Chapter X

Toward Socially Intelligent Interviewing Systems

Natalie K. Person, Sidney D'Mello, and Andrew Olney

Advances in technology are changing and personalizing the way humans interact with computers, and this is rapidly changing what survey researchers need to consider as they design the next generation of interviewing technologies. The nearly geometric growth rate of processing power, memory capacity, and data storage are allowing designers to endow software and hardware with capabilities and features that seemed impossible less than a decade ago. Among the innovations of greatest relevance for survey designers are intelligent systems: computers, robots, and software programs that are designed to exhibit some form of reasoning ability. The concept of an intelligent system does not necessarily imply that the system thinks independently; however, it does imply that the system has been programmed to respond in intelligent, adaptive ways to specific kinds of input.

In this chapter, we discuss how technology is already being used in interactive dialogue systems to optimize user input and to adapt to users in socially intelligent ways. Most of the methods and technologies that we will be discussing have been incorporated in systems other than automated survey systems; however, we believe it is only a matter of time before the social intelligence advances that have been made and implemented in other intelligent systems will be proposed for survey technologies—for better or worse. The extent to which what we discuss here ought to be extrapolated to survey systems depends crucially on the parallels between what these systems and potential interviewing systems do; we will return to this issue along the way and at the end, but there is sufficient overlap to make the comparisons valuable.

Thinking about automated interviewing systems with greater social intelligence is worthwhile given how increasingly important intelligent systems are becoming in our everyday experiences. Intelligent systems aid us in making travel arrangements, in finding phone numbers and addresses when we dial directory assistance, and in using our word processing applications on our personal computers. In some of our past work, we have been particularly interested in exploring how intelligent systems that manifest aspects of social intelligence affect user's behaviors and cognitive states. In particular, we have explored how systems that possess differing degrees of social agency affect language production when users are asked to disclose personal and emotionally-charged information. We have also examined how users' affective states can be measured in intelligent tutoring systems, which are similar to automated survey interviews in certain important ways: both are interactive tasks in which the system ("tutor" or "interviewer") drives the dialogues, presents prompts and cues, and must motivate users (learners or respondents) to continue the activity and provide thoughtful responses. It is plausible that systems that can adapt or mirror users' affective states (Cassell & Miller, Chapter XX in this volume) will lead to more positive user perceptions, greater trust between the user and the system, and possibly greater learning gains. For example, an intelligent tutoring system that can detect when learners are frustrated, angry, or confused can adjust its teaching and communicative style accordingly. Similarly, a system that can sense when learners are pleased will know what and when particular actions and responses are desirable (Picard, 1997).

Unfortunately, detecting and responding to user affect are difficult endeavors for several reasons. First, humans themselves often have difficulty detecting the precise emotions of other humans (Graesser, McDaniel, Chipman, Witherspoon, D'Mello, & Gholson, 2006). This may be due to the tremendous variability between individuals in their outward expression of emotion. Second, although detecting user emotion is a difficult task, having a system select and deliver a socially appropriate response is a separate and equally difficult task. Simply linking canned responses to particular user affective states is probably not enough to enhance the quality of the interaction or have a significant effect on user performance. Third, it is unclear whether certain media are more effective in eliciting affective responses from users and responding to users than others. For example, users may respond very differently to animated agents that are visually and verbally responsive than to a system that only provides text responses (CITE REF FOR FORTHCOMING STUDIES BY THE EDITORS).

In this chapter, three issues that are relevant to social intelligence and that have been studied in the context of interactive dialogue systems will be addressed. We have chosen to focus on issues that are receiving considerable attention in other research areas and that we feel are pertinent to the next generation of survey systems. The first issue involves *social agency*. We will discuss whether users communicate differently with systems that incorporate some form of social agency (e.g., animated agents) compared to those that do not. The second issue involves *detecting users' affective states* by monitoring bodily movements and using paradata, i.e. data about the user's process in performing an interactive task. We will discuss some of the state-of-the art technologies and methodologies for detecting users' affective states during the course of a real-time dialogue. The third issue is concerned with designing systems that are *socially responsive to users*. Is it the case that dialogue systems should always adhere to social conventions and respond to users in polite and socially appropriate ways? It is our belief that interactive dialogue systems that possess forms of social intelligence will greatly enhance the interactive experience for users and will yield more informative data sets for researchers. As mentioned earlier, the methods and technologies that we will be discussing have yet to make their way into survey technologies; however, in each section of the chapter we will discuss how, with slight modifications, these methods and technologies could be used by the survey community.

Social Agency

Nass and colleagues have reported on numerous occasions that humans, mostly without realizing it, apply implicit rules of human-to-human social interaction to their interactions with computers. Computers that exhibit more human-like behaviors (e.g., those that include animated characters or voices) are subject to greater social expectations from their users than those with less human-like features (although Nass and colleagues argue that even line drawings can evoke social responses, e.g. Nass, 2004; Reeves & Nass, 1996). As a result, when computers violate rules of social etiquette, act rudely, or simply fail to work at all, users become frustrated and angry. Interestingly, users do not attribute their negative emotions to their own lack of knowledge or inadequate computer skills, but instead direct them towards the computer and its inability to sense the users' needs and expectations (Miller, 2004; Mishra & Hershey, 2004, Nass, 2004; Reeves & Nass, 1996). Such claims have numerous implications for the way dialogue systems are designed, especially those that attempt to create the illusion of social agency by including voices or animated agents with human-like personas.

Animated agent technologies have received considerable attention from researchers who are interested in improving learning environments and automated therapeutic facilities (Baylor, Shen & Warren, 2004; Marsella, Johnson, & LaBore, 2000). However, it is still unclear whether such technologies improve user performance, e.g., contribute to learning gains or offer substantial benefits (therapeutic or otherwise), over environments without agents. It's worth noting that the ways in which animated agents are being used and studied in other domains could certainly inform their use in automated survey systems. For example, some intelligent tutoring systems use animated agents as tutors. These tutoring agents engage students in conversations in attempts to get students to provide information or construct knowledge. The tutor and student work persistently to negotiate meaning and establish common ground. The dialogue of tutoring interactions has some parallels to what transpires in interviewer-administered survey interviews. Respondents are often required to provide information to questions that can contain confusing or unfamiliar terminology (or even ordinary words that respondents are conceptualizing differently than the survey designers), and common ground has to be negotiated before the respondent can supply the information accurately (Conrad & Schober, 2000; Schober, Conrad, & Fricker, 2004). To date, the results from most studies in which animated agent vs. no agent comparisons have been made indicate that users tend to like environments with agents more than environments without agents. This effect is known as the Persona Effect (Andre, Rist, & Müller, 1998; Lester et al., 1997). However, there is only preliminary evidence that animated agents actually contribute to the main goals (learning, therapeutic, etc) of the systems they inhabit (Atkinson, Mayer, & Merrill, 2004; Moreno, Mayer, Spires, & Lester, 2001).

