/A

Table 1: Number of subjects in each evaluation method in the Job-Candidate and Britain-Zimbabwe domains.

different human subjects we had per each method we evaluated. Each subject served only one specific role in the negotiations (either the employer (Britain) role or the job candidate (Zimbabwe) one).

Each simulation was divided into two parts: (i) using one of the training method described above, and (ii) negotiating against the standardized agent. Prior to the experiments, the subjects were given oral instructions regarding the experiment and the domain. The subjects were instructed to play based on their score functions and to achieve the best possible agreement for them. While the subjects knew that they will negotiate twice, they did not know in advance against whom they played (whether it is a human negotiator or an automated one).

3.5. Evaluating the Negotiation Skills

To avoid bias and subjective measures, in order to evaluate the different training methods that we propose, all people negotiated against a standardized objective agent. In addition, a control group was used to compare the different results. This group consisted of people that had not undergone any training before negotiating against the standardized agent. While well-designed questionnaires may be constructed to allow providing objective and useful insights, some papers (e.g., [4, 19]) rely on questionnaires, which are subjective. For example, subjects were asked how they evaluated their negotiation experiment, whether they believe they are better trained now and the sort.

We used the *QOAgent* as the standardized agent. The *QOAgent* is an automated negotiator which has been previously shown to be an efficient negotiator against human counterparts [15], to evaluate the methods we applied.

Method	Role	Average	Std.	p-value
Control Group	Employer	431.78	80.83	
	Job Can.	320.5	112.71	
	Britain	335.33	194.62	
	Zimbabwe	-320.07	274.42	
Training via Human Negotiation	Employer	448.56	66.08	0.25
	Job Can.	383.83	112.73	0.05
	Britain	366.45	198.65	0.32
	Zimbabwe	-268.7	301.93	0.3
Training via Agent Design	Employer	466.84	46.26	0.06
	Job Can.	391.53	76.75	0.02
	Britain	422.93	162.77	0.09
	Zimbabwe	N/A		
Training via Automated Negotiator	Employer	468.6	38.94	0.04
	Job Can.	433	102.84	0.002
	Britain	301.22	182.14	0.3
	Zimbabwe	-44.6	196.19	< 0.002

Table 2: Comparison of the average scores and standard deviation of human negotiators using different training methods and the control group.

3.6. Results

Table 2 summarizes the average scores achieved by the human negotiators in all of the experiments condition when matched with the *QOAgent* - either after going one of the training procedures (role playing with people, training via agent design and training with another automated negotiator) or not (the control group – people that were only matched with the standardized agent). The table also presents the statistical significance of the results, by applying the *t-test* analysis, compared to the control group.

We begin the first training method in which people were matched with other people. The results demonstrate that the classical training method of role playing between humans allows the human negotiators to achieve higher scores (448.56

and 383.83 for the employer and job candidate roles, respectively; 366.45 and -268.7 for the Britain and Zimbabwe roles, respectively) compared to the control group. However, this is only significant for one of the roles and only in the Job-candidate domain (the job candidate role, with p -value < 0.05).

Next we tested the results of the training method using agent design. Similar to role playing, the results are also better than the control group, in terms of average scores, but allowed to achieve even significantly better results in one setting. The average scores of the people in the training group were 466.84 (p -value < 0.06) and 391.53 (p -value < 0.02) for the employer and job candidate roles, respectively, compared to 431.78 and 320.5 for the control group. In the Britain-Zimbabwe domain this method allowed an increased score, though not significant.

Finally, we evaluated the training method in which negotiation was done with another automated agent. In this training method, as opposed to all other methods, we find that the average score obtained by people is significantly higher in all cases, but one (the Britain role, in which the results, though not significant, are actually worse than the control group) (468.6 compared to 431.78 with a p -value < 0.04 for the employer role, 433 compared to 320.5 with a p -value < 0.002 for the job candidate role, and -44.6 compared to -320.07 with p -value < 0.002 for the Zimbabwe role).

