T. Koschmann, D. Suthers, & T.W. Chan (Eds.), *Computer Supported Collaborative Learning 2005: The Next 10 Years!*
Mahwah, NJ: Lawrence Erlbaum Associates. ©2005 International Society of the Learning Sciences.

We Learn Better Together: Enhancing eLearning with Emotional Characters

Heidy Maldonado[1], Jong-Eun Roselyn Lee[2], Scott Brave[2], Cliff Nass[2], Hiroshi Nakajima[3], Ryota Yamada[3], Kimihiko Iwamura[3], Yasunori Morishima[4]

1. School of Education, Stanford University. heidym@cs.stanford.edu
2. Communications Department, Stanford University. {jroselyn, brave, nass}@stanford.edu
3. Control Technology Lab, Omron Corporation. nak@ari.ncl.omron.co.jp, {ryamada, kiwamura}@oas.net
4. Division of Languages, International Christian University. morishima@icu.ac.jp

Abstract. In this paper we explore a new direction for pedagogical computer characters, which we believe will maximize students' learning gains and enjoyment. To the traditional scenario where students interact primarily with a single coach or tutor character on-screen, we introduce the addition of both a social, animate colearner, and the student's own avatar character. Variations of the colearner's attributes, informed by research literature on human partners, are explored through an online testbed application of English language idioms. Results from an experimental study with 76 Japanese college students reveal that cooperative colearners have a positive impact on students' performance and experience, as well as increasing perceptions of the character's intelligence and credibility. Findings provide grounding for a fruitful new direction for pedagogical characters, where students learn alongside emotional companions.

Keywords: Pedagogical computer characters, evaluation of computer characters, language learning, colearners.

INTRODUCTION

Artificial intelligence applications in education most often appear as tutoring systems, diagnosing students' misconceptions on subjects such as algebra through problem solving activities and prearranged prediction patterns. Recently a new breed of agents has begun to surface in everyday applications: these agents appear on screen as embodied entities – whether humans, or anthropomorphized objects and animals – facilitating our interactions with software applications, navigating menus and web-pages, offering tips and customizing our online purchases. These embodied agents converse and interact with humans through text bubbles, sometimes accompanied by sound clips of their utterances, often expressing colorful personalities through animated gestures.

Drawing inspiration from traditional cartoon animation and comic layout processes, these characters resemble actors in their emotional range and improvisational responses. From the Microsoft helper agent, to company spokesperson, to customer service representatives on Dell.com and buy.com, computer characters are rapidly gaining prevalence in our everyday online activities, and with good reason. Computer characters have been shown to be effective proxy sellers, customer service representatives, and teachers, by engaging our unconscious social nature through life-like language, presence and behaviors (Reeves and Nass, 1996; Thomas and Johnston, 1981; Reeves, 2001). Of particular importance for the educational domain, characters, when properly deployed, may generate additional interest and motivation in the content presented (Moreno et al, 2000), and influence a proactive change in everyday behavior (Bickmore et al, 2004).

Pedagogical characters have appeared in language learning applications (Hayes-Roth et al, 2002), middle school curricula (Lester et al, 1997; Moreno et al, 2000; Biswas et al., 2004), oral storytelling (Kehoe et al, 2004), corporate training (Extempo, 2004), health behavior change interventions (Bickmore et al, 2004), and even military instruction (Johnson et al, 2004).

Previous research and industry applications on embodied conversational computer characters in pedagogical domains has conceptualized the interaction between the human and the computer character as a one-to-one tutoring or coaching intervention. The learner primarily interacts with one character on the screen at a time; when more than one character is present simultaneously in the application, their roles tend to be supplementary or supportive ones, providing background tips for the interaction as an articulate "Help" menu option or enriching the background of the activity as non-interactive extras, often called non-player characters.

We believe that giving the student a virtual presence in the environment, and enriching the learner's world with a colearner will lead to greater gains in learning and enjoyment of the educational application. To that end, we have designed and built an online application testbed, where we vary characteristics of the colearner to evaluate research-based predictions of factors that maximize achievement and enjoyment. We report on the first experimental study using this eSchool application, focused on teaching Japanese college students American idioms. The results offer great promise for breaking free of the traditional paradigm of a single interactive character on the screen, while revealing further research questions for pedagogical characters .

In the archetypical pedagogical character application, a computer character on the screen presents material, situations and questions, through images, videos, text and voice. The character may be embodied, or represented on screen as photographic images and videos, photorealistic 3d images, and even 2d anthropomorphized animal characters. To respond, human learners type comments, or choose from options presented by menus or buttons. Feedback on the learner's performance is then delivered through the characters, such as an explicit rating or goal achievement scale, or through the environment itself. For example, in certain simulations, the learner's behavior impacts the scenario illustrating, through the on-screen characters' behaviors, potential consequences and nuances of interpretation, enriching traditional numerical ratings with affective performances. This feedback can be both formative, as the learner progresses through the interaction, and often summative at the completion of the curriculum.

Computer characters exhibiting realistic behaviors in the pedagogical arena tend to follow three primary models, mirroring Taylor's 1980 taxonomy of computer usage in schools: tutor, tool and tutee. Characters may be cast as expert teachers or coaches, presented as role-play partners in simulations of real-world situations, or act as learners whom the human student teaches. Most often, computer character inhabit the role of expert or knowledgeable teacher (Hayes-Roth et al, 2002b; Baylor et al, 2004; Lester et al, 1997; Moreno et al, 2000). While these characters may be embodied as older or of a similar age to the human student, they interact with the students as experienced coaches and erudite tutors.

