Tech., Inst., Cognition and Learning, Vol. 2, pp. 291-314
Reprints available directly from the publisher
Photocopying permitted by license only

THE PSYCHOMETRIC STRUCTURE OF PEDAGOGICAL AGENT PERSONA

JEEHEON RYU[1] AND AMY L. BAYLOR[2*]

1 Institute of Educational Research, Korea University, Seoul Korea

2 Department of Educational Psychology and Learning Systems
Florida State University
307 Stone Building, Tallahassee, FL 32306, USA
(850) 644-5203 baylor@coe.fsu.edu

This paper describes the development and validation of an instrument for measuring learner perception of pedagogical agent persona. After collating items from other empirical studies, exploratory and confirmatory factor analyses were conducted with three samples of undergraduate students working within the Multiple Intelligent Mentors Instructing Collaboratively (MIMIC) agent-based research environment. The final model identified four pedagogical agent persona factors (Credible, Facilitating Learning, Engaging, and Human-like) and two latent variables (Informational Usefulness and Affective Interaction). Results suggest that there are two primary aspects to student perception of a pedagogical agent's persona: 1) its role as a knowledgeable instructor that facilitates learning; and, 2) its provision of affective, human-like interactions in the process. The final model is discussed with respect to its implications for pedagogical agent research and computer-based instructor characteristics.

Keywords: pedagogical agents, human computer interaction, affective computing, persona effect

*Corresponding author

There is a growing interest among computer scientists and other interdisciplinary researchers in developing more lifelike pedagogical agents (animated computer-based interface characters) to create more effective and interactive learning environments. The presence of such life-like characters has been found to have a positive effect on learners' interactive experience. By engaging the learner, such agents can create more meaningful learning experiences (Baylor, 2000; Lester *et al.*, 1999; Towns *et al.*, 1998) and can in turn positively impact learning performance (Atkinson, 2002; Baylor, 2002; Moreno *et al.*, 2001). Their capacity to facilitate more interactive and meaningful communication between learners and the computer contributes to their effectiveness in supporting learning and motivation (Craig *et al.*, 2002; Lester *et al.*, 1999; Moreno *et al.*, 2000).

While researchers continue to determine the particular affordances and constraints of pedagogical agents for learning, it is already evident that pedagogical agents can have a strong positive effect on learners' perceptions of the learning experience. Baylor and colleagues (Baylor, 2001; Baylor & Kim, 2005; Baylor & Ryu, 2003) consistently found that students perceived 3-dimensional animated agents within the MIMIC agent-based environment to be useful, credible, and worthy of their attention. It is speculated that one key advantage of anthropomorphic agents is that their human-like characteristics help create a more positive learning experience. For instance, appealing facial expression and emotive features make lifelike pedagogical agents more communicative with learners and simultaneously provide a strong motivating effect (Baylor, *et al.*, 2005; Elliott *et al.*, 1999; Towns *et al.*, 1998). Similarly, learners collaborating with agents reported increased motivation, even when the agents did not necessarily promote the learning outcomes (Baylor, 2005; Moundridou & Virvou, 2002).

In an effort to measure learners' perceptions of pedagogical agent persona features, there have been several studies that employ scales based on user ratings. Lester and colleagues have conducted extensive studies on the affective impact of lifelike pedagogical agents (Lester *et al.*, 1997; Towns *et al.*, 1998). In one particular study, the researchers used 18 questions to measure learners' perception and found that learners perceived the agent as being very helpful, credible, and entertaining. They concluded that the presence of an animated agent had a strong positive effect on learners' perception of the learning experience.

In addition, van Mulken and colleagues (1998) investigated how the agent's referential expressions affected how credible and entertaining the learners perceived the agent to be, based on 13 questions. Their focus was to

identify how learners perceived the agent's comments when the agent's persona (its human-like aspects) was present or absent. Two different treatments were implemented that differed by the presence of the agent persona. Several significant agent the persona effects were reported. First, the participants in the persona condition showed significantly more positive affect toward the agent than those in the non-persona condition. In addition, participants in the persona condition found the agent to be significantly more entertaining. Also, those in the agent persona condition were significantly better able to concentrate on the information. Other researchers (Andre *et al.*, 1999b) replicated these results with the same question items used in the previous study (van Mulken *et al.*, 1998).

Moundridou and Virvou (2002) assessed learner perceptions of agents with seven question items. They found that the presence of the agent had positive effects on perception of task difficulty. Participants in the agent condition rated the task significantly less difficult than those in the non-agent condition. Specifically, learners working with an agent found it easier to solve problems, despite the fact that there were no differences in learning outcomes. The participants in the agent condition also felt that the system was more useful in helping them improve their problem-solving skills. Craig and colleagues (2002) assessed agent enjoyability through one item and found a positive trend for the persona effect (but only approaching statistical significance). Baylor and Ryu (2003) developed and tested three sub-scales (person-like, engaging, and credible) of pedagogical agent persona while investigating the effects of agent image and animation, but reliabilities of the sub-scales were relatively low (e.g., .68 to .74) and no validation process was employed. A main concern across all of these studies is that the items employed to assess agent persona were not evaluated and had no established reliability and validity.

Recently, Van Eck and Adcock (2003) raised this issue of establishing a reliable and valid scale to measure agent persona effects. They rigorously conducted a factor analysis and constructed the Attitude Toward Agent Scale (ATAS), which consists of two sub-scales: 1) the pedagogical efficacy of the agent; and 2) the agent's attitude toward teaching. To establish the scale, they adapted questions from a human teacher rating scale that assessed how effectively human teachers teach students. However, there are some potential problems with this method. A pedagogical agent is not a human teacher but is a computer-based animated interface character. Consequently, students might have different perceptions and expectations of a human teacher as compared to a pedagogical agent. For instance,

acceptance and/or believability of the agent may differ depending on agent appearance (e.g., represented by a static image vs. animation) (van Mulken *et al.*, 1998) or level of realism (Baylor, Shen, & Huang, 2003). Students also may be affected by the agent's voice, whether it is machine-generated or human (Atkinson, *et al.*, 2005; Mayer *et al.*, 2002; Moreno & Mayer, 2002; Moreno *et al.*, 2000). Since these characteristics may not be of concern with a human instructor, adapting human instructor questions may not be appropriate for assessing agent-based instructors (i.e., pedagogical agents) and may miss some unique features of human-computer interaction. For example, nonverbal communication (e.g., deictic gestures, expression, natural movements) influences learners' perception of pedagogical agents (Baylor, *et al.* 2005). However, such a factor is not as relevant for human instructors because their movements and gestures are already natural.