We continued to test whether some of the training methods that we evaluated were better than others. To this end we did a cross-reference comparison between the results of the people in each group when matched with the standardized automated negotiator after undergoing training. That is, we compared (a) the groups of people training via human negotiations and those trained via agent design, (b) training via automated negotiator versus training via agent design, and (c) training via automated negotiators compared to training via human negotiations. The comparison revealed that while some training methods enabled the negotiators to achieve higher scores than others, the results were not significant in most roles or training method, and thus we cannot state that one training method is superior to others.

3.7. Discussion

The experimental results show that in the vast majority of the cases we studied, using automated agents—whether designed by experts or the subjects themselves—improved people’s performance when negotiating in two different domains compared to the more traditional approach of training with people. Surprisingly, no significant differences were found between the different training methods. Yet in all of these methods higher scores were achieved when the automated agents

were involved as opposed to training with other humans, with the exception of the diplomatic negotiation domain. Our hypothesis for this exception is the specific characteristics of this role in the Tobacco domain. Britain has more leverage on the other side, which made people, after negotiating with another automated agent, to concede faster than they should have. Though the results are not significant, it is quite obvious that training via automated negotiators is a much more simpler task and less time consuming than training via agent design. Since we did not focus in this paper on the design of an automated negotiator we did not experiment with other automated agents other than the *QOAgent* and the *KBAgent*.

4. Coordination Game Settings

In this section we extend our study to three-player coordination games. Our setting of choice was a three player constant-sum game with simultaneous moves called the Lemonade Stand Game (LSG), originally proposed as a test-bed for the evaluation of opponent modeling and machine learning techniques [25]. In this game there are twelve possible actions for each player, representing possible locations to setup a lemonade stand on the beach of an island. The twelve locations are uniformly spread around the perimeter, in a similar way to the hours on the face of a clock. Players choose their locations on the board simultaneously. The players in the game represent lemonade vendors who compete to serve customers in their vicinity. The utility to each player is the sum of its distance from the nearest player in a clockwise direction and the nearest player in a counter-clockwise direction. Distances are measured by counting the number of positions between players. Actions are taken simultaneously by all players. This is analogous to the profit each player would make if customers buying lemonade are uniformly distributed around the island. If more than one player is positioned in the same location, there is a “collision”, and the players share the profit incurred in that location. In this case, both of the two players in the “collision” receive a score of 6 (because there is a distance of 12 between their positions and the position of the third player), whereas the third player receives a score of 12. If all three players are positioned in the same position, then each receive a score of 8.

A snapshot of the GUI for playing the LSG is shown in Figure 1. The snapshot is shown for a particular round from the perspective of the player whose position is represented by the red disc at 6 o’clock. This player scored 10 points in this round. This is because the distance of the player from the “green player” at 2 o’clock is four positions, and the distance of the player from the “blue player” at 12 o’clock was 6 positions. The score shown in the upper right-hand-side of the