The second category, where characters are presented as partners in role-play simulations is prevalent for on-the-job training and eLearning situations, where the learner may be older and more experienced. In these cases, rather than an expert teacher, computer characters act as colleagues and coworkers. These characters foster learning through realistic role-play scenarios as the human student practices concepts, strategies, and behaviors (Extempo, 2004; Johnson et al., 2004; Aldrich's Simulearn company). A third mode of interaction with computer characters prevalent today presents the computer character as a learner, progressing alongside the human student (Maldonado et al, 2004), or being taught by him or her (Biswas et al, 2004). Peer characters, unlike the expert coach and tutors, are perceived as possessing as much content knowledge on the subject as the human learner, and often less. In cases where students teach the peer character, the cyclical act of preparing content for teaching, and successfully communicating it becomes the learning experience in itself.

LEARNING WITH CLASSMATES

In the last twenty years since Taylor published his taxonomy, we have developed the technologies to implement and render believable, animate characters, yet we have made comparatively little progress in developing applications for these characters that maximize the social and emotional relationship with human interactors. For each of the three modalities described above the human student tends to be implicitly present alongside the character, rather than visually embodied on screen, and the interaction is limited to exchanges between the human learner and a single computer character. Graphical representations of the human users that appear on-screen, are often referred to as avatars, which can be directed in their interactions with other computer characters in the environment, whether autonomous or directed avatars themselves. Despite few occasional appearances in educational applications (Maldonado 1998a, 1998b; Johnston et al, 2004), avatars frequently appear on many commercial sociable applications (Clanton et al, 2003; DiPaola, 1999; Wright, 2003), with varying degrees of autonomy: users may direct the avatar to perform specific behaviors and utterances or, at a higher level, indicate a preference or direction and let the avatar fill in with appropriately corresponding behaviors, gestures, onomatopoeias and comments.

We propose a radical reframing of the learning context in which students interact with animated pedagogical computer characters, which we believe will maximize the students' learning gains and enjoyment. In the eSchool application we have developed, students interact with *two* computer characters at the same time, a teacher and a peer fellow student, within one screen. Students are themselves represented through an avatar, whom they can direct in emotional and subject matter responses. The avatar's conversational behaviors and gestures are autonomously derived from the directions given by the human learner and the learner's answers to the teacher's questions. Research suggests this combination of high-level directions and autonomous behavior is perceived as more natural than avatars whose behaviors are minutely controlled by the users, leading to increases in perceived expressiveness of the conversation, and greater sense of user control (Cassell et al, 1999).

Figure 1 shows the eSchool interface for handheld computers from 2003 using cartoon characters, and Figure 2, the current interface on a PC using photographic images at the exact same moment in the interaction. In both cases the teacher character is located on the upper right hand side of the screen, above the chalkboard, which, in turn, displays the multiple choice interpretations the student must decide between, mimicking the learning space of a classroom. In Figure 1, the students' avatar is Neko the cat, and the colearner is Taro the Tiger, while in Figure 2 the students' avatar is Susy and the colearner is Ryota.

Each of the three interactors has their own emotions and embodiment, built for adaptation across platforms and conditions of analysis. The eSchool environment is written in Java, allowing students to conduct all their interactions online, perhaps across different days and computers as they progress through the lessons; the system will record the state for every student that logs in through a database driven registration process.

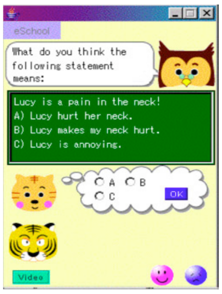


Figure 1: eSchool 2003 for Handhelds featuring cartoon characters

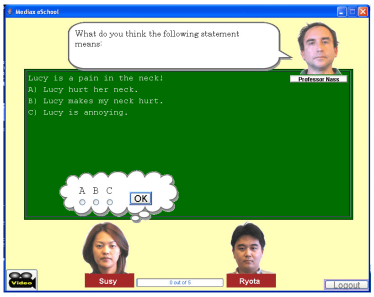


Figure 2: eSchool 2004 for the PC featuring photorealistic characters

Figures 1 and 2 show only some of the possibilities available within eSchool: students can choose the look and feel of their avatar and their colearner from a wide range of embodiment options: nine different anthropomorphized animals in 2003 and ten photorealistic humans in 2004, five male and five female. The actions the avatar and colearner take affect each other’s emotions accordingly, and this change is reflected in their facial expressions, following the model of Ekman’s (1992) universal or basic emotions: anger, sadness, disgust, surprise, fear, and enjoyment or happiness. Answering a question right, for example, increases the student’s confidence and happiness, and also affects the emotions of the autonomous peer colearner, depending on his or her personality traits. The emotion model is grounded on personality traits, and the underlying architecture that modulates the emotions of all three characters is a novel, patent-pending design, product of a unique collaboration between Omron Corporation and Stanford University (for more implementation details, please see Morishima et al, 2004; Nakajima et al, 2004), and allows for a synchronous learning experience as well. That is, two students can interact through their avatars online in eSchool, although for the purposes of our experiments we have isolated the behaviors under analysis by modifying specific attributes of the colearner as an autonomous character. Unlike human confederates sometimes used in experimental studies, characters can reliably showcase the same pattern of behavior, with slight improvisations, for every user.

The colearner, or classmate, as its name indicates, is learning alongside the human student and has no additional knowledge on the subject nor explanations to provide. It is not privy to additional information than the student views on the screen, and is called to answer questions as often as the student, on average. The colearner exchanges friendly banter with the student’s avatar before the teacher poses a question, or after the answer is revealed. We have introduced this presence in the learning environment to determine the effects that learning alongside a social animate character has upon the student’s performance and attitudes, rather than using the colearner to elicit nuanced explanations from the student, or model understanding, as is the case with some intelligent tutoring systems (such as Goodman, et al., 1998).