Overall, there is a need for a pedagogical agent persona scale to assess the specific features of anthropomorphic pedagogical agents that facilitate their instructional persona. Such an instrument should be founded on the theoretical rationales for the pedagogical persona effect and must be valid and reliable. In a series of studies, described next, we 1) identify constructs contributing to the pedagogical agent persona effect; 2) select the model that best explains the relationships between the constructs, among alternative models, through a confirmatory factor analysis with a cross-validation method; and, 3) provide a reliable and validated instrument to measure the pedagogical agent persona. After identifying the constructs of the Agent Persona Instrument (API) through factor analysis, we will describe how the constructs were validated and adjusted after implementing the instrument to two additional groups of students.

METHODS

Three Samples

Three different samples of participants were selected for this study. The first sample, Initial Group (I-Group), was selected to construct an initial factor model. The participants of I-Group (N=80) were undergraduates in a computer literacy course. Two additional groups were included for Confirmatory Factor Analysis (CFA): Agent Role Group (AR-Group) and Agent Image Group (AI-Group). The participants of the AR-Group (N=72) worked with one of three types of agents, differing by role (Expert, Motivator, and Mentor). The participants of the AI-Group (N=188) worked with one of eight agents differing by gender (male, female), ethnicity (Caucasian, African American), and degree

of realism (realistic, cartoon-like). Two different sets of agents were used for the CFA to better cross-validate the identified factor model.

The sample size of the AR-Group (N=72) was over the minimum sample size (N=50) but did not meet the recommended sample size for CFA to conduct the Maximum Likelihood Estimation (MLE). The sample size of AI-Group (N=188) was acceptable within the 100-200 recommended range (Hair *et al.*, 1998). Since the sample size of the AR-Group was smaller than recommended, several considerations were taken during the data analysis. First, a comparison strategy between the AR-Group and AI-Group was used for the model evaluation. We observed how the model fit indices of the AR-Group and AI-Group differed. Second, ominous model fit indices were reported across the AR-Group and AI-Group. Third, the root mean square error of approximation (RMSEA) was estimated because it reflects a correction of the given sample size (Steiger, 1990). Last, the nonnormed fit index (NNFI) and incremental fit index (IFI) were considered as measures of model fit because they are relatively unaffected by sample size (Hoyle & Panter, 1995; Hu & Bentler, 1995).

Participants

The participants of the I-Group consisted of 80 undergraduate computer literacy students (30% Male and 70% Female), with an average age of 19.48 years (SD=1.64). This sample was used to conduct an exploratory factor analysis to identify an initial model. The AR-Group consisted of 72 (12.5% Male and 87.5% Female) pre-service teachers, with an average age of 19.64 years (SD=3.96). The AI-Group consisted of 188 (21.8% Male and 78.2 % Female) computer literacy students, with an average age of 20.51 years (SD=2.77). Approximately 35% of the participants in the AI-Group were undergraduates at a historically Black university. The participant samples are listed in Table 1.

Instrument Items

To develop an item pool for the instrument we collected the instruments used in other studies that investigated pedagogical agent persona (Andre *et al.*, 1999b; Baylor & Ryu, 2003; Eck & Adcock, 2003; Lester *et al.*, 1997; Moundridou & Virvou, 2002; van Mulken *et al.*, 1998), together with the ATAS (Eck & Adcock, 2003). After collating all 66 items from these studies, we selected appropriate items from the initial item pool and revised them where necessary, deleting duplicate items and those that did not specifically measure agent persona. For instance, some items asked about

Sample	*Participants*	*Pedagogical Agents*
I : Initial Group	80 Computer literacy students	Role Differentiation: Expert, Motivator, Mentor 
AR : Agent Role Group	72 Pre-service teachers	Role Differentiation: Expert, Motivator, Mentor 
AI : Agent Image Group	188 Computer literacy students	Image Differentiation: Realism, Gender, and Ethnicity CFC* CFR AFR AFC  CMR CMC AMR AMC  *First Letter: C = Caucasian A = African-American Second Letter: F = Female M = Male Third Letter: C = Cartoon R = Real

TABLE 1
Summary of Participant Samples and Pedagogical Agent Conditions

usability of the learning system rather than perception of the agent. The final initial instrument was comprised of 38 items with a 5-point Likert scale, with 1 for "*strongly disagree*", 2 for "*disagree*", 3 for "*neutral*", 4 for "*agree*", and 5 for "*strongly agree*".

Procedure

As shown in Table 1, participants in the I-Group and AR-Groups were randomly assigned to work with one of three pedagogical agents that differed by role: Expert, Motivator, and Mentor. The participants of the I-Group were instructed to listen to the agent's instruction about how to plan a lesson. After listening to the agent's instruction, participants rated the agent according to the 38 question items. The average time to complete the procedure was around 20 minutes.

Participants in the AR-Group and AI-Group worked within the Multiple Intelligent Mentors Instructing Collaboratively (MIMIC) environment to develop a complete instructional plan with the assistance of the pedagogical

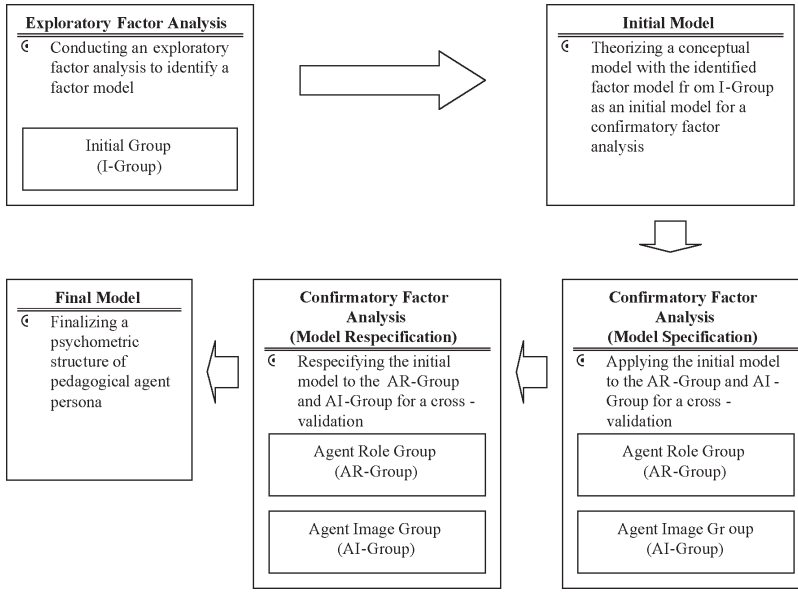

FIGURE 1
Model Saturation Process.

agent. For the AR-Group, the same three agents were provided and were randomly assigned. However, participants in the AI-Group were assigned to one of eight agents that differed by gender (male, female), ethnicity (Caucasian, African American), and degree of realism (realistic, cartoon-like). After completing the learning task within MIMIC, the participants responded to items regarding agent persona. The average time for completing the task and questions was approximately 45 minutes.

