

**Group dynamics in mixed human | agent environments:
Betty, the Teachable Agent, is the next contestant on the Triple A Game Show!**

Paula Wellings
Communication 266
03/19/2005

This paper is an exploration of the challenges and opportunities available when an educational technology system that accommodates one human and one embodied agent becomes a system that enables the interaction of many humans and many agents. Under the direction of Dan Schwartz, Stanford's AAA Lab develops software that lets children teach virtual agents in order to learn scientific concepts and content. Based on the adage that one learns more when one teaches, the AAA Lab's Teachable Agents (TAs) reason about academic tasks based on the instruction provided to them by their human teachers. Through watching their agents perform these tasks, students learn how well they have taught their agents and can address any gaps in knowledge their agents may have. Unlike pedagogical agents that are primarily demonstrative, such as intelligent tutoring systems, "a teachable agent...has little or no *a priori* knowledge to demonstrate. The interface between the agent and the user serves two purposes, that of allowing the user to teach the agent, and that of engaging the user's attention through the agent's personality and interactivity" (p. 27, Davis et. al.).

One of the Lab's TAs, Betty interacts with her human teacher via a concept map that represents her "brain" (see Figure 1). The human teacher teaches Betty by linking together concepts via cause-effect relations in Betty's brain. Teachers can then ask Betty to explain her reasoning based on her concept map and Betty animates her *thought process*. Teachers can also ask Betty questions and send Betty to take a quiz where she is given feedback about her knowledge. Based on quiz results, the teacher can improve Betty's knowledge by altering the concept map.

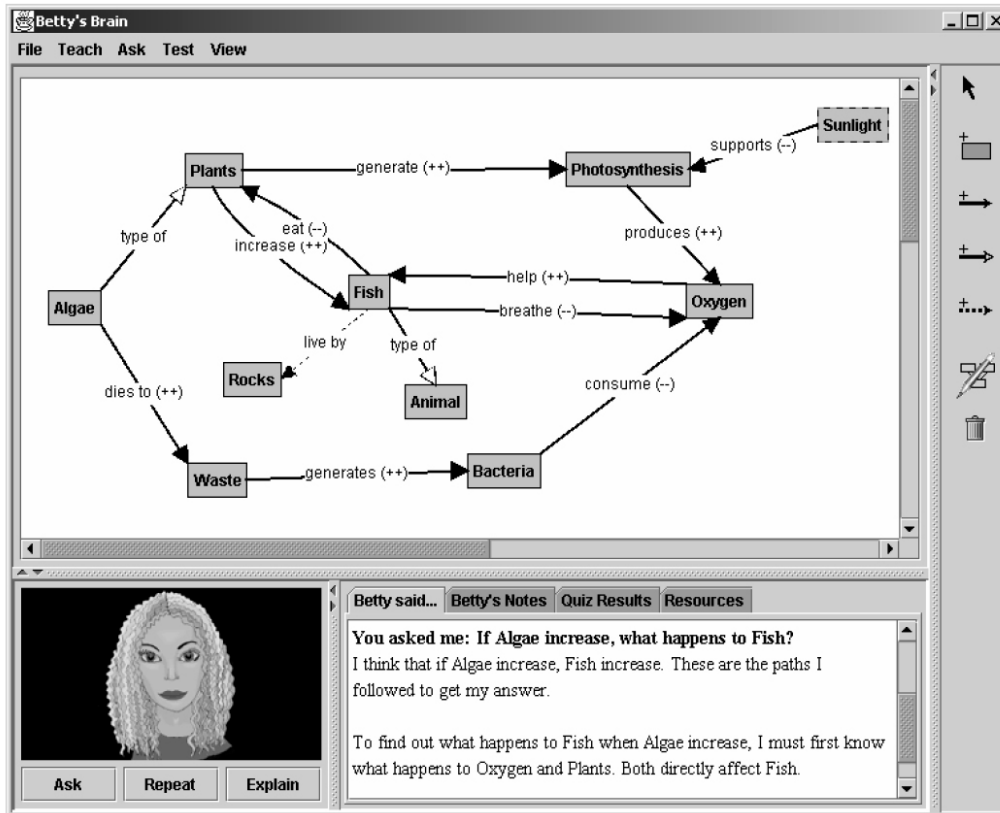


Figure 1: Betty's Brain Interface (p. 27, Davis et. al.)