The design, implementation, and maintenance of animated agents for any kind of environment (learning, therapeutic, survey, or otherwise) is expensive and time-consuming. Before such resources are invested, it seems worthwhile to determine whether users do indeed communicate differently with systems that have some form of social agency versus those that do not. After all, it may be that the presence of a social agent (human or animated) is viewed as face-threatening. If this is the case then in a survey about a sensitive topic an animated agent might collect less candid answers than a textual interface; this is why Cassell and Miller (Chapter X) conclude that agents are probably not promising as survey interviewers. Or the opposite may be true: users may be more inclined to learn and/or disclose more when some form of social agent is present. To the extent that a survey interviewer can reassure a respondent that many people give similar answers and all answers are confidential, so an animated agent might collect more candid responses than an interface without a human-like presence.

In a recent study following this line of inquiry, Person et al. (2005) investigated how different types of social agency affect language production in interviews with college students. Fifty-nine college students from two institutions participated in interviews in which they answered questions about their alcohol consumption behaviors, their attitudes and beliefs about alcohol use, their personal lives (e.g., What do they do in their free time? What kinds of things do they do when they are hanging out with friends?), and their personal and family histories of alcohol-related problems. The participants in the study were assigned to one of four social agency conditions: (1) animated agent (AA), (2) text-only (TO), (3) instant messaging (IM), and (4) human-to-human (HH). In the AA condition, an animated agent conducted the interview and the participants typed their responses. In the TO condition, the interview questions appeared one at a time on a computer monitor, and students typed their responses. In the IM condition, a human interviewer presented the questions to the participant via an instant messaging program (i.e., Instant Messenger), and participants in the IM condition were aware that they were communicating with another human. In the HH condition, a human conducted the interview in a face-to-face setting. The interviews in the HH condition were recorded and transcribed. The 12 interview questions were the same in all four conditions, and the interviewer in the social agency conditions (AA, IM, and HH) was always a male or a male persona (in the condition with the animated agent). The length of the interview sessions ranged from approximately 10 minutes to 45 minutes.

The participants' answers were analyzed with Wmatrix, a corpus analysis and comparison software tool (Rayson, 2003, 2005). Wmatrix provides a web interface to the UCREL semantic analysis system (USAS) and CLAWS word-tagging corpus annotation tools, and to standard corpus linguistic methodologies such as frequency lists, word concordances, and grammatical and semantic category parsing. The results indicated differences between the conditions in terms of disclosure amount, semantic richness, and other linguistics features. For disclosure amount, although the means for the total numbers of words did not differ across conditions, TO = 660.6, AA = 622.8, HH = 606.2, and IM = 552.7), there were differences in disclosure amounts between the two computer conditions (TO and AA, Total words mean = 641.7) and the two human conditions (IM and HH, Total words mean = 578.5). That is, the participants disclosed more in the conditions in which the interviewer was a computer than they did in the conditions where the interviewer was a human. There were also differences between the conditions for particular semantic categories. Of interest were semantic categories related to drinking, family, relationships, religion, and emotion. Overall, it was found that participants used more semantically rich and emotion language in the AA and IM than in the TO and HH conditions.

One interpretation of these findings is that social agency is important for disclosing information, especially for information that is rich in content and emotion. However, such disclosure occurs less in human-to-human interviews, perhaps because the face-to-face interactions are too face-threatening. These findings indicate that animated agents may have a future in forthcoming survey systems given that, at least in the current case, animated agents do not seem to interfere with participants' willingness to disclose information (in contrast to Cassell and Miller's, Chapter X, predictions). In fact, animated agents may turn out to be preferable to both humans and textual interfaces for collecting personal and sensitive information in interviews.

Users' Affective States

Research in the relatively new area of Affective Computing is providing tremendous insight for the next generation of intelligent systems. Generally speaking, Affective Computing focuses on creating technologies that can monitor and appropriately respond to the affective states of the user (Picard, 1997). Evaluating the emotions of others is a skill that allows humans to function appropriately and effectively in social interactions. Mishra and Hershey (2004) reported that systems that interpret users' emotions and exhibit appropriate social responses are likely to result in increased usability and productivity on the part of the human user. In order to compensate for the lack of social skills of today's personal computers, affective computing has emerged as a new and exciting research area that attempts to bridge the gap between the emotionally expressive human and the socially challenged computer. Affective computing can be considered to be a subfield of human-computer interaction (HCI), where the affective states of a user (feelings, moods, emotions) are incorporated into the decision cycle of the interface in an attempt to develop more effective, user-friendly, and naturalistic applications (Bianchi-Berthouze & Lisetti, 2002; Prendinger & Ishizuka, 2005; Whang, Lim, & Boucsein, 2003).

Though such interfaces are being developed for a wide variety of systems, including affect-sensitive robots (Rani, Sarkar, & Smith, 2003), one particular area that has emerged as a leader in the development of affect-sensitive interfaces is the creation of intelligent tutoring systems and artificial peer learners (Conati, 2002; D'Mello et al., 2005; Kort, Reilly, & Picard, 2001; Litman & Forbes-Riley, 2004). Tutoring systems attempt to incorporate the affective states of a learner into their pedagogical strategies to (presumably) increase learning. This work parallels attempts to capture paradata in today's on-line surveys, in which information like timing of mouse clicks and keyboard entries tell something about the respondent's experience answering questions, either for subsequent additional analysis (e.g., Couper, 2000) or to change the system's behavior with the respondent (e.g., Conrad, Schober & Coiner, 2007), although thus far the work on surveys has not focused on respondent affect.

In general, an affective interaction involves the immersion of a user into an affective loop or a *detect-select-synthesize cycle*. This involves the *identification* and then the real-time *detection* of the user's affective states relevant to the domain, the *selection* of appropriate actions by the system to optimize task efficiency, and the *synthesis* of emotional expressions by the system so that the user remains engaged and the interaction is not compromised. From a broad perspective, the players in the affective loop are a user and a computer interface. However, within the context of particular domains the roles of the players become more specific. For example, within the context of intelligent learning environments, the user becomes a learner and the interface becomes an artificial tutor (D'Mello et al., 2005) or a learning companion (Kort, Reilly, & Picard, 2001). Similarly, implementing the affective loop in a dialogue based survey system would involve merging affect models for the respondent and interviewer. The respondent model would first be concerned with *identifying* the prominent affective states that a respondent experiences when participating in a survey interview. Of course, these affective states could vary depending on the nature of the information being collected in the survey or interview (e.g., ketchup brand preferences versus risky sexual behavior). Once this has been established, the manner in which respondents manifest these affective states can be investigated, thus providing the foundations for the development of an automatic affect *detection* system. The interviewer model would take into account how effective human interviewers *adapt* their behaviors to handle the emotions of the respondent (as when interviewers use a more empathetic tone or move on to the next question when a respondent seems overanxious or challenged). Finally, embodied conversational agents that simulate human interviewers would have to be programmed to *synthesize* affective elements through natural body movements such as the generation of facial expressions, the modulation of speech, and the inflection of posture.