The focus of our eSchool project on the coleaner is an innovative aspect of our character-based system, a research direction we believe fruitful to explore for computer supported educational interventions in a wide range of contexts, from eLearning seminars, distance education programs, to after school activities, in-class complementary modules, and behavioral change interventions. Thus far, experimental research has focused primarily on the modality of the exchanges between the teacher character and the human student, as well as the character’s believability: text vs. voice (Moreno et al, 2000), photo-realistic vs. drawn representations (Baylor, 2004); two vs. three dimensional presence (Shinozawa et al, 2003); human-like vs. anthropomorphic animals (Moreno et al, 2000). Some studies have varied the degree of animated behaviors (Lester et al, 1997; Moreno et al, 2000), and others have explored ethnic and gender combinations between the learner and the character tutor (Nass et al, 2003; Baylor, 2004). In contrast, we are conducting several experiments evaluating the effects that different dimensions of the colearner have on the students’ performance. Among the colearner dimensions we are studying are social ones (competitiveness and cooperativeness), personality traits (degrees of introversion and extroversion), and performance (high- to low-achiever). We are seeking out instances where colearners influence students’ learning outcomes and attitudes through their behaviors, with the aim of gradually expanding the focus to study how these behavioral expressions relate to the students background, performance, personality profile and preferences.

PILOT APPLICATION

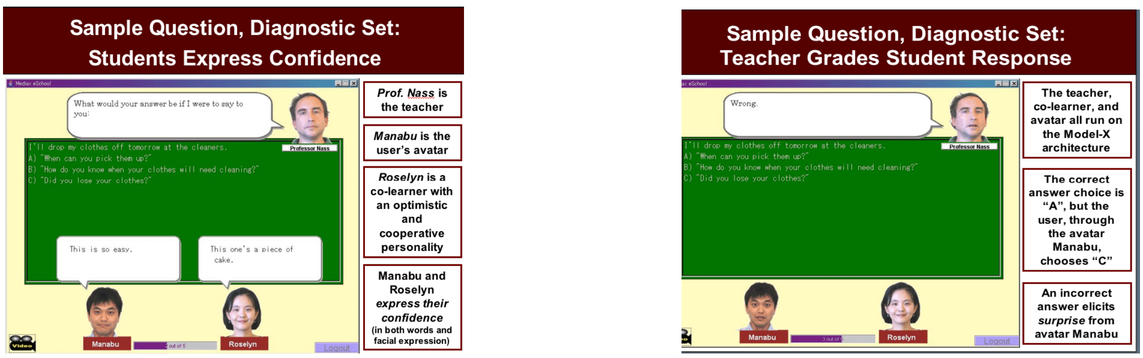
We inaugurated the colearner research field by isolating the effects of the autonomous colearner in an experimental study with three conditions: in the first of these, the colearner was not present, and the student interacted directly with the teacher character. We call this control condition the “No Colearner Present” condition. In the second condition, the colearner was present but did not interact with the student. While it answered questions when called on, in this “No Emotion Colearner” condition, the colearner character did not exchange or respond to any social banter with the student avatar in neither words nor gestures. Researchers such as Zajonc (1965) have noted that the presence of an audience tends to enhance performance and inhibit learning, although even he points out that some studies suggest an increase in learning scores in the presence of other people who learn with the individual. With the "No Emotion Colearner" condition we sought to study how the mere social presence of the colearner, even when it did not interact directly with the student, would affect the learning experience.

Research suggests that studying in a cooperative group leads to greater learning gains than individual or competitive conditions, with improvements in critical thinking skills (Stockdale and Williams, 2004; Skon et al, 1981), and to greater gains in motivation (Bickmore, 2005; Wentzel, 1997). Accordingly, our third condition featured a cooperative colearner. While most previous research attributes the performance gains of students in a cooperative group to shared reasoning and dialogue, we are interested in exploring the effects that colearners' social banter may have on the student's performance, even when it is devoid of educational content relevant to the learning activity at hand. Therefore, our third condition is the "Cooperative Colearner" condition, where the colearner's social nature was manifested through utterances directed at the student's avatar, making compliments and showing concern, particularly as the difficulty level of the questions increased or when the student missed a question. For example, when the student through his or her avatar answers a question correctly, the collaborative colearner may utter one of ten supportive utterances, such as "I knew you'd get it right!" or "That was hard, and you got it!" If the student's avatar had answered incorrectly, the collaborative colearner may say: "This is very tough," "I didn't know that one either," "You'll get the next one!," "I would have given the same answer, this is hard," among other possibilities.

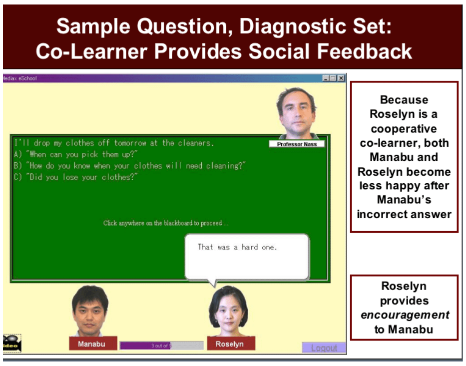
In addition to supportive utterances, in the "Cooperative Colearner" condition, the colearner also expresses support through appropriate changes in its facial expression. By creating a sense of unity between the self and the other, cooperativeness promotes friendship and perception of social support, which in turn may contribute to enhancing social relationship and performance both in dyadic and group interaction (Argyle, 1989; Argyle 1991; Deutsch, 1949). Because we sought to isolate the effects of the colearner character's emotive, the teacher character in this evaluation did not exhibit any emotions, in his speech or expression.

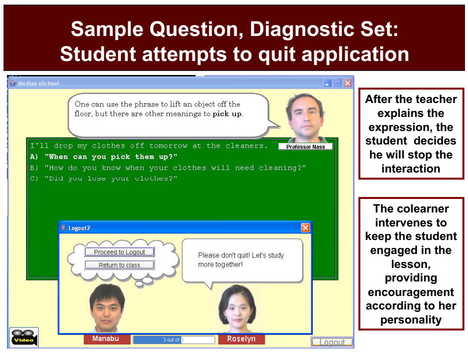
For the current eSchool web-based implementation, we have developed an intermediate English language lesson, accompanying video, and evaluation instruments, primarily aimed at Japanese college students enrolled in English language courses at Japanese universities. Students progress through fifteen multiple-choice questions about American idioms, a topic whose relevance to the target population was predetermined through a focus group study. Our target population has already mastered grammatical and syntactical nuances of the language and is interested in improving their understanding of colloquialisms commonly used by their American counterparts. As Figure 1 highlights, the interface design of emotive cartoon characters for eSchool 2003, was explicitly developed in Japan to be appropriate for the target audience, in terms of aesthetics, experience, culture and age. As we shift our focus globally with the online release of eSchool 2004, available to students anywhere, we have redesigned the interface for appeal to a broader audience, with photorealistic characters, and an emphasis on emotional expressions based on globally recognizable features (Ekman, 1992).