Relevant finding about agent presence and human expectations have been gathered from creating and studying the TA Betty. It has been found that agent indicators of presence such as speech lead to increased learning behaviors when students use Betty. “On average, we found that students using the teachable agent with voice spent 2.5 times more time with the agent than students using the agent without voice (p.32, Davis et al. 2003). It has also been found that students expect their agents to respond to their teaching activities and demonstrate initiative. Students in one study with Betty stated that they “would like Betty to participate more actively during the teaching phases, i.e., they wanted Betty to exhibit characteristics of a good student and be a more *active learner*” (p.31, Davis et al. 2003).

Previous studies with Betty have also identified specific findings regarding student learning behavior. Younger children, often less experienced in the role of teacher need assistance in becoming better teachers. An updated version of Betty has added a mentor teacher to support novice teachers. Another finding, that has lead to additional features for Betty, is that while the goal of the concept map is to assist students and agents in understanding multi-link relationships between concepts, after students have created their initial concept map they tend to use the quiz feature to facilitate a “test, modify, and retest” strategy focused on answering single questions and addressing single links between concepts. The result of this strategy is that students create new wrong linkages while correcting single link concepts. To address this scenario, a version of Betty has been created that incorporates self-regulated learning algorithms that cause Betty to refuse to take another quiz until she considers herself to be adequately taught.

A related significant finding has been the use of Betty in a classroom setting where students can see the concept maps of other agents broadcast on a screen at the front of the class. In this Front of the Class (FoC) version of Betty, students’ revisions to their agents’ concepts maps improved, with a greater emphasis on solving multi-link questions (see Figure 2).

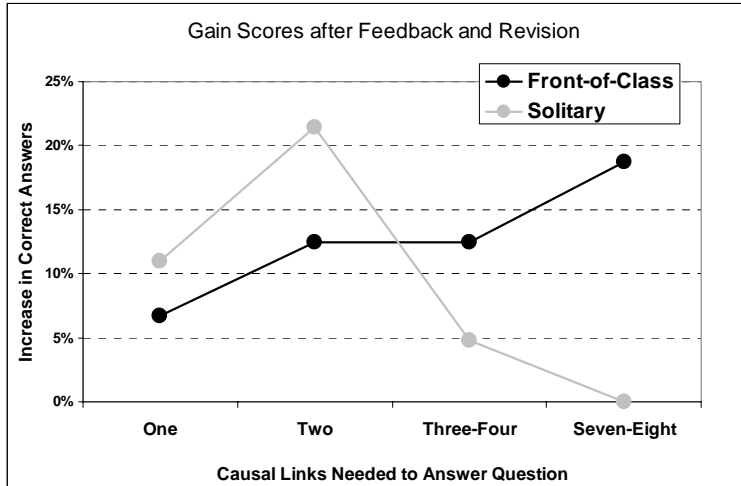


Figure 2. Changes to the quality of Betty maps when assessed by asking all possible questions. Students in both conditions improved their maps after seeing Betty answer a few quiz questions. Students who saw their agents at the front of the class improved the overall structure of their map more (Schwartz et. al.).