It should be noted that the aforementioned description of an affective loop between an artificial interviewer and a human respondent is broad enough to encompass the entire gamut of the affective interaction. Of course, there may be circumstances during survey interviews in which overly expressive displays of emotions by artificial interviewers would be inappropriate or could bias the respondent's answer. Nevertheless there is important information to be gained by simply monitoring the affective states of the respondent. For example, even if no reactive strategy is employed, simply recording the emotions of the respondent may be beneficial for offline analyses (e.g., Mossholder, Settoon, Harris, & Armenakis, 1995).

In the subsequent sections we highlight the major steps involved in designing an affect-sensitive interviewer. In order to appropriately ground the discussion we consider a project at the University of Memphis that directly tackles this issue by attempting to transform an existing intelligent tutoring system, AutoTutor, into an affect-sensitive tutor. AutoTutor is a one-on-one tutoring system that helps learners construct explanations by interacting with them in natural language and by helping them solve problems in simulation environments (Graesser, Chipman, Haynes, & Olney, 2005; Graesser, Person, Harter, & TRG, 2001). The AutoTutor interface includes an animated conversational agent that speaks the content of AutoTutor's turns and displays simple facial expressions and rudimentary gestures. Students type their answers on the keyboard. AutoTutor provides *feedback* to the student (positive, neutral, or negative feedback), *pumps* the student for more information ("What else?"), *prompts* the student to fill in missing words, gives *hints*, fills in missing information with *assertions*, identifies and corrects *misconceptions* and erroneous ideas, *answers* students' questions, and *summarizes* topics. A full answer to a problem is eventually constructed by the student and AutoTutor and normally takes between 30 and 200 conversational turns.

The AutoTutor interface had 5 windows, as illustrated in Figure 1. Window 1 (top of screen) has the main question that stays on the computer screen throughout the conversation that involved answering the question. Window 2 (bottom of screen) is affiliated with the learner's answer in any one turn and Window 3 (below agent) echoes whatever the learner types in along with the responses of the tutor. Window 4 (left middle) hosts an animated conversational agent that speaks the content of AutoTutor's turns. The talking head had facial expressions and some rudimentary gestures. Window 5 (right middle) is either blank or had auxiliary diagrams.

INSERT FIGURE 1 ABOUT HERE

AutoTutor has been tested in several experiments on approximately 1000 students in computer literacy and physics courses. Significant learning gains were obtained in all of these experiments (with an average effect size of .8), particularly at the level of deep explanations as opposed to shallow facts and ideas (Graesser et al., 2004). AutoTutor has also been evaluated on the conversational smoothness and the pedagogical quality of its dialog moves in the turn-by-turn tutorial dialog (Person, Graesser, & TRG, 2002). A *bystander Turing test* on the naturalness of AutoTutor's dialog moves indicated that bystanders (i.e., research participants who read AutoTutor and human-to-human tutoring transcripts) were unable to discriminate AutoTutor dialog moves from the dialog moves of a real human tutor.

The general assumption behind the affect-sensitive AutoTutor project is that endowing the tutor with the ability to incorporate the learner's affective states into its pedagogical strategies (in addition to their cognitive states which are already tracked by the tutor) will result in even greater learning gains for students. The project integrates state-of-the-art, non-intrusive, affect-sensing technology with the AutoTutor program in an endeavor to classify emotions on the basis of facial expressions, gross body movements, and conversational cues. Figure 2 illustrates the setup of a recent study involving learners (N = 28) interacting with AutoTutor is computer literacy while the sensors silently capture their bodily data (Graesser et al., 2006). Since many of the emotions tracked in the study may occur within an interviewer-respondent interaction, the methods and results achieved in the development of the affect-sensitive AutoTutor can potentially be used as a model for developing the first affect-sensitive interviewer.

INSERT FIGURE 2 ABOUT HERE

Identifying the Affective States

Ekman and Friesen (1978) have proposed 6 *basic* emotions that are ubiquitous in everyday experience. These include fear, anger, happiness, sadness, disgust, and surprise. However, many have called into question the adequacy and validity of a theory of emotion that is based on these six basic emotions alone (e.g. Kort, Reilly, & Picard, 2001). For example, within the context of a survey it is unlikely that a respondent will experience particularly strong states of fear, happiness, or sadness. Similarly these basic emotions rarely occur during learning, at least not with AutoTutor. In a series of exploratory studies that utilized online measures of affect such as observational and emote-aloud protocols, as well as offline judgments of emotions by multiple judges, Graesser and his colleagues have identified a set of affective states that typically *do* play a significant role in learning: boredom, flow (or engagement), confusion, and frustration (Craig, Graesser, Sullins, & Gholson, 2004; D'Mello et al., 2006; Csikszentmihalyi, 1990; Graesser, et al., 2006; Graesser et al., in review).

It is reasonable to presume that the affective states of boredom, engagement, confusion, and frustration may also play important roles in the interviewer-respondent relationship. For example, boredom may occur if the survey involves an impersonal or uninteresting topic and may become more pronounced over an extended period of time. Confusion may be prevalent when respondents are not familiar with terminology or do not understand the question. Discomfort may transition into frustration if the questions are very sensitive in nature or are not deemed to be appropriate by the respondent. In the ideal, a respondent could even be engaged in the survey and have a *flow* experience, where they are so engrossed that time and fatigue disappear (Csikszentmihalyi, 1990), although declining participation and completion rates in surveys argue against this being at all common.

Methods for detecting user emotion

An emotionally-sensitive learning environment, whether it be human or computer, requires some degree of accuracy in classifying the learners' affect states. Over the last decade, efforts have been launched to automatically detect emotions by means of a variety of signal detection algorithms operating on a host of sophisticated sensors. Although the use of physiological signals have been relatively successful in emotion detection (Nakasone, Prendinger, & Ishizuka, 2005; Rani, Sarkar, & Smith, 2003; Whang, Lim, & Boucsein, 2003), these methods rely on obtrusive sensing technologies such as skin conductance, heart rate monitoring, etc. Although acceptable in some domains, once users habituate to the presence of obtrusive sensors, they may not be optimal in learning or survey environments because this approach could cause distraction and task interference (not to mention they are very expensive and inaccessible to most end-users). Therefore, we recommend the use of non-intrusive bodily sensors such as cameras that track facial features and microphones that monitor speech contours (which are also less expensive and more accessible). In fact the majority of current affect detection systems track facial features (Cohn & Kanade, in press; Oliver, Pentland, & Berand, 1997) and acoustic-prosodic vocal features (Bosch, 2003; Fernandez & Picard, 2004; Shafran & Mohri, 2005). Additionally, some pioneering research has also focused on affect detection from posture patterns (Mota & Picard, 2003). The AutoTutor team has used non-intrusive sensors such as cameras and a posture sensor, along with discourse cues that may convey affect. In the following sections, we describe how affective states have been measured in AutoTutor via discourse cues, facial expressions of users, and a posture sensor.