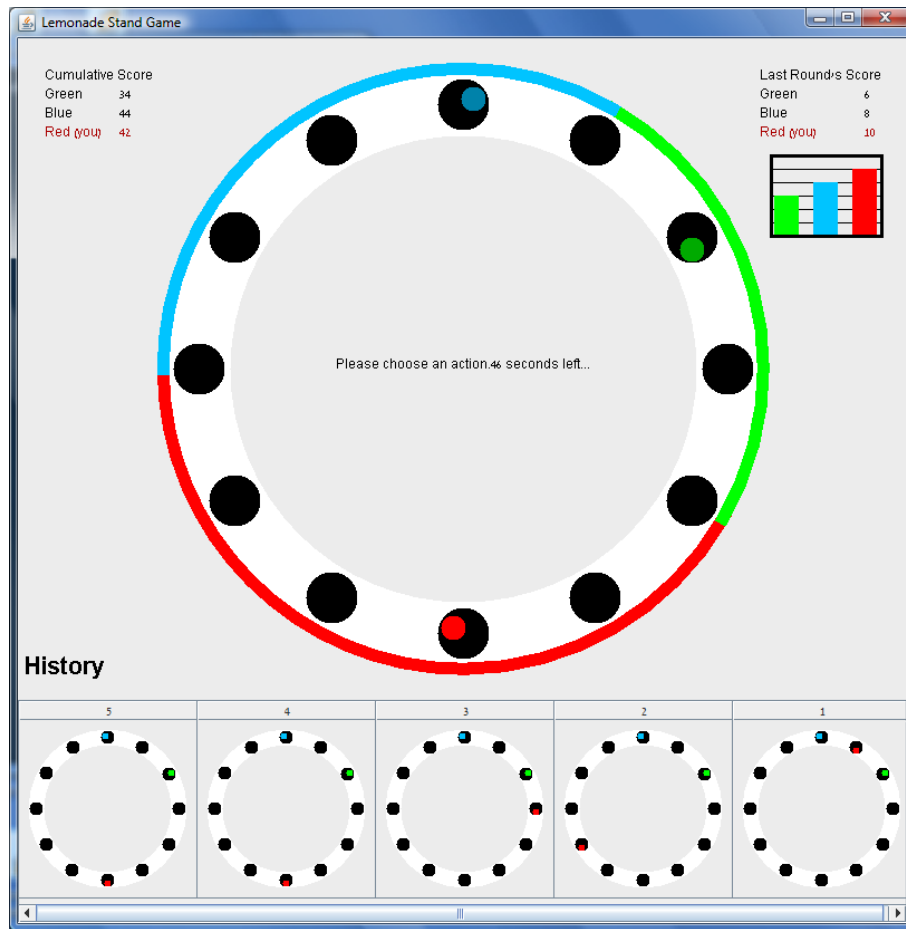


Figure 1: A snapshot of the Lemonade Stand Game

figure in both table and graph formats. In addition, the cumulative score for all players is displayed on the upper left-hand-side of the figure. The bottom part of the Figure is a history panel in which participants can observe the results of all of the prior rounds of the game.

The advantage of using the LSG as a setting for training people’s play is twofold. First, its rules are simple and intuitive. However, it is challenging for people to play. The game is inherently competitive, meaning that when a player gains from playing a particular strategy, other players necessarily lose. However, It also allows for cooperation between players to coordinate an “attack” on the third player. To succeed, players not only need to reason about what strategy the

other players are using but also to try to influence their strategy over time. In addition, player's outcomes are affected by the strategies of the two other players. This makes it more difficult for people to identify and learn beneficial strategies to play during training. Second, there is a publicly available library of agents designed by experts to compete in an annual tournament.¹

4.1. Empirical Methodology

We recruited 56 undergraduate students from Ben-Gurion university to play the game. Ages ranged from 24 to 30, 58% males and 42% females. All subjects played 90 rounds of the lemonade stand game divided into three epochs of thirty games each. The first two epochs (called "training epochs") were used to train the people to play the game, and their performance was measured when playing the final "testing epoch". The player configuration in the training epochs varied as follows. In the "all-human" training condition, the player configuration included only human players. In the "single-agent" training condition, the player configuration included two human players and a single agent player. In the "two-agent" training condition, the configuration included a single human player and two agent players. The player configuration for the testing epoch included two human players and a standardized agent player. In the "agent design" condition, subjects were trained by designing an automated negotiator.

To choose the standardized and training agent we ran an independent tournament which evaluated the agent entries submitted to the 2009 and 2010 agent competitions. The tournament consisted of 30 rounds, the same number of rounds in the testing epoch.² The winner of the tournament, called *EA*² (which also won the 2009 competition) was chosen to be the standardized agent for testing performance after training. The training agents were chosen to be the second and third runner ups in the competition. All results reported as significant in the following section were confirmed in the $p < 0.05$ range using single-factor ANOVA tests. Subjects were randomly divided into the single- and two-agent training conditions, as well as a baseline condition in which subjects played a single testing epoch with the standardized agent. Altogether, there were 16 games played in the no-training condition, 19 games played in the all-human and single-agent training conditions, and 12 games played in the double-agent training condition.

¹The rules of the game are commonly changed every year to promote research, such as varying the utility function or the number of games played in succession with the same agent players. In this study we confined ourselves to the original game description used in the 2009-10 competitions.

²The actual tournament ran 1000 rounds.