The Triple A Challenge Game Show

As an attempt to leverage the role that social display and knowledge sharing provides to students, the Triple A Challenge Game Show is an extension of the FoC Betty system. The game show puts many human teachers and TAs in a quiz show-like virtual environment. The TAs compete against each other based on the knowledge they have acquired from their human teachers. While previous versions of TAs are single agents with one human teacher, in this case, there are multiple humans and multiple agents playing the same game. Each human has a direct relationship to an agent, playing the role of the agent's coach within the quiz show game world and wagering on his or her agent's performance. Each human also has relationships with the other humans playing the game, as friends, as potential competitors and collaborators. Each of the TAs has a relationship to their human coach and to the other agents which they compete against in the game show. All of these relationships introduce a new element into the virtual environment, virtual group dynamics.

Multiple agents and group dynamics

Previous work looking at group dynamics in multi-human/agent environments reveal that similar to physical environments, group dynamics exist in these virtual environments. In Slater et. al.'s (2004) study of the fear of public speaking in a virtual environment, people who were typed either confident or phobic about public speaking addressed a virtual audience composed of five animated agents who were either attentive or inattentive. The study found that although "the avatars had low representational quality, with pre-programmed behaviors independent of the actual behavior of the speaker, people still responded according to type" (p. 8, Slater et. al.). While this study focuses on addressing phobias associated with public speaking, it is an example of research where one human avatar is placed within a social dynamic constructed by multiple agents.

Blascovich et. al.'s (2002) work on conformity and social comparison also considers the effect that multiple agents have on human behavior. In a virtual blackjack simulation, one human played blackjack with two agents in either low-, same-, or high-betting norm conditions. The study found a conformity effect such that the human participant would gravitate towards to the betting norms demonstrated by the two agents.

Is there a difference between individual co-presence and group co-presence?

Blascovich's model of social influence within virtual environments (see Figure 3) argues that agents require a high level of realism in order to exert social influence on human participants while avatars require a much lower level of realism to exert the same level of

social influence. Based on the work cited above, including Blascovich's blackjack study, it is worth considering if Blascovich's model is equally applicable to environments where multiple agents are engaged in social interaction with each other and with human avatars.

Slater's findings specifically note that agents with a low level of realism evoked similar levels of social influence upon the public speakers as they had experienced in previous public speaking experiences. While Blascovich's blackjack study found stronger normative effects in conditions where people played with other human avatars, there was also significant effects in the human-agent condition.

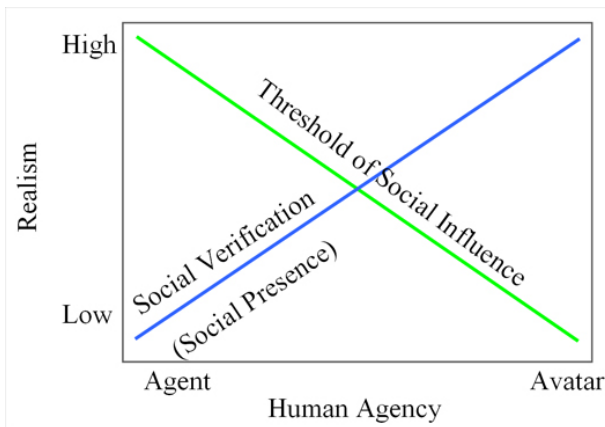


Figure 3: Model of Social Influence within Virtual Environments (Blascovich et. al.,2004).

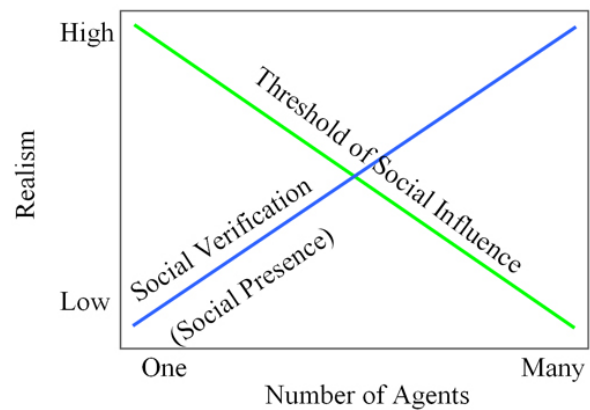


Figure 4: Potential hypothesis for group co-presence scenarios.