Affect detection from conversational cues

One of the first channels that we explored to detect affect was the discourse that occurred in AutoTutor's natural language, mixed-initiative dialogues. Although dialogue has traditionally been a relatively unexplored channel for affect/emotion detection, it is a reasonable information source to explore because dialogue information is recorded in virtually all human-computer conversations and is inexpensive to collect. In addition to detecting the emotions of a student or respondent, dialogue features can be used to infer additional task-specific information, some of which may be relevant to a tutoring or survey domain. In addition to Affective Computing innovative uses of dialogue have surfaced from research on the identification of problematic points in human-computer interactions (Batliner, Fischer, Huber, Spilker, and Noth, 2003; Carberry, Lambert, and Schroeder, 2002; Walker, Langkilde-Geary, Hastie, Wright, & Gorin, 2002). For example, Carberry et al. (2002) proposed an algorithm to recognize doubt by examining linguistic and contextual features of dialogue in conjunction with world knowledge. Recognizing doubt is certainly important in tutoring, but recognizing doubt could allow a survey interviewing system to help the respondent understand a question or carry out the task, or to know when to reassure the respondent that he or she is doing a good job.

In AutoTutor the efficacy of tutorial dialogue as a predictor of the affective states of a learner was investigated by extracting a variety of dialogue features from AutoTutor's text based log files and then connecting these with the emotions of the learner (D'Mello, Craig, Witherspoon, McDaniel, & Graesser, in press). The features included temporal measures such as response time and time into the session, measures of response verbosity (number of words), and assessments of student's ability. Other measures that influenced the affective states of the learner were measures of tutor directness (e.g., how much information does the tutor explicitly provide the learner), and tutor feedback (i.e. positive or negative), which is manifested in the verbal content, intonation, and a host of other non-verbal conversational cues of the embodied conversational agent that personifies the tutor.

Although the features of dialogue we analyzed were specific to AutoTutor, a similar set of features would presumably be expected in an interactive dialogue-based survey system. The lower level features specific to AutoTutor can be generalized to generic categories of dialogue features such as temporal assessments, response verbosity, respondent answer quality, answer attempts, and interviewer clarifications. Once these features are specified and the respondent affect model developed, observed relationships between dialogue and affect can be applied to the survey domain. In AutoTutor, for example, boredom occurs later in the session and after multiple attempts to answer a question. Alternatively, confusion occurs earlier in the session, with slower and shorter responses, and with frozen expressions (e.g. "I don't know", "What did you say?). Engagement seems to occur earlier on in the session and involves longer responses. Additional details on these and other patterns are reported in D'Mello et al. (in press).

Based on the relationships that emerged between dialogue features and emotions that occur during learning, AutoTutor researchers developed a system that could successfully differentiate the various affective states from a neutral baseline state. In particular, accuracy rates of 69%, 68%, 74%, 71%, and 78% in detecting boredom, confusion, delight, flow, and frustration, respectively, from a neutral affective were obtained (D'Mello & Graesser, 2006; D'Mello et al., in press). These results support the notion that dialogue is a reasonable source for measuring the affective states that a learner is experiencing.

Affect detection from posture patterns

Monitoring posture to infer affect is interesting because it is rarely the case that posture is intentionally controlled by humans. Tracking Posture patterns provide an added advantage when compared to facial expressions and gestures, because these motions are ordinarily unconscious, unintentional, and thereby not susceptible to social editing. Ekman and Friesen (1969), in their studies of deception, used the term *nonverbal leakage* to refer to the unsuccessful attempts of liars to disguise deceit through less controlled channels such as the body movements compared to more controlled channels like facial expressions (for more on detecting deception, see Hancock, this volume).

There have been a few studies that have documented the importance of posture in expressing affect (e.g. Coulson, 2004; Schouwstra & Hoogstraten, 1995; Wallbott, 1998). Recently, D’Mello, Chipman, and Graesser (under review) have demonstrated that posture can be a viable channel to discriminate between the affective states of boredom (low engagement) and flow (high engagement). Being able to discriminate levels of engagement has utility in survey domains; presumably many breakoffs (premature terminations of the interview by respondents) are due to respondent boredom (for evidence see Conrad, Couper, Tourangeau, & Peytchev, 2005). Establishing and maintaining the engagement of respondents is especially critical in situations with high degrees of user control, such as in automated telephone surveys and computer-based surveys where respondents can easily break off. For instance, with web-based surveys, individuals are one-mouse-click-away (or one switch of windows or departure from the room) from ending the session; desktop activities like clicking and typing won’t be useful for detecting boredom because a bored respondent won't produce any. Therefore, detecting via other means when a respondent is transitioning into boredom and proactively engaging the respondent may help prevent respondent attrition.

D'Mello and colleagues (D'Mello, Chipman, & Graesser, in review) used the Body Pressure Measurement System (BPMS), developed by Tekscan™ (1997), to track learners' posture when interacting with AutoTutor. The posture sensor is a thin pressure pad (or mat) that can be mounted on a variety of surfaces. In this study, it was placed on the seat and back of a chair that the learners sat on while interacting with AutoTutor. The pad is paper thin with a rectangular grid of sensing elements, each providing pressure in millimeters of mercury. High level posture features such as net pressure and arousal levels were tracked and their relationships to boredom and flow were determined by point biserial correlational analyses. The results indicated that boredom was manifested in two distinct forms. The first is consistent with the preconceived notion of boredom in which a learner stretches out, lays back, and simply disengages. However, a counter-intuitive finding was that *boredom* was associated with a form of restlessness manifested by rapid changes in pressure on the seat of the chair. The affective state of *flow* was associated with a heightened pressure in the seat of the chair with minimal movement. This may imply that the learner is mentally engaged in absorbing the material and thereby devotes a smaller amount of cognitive processing towards trivial bodily motion.

Machine learning algorithms were also quite successful in segregating these emotions. In particular, a nearest neighbor classifier (Aha & Kibler, 1991), perhaps one of the simplest classification schemes, achieved an accuracy of 78% in discriminating between boredom and engagement. In addition, this and other classifiers were also moderately successful in segregating each of the other target emotions from a neutral state on the basis of the posture features alone. The accuracy scores were 70%, for boredom, 65% for confusion, 70% for delight, 74%

for flow, and 72% for frustration. Although the classification accuracy rates could be improved, this exploratory research has highlighted the efficacy of monitoring the posture of the user as a viable channel to infer complex mental states.

Affect detection from facial features

Ekman and Friesen (1978) highlighted the expressive aspects of emotions with their Facial Action Coding System. This system specifies how “basic emotions” can be identified by coding specific facial behaviors and the muscle positions that produce them. Each movement in the face is referred to as an *action unit* (or AU). There are approximately 58 action units. Patterns of AUs were used to identify the emotions of happiness, sadness, surprise, disgust, anger, and fear. Developing systems to automatically detect the action units, however, is a challenging task because the coding system was tested primarily on static pictures rather than on expressions that change in real time.

