



This is to certify that the
thesis entitled


**AN ELECTRONIC QUIT SMOKING COACH:
DESIGNING COMPUTER AGENTS AS PERSUASIVE
SOCIAL ACTORS**

presented by

GERAUD LUIS PLANTEGENEST

has been accepted towards fulfillment
of the requirements for the

M.A. degree in **TELECOMMUNICATION,
INFORMATION STUDIES,
AND MEDIA**



Major Professor's Signature

Date

3/1/05

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
MAR 25 2013		

**AN ELECTRONIC QUIT SMOKING COACH:
DESIGNING COMPUTER AGENTS AS PERSUASIVE SOCIAL ACTORS**

By

Geraud Luis Plantegenest

A THESIS

Submitted to
Michigan State University
In partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Telecommunication, Information Studies and Media

2005

ABSTRACT

AN ELECTRONIC QUIT SMOKING COACH: DESIGNING COMPUTER AGENTS AS PERSUASIVE SOCIAL ACTORS

By

Geraud Luis Plantegenest

An online electronic quit smoking coach takes the place of structured phone-based human coaches. Using the same interview protocols as in person and telephone human interventions, online agents can potentially promote health and make services which persuade and provide motivation available to anyone, anytime and anywhere.

The thesis study reports on development and the first test of this new system, implemented using coaches with two different personalities, friendly and stern. The project explores the potential of the Internet as the medium to deliver the simulated smoking cessation experience and looks at reactions on the part of a small number of users to different aspects of the electronic agent. It is a prelude to a large scale study to test the effectiveness of electronic online coaches in helping people quit smoking through a new web-based approach.

Copyright by
GERAUD LUIS PLANTEGENEST
2005

This thesis is dedicated to my dear parents and grandparents. Thank you for all your support during my education and the always present encouragement to better myself as a professional and as a person.

This thesis is also dedicated to my wife, Tania, and my kids, Shenika and Ederick, thank you for all your unconditional love, support and motivation.

ACKNOWLEDGEMENTS

Thanks to Dr. Carrie Heeter for all her support and guidance from the conception of this thesis to its completion. Her patience and ability to communicate ideas, and her extensive knowledge and experience in designing for the user were an essential part in making this thesis project more interesting and one of higher quality.

Thanks to Dr. James Anthony for being the keystone that helped make this thesis a reality. His broad vision, ideas and feedback made this project fun and challenging. Thanks to Professor Brian Winn for his invaluable time, programming skills and advice on key technical challenges the project posed.

Thanks to Dr. Wadland and Dr. Holtrop for adapting and revising key parts of the quit smoking protocol to be in tuned with the target audience. Thanks to all the Epidemiology post-doctoral fellows with whom I spent many hours interchanging ideas about the project. I have great appreciation and value for your perspectives. Thanks to Priyanka Rao and Mashell Lawson, two interns who were very supportive and resourceful during the initial stage and development of this project.

Thanks to Dr. Frank Biocca, Zena Biocca and all the amazing and talented people in the M.I.N.D lab. You gave me a place to call my own, the opportunity to

explore and key guidance to further develop my academic potential. Thanks to Professor Darcy Greene for her input, and ideas. Thanks to Jason Justman for his programming skills and for the time invested to make sure that things were ready when they needed to be. Also, thanks to the staff of the Communication Technology Lab for their help and prompt response to my calls for help when technology would not collaborated.

Finally, thanks to my wife, Tania and my kids, Shenika and Ederick for their endless support, love, motivation and laughter. I love you with all my heart.

TABLE OF CONTENTS

LIST OF TABLES.....	ix
LIST OF FIGURES.....	x
 CHAPTER 1	
A REVIEW OF THE LITERATURE ON IMPACTS OF AGENT REPRESENTATIONS.....	1
1.1 OVERVIEW.....	1
1.2 EFFECTIVENESS OF COMPUTER AGENTS AS LEARNING TECHNOLOGIES.....	1
1.3 FACILITATING LEARNING THRU MULTIMEDIA.....	5
1.4 EFFECTS OF IMAGE MOTION AND STATIC IMAGES.....	9
1.5 SOCIAL EFFECTS OF IMAGE AND VOICE IMPACT OF PEDAGOGICAL AGENTS IN LEARNING AND MOTIVATION.....	17
1.6 AGENTS AS SOCIAL PERSUADERS.....	22
 CHAPTER 2	
AGENT TECHNOLOGIES.....	31
2.1 OVERVIEW.....	31
2.2 TECHNOLOGIES.....	32
2.2.1 Oddcast Sitepal.....	32
2.2.2 Artificial Life (ALife).....	34
2.2.3 Extempo.....	37
2.2.4 Pulse 3D.....	39
2.2.5 Microsoft Agent.....	40
2.2.6 Famous3D.....	41
2.3 STRIKING THE IRON WITH THE RIGHT HAMMER.....	43
 CHAPTER 3	
PRACTICE- AND COMMUNITY -BASED SMOKING CESSATION COUNSELING PROGRAMS: "I'D RATHER COPE THAN SMOKE".....	47
 CHAPTER 4	
APPLYING CAPTOLOGY TO SMOKING CESSATION: AN ADAPTABLE APPROACH.....	51
4.1 OVERVIEW.....	51
4.2 A FUNCTIONAL SHIFT ON PERSUASION.....	51
4.3 HOW COMPUTERS PERSUADE.....	54
4.4 FUTURE TRENDS.....	60

CHAPTER 5	
PRODUCTION PROCESS.....	64
5.1 INFORMATION DESIGN.....	64
5.1.1 SCRIPT ADAPTATION.....	64
5.2 AGENT DESIGN.....	67
5.2.1 COMPUTER AGENTS.....	67
5.3 SURVEY DESIGN.....	69
5.4 USER INTERFACE DESIGN.....	73
5.4.1 INTERFACING THE USER AND THE COACH.....	73
5.4.2 NON-VISIBLE FUNCTIONS.....	74
5.5 DELIVERY OF LEARN TO QUIT.....	76
 CHAPTER 6	
RESEARCH METHODS AND FINDINGS.....	78
6.1 METHODS.....	78
6.1.1 AGENT (COACH) DEVELOPMENT.....	78
6.1.2 SAMPLE SIZE, SUBJECTS AND RECRUITMENT.....	79
6.1.3 INSTRUMENT.....	80
6.1.4 SURVEY ADMINISTRATION.....	82
6.2 FINDINGS.....	84
6.2.1 GROUP PARTICIPATION AND RETENTION.....	84
6.2.2 COACH EVALUATION.....	86
6.2.3 DATA INTERPRETATION.....	96
 CHAPTER 7	
CONCLUSION.....	103
 APPENDICES:	
A. DIRECTOR SERVER COMMUNICATION SCRIPT.....	108
B. RECRUITMENT POSTER.....	110
C. ELECTRONIC QUIT SMOKING COACH CONSENT FORM.....	111
D. SURVEY QUESTIONS- COACH COMPONENT.....	114
E. LEARN TO QUIT OFFICIAL LETTER OF INVITATION.....	117
 BIBLIOGRAPHY.....	120

LIST OF TABLES

Table 1	
Overall Subject Participation between Sessions and Subject Retention per Session.....	84
Table 2	
Combined Group Demographics for Stern and Friendly Coach.....	86
Table 3	
Overall Means and Comparisons of Means between Stern and Friendly Coaches.....	88

LIST OF FIGURES

FIGURE 1: Sitepal Online Interface for Designing Agents.....	33
FIGURE 2: Alife Online Demo of Luci, a Website Tour Guide.....	35
FIGURE 3: A Sample of the Extempo Expert Agents.....	38
FIGURE 4: A Pulse 3D Veeper Agent.....	39
FIGURE 5: Peedy, a Microsoft Agent.....	40
FIGURE 6: Famous3D Talking Head Models.....	41
FIGURE 7: Shockwave Player Adoption Statistics.....	46
FIGURE 8: Cigarettes' Ads with Hip-Hop DJ's and Dancers that Critics Say Targets Black Youth.....	52
FIGURE 9: A Virtual Bar Simulation Shows the Impact of Alcohol in the Human Body.....	53
FIGURE 10: Avatar Representation of Male and Female Characters Created by Users of the Adobe Atmosphere Software.....	57
FIGURE 11: Electronics Gadgets.....	59
FIGURE 12: Information Structure of Learn to Quit Application.....	65
FIGURE 13: Computer Agents.....	68
FIGURE 14: Agent's Poses and Expressions.....	70
FIGURE 15: The Learn to Quit Web Page Dissected.....	72
FIGURE 16: Learn to Quit User Interface.....	73
FIGURE 17: Shockwave Movie, Backend System, and TTS Engine Integration.....	76

CHAPTER 1

A REVIEW OF THE LITERATURE ON IMPACTS OF AGENT REPRESENTATIONS

1.1 OVERVIEW

In this chapter I present a literature review that seeks to understand what empirical research tells us about the design and implementation of pedagogical and other helpful computer agents in web based interactive environments. Specifically, I wanted to look at what previous studies have done with anthropomorphic agents to better inform the design of the web-based quit smoking electronic coach.

1.2 EFFECTIVENESS OF COMPUTER AGENTS AS LEARNING TECHNOLOGIES

In their book *The Media Equation*, Reeves and Nass (1996) explain that our brain has not evolved yet enough to catch up with communication media. As a result of the newness of computers and the slow speed of evolution, we will tend to extend our reactions to other humans onto media and orient ourselves socially to computers and software, and we do so even more when computers attempt to personify human characteristics (David, Lu, Kline, and Cai, 2002). Numerous research studies support this theory, known by the same name as the book (e.g. Reeves and Nass, 1996).

Anthropomorphic pedagogical computer agents are still in their early developmental stages, yet, current research seems to indicate that they have great potential to be used as new instructional technologies. Two agents developed by the Center for Advanced Research in Technology for Education (CARTE), Steve and Adele, are examples of successful pedagogical agents (Johnson, Rickel, & Lester, 2000). Steve was designed to instruct students on naval training tasks inside an immersive virtual environment, while Adele uses a more conventional interface to provide web-based medical education materials to students using a web browser. Johnson et al. (2000) hypothesize that a positive effect of creating agents as anthropomorphic entities is that they can be perceived as more likable and engaging, especially when they are made lifelike. Johnson et al. caution on the use of unnatural behaviors that can be expressed by lifelike agents interacting with the users, because these behaviors will more likely have a negative impact in the overall experience of the user by calling attention to themselves and resulting in repetitive distractions to the user.

Computer agents can be created with rich personalities to motivate, express concern and show their enthusiasm to users. It is understood that motivation is a key ingredient in learning (Johnson et al., 2000). Some of the key capabilities that agents can possess are verbal and non-verbal feedback. Johnson et al. suggest that by using verbal and non-verbal feedback, agents are able to convey emotions and by doing so; they can create a more compelling and enjoyable experience. This enjoyment can in turn benefit the learning experience

and engage the learner to interact more with the agents and for prolonged periods of time.

A formal empirical study with 100 middle school students was conducted by Lester and colleagues (1997) to evaluate the effects of animated pedagogical agents on students' problem solving using a learning environment called Design-A-Plant and an animated pedagogical agent by the name of Herman the Bug. The environment is designed to support botanical anatomy and physiology characteristics needed for the sustained survival of plants. The agent Herman the Bug was implemented in the study with different versions and each version provided a different level of interaction to advise students as they designed plants that could thrive under different environmental conditions. Findings from the study indicated that the presence of animated pedagogical agent produced better problem-solving performance in the learning environment. Lester et al. also found that agents with multiple modalities and multiple levels of advice were perceived as more expressive and believed to be clearer when communicating.

Despite the growing interest shown by researchers for formulating new applications and envisioning new ways to apply agents as pedagogical vehicles for learning, there exists some controversy about the anthropomorphization of computer interfaces. In a discussion about Direct Manipulation vs. Interface Agents (Shneiderman & Maes, 1997), Shneiderman, a supporter of Direct Manipulation, argues that humanizing the interface takes away sense of

accomplishment a user may have if a computer agent is doing the work for them. In addition, he points out that predictability and user control are central items in the design of Direct Manipulation interfaces, both of which he sees as being greatly reduced when computer agents are introduced in an application.

In contrast, Maes who is an advocate for computer agents ascertains that the challenge lies in creating a solid 'user-agent interface' when introducing agents in an application. She explains:

...we need to take care of these two issues: Understanding and control.

Understanding means that the agent-user collaboration can only be successful if the user can understand and trust the agent, and control means that users must be able to turn over controls of tasks to agents but users must never feel out of control (p. 54)

In a review of empirical research on the impact of animated interface agents (Dehn & Mulken, 2000) found that empirical investigations of the effects of animated agents are small in numbers and are hard to compare because they differ with regard to the measured effects. However, they sustain that differing views on the use of computer agents for human-computer interaction, are potentially testable, and suggest that future studies should "(1) overcome the methodological shortcomings of existing studies and (2) take a more fine-grained

perspective on the effect of employing animated agents on the user's motivation and cognition" (Dehn & Mulken, 2000, p. 17).

1.3 FACILITATING LEARNING THRU MULTIMEDIA

Mayer (2001) studied how presentational formats of media affected the retention of knowledge and how it facilitated the transfer of knowledge. In particular, Mayer studies looked at whether two representational formats were better than one and whether it was better to present words and picture rather than words alone.

Mayer discussed two theories that have opposing views about the how the presentation of information should take place (1) The information delivery theory and (2) The cognitive theory of multimedia learning. To illustrate the theories Mayer gives the example of how a bicycle pump works by presenting the information in both formats; text and a drawing. The text describes how the pump works and the pictorial representation presented in two line drawings is assumed to be the equivalent presentation of what is described in the text. The fundamental question is whether the words described in the text convey the same basic information as the picture. Are the two modes of informational presentation equivalent? (Mayer 2001)

According to the information delivery theory, they are, because words and pictures are seen only as two different vehicles that present the same

information. On the other hand, the cognitive theory of multimedia argues that they are not equivalent, because these two different representations activate "...different knowledge representation systems in learners – a verbal channel and a visual channel" (p. 65).

A series of nine separate tests were performed to compare retention and transfer performance of information presented in the previously mentioned formats. Participants of the studies were divided into two groups. One group was called the multiple-representation group and received words and pictures, or a narrated animation. The other group was called the single-representation group and received text alone, or narration alone. The content of the presentations for the experiments was based on scientific explanations that varied from how pump or brakes worked to how lightning storms developed.

In six of the nine tests conducted it was found that the multiple-representation group was able to recall more of the presentations than did the single representation group. The implication for retention of information using multimedia is pointed out by Mayer as follows: "... students perform better on verbal retention when learning with text and illustrations or narration and animation than when they learn with text alone or narration alone" (p.73).

Mayer pointed out that the cognitive theory of multimedia is upheld but not without some contradictions. The effect size found was moderate. Recalling the

information in the presentation produced only about 23% more of the steps from participants in the multiple-representation group than the participants in the single representation. Another important finding is that adding animation did not help retention, but in general adding illustrations to text resulted in better retention.

The multimedia effect for transfer was the main focus of the research. Mayer (2001) explains, "...transfer performance is a reflection of how well students understand an instructional message" (p. 75). Results from the studies indicated that the multiple-representation group performed better than the single-representation group on the transfer test. This translated into effect sizes that were large and consistent with the multiple-representation group generating a median of 89% more creative solutions than did the single-representation group. In general, the results obtained were more consistent with the cognitive theory of multimedia and very much the opposite with the information-delivery theory, which suggested that there would be no differences between the two groups.

The practical implications of the study show that the multimedia presentation format is effective for learning. "...at least in the case of scientific explanations...." (p.79). The implications for multimedia design, is that the multimedia effect facilitates the creation of more effective messages in which both words and pictures coexist to enhance the students learning. However, Mayer recognizes that the multimedia design principle which states that we

should "...present words and pictures rather than words alone" (p.79) is a bit too ambiguous and therefore needs some clarification. He asks the following questions to address this ambiguity, "What kinds of pictures should be added, how should they be added, and when should they be added?" (p.80)

In order to make the multimedia principle more robust; Mayer went on to define and test empirically a set of principles that he suggested should be used in conjunction with the multimedia principle to inform the design of messages and thus create more effective messages. These are some of the most relevant principles:

Spatial Contiguity Principle: Students learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen.

Temporal Contiguity Principle: Students learn better when corresponding words and pictures are presented simultaneously rather than successively.

Coherence Principle: Students learn better when extraneous material is excluded rather than included. The coherence principle can be broken into three complimentary versions: (1) student learning is hurt when interesting but irrelevant words and pictures are added to a multimedia presentation; (2) student learning is hurt when interesting but irrelevant sounds and music are added to a

multimedia presentation; and (3) student learning is improved when unneeded words are eliminated from a multimedia presentation.

Modality Principle: Students learn better from animation and narration than from animation and on-screen text; that is, students learn better when words in a multimedia message are presented as spoken text rather than printed text.

Redundancy Principle: Students learn better from animation and narration than from animation, narration, and text. The theoretical rationale for this claim is that when pictures and words are both presented visually (i.e. as animation, and text), the visual channel can become overloaded.

The empirical results from the studies in which these five principles were tested showed consistently that the effect for transfer and retention of information in multimedia messages was improved when these conditions were present.
(Mayer, 2001)

1.4 EFFECTS OF IMAGE MOTION AND STATIC IMAGES

Ravaja (2004) conducted a study to find out the effects of a moving-face versus a static-face newscaster that delivered financial news to a mobile device (pocket PC) with a small screen. Ravaja examined the modulating influence that the moving vs. the static face had on emotion, memory performance and attention-related subjective and physiological responses. Subjective data was

collected by self-report of participants and physiological data was recorded using electrocardiogram (ECG), electrodermal activity (EDA) and facial electromyography (EMG). Citing previous work done by P.J Lang and coworkers, Ravaja (2004) explains how measuring physiological data helps to interpret the arousal (excited or calm) and valence (pleasant or unpleasant) dimensions which are the coordinates of a dimensional theory of emotion that believes that all emotions are located in a two-dimensional space, namely, the coordinates of valence and arousal:

...Using a paradigm in which stimuli are presented for 6 s (with an interstimulus interval of about 20 s), P.J Lang and co-workers have shown that both affective (still) pictures and sounds elicit patterns of phasic physiological response that vary systematically as a function of the normative valence and arousal of the stimuli (p.109).