The purpose of implementing different agents for the AR-Group and AI-Group was to cross-validate the factor model identified from the I-Group. We assumed that if a factor model fit to the different sets of agents, the model would be evaluated as more robust.

Model Saturation Process

Analysis consisted of four phases to identify the psychometric structure of the pedagogical agent persona as illustrated in Figure 1. First, an exploratory factor analysis was conducted to identify a factor model of agent persona from the I-Group. In the factor analysis, principal component analysis and maximum likelihood analysis were applied. Three different

factor models were compared, and a five-factor model was determined to be the most appropriate factor model.

Second, we theorized a conceptual structure of the five factors from the exploratory factor analysis. Since no prior study had examined the theoretical structure of agent persona, we reviewed the literature regarding how students perceived human instructors. From the literature review, we identified two main characteristics of human instructors: ability (skills, knowledge, and experience) and personality (emotional expression and communication). Each of these characteristics was mapped to the five factors, which was used as an initial model.

Third, the initial conceptual model was assessed by confirmatory factor analysis for both the AR-Group and AI-Groups to verify its psychometric structure. Results from both samples were used for a model specification (see Figure 1) and analyzed with PRELIS 2 and LISREL 8 using the SIMPLIS command for structural equation modeling (Joreskog & Sorbom, 1993). Participant responses were input into PRELIS 2 to create a covariance matrix that was input to LISREL 8 to evaluate model estimates by maximum likelihood estimation (MLE).

Six measures of model fit indices were considered for model fit as suggested by Hoyle and Panter (1995) while considering sample size (Steiger, 1990). As absolute fit indices, Chi-square and the goodness-of-fit index (GFI) were calculated. However, given that the absolute fit indices can be affected by sample size and the distribution of variables (Hoyle & Panter, 1995; Steiger, 1990), several other indices were reported and compared: The root mean square error of approximation (RMSEA), the comparative fit index (CFI), the nonnormed fit index (NNFI), and the incremental fit index (IFI). For GFI, CFI, NNFI, and IFI, values can range from 0 to 1.0, with 0 indicating a poor fit, and 1.0 indicating a perfect fit. Generally, for the indices, values above .90 are considered as an acceptable cutoff indicator of good model fit. However, we took a more conservative aspect by setting the cutoff criterion at .95 (Newsom, 2001). Regarding RMSEA, the test criterion was set at .05 (Tate, 1998). CFI, IFI, and RMSEA were included for the analysis because they are less sensitive to sampling characteristics. This was important to consider for this study since the AR-Group and AI-Group had different sample sizes.

Fourth, after it was identified that the conceptual model failed to provide a solid explanation for the AR-Group and AI-Group, a model respecification was performed with the same measures of model fit indices. From this, the final four-factor model was specified.

I01	Agent was believable.
I02	Agent was enjoyable.
I03	Agent was entertaining.
I04	Agent was enthusiastic.
I05	Agent was expressive.
I06	Agent was friendly.
I07	Agent was helpful.
I08	Agent was human-like.
I09	Agent was instructor-like.
I10	Agent was intelligent.
I11	Agent was interesting.
I12	Agent was knowledgeable.
I13	Agent was motivating.
I14	Agent was personal.
I15	Agent was trustworthy.
I16	Agent was understandable.
I17	Agent was useful.
I18	Agent encouraged me to reflect what I learning.
I19	Agent encouraged me to think for myself.
I20	Agent focused me on the relevant information
I21	Agent helped me to concentrate on the presentation.
I22	Agent helped the instruction make sense.
I23	Agent improved my knowledge of the content.
I24	Agent kept my attention.
I25	Agent led me to think more deeply about the presentation.
I26	Agent made the instruction interesting.
I27	Agent presented the material effectively.
I28	Agent showed emotion.
I29	Agent talked as if we were having a face-to-face conversation.
I30	Agent has a personality.
I31	Agent was concerned with whether I learned the material.
I32	Agent was essential for the information to make sense
I33	Agent was interested in the content.
I34	Agent was tuned to my needs.
I35	Agent's advice was helpful.
I36	Agent's emotion was natural.
I37	Agent's language was easy to understand.
I38	Agent's movement was natural.

TABLE 2
Initial List of Items, by Item Number.

RESULTS

Exploratory Factor Analysis

Several steps were performed to specify the initial factor model. First, two factor analyses for the I-Group (N=80) were conducted to ascertain the factorial structure for the 38 question items, as listed in Table 2.

Principal component analysis and maximum likelihood analysis were performed on the raw scores of the I-Group. From the principal component analysis seven eigenvalues larger than one were found (the eighth eigenvalue was .88). This suggested that a seven-factor model was an appropriate factorial structure according to the "Kaiser rule." However, the location of the "knee in the scree" plot of eigenvalues indicated a four-factor model. Based on this second consideration, a series of three different models, including three, four, and five factors, were evaluated. Table 3 shows the result of model fit with three, four, and five factor models.

The explained variances were 60.86%, 65.10%, and 68.76% for the three, four, and five factor models respectively. The maximum likelihood procedure was performed for the three, four, and five factor models to estimate the initial factor model. However, there were no sound indices showing model parsimony and adequacy of fit from the maximum

TABLE 3
Model Fit Results

Number of Factors	Explained Variance from Principal Component Analysis	Number of Residuals >.05	p-value for Model Fit Test
3	60.86	297 (42%)	< .001
4	65.10	244 (34%)	< .001
5	68.76	179 (25%)	< .001

likelihood analysis. Number of residuals did not show parsimony corresponding between observed correlations and estimated correlations implied by the models. Further, the adequacy of fit index did not indicate a correct model specification.

Although none of the results from the maximum likelihood analysis yielded a clear picture to support any of the three suggested factor models from the principal component analysis, the five-factor model was selected based on several considerations. First, the five-factor model explained 68.76% of the total variance of the 38 question items. Second, the commonalities of all items were well explained, ranging from 0.31 to 0.74.