AutoTutor researchers are currently exploring some of the technical challenges associated with the automated detection of the facial expressions. As an initial step, two trained judges coded a sample of the observed of facial expressions using the action units. The coding yielded a set of 12 prominent action units that were correlated with the emotions, thus reducing the requirements of an automated facial expression measurement system (McDaniel, D’Mello, King, Chipman, Tapp, & Graesser, in review) . This research has also elucidated important patterns in how learners display particular emotions. Confusion and delight are highly animated affective states and are easily detectable from facial expressions. From an evolutionary perspective, this might suggest that humans use their face as a social cue to indicate that they are confused, which helps them recruit information or resources to alleviate their confusion (see also Conrad, Schober & Dijkstra, 2007, for discussion of facial confusion cues, including gaze aversion, in survey interviews). Delight is another emotion that is also readily expressed on the face, perhaps because it is a positive emotion. However, it appears that learners do not readily display frustration, perhaps due to the negative connotations associated with this emotion. This finding is consistent with Ekman's theory of social display rules, which state that social pressures may result in the disguising of negative emotions such as frustration.

The accuracy scores for detecting affect on the basis of facial features alone are promising. Classification accuracies for detecting delight were the highest (90%), boredom the lowest (60%), and confusion (76%) and frustration (74%) were in between. These results support the conclusion that classifiers are more successful in detecting emotions that are manifested with highly animated facial activity, such as delight, than emotions that are more subtly expressed (e.g., boredom which is easily confused with neutral).

It is important to acknowledge that the facial action units used in the emotion classification analyses described above were not automatically extracted as would be required for a real-time affect-sensitive system. Instead they were rated by human coders which adversely effects the applicability of these results. Fortunately, stemming from a NSF funded workshop in 1992 (Ekman, Huany, Sejnowski, & Hager, 1992), the last decade has been ripe with research geared towards developing automated facial expression recognition systems (e.g. Bartlett, Hager, Ekman, & Sejnowski, 1999; Cohn & Kanade, in press; Fasel & Luttin, 2000; Pantic & Rothkrantz, 2000). We are currently investigating the possibility of integrating these computational systems to aid in inferring the affective states of the learner.

Socially Responsive Systems

We have already addressed the importance of social agency and detecting users' affect states in the design of socially intelligent systems. In the previous section, we discussed in detail methods for detecting users' affective states and mentioned that part of the affective loop requires the system to provide a synthesized response to user's current emotional and cognitive states. Simply put, systems should respond to users in socially and contextually appropriate ways. This begs the question of what actually constitutes a socially and contextually suitable response. Is it the case that a canned response to a user who is experiencing particular emotions (e.g., confusion and frustration) will be equally effective for all users experiencing the same emotions? Should computers avoid negativity and always respond in positive, upbeat ways? Gong (2002) reported that users tend to prefer systems that always express positive affect (e.g., are very polite or upbeat) versus ones that may display negative affect when it is contextually appropriate. This finding suggests that users are sensitive to face-threatening speech from computers. Person et al. (1994) reported that human tutors often provide ambiguous feedback to students to minimize face-threatening impositions. However, Person et al. also hypothesized that a tutor who caters solely to students' emotions by strictly adhering to politeness and conversational norms may thwart the student's efforts to discard misconceptions and to interpret clearly the intention of the tutor's feedback. Similar tradeoffs between maintaining rapport and collecting high quality responses can occur in survey interviews.

AutoTutor researchers conducted a study to determine whether altering AutoTutor's personality on politeness-related dimensions affects student learning gains (Person, Burke, & Graesser, 2003). An experimental version of AutoTutor, called RudeTutor, was designed to provide face-threatening feedback to users by flouting conversational and politeness norms and by only providing neutral or negative feedback to student responses. Below is an excerpt from a RudeTutor tutoring session.

RudeTutor: The sun exerts a gravitational force on the earth as the earth moves in its orbit around the sun. Does the earth pull equally on the sun? Explain why.

Student:	The sun pulls more.
RudeTutor:	I know that you know more than that. Say more.
Student:	The sun pulls more on the sun than vice versa.
RudeTutor:	Did you pay attention in class? How does Newton's third law of motion apply to this situation?
Student:	The sun and earth are both in motion.
RudeTutor:	No. Go back and answer the question completely. Can't you add anything?
Student:	The earth pulls on the sun and the sun pulls on the earth?
RudeTutor:	Do you know what you're talking about. I'm pretty sure you can't get this but, Newton's third law refers to the forces exerted by one body on another . . . . .?
Student:	body
RudeTutor:	another body. You really need some help. Try this. When two objects of different masses hit each other, which one exerts the greater force on the other, the heavy object, the light object, or are the forces equal?
Student:	They are equal.

Participants in the study were randomly assigned to interact with either AutoTutor or RudeTutor. Learning gains were assessed for both groups and user perception data were collected. The results indicated that participants who interacted with RudeTutor rated the system very negatively with regard to politeness, effectiveness, and likeability compared to the participants who interacted with the regular AutoTutor. However, it was also the case that all students experienced equivalent positive learning gains, regardless of whether they interacted with RudeTutor or AutoTutor. That is, students who were subjected to face-threatening responses and who received excessive negative feedback learned just as much as those who interacted with the regular, polite AutoTutor. One particularly interesting result involved participants' overall enjoyment of the tutoring sessions: some participants enjoyed interacting with RudeTutor very much, while others seemed to be genuinely offended by its lack of politeness.

The results of the RudeTutor study generated more questions than they answered. For example, how are users' affective and cognitive states affected by the tutor's seemingly inappropriate responses? What is the nature of the individual differences that result in different enjoyment levels for the students? In human-to-human conversations, the affective states of the participants are determined not only by the affective nature of the conversational turns but also by the personalities of the participants. Human-to-computer interactions function much the same way because humans readily and subconsciously apply social expectations to computers irrespective as to whether or not it is appropriate to do so (Reeves & Nass, 1996). Although many investigations have focused on how personality is related to a variety of learning related factors (e.g., achievement, learning approaches, specific aptitudes), and personality researchers have theorized for years about the interaction between teacher and student personality (Mills, 1993, Zhang, 2003), little or no research has attempted to investigate the dynamics of tutor personality (or responsiveness), student personality, and affective states with a high level of experimental precision.

Conclusions

The work described here demonstrates that our vision of developing a socially intelligent, affect-sensitive tutor is not merely an exercise in science fiction, and that some components of social intelligence are already being implemented in tutoring systems. In our view, when our approaches are coupled with serious knowledge engineering efforts, several of the methods and computational algorithms that we have used or developed may be applicable for creating socially intelligent, affect-sensitive, dialogue-based, survey systems. Although, tutoring interactions are not precisely the same as survey interactions; learner motivation to interact with a tutor is often quite different from respondent motivation to interact with an interviewer, and the flow of the tutorial dialogue follows different routes than survey dialogue. Nonetheless, the similarities are great enough that much can be extrapolated, and with existing (and increasingly less costly) technologies.

In particular the use of dialogue and facial expressions could be used to effectively and reliably measure the emotions of a respondent. Although we have used an expensive, hand crafted camera for facial feature tracking, recent technology makes this possible with a simple web cam, thereby significantly reducing the associated equipment costs. In fact, certain laptop manufactures now market their systems with integrated cameras that are optimized to perform on specialized hardware. Noise reduction microphones are also routinely shipped with contemporary laptops. Although we did not initially track acoustic-prosodic features of speech to infer the affective state of a learner, the literature is rich with such efforts (see Pantic & Rothkrantz, 2003 .for a comprehensive review).