Thirty-six young adults were recruited for this study. Participants wore attached electrodes and were given a Fujitsu Stylistic Pen Tablet to hold on their lap. The tablet was used to display a simulated Pocket PC appearing with its natural size in the center of the display while the rest of the display was kept in black. The experiment had two phases, during the first phase participants were asked to hear and view all the news bulletins in a certain order and to keep her eyes focused on the image of the newscaster while each clip was being played.

After viewing all the clips, the electrodes were removed and subjects were given a questionnaire to measure the amount of knowledge measure from the messages. The second phase involved having the participant viewing and hearing all the news bulletin for a second time with the difference that after each message the participant was asked to rate the message on several dimensions as well as the emotional response to the message the subject felt. Experiment results showed that as Ravaja hypothesized, moving face images elicited greater self-reported arousal than static-face images. Also, self-reported pleasure was found to be higher for moving-face messages than their static-face counterparts. However, physiological data, specifically the EMG data was ambiguous with regards to the higher pleasure reported by participants. Analysis of the EMG data suggested that in some respect, static-face bulletins elicited a more pleasant emotional response as compared to moving-face bulletins.

The reasons and implications for the ambiguity and dissociation of these two measured emotional responses remained unknown to the researchers. Further, Ravaja (2004) stated that "In regard to self-report, moving face messages tended to be rated as more interesting than static-face messages, but the difference narrowly failed to reach statistical significance" (p. 127-128).

Practical implications of the study findings suggests that when designing applications for mobile devices that use small screen or have limited bandwidth capabilities, designers need to consider the real objective of their media

message and how to best make an impact on the end user and whether that impact is enhanced by presenting a moving- face image while sacrificing the limited bandwidth.

Earlier studies by Detenber and Reeves (1996) investigated the effects of image motion and static images and found that still images generated more arousal than images with motion. The researchers were intrigued by the findings because they actually expected to see some effect on arousal with regard to motion. Detenber and Reeves explain the contradiction:

...The result may seem counterintuitive until one considers more carefully the nature of the manipulation. The still versions of the images were *stilled* moving pictures. Although the stimuli were selected to avoid a drastic or dynamic motion, there was motion in each on the initial 6-second clips. So, when the still versions were created by freezing one frame, there was often a sense of implied motion (p.79).

They suggested that these particular still images invited the user to interpret related motion to the image in a way that helped them to fill in the missing information. The users' search to connect the dots of what motion was there in the image before it was frozen and what motion comes after the current frame may have resulted in the activation of a stronger emotional response. In contrast, when a moving image is being presented the information is more

complete and demands greater cognitive resources which may result in the viewer having limited access to these resources to be able to speculate (Detenber & Reeves, 1996).

In a study that examined the effects that a synthesized talking face would have on subjects' responses in the context of a questionnaire study; Walker, Sproull, and Subramani (1994) compared subjects who answered questions presented via text display on the screen vs. subjects who answered the same questions spoken by a talking face.

Walker et al. used a between-subject design and randomly assigned a group of 49 people to one of three conditions:

- (1) Question spoken by a face with a neutral expression.
- (2) Question spoken by a face with a stern expression.
- (3) Text only.

It was found that subjects who interacted with the talking face format, spent more time in answering the questions, made fewer mistakes and also wrote more comments. In addition, talking faces were found to be more engaging, this was measure by the time people spent answering questions and in how much they wrote in response to open-ended questions.

Differences between expressions of the talking faces were also found; the sterner face resulted in more engagement than the neutral face. Participants that interacted with the sterner face made fewer mistakes, spent more time and wrote more comments. However, Walker et al. (1994) pointed out that "...engagement does not mean liking" (p. 89). In fact, subjects assessed the stern face less positively than the neutral face. It was estimated that the contradiction of paying more attention and spending more time with a face the user did not like rather than with the more neutral face may be founded on what social psychologists have found about the presence of another person Walker et al. explains:

...the presence of another person usually serves to increase arousal and motivation on the part of someone asked to perform a task [35]. This can lead to improved performance if the task is not very complex or to degraded performance if the task is complex. The presence of another person apparently produces evaluation reminders and therefore leads people to try harder (p. 89)

With this reasoning in mind, the researchers assumed that in this particular study, the sterner face, which was characterized by having more expressions, had in fact produced the most evaluation reminders than the other two conditions.

It seems important that the agent design and the presentation of contextually appropriate facial expressions and gestures are key factors to achieve the communication objective the agent is intended for. Johnson et al. (2000) points out that “engaging, lifelike pedagogical agents that are visually expressive could clearly communicate problem-solving advice and simultaneously have a strong motivation effect on learners” (p 71).

Context appropriate facial expressions can exploit the potential cognitive resources of the visual channel; thus, making the agent more “present”, and visually appealing to the user while retaining variability and surprise factor over time. For example, agents can use body language and gestures to express happiness, joy, excitement, empathy, sadness, fear, etc. This is especially important when users need to interact with agents over extended periods of time, because the more complex the agent behavior is, the more difficult it is to be quickly induced. As a result, this helps the agent to retain believability over time while providing the users with advice and encouragement by drawing from what Johnson et al. (2000) refers to as “...a rich repertoire of emotive behaviors....” (p 71); as the interaction between the learner and agent is taking place.

The relationship between agent persona characteristics and the effects it has on learner perceptions has also been studied by testing the role of static image and animation. Baylor and Ryu (2003) asked 75 pre-service teachers enrolled in an “Introduction to Educational Technology” course to developed an

instructional plan for a case study within an agent based environment called MIMIC (Multiple Intelligent Mentors Instructing Collaboratively). The MIMIC environment was designed to organize the participant's instructional planning processes into four main phases, namely, case study, blueprints, planning and assessment. The role of the agent within MIMIC was to serve as a mentor. The study consisted of two contrasting comparisons: (1) comparing presence/absence of agent image, and (2) comparing static vs. animated agent images. Baylor and Ryu (2003) described four persona characteristics the agent should possess:

- (1) Engaging
- (2) Person-like
- (3) Credible and
- (4) Instructor-like

The findings indicated that in the overall evaluation of the four persona characteristics, animation was found to be beneficial, but not always as the single best implementation. Also, the "presence" of an image (static or animated) was found to be "critical" to facilitate the agent credibility. Perception of the agent as instructor-like was found to be affected strongly when animation was present. Similarly, perceptions of the agent as engaging and person-like were also improved by animation, however person-like was just as improved by no image (Baylor & Ryu, 2003). Some practical guidelines that came out of this study seem

important to take into account when deciding the most viable and appropriate implementation of the technology:

...Overall, in all aspects of the pedagogical agent persona, animation is listed as one of (if not the only) most effective technological implementation. Consequently, if cost of implementation is not an issue, animation could lead to optimal perceptions for these four pedagogical agent persona features. For an agent-based pedagogical system where credibility is the most important issue, either static or animated image will lead to equally good results (p. 24)

In terms of implementing these proposed guidelines into practical use when designing an agent-based environment; one should pay particular attention to the desired learning outcome of the instruction and choose the approach that best addresses the needs of the interaction.

1.5 SOCIAL EFFECTS OF IMAGE AND VOICE IMPACT OF PEDAGOGICAL AGENTS IN LEARNING AND MOTIVATION

In one of the most widely cited meta-analytic reviews of research on the physical attractiveness stereotype, Eagly, Ashmore, Makhijani, and Longo (1991) looked at a number of studies that focused on the physical attractiveness stereotype to assess how "...strong and robust..." (p.119) the beauty-is-good stereotype really was. The beauty-is-good stereotype or also known as the

physical attractiveness stereotype suggests that "...attractive people lead perceivers to make strong inferences of personality goodness.... (p. 109). Eagly et al. pointed out that the general consensus among psychologists has been that the stereotype is very robust.

However, the review they performed demonstrated that the beauty-is-good stereotype was not as strong or generalizable as it is often suggested. The findings of their review indicated that the average magnitude of beauty-is-good effect was moderate and its strength varied among all studies reviewed. It is important to explain that the criteria Eagly et al. used for the selection of the articles was based only on studies that examined the physical attractiveness stereotyping in environments that were role-free and in which participants were able to evaluate the attractiveness or un-attractiveness of the target people without there being any sort of relation linked by social roles (e.g., coach-athlete, teacher-student).

A different study conducted by Kim, Baylor and Reed (2003) on 109 undergraduates examined the effects of image and voice of computer agents on student perceptions and learning. Contrary to the example of the meta-review by Eagly et al. (1991), this study established social roles between the subjects (perceivers) and the agent (target model). The agents were developed using different images to created simulated instructors with the roles of expert and mentor. The expert image wore a suit and looked stern while the mentor agent

wore casual clothing and look comfortable. Two types of voices were used, a strong voice that was authoritative, assertive and enthusiastic and a calm voice that was soft, nice and kind. The voices were delivered in two formats, (1) computer-generated and (2) human voice.

The findings of the study revealed that the agent image has a significant affect on the participants' perception towards the agent role. As Kim et al. (2003) explain: "...The perceived role of the mentor-like image was as a motivator, whereas the perceived role of the expert-like image was as an expert. Similarly, students assessed the mentor-like image as significantly more motivating than the expert-like image" (p. 4). With respect to the voices the following results were revealed: "...student perceived both human voices (strong and calm) as significantly more affective, affable, credible, and facilitating their learning than the computer generated voice" (p. 4). On the other hand, the agents with strong voices (computer-generated and human) were perceived to be significantly more motivational.

Furthermore, the results suggests that the type of selection designers make for images and voices to be used in agent-based applications could be taken as "...important predictors" (Kim et al., 2003, p. 4) on how motivated the students might be when interacting with the application in the learning environment.

More recent studies conducted by Baylor (2004) implemented two different research designs to study the potential of pedagogical agents to motivate learning by addressing the diversity of learners and the diversity of learning situations. Baylor suggests that because agents are acting as "...a social human-interface" (p. 1), they have unique capabilities that allow designers to be more flexible when designing characters that facilitate learning.

The first study (see Baylor, Shen, & Huang, 2003) recruited 183 undergraduate students enrolled in a computer literacy course. Eight agents were created and each one differed in gender (male, female), ethnicity (African American, Caucasian), and realism (realistic, cartoon). Students were asked to select one of these eight agents. The goal was to examine how the learners' gender and their ethnicity influenced the choices they made for the agents, it also looked at how they perceived the persona of the chosen agents. The content used for interaction between the learner and the agents was related to coping with college life. The results of the first study found that in general learners chose the agent based on the agent demeanor (e.g., behavioral attributes). African American participants were found to choose in greater proportion agents with the same ethnicity and reported more positive attitudes about the chosen agent after learning from it. Given the available choices between realistic and a cartoon-like agent; female's learners were more likely to choose the cartoon-like agent than male participants.

In the second study (see Baylor & Kim, 2003) 139 students were randomly assigned to one of four conditions. The conditions were based on ethnicity (African American, Caucasian) and gender (male, female). This study used an in between-subjects design and focus on agent gender and ethnicity. Baylor (2004) found that the agent gender has a main effect; in contrast with female agents, male agents were found to be more extraverted and agreeable. In addition, with respect to learning and motivation, the students reported more satisfaction with the performance of male agents. Baylor's assessment of an interaction effect between the agent ethnicity and the student ethnicity revealed the following "...students working with agents of the same ethnicity perceived the agents to be significantly more engaging and affable" (p. 1). As previous studies have suggested (e.g., Baylor and Kim, 2003), the preliminary results from these two studies revealed that the "...agent *image*" (p. 1) can indeed play a key role in the motivation of learners.

David et al. (2002) conducted a study employing an anthropomorphic agent situated within a quiz environment, in which the agent offered help to the students while taking the quiz. The agent was endowed with two simple anthropomorphic cues, a name and passport-size picture of a male. The goal of the study was to examine the perceptions of fairness participants felt when the agent offered help that either led them to success or failure. The findings suggested that the anthropomorphic cues reduced the perception of fairness of

the quiz; in particular more marked differences were observed when *male* anthropomorphic cues were added to the help agent.

Clearly these findings are somewhat contradictory to the results found in Baylor (2004) which indicated that male agents were found to be more extraverted and agreeable. This difference may be due to the context in which the agents were used. In one study, the agent is actively involved in helping the user make decisions by providing hints to the user that are related to the question being answered in the quiz (e.g., David et al., 2002). Baylor's study framed the agent in a totally different way that was not related to success or failure of the participant as a result of the interaction with the agent (e.g., coping with college life). This implies the importance of paying particular attention to the design of agents that are suitable to the tasks, content and environment as well as the selection of instances where an agent should be present or not present in order to maximize effectiveness and performance of the agent, and minimize unintended effects on participants' perceptions.

1.6 AGENTS AS SOCIAL PERSUADERS

Kim and Baylor (In review) used social cognitive perspectives such as distributed cognition, social interaction, and social cognitive theory to create a conceptual framework that explores the significance and potential of pedagogical agents based on social cognitive literature. The social cognitive framework looks at the particular advantage of designing pedagogical agents as learning

companions, also known as PALs. The main difference between a “pedagogical” agent and “learning companion” agents is that pedagogical agents guide the learner through the learning process, and learning companion agents work with the user in collaboration towards learning (Baylor, 2001). Kim and Baylor (In review) described a number of potential benefits that although hypothetical, they represent guidelines that could serve to further the empirical research regarding the efficacy of PALs.

The value of learning with peers is clearly present and emphasized in the social cognitive theory. Interaction among peers causes “... a great deal of psychological modeling” (Kim & Baylor, In review, p. 5). Thus, peers that can identify with each other more easily can be potentially more open to motivation and learning from other peers who may be more “skilled”. Kim and Baylor hypothesized that this same potential can be harnessed with PALs to take advantage of “...the benefit of peer interaction” (p. 5). They point out that empirical research supports findings that peers as partners are often more effective than their adult counterparts in learning and motivation in a variety of subject areas and age groups.

Furthermore, Kim and Baylor believe that PALs are able to overcome some of the limitations of human peers. For example, researchers have argued that free and open forums foster the proactive exchange of ideas among peers, but this is not always the case, “...this type of transactional intellectual

partnership is not always possible to implement in classrooms for political reasons” (p 5). Indeed, humans are complex beings and not necessarily follow the proposed or established channels for communication, even when we know that doing so may be beneficial. They proposed that PALs can be designed with exactly those peer characteristics that will make a peer more desirable to work with, which are not easy to manipulate in real human peers. As they state:

PALs could be designed as peer tutor or peer tutees with different levels of competency. The gender of PALs can be matched with learners’ gender to facilitate collaboration and to change their attitudes toward computing. Or the number of pals can be adjusted to consider different instructional strategies and contexts (p. 5)

Whether agents are interacting with the user as PALs or other roles (e.g., assistant, mentor), one of the most important factor to consider when designing an effective user experience is the relationship between the user and the agent. In order for the user to learn, a certain level of confidence in the computer product (agent-based environment) needs to be built from the ground up. Baylor (2001) proposes that the agent must be believable to inspire that confidence in the learner. To inspire trust; “the agent must demonstrate that it is competent, and a motivational attachment must be developed between the learner and agent” (p.10). The agent must be able to show control and competency over the

content and learning process, whether he is guiding the user through it or doing peer collaboration "...so as to earn the learner's trust and confidence" (p.10).

Fogg (2003) coined the term "Captology", an acronym for computer as persuasive technologies. Fogg is an experimental psychologist whose works has focused on the design, research and analysis of interactive computing products that are created with the purpose of attempting to change people attitudes and behaviors. He asserts that in many cases computing products need to be perceived credible in order to be persuasive and successful; and in those cases we need to pay attention and understand how the product's credibility affects the human-computer interaction. He defines credibility as a perceived quality with two dimensions: trustworthiness and expertise. Computing products that have high levels in these two dimensions are more likely to produce higher credibility perceptions in the users and therefore have more potential to persuade. Fogg hypothesizes that if one dimension is very strong, the other dimension may be perceived as credible as well, due to what he calls the "halo effect"(p. 92). However, he points out that the opposite also applies, if one dimension is known to be weak, overall credibility of the product may suffer regardless of the other dimension.

Baylor (2001) acknowledged the development of a social relationship between learner and agent as a fundamental process to create a rich environment where interaction and learning can be promoted. Along this line, we

must take into account that social relationships between agent-learner can take many shapes (e.g.; coach, assistant, learning companion) and in a wide variety of contexts. These two factors alone can help determine the approach that should take place to foster the product's interaction. Fogg (2003) proposed a list of seven contexts in which, he believes credibility is essential to achieve the human-computer interaction desired:

Credibility matters when computers:

- 1) Instruct or advise users
- 2) Report measurements
- 3) Provide information and analysis
- 4) Report on work performed
- 5) Report about their own state
- 6) Run simulations
- 7) Render virtual environments

These contexts are not mutually exclusive and given the type of interaction desired for the agent system, several of these categories are likely to be integrated in one single product. Fogg (2003) considers that as designers take into consideration and "...understand and differentiate among these contexts and categories..." (p. 141) computer products will be designed and perceived as more credible and trustworthy. Certainly, we can appreciate that the list has

practical pointers to take into consideration when defining the type of social relationship and contexts in an agent-based environment.