Since the items were anticipated to show inter-correlations, an oblique rotation method was used to obtain the rotation result. Strength of association between items and subscales were examined to determine if any items should be deleted. The criterion was set at .70, a more conservative level, as used in a similar study (Eck & Adcock, 2003). Eight items were removed across the five factors. Table 4 shows the strength of association between items and subscales. Item number 19 and 22 were deleted from the first factor. Item number 35, 32, and 1 were removed from the second factor. Item number 37 and 29 were deleted from the third factor. Item number 16 was removed from the fifth factor.

All remaining items loaded strongly on the five factors. After deleting the 8 items, an exploratory factor analysis was re-conducted to evaluate the revised set of items for the five-factor model. The explained variance of the revised factor model was improved from 68.76% to 72.66%. The five factors were labeled to represent the agent persona features of *Facilitating Learning*, *Credible*, *Human-like*, *Mentor-like*, and *Engaging,* and are described in more detail below.

- Facilitating Learning: This factor consists of 10 items, relating to how the agent facilitates learning and reflection.

TABLE 4
Strength of Association

Question Items	Factor of Agent Persona Instrument				
	Facilitating Learning	Credible	Human-like	Mentor-like	Engaging
i25	.800	-.426	.448	-.281	.279
i26	.796	-.301	.554	-.337	.512
i18	.786	-.410	.105	-.354	.549
i24	.766	-.218	.447	-.455	.460
i27	.762	-.635	.309	-.419	.368
i21	.761	-.401	.508	-.476	.241
i20	.744	-.665	.188	-.323	.319
i23	.729	-.626	.266	-.411	.342
i11	.711	-.505	.429	-.192	.679
i02	.709	-.288	.386	-.344	.693
i19	.699	-.437	.158	-.485	.449
i22	.696	-.628	.284	-.593	.263
i12	.360	-.890	.270	-.263	.083
i10	.471	-.871	.246	-.284	.142
i17	.504	-.840	.302	-.159	.387
i07	.477	-.836	.283	-.447	.328
i09	.174	-.759	.360	-.133	-.077
i35	.642	-.666	.416	-.516	.537
i32	.540	-.598	.179	-.361	.347
i01	.466	-.551	.248	-.217	.180
i30	.389	-.076	.802	-.443	.352
i36	.457	-.405	.753	-.364	.454
i08	.291	-.446	.750	-.177	.078
i38	.344	-.423	.717	-.236	.400
i28	.345	-.160	.707	-.549	.606
i37	.281	-.196	.666	-.534	.364
i29	.413	-.336	.560	-.377	.328
i34	.469	-.438	.360	-.768	.328
i33	.391	-.585	.320	-.720	.439
i15	.369	-.247	.357	-.719	.402
i14	.475	-.095	.526	-.716	.437
i31	.466	-.354	.397	-.707	.335
i05	.368	-.291	.352	-.424	.853
i04	.323	-.008	.291	-.604	.798
i03	.566	-.165	.265	-.009	.739
i13	.646	-.359	.295	-.426	.717
i06	.240	-.020	.431	-.465	.715
i16	.424	-.316	.499	-.595	.605

- Credible: This factor consists of five questions, relating to the value of the advice or instruction from the agent.
- Mentor-like: This factor consists of five questions, reflecting the agent's interaction with the learner.
- Human-like: This factor consists of five questions, addressing the naturalness of the agent's nonverbal communication with respect to personality and emotional expression.

- Engaging: This factor consists of five questions, relating to the level of motivation provided by the agent.

Reliability analysis was conducted to evaluate how the remaining question items were consistent within the subscales. The overall reliability of Cronbach's alpha was estimated at 0.96 for the I-Group. The internal consistency of items for each subscale showed high reliabilities: 0.94, 0.92, 0.87, 0.86, and 0.86 for *Facilitating Learning*, *Credible*, *Mentor-like*, *Human-like,* and *Engaging* respectively. This indicated that items within each factor of the five-factor model were very reliable.

Initial Model: Five-factor Model

In the second phase, these five factors were reviewed according to other theoretical and empirical evidence to build a theorized psychometric structure of pedagogical agent persona. We reviewed literature with respect to student perception of human instructors and identified two major instructor characteristics: 1) ability and 2) personality (Beishuizen *et al.*, 2001). Table 5 summarizes how these characteristics link to the five pedagogical agent factors of the initial model.

Instructor ability includes skills, knowledge, and experience, incorporating both theoretical knowledge (i.e., richness of knowing the content) and practical knowledge (i.e., handling complex and difficult instructional situations) (Kessels & Korthagen, 1996). Specific skills include presenting learning materials in a logical way with relevant concepts, while considering students' learning style. These teaching skills correspond to the factor, *Facilitating Learning*, given that these items relate to how the agent facilitates the process of the student's learning.

TABLE 5
Instructor Characteristics and Relationship to Agent Factors.

Instructor Characteristics		**Agent Factors**	**Construct**
Ability	Skills	Facilitating Learning	Informational Usefulness
	Knowledge (content & curricular)	Credible	
	Experience (pedagogical content knowledge)	Mentor-like	
Personality	Emotional expression	Human-like	Affective Interaction
	Communication	Engaging	

Instructor knowledge and experience (two other components of instructor ability) are comprised of three types of knowledge, (Shulman, 1986) : 1) Content knowledge, 2) Curricular knowledge, and 3) Pedagogical content knowledge. Content knowledge refers to the instructor's knowledge of the facts and structure of the content domain. Curricular knowledge refers to the in-depth knowledge of the domain which allows for providing alternative approaches to the content. Pedagogical content knowledge is an integrated knowledge of instructional strategies best related to the content, such as analogies, illustrations, and demonstrations. The content/curricular knowledge components correspond to the *Credible* factor as it relates to the instructor's knowledge and expertise in the content area. The experience component corresponds to the *Mentor-like* factor because the items reflect student and agent interaction (e.g., pedagogical content knowledge), also similar to instructor practical knowledge, as described by Kessels and Korthagen (1996).

Instructor personality is related to emotional expression and communication (Beishuizen *et al.*, 2001). Emotional expression corresponds to the *Human-like* factor, which refers to nonverbal communication such as facial expression, body-gesture, natural movement, and voice. Given that nonverbal communication always takes place in any social interaction (Guye-Vuilleme *et al.*, 1999; Prendinger & Ishizuka, 2001), it is a crucial factor for building a social interaction among learners and agents. Furthermore, how human-like the agent appears impacts whether the learner perceives the interaction as enjoyable and compelling. The *Human-like* factor is thus highly related to effective interaction and instructor personality.