The idioms covered in eSchool lesson plan currently include: "being a pain in the neck," "to drop off an item," "to drop in," "to drop out," "getting cold feet," "getting up the nerve," "being tongue tied," "being chicken," "being a scaredy cat," "being a couch potato," "pulling someone's leg," "hitting the books," and "hitting the town." These idioms are organized in three groups of five questions each, according to their difficulty and grammatical similarity. Progress through the diagnostic, basic, or advanced question sets, is represented graphically on the screen through a progress bar between the images of the avatar and colearner, as can be seen in Figure 3.



Figures 3a and b: Typical question-answer interaction in eSchool 2004, featuring a cooperative colearner. Highlights the avatar's emotional reactions to the questions posed by the teacher.





Figures 3c and d: Typical question-answer interaction in eSchool 2004, featuring a cooperative colearner. Showcases the cooperative comments and expressions of the colearner.

After the initial diagnostic questions, a video featuring an American college student discussing graduation with a Professor is played, using several typical idiomatic expressions. Students are able to see the video as many times as they wish, in its entirety or by choosing segments to review, and we record the viewing pattern for post-experience comparisons. Ten questions of increasing difficulty follow, based on the expressions covered in the video. The student's answers to each question are stored in the eSchool database for comparison as well, and affect both the confidence and emotional state of the characters. For example, in Figure 3a we see both the avatar and the colearner autonomously making statements of confidence, after the teacher has posed a question, based on their internal emotion model.

The teacher character then calls on the student or the colearner to reveal their answer on screen, determines its appropriateness for the question, and explains the origin or usage of the expression. On average, the teacher will call on the learner for 60 percent of the questions, while the remaining 40 percent will be answered by the colearner, which has been programmed to have a 50 percent chance of guessing the correct response each time s/he is asked to reveal the answer. Independently of the correctness of the answer revealed, the same teacher explanation will be seen by every student and colearner, to ensure that every participant is exposed to each aspect of the lesson content. For example, in the previous question scenario, the teacher explanation is as follows: "To drop something off is to leave something in a place for others to pick up. One can use the phrase to lift an object off the floor, but there are other meanings to *pick up*. One can use it to mean collecting an object from a place where it has been left, or where it is being fixed."

In the previous example, avatar Manabu reveals his choice when called to present an answer. Given the high level of confidence we can observe in Figure 3a, when the student's respond choice is incorrect, avatar Manabu expresses surprise, as we can see in Figure 3b; this emotional expression is appropriately and autonomously produced by the emotion engine, rather than a direct input from the student. Immediately following Manabu's surprise, the colearner makes a socially appropriate cooperative comment, captured in Figure 3c ("That was a hard one"). When the discouraged student tries to quit the application, the colearner intervenes to entice the student to continue the lesson, in a socially appropriate manner. In Figure 3d we can see the choices on avatar Manabu's bubble ("Proceed to logout" and "Return to class") and colearner Roselyn's plea to continue: "Please don't quit! Let's continue studying together!"

EVALUATION AND RESULTS

For our first evaluation of the eSchool system and underlying emotion-generation architecture, we partnered with an English language college class at International Christian University, in Tokyo, Japan, and sought to determine the impact of the colearner character on the students' understanding, recall, recognition, and motivation. Seventy-eight students were randomly assigned to one of three conditions; of those that chose to respond, 25 were male and 51 were female. Of these 76 respondents, 25 students interacted with the "Cooperative Colearner," a different group of 25 interacted with the "No Emotion Colearner," and a third group of 26 students interacted with the eSchool system in the "No Colearner Present" condition. Students interacted with the eSchool system at their school's computer lab, and then answered an online questionnaire. The entire process of interaction within the eSchool environment and questionnaire response lasted approximately an hour. The questionnaire covered attitudinal responses to the software system, probed their perceptions of the colearner character, if there was one present, and included a learning assessment metric with open-ended questions. These latter questions are of particular interest for our evaluation, as we sought to differentiate a gain in understanding and learning from rote memorization, recognition, and chance guesses, that plague typical recall multiple-choice assessment instruments. Students were asked to fill in the blanks in 11 sentences using some of the colloquial idioms covered in their interaction, completing grammatically correct sentences similar in meaning and structure to those presented through the lessons. For example, one of the fill-in-the-blank questions was "How could you believe what he said? He was just ________________." (The correct answer would be "pulling your leg".)

Before analyzing the results, two preliminary checks were conducted to certify the study results. Given concerns on the differential number of participants' by gender, a contingency-table Chi-square test was conducted to determine that the gender did not have particular effects on the intervention results. The observed Chi-square value indicated that the participants' gender was balanced across conditions (χ^2 (2) = 1.59, ns.) Because of the importance placed on the learner's feelings of being supported and cared for in the literature and in our research-based design, we included items on the questionnaire to validate our belief that our cooperative colearner was perceived as such. The participant's ratings of the colearner characters as "cooperative," "warm," and "caring" on a 10-point scale were compared based on an additive index of the three items. Because the control condition did not feature a colearner character, it was excluded from this comparison.