In fact, the only system we have discussed that may not readily be useful for the survey domain would be the pressure sensitive chair. However, with some innovation, some of its features can be approximated with a visual image captured by the camera. For example, the distance between the tip of the nose of a user to the monitor can be used to operationally define a leaning forward posture. The movement of the head can be used to approximate arousal. Also some recent evidence suggests that posture features may be redundant with the facial expressions and dialogue features, thereby eliminating the need for an expensive pressure sensor (D'Mello & Graesser, in press).

Likewise agent technologies have matured considerably. Not only have agent technologies become easier to use, but they also have greatly increased in realism. Nowhere is this more evident than in the Microsoft Agent technology (Microsoft, 1998). Cutting edge ten years ago, Microsoft Agent's talking heads, and the infamous Clippy paperclip, have been eclipsed by the new wave of 3D full body agent technologies from Haptek™ and the game Unreal Tournament 2004. These agent technologies are already in use by current tutoring systems (Graesser et al., 2005; Johnson et al., 2005) and have advanced tool suites that allow the agent's behavior and emotions to be customized. These agent technologies, already proven in the tutoring domain. can be easily extended to provide socially responsive agents in the survey domain. The real progress will come not in deploying existing affect detection and production models but instead in basic research that improves the accuracy of these models with respect to individual differences.

Acknowledgements

We thank our research colleagues in the Emotive Computing Group and the Tutoring Research Group (TRG) at the University of Memphis (http://emotion.autotutor.org). Special thanks to Art Graesser. We gratefully acknowledge our partners at the Affective Computing Research Group at MIT. We thank Steelcase Inc. for providing us with the Tekscan Body Pressure Measurement System at no cost. This research was supported by the National Science Foundation (REC 0106965, ITR 0325428, and REC 0633918) and the DoD Multidisciplinary University Research Initiative administered by ONR under grant N00014-00-1-0600. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF, DoD, or ONR.

References

Aha, D., & Kibler, D. (1991). Instance-based learning algorithms. *Machine Learning, 6,* 37-66.

André, E., Rist, T., & Müller, J. (1998). WebPersona: A life-like presentation agent for the world-wide web. *Knowledge-Based Systems, 11,* 25-36.

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2004). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology*. Elsevier, Inc.

Batliner, A., Fischer, K., Huber, R., Spilker, J., & Noth, E. (2003). How to find trouble in communication. *Speech Communication, 40,* 117–143.

Bartlett, M.S., Hager, J.C., Ekman, P., and Sejnowski, T.J. (1999). Measuring facial expressions by computer image analysis. P*sychophysiology, 36,* 253-263.

Baylor, A. L., Shen, E., & Warren, D. (2004). Supporting learners with math anxiety: The impact of pedagogical agent emotional and motivational support. *ITS 2004 Workshop Proceedings on Social and Emotional Intelligence in Learning Environments.* Maceio, Brazil: Springer-Verlag.

Bianchi-Berthouze, N. & Lisetti, C. L. (2002). Modeling multimodal expression of users affective subjective experience. *User Modeling and User-Adapted Interaction 12 (1),* 49-84.

Bosch, L. T. (2003). Emotions, speech, and the ASR framework. *Speech Communication 40 (1-2),* 213-215.

Carberry, S., Lambert, L., & Schroeder, L. (2002). Toward recognizing and conveying an attitude of doubt via natural language. *Applied Artificial Intelligence 16(7),* 495-517.

Cassell, J. (2007). Is it self-administration if the computer gives you encouraging looks? In M. Schober & F. Conrad (eds.) *Envisioning the Survey Interview of the Future*. Wiley.

Cohn, J. F. & Kanade, T. (in press). Use of automated facial image analysis for measurement of emotion expression. In J. A. Coan & J. B. Allen (Eds.), *The handbook of emotion elicitation and assessment. Oxford University Press Series in Affective Science.* New York: Oxford.

Conati C. (2002). Probabilistic assessment of user's emotions in educational games. *Journal of Applied Artificial Intelligence, 16,* 555-575.

Conrad, F. G., & Schober, M. F. (2000). Clarifying question meaning in a household telephone survey. *Public Opinion Quarterly, 64,* 1-28.

Coulson, M. (2004). Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior, 28,* 117-139.

Craig, S. D., Graesser, A. C., Sullins, J., & Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning. *Journal of Educational Media, 29,* 241-250.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper-Row.

D'Mello, S. K., Craig, S. D., Gholson, B., Franklin, S., Picard, R., & Graesser, A. C. (2005). Integrating affect sensors in an intelligent tutoring system. In Affective Interactions: The Computer in the Affective Loop Workshop at 2005 *International conference on Intelligent User Interfaces* (pp. 7-13). New York: AMC Press.

D'Mello, S., & Graesser, A. C. (2006). Affect detection from human-computer dialogue with an intelligent tutoring system. In J. Gratch et al. (Eds.), *IVA 2006, LNAI 4133* (pp. 54 – 67). Berlin Heidelberg: Springer-Verlag.

D'Mello, S., & Graesser, A. C. (in press). Mind and Body: Dialogue and Posture for Affect Detection in Learning Environments. Proceedings of the 13th International Conference on Artificial Intelligence in Education (AIED 2007).

D'Mello, S. K., Craig, S. D., Sullins, J., & Graesser, A. C. (2006). Predicting affective states through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue. *International Journal of Artificial Intelligence in Education, 16,* 3-28.

D'Mello, S. K., Chipman, P., & Graesser, A. C. (in review). *Posture as a predictor of learner's affective engagement.*

D'Mello, S. K., Craig, S. D., Witherspoon, A. M., McDaniel, B. T., & Graesser, A. C. (in press). Automatic detection of learner's affect from conversational cues. *User Modeling and User Adapted Interaction.*

Ekman, P. & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32,* 88-105.

Ekman, P., & Friesen, W. V. (1978). *The facial action coding system: A technique for the measurement of facial movement*. Palo Alto: Consulting Psychologists Press.

Ekman, P., Huang, T.S., Sejnowski, T.J., & Hager, J.C. (July 30 to August 1, 1992). *Final report to NSF of the planning workshop on facial expression understanding,* Washington, DC: NSF. *http://face-and-emotion.com/dataface/nsfrept/overview.html*

Fasel, B. & Luttin, J. Recognition of asymmetric facial action unit activities and intensities. *Proceedings of the International Conference on Pattern Recognition (ICPR 2000),* Barcelona, Spain.

Fernandez, R., & Picard R. W. (2004). Modeling driver's speech under stress. *Speech Communication, 40,* 145-159.

Gong, Li. (2002). *Towards a theory of social intelligence for interface agents*. Paper presented at Virtual Conversational Characters: Applications, Methods, and Research Challenges, Melbourne, Australia.

Graesser, A.C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions in Education 48,* 612-618.

Graesser, A.C., Chipman, P., King, B., McDaniel, B., & D'Mello, S. (in review). *Emotions and learning with AutoTutor*.

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers, 36,* 180-193.