Other research indicates positive instructor personality traits include characteristics such as "kind," "enthusiastic," "friendliness," and possibly "attractive" (Lowyck, 1994). These are closely related to affective/motivational perceptions (e.g., agent friendliness and enthusiasm) as assessed in the *Engaging* factor.

In conclusion, we defined two constructs to represent the relationships of the five factors for the initial model (shown in Figure 2). The first construct reflects the ability perspective of a good teacher consisting of the *Facilitating Learning*, *Credible*, and *Mentor-like* factors, labeled as "Informational Usefulness." The second construct reflects the affective features (e.g., personality, expression, communication) of the agent, including the *Engaging* and *Human-like* factors, and labeled as "Affective Interaction."

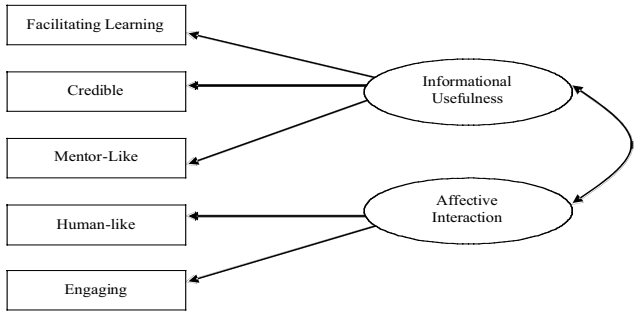

FIGURE 2
Path Diagram of the Initial Model.

Model Specification of the Initial Model (Five Factor Model)

Confirmatory factor analysis was conducted by using the maximum likelihood method to estimate theoretical validity of the initial five-factor model as applied to two different sets of agents (AR-Group, N=72; and AI-Group, N=188). Results indicated that the initial five-factor model was unacceptable for both samples (χ^2=29.148, p<.001; χ^2=28.890, p<.001).

The model fit indices of the AR-Group, GFI, CFI, NNFI, IFI, and RMSEA were 0.887, 0.875, 0.688, 0.879, and 0.256 respectively. The criterion for GFI, CFI, NNFI, and IFI was set at 0.95; however, none of the indices met the criterion. Also, RMSEA was greater than 0.05, and it was assessed as unacceptable. As a result, the model fit was concluded to be unacceptable.

For the AI-Group, GFI, CFI, NNFI, IFI, and RMSEA were 0.935, 0.963, 0.908, 0.964, and 0.196. Unlike the result from the AR-Group, NNFI and IFI were greater than 0.95, indicating that they were acceptable indices. Even so, the initial model for the AI-Group was evaluated as unacceptable when reviewing other indicators of goodness of model fit. Figure 3 and 4 shows the results of initial model for the AR-Group and AI-Group with maximum likelihood estimates of correlation coefficients. Labeling the latent variables was postponed until saturating the final model.

Model Respecification

Since the initial model failed to fit the data for the AR-Group and AI-Group, a model respecification was performed. For the model respecification, we followed a post-hoc modification process to re-specify

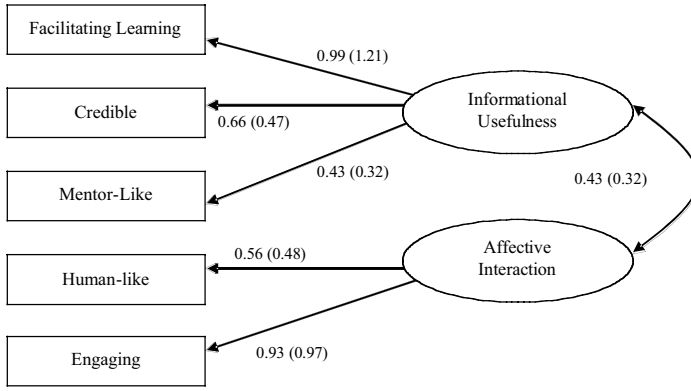

FIGURE 3
The Correlation Coefficient and R-square of the Initial Model for the AR-Group (R-square are in parenthesis).

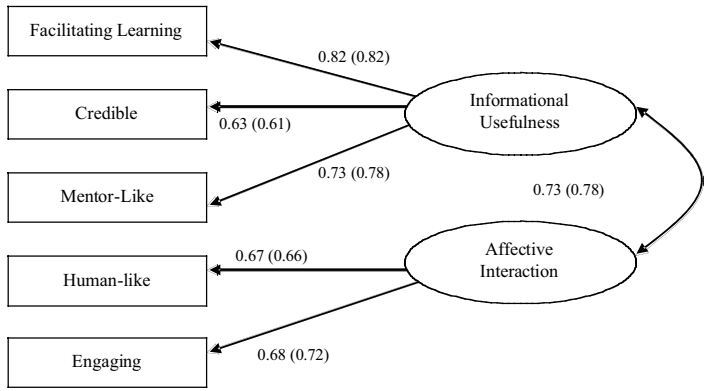

FIGURE 4
The Correlation Coefficient and R-square of the Initial model for the AI-Group (R-square in parenthesis).

the target model (Hoyle & Panter, 1995).

During the model respecification, two aspects were considered: a) the correlation coefficients of the constructs from the initial model; and, b) R-square of each construct from the initial model. As shown in Figure 3, the *Mentor-like* factor had the lowest correlation coefficient and R-square of the AR-Group, and, as illustrated in Figure 4, the *Credible* factor had the lowest correlation coefficient and R-square of the AI-Group.

Since there was inconsistency between the AR-Group and AI-Group, two alternative models were tested for both groups: 1) Four-factor model I, that eliminated the *Mentor-like* factor; and, 2) Four-factor model II, that eliminated *Credibility.* Each of these two alternative four-factor models were evaluated, as described next.

Four-factor Model I

The chi-square statistics for the AR-Group and AI-Group were estimated at 0.066 (p=0.80) and 3.373 (p=0.07) respectively. It was concluded that the respecification process improved the model-fit to the data. This revised model was evaluated as overall acceptable for both the AR-Group and AI-Group. Furthermore, the theoretical structure of the agent persona factors was assessed as robust because it successfully fit to both samples, which used different sets of agents.

