The Politeness Effect: Pedagogical Agents and Learning Outcomes

Ning Wang[1], W. Lewis Johnson[2], Richard E. Mayer[3], Paola Rizzo[4], Erin Shaw[5], Heather Collins[6]

[1]Institute for Creative Technologies
University of Southern California
13274 Fiji Way,
Marina del Rey, CA 90066 USA
Email: nwang@ict.usc.edu

[2]Information Sciences Institute
University of Southern California
4676 Admiralty Way,
Marina del Rey, CA 90066 USA
Email: johnson@isi.edu

[3]Dept. of Psychology
University of California, Santa Barbara
Santa Barbara, CA 93106-9660 USA
Email: mayer@psych.ucsb.edu

[4]Dept. of Computer Science
University of Rome "La Sapienza"
Via Salaria 113, 00198 Roma Italy
Email: pr04@libero.it

[5]Information Sciences Institute
University of Southern California
4676 Admiralty Way,
Marina del Rey, CA 90066 USA
Email: shaw@isi.edu

[6]Dept. of Psychology
University of California, Santa Barbara
Santa Barbara, CA 93106 USA
Email: hcollins@umail.ucsb.edu

Corresponding Author:

Ning Wang

[1]Institute for Creative Technologies
University of Southern California
13274 Fiji Way,
Marina del Rey, CA 90066 USA
Email: nwang@ict.usc.edu
Tel: +1(310) 574-5700
Fax: +1(310) 574-1631

Abstract

Pedagogical agent research seeks to exploit Reeves and Nass's Media Equation theory, which holds that users respond to interactive media as if they were social actors. Investigations have tended to focus on the media used to realize the pedagogical agent, e.g., the use of animated talking heads and voices, and the results have been mixed. This paper focuses instead on the manner in which a pedagogical agent communicates with learners, i.e., on the extent to which it exhibits social intelligence. A model of socially intelligent tutorial dialog was developed based on politeness theory, and implemented in an agent interface within an online learning system called Virtual Factory Teaching System. A series of Wizard-of-Oz studies was conducted in which subjects either received polite tutorial feedback that promotes learner face and mitigates face threat, or received direct feedback that disregards learner face. The polite version yielded better learning outcomes, and the effect was amplified in learners who expressed a preference for indirect feedback, who had less computer experience, and who lacked engineering backgrounds. These results confirm the hypothesis that learners tend to respond to pedagogical agents as social actors, and suggest that research should focus less on the media in which agents are realized, and place more emphasis on the agent's social intelligence.

The Politeness Effect: Pedagogical Agents and Learning Outcomes

1. Introduction

Researchers have for several years been investigating the potential of pedagogical agents to promote learning. One of the most influential papers in this area was a study by Lester, Converse, Kahler, Barlow, Stone and Bhogal (1997) that demonstrated a Persona Effect, in which learning was facilitated by an animated pedagogical agent that had a life-like persona and expressed affect. The rationale for this research has been the media equation hypothesis of Reeves and Nass (1996). The media equation holds that people tend to respond to interactive media much as they do to human beings. That is, they respond as if the media were social actors.

A number of pedagogical agent investigations have since been conducted, seeking to understand the Persona Effect in more detail, and replicate it in a range of learning domains (Johnson et al., 2000; Moreno et al., 2001). The results of these studies have been mixed. For example, a study by André, Rist and Müller (1998) showed that animated agents reduce the perceived difficulty of the material being learned, and a study by Bickmore (2003) showed that subjects liked an animated agent that responded socially to them, but neither study reported significant learning gains. Moreover, several studies (Atkinson et al., 2005; Graesser et al., 2003; Mayer et al., 2004; Mayer et al., 2003; Moreno and Mayer, 2000, 2004) indicated that the agent's voice was the significant contributor to learning outcomes, not the animated persona. Thus the Persona Effect is at best unreliable, and may in fact be a misnomer if the animated persona is not the primary cause of the learning outcomes.

This paper examines a different approach to applying the media equation hypothesis to intelligent tutoring. If as Reeves and Nass (1996) suggest learners respond to pedagogical agents as if they were social actors, then the agents' effectiveness should depend upon whether or not they *behave* like social actors. The agents should be socially intelligent