Graesser, A.C., McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S. K., & Gholson, B. (2006). Detection of emotions during learning with AutoTutor. In R. Son (Ed.), *Proceedings of the 28th Annual Meetings of the Cognitive Science Society* (pp. 285-290). Mahwah, NJ: Erlbaum.

Graesser, A. C., Person, N. K., Harter, D., & the Tutoring Research Group. (2001). Teaching tactics and dialogue in AutoTutor. *International Journal of Artificial Intelligence in Education, 12,* 257-279.

Johnson, L., Mayer, R., Andre, E., & Rehm, M. (2005). Cross-cultural evaluation of politeness in tactics for pedagogical agents. *Proceedings of the 12th International Conference on Artificial Intelligence in Education.*

Kort, B., Reilly, R., & Picard, R. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy—building a learning companion. In T. Okamoto, R. Hartley, Kinshuk, & J. P. Klus (Eds.), *Proceedings IEEE International Conference on Advanced Learning Technology: Issues, Achievements and Challenges* (pp. 43-48). Madison, Wisconsin: IEEE Computer Society.

Lester, J., Converse, S., Kahler, S., Barlow, T., Stone, B., & Bhogal, R. (1997). The persona effect: Affective impact of animated pedagogical agents. *Proceedings of CHI '97* (pp. 359-366). ACM Press.

Litman, D. J. & Forbes-Riley, K. (2004). Predicting student emotions in computer-human tutoring dialogues. *Proceedings of the 42nd annual meeting of the Association for ComputationalLlinguistics* (pp. 352-359). East Stroudsburg, PA: Association for Computational Linguistics.

Marsella S., Johnson, W. L., & LaBore, K. (2000). Interactive pedagogical drama. *Proceedings of the 4th International Conference on Autonomous Agents.*

McDaniel, B. T., D'Mello, S. K., King, B. G., Chipman, P., Tapp, K., & Graesser, A. C. (in review). Facial Features for Affective State Detection in Learning Environments.

Microsoft Corporation. (1998). *Microsoft Agent Software Development Kit with CD-Rom*. Microsoft Press, Redmond, WA.

Miller, C. (2004). Human computer etiquette: Managing expectations with intentional agents. *Communications of the ACM, 47,* 31-34.

Mills, C. (1993). Personality, learning style, and cognitive style profiles of mathematically talented students. *European Journal for High Ability, 4,* 70-85.

Mishra P., & Hershey, K. (2004). Etiquette and the design of educational technology. *Communications of the ACM, 47,* 45-49.

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

Mossholder, K. W., Settoon, R. P., Harris, S. G., & Armenakis, A. A. (1995). Measuring emotion in open-ended survey responses: An application of textual data analysis. *Journal of Management,* Vol. 21, No. 2, 335-355.

Mota, S. & Picard, R.W. (2003). Automated posture analysis for detecting learner's interest level. *Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction.*

Nakasone, A., Prendinger, H., & Ishizuka, M. (2005). Emotion recognition from electromyography and skin conductance. *Proceedings of the Fifth International Workshop on Biosignal Interpretation* (pp. 219-222). Tokyo, Japan: IEEE.

Nass, C. (2004). Etiquette and equality: Exhibitions and expectations of computer politeness. *Communications of the ACM, 47,* 35-37.

Oliver, N., Pentland, A., & Berand, F. (1997). LAFTER: A real-time lips and face tracker with facial expression recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 123-129). San Juan, Puerto Rico: IEEE.

Pantic, M. & Rothkrantz, M. (2000). Expert system for automatic analysis of facial expression. *Image and Vision Computing, 18,* 881-905.

Pantic, M. & Rothkrantz, L. J. M. (2003). Towards an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE, Special Issue on Multimodal Human-Computer Interaction, 91(9),* 1370-1390.

Person, N. K., Burke, D. R., & Graesser, A. C. (2003). RudeTutor: A face-threatening agent. *Proceedings of the Society for Text and Discourse Thirteenth Annual Meeting. Madrid, Spain*.

Person, N. K., Graesser, A. C., & The Tutoring Research Group (2002). Human or computer: AutoTutor in a bystander Turing test. In S. A. Cerri, G. Gouarderes, & F. Paraguacu (Eds.) *Intelligent Tutoring Systems 2002 Proceedings* (pp. 821-830). Berlin: Springer-Verlag.

Person, N. K., Kreuz, R. J., Zwaan, R., & Graesser, A. C. (1994). Pragmatics and pedagogy: Conversational rules may inhibit effective tutoring. *Cognition and Instruction, 2,* 161-188.

Person, N. K., Petschonek, S., Gardner, P. C., Bray, M. D., & Lancaster, W. (2005). *Linguistic features of interviews about alcohol use in different conversational media*. Presented at the 15th Annual Meeting of the Society for Text and Discourse. Amsterdam, The Netherlands.

Picard, R. W. (1997). *Affective computing*. Cambridge, MA: MIT Press.

Prendinger, H. & Ishizuka, M. (2005). The empathic companion: A character-based interface that addresses users' affective states. *International Journal of Applied Artificial Intelligence, 19(3,4),* 267-285.

Rani, P., Sarkar, N., & Smith, C. A. (2003). An affect-sensitive human-robot cooperation: Theory and experiments. *Proceedings of the IEEE Conference on Robotics and Automation* (pp. 2382 – 2387). Taipei, Taiwan: IEEE.

Rayson, P. (2003). Wmatrix: A statistical method and software tool for linguistic analysis through corpus comparison. Ph.D. thesis, Lancaster University.

Rayson, P. (2005) Wmatrix: a web-based corpus processing environment. Retrieved from Lancaster University Computing Department Web site: http://www.comp.lancs.ac.uk/ucrel/wmatrix/.

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

Schober, M. F., Conrad, F. G. & Fricker, S. S. (2004). Misunderstanding standardized language. *Applied Cognitive Psychology, 18*, 169-188.

Schouwstra, S., & Hoogstraten, J. (1995). Head position and spinal position as determinants of perceived emotional state. *Perceptual and Motor Skills, 81,* 673-674.

Shafran, I., & Mohri, M. (2005). A comparison of classifiers for detecting emotion from speech. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing* (pp. 341-344). Philadelphia, PA: IEEE.

Tekscan (1997). *Tekscan body pressure measurement system user's manual*. Tekscan Inc., South Boston, MA, USA.

Walker, M. A., Langkilde-Geary, I., Hastie, H. W., Wright, J., & Gorin, A. (2002). Automatically training a problematic dialogue predictor for a spoken dialogue system. *Journal of Artificial Intelligence Research, 16,* 293–319.

Wallbott, N. (1998). Bodily expression of emotion. *European Journal of Social Psychology, 28,* 879-896.

Whang, M. C., Lim, J. S., & W. Boucsein, W. (2003). Preparing computers for affective communication: A psychophysiological concept and preliminary results. *Human Factors*, *45,* 623-634.