Other indicators showed similar results. For the AR-Group, GFI, CFI, NNFI, IFI, and RMSEA were 1.000, 1.000, 1.038, 1.006, and 0.000 respectively. GFI, CFI, NNFI, and IFI were thus greater than .95 (the predetermined cutoff value) and RMSEA was acceptable (<.05) . For the AI-Group, GFI, CFI, NNFI, IFI, and RMSEA were 0.991, 0.995, 0.969, 0.995, and 0.112 respectively. GFI, CFI, NNFI, and IFI were greater than .95, but RMSEA was unacceptable (> .05). Although RMSEA of the AI-Group did not satisfy the criterion, the overall evaluation of Four-factor Model I indicated that it fit well to the data.

Four-factor Model II

The chi-square statistics for the AR-Group and AI-Group were 6.320 (p<.05) and 3.547 (p=.06) respectively. The chi-square indicated inconsistent results for the two groups. While the four-factor model II showed overall goodness of fit for the AI-Group, it did not fit well to the AR-Group. The other model fit indicators similarly showed inconsistent results. For the AR-Group, GFI, CFI, NNFI, IFI, and RMSEA were 0.959, 0.959, 0.755, 0.961, and 0.267 respectively. GFI, CFI, and IFI were greater than .95, but IFI and RMSEA were unacceptable. For the AI-Group, GFI, CFI, NNFI, IFI, and RMSEA were 0.991, 0.995, 0.971, 0.995, and 0.116 respectively, all of which met the criterion of .95, yet RMSEA was unacceptable. Table 6 shows the model fit indices for four-factor models I and II.

TABLE 6
Model Fit Indices of Four-factor Model I and II for Model Respecification

Alternative Models		Chi-square	GFI	CFI	NNFI	IFI	RMSEA
Four-factor Model I	AR-Group	0.066	1.000	1.000	1.038	1.006	0.000
	AI-Group	3.373	0.991	0.995	0.969	0.995	**0.112**
Four-factor Model II	AR-Group	**6.320***	0.959	0.959	**0.755**	0.961	**0.267**
	AI-Group	3.547	0.991	0.995	0.971	0.995	**0.116**

(Note: The bold and underscored numbers indicate unacceptable values for the model fit indices.)
(Four-factor Model I: Facilitating Learning, Credible, Human-like, and Engaging)
(Four-factor Model II: Facilitating Learning, Mentor-like, Human-like, and Engaging)
(* $p < .05$)
(Chi-square = Minimum fit function, GFI=Goodness of fit index, CFI= comparative fit index, NNFI=Bentler-Bonett nonnormed fit index, IFI= Incremental fit index, RMSEA= Root mean square error of approximation)

Final Model

Although the four-factor model II showed a good model fit to the AI-Group data, it did not fit well for the AR-Group data. We determined that the four-factor model I was a better fit to the data because it relatively successfully fit to both two samples with different sets of agents. We identified the four-factor model I as the final model to explain the psychometric structure of pedagogical persona effects, as illustrated in Figure 5.

The reliability of each observed variable ranges from 0.225 to 2.383 for the AR-Group and from 0.639 to 0.898 for the AI-Group. The correlations among the latent variables were estimated at 0.43 and 0.91 for the AR-Group and AI-Group respectively. The correlations were all positive with magnitudes consistent with expectations, based on the theory of instructor characteristics and relationship to agent factors. As shown in Table 8, the assessed Cronbach's alpha of items for each factor indicated that the items showed very reliable consistency within the factors. The final instrument, the Agent Persona Instrument (API), is included in Appendix A.

DISCUSSION

Results provide empirical support for the psychometric structure of pedagogical agent persona, together with a reliable and validated instrument for its assessment. The four factors that emerged include *Facilitating Learning*, *Credible*, *Human-like*, and *Engaging* with two latent variables —

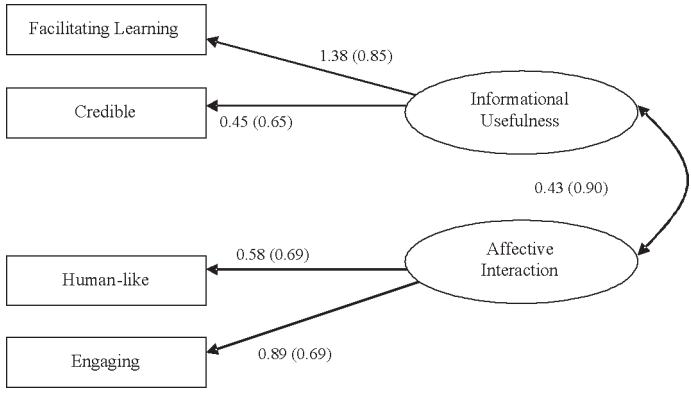

FIGURE 5
The Correlation Coefficient of the Final Model for the AR-Group and AI-Group (The numbers in parenthesis show the results of the AI-Group).

Informational Usefulness and Affective Interaction. As theorized in establishing the initial model of pedagogical agent persona, it was confirmed that the underlying factors of pedagogical agent persona are similar to those of human instructors, even though the actual items to assess them differ. Two characteristics of human instructors – ability and personality (e.g., Beishuzen *et al.* 2001) – correspond well with the pedagogical agent persona factors elicited here. Specifically, *Facilitating Learning* and *Credible* constituted the construct of Informational Usefulness, which corresponds with instructor ability. The factors of *Human-like* and *Engaging* constituted the construct of Affective Interaction, which corresponds with instructor personality.

Informational Usefulness: Agent as Knowledgeable Instructor

To understand how the factors in the Informational Usefulness construct interrelate, further investigation was conducted to discover why the AR-Group and AI-Groups differed with respect to fitting the factors to the model; specifically, the *Mentor-like* factor had the lowest correlation for the AR-Group whereas the *Credible* factor had the lowest correlation for the AI-Group. We have two possible explanations for these differences.

First, the differences could be attributed to the nature of the instrument items. Interestingly, most of the items used for *Mentor-like* were revised from the ATAS scale (Van Eck & Adcock, 2003), based on items for a *human* tutor.

These items focused on individualized feedback to students, which were more appropriate for a human teacher or to a more adaptive agent than the interface agents employed here. This could explain why the *Mentor-like* factor had the lowest correlation coefficient for the AR-Group. A second possible explanation involves the fact that two different sets of agents were employed. The AR-Group was exposed to three agents differing by role with each agent providing different messages, from direct information (Expert agent) to motivational support (Motivator agent) to a combination of both (Mentor agent). In contrast, the AI-Group was exposed to agents with *identical messages*, differing only by image. Thus, the AR-Group responses might be more sensitive than those of AI-Group because they were exposed to different messages from the agents. Consequently, it would be expected that credibility would play a larger factor for this group (and less for the AI-Group where the agents had identical messages) given that it pertains primarily to the agent's *message*. To test this explanation, these factors should be assessed again with new agents that have different messages.