A possibility that emerges from these findings is the exchange of the Human Agency axis with a Number of Agents axis (see Figure 3 and 4). Related to this possibility is Asch's (1953) experimental work on the relationship between group size and group pressure. In a series of experiments, a lone participant is "placed in the position of a *minority of one* in the midst of a *unanimous majority*" (p. 152, Asch). The majority is a group of confederates who

unanimously provide the wrong answer in a shared forum to a simple measurement problem. Of interest is the response given by the lone participant. In 32% of cases participants chose an incorrect answer even though the problem had an obvious correct answer. Asch then went on to manipulate the size of the unanimous majority to see how the size of the majority impacted the number of errors made by the lone participant. He found that the effect was in full force with a majority of three and larger group sizes did not increase the error rate. This study introduces the concept of group pressure as an attribute of group size. Additional work by Lau, and Murnighan (1998) suggests that as groups grow the probability of subgroups, smaller groups within the larger group, increases and may lead to additional group effects. Looking back at Figure 3 and Figure 4, it is relevant to consider what might be the potential relationship between the threshold of social influence and group pressure.

Assuming for the moment that a group dynamic effect is evoked in multiple agent environments, what sorts of group phenomena are likely to emerge? From Kelly and McGrath's work on social entrainment, there are a number of low level adjustments people should make in multiple agent environments. "Strauss (1965) has...noted that each member of a social system brings a set of preferences for activity levels and time allotments into the system and that their joint activity patterns will reflect a negotiated temporal order. Engaging in social interaction has also been show to modify an individual's activity level in terms of gaze patterns, voice amplification, and other temporal features of expressive behavior such that complementary or compensatory patterns emerge" (p. 396, Kelly & McGrath). Beyond these low level responses to group interaction, other relevant features of

group dynamics become available for manipulation within a group composed in part of embodied agents.

Blascovich et. al. (2004) tout the opportunities immersive virtual environments (IVEs) afford social psychologists. The ability to easily manifest synthetic environments and replicate studies are features that make IVEs a powerful scientific tool. From the perspective of designing learning environments for education, there is the additional opportunity of being able to directly apply research findings to designed educational experiences. Software systems that evoke relevant psychological and learning findings can be more directly transferred to educational software environments than findings that must be translated via interfaces and technology. As it is possible to unsuccessfully translate good findings into badly designed interfaces, skipping this translation step is highly advantageous.

A current area of interest in the Organizational Behavior literature that is relevant to the Triple A Challenge is Edmondson's linking of team psychological safety with team learning behavior. Edmondson defines team psychological safety as "a sense of confidence that the team will not embarrass, reject, or punish someone for speaking up. The confidence stems from mutual respect and trust among team members" (p. 354, Edmondson). She identifies team learning behavior as team members "seeking feedback, discussing errors, seeking information and feedback from customers and others" (p. 354, *ibid.*). The Triple A Challenge currently has a mismatch between a competitive game-like environment where agents respond to questions in a quiz-like manner and the intellectual task of teaching agents through building and revising complex multi-link concept maps. It may be that the manner

in which agents interact with each other and with their human coaches can create the social norms in which human coaches then operate. It will be interesting to find out if agents who collectively act in a flippant and competitive manner evoke different types of learning behavior in their human coaches than agents who are generally non-defensive, collaborative, and supportive.

Previous work by Rokokai et. al. on virtual peers suggest that students can develop new skills when working with a slightly more competent agent. “Children model literacy skills from a competent partner. Sam acts as that competent partner as she tells stories using more advanced forms of linguistic expression (specifically, Sam uses cohesive pronouns, connectives, and tense and temporal markers in telling her story)”(p. 354, Rokokai et.al.). In a similar way, it may be possible for TAs to model group learning behaviors that lead to enhanced learning outcomes while maintaining the atmosphere of *playing* the Triple A Challenge.