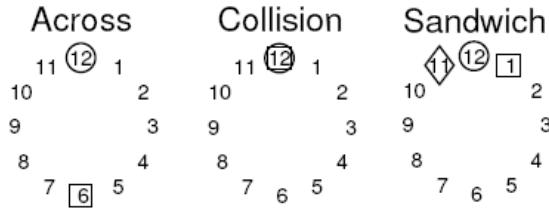


Figure 2: key outcomes in the Lemonade Stand Game. Across (left); Collision (middle); Sandwich (right)

4.2. Results

We first compare people’s play in the various conditions to that of the standardized agent. Figure 3 shows the average aggregate performance of people and of the standardized agent in the testing epoch. As shown by the Figure, the EA^2 agent significantly outperformed people in the all-human and all-human training conditions. However, the difference in performance between the EA^2 agent and people was not significant in the single-agent and double-agent training conditions. This effect was also consistent for the condition in which people designed their own agent.

We conjectured that the reason for the lack of difference between the agent and people’s performance was that people learned to play beneficial strategies from interacting with agents in the training epochs. To examine this, we analyzed people’s behavior in the game using key tactics used to signal and communicate with other players. In one of these tactics, called “Stick”, a player chooses to remain in the same location in two consecutive rounds. In another tactic, called “follow”, a player chooses a location that is directly across the location of another player in the previous round. The situation in which both players are positioned directly across from each other results is called “across”. This is a form of cooperation which guarantees a payoff of between 6 and 12 to each of the players, while the third player receives a payoff of 6. This outcome is stable (no player has incentive to deviate), and is one of the multiple pure Nash equilibria of this game. It was the most common form of cooperation achieved by agents playing the game. In contrast, the outcome of “sandwich”, in which one of the players is positioned directly between the two other players, provides a higher outcome of 11 for both of the agents positioned at the edges and a low score of two for the agent in the middle. However, this strategy is not stable for the low scoring agent, who can do better by moving out of the sandwich.

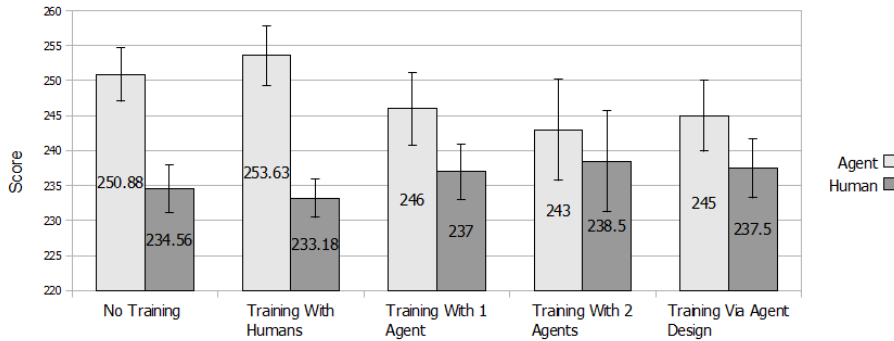


Figure 3: Performance Comparison: People versus the standardized EA^2 agent

	All-Human			Two-Agent		
	Follow	Stick	Across	Follow	Stick	Across
People	8.82	7.13	9	17	21	25.33

Table 3: Number of “follow” and “stick” strategies used by human players in the testing epoch

Table 3 shows the frequency of follow and stick strategies in the testing epoch for the all-human and double-agent training condition. As shown by the table, people engaged in significantly more follow, stick and across strategies in the two-agent training condition (shown in boldface) than in the single-agent training condition. The difference between the number of follow and stick strategies played by people in the single-agent and no-training condition exhibited a similar pattern. Therefore, we attribute the improvement in people’s play in the single- and double-agent condition to their increased use of cooperative strategies. Lastly, as shown by Figure 3, people were not able to outperform the agent after training. We attribute this to the inherent difficulty of playing the state-of-the-art agent for this game.