Regardless, the factors of *Facilitating Learning* and *Credible* clearly relate to enhancing learning, and the pedagogical agent's message and/or content design is a critical component of its persona. The knowledge that the pedagogical agent provides also may lead the learner to interpret it as useful and helpful, even when it does not directly contribute to significantly better performance (e.g., Andre, Rist et al, 1999). Overall, a key feature of the pedagogical agent persona is that it is perceived as a knowledgeable instructor; thus, its message as it facilitates learning and its corresponding credibility are important to facilitate learners' thinking, reflection, and understanding.

Affective Interaction: The Social Presence of the Pedagogical Agent

Another key construct of the pedagogical agent persona psychometric structure is Affective Interaction, which corresponds to human instructor personality and communication. Among the factors of this construct, the *Human-like* factor is unique to the human-computer interaction features of an interface agent. The *Human-like* factor assesses how well the agent communicates naturally, including emotional expression and nonverbal communication. This lifelike quality of the agent has been shown to enhance its believability and produce a more natural conversational manner (Andre *et al.*, 1999a; Baylor *et al.*, 2005). It also supports results that show that anthropomorphic or lifelike agents lead to more enjoyable and engaging learning (Baylor & Ryu, 2003). Essentially, the *Human-like* factor of the

pedagogical agent persona is what makes it figuratively "real."

Once the pedagogical agent is perceived as virtually "real," the *Engaging* factor describes the positive social presence of the agent with the learner. The social presence of the agent can thus be characterized as its "being there" for the learners and motivating them. Consequently, as represented by the question items for the *Engaging* factor, the agent should be expressive, enthusiastic, entertaining, motivating, and friendly. The characteristics of the *Engaging* factor thus represent the social richness of the communication channels (Whitelock *et al.*, 2000) and play an important role to provide "personality" to the agent and enhance the learning experience. By being engaging, the agent personality can facilitate active participation in learning, boost students' esteem, and support the student-teacher social relationship (Makhanya, 2002).

Overall, to establish social presence with the learners, the pedagogical agent's personality should be designed in a manner that is easily interpretable. One way is through personalizing the message (e.g., Moreno *et al.*, 2001) or by including encouraging/motivating messages such as "You did a good job," "It is an excellent question," and "You can do this task." Even so, agent nonverbal communication is important to provide a rich sense of personality through emotional expression, deictic gestures, and voice (Andre *et al.*, 1999a). Thus, both the *Human-like* factor and *Engaging* factors shape the pedagogical agent's social presence and personality.

LIMITATIONS

There are two major limitations which could impact the potential generalizability of the model. First, only undergraduate learners participated, and it is possible that learners of different ages (particularly children) may perceive agent-based instructors differently. Second, the task was limited to instructional planning, which is ill-structured task and requires higher order thinking. Given that the pedagogical agents were not adaptive (e.g., intelligent) in these studies, participants may have been more skeptical of their value than if the agents had provided more guidance. Further, the applicability of the model to tasks in other domains, such as conceptual learning, is not clear; however, recent evidence suggests that the API is valid within the domains of attitudinal and procedural learning (Baylor, *et al.*, 2005).

CONCLUSION

The most promising finding of this study is the identification of two constructs — Informational Usefulness and Affective Interaction – as comprising the psychometric structure of pedagogical agent persona. Informational Usefulness reflects the pedagogical agent role as a knowledgeable instructor, and Affective Interaction reflects the agent's social presence and personality.

Informational Usefulness and Affective Interaction mutually affect the agent persona as Silverman *et al.* (2001) explained:

> *Emotive and animated pedagogical agents appear irrelevant when the situation and/or other characters make it clear what is required of the user. We call [it] "situational dependency" that animated pedagogical agents are helpful only where needed to boost companionship, increase entertainment value, and reduce complexity. Thus if the situation is already understandable and companionship and entertainment offer little added value or if there is a mismatch between the virtual agent and user's personalities, then there is will be no persona effect and one can safely omit the pedagogical agent from that situation (p. 227). In other words, if the information is not useful or the affective interaction is not necessary, the presence of a pedagogical agent may not be enough to yield significant outcomes.*

This situational dependency of the pedagogical agent persona may help explain why there are relatively few studies (Atkinson, 2002; Baylor & Kim, 2005; Moreno *et al.*, 2001) where an anthropomorphic pedagogical agent had significant positive effects on learning, whereas many studies have reported positive motivational outcomes (Andre *et al.*, 1999b; Baylor, 2005; Baylor *et al.*, 2003; Baylor & Ryu, 2003; Moundridou & Virvou, 2002; van Mulken *et al.*, 1998). If the pedagogical agent does not enhance learning or if the difficulties of learning the information were beyond the agent's capacity to help, the agent may be able instead to enhance the learners' attitudes toward the task. Thus, the two-dimensional structure of the pedagogical agent persona suggests that it is substantive to address both task difficulties as well as address learner characteristics.

REFERENCES

Andre, E., Klesen, M., Gebhard, P., Allen, S., & Rist, T. (1999a). *Integrating models of personality and emotions into lifelike characters.* Paper presented at the the Workshop on Affect in Interactions - Towards a New Generation of Interfaces, Siena, Italy.

Andre, E., Rist, T., & Muller, J. (1999b). Employing AI methods to control the behavior of animated interface agents. *Applied Artificial Intelligence, 13*(4/5), 415-448.

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94*, 416-427.

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology, 30*, 117-139.

Baylor, A. L. (2000). Beyond butlers: Intelligent agents as mentors. *Journal of Educational Computing Research, 22*(4), 373-382.

Baylor, A. L. (2001). *Investigating multiple pedagogical perspectives through MIMIC (Multiple Intelligent Mentors Instructing Collaboratively).* Proceedings of Artificial Intelligence in Education (AI-ED) International Conference, San Antonio, Texas.

Baylor, A. L. (2002). Expanding preservice teachers' metacognitive awareness of instructional planning through pedagogical agents. *Educational Technology Research & Development, 50*(2), 5-22.

Baylor, A. L. (2005). *The impact of pedagogical agent image on affective outcomes.* Proceedings of workshop "Interactions: The Computer in the Affective Loop" at the Intelligent User Interface International Conference, San Diego, CA.