Fogg has explored the role of computing products as persuasive social actors building on the empirical evidence that people respond to computers socially (e.g., Reeves & Nass, 1996). He asserts that this social response has "...significant implications for persuasion" (p. 90). He conducted a couple of studies in the mid 1990's to find out the persuasive power of a computer product that used the principle of similarity. He defines the principle of similarity the following way: "People are more readily persuaded by computing technology products that are similar to themselves in some way" (p. 99).

The first study called "the personality study" looked at how people would respond to computer with personalities. For the study, two computer personalities were created; one dominant, and the other submissive. A total of 48 participants out of 200 students were chosen as the final group, because they matched the criteria of being and the extremes of these two personalities. Participants were mixed and matched randomly with the computer personalities.

The second study called "The affiliation study" was composed of 56 participants, mostly students a few people from the Silicon Valley community. The study looked at the persuasive power of similarity in affiliation, "...specifically, the persuasive impact of being part of the same group or team"

(p. 99). In this study, two computers were setup; one was identified with a label as being the teammate of the participants, and the other was given no label. In addition, the computer labeled as the “teammate” had a blue color frame around the monitor which matched the color of a wristband given to the participant working with that computer. The no-label computer also had a blue framed around its monitor, but participants working with this computer were given green wristbands instead.

All participants had the same interaction with both computers; the only difference was the assignment of the computer as a teammate, or to the one with no-label. Analysis of the results revealed that people working with the “teammate” computer reported the computer being more similar to them, and thought that the computer was smarter and offered better information. Fogg (2003) also measured the participants’ behavior and found out that “... teammate computers were more effective in changing people’s behavior” (p. 99), because people were more likely to choose the problems solution that the “teammate” computer offered to them.

The findings from these studies indicated that participants preferred to work with a computer they perceived as having a similar personality style or affiliation as that of the participant. The practical implications of these results suggest that designers can exploit this similarity principle to create more effective

persuasive products by matching the general characteristics of the target audience or addressing similarities in some other way.

Fogg cautions about the used of psychological social cues when designing persuasive computer applications “...when you turn up the volume on the “social” element of a persuasive technology product, you increase your bet: you either win bigger or lose bigger, and the outcome depends on the user” (p 114). He manifests that designers must be able to understand and make appropriate use of those cues to make a positive impact. Otherwise, as designers, we stand to lose our product’s credibility and make our users feel miserable.

He believes that there are some areas where people are more likely to embrace and welcome explicit social cues in the product. He explains:

In general, I believe it’s appropriate to enhance social cues in leisure, entertainment, and educational products (smart toys, video games, kid’s learning applications). Users of such applications are more likely to indulge, accept, and perhaps even embrace an explicit cyber social actor either embodied or not”(Fogg, 2003, p. 115)

On the other hand, he argues that social cues should not be enhanced when the goal of the product is to improve efficiency (e.g., buying clothes, books,

typing a paper in word processor). He believes that adding a “social” element to these types of interactions will do more harm than good; slow things down and decrease the efficiency of the user, resulting in people being disturbed and annoyed.

The bottom line is that we must learn to identify where and when to make our application more socially explicit or less socially explicit and differentiate the level of “humanness” our product should have, keeping in mind the needs of the target users and making sure that their needs are covered, rather than letting the complex cloud of technology drive the design of the application. Through “interactivity”, the computer offers designers many advantages to be creative and flexible in the design of products by employing the power of the technology chosen. This modality can be harnessed to match people’s preferences of interactive experience (e.g., text, visual, audio) and allows us to deliver the optimal persuasive impact to users.

CHAPTER 2

AGENT TECHNOLOGIES

2.1 OVERVIEW

A number of commercial web-based agent software programs appeared in the past few years opening new ways of using the Internet to deliver agent-enhanced online content ranging from marketing, education, to hosting sites that deliver real time customer service. Penelope Patsuris (2004) from Forbes.com magazine reports on the growing use of online characters and their impact on online sales:

People feel more comfortable interacting with an onscreen character, even if it isn't a video, and companies know this, which is why virtual characters are playing a growing role in online sales and marketing. A study... by Byron Reeves, Director of the Center for the Study of Language and Information at Stanford University, found that a virtual customer support character on Buy.com cut the cost of support as a percentage of revenue from 2.06% to 0.57%.

This chapter analyzes the six products located through consultation with experts and through web searches that seemed to hold the most promise for possible use in the planned online quit smoking coach application.

2.2 TECHNOLOGIES

2.2.1 Oddcast Sitepal

Oddcast Sitepal allows users to create web animated characters delivered using flash animation. Characters are presented in 2D and have sets of built in customizable features to allow designers to modify a number of agent characteristics to fit their needs. Features like age, skin, gender, hair, eye color, clothing, accessories and make-up can be modified through an easy to use, intuitive interface. Users can add their own audio message using a Text-to-Speech (TTS) option or they can record the message with their own or an actor's voice.

The Sitepal characters are created using an online application; they are server-based, and thus “live” online. No software needs to be downloaded or installed in the computer in order to design them, but they cannot be used in offline applications. Creating a Sitepal character has a three step process, (1) Design the character (e.g., gender, hair, skin, age, etc) (2) Record your own voice message, use TTS or select from a pre-recorded set of messages to talk back to the users and (3) publish the resulting lip-synced, animated agent plus audio or speech directly to a web page by pasting code into an HTML page.

Sitepal works well in Mac or Windows machines, requires not special plugins, it is small in file size and optimized for 56k modems. Messages recorded



Figure 1: Sitepal online Interface for designing agents
<http://www.oddcast.com/sitepal/>

are automatically lip-synched into the animated character. Sitepal also offer an Artificial Intelligence (AI) engine that can enable the agent to respond to questions the user asks of them. Users ask questions by typing text, and the character responds with TTS. However, at this point in time the feature requires some basic java programming and has other limitations, one of them being that the agent knowledge base use by the character is already built in and cannot be customized to fit a particular or specialized knowledge base. Given the preset condition of the knowledge base, its real time AI query-response is limited to be use only in the English language for now. Oddcast plans are to remove these two limitations in later versions of the Sitepal online application.

The TTS feature can be used with dynamic or static text. With the dynamic feature the character can be instructed to speak when a link is clicked, or when the mouse rolls over a certain area of the page, these dynamic behaviors are achieved through JavaScript. Oddcast has also developed integration with many backend programming languages (e.g., PHP, ASP, JSP, Python, etc.).

Even though the application has great promises for rapid development of characters with voice, accessories, age, gender, etc., it is still limiting in terms of having the users only modify certain characteristics of the preset characters with materials already built into the application, which does not give any room for developers to create their own truly custom agents and integrate it with the character functionality. Because the application lives online, developers cannot fully own the sitepal software, instead, they are required to select the type of account they want to license, and based on their selection they will pay a renewable fee. An intro account would license the application for one month, and a premium account includes a three year license. Licensing the AI technology costs more than simply licensing the animated lipsynch and speech function.

2.2.2 Artificial Life (ALife)

ALife focus on offering uniquely conversational smart bot-based products for customer service, consultative selling, web site navigation, automated e-mail

response and financial portfolio management. The company offers a suite of products called the ALife-SmartBot suite to build agents that are able to engage people in conversation using natural language. Among some of the applications ALife agents are being used we find friendly salesperson, website tour guides, personal assistant, and personal teacher.

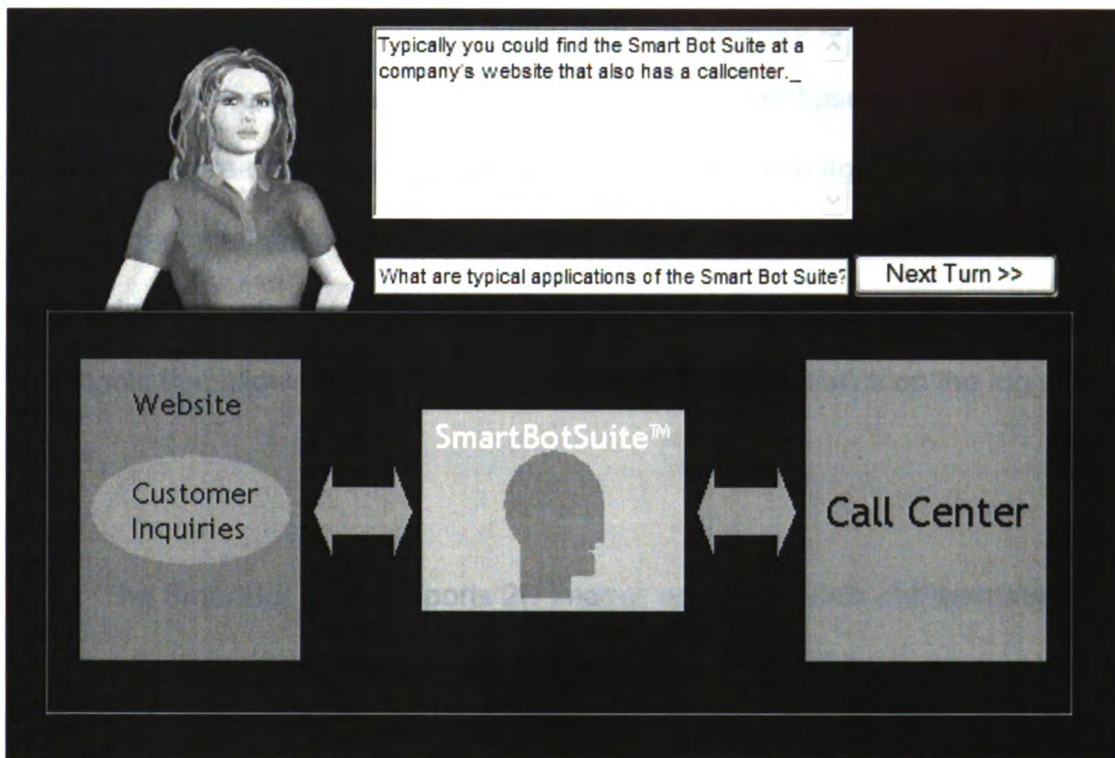


Figure 2: ALife online demo of Luci, a website tour guide
<http://www.artificial-life.com/en/demos/SBSDemo/>

ALife agents make use of an AI engine to communicate in real time with online users. ALife agents are able to direct users to specific web pages; answer questions and even make recommendations. The agents are able to access any industry standard relational database to retrieve dynamic information without need of updating the AI knowledge base of the agent. The SmartBot suite also

provides utilities to create log files of all agent sessions and it is able to provide complete scripts of dialogs between the user and the agent. This is a feature that depending on the focus being given to the application can become a valuable asset for collection of information.

Perhaps, one of the unique features of this software is that it is capable of switching from the agent-based conversation to a human-based conversation when certain pre-determined conditions are met. The Alife agent will provide the human agent a summary of the conversation that has already taken place as well as any data recorded during the user previous visits. The program also includes other tools that allow the developers to run reports and statistics on the logged scripts.

The SmartBot suite supports 2D images and 3D models and animated mannerisms seem to be limited only by the amount of creativity developers put into it. Developers can make use of a bot's animation kit to give full freedom of movement to the character or just as well used the more subtle approach of a wink and a smile.

At the time of this research, it does not seem Alife characters support any type of TTS technology, voice recording options, or lip-synching from text. Thus, the most prominent channel the agent uses to communicate with the users is through text-based chat boxes. Overall, this product has a wide variety of tools

and features, and allows greater flexibility to customized characters. On the other hand, the learning curve to develop ALife agents with the software seems to be quite steep.

2.2.3 Extempo

The Extempo application employs expert 2D and 3D photorealistic characters designed for e-learning applications. Through natural language conversation, expert characters can interact, act like people, and engage users by offering advice and guidance. These characters are often integrated with Internet-based corporate training content and are currently being used to train company employees in management and leadership skills, customer service, product information and business ethics.

The technology and tools facilitate the users to build three different kinds of expert agents: expert coaches, expert role-players and expert guides. Expert coaches are one-on-one training coaches whose objectives are to help users master web-based content. Expert role-players can act as practice partners in role-play exercises, and expert guides can be implemented to walk the user through training content (e.g., company procedures, policies).

Users are able to author their own agents with a graphical toolkit, integrate their content and interface with any database to retrieve or record information. Utilities for administering learner performance and analysis are also provided.



Figure 3: A sample of the Extempo expert agents
<http://www.extempo.com/index.html>

This application is AI intensive, and will require the user to develop an extensive knowledge base for their characters. However, some default characters already built into the application have knowledge bases that can be modified to accommodate new knowledge content. Extempo does not support any type of motion, or video capture, but it does offer support for Emotional Markup Language (EML) to allow developers to program characters to portray common emotions (e.g., sad, happy, afraid, surprised, etc). In addition, the company provides a range of services from e-learning consulting to training users on character authoring, content integration and learner analysis/ administration as well as offering customized content development.

2.2.4 Pulse 3D

Pulse 3D had the most realistic and high quality visuals and very good lip synch. Their characters, called Veepers, can be based on photographs of actual humans. They have proprietary software for converting photographs into animated agents. They license the creation software (Veepers Integrated Development Environment) – and they offer custom character design services. Although the description claims to be able to do text to speech, this actually need to occur before delivery, because the system converts text to speech and turns it into a .wav file, which then needs to be downloaded on demand. Download times are as long for TTS as for recorded voice, since both effectively download .wav files. And, dynamic use of text to speech is not dynamic in real time.



Figure 4: A Pulse 3D Veepers agent
<http://www.pulse3d.com/>

No plug-ins are needed for users to view Veepers, which can be delivered over the web or cell phone. Once created, Veepers can be deployed as stand

alone characters or web-based ASP. To include dynamic as opposed to static, pre-planned speech requires licensing Veepers WebService.

If cost were not so prohibitive, this would have been the technology of choice as far as quality of visuals and customizability of character design. The AI was weak and dynamic text to speech was not possible.

2.2.5 Microsoft Agent

Microsoft Agent is a commonly used agent technology for research. The four present agents are cartoon characters – Peedy the parrot, Genie, Merlin, and Robby. An advantage of this software is that the developer can position Microsoft Agents anywhere on a web page or on the Windows desktop. A variety of complex licensing and distribution agreements need to be completed for various Microsoft technologies in order to distribute and use Microsoft Agent in an application. The agents only work on Windows computers, and running Microsoft Agents requires downloading and installing the Microsoft Agent software and configuring it correctly. These limitations did not meet the long term needs for the Online Quit Smoking coach.



Figure 5: Peedy, a Microsoft Agent
<http://www.microsoft.com/msagent/>

2.2.6 Famous3D

With Famous3D netFace [complete] software application web and multimedia designers have a great deal of flexibility to create virtual characters. Using netFace [complete], designers can drive the delivery of their content by using text, audio or video input. The application can handle imported 2D graphics and photorealistic 3D models that can be easily deployed on the web by embedding a shockwave movie file into an HTML page or using a proprietary Famous3D streaming plugin called famous3D Viewer with a file size of 165Kb, downloadable in approximately 30 seconds using a 56K modem.

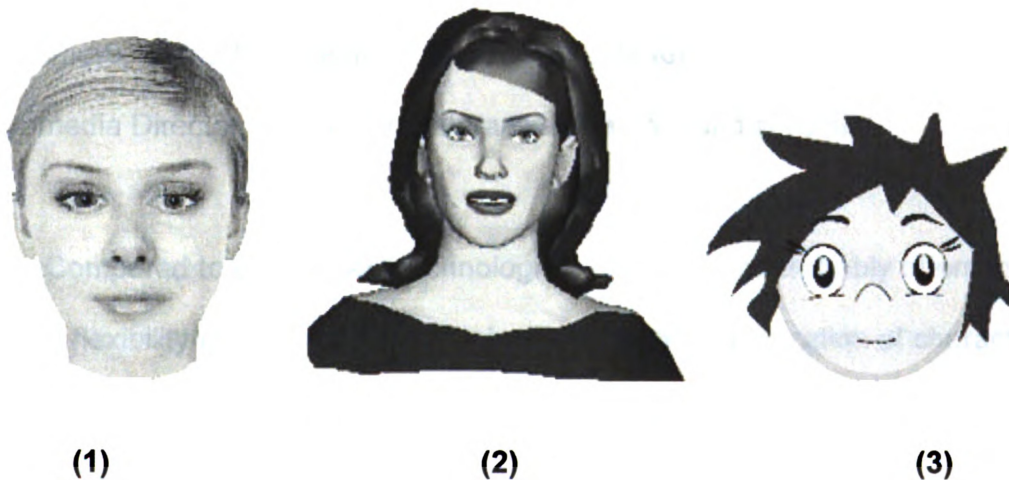


Figure 6: Famous3D Talking head models, from left to right: (1) Photorealistic model (2) 360 degree model (3) 2D Cartoon model
<http://www.famous3D.com>

When text is used as an input, several sources are supported including XML files, content from databases, AI engine, or a human typing the text directly. The content taken from any of these sources is sent to an animation engine which works in conjunction with a TTS engine to produce the voice synthesis,

once the audio file is created by the TTS engine, it is passed back to the animation engine which then uses that information to create the appropriate agent's facial movements and lipsynchs the audio. In addition, if a recorded human voice is the source, then a similar process takes place in which the voice is automatically analyzed and the appropriate facial movements are created.

Famous3D netFace [complete] also supports EML and multilingual TTS engines, and perhaps one of the most compelling features of the program is the ability to capture motion data from a video input and apply it to a 3D talking head model that replicates gestures, head movements and synchronizes speech. Furthermore, characters create in netFace [complete] can be exported to Macromedia Director to add more interactive features and controls.

Compared to other agent technologies, this software probably offers the greatest flexibility in terms of inputs, animations, and customization of characters; however, there are still some technological obstacles that will need to be resolved before the application can be made fully compatible with all web platforms. For example, at the present time, the famous3D viewer can only be used in PC's with Internet Explorer; and it does not support other browsers or operating systems.