Next steps and research opportunities

This paper argues that a significant design issue for the evolution of TAs within the Triple A Challenge is emergent and manufactured group dynamics. The challenge the Triple A Challenge now faces is how best to manifest these group dynamics between agents and humans. The reality of the Teachable Agent project, as with many non-commercial educational software applications is that resources are limited for programming and production values. Initial forays into virtual group dynamics are unlikely to see humans and agents interacting with parity, as we do not currently have the resources to explore the use of

sophisticated language recognition AI software attuned to group and agent to agent interactions.

Given our existing data collection methods associated with how student coaches interact with their TA, the most obvious areas to begin an exploration of virtual group dynamics is by expanding existing agent-human coach interactions and exploring agent-to-agent communication. Agent-to-agent interactions have the potential to enhance the perception of presence within the game and provide us a method to model teaching strategies and collaboration behaviors to human participants.

Teachable agent: agent or avatar?

In order to determine how agents should act towards each other it is useful to develop a deeper understanding of what exactly a teachable agent is. Unlike more autonomous agents, TAs straddle the boundary between agent and avatar in a unique way. Bailenson and Blascovich's (2004) definition of agents and avatars provides a vantage point from which to consider the challenge of understanding was exactly is a Teachable Agent.

“When a given digital representation is controlled by a human, it is an avatar, and when it is controlled by a computational algorithm it is an embodied agent. Central to the current definition is the ability for real-time behavior, in that the digital representation exhibits behaviors by the agent or human as they are performed.” (p. 64, Bailenson, J. N. & Blascovich).

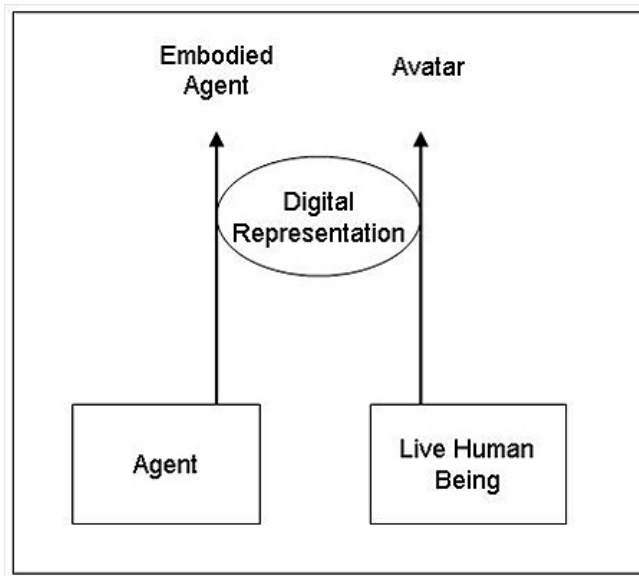


Figure 5: Bailenson, J. N. & Blascovich, J. (2004).

Figure 5 depicts a vision of agents and humans where there is little commingling of agency. TAs on the other hand operate with knowledge taught to them by a human but must also be autonomous enough to maintain the human's sense of teaching another being, contrasted with the experience of programming a computer. The teachable agent is thus both an autonomous agent and a repository of knowledge provided by the human teacher. The agent acts based on its own algorithms but also incorporates the *algorithms* of its teacher. Balancing these two characteristics is crucial to how human teachers interpret their task as a teacher while also seeking to interpret understanding the impact their instruction has on the TA student.

As Persson et. al. (2000) argue, the process of viewing an entity as a human-like being involves an emergent process of anthropomorphism where people make observations about an entity's appearance and behavior and attribute the entity with traits, social roles, and

emotions. This point is also elegantly illuminated by Braitenberg's description of machines with simple internal structures that appear to move with intentionality. By observing a simple vehicle that responds to sensor input, attitudes such as fear, aggression, love and values can be interpreted from the vehicle's behavior. Human's willingness to attribute such behavioral characteristics to animate, but non-biological, entities speaks significantly to our willingness to attempt to understand and interpret animate actions and behaviors. In both of these cases, anthropomorphism is in part an activation of the observing human's existing social schema. People see a behavior and appearance and map this information into their previous knowledge of such behaviors and appearances.