Baylor, A. L., & Kim, Y. (2005). Simulating pedagogical agent instructional roles. *International Journal of Artificial Intelligence in Education.15(1).*

Baylor, A. L., Kim, S., Son, C., & Lee, M. (2005). *Designing effective nonverbal communication for pedagogical agents.* Proceedings of Artificial Intelligence in Education (AI-ED), Amsterdam.

Baylor, A. L., & Ryu, J. (2003). Does the presence of image and animation enhance pedagogical agent persona? *Journal of Educational Computing Research, 28*(4), 373-395.

Baylor, A. L., Shen, E., & Huang, X. (2003). *Which pedagogical agent do learners choose? The effects of gender and ethnicity.* Proceedings of E-Learn (World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education), Phoenix, Arizona.

Beishuizen, J. J., Hof, E., Putten, C. M. v., Bouwmeester, S., & Asscher, J. J. (2001). Students' and teachers' cognitions about good teachers. *British Journal of Educational Psychology, 71*, 185-201.

Craig, S. D., & Gholson, B. (2002). *Does an agent matter? The effects of animated pedagogical agents on multimedia environments.* Proceedings of ED-MEDIA 2002: World Conference on Educational Multimedia, Hypermedia and Telecommunications., Norfolk, VA.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94*(2), 428-434.

Eck, R. V., & Adcock, A. B. (2003). *Reliability and factor structure of the Attitude Toward Agent Scale (ATAS).* Paper presented at the American Educational Research Association (AERA), Chicago, IL.

Elliott, C., Rickel, J., & Lester, J. (1999). Lifelike pedagogical agents and affective computing: An exploratory synthesis. In M. W. M. Veloso (Ed.), *Lecture notes in artificial intelligence* (Vol. 1600, pp. 195-212). Springer-Verlag, Berlin.

Ertmer, P., & Russell, J. (1995). Using case studies to enhance instructional design education. *Educational Technology*, 23-31.

Guye-Vuilleme, A., Capin, T. K., Pandzic, I. S., Thalmann, N. M., & Thalmann, D. (1999). Nonverbal communication interface for collaborative virtual environments. *Virtual Reality, 4*, 49-59.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis*. Upper Saddle River, New Jersey: Prentice Hall.

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications*. Thousand Oak, California: SAGE.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76-99). Thousand Oak, California: SAGE.

Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education, 11*, 47-78.

Joreskog, K. G., & Sorbom, D. (1993). *Lisrel 8: Structural equation modeling with the simplis command language*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Kessels, P. P. A. M., & Korthagen, F. A. L. (1996). The relationship between theory and practice: Back to the classic. *Educational Researcher, 25*(3), 17-22.

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhoga, R. S. (1997). *The persona effect: Affective impact of animated pedagogical agents.* Paper presented at the CHI'97 Human Factors in Computing Systems.

Lester, J. C., Voerman, J. L., Towns, S. G., & Callaway, C. B. (1999). Deictic believability: Coordinated gesture, locomotion, and speech in lifelike pedagogical agents. *User Modeling and User-Adapted Interaction, 9*(1-2), 1-44.

Makhanya, M. S. (2002). What do teachers do? A qualitative analysis of the role of the teacher. *Systemic Practice and Action Research, 15*(2), 123-144.

Mayer, R. E., Sobko, K., Fennell, S., & Mautone, T. (2002). *Social cues in multimedia learning: Role of speaker's voice and conversational style.* Paper presented at AERA annual conference.

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94*(1), 156-163.

Moreno, R., Mayer, R. E., & Lester, J. C. (2000). *Life-like pedagogical agents in constructivist multimedia environments: Cognitive consequences of their interaction.* Proceedings of ED-MEDIA 2000, Montreal, Canada.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19*(2), 177-213.

Moundridou, M., & Virvou, M. (2002). Evaluating the persona effect of an interface agent in a tutoring system. *Journal of Computer Assisted Learning, 18*, 253-261.

Newsom, J. T. (2001). Some clarifications and recommendations on fit indices. Retrieved May 5th, 2003, from http://www.upa.pdx.edu/IOA/newsom/semclass/ho_fit2_03.doc

Prendinger, H., & Ishizuka, M. (2001). *Simulating affective communication with animated agents.* Proceedings of International Conference on Autonomous Agents, Montreal, Canada.

Shulman, L. S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 3-36). Newbury Park, CA: Sage Publications.

Silverman, B. G., Holmes, J., Kimmel, S., Branas, C., Ivins, D., Weaver, R., *et al.* (2001). Modeling emotion and behavior in animated persona to facilitate human behavior change: The case of the heart-sense game. *Health Care Management, 4*, 213-228.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*, 173-180.

Tate, R. (1998). *An introduction to modeling outcomes in the behavioral and social sciences*. Edina, Minnesota: Burgess International Group.

Towns, S. G., FitzGerald, P. J., & Lester, J. C. (1998). *Visual emotive communication in lifelike pedagogical agents.* Proceedings of International Conference on Intelligent Tutoring Systems, San Antonio, TX.

van Mulken, S., Andre, E., & Muller, J. (1998). *The persona effect: How substantial is it?* Proceedings of HCI-98, Berlin.

Whitelock, D., Romano, D., Jelfs, A., & Brna, P. (2000). Perfect presence: What does this mean for the design of virtual learning environments? *Education and Information Technologies, 5*(4), 277-289.

APPENDIX A

Agent Persona Instrument (API) : Final Instrument

All items should be presented with a 5-point Likert scale, ranging from 1=Strongly disagree to 5=Strongly agree.

Facilitating Learning (10 items)

The agent led me to think more deeply about the presentation.
The agent made the instruction interesting.
The agent encouraged me to reflect what I was learning.
The agent kept my attention.
The agent presented the material effectively.
The agent helped me to concentrate on the presentation.
The agent focused me on the relevant information.
The agent improved my knowledge of the content.
The agent was interesting.
The agent was enjoyable.

Credible (5 items)

The agent was knowledgeable.
The agent was intelligent.
The agent was useful.
The agent was helpful.
The agent was instructor-like.

Human-like (5 items)

The agent has a personality.
The agent's emotion was natural.
The agent was human-like.
The agent's movement was natural.
The agent showed emotion.

Engaging (5 Items)

The agent was expressive.
The agent was enthusiastic.
The agent was entertaining.
The agent was motivating.
The agent was friendly.