2.3 STRIKING THE IRON WITH THE RIGHT HAMMER

When conducting the research of these agent technologies, I looked at certain products features to help determine which one of these programs (if any) could be the best design package to help us reach the greatest majority of users and yet maintain the greatest flexibility for development, authoring, customization, and integration with the backend system upon which the online electronic quit smoking coach is built. A great deal of attention was paid to whether software will support both low and high bandwidth delivery, its multimodal capabilities for input and output of content, the ease of customization and authoring of agents, the learning curve for designers, cost of software and development time.

What I found in reviewing these set of software applications was a mix of desirable features scattered across the different products available. However, two characteristics were found in common in all these applications either, (1) you could buy and own the software, but the price for this commodity was very high, almost to the point to declare it prohibitive for most money-minded budgets doing exploratory research with this technology, or (2) the software will live online in a remote server (not own by you), you would still pay a license fee and the license fee would still be relatively expensive depending on the level of access and tools you desired, you are still limited in how much you can customized your characters to fit your project's particular needs, and given than you do not have absolute control of the environment where it resides, you have to be willing to

jump many more technology obstacles to get it to integrate with your backend systems.

Using Extempo for a single small scale (less than 50 users) trial would have been at least \$15,000. Conversations with Pulse 3D resulted in a quote of a “special academic rate” of more than \$100,000 per year licensing fees.

The review of these commercialized agent technologies and some of the limitations found seem to support what researchers have found regarding web-based agent design and the issues still present with regard to platforms and networking of these technologies. As Johnson et al. (2000) explain: “...integration of animated pedagogical agents into Web-based learning materials inevitably entails working around the latencies...” (p. 73). Furthermore, they stressed that in order for an agent design to be successful, it must take into consideration first the capabilities of the intended platform and network where it would be used (Johnson et al., 2000). Kim and Baylor (In review) also point out the many constraints still available in designing agents based environments, “while there are some ready-made agents available...it is usually necessary for researchers to develop their own...for their particular research questions” (p 23).

No doubt, it was beneficial to look at these applications to get a feel about the current state of the technology; on the other hand, considering the various constraints of the software reviewed and the particular needs of this project with

regards to platform and network where it needed to be integrated, we decided the best approach was to develop our own agent software. To do this, we looked at Macromedia Director MX as the authoring environment. Director offered great flexibility to be creative in designing and customizing agents and the user interface; it also supported TTS and synchronized text captioning. Through Lingo (Director's programming language) designers have the ability to create interactive interface controls to let users manipulate the multimedia content. Furthermore, Director's Lingo allows for easy integration to the Online Quit Smoking Coach backend database system to insert and retrieve dynamic information on real time.

Director content can be exported for web delivery using the Macromedia Shockwave player, which is considered the web standard for powerful multimedia playback with a 55.5% penetration of Internet viewers. This represents a number of 289.8 Million people according to a NPD Research online survey conducted in September 2004 (see Figure 7). Among some of the most common uses of Shockwave content being delivered on the web we find 3D games and entertainment, interactive product demonstrations, and online learning applications.

In addition, Shockwave player supports many widely used operating systems (e.g., Windows, Mac) and browsers (e.g., Internet Explorer, Netscape); all of these features made Director an excellent choice as a digital media tool to

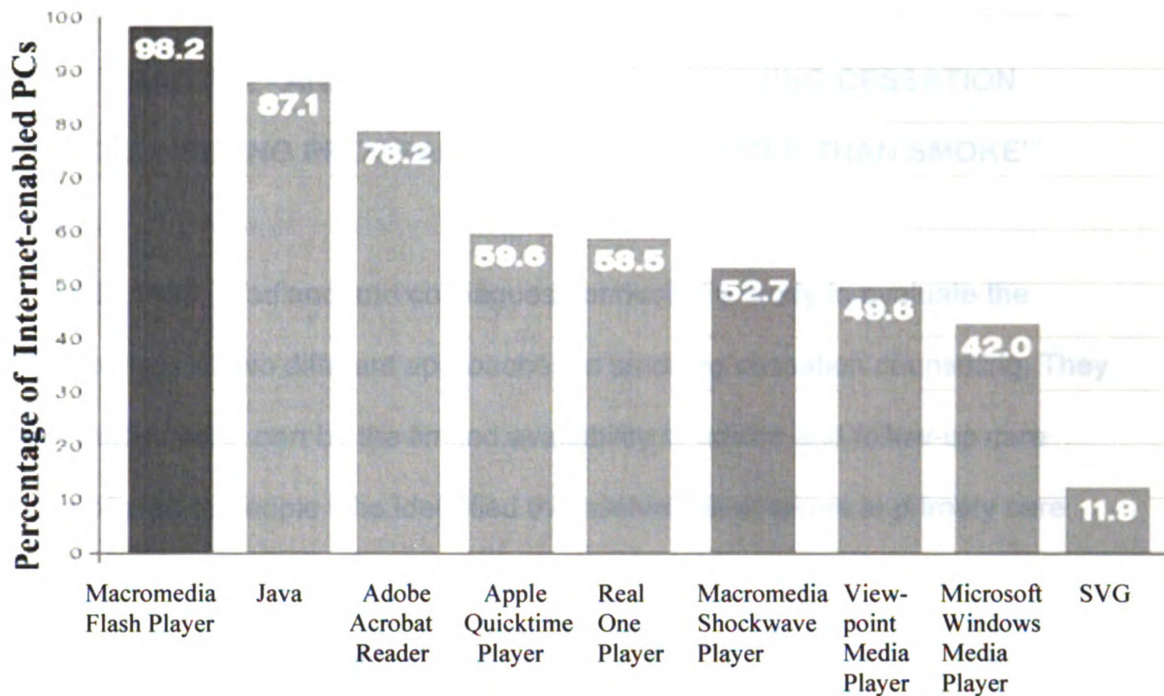


Figure 7: Shockwave player adoption statistics
http://www.macromedia.com/software/player_census/shockwaveplayer/

work with. It allows the designer to have rapid access to all levels of the production; while maintaining modularity (i.e. agent character, user interface controls, TTS, and server communication). As a result, rapid prototyping and testing can take place with the least amount of time and money.

CHAPTER 3

PRACTICE- AND COMMUNITY-BASED SMOKING CESSATION COUNSELING PROGRAMS: "I'D RATHER COPE THAN SMOKE"

In 1999, Wadland and colleagues conducted a study to evaluate the effectiveness of two different approaches to smoking cessation counseling. They were motivated in part by the limited availability of advice and follow-up care being offered to people who identified themselves as smokers in primary care practices. Guidelines on smoking cessation as indicated by the Agency for Health Care Policy and Research (AHCPR), recommend that physicians should provide both brief advice and follow-up care for all smokers (Wadland et al., 1999); yet as the researchers pointed out "... the actual implementation of these guidelines is limited" (p. 711).

The study tested two different smoking cessation counseling approaches in two settings: practice-based and community-based. Both approaches employed the same number of counseling sessions consisting of 1 recruitment/intake session followed by 6 counseling sessions. The main difference between the two approaches was that in the practice-base approach, counseling was provided by nurses and telephone counselors and in the community-based approach, counseling was provided by telephone counselors only. The great majority of people who received medical services in the practices selected were low-income or Medicaid patients. For the practice-based

approach, participants were identified by physicians during their office visits to the primary care facilities and if interested, they were referred to a nurse to obtain consent for the study to begin taking an intake survey session. Likewise, participants for the community-based approach were recruited in the same communities where the medical facilities of the practice-based approach were located. Because the nature of the community-based approach was a “telephone only” interaction, recruitment of participants was conducted thru newspaper ads, and public service announcements. Also, a hotline number was given, so that interested participants could call-in to begin the program.

Physicians were trained according to the smoking cessation recommended guidelines of the AHCPR, which lays out five major steps needed to achieve a successful intervention. Together, the steps form a group called the “5 A’s”; namely the five A’s are: ask, advise, assess, assist, and arrange. To ask, is to identify and document the tobacco use for every patient; to advise, is to urge the tobacco user in a clear, strong and personalized way to quit; to assess, is to determine if the smoker is willing to make a quit attempt at the time; to assist, is to provide the willing patient with counseling and pharmacotherapy; to arrange, is to schedule follow-up contact by person, or by telephone, within a few days after the quit date (United States Department of Health and Human Services, *undated*).

The counseling sessions offered by nurses and telephone counselors were driven by the “I’d Rather Cope than Smoke” computer application developed by the researchers. The application is based on a model of counselor support and relapse prevention (RP) and given the nature of the computer-driven counseling sessions; appropriate training on relapse prevention, computer skills, as well as individual case management was provided to nurses and telephone counselors.

The “I’d Rather Cope than Smoke” application employs computer prompts to cue the counselors with questions they should ask the participants. The software is programmed to adapt questions according to the answers given by the participants during the session. It has the capacity to retrieve information from previous sessions which counselors can use to provide more effective relapse prevention therapy. Because the program’s main target audience was low-income smokers, many of the application questions were formulated to target that group.

At the heart of the program, there are 70 relapse situations, and 90 coping responses, which counselors used to ask smokers about relapse situations and to help develop customized coping strategies to deal with those situations. The breakdown of the counseling intervention sessions has the following intervals: 1, 3, 7, 14, 30, 60, 90, and 180 days after the quit date set during the intake survey. Of these, the first six sessions are very active counselor-patient sessions, in

which relapse situations and coping responses are actively discussed, and the latter two (90, 180 days) are follow-up sessions.

The results obtained in the study indicated a quit rate (7 days smoke-free status) at 3 months of 8.1% in the practice-based group and a 21% in the community-based (telephone only) group (Wadland, Stoffelmayr & Ives, 2001). In a more recent article about the examination of the process of RP therapy designed to aid smoking cessation Stoffelmayr, Wadland, and Pan (2002) found that patients would have better chance for positive outcomes by completing a greater number of counseling sessions and by discussing a greater number of coping responses per session. To stress the importance of these findings they explain that “training in coping is the essence of RP therapy” (p. 1357).

Wadland et al. (1999) considered that the results of the study (i.e. practice-based vs. community-based approach) validated the practicality and effectiveness of the telephone-based approach when customized counseling is provided, specifically with regards to low-income smokers, given that they were the primary audience of the program. Taking this study as the basis, the researchers encourage further research should continue in this area, and specifically, on how these relapse prevention approaches can be modified and adapted to reach and meet the needs of other special groups or audiences.

CHAPTER 4

APPLYING CAPTOLOGY TO SMOKING CESSATION: AN ADAPTABLE APPROACH

4.1 OVERVIEW

Based on the suggestions formulated by Wadland et al. (1999) to further the research of new intervention approaches to smoking cessation programs and the success they had with their computer-driven, telephone-based counseling program; it seemed plausible to take that challenge and conduct research in that area based on my interests on online education, human-computer Interaction and applications of new media technologies. Thus, with a focus on healthcare and based on the architecture of the existing “I’d Rather Cope than Smoke” software application developed by Wadland and colleagues; I decided to conduct the research and production of this thesis by drawing from the Captology framework to help devise a new approach to a smoking cessation program.

4.2 A FUNCTIONAL SHIFT ON PERSUASION

If we recall, Captology is the study of computers as persuasive technologies and it focuses on the design, research, and analysis of interactive computing products that are endogenously created to change people’s attitudes or behaviors (Fogg, 2003). Certainly, the mere use of the words ‘persuasive’ and the attempt to change people’s ‘attitudes’ or ‘behaviors’ would raise an eyebrow, or two, as being an unethical method and use of the technology; but

this is not necessarily truth in all cases. Fogg explains in his book that his vision and the purpose for putting this new body of knowledge in text is so that we can have "...a collective understanding of persuasive technology so computing products can be created to influence people in ways that enhance quality of life" (p. 251). Fogg is aware that persuasion can be unethical; in fact he devotes a whole chapter in his book (see Fogg 2003) to discuss the ethics of persuasive technology. One just have to take a brief look to find plenty of unethical persuasive and deceptive examples out there, to cite a specific example that has a clear relationship to this project, we just have to look at the ads and commercial that try to persuade people to smoke, particularly those designed to reach the more vulnerable groups like teenagers and specific ethnic groups (see Figure 8).

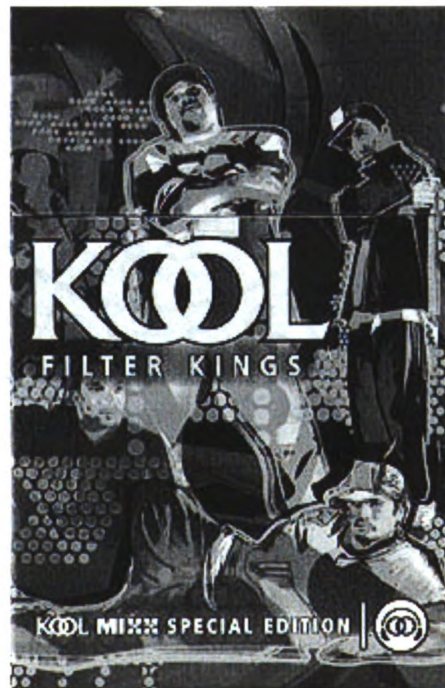


Figure 8: Cigarettes' ads with hip-hop DJ's and Dancers that critics say targets black youth
<http://www.courier-journal.com/>

On the other hand, there are already some positive and empowering implementations of persuasive technologies that have proven to be quite effective in reaching their objective. For example, alcohol 101 plus, is an interactive CD-ROM that helps colleges students make safe and responsible decisions about alcohol. One of the segments in the program portrays a virtual bar environment in which students can interact to see how various factors can affect the person's blood alcohol concentration (BAC). Users can enter values such as gender, weight, quantity and type of beverage, and see a simulation that monitors the BAC level, and the time it will take for the BAC level to return to normal in relation to the number of drinks consumed (see Figure 9).



Figure 9: A virtual bar simulation shows the impact of alcohol in the human body
<http://www.alcohol101plus.org/>

Just as we find excellent applications such as alcohol 101 plus, we can in a similar way, make use of that very same concept of 'persuasion' to promote social good by designing applications that harness the power of interactive

technology. The ethics of persuasive technology are clearly an area of concern and one we should pay close attention. Fogg believes that education is the key and proposes that at least three things will be required to make the leap forward in this new area of study: (1) awareness of persuasive technologies must be raised among the general public, (2) designers must be encouraged to follow guidelines for creating ethically sound interactive products, and (3) action must be taken against individuals and organizations that use persuasive technology for exploitation (Fogg, 2003).

Primarily, I expect to use the theoretical and practical framework of Captology to draw from empirical studies, concepts and principles of persuasive technology to help inform the design of a new interactive software application that explores: (1) the integration of persuasive computer agents acting as social actors in place of human counselors to offer smoking cessation counseling, and (2) employs the Internet as the medium that delivers the information through a web-based interface in place of the telephone-based system, both of which were originally present (i.e., counselors and telephone system) in the “I’d Rather Cope than Smoke” application.

4.3 HOW COMPUTERS PERSUADE

One of the fundamental reasons for using computers as communication technologies (i.e., Internet, WWW, and agents) to frame this experimental research is because they offer a great number of advantages over traditional

media. Particularly, computers have the capacity of being 'interactive'; this implies that a two-way communication can be established between computer and users. This interaction takes place when the user executes a number of actions intended to communicate with the computer, for example, mouse clicks, typed commands, voice commands, etc. At the same time, the computer system can provide immediate feedback, guidance and verification of input data to the user by presenting integrated multimedia information in many different forms (e.g., text, pictures, audio, video, graphics, and animation).

The flexibility of computer technology to adapt to different presentation modes allows designers to potentially match people's preference for a visual, auditory, or textual experience with a high personalization factor and a low implementation cost. Fogg (2003) points out that "... as a general rule, persuasion techniques are most effective when they are interactive" (p. 6). Compared to traditional media, users can exercise a great deal of latitude when using an interactive product; for example, users have more individual control over the medium with regards to exposure time, length and the type of content they choose to view. Although other media such as newspapers, magazines, radio and television also offer control of exposure time and length, they don't give the user any control over the content. (Martin & Leckenby, 1999). Thus, by effectively integrating these presentation modalities, an optimal persuasive impact can be created and meaningful learning that increases retention of information and transfer of knowledge can take place (Fogg 2003, Mayer 2001).

Computers also have some clearly perceived attributes that give them an advantage over human persuaders. Specifically, Fogg (2003) identifies six of these distinct advantages. Computers can:

- (1) Be more persistent than human beings
- (2) Offer greater anonymity
- (3) Manage huge volumes of data
- (4) Use many modalities to influence
- (5) Scale easily
- (6) Go where humans cannot go or may not be welcome

To see the practical implementations of these attributes, we can apply them into the context of the web-based application to help people quit smoking. Computers can be more persistent than a human being on the phone because they don't need to eat or sleep; they don't get frustrated or angry, and can be available 24 hours 7 days a week. The application can offer anonymity by asking the user to create a username and password without need to declare or input any sensitive personal data into the system. Fogg, points out that "anonymity ... is important when people are experimenting with new attitudes and behaviors" (p. 8). An online text-based chat room serves as an example of anonymity; there, users can create a fictitious self and be whoever they want to be. A similar experience takes place in a virtual chat room, where avatars identities are chosen as the person's representation in the virtual world (see Figure 10).



Figure 10: Avatar representation of male and female characters created by users of the Adobe Atmosphere software.
<http://www.atmospherians.com/index.html>

For the web-based quit smoking application, the ability to easily store, manipulate and retrieve volumes of data is a great asset to users and researchers alike. The application can make use of the data it has acquired on previous interactions with the user, and draw from that data to provide the user with the right piece of information or reference at the right time; doing this effectively has the persuasive power to promote the behavior change intended by the application.