As teachers of TAs, students will attempt to interpret the actions and changes in the behavior of their TAs to determine the success of their instruction. What is new in the virtual group dynamics environment is that teachers now have the opportunity to see how their TAs act, not only in direct response to them, but also in response to other agents and in the future potentially other teachers. The virtual group dynamics environment is an exciting arena in which to consider not only individual co-presence but also group co-presence. Work in this area promises to be productive both for social psychological research efforts and for efforts in constructing new educational technology systems.

Next steps for the Triple A Challenge will be to create a working model of agent-to-agent interaction within the context of the game show. Manipulations to this model will enable us to see better understand what educational possibilities group co-presence can evoke. We are hopeful that the presence of multiple agents can create social contexts in which students

engage in optimal learning behavior and maximize their individual and group learning experiences.

References:

- Asch, S. E. (1953). Effects of group pressure on the modification and distortion of judgments. In Cartwright & Zander (Eds.), *Group Dynamics: Research and Theory*. Evanston, IL: Row, Peterson & Company.
- Bailenson, J. N. & Blascovich, J. (2004). Avatars. In W.S. Bainbridge's (Eds.) *Encyclopedia of Human-Computer Interaction*, Berkshire Publishing Group, 64-68.
- Bailenson, J. N., Swinth, K. R., Hoyt, C. L., Persky, S., Dimov, A., and Blascovich, J. (2004, in press). The independent and interactive effects of embodied agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in Immersive Virtual Environments. *PRESENCE: Teleoperators and Virtual Environments*.
- Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., & Bailenson, J.N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13, 103-124.
- Braitenberg, V. (1984) *Vehicles: Experiments in Synthetic Psychology*. Cambridge, MA: MIT Press
- Davis, J. M., K. Leelawong, K. Belyne, R. Bodenheimer, G. Biswas, N. Vye and J. Bransford (2003). Intelligent User Interface Design for Teachable Agent Systems: A Demonstration Abstract. *International Conference on Intelligent User Interfaces* (pp. 320). Miami, Florida, The Association for Computing Machinery.
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44, 350-383.
- Kelly, J. & McGrath, J. (1985). Effects of time limits and task types on task performance and interaction in four person groups. *Journal of Personality and Social Psychology*, 49, 395-407.
- Lau, D.C. & Murnighan, J.K. (1998). Demographic diversity and faultlines: The compositional dynamics of organizational groups. *Academy of Management Review*, 23, 325-340.
- Persson, P., Laaksolahti, J. & Lönnqvist, P (2000) Anthropomorphism - a multi-layered phenomenon, Proceedings of *AAAI Fall symposium, Socially Intelligent Agents - The Human in the Loop*, North Falmouth, MA.
- Rousseau, D., and Hayes-Roth, B., *Interacting with Personality-Rich Characters*, Stanford University Knowledge Systems Laboratory Technical Report KSL-97-06, September 1997.

Ryokai, K., Vaucelle, C., Cassell, J.(2002). "Literacy Learning by Storytelling with a Virtual Peer". Proceedings of Computer Support for Collaborative Learning. January 7-11, Boulder, CO, pp. 352-360.

Schwartz, D, Blair, K., Biswas, G., Leelawong, K., and Davis, J.(in preparation). Animations of thought: Interactivity in the teachable agent paradigm. To appear in R. K. Lowe & W. Schnotz (Eds), *Learning with animation: Research and implications for design*. New York: Cambridge University Press.

Slater, M., Pertaub, D., Barker, C., and Clark, D. (2004) An Experimental Study on Fear of Public Speaking Using a Virtual Environment. *3rd International Workshop on Virtual Rehabilitation*. Lausanne, Switzerland.