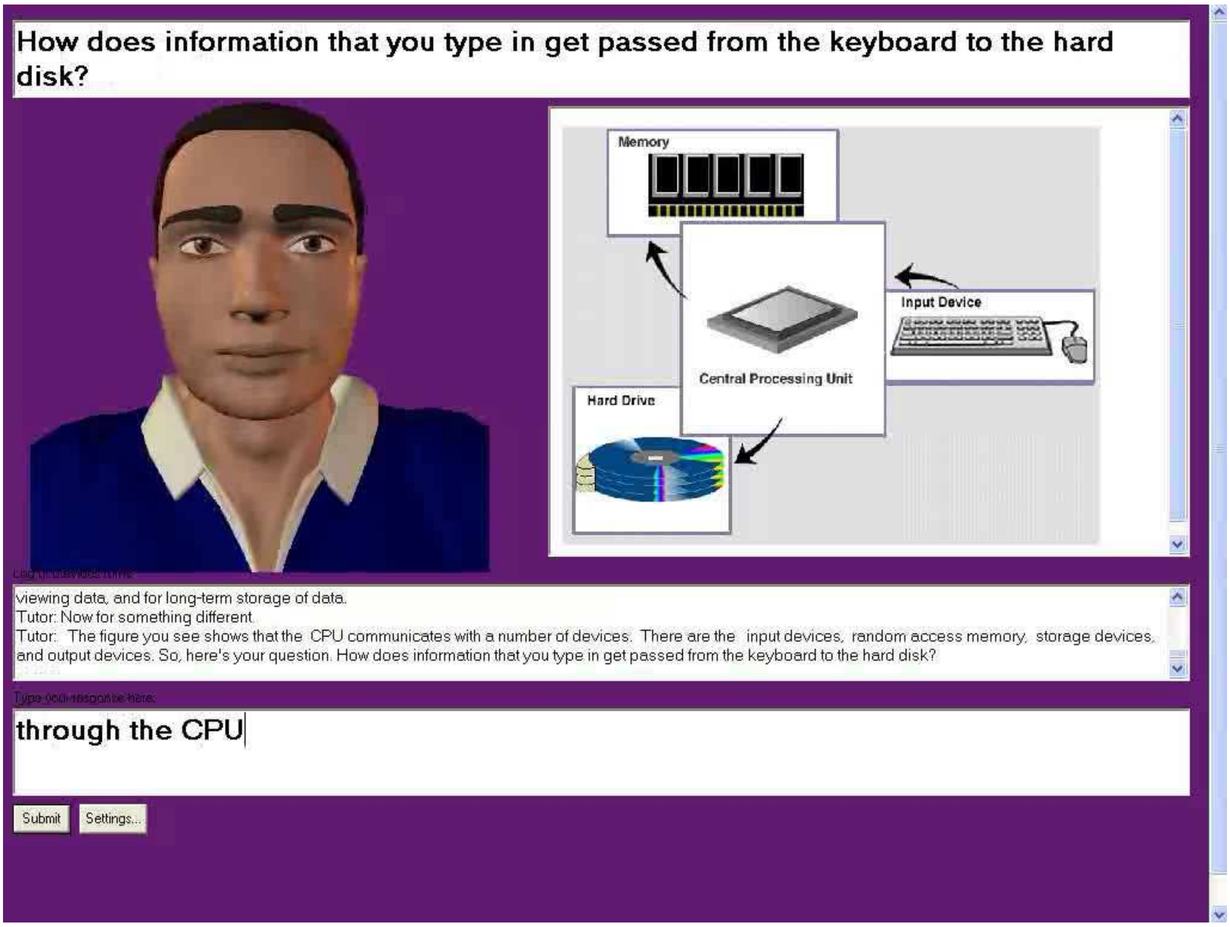
Wilson, M. (1987). MRC psycholinguistic database: Machine usable dictionary. *Technical Report*. Oxford: Oxford Text Archive.

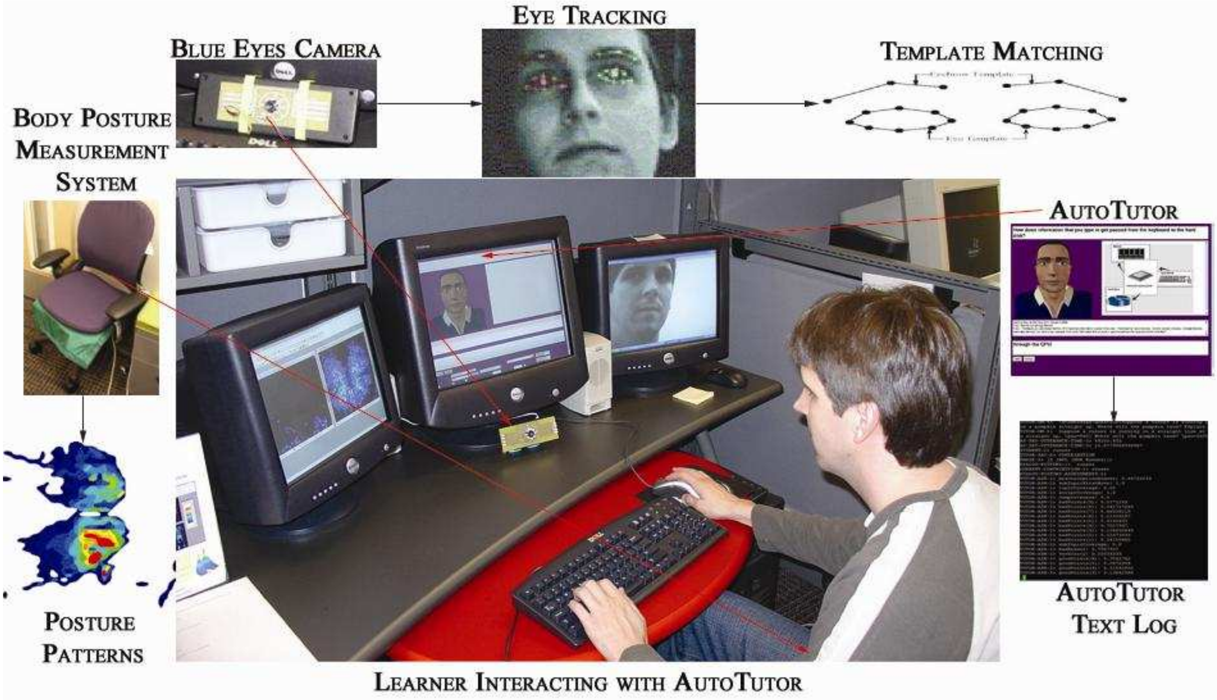
Zhang, L. (2003). Does the big five predict learning approaches? *Personality and Individual Differences*, *34,* 1431-1446.

Figure Captions

Figure 1. The AutoTutor interface

Figure 2. Sensors used to track diagnostic data while learner interacts with AutoTutor


How does information that you type in get passed from the keyboard to the hard disk?
Memory
Input Device
Central Processing Unit
Hard Drive
viewing data, and for long-term storage of data.
Tutor: Now for something different.
Tutor: The figure you see shows that the CPU communicates with a number of devices. There are the input devices, random access memory, storage devices, and output devices. So, here's your question. How does information that you type in get passed from the keyboard to the hard disk?
through the CPU
Submit
Settings...



Note. The left and right monitors are turned off during actual data collection