On the other hand, researchers also benefit as data gathering in the web-based application is instantaneous and data collected is much easier to manipulate and analyzed than the traditional pencil and paper method. Boyer, Olson, and Jackson (2001) conducted a study about advantages and disadvantages of computer surveys over the traditional print surveys and explain some of the unique potentials:

Another... advantage of electronic surveys is the ability to include pictures, special formatting, audio or video links along with straight text. Researchers can use these features to emphasize or draw attention to critical aspects of a question or ask a new type of question (p. 4)

As described before, computing modalities can be applied to persuade effectively by presenting information in many different formats, giving the computer an advantage that can't be matched by a human being. Only in the digital world can we integrate all these media elements into a harmonized unit that plays the experience in unison (Fogg 2003). Once a design model for the smoking cessation application has been developed, then replication can take place, and with relative ease, we could copy a computer agent acting as a relapse prevention coach and adapt its personality and appearance to match those of a particular target group. At the core of the interaction between the computer agent and the user, the same technologies are still being used, only that now we have the option to scale our software to one or more agents to match particular needs and maximize the persuasive impact.

Computer products are potentially ubiquitous (i.e., existing or being everywhere). More and more, we see seamless connections being created to allow us to reach the 'data' we care about in ways we never imagined or thought possible. Electronic gadgets like cell phones, PDA's, MP3's players, Portable Media centers, smart watches (see Figure 11) represent only a small subset of

these devices, and are indeed some of the most common objects we see in our everyday life, but as Fogg explains, not everything can be seen with the naked-eye anymore:

With the growth of embedded computers, computing applications are becoming commonplace in locations where humans-persuaders would not be welcomed, such as the bathroom or bedroom, or where humans cannot go (inside clothing, embedded in an automotive system, or implanted in a toothbrush) (p. 10)



Figure 11: Electronics gadgets: (1) The new Razor cell phone by Motorola (2) A Fossil Smart Watch, (3) The popular IPOD by Apple

So, an appropriate way to think of how we can take advantage of this omnipresence of computers to extend the reach of the smoking cessation application could be to 'scale' it and adapt it to fit the hardware and software supported in these electronic devices. Indeed, that is a valid approach but there

are also other ways that we can reach and persuade users without sacrificing quality or content. For example, we could use the current state of the technology and have the computer agent send an instant message in the form of audio, text, or video to the user's PDA or Smart Watch and remind them of the goals for the program and when they need to log back to the website for their next counseling session.

A very simple action like this can have an equally important persuasive impact, because it reaches out to the user with a high personalization factor, and we should not rule out the likely possibility that the message arrives at a time and place where the person needs the most (like when a user is having a strong urge to smoke). Thus, we can also use the 'pervasive' quality of computers to communicate our message integrating various technologies to create a seamless experience.

4.4 FUTURE TRENDS

At the heart of this production thesis are two elements that motivate the research, namely; education and promoting healthcare; Fogg (2003) identifies these two domains (among others) as having the greatest potential for quick growth in terms of the persuasive technologies landscape because of the financial benefits it can bring. Healthcare is one of the areas seen as "a vertical market that is most likely to leverage and invest in persuasive technology to promote health (p 245); but why? According to Fogg, insurance companies and

health institutions are aware that many of the problems their patients develop are due to their behavior, for example: “smoking contributes to heart disease, unprotected sex increases risk for contracting HIV, failure to treat diabetes leads to a host of health problems” (p 245). He predicts that these institutions will realize that a great potential for financial gain is possible by proactively investing in the development of applications that not only support health (as many of them already do); but that are actually designed to motivate or persuade people to change their behaviors and live better, longer and healthier lives. The direct benefit for insurance companies and healthcare providers is that, as people become more knowledgeable and learn to take better care of their health; they will save money and boost their profits (Fogg 2003).

The ability to adapt software to match people's preferences and to customize it so that content can be presented in various forms and with various levels of knowledge and skills, makes the education domain a fertile ground for creating persuasive technologies that can (as Fogg points out) “motivate people to initiate a learning process, to stay on task, and then to review material as needed” (p. 246).

To put into perspective the ability to adapt software to meet the intended educational goal and tailor the experience to the user; we can think of the smoking cessation program as having a set of computer agents, each one of them is potentially different in their personality, physical appearance, gender, and

ethnicity; yet each one have the same amount of expertise in the content domain (see Baylor, 2004; Baylor & Kim, 2003; Baylor, Shen, & Huang, 2003). At the core, we have the same basic model of an agent being tailored to potentially match a learner's choice of how they would like to receive the interactive experience. This list of choices empowers the user with some control over the interactive experience; by anticipating the user's needs and wants, we make our software more polite (Cooper, 1999).

These adaptable agents' characteristics represent only a small subset of the type of interactive persuasive elements that can be used to help motivate and influence people to start a program and stay on track. As the technology becomes more sophisticated, we can expect to increase this level of personalization and be more effective; for example, as the user and agent interact over time, the agent gets to know more and more about the user, it is able to build upon this knowledge to adjust its educational approach to be more effective in reaching the practical goals of the system and the personal goals of the learner.

Cooper (1999) points out very eloquently the importance of designing software that meets and satisfies the user's personal goals, and the corporate goals (i.e. the designers' goals): "there is a close parallel between corporate and personal goals: Both are at the highest expressions of goals for their respective

owners. Neither can be slighted. Software that fails to achieve either one will fail”
(p.157).

CHAPTER 5

PRODUCTION PROCESS

5.1 INFORMATION DESIGN

5.1.1 SCRIPT ADAPTATION

For the purpose of this production thesis the original smoking cessation program “I’d Rather Cope than Smoke” by Wadland et al. (1999) was renamed to “Learn to Quit” and a substantial number of adaptations were made to fit with our agent-based approach and a new target audience.

The original target audience for the “I’d Rather Cope than Smoke” program was Medicaid patients. The majority of these patients came from very low-income families; therefore, a great number of questions and scripts in the original program were designed to match that particular target group, their living and working conditions. Since our new target audience for the Learn to Quit approach were college students; we adapted many of these questions to match the new target group environment. A great majority of the changes to questions took place in the first session (intake survey) while the majority of computer-driven scripts used by the telephone counselors in the original program were changed to match the culture and language of the new target group. The exception to this were most dialogues related to the relapse prevention therapy counseling, because much of that content was transferable and required little to no change to be effective with the new audience.

The Learn to Quit agent intervention approach was divided into the following information modules: Program Introduction, First Session, Second Session, Third Session, and a brief Evaluation of Agent section at the end of every session (see Figure 12). Each session was designed to last between 20-25 minutes.

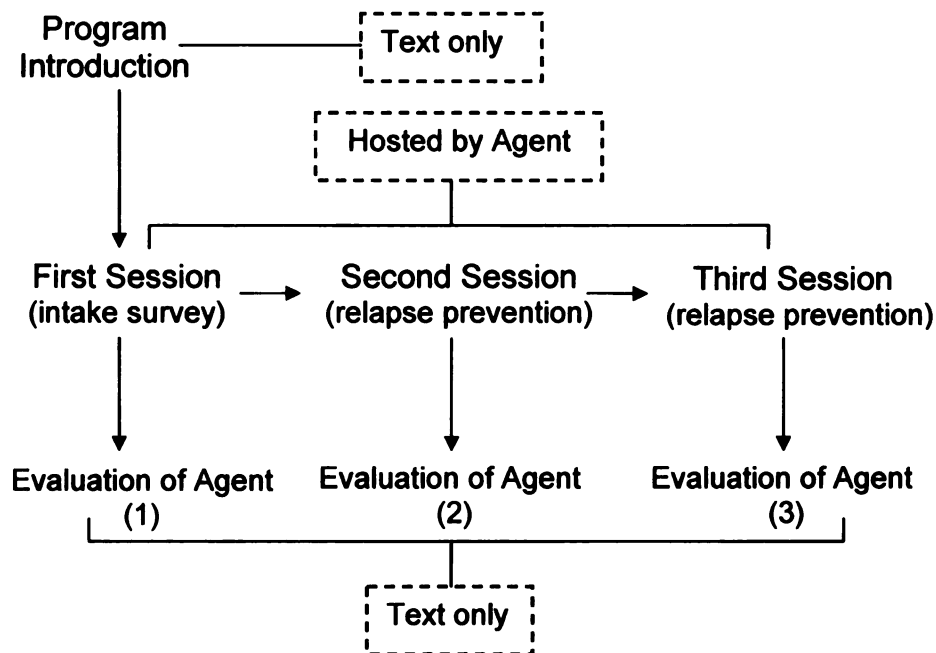


Figure 12: Information Structure of Learn to Quit Application

- **Program Introduction**

The introduction section welcomes the user to the program. The information is presented in text to describe the project, it provides instructions about how to work with the program, and explain the benefits for participating. Participant's consent to take part in the study is obtained at this point. Once the participant has given consent to participate, they are linked to a new screen where the intervention begins.

- **First Session**

The first session introduces the agent character, the agent is used to coach and guide the user through this session and all the remaining sessions. The intake survey is primarily designed to collect demographic data about the participant's smoking habits, emotional state, and readiness to quit smoking. It also introduces motivational planning material that allows the participant to see how relapse prevention counseling will work in successive sessions.

- **Second and Third session**

Session two and three are very different from the first session because they build upon information acquired in session one to put into practice the relapse prevention counseling. These sessions are designed so that the computer agent can assess whether the participant has smoked or not since the last session and to adapt his presentation of material and questions in relation to that status. The goals of these sessions are to teach the participant to recognize situations that tempt or have tempted them to smoke and to plan ahead to cope with those situations or other potential tempting situations. Once, the tempting situations are identified, participants are encouraged to select coping strategies from a list that can be applied to deal with the tempting situations.

- **Evaluation of Agent section**

This section is incorporated at the end of each session and is presented to the user after he/she has completed the main part of the intervention session

with the agent. This section is presented only in text and in the form of a short survey to collect data about the participants' reaction while interacting with the online quit smoking agent.

5.2 AGENT DESIGN

5.2.1 COMPUTER AGENTS

Two anthropomorphic computer agents acting as quit smoking coaches were created to interact with the users (see Figure 13). The agents were designed and embodied with still images to portray two different personas; John, is a stern and tough guy who used to be a U.S. Marines drill sergeant and now performs the role of a coach in the quit smoking program. John knows that his experience in the armed forces and discipline are two key assets that he can bring to the table to help people quit smoking. John, does not like to be rude; but he also doesn't like to spend a lot of time fooling around, he knows that people are there for a reason, so he likes to cut down to the chase and get to work on discussing smoking situations, and help participants to find strategies to solve them.

The second agent, Michael is a more down-to-earth friendly guy, he is always smiling and a bit more sensible in the way he communicates. Michael used to smoke but through self motivation, determination and the support of a smoking cessation program, he managed to quit smoking. He felt so good about rediscovering his new self again that he believed he could make a real difference

by becoming a smoking cessation coach himself. He wants to be there for people who want to quit, he wants others to know that if he could do it, they also can.

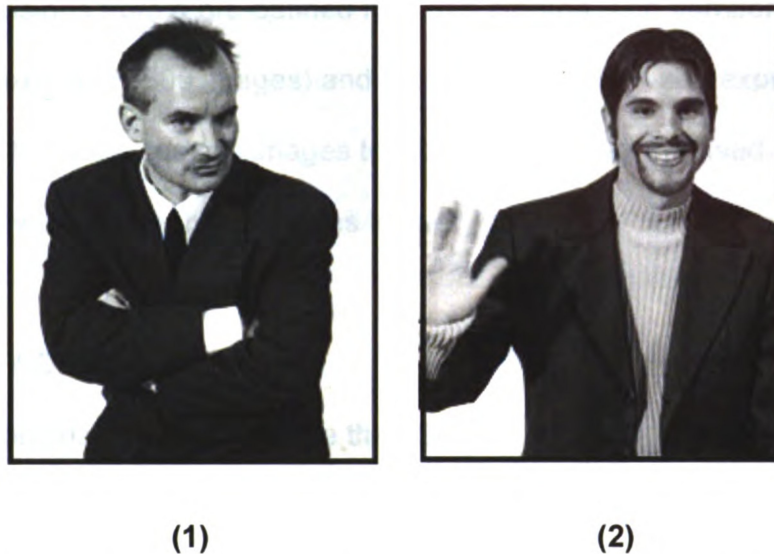


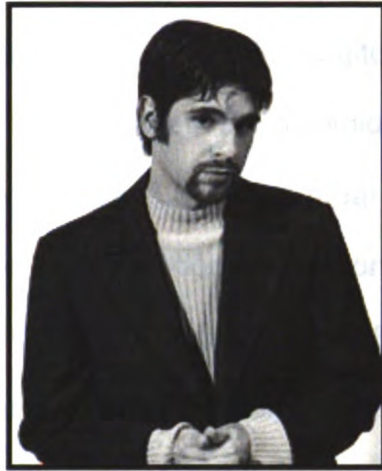
Figure 13: Computer Agents: (1) John and (2) Michael

Each agent has a series of still images to update their body and facial expressions according to the interaction with the user during the session (see Figure 14). The images for the agents needed to meet two specific criteria: (1) they agents needed to appear to be in the age range between 24-30 years old in order to be able to act in the role of coaches that were mature enough to know what they were talking about, and yet, not too old to be out of synch with the values and lifestyle of the target audience, (2) the agents needed to have the looks of having a stern personality and a friendly approachable personality.

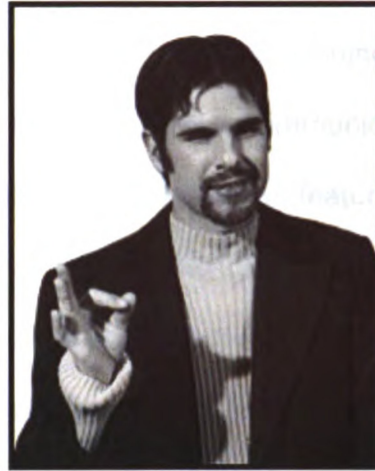
I consulted with a local model agency to search and hire two male models that met these criteria. Hiring the models and a photographer gave me the flexibility to setup a photo shoot to obtain images that would match body and facial expressions from a pre-defined list I had created. The number of images taken (approximately 500 images) and the variety of poses and expressions allowed me to match specific images to the scripts the agents used. It also gave me the ability to update these images in every session.

5.3 SURVEY DESIGN

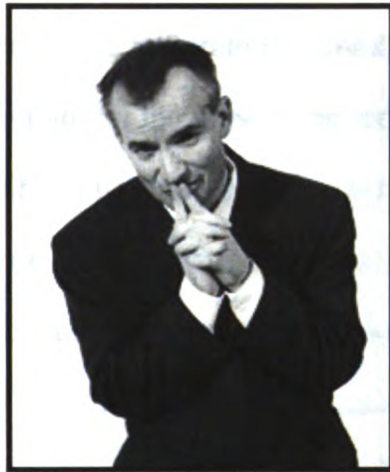
The original “I’d Rather Cope than Smoke” program used computer-driven survey questions that counselors asked their patients as they appear on the screen; as a result, participants of the program never got to see the questions, even though they answered them. Telephone counselors were responsible to enter every answer into the system. In order to make this work for online delivery, we needed to go from a third person delivery to a first person implementation and make all questions visible to the participant on the computer screen.



(1)



(2)



(3)



(4)

Figure 14: Agent's Poses and Expressions: (1) Michael talks about depression and stress, (2) Michael congratulates user with a wink and hand gesture, (3) John discuss the user's smoking history, (4) John talks about importance of finding a support person.

To achieve this goal, we needed a development environment capable of creating electronic surveys that could incorporate a wide range of question types

(e.g., multiple choices, Likert scale, semantic differential, and essay). We decided to use an experimental longitudinal survey engine (LSE) being developed by Department of Epidemiology at MSU and the Communication Technology Lab. The LSE has a great number of survey design features that would make the adaptation of questions to the new environment relatively easy using its web-based interface. The LSE also possessed some advance features not found in the original smoking cessation program; including the capacity to integrate multimedia elements, and linking to external reference material among others.

The ability to integrate multimedia elements in the LSE was a very important feature for this project, because that is how the computer coach interacts with the participant. The coach's messages are divided into a series of self-contained shockwave movies (about 25 per session) that are small in size and quick to load on a web page, even over dial-up connections. In order to understand how the coach interacts with the user, it may be useful to see how the web page is divided (see Figure 15). Basically, there are three sections on the web page: (1) a header, (2) a shockwave movie, and (3) survey questions. The header indicates the current session number; the self-contained agent movie is loaded onto the web page as a media element along with all the related survey questions the user will answer in that page.

Thus, once the shockwave movie is loaded, the coach can speak to the user in relation to the questions in that page, and the user is able to select the answer directly from the choices shown.