For this manipulation check, the participants in "Cooperative Colearner" condition (M= 6.04, SD= 2.84) rated their colearner as more cooperative than those who were in "No Emotion Colearner" condition (M= 3.12, SD= 1.88), and the mean difference was statistically significant (with equal variances not assumed according to the Levene test for equality of variances) t (43.46) = 4.38, $p < 0.001$. Second, with respect to the rating of the colearner character being "warm", the mean score of those who participated in the idiom lesson in "Cooperative Colearner" condition (M= 6.62, SD= 2.48) was significantly higher than the participants in "No Emotion Colearner" condition (M= 4.80, SD= 2.31), with equal variances assumed, t (49) = 2.70, $p < 0.01$. In addition, the participants in "Cooperative Colearner" condition gave a significantly rating of the colearner character as "caring" (M= 5.31, SD= 2.21) than those who were in "No Emotion Colearner" condition (M= 3.38, SD= 1.86), t (50) = 3.40, $p < 0.01$. From this analysis, we conclude that participants accurately perceived the substantial differences in the treatment variable, and interpreted the emotive colearner to be "cooperative," "warm," and "caring," as designed in the "Cooperative Colearner" condition.

Yet as characters in books, theater and television demonstrate, viewers can attribute cooperation and friendliness to performers without perceiving these attributed feelings as directed to themselves, the audience. Therefore, we set out to ascertain that the students felt supported by the colearner throughout the interaction. Since studying in a cooperative group leads to greater learning gains than individual or competitive conditions, with improvements in critical thinking skills (Stockdale and Williams, 2004; Skon et al, 1981), we were interested in whether our cooperative colearner was perceived as cooperative *with* the participants when compared with the unemotional colearner.

As feelings of support, and of being cared for have profound effects in cognition, emotion and even physiology, effects that are particularly relevant in educational settings where motivation is key (Bickmore, 2005; Wentzel, 1997), an index of "feelings of being supported" was created based on three items ("Not Alone," "Praised," and "Supported"). The participant rated how well the given adjectives described their feelings during their interaction with the colearner on a 10-point Likert scale: constructed to indicate that the higher the score, the more emotional support experienced. A factor analysis showed that the three items were loaded on a single factor, and the reliability test also indicated that the index could be reliably used (Cronbach α = 0.74). Hence, the "Cooperative Colearner" condition (M= 6.04, SD= 1.75) and the "No Emotion Colearner" condition (M= 3.90, SD= 1.07) were contrasted based on an independent sample t-test. The result (equal variance assumed according to the Levene test) revealed that the participants in "Cooperative Colearner" condition gave significantly higher ratings in feelings of being supported, t (50) = 4.47, $p < 0.001$, as can be seen in Figure 5. This result confirmed our prediction that the participants in "Cooperative Colearner" condition would experience the feelings of being supported to a greater degree than those who participated in the idiom lesson with a colearner character that neither uttered nor showed emotional expressions. Moreover, feeling supported and interacting with a cooperative colearner also impacted significantly students' learning during the intervention.

For the learning assessment, we concentrate on the eleven open-ended questions posed, as they signal a deep dominion of the idiomatic expressions presented interactively through the eSchool lessons. A preliminary analysis of the fill-in-the-blank responses led us to drop two questions due to misleading or confusing cues, as more than 80 percent of the respondents used content outside the lesson to answer. Therefore, only the responses to the remaining nine open ended questions were used in this analysis. Students typed into the text-boxes provided expressions that they believed would make the partial sentences presented sensible and grammatically correct, which resulted in wide variations of answer format. Each answer to the open-ended questions was assessed on a five-point scale, which ranged from "0" (left blank or irrelevant) to "4" (perfect answer), so that the maximum score possible in this section of the questionnaire was 36 points. Participants received partial points depending on the number of grammatical or spelling mistakes made per answer; two coders worked independently on the data set to ensure reliability. As few participants received scores above 25 points or below 12, we considered this assessment measure to be of adequate difficulty for our sample population.

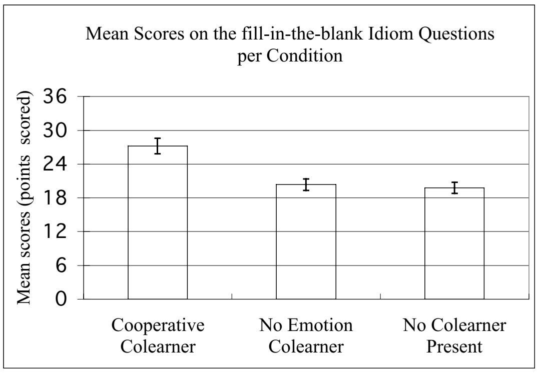


Figure 4: Participants' mean scores on the fill-in-the-blank questions.

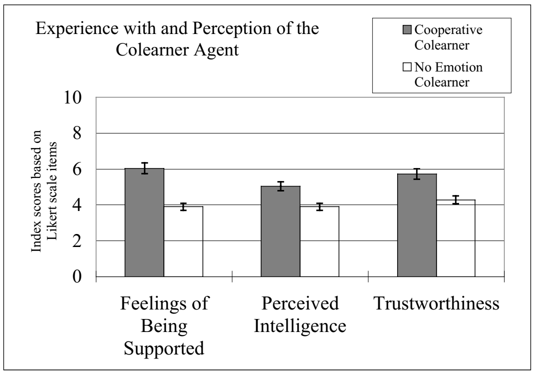


Figure 5: t-test comparisons between "Cooperative Colearner" and "No Emotion Colearner"

When participants' scores on these nine items were compared across the three conditions, the One-way ANOVA test indicated that the means differ significantly, $\underline{F}$ (2, 63) = 4.80, $\underline{p}$ < .05, $\underline{\eta}^2$_ = .13. In order to examine the differences more specifically, a post-hoc comparison was conducted using Scheffé test, one of the most conservative post-hoc comparison techniques. The result showed that the participants in "Cooperative Colearner" condition (M= 27.21, SD= 8.03) performed significantly better on the fill-in-the-blank questions compared to those who were in "No Emotion Colearner" condition (M= 20.35, SD= 9.97) and in the control "No Colearner Present" condition (M= 19.79, SD= 7.26), as can be seen in Figure 4.