4.3. Comparing People’s Performance Across Conditions

We now compare people’s performance scores across the various conditions. The number of follow, stick and across outcomes for people in the two-agent training condition was significantly higher than in the no-training and all-human training condition. Thus people learned to be more cooperative when training with two agents. As shown by Figure 3, people’s average score in the two-agent training condition (238 points) was higher than in the no-training (234 points) and single-agent training condition (237), but this difference was not significant.

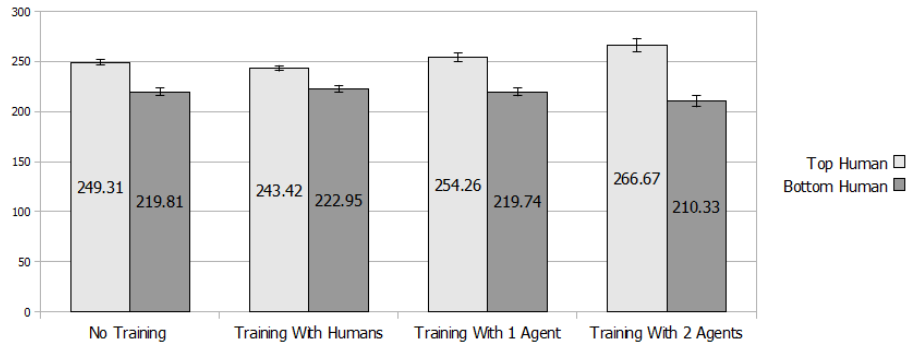


Figure 4: Performance Comparison: People versus people

To explain this discrepancy we distinguish between the top- and low-scoring human players in each game (the people who scored the highest and lowest scores in each game). Because the LSG is constant sum game, one player’s win is another player’s loss. In the context of the lemonade stand game, this means that when two players coordinate and play across, they necessarily earn more points than the player that is left out. We hypothesized that top-scoring human players coordinated more often with the standardized agent than low-scoring players, allowing them to outperform the low-scoring players.

Figure 4 shows the average performance of low- and top-scoring human players in all conditions. As shown by the Figure, the top-scoring players in the two-agent training condition outperform top-scoring players in the one-agent and all-human training conditions. Also shown in the figure is that there is no difference in the performance of low-scoring agent across conditions. Table 4 shows the frequency of across outcomes in the two-agent training condition for low- and top-scoring human players. As shown by the table, the top-scoring players played significantly more across outcomes than low-scoring players in each of the conditions. In addition, top-scoring players achieved significantly more across outcomes in the two-agent training condition than in the all-human and single-agent training condition. This confirms our hypothesis, in that the success of top-scorers in the game is attributed to their increased coordination with the standardized agent (rather than the other human player).

5. Conclusions

In this paper we presented an extensive experimentation to answer the question whether simulation improves people’s performance in tasks requiring negotiation

	Follow	Stick	Across
low-scoring	16	15	14.92
top-scoring	26.6	20	25.3

Table 4: Number of “follow” and “stick” strategies used by top- and low-scoring human players in the two-agent training condition

and coordination skills. Our results showed that training human subjects with automated agents designed by researchers or the subjects themselves improved their performance when compared to the more traditional method of using training with other people. These results suggest that agents can be used as tools for training people in our settings of choice, resulting in considerable savings in cost and effort as compared to using people for training purposes. In future work we will generalize our approach by testing people’s performance on a different domain in which they were trained.

6. Acknowledgements

This research is supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office under grant number W911NF-08-1-0144, under NSF grant 0705587 and by ERC grant #267523. Y.G. was supported in part by Marie Curie grant number #268362.

References

- [1] Y. Bachrach, P. Kohli, T. Graepel, Rip-off: playing the cooperative negotiation game, in: Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems, pp. 1179–1180.
- [2] D. Butler, Air Gondwana: Using ICT to create an authentic learning environment to teach basic negotiation skills, in: Proceedings of the 32nd Higher Education Research and Development Society of Australasia Annual Conference, pp. 53–64.
- [3] A. Byde, M. Yearworth, K. Chen, C. Bartolini, AutONA: A system for automated multiple 1-1 negotiation, in: Proceedings of the 2003 IEEE International Conference on Electronic Commerce, pp. 59–67.