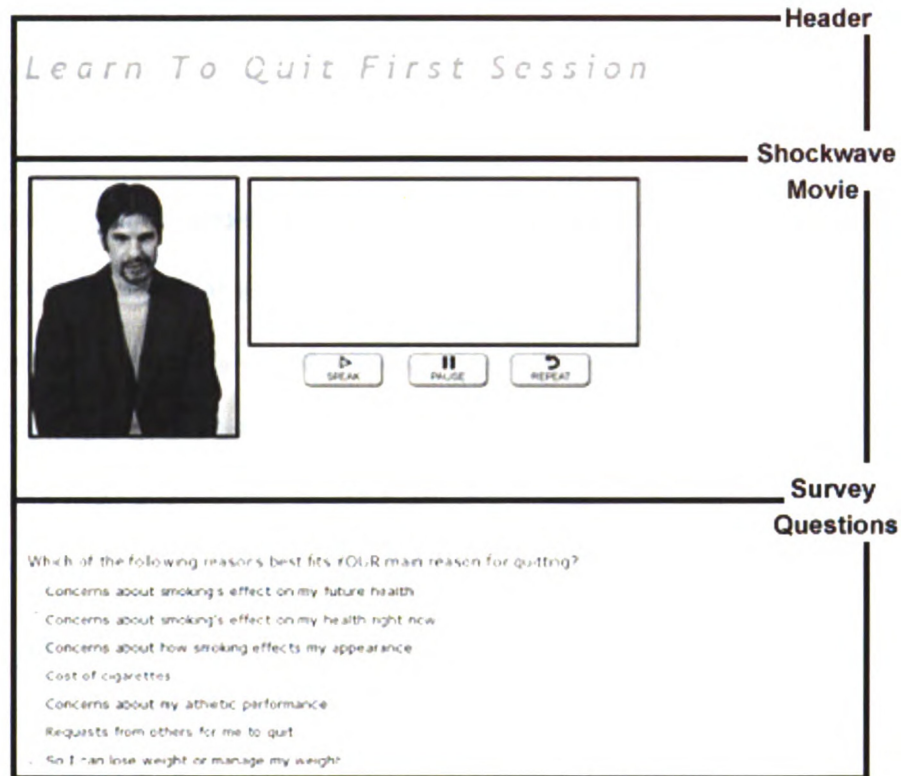


Figure 15: The Learn to Quit webpage dissected

Further, the LSE allows the designer to query data from previous questions or surveys and to program the survey with page skips (branching) in between pages. The ability to program page skips gave me the flexibility to adapt the agent messages and his questions to present new material based on the user's answers, resulting in a more customized experience for the user.

5.4 USER INTERFACE DESIGN

5.4.1 INTERFACING THE USER AND THE COACH

Macromedia Director MX was used as the authoring environment to design the user-agent interaction component of the program. All the images and graphics interface controls were created, touched-up, resized and exported to a web delivery format using Adobe Photoshop 7.0. The interface was designed to be simple and clean to avoid cluttering the screen's real estate space; thus only the required elements that would allow the user to have control of the interaction with the agent are present. As the authoring environment, Director MX served a very important role in the integration of various functional features of the user interface. The user interface is divided into three areas: (1) The image pane, (2) the captioning pane, and (3) the interactive control buttons (see Figure 16).

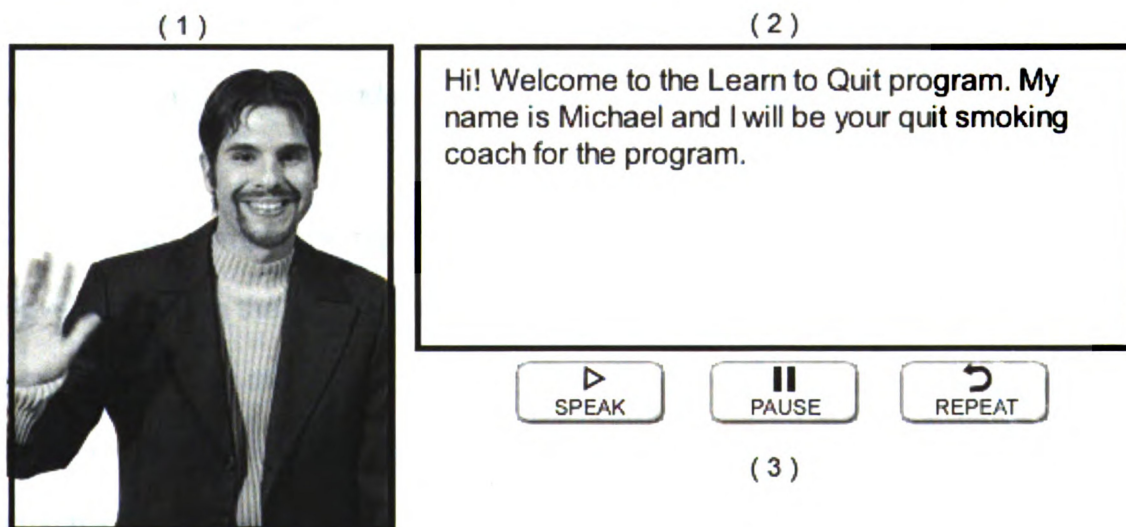


Figure 16: Learn to Quit User Interface

The image pane is located on the left side of the interface; images appearing in this window are updated to match with body position and facial expressions the message being communicated to the user. The captioning pane,

located to the right of the interface displays all the animated text spoken by the agent. This text is presented in a left-to-right, top-to-bottom format and changes constantly to keep a synchronized display of the text with the audio. Lastly, interactive control buttons are represented using familiar iconic representations that are both, visual and textual to minimize confusion.

5.4.2 NON-VISIBLE FUNCTIONS

There are other functions working in conjunction with the interface that are not so visible, and yet equally important to achieve the seamless experience we required with this application. I employed the accessibility behavior library built in Director MX to enable the interface to communicate with the user through audio using the text-to-speech (TTS) engine that resides in the user's computer system. The accessibility library also provided the behaviors that worked alongside the TTS to provide synchronized text captioning.

Another important component is the server communication script developed using the Lingo programming language available in Director. This script lives inside the shockwave movie and it is used when the movie needs to retrieve dynamic data from the LSE database and output the data through the coach's voice. For example, the agent can query the LSE database to retrieve the user's name, and address him/her directly by his/her first name using TTS. With this server communication script, the agent can query any previously recorded session information and speak it back to the user. Thus, we have the

benefit of being able to talk in real time with the database to retrieve dynamic information and the potential to deliver highly personalized and usable information to the user. It may be useful to go over this process to understand what takes place when the movie requests information from the database and how the agent communicates it to the user (see Figure 17).

The Process:

- (1) The shockwave movie is loaded on the web page, if the movie requires data from the database, then the server script in the movie is activated.
- (2) The script triggers a data query to the LSE database.
- (3) Once the data is found, it is returned back to the script, and inserted in a text file.
- (4) The text file is send to the TTS engine in the user's system.
- (5) The TTS engine analyses the text to generate the audio, and
- (6) Sends the analyzed data back to the movie script,
- (7) The data is used to synch the text with the audio.
- (8) and (9) The audio and synched-text are output simultaneously to the computer screen and the user's speakers/headphones.

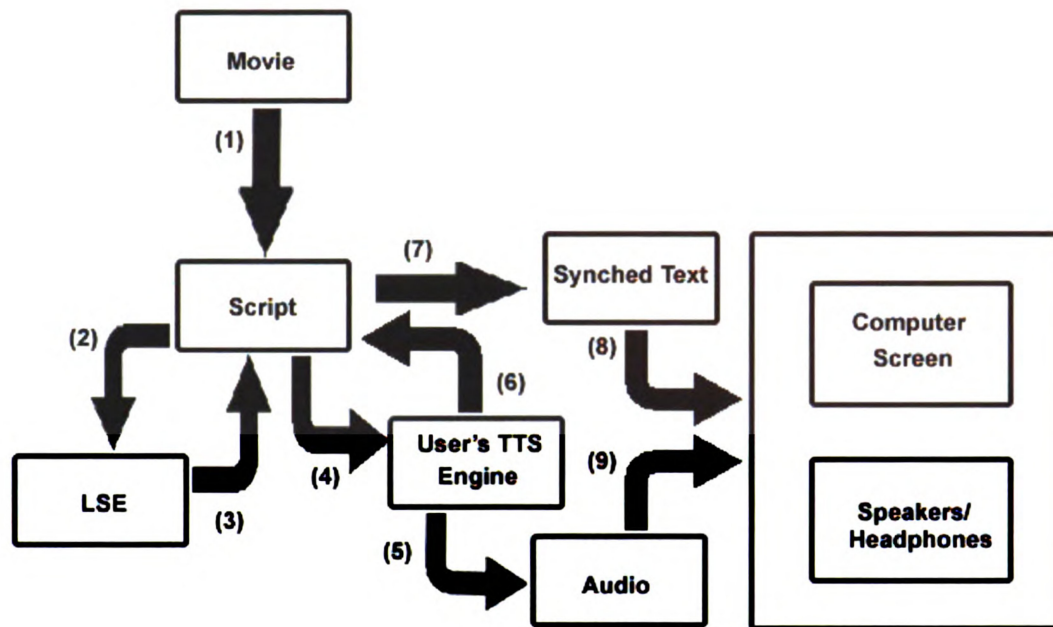


Figure 17: Shockwave movie, backend system, and TTS engine integration

Please refer to APPENDIX A: DIRECTOR SERVER COMMUNICATION SCRIPT for more details.

5.5 DELIVERY OF LEARN TO QUIT

The Learn to Quit application is completely web-based and can be accessed in Macintosh and Windows platforms through the most popular web browsers (e.g., Internet Explorer, Netscape Navigator). Because of its web format, it has the potential to reach an unlimited amount of target users who have access to an Internet connection, be it broadband or dial-up. All the agent movies were exported into multimedia web deliverable shockwave movies integrated with the Speech Xtra. The Speech Xtra adds special commands to the Lingo programming to enable Director's text-to-speech capability. The Xtra is from

Macromedia, and it is considered a trusted source. Therefore, when a user encounters a shockwave movie that employs this Xtra, the Xtra is downloaded automatically and the user does not have to interact with any dialogs boxes for it to be installed. The Xtra is approximately 45K on Windows and 35K on the Macintosh.

CHAPTER 6

RESEARCH METHODS AND FINDINGS

6.1 METHODS

The main objective of this study was to look at the reactions on a small number of users to different aspects of electronic agents acting as social actors to potentially promote health and make smoking cessation services available to anyone, anytime, and anywhere using web-based technology. Specifically, this research attempts to measure the users' reactions in relation to agent credibility, agent helpfulness, and general user rating of agent as a coach. Data was also analyzed to track participation of subjects in sessions and retention over the intervention period.

6.1.1 AGENT (COACH) DEVELOPMENT

The electronic agents for the study were designed as two supportive coaches and differentiated by their images (stern vs. friendly) and personalities (authoritative vs. kindhearted). The fundamental nature of the agents' personalities was communicated through the use of voice, text and images.

A text-to-speech engine already installed in the user's computer system was used to process and generate speech by sending a text file to the engine; this helped to make loading of the application faster. The voice the users heard during the delivery of the agent application depended on the voices they currently

had installed in their computer systems. To try to compensate for not knowing what type of voices and other multimedia components were installed and selected by default in the user's system; a step-by-step system setup webpage was created and participants were asked to visit the page before the start of the study.

The setup page showed participants the minimum requirements for Windows and Mac systems to run the application and provided instructions on how to download the Macromedia Shockwave player and how to activate the speech engine in their system to select a "male" voice to use with the agent.

The images of the coaches were portrayed as males in their mid to late 20's. The stern coach wore a suit and looked more formal while the friendly coach wore more casual clothing, and looked more comfortable.

6.1.2 SAMPLE SIZE, SUBJECTS AND RECRUITMENT

The sample size used for the research of this production thesis was set at 25. This number was considered sufficient to do initial testing of the online agent and find out the reactions to the look, feel and sound of the online coach. A total of 20 participants were recruited for the study. 20 was a large enough number to seriously test the system and to gain a general sense of what to retain before launching a larger scale online intervention. Participants were males and female MSU students 18 years of age and older who smoked and were thinking of

quitting smoking. They were recruited through posters posted around the MSU campus in student's dorms, lecture halls and popular spots like cafes and dining halls. Subjects were offered a \$10 gift certificate to download online music upon completion of the study (Please refer to APPENDIX B: RECRUITMENT POSTER for more details). Interested participants were asked to contact the project coordinator whom in return replied back with detailed information about participation in the study.

6.1.3 INSTRUMENT

The intervention/instrument is an adaptation of the over-the-phone quit smoking protocol (see Wadland et al., 1999). Similar questions and answers were used, but instead of being delivered by a human, over-the-phone coach, they were delivered by a computer "coach" represented by a photograph and speech to text synthesis. Rather than speaking their answers into the phone, participants responded directly over the Internet.

There were three 20-25 minute sessions scheduled over a period of two weeks (November 16 - 26, 2004). Participants were asked (via email to the entire group of volunteers, with their emails in the BCC area to maintain confidentiality) to log in for the first session starting on November 16, 2004. The schedule of the quit smoking date and two remaining sessions depended on the date the participant completed the first session. This gave us the flexibility of having multiple quit dates and session dates for participants over the two week period so

that not everyone was tied to the same schedule. Thus, sessions were schedule as follows:

- Quit smoking date was set 4 days after the first session
- Second session was scheduled 1 day after the quit date
- Third session was scheduled 3 days after the quit date

Participants were randomly assigned to work with one of the two agents. At the scheduled date, participants logged in to participate in the first session in a specified online site and were presented with an online consent form (please find the entire consent form in APPENDIX C: ELECTRONIC QUIT SMOKING COACH CONSENT FORM).

Participants were given two sets of questions for each session. One set was related to the quit smoking intervention questions, part of the protocol (I'd Rather Cope than Smoke) program used by telephone "coaches" to help people who have decided to quit smoking actually do so. In addition to that interaction, there were additional posttest questions at the end of the sessions which asked mostly close-ended questions about reactions to the coach.

The questions related to the reactions of users when interacting with the coach was the focus of the study, and not the answers to the questions about smoking behaviors (see APPENDIX D: SURVEY QUESTIONS- COACH

COMPONENT). Both sets of questions were delivered through the internet-based survey.

6.1.4 SURVEY ADMINISTRATION

The survey was administered using the anonymous longitudinal survey engine (LSE) jointly developed by the Michigan State University Department of Epidemiology and the Communication Technology Laboratory.

The LSE has a variety of possible settings. The setting used for this project required subjects to log on the first time using a coupon number. When they log on, they were asked to auto generate an ID and password so that there was no recorded association of online identities with the real person name. No one but they knew their ID and password. They used the same ID and password to log in for subsequent sessions.

In order to control access to the survey, volunteers were asked to email the study coordinator indicating their interest in participating. The email addresses were collected by the study coordinator who shared them with an administrative assistant in Epidemiology who had no access to LSE. This person composed and sent emails to each participant, informing them of their personal coupon number and the URL to log on (see APPENDIX E: LEARN TO QUIT OFFICIAL LETTER OF INVITATION for details). The administrative assistant then proceeded to delete the emails after sending them, so that neither she nor

anyone else associated with the study knew which subjects were associated with each coupon number.

When participants accessed the website for the first time and created their ID and password; they were presented with the online consent form. Once they read and agreed to the information, they clicked on a NEXT button and began the first session.

The study coordinator sent email reminders to the group of participants as the schedule dates for the two follow-up sessions approached (with their emails in BCC mode so that identities were not shared with other participants). The coordinator could not associate which subjects from the participant group had actually completed the first wave/second wave of the intervention, because each participant had logged in with a coupon number and created their own online identity. Thus, the anonymous nature of the study was kept intact at all times. The emails were brief and include the URL of the site to remind users where to login as well as link to the online session calendar. This same process occurred to remind them of the third (and final) intervention.

6.2 FINDINGS

6.2.1 GROUP PARTICIPATION AND RETENTION

Out of the group of 20 participants who had expressed interest and who were sent coupons to participate in the study, only 8 people actually activated the coupons and logged in to the site to participate in the first session. Of the 8 participants, 4 were in the friendly coach group, and 4 were in the stern coach group. Participation and retention of subjects across all three sessions were analyzed and revealed the following (see Table 1): In session 1, 3 out of 4 participants interacting with the stern coach completed the session. In the friendly coach group, 2 out of 4 people completed session 1.

Table 1

Overall Subject Participation between Sessions and Subject Retention per Session

Coach	Session #	# of people who logged in	Participation drop between session	# of people who completed the session	Participant retention per session
Stern	1	4	n/a	3	75%
	2	2	33.4%	1	50%
	3	2	0%	0	0%
Friendly	1	4	n/a	2	50%
	2	1	50%	1	100%
	3	1	0%	0	0%

Source: Web Server data books for Learn to Quit Session 1, 2, 3.

In session 2, only 2 individuals out of the 3 who had completed session 1 logged back to interact with the stern coach, registering a 33.4% decline in participation; and of those 2 participants only 1 completed the session. This represents 50% participant retention for the stern coach in session 2. The friendly coach also experienced a decline in participation (50%) when only 1 out of 2 participants logged back and completed the session.

In Session 3, again 2 participants logged in (no decline in participation) to interact with the stern coach; but neither one completed the session. The friendly coach also registered 1 participant logged in (no decline in participation); but did not complete the session. Thus, there was no decline in the number of people participating from session 2 to session 3 for both coaches; but we had 0% participant retention in the last session.

At the end of the study, it was found that out of the original 8 people who activated the coupons, only 6 completed at least 1 out of 3 sessions. The great majority (83.3%) of those who completed the sessions did so in session 1 and the rest (16.7%) were in session 2. The demographic data for participants was obtained during session 1 and combined for the two groups (stern coach and friendly coach). Data was obtainable for only 5 out of the 6 participants (1 participant did not answer any questions in session 1; but completed session 2). This data revealed the following about the participants: the mean age was 21 years old; 80% were female, 40% were Asian, 20% Caucasian; 20% Indian; and

20% African American. Most participants (80%) were in their Senior year at school (see Table 2).

Table 2

Combined Group Demographics for Stern and Friendly Coach

Demographics	Stern and Friendly Coach Group N= 5
Age, years	21.2
Gender	80% female 20% male
Ethnicity	
Caucasian	20%
African American	20%
Asian	40%
Indian	20%
College level	80% senior 20% junior

Source: Web Server data book for Learn to Quit Session 1

6.2.2 COACH EVALUATION

The data obtained for the coach evaluation component of the intervention sessions was only obtained for the participants who completed the sessions.

Thus, we find that session 1 yielded the most coach evaluations (n= 5); session 2 yielded one evaluation per coach, and session 3 yielded no coach evaluation.