For learning at least, it would seem that the mere appearance of an additional face does not trigger the same level of motivation unless it expresses emotions. And interacting with an unemotional colearner also diminished enjoyment of the system: in the attitudinal questionnaire participants filled out, six 10-point Likert scale items probed their feelings about the eSchool environment. Students were asked to quantify their agreement with statements such as "I would recommend this software system to other people," and "I would use this software again." The analysis of covariance of these items covering participant's impression of the system showed that the participants felt it was significantly less enjoyable to work with the system in the "No Emotion Colearner" condition, compared both to the "Cooperative Colearner" condition and the "No Colearner Present" condition [F(1, 76) = 5.25, p < .03]

Participants in the "No Emotion Colearner" condition also rated their colearner as less trustworthy and intelligent than those in the "Cooperative Colearner" condition. Since credibility and intelligence are considered key in persuasion and motivation, it should come as no surprise that these participants did not perform as well in the post-experience evaluation. The index for measuring the perceived trustworthiness of the colearner character was created based on three 10-point Likert scale items describing the colearner character ("Honest," "Sincere," and "Trustworthy"). A factor analysis demonstrated that there was a single factor extracted, and a reliability test calculated revealed that the index could be considered as a reliable measure (Cronbach $\underline{\alpha}$ = 0.82). A t-test revealed that the "Cooperative Colearner" condition (M= 5.72, SD= 1.76) and the "No Emotion Coleaner" condition (M= 4.28, SD= 1.88) that there was a significant difference between the conditions (equal variance assumed according to the Levene test). As can be seen in Figure 5, the cooperative colearner was perceived as more credible and trustworthy than the no-emotion colearner, $\underline{t}$ (49) = 2.28, $\underline{p}$ < 0.01.

In order to examine the perceived intelligence of the colearner character, an index, which consisted of three 10-point Likert scale items ("Smart," "Intelligent," and "Confident"), was created. A factor analysis showed that a single factor could be extracted from the three items, and a reliability test calculated for the three items demonstrated that the index could be used as a reliable measure (Cronbach $\underline{\alpha}$ = 0.70). When the "Cooperative Colearner" condition (M= 5.04, SD= 1.93) and the "No Emotion Colearner" condition (M= 3.90, SD= 1.07) were compared, the t-test result (equal variance not assumed according to the Levene test) revealed that there was a significant difference, as can be seen in Figure 5. The "Cooperative Colearner" character received significantly higher ratings in terms of perceived intelligence, $\underline{t}$ (38.91) = 2.64, $\underline{p}$ < 0.05.

FUTURE DIRECTIONS

Now that we know colearners – even autonomous ones – can impact students' performance, we are looking forward to following up this experimental study to determine other characteristics of colearners that may also contribute towards enjoyment of the experience and learning gains. An overarching goal is to find out if it will be possible to design a colearner in the future that will maximize learning for every student it interacts with, regardless of age, culture, subject matter, preferences, or personality traits. Perhaps such as super-colearner is not feasible, yet we would be delighted to shed additional light into the characteristics of successful dyads in academic environments, real or virtual. Through this new line of research with pedagogical characters we may be able to realize the promise of personalized learning through companions, recommending a combination of colearner attributes to maximize enjoyment and achievement based on a short personality quiz and background profile.

There are many factors that difficult arranging and evaluating such matches in real-world classrooms: among them, students social networks and status differentials, heterogeneity of student responses and high levels of distractions. Heterogeneous groupings in classrooms are often perceived as detrimental for some of the group members, whether in terms of racial and social tensions, smaller performance gains, and additional distractions, among others (Carter and Jones, 1994). Yet in online teaching environments it may be possible to customize the colearner to best fit the needs of every learner, and to progressively adapt should these change.

Within our group we will continue to explore dimensions of the colearner, cautiously controlling for interaction effects among these. We are already planning to evaluate four of these dimensions shortly: personality traits, cross-cultural comparisons, effects of other types of social banter beyond cooperative comments and of the colearner's performance. As a first step, we have continued strengthening and adding value to our eSchool platform, so that future experiments will examine potential matches and mismatches of the colearner's personality traits with those of the participant. Earlier research has showed preferences for interactions with characters whose personality matches the experiment participants' (Reeves and Nass, 1996), and we would like to extend this research by evaluating whether students learn more when they are matched with the colearner or mismatched, even if their enjoyment of the experience is diminished.

A second direction that we are very interested in pursuing within our cross-cultural team is that of cross-cultural comparisons. As the Third International Mathematics and Science Study (Stigler et al, 1999) videos demonstrate, countries vary widely in their educational values and practices. While the experiment described above showed remarkable gains in learning and enjoyment for the cooperative colearner, it is possible that these results are culturally-specific. Given the strong collectivism of Japanese society, would students from more individualistic cultures respond as strongly to the "Cooperative Colearner" as the Japanese students in our sample?

A recent experimental study within our lab (Ju et al, 2005) seems to indicate that American students at least, also enjoy interacting with cooperative colearners more than with competitive colearners. Fourty-four undergraduates participated in a follow-up to our pilot study, using a web-based prototype similar to eSchool in their use of a colearner, avatar, and teacher. However, in this emulation of eSchool, the characters appeared on screen next to the chalkboard as stick figures, completely scripted and devoid of facial expressions. Rather than evaluating American idioms in a population immersed in their usage, this stick figure study focused on morse code lessons, and reported promising results for exploring performance variations in the colearners. This was one of the directions we were originally interested in pursuing: as with student dyads (Carter and Jones, 1994), participants with high-achieving colearners performed significantly better in this study than those paired with low achieving colearners. We are very interested in replicating these results, both those varying the colearner's performance and social banter, and hope through this research to enrich the understanding of how the interaction of these different characteristics of the colearners impact participants' learning and enjoyment. Seeking to build a robust corpus of characteristics of colearners, in our next experiment we will explore combinations of cooperative and competitive colearners with introverted and extroverted traits, to determine possible interactions between these two characteristics.