- [4] D. Druckman, N. Ebner, Onstage or behind the scenes? relative learning benefits of simulation role-play and design, *Simulation & Gaming* 39 (2008) 465–497.
- [5] G. Hofstede, L. De Caluwé, V. Peters, Why simulation games work-in search of the active substance: A synthesis, *Simulation & Gaming* 41 (2010) 824–843.
- [6] C.M. Jonker, V. Robu, J. Treur, An agent architecture for multi-attribute negotiation using incomplete preference information, *Autonomous Agents and Multi-Agent Systems* 15 (2007) 221–252.
- [7] E. Kamar, Y. Gal, B.J. Grosz, Modeling user perception of interaction opportunities for effective teamwork., in: *Proceedings of the International Conference on Computational Science and Engineering*, pp. 271–277.
- [8] P. Kenny, A. Hartholt, J. Gratch, W. Swartout, D. Traum, S. Marsella, D. Piepol, Building interactive virtual humans for training environments, in: *Proceedings of Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)*, pp. 1–16.
- [9] G.E. Kersten, S.J. Noronha, WWW-based negotiation support: design, implementation, and use, *Decision Support Systems* 25 (1999) 135–154.
- [10] S. Kraus, P. Hoz-Weiss, J. Wilkenfeld, D.R. Andersen, A. Pate, Resolving crises through automated bilateral negotiations, *Artificial Intelligence* 172 (2008) 1–18.
- [11] R. Lennon, A. Sharland, M. Gonzalez, International negotiation simulations: An examination of learning processes and outcomes, *College Teaching Methods & Styles Journal (CTMS)* 2 (2011) 43–52.
- [12] J. Lim, Multi-stage negotiation support: a conceptual framework, *Information and Software Technology* 41 (1999) 249–255.
- [13] R. Lin, S. Kraus, Can automated agents proficiently negotiate with humans?, *Communications of the ACM* 53 (2010) 78–88.
- [14] R. Lin, S. Kraus, T. Baarslag, D. Tykhonov, K.V. Hindriks, C.M. Jonker, *Genius: An integrated environment for supporting the design of generic automated negotiators*, *Computational Intelligence* (In Press).

- [15] R. Lin, S. Kraus, J. Wilkenfeld, J. Barry, Negotiating with bounded rational agents in environments with incomplete information using an automated agent, *Artificial Intelligence* 172 (2008) 823–851.
- [16] R. Lin, Y. Oshrat, S. Kraus, Investigating the benefits of automated negotiations in enhancing people’s negotiation skills, in: *Proceedings of the Eighth International Conference on Autonomous Agents and Multi-Agent Systems*, pp. 345–352.
- [17] M.J. Osborne, A. Rubinstein, *A Course In Game Theory*, MIT Press, Cambridge MA, 1994.
- [18] Y. Oshrat, R. Lin, S. Kraus, Facing the challenge of human-agent negotiations via effective general opponent modeling, in: *Proceedings of the Eighth International Conference on Autonomous Agents and Multi-Agent Systems*, pp. 377–384.
- [19] W.H. Ross, W. Pollman, D. Perry, J. Welty, K. Jones, Interactive video negotiator training: A preliminary evaluation of the mcgill negotiation simulator, *Simulation & Gaming* 32 (2001) 451–468.
- [20] B.A. Starkey, E.L. Blake, Simulation in international relations education, *Simulation & Gaming* 32 (2001) 537–551.
- [21] L.E. Susskind, J. Corburn, *Using Simulations to Teach Negotiation: Pedagogical Theory and Practice*, Working Paper 99-1, Program on Negotiation at Harvard Law School, 1999.
- [22] E.M. Thiessen, D.P. Loucks, J.R. Stedinger, Computer-assisted negotiations of water resources conflicts, *Group Decision and Negotiation* 7 (1998) 109–129.
- [23] D. Traum, S. Marsella, J. Gratch, J. Lee, A. Hartholt, Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents, in: *Proceedings of the 8th International Conference on Intelligent Virtual Agents*, pp. 117–130.
- [24] A. van Wissen, Y. Gal, B. Kamphorst, M. Dignum, Human-agent team formation in dynamic environments, *Computers in Human Behavior* 28 (2012) 23–33.