Thus, session 1 data was used to perform the analysis of user reactions to the agents. Analysis of session 1 data includes the overall means and comparisons of mean between stern and friendly coaches (see Table 3).

Participants responses were collected using Likert Scale questions in the form of numerical scales (i.e., 1, 2, 3, 4, 5) where 1= strongly agree; 2= Agree; 3=neutral; 4=disagree; and 5= strongly disagree, and in the form of Semantic Differential questions to measure directionality of reactions. For example, one of the questions rates the coach in this scale as follows:

The coach was:

Friendly ___ ___ ___ ___ ___ ___ ___ Unfriendly
 3 2 1 0 1 2 3

In this scale usually the position marked 0 is neutral, the 1's positions are labeled "slightly," the 2's positions "quite," and the 3's positions "extremely".

Table 3**Overall Means and Comparisons of Means between Stern and Friendly Coaches**

Questions	Scales L= Likert (1 to 5) SD=Semantic Differential (1 to 7)	Overall Means	Means Comparison between Stern and Friendly coach
Credibility			
1. The coach was knowledgeable about quitting smoking	L	1.80	Friendly Coach: 2.5 Stern Coach: 1.33
2. The coach was intelligent.	L	2.20	Friendly Coach: 3.0 Stern Coach: 1.67
3. The coach knows a lot about ways to quit smoking.	L	2.40	Friendly Coach: 2.50 Stern Coach: 2.33
4. The coach understands what is hard about quitting smoking.	L	2.00	Friendly Coach: 2.5 Stern Coach: 1.67

5. The coach was helpful.	L	2.20	Friendly Coach : 3.0 Stern Coach: 1.67
6. The coach knows how to ask questions that get me to think about my behaviors	L	2.20	Friendly Coach: 3.50 Stern Coach: 1.33
7. The coach improved my knowledge	L	2.40	Friendly Coach: 3.50 Stern Coach: 1.67
8. The coach helped me to concentrate on the information	L	1.80	Friendly Coach: 2.0 Stern Coach: 1.67
9. The coach presented information effectively	L	1.80	Friendly Coach: 2.5 Stern Coach: 1.33
Coach Attitude			
10. The coach cares whether I succeed in quitting smoking.	L	2.20	Friendly Coach: 3.50 Stern Coach: 1.33

11. The coach wants me to stop smoking.	L	2.00	Friendly Coach: 3.0 Stern Coach: 1.33
12. The coach cares about me.	L	2.20	Friendly Coach: 3.0 Stern Coach: 1.67
13. The coach will be disappointed if I do not quit smoking.	L	2.60	Friendly Coach: 2.50 Stern Coach: 2.67
14. I would prefer to have a coach who is strong and firm with me.	L	2.60	Friendly Coach: 2.0 Stern Coach: 3.0
15. I would prefer to have a coach who is gentle and kind.	L	1.60	Friendly Coach: 1.50 Stern Coach: 1.67
Rate the Coach			
16. I liked to learn from the coach.	L	3.00	Friendly Coach: 3.50 Stern Coach: 2.67

17. I would prefer to have a live human coach text chatting with me over the internet.	L	3.20	Friendly Coach: 1.50 Stern Coach: 4.33
18. I would prefer to work in-person with a human coach	L	2.40	Friendly Coach: 2.0 Stern Coach: 2.67
19. I would prefer to have a computer coach.	L	2.60	Friendly Coach: 4.0 Stern Coach: 1.67
20. The coach was Friendly or Unfriendly	SD	2.00	Friendly Coach: 3.0 Stern Coach: 1.33
21. The coach was Cold or Warm	SD	5.40	Friendly Coach: 6.0 Stern Coach: 5.0
22. The coach was Impersonal or Personal	SD	3.80	Friendly Coach: 2.5 Stern Coach: 4.67

23. The coach was Sociable or Unsociable	SD	3.20	Friendly Coach: 4.50 Stern Coach: 2.33
24. The coach was Stern or Forgiving	SD	5.20	Friendly Coach: 5.0 Stern Coach: 5.33
25. The coach was Gentle or Strong	SD	3.00	Friendly Coach: 3.0 Stern Coach: 3.0
26. The coach was Sensitive or Insensitive	SD	2.80	Friendly Coach: 4.0 Stern Coach: 2.0
27. The coach was Understanding or Not understanding	SD	2.40	Friendly Coach: 3.0 Stern Coach: 2.0
28. The coach was Harsh or Gentle	SD	5.20	Friendly Coach: 5.0 Stern Coach: 5.33

29. The coach was Strong or Weak	SD	3.80	Friendly Coach: 4.0 Stern Coach: 3.67
About the Coach's Voice			
30. The coach's voice was Easy to understand or Hard to understand	SD	2.40	Friendly Coach: 4.0 Stern Coach: 1.33
31. The coach's voice was Pleasant or Unpleasant	SD	4.40	Friendly Coach: 6.0 Stern Coach: 3.33
32. The coach's voice was Human or Inhuman	SD	4.20	Friendly Coach: 5.50 Stern Coach: 3.33
33. The coach's voice was Hard to Listen to or easy to Listen to	SD	5.20	Friendly Coach: 3.50 Stern Coach: 6.33
34. The coach's voice was Nice or Not nice	SD	3.40	Friendly Coach: 4.0 Stern Coach: 3.0

35. Hearing the coach speak made him seem less human.	L	4.00	Friendly Coach: 2.50 Stern Coach: 5.0
36. I would strongly prefer a computer coach with a natural human voice.	L	2.60	Friendly Coach: 1.50 Stern Coach: 3.33
37. I would prefer a computer coach that is only text on screen, no pictures or voice.	L	4.40	Friendly Coach: 3.50 Stern Coach: 5.0
38. I like seeing pictures of the coach.	L	2.60	Friendly Coach: 3.0 Stern Coach: 2.33
39. I would prefer to just read and not hear the coach.	L	3.80	Friendly Coach: 3.0 Stern Coach: 4.33
40. I would prefer to hear the coach and NOT read what he is saying.	L	4.00	Friendly Coach: 2.50 Stern Coach: 5.0

41. I rated this coach like I would rate a real-life coach.	L	3.00	Friendly Coach: 3.50 Stern Coach: 2.67
42. I would be likely to use a computer coach to support and guide me in other health related interventions.	L	2.60	Friendly Coach: 3.5 Stern Coach: 2.0
43. The internet connection for the computer I am using is:	1. Campus Internet 2. DSL 3. Cable Modem 4. Dial-up Modem	2.80	Friendly Coach: 2.5 Stern Coach: 3.0
44. The computer I am using is located:	1. At Home 2. At work 3. In my dorm room 4. In a public lab 5. Somewhere else.	1.00	Friendly Coach: 1 Stern Coach: 1.0

Source: Web Server data book for Learn to Quit Session 1

6.2.3 DATA INTERPRETATION

With regards to the coach credibility it was found that participants in both groups agreed that the coaches were knowledgeable about the quitting smoking process and that they knew a lot about ways to quit smoking.

The stern coach group perceived their coach to be more intelligent and with a good understanding about the hardships of quitting smoking that did the participants in the friendly coach group, who were mostly neutral about the coach intelligence (rated 3 on scale of 5) and did not quite agreed about the coach's understanding of how hard it was to quit smoking with one person rating the coach 1 (strongly agree) and the other person gave a rating of 4 (disagree). Overall the stern coach received more positive reactions as been more credible than did the friendly coach.

Regarding agent helpfulness, the overall mean between the two groups were mostly on the positive side for both coaches. Both groups agreed that the presence of the coach had helped them to concentrate in the information.

However, a few instances are found in which the friendly agent participants had some marked differences with respect to the friendly coach helpfulness. Specifically, a 50/50 split evaluation is appreciated when both participants in the friendly group gave their rating for the agent being helpful, one person gave a rating of 1(strongly agree) and the other rated the agent 5

(strongly disagree). A similar result was found when the friendly coach participants rated whether the agent presented information effectively or not. One person strongly agreed and the other disagreed. As a result, the friendly coach received a neutral mean (3.5) reaction to whether or not the participants' knowledge had improved as a result of their experience with the coach.

On the other hand, the stern coach group participants strongly and consistently believed that the stern coach knew how to ask questions to get them thinking about their smoking behaviors. They strongly believed that the coach helped them to concentrate in the information and that their coach had presented the information more effectively. They found the coach to be extremely helpful and believed that their knowledge about quitting smoking had improved as a result of the interaction with the stern coach.

With regards to how the participants perceived the attitudes of the coaches, it was found that among the two coaches, the stern coach was perceived to be more caring about the participants and their success with quitting smoking than the friendly coach. Although, both groups did feel that their coaches would be disappointed if they did not quit smoking; the stern coach was found to have expressed more strongly that he wanted the participants to stop smoking.

When asked what type of attributes they would prefer to have in a coach, the overall response was that they would like a coach who was strong and firm, yet; gentle and kind at the same time. An interesting finding is that when we formulated the questions about having a stronger or firmer coach, or a gentle or kind coach; we intended them to be the opposite to each other but it turned out that people did not interpret them to be that way. Instead, participants believed that all these attributes (stronger, firm, gentle, kind) should be part of an agent's personality.

When participants were asked whether they liked learning from the coach, it was found that most of the stern coach group liked learning from the coach (group mean=2.67) compared to the participants in the friendly group coach (group mean=3.5) who were between neutral or disagreeing that they liked to learn from the coach.

Another question asked if they would prefer to work with a real human text-chatting with the participant over the Internet. The stern group participants said they would strongly prefer working with a computer coach than text-chatting with a live human coach over the Internet. Friendly coach participants, on the other hand, expressed that they would strongly prefer a live human text-chatting than the electronic coach. When given the option of working in person with a real-human coach, then we found some consistency among the two groups (overall mean=2.40) who said they would prefer working in person with a live

human coach than any of the other two choices (i.e., text-chatting or a computer coach).

Overall mean ratings suggested that the coaches were perceived as strong, slightly gentle, sensitive and understanding about the participants' smoking problems. However, there were some particular differences found among how the two groups evaluated their coaches. The stern coach group found their coach to be stronger, more sensitive, and more understanding of their problems that did the participants in the friendly group.

Both coaches were rated as being friendly; but of the two coaches, the stern coach was found to be quite friendly (mean= 1.33) on the semantic differential scale and the friendly coach slightly friendly (mean=3.0) on the same scale. Regarding to how warm the personalities of the coaches were, the friendly coach was rated as being quite warm and the stern coach was rated as having a slightly warm personality. Due to the stern nature of the stern coach, one would have expected that his personality would be found to be less warm than the friendly coach.

When asked about how sociable the coaches were, it was found that an overall mean of 3.20 on a semantic differential scale suggested that participants thought the coaches were slightly sociable. Comparing responses within the groups we find a 50/50 split in the responses of the friendly coach group with one

participant rating the coach 3 (slightly sociable) and the other person rated the coach 6 (quite unsociable). On the other hand, the stern coach was found to be quite sociable with a group mean of 2.33.

With regards to the reactions to the computer generated coach's voice, the stern coach group strongly disagreed that hearing the coach speak made him less human. In fact, they felt they coach seemed more human because it had a voice that spoke to them. On the other hand, participants of the friendly group coach believed that hearing the coach speak made him less human.

Overall the stern coach voice was found to be nice, pleasant and easy to listen to, while the friendly group only found the coach voice slightly easy to listen to, and rated the voice of the coach as being quite unpleasant. It is possible that the differences about how pleasant or unpleasant the voices depended upon the type of voice engine and voices installed in the user's computer system.

When asked if they would prefer a computer coach with a natural human voice, participants of the friendly group expressed that they would strongly prefer working with a coach with a natural human voice. On the other hand, most participants in the stern group have mixed reactions to having a more natural human voice, 2 out of 3 disagreed to preferring working with a coach with a more natural human voice.

With regard to their preferences for delivering the web-based coaches, most participants in both groups disagreed to having a coach that was just text on the screen. Most of the subjects liked seeing the coach pictures on the screen. However, a high standard deviation in both groups (friendly=2.8, stern=2.3) with regard to pictures on the screen suggested a lot of variation in the answers. Looking at the data in more detail, it was found that two persons in the stern group said they strongly liked seeing pictures of the coach and one person strongly disagreed. In the friendly group, a similar rating was found with one person saying that they strongly agreed and one person strongly disagreed.

The overall mean between the groups (4.0) on the Likert scale suggested that most participants would also disagree to having a coach with voice and not text. In particular, the stern group coach strongly disagreed to having only voice and no text. However, at least one participant in the friendly group seemed to prefer the idea of just being able to hear and not read any text. What we can gather from the overall reactions of most participants is that they were more inclined to preserving the three coach attributes that were presented to them during the intervention (i.e., text, voice and images).

When asked whether they would rate their coach as a real-life coach; the stern coach group agreed that their coach felt more like a real-life coach (mean=2.67) compared to the friendly coach group who remained mostly neutral in their rating (mean=3.5).

Most participants expressed they were likely to use a computer coach in other health interventions (overall mean= 2.60). Of the two groups the stern coach group was more likely to want to use a coach for other health intervention (mean=2.0) than the friendly coach group (mean=3.5).

Data about the location and the type of connection subjects used to participate in the study revealed that most participants accessed the intervention from their homes using either a DSL or a cable modem connection.

CHAPTER 7

CONCLUSION

The results obtained in this pilot study provide some optimism about the potential and feasibility of employing online coaches to deliver online health-related interventions effectively. Although a very small sample of participant's data was collected in the study, the data does provide some small hints about the reactions of the users and suggest that online smoking cessation coaches can be perceived as credible, knowledgeable and helpful in providing information to help people quit smoking. Being perceived as credible and as having expertise on the subject matter (i.e., smoking cessation) is important, because it increases the persuasive power of the agent to motivate and guide the users who want to quit smoking to actually do so.

The results obtained suggest that overall people had more positive reactions to the stern coach personality than to the friendly coach personality. These preliminary results are indeed interesting; given the nature of the personality embodied in the friendly coach by the designer, it was expected that people would have more affinity towards a coach that was more easy-going, approachable and gentle. On the other hand, there have been empirical studies that seem to support what our results suggested. These studies found that sterner agents were also perceived as more credible and more engaging (e.g., Walker et al., 1994 and Baylor et al., 2003). Participants in this pilot study also

expressed that overall they would prefer a coach that was stronger and firmer and at the same time, gentle and kind.

This pilot study is not without its own limitations. A retrospective review of the design process of the intervention was useful to identify some key issues.

One problem was the development process of the online survey sessions using the current state of the Longitudinal Survey Engine (LSE) technology. Adapting the existing session of the “ I’d Rather Cope than Smoke” program to the LSE proved to be a very challenging task that required a great deal of time and effort.

Mainly, the design environment of the LSE did not quite lend itself to the authoring with ease of very large surveys as the ones needed to be developed in this application. For example, as the survey grows larger, the time that it takes to load the survey, to insert a question or perform any other action increases considerably; the end-result is that a lot of time is wasted for every action that is requested. The longest survey designed for this application was 245 pages long with a total of 288 questions. With a survey this long, adding a new question, a new page skip or linking a media asset took anywhere from 1 min to 2 min per item requested.

An alternative to speed up development of the survey session may be to explore with the LSE dividing the long session surveys into a series of smaller

sub-sections, and create a survey for each sub-section and then link them together. This way every sub-section survey would have a much smaller number of pages and questions; as a result, the loading time of surveys is considerably reduced and development speeds up. Another alternative is to explore the potential of using Macromedia Director to also author the development of the surveys items and combine this with the agent delivery all in one self-contained unit.

With respect to the participant's motivation and commitment to complete all three sessions of the intervention, I believed that the reward system may have had some influence on the final participation. A \$10 gift certificate for online music may not be the most motivating reward for our target audience (college students) if one takes into account that college students are very resourceful and that they may in fact use many other channels to obtain free online music.

Another issue that surfaced as a result of doing the design of the study anonymously during this stage of the pilot study is that we had no mechanism to meet or communicate directly with the people who participated in the study, other than sending emails to everyone in the group. Being able to conduct one-on-one interviews or focus group discussions with participants may have yielded more insights about what kind of problems they experienced with the web-based agents and the ability to collect more qualitative data on their reactions to the agents.

However, I was also able to see some of the benefits of doing the survey anonymously. It prompted me to think and tackled early on some of the design challenges when designing for anonymous participants and how to recruit and keep them interested in the program. For example, adjusting the content to the potential recruitment of participants (MSU students vs. other college students), typed and amount of incentives (online music vs. online food), timing for launching the study (holidays, mid terms, final exams), preparation and scheduling of sessions (7 days vs. 4 days for quit date) and creating and maintaining a flexible timeline for the participants so that not everyone was tied to the same schedule. All these design decisions were explored in detail in order to implement the web-based anonymous study.

These preliminary results suggest that further development and research of the online agents should be conducted taking into considerations the design revisions resulting from this pilot study. Particular attention should be paid to the recruitment and retention of a greater number of users for the longitudinal study. Additional research that builds upon the approach adopted in this pilot study would help to test the validity of the results found in this study and what these suggested about the user's reactions to the stern and friendly personality.