In terms of the cross-cultural direction, while the stick-figure study holds promise, rather than conducting experiments in multiple worldwide sites to determine cultural preferences, we are looking at abstractions from the rich literature of the international math and science education comparative studies (Stigler et al, 1999), as well as psychology and corporate training (Hofstede, 2003) to determine which aspects or dimensions of cultures may be applicable to, and general enough for, our educational goals.

CONCLUSION

The experiment reported in this paper, as well as those we hope to conduct shortly, have three inherent limitations that we would like to resolve gradually. Firstly, a one-time interaction may show an immediate spark in the learning gains that may not be maintained. We, as designers of promising educational software, hope it is indicative of long-term learning and transfer, and would like to re-evaluate the students that interact with eSchool in future experiments several days and weeks after the initial interaction, to determine these potential longer term learning gains. Because motivation effects can impact attainment, we are also very interested in a longitudinal study, where students would interact through the eSchool environment for several weeks, perhaps as part of a college course. If these interactions were voluntary for a sample of the subject population, we could explore whether colearners make the system enticing enough for students to progress on their own.

Secondly, we are considering how to best expand the colearner paradigm to other domains where the social aspect of our interactions with emotional characters can be leveraged, extending it beyond language learning. As mentioned, pedagogical characters have been successfully deployed in middle school curricula applications (Lester et al, 1997; Moreno et al, 2000; Biswas et al., 2004), oral storytelling (Kehoe et al, 2004), corporate training (Extempo, 2004), health behavior change interventions (Bickmore et al, 2004), and even military instruction (Johnson et al, 2004). We are eager to explore other domains where characters offer a unique advantage to traditional drill-and-practice software, motivating and entertaining students through their emotions and improvisational behaviors. We hope to replicate the colearner effects as we break-free from the classroom emulation into richer environments, where other interactive dimensions can be explored.

Among the suggestive findings we report are the implications for characters beyond pedagogical domains. Given the gains in enjoyment, credibility and perceptions of intelligence, character designers in domains where these three areas are critical may want to consider adding a cooperative dimension to their characters. From information kiosks to customer service representative, help and sales characters could all benefit from supporting their clients through cooperative behaviors.

Lastly, the world of pedagogical characters has rarely been as populated as in our application, yet we would like to explore interactions beyond the dyad, as well as those where the user is implicitly embodied. By evaluating the effects of the avatar's presence on social exchanges and achievement we hope to grow our understanding of how colearners improve learning environments, and to translate those findings to applications with multiple colearner characters. While our innovative focus thus far has been on dyadic interactions, the educational literature suggests that small group learning may be as productive, if not more. Nowadays, most software platforms and applications are capable of handling several characters operating simultaneously, which brings closer the potential for future evaluations with multiple colearners. We welcome the challenge of replicating the achievement and enjoyment gains colearners bring without losing our human participants in the virtual melee.

ACKNOWLEDGMENTS

We thank everyone that made this collaboration possible through Media X at Stanford University, and everyone who was involved with the design and implementation of the system and pilot studies, as well as the Kozmetsky Global Collaboratory Research Assistantship. Among them, Professors Roy Pea, Byron Reeves, and Terry Winograd; students Wendy Ju, Gary Tan, Elizabette Amaral; administrators Kathy Lung and Joan Ferguson from Stanford; Omron personnel, Shigeyasu Kawaji, Thomas Kurian, Ritsuko Nishide, Masaki Aarao, Hideki Kunii, Tatsuke Tateishii, Yuriko Iwashita, and from eBase, Wataru Hasegawa and Keiko Hayashi.

REFERENCES

Aldrich, C. (2004) *Simulearn's Virtual Leader*. Retrieved November 10, 2004 from http://www.simulearn.net/leadershiptraining.html

Argyle (1989). *The social psychology of work*. Harmondsworth: Penguin.

Argyle (1991). *Cooperation: The basics of sociability*. New York: Routledge.

Baylor, A.L. (2004) Designing pedagogical agents to address diversity in learning. Proceedings of the Sixth International Conference of the Learning Sciences, 2004. Lawrence Earlbaum & Associates. Mahwah, NJ. p. 586

Bickmore, T., Gruber, A., and Picard, R. (in press) Establishing the Computer-Patient Working Alliance in Automated Health Behavior Change Interventions. *Patient Education and Counseling*.

Biswas, G., Leelawong, K., Belynne, K., Viswanath, K., Vye, N., Schwartz, D., Davis, J. (2004) Incorporating Self Regulated Learning Techniques into Learning by Teaching Environments. The Twenty Sixth Annual Meeting of the Cognitive Science Society, (Chicago, IL).

Carter, G., Jones, M. G. (1994) Relationship Between Ability-Paired Interactions and the Development of Fifth Grader's Concepts of Balance, *Journal of Research in Science Teaching* **31**, 8, 847-856

Cassell, J. and Vilhjálmsson, H. (1999). Fully Embodied Conversational Avatars: Making Communicative Behaviors Autonomous. *Autonomous Agents and Multi-Agent Systems* **2**, 1, 45-64.

Clanton, C., Ventrella, J. (2003) Avatar-centric Communication in There. Presentation at the People, Computers and Design seminar, Stanford University (April 4, 2003).

Deutsch, M. (1949). A theory of cooperation and competition, *Human Relations*, 2, 129-139.

DiPaola, S. (1999) Internet-based Interactive Character Design: From Agents to Avatars. Presentation at the People, Computers and Design seminar, Stanford University (December 3, 1999).

Ekman, P. (1992) An argument for basic emotions. *Cognition and Emotion*, 6, 169-200.

Extempo Systems Inc (2004). *Extempo Expert Agents - People Skills Workshop*. Retrieved November 10, 2004 from http://www.extempo.com.

Goodman, B., Soller, A., Linton, F., and Gaimari, R. (1998) Encouraging student reflection and articulation using a learning companion. *International Journal of Artificial Intelligence in Education*, 9, 3-4.