- [25] M. Zinkevich, M. Bowling, M. Wunder, The lemonade stand game competition: solving unsolvable games, *ACM SIGecom Exchanges* 10 (2011) 35–38.

Appendix A. Score Tables

The following tables present the score tables for both negotiators, in both domains, from which the score of an agreement is calculated. While the human subject is given her own score table at the beginning of the negotiation, she is also given three additional score tables which model the different possible types of her opponent.

Appendix A.1. The Job Candidate Domain (i) Short-Term, (ii) Long-Term and (iii) Compromise Orientation Score Tables

OUTCOMES	Job Candidate Outcome Weight / Importance			Employer Outcome Weight / Importance		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Salary	20%	30%	15%	20%	15%	10%
7,000 NIS	3	2	3	8	7	7
12,000 NIS	6	6	5	6	6	6
20,000 NIS	8	9	6	3	3	4
No agreement	0	0	0	0	0	0
Job Description	15%	25%	20%	20%	30%	20%
QA	2	-2	2	4	2	3
Programmer	4	3	4	6	6	6
Team Manager	5	6	6	4	3	4
Project Manager	6	8	8	2	1	3
No agreement	0	0	0	0	0	0
Leased Car	20%	5%	10%	10%	10%	10%
Without leased car	-5	-5	-2	3	4	5
With leased car	5	5	2	-2	2	4
No agreement	0	0	0	0	0	0
Pension Fund	10%	5%	10%	10%	10%	10%
0% pension fund	-2	-2	-2	3	6	6
10% pension fund	3	4	3	4	4	4
20% pension fund	5	6	5	3	3	3
No agreement	0	0	0	0	0	0
Promotion Possibilities	5%	25%	35%	10%	20%	20%
Slow promotion track	4	1	-2	3	8	6
Fast promotion track	5	5	5	3	5	4
No agreement	0	0	0	0	0	0
Working Hours	30%	10%	10%	30%	15%	30%
10 hours	3	3	4	8	8	9
9 hours	5	4	5	6	6	6
8 hours	7	5	6	3	4	3
No agreement	0	0	0	0	0	0
Time effect	-8	-8	-8	-6	-6	-6
Status Quo	160	135	70	240	306	306
Opting out	150	75	80	210	150	215

Appendix A.2. The Britain-Zimbabwe Domain (i) Short-Term, (ii) Long-Term and (iii) Compromise Orientation Score Tables

OUTCOMES	Zimbabwe Outcome Weight / Importance			Britain Outcome Weight / Importance		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Size of Fund	50%	10%	20%	50%	10%	30%
\$100 Billion	9	5	6	-5	1	2
\$50 Billion	2	2	4	2	3	4
\$10 Billion	-5	-3	2	10	6	6
No agreement	-8	-6	-2	7	-1	-2
Impact on Other Aid	30%	10%	20%	30%	10%	30%
No reduction	8	6	3	-4	1	0
Reduction is equal to half of the fund size	0	0	0	4	2	3
Reduction is equal to the fund size	-3	-3	-2	10	3	5
No agreement	-5	-4	-4	-7	0	-2
Trade Policy	10%	30%	30%	10%	30%	10%
Zimbabwe will reduce tariffs on imports	-6	-3	-4	3	4	5
Zimbabwe will increase tariffs on imports	3	6	4	-3	-6	-6
Britain will increase imports	7	8	10	-4	-8	-5
Britain will reduce imports	-8	-9	-8	4	6	4
No agreement	0	0	0	0	0	0
Forum on Other Health Issues	10%	50%	30%	10%	50%	30%
Creation of fund	9	8	7	-8	7	4
Creation of committee to discuss creation of fund	3	5	5	2	4	7
Creation of committee to develop agenda	-5	-6	3	6	-2	1
No agreement	-6	-8	-3	1	-4	-2
Time effect	-16	-16	-16	12	12	12
Status Quo	-610	-500	-210	150	-210	-180
Opting out	-530	-520	-240	-105	-240	-75