APPENDICES

APPENDIX A

DIRECTOR SERVER COMMUNICATION SCRIPT

--Parse Questions--

```
global query

on exitFrame me

    -- setup the server defaults
    theURL = member("hostname").text & "/data_interchange.php"

    -- parse the text
    txt = member("query text").text
    the itemDelimiter = "%"
    size = the number of items in txt

    -- build queries
    query = [:]
    repeat with i = 1 to size
        if i mod 2 = 0 then
            setaprop query, txt.item[i], []
        end if
    end repeat

    -- begin calls to LSE
    max = query.count()
    repeat with i = 1 to max
        theprop = getpropat(query,i)
        querystr = "%"&theprop&"%"
        postList =
["c_session":member("sessionid").text,"data_retrieval_reques
t::marker":querystr]
        netid = _movie.postnettext(theURL, postList) -- start
net connection and store netid
        setaprop query, theprop, [netid,false,EMPTY]
    end repeat

end
```

--Connect to Server--

```
global query

on exitframe me
    -- see if the net connection is done

    max = query.count()

    ready = true
    repeat with i = 1 to max
        netid = query[i][1]
```

```

    if netDone(netID) then
        query[i][2] = true
        query[i][3] = netTextResult(netID)  -- get CGI data
    else
        ready = false
    end if
end repeat

if ready then
    -- parse the text
    txt = member("query text").text
    the itemDelimiter = "%"
    size = the number of items in txt

    -- insert results
    newtxt = EMPTY
    repeat with i = 1 to size
        if i mod 2 = 0 then
            newtxt = newtxt & query.getaProp(txt.item[i])[3]
        else
            newtxt = newtxt & txt.item[i]
        end if
    end repeat
    member("output text").text = newtxt
    voiceSpeak(newtxt)
    go to "results"
else
    go to the frame
end if
end

```

APPENDIX B
RECRUITMENT POSTER

Get Expert Help to Quit Smoking and Earn

\$10 FREE MUSIC



PARTICIPANTS FOR QUIT SMOKING STUDY WANTED!



Comm Tech Lab
Michigan State University

We create innovative learning experiences
which elegantly integrate technology.

Who we would like to talk to?

We are looking for research subjects who are between the ages of **18-30**, are **MSU students**, and who currently **smoke** and are **thinking about quitting smoking**. You get a gift certificate for \$10 from the popular iTunes Music store for participating. If you know other people that fit the above criteria, please let them know about this too!

What is involved?

The main experiment involves accessing a **website anonymously from the comfort of your own residence hall, home or other computer of your choice** to interact with an online virtual coach, whose role is to support and guide you in the quitting process. You will be required to login to three scheduled sessions of approximately 20-25 minutes of your time. You are free to skip any question you are not comfortable answering, but to earn the free music download, you must log in and navigate through all three sessions.

If you are interested in this opportunity, please contact **A.S.A.P Geraud Plantegenest** at plantege@msu.edu to get more information about how to participate in the study.

The study starts on Tuesday November 16, so hurry up and sign up for your slot!

APPENDIX C

ELECTRONIC QUIT SMOKING COACH CONSENT FORM

Thank you for logging in to the Electronic Quit Smoking Coach Web Site!

The purpose of this first page is to explain your rights and protections as a participant in human subjects research. If, after reading this informed consent statement, you consent to participate, then please follow the instructions to begin the experience.

Instructions

You are being asked to participate in a prototype computerized version of a clinically tested quit smoking program designed to help persons cope with the challenges they encounter when quitting smoking. The program consists of three sessions arranged on different dates in order to maximize the help the coach can provide you when you need it the most. Please read this consent form thoroughly.

Project Description

The purpose of this thesis is to create an online environment for college students between the ages of 18 and 30 years old who currently smoke and want to learn how to quit and stay quit. The website makes use of a virtual agent whose role would be to serve as your support coach to guide and help you in the process of quitting smoking.

Survey Procedure

Participation in the survey is voluntary. This survey will be conducted only with your consent. You may stop participating in the survey at any time for any reasons. You don't have to answer any questions you do not want to. Each survey will take about 20-25 minutes to complete depending on the connection speed of your computer system and hardware.

The surveys are divided into two areas. The first area focus on your smoking history and habits, the information you provide is used by the coach to help assess your situation and to provide you with the necessary coping strategies you need to succeed in your attempt to quit smoking. The second area of the survey will focus on questions about the application itself. The information you provide in the second part of the survey is important to help us understand how effective this new form of web coaching for health behavior will be.

Benefits

The study will give you an opportunity to learn coping strategies that can be applied when you are trying to quit smoking. The study also gives you a great opportunity to provide your personal input to influence the content and presentation of future online coach applications. After gathering and analyzing

the data we collect, we hope to have a better understanding of how to design more effective online coaches that can help you and others to quit smoking.

Award

At the end of the third session with the online coach, you will receive an online link to redeem an anonymous \$10 online gift certificate enabling you to download any songs you wish from the popular site iTunes Music Store. You are free to skip any question you are not comfortable answering, but to earn the free music download, you must log in and navigate through all three sessions.

Foreseeable Risks or Discomforts

Participations in this research poses minimal risks to you. It is possible that you may feel awkward about expressing your feelings and opinions and experience fatigue during the sessions with the online coach. The length of time you spend on answers is entirely up to you. If you feel uncomfortable at any time during the session, you may immediately withdraw by closing you browser window with no questions asked.

Confidentiality and Anonymity

Your privacy will be protected to the maximum extent allowable by law. Your name and online identity will not be associated with any reporting of the findings.

Your responses to the online questions will be completely anonymous. You have logged in using an assigned coupon number. The administrative assistant who assigned the coupon number does not have access to the data. And she destroyed the invitation emails as soon as she sent them. So, no record exists linking your answers to you.

During data collection, your answers will be stored on a password protected computer in an alarm-protected laboratory, until the data is downloaded by the investigators and stored on their password-protected computer for analysis.

If you have any questions or concerns regarding your rights as a study participant, or are dissatisfied at any time with any aspect of this study, you may contact- anonymously, if you wish - Peter Vasilenko, P.h.D., Chair of the University Committee on Research Involving Human Subjects (UCRIHS) by phone: (517) 355-2180, fax: (517) 432-4503, email: ucrihs@msu.edu, or regular mail: 202 Olds Hall, East Lansing, MI 48824.

Contact Persons

If you have any questions about this study, please contact the investigators.

Principal Investigator Carrie Heeter 253 Communication Arts Building Michigan State University East Lansing, MI 48824 517-353-5497 heeter@msu.edu	Secondary Investigator Geraud Plantegenest 253 Communication Arts Building Michigan State University East Lansing, MI 48824 517-353-5497 plantege@msu.edu
--	--

Statement of Consent

You indicate your voluntary agreement by clicking on the NEXT button below to open the survey.

APPENDIX D

SURVEY QUESTIONS- COACH COMPONENT

(These questions appeared at the end of each session, immediately following the quit smoking questions with the electronic coach.)

Congratulations! You have just completed the (first/second/final) quit smoking interview.

Please, the researchers would greatly appreciate a little more of your time, to give us feedback about the electronic coach. These questions are text only and you can answer at your own pace. Thank you!!

1. The coach was knowledgeable about quitting smoking.
Strongly Agree 1 2 3 4 5 Strongly disagree
2. The coach was intelligent.
Strongly Agree 1 2 3 4 5 Strongly disagree
3. The coach knows a lot about ways to quit smoking.
Strongly Agree 1 2 3 4 5 Strongly disagree
4. The coach understands what is hard about quitting smoking.
Strongly Agree 1 2 3 4 5 Strongly disagree
5. The coach was helpful.
Strongly Agree 1 2 3 4 5 Strongly disagree
6. The coach knows how to ask questions that get me to think about my behaviors.
Strongly Agree 1 2 3 4 5 Strongly disagree
7. The coach improved my knowledge.
Strongly Agree 1 2 3 4 5 Strongly disagree
8. The coach helped me to concentrate on the information.
Strongly Agree 1 2 3 4 5 Strongly disagree
9. The coach presented information effectively.
Strongly Agree 1 2 3 4 5 Strongly disagree
10. The coach cares whether I succeed in quitting smoking.
Strongly Agree 1 2 3 4 5 Strongly disagree

11. The coach wants me to stop smoking.
Strongly Agree 1 2 3 4 5 Strongly disagree
12. The coach cares about me.
Strongly Agree 1 2 3 4 5 Strongly disagree
13. The coach will be disappointed if I do not quit smoking.
Strongly Agree 1 2 3 4 5 Strongly disagree
14. I would prefer to have a coach who is strong and firm with me.
Strongly Agree 1 2 3 4 5 Strongly disagree
15. I would prefer to have a coach who is gentle and kind.
Strongly Agree 1 2 3 4 5 Strongly disagree
16. I liked to learn from the coach.
Strongly Agree 1 2 3 4 5 Strongly disagree
17. I would prefer to have a live human coach text chatting with me over the Internet.
Strongly Agree 1 2 3 4 5 Strongly disagree
18. I would prefer to work in-person with a human coach.
Strongly Agree 1 2 3 4 5 Strongly disagree
19. I would prefer to have a computer coach.
Strongly Agree 1 2 3 4 5 Strongly disagree
20. The coach was:
- | | | | |
|-----|---------------|---------------|-------------------|
| 21. | Friendly | 1 2 3 4 5 6 7 | Unfriendly |
| 22. | Cold | 1 2 3 4 5 6 7 | Warm |
| 23. | Impersonal | 1 2 3 4 5 6 7 | Personal |
| 24. | Sociable | 1 2 3 4 5 6 7 | Unsociable |
| 25. | Stern | 1 2 3 4 5 6 7 | Forgiving |
| 26. | Gentle | 1 2 3 4 5 6 7 | Strong |
| 27. | Sensitive | 1 2 3 4 5 6 7 | Insensitive |
| 28. | Understanding | 1 2 3 4 5 6 7 | Not understanding |
| 29. | Harsh | 1 2 3 4 5 6 7 | Gentle |
| 30. | Strong | 1 2 3 4 5 6 7 | Weak |
31. I rated this coach like I would rate a real-life coach.
Strongly Agree 1 2 3 4 5 Strongly disagree
32. I would be likely to use a computer coach to support and guide me in other health related interventions.
Strongly Agree 1 2 3 4 5 Strongly disagree

33. The coach's voice was:

- | | | | | | | | | | |
|-----|--------------------|---|---|---|---|---|---|---|--------------------|
| 34. | Easy to understand | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Hard to understand |
| 35. | Pleasant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Unpleasant |
| 36. | Human | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Inhuman |
| 37. | Hard to listen to | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Easy to Listen to |
| 38. | Nice | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Not nice |

39. Hearing the coach speak made him seem less human.

Strongly Agree 1 2 3 4 5 Strongly disagree

40. I would strongly prefer a computer coach with a natural human voice.

Strongly Agree 1 2 3 4 5 Strongly disagree

41. I would prefer a computer coach that is only text on screen, no pictures or voice.

Strongly Agree 1 2 3 4 5 Strongly disagree

42. I like seeing pictures of the coach.

Strongly Agree 1 2 3 4 5 Strongly disagree

43. I would prefer to just read and not hear the coach.

Strongly Agree 1 2 3 4 5 Strongly disagree

44. I would prefer to hear the coach and NOT read what he is saying.

Strongly Agree 1 2 3 4 5 Strongly disagree

45. The Internet connection for the computer I am using is:

Campus Ethernet DSL Cable Modem Dialup Modem

46. The computer I am using is located

At Home At Work In a Public Lab Somewhere Else

Thanks again so very much for your valuable opinions. This completes Session (1/2/final) of the Electronic Quit Smoking Program. Have a great day!

APPENDIX E

LEARN TO QUIT OFFICIAL LETTER OF INVITATION

Dear Participants,

Today is the date that we require you to Login to take part of the first session with the online Learn to Quit Smoking Coach. We are very excited and thrilled to have you as a participant in this study and we all look forward to hearing what you have to say about your experience with the online coach. Your input is very important to us.

It is also very important that you have completed the setup for your computer system prior to going to the website to interact with the online coach. So, please if you have not done this yet, visit the following url: http://www.msu.edu/~plantege/system_setup/setup.htm and follow the simple instructions to get your system ready.

To protect your anonymity during the study, this email is being sent to you by an administrative assistant from the Department of Epidemiology at MSU. This person has no access to any other data that may identify you as you start the study. This person will destroy the copy of the email being sent to you, immediately after she has sent it.

Likewise, the researchers of this study will have no access to any of these emails being sent, nor the coupon numbers being assigned to any participants. Therefore, there is not possible way to link your real identity to your online identity; this guarantees that your participation in this study remains completely anonymous at all times, as required and approved by university guidelines.

If at any given time, you experience any technical problems either during setup of your system or during the actual study, please contact Geraud Plantegenest at (517) 974 45 67 as soon as possible. He will be standing-by to help participants.

Tell him you are a participant in the ANONYMOUS Quit Smoking Study and you want help setting up your computer or explain the problem you encountered. Please DO NOT give him your name or identifying information. We want to help you set up or troubleshoot your computer but we do not want to know your identity.

Here are some important points about using the technology that you should keep in mind when participating in the study and while interacting with the online coach and answering survey questions.

- Once you leave a page by clicking the 'next' link that appears on every page of the site, you will not be able to come back to the same page

again. Therefore, please do not attempt to use the 'Back Button' of your browser. Doing so, will regretfully take you out of the online site and you will be required to log back in to the site and start from the beginning.

To Login to the Learn to Quit site please visit the following url:

Your coupon number is:

Once you are on the site, click on the link '**I have a coupon number**' and follow the instructions to create your unique and anonymous username.

Thank you and we look forward to having you at the Learn to Quit site today!

The Learn to Quit Team.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Baylor, A. L. (2001). Permutations of control: Cognitive considerations for agent-based learning environments. *Journal of Interactive Learning Research*, 12(4), 403-425.
- Baylor, A. L. (2004). *Designing pedagogical agents to address diversity in learning*. Paper presented at the International Conference of the Learning Sciences, Los Angeles, California.
- Baylor, A. L., & Kim, Y. (2003). *The Role of Gender and Ethnicity in Pedagogical Agent Perception*. Paper presented at the E-Learn (World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education), Phoenix, Arizona.
- 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*. Paper presented at the E-Learn (World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education), Phoenix, Arizona.
- Boyer, K. K., Olson, J. R., & Jackson, E. C. (2001, July). Electronic Surveys: Advantages and Disadvantages Over Traditional Print Surveys. *Decision Line*, 32, 4-7.
- Cooper, A. (1999). *The Inmates Are Running the Asylum: Why High-Tech Products Drive Us Crazy and How to Restore the Sanity*. Sams.
- David, P., Lu, T., Kline, S., & Cai, L. (2002, October 9-11). *Social Effects of an Anthropomorphic Help Agent: The Fairness of Man and Machine*. Paper presented at the Fifth International Workshop, Presence, Universidade Fernando Pessoa, Porto, Portugal.
- Dehn, D. M., & Mulken, S. V. (2000). The impact of animated interface agents: a review of empirical research. *International Journal Human-Computer Studies*, 52, 1-22.
- Detenber, B. H., & Reeves, B. (1996). A Bio-Informational Theory of Emotion: Motion and Image Size Effects on Viewers. *Journal of Communication*, 46(3), 66-84.

- Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What Is Beautiful Is Good, but ...: A Meta-Analytic Review of Research on the Physical Attractiveness Stereotype. *Psychological Bulletin*, 110, 109-128.
- Fogg, B. J. (2003). *Persuasive Technology: Using Computers to Change What We Think and Do*. Amsterdam: Morgan Kaufmann.
- 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.
- Kim, Y., & Baylor, A. L. (In review). A Social Cognitive Framework for Designing Pedagogical Agents as Learning Companions. *Educational Technology Research and Development*.
- Kim, Y., Baylor, A. L., & Reed, G. (2003). *The Impact of Image and Voice with Pedagogical Agents*. Paper presented at the E-Learn (World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education), Phoenix, Arizona.
- Lester, J. C., Converse, S. A., Stone, B. A., & Kahler, S. E. (1997). *Animated Pedagogical Agents and Problem-Solving Effectiveness: A Large-Scale Empirical Evaluation*. Paper presented at the Eighth World Conference on Artificial Intelligence in Education.
- Martin, W., & Leckenby, J. D. (1999). *New Media Development: The Internet vs. Traditional Media* (Presentation). Austin: The University of Texas at Austin.
- Mayer, R. E. (2001). *Multimedia Learning*. Cambridge ; New York: Cambridge University Press.
- Moon, Y. (1998). Impression management in computer-based interviews. *Public Opinion Quarterly*, 62(4), 610.
- Patsuris, P. (2004). *Look Who's Talking*, from http://www.forbes.com/2004/01/06/cx_pp_0105pulse_print.html
- Ravaja, N. (2004). Effects of Image Motion on a Small Screen on Emotion, Attention, and Memory: Moving-Face vs. Static-Face Newscaster. *Journal of Broadcasting & Electronic Media*, 48(1), 108-133.
- Reeves, B., & Nass, C. (1996). *The Media Equation: How people treat computers, television, and new media like real people and places*. Cambridge: Cambridge University Press.

- Shneiderman, B., & Maes, P. (1997). Direct Manipulation vs. Interface Agents: Excerpts from debates at IUI and CHI 97. *Interactions*, 42-61.
- Stofellmayr, B., Wadland, W. C., & Pan, W. (2003). An examination of the process of relapse prevention therapy designed to aid smoking cessation. *Addictive Behaviors*, 28, 1351-1358.
- United States Department of Health and Human Services (undated) *Five Major Steps to Intervention (The "5A's")*., retrieved January 25, 2005.
<http://www.surgeongeneral.gov/tobacco/5steps.htm>
- Wadland, W. C., Stofellmayr, B., Berger, E., Crombach, A., & Ives, K. (1999). Enhancing Smoking Cessation Rates in Primary Care. *The Journal of Family Practice*, 48(9), 711-718.
- Wadland, W. C., Stofellmayr, B., & Ives, K. (2001). Enhancing Smoking Cessation of Low-Income Smokers in Managed Care. *The Journal of Family Practice*, 50(2), 138-144.
- Walker, J. H., Sproull, L., & Subramani, R. (April 24-28, 1994). *Using a Human Face in an Interface*. Paper presented at the Human Factors in Computing Systems, Boston, Massachusetts.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03062 7172