Hayes-Roth, B., Knodt E., and Maldonado. H. (2002) Language learning with an adaptive coach and lifelike conversation partners. Computer Assisted Language Instruction Consortium (CALICO) Courseware Showcase. March 28, 2002: University of California at Davis, CA.

Hayes-Roth, B., Maldonado H., and Moraes M. (2002) Designing for diversity: Multi-cultural characters for a multi-cultural world. Proceedings of IMAGINA 2000 (Monte Carlo, Monaco). 207-225.

Hofstede, G. (2001) *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*, SAGE Publications.

Kehoe, C., Cassell, J., Goldman, S., Dai, J., Gouldstone, I., MacLeod, S., O'Day, T., Pandolfo, A., Ryokai, K., Wang, A. (2004) Sam Goes to School: Story Listening Systems in the Classroom. In Kafai, Y., Saldoval, W., Enyedy, N., Nixon, A.S., Herrera, F. (Eds.) Embracing Diversity in the Learning Sciences: Proceedings of the Sixth International Conference of the Learning Sciences (ICLS 2004), 613. Mahwah, NJ: Lawrence Erlbaum Associates.

Johnson, W. L., Choi, S., Marsella, S., Mote, N., Narayanan, S., Vilhjálmsson, H. (2004). Tactical Language Training System: Supporting the Rapid Acquisition of Foreign Language and Cultural Skills. Proceedings of InSTIL/ICALL-NLP and Speech Technologies in Advanced Language Learning Systems, June 17-19, Venice, Italy.

Ju, W., Nickell, S., Eng, K., Nass, C. (2004) Influence of Colearning Agent Behavior on Learner Performance and Attitudes. Extended Abstracts of CHI 2005. ACM Press: Portland, OR.

Lester, J. C., Convers, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997) The persona effect: Affective impact of animated pedagogical agents. Proceedings of CHI '97, ACM Press: Atlanta, GA. 359-366.

Maldonado, H., Picard, A., Doyle, P., Hayes-Roth, B. (1998) Tigrito: A multi-mode interactive improvisational agent. Proceedings of the International Conference on Intelligent User Interfaces '98, ACM Press, 29-32.

Maldonado, H., A. Picard, and B. Hayes-Roth (1998) Tigrito: A high affect virtual toy. Summary of CHI '98 (Los Angeles, CA), ACM Press. 367-368

Maldonado, H. and Hayes-Roth, B. (2004) Toward Cross-Cultural Believability in Character Design. In Payr, S. and Trappl, R. (Eds.). *Agent culture: Designing virtual characters for a multi-cultural world*, **7**, 143-175. Lawrence Erlbaum Associates, Mahwah, NJ.

Moreno, R. and Mayer R.E. (2000). Pedagogical agents in constructivist multimedia environments: The role of image and language in the instructional communication. *American Educational Research Association Annual Meeting* (New Orleans, LA, 2000).

Morishima, Y., Nakajima, H., Brave, S., Yamada, R., Maldonado, H., Nass, C., Kawaji, S. (2004) The Role of Affect and Sociality in the Agent-Based Collaborative Learning System. In E. Andre, et al. (Eds). *Affective Dialog Systems: Tutorial and Research Workshop* (ADS2004). Springer-Verlag New York Inc.: New York, NY. 265-275.

Nakajima, H. Morishima, Y., Brave, S., Yamada, R., Maldonado, H., Nass, C., Arao, M., Kawaji, S. (2004) Toward an Actualization of Social Intelligence in Human and Robot Collaborative Systems. Paper presented at the IROS 2004, IEEE/RSJ International Conference on Intelligent Robots and Systems on Oct 2, 2004

Nass, C. and Najmi, S. (2003) Race vs. culture in computer-based agents and users: Implications for internationalizing websites. Stanford University: Stanford, CA.

Reeves, B. and Nass C. (1996) The media equation: How people treat computers, televisions, and new media like real people and places. 1996, New York: Cambridge University Press.

Reeves, B. (2001) *Conversational agents online: Automating human social intelligence in web transactions.* Invited talk, Symbolic Systems Forum, Stanford University, CA. (November 15, 2001).

Shinozawa, K., Reeves, B., Wise, K., Lim, S., Maldonado, H. (2003) Robots as New Media: A Cross-Cultural Examination of Social and Cognitive Responses to Robotic and On-Screen Agents. Proceedings of the 53rd Annual Conference of the International Communication Association, Information Systems Division. San Diego, CA. 998-1002.

Skon, L., Johnson, D.W., & Johnson, R. (1981). Cooperative peer interaction versus individual competition and individualistic efforts: Effects on the acquisition of cognitive reasoning strategies. *Journal of Educational Psychology, 73*, 83-92.

Stigler, J., Gonzalez, P., Kawanaka, T., Knoll, S., and Serrano, A. (1999) The TIMSS Videotape Classroom Study: Methods and Findings from an Exploratory Research Project on Eighth-Grade Mathematics Instruction in Germany, Japan, and the United States. National Center for Educational Statistics, Office of Educational Research and Improvement, U.S. Dept. of Education.

Stockdale, S. and Williams, R. (2004) Cooperative learning groups at the college level: differential effects on high, average, and low exam performers. *Journal of Behavioral Education,* **13**, 1, 37-50.

Taylor, R. (Ed) (1980) The computer in school: tutor, tool, tutee. Teachers College Press, New York, NY.

Thomas, F. and O. Johnston (1981) The illusion of life: Disney animation. New York: Hyperion Books.

Wentzel, K. (1997) Student Motivation in Middle School: The Role of Perceived Pedagogical Caring. *Journal of Educational Psychology*, **89**, 3, 411-419.

Wright, W. (2003) Design lessons from The Sims and SimCity titles. Presentation at the People, Computers and Design seminar, Stanford University (May 2, 2003).

Zajonc, R.B. (1965). Social facilitation. *Science,* 149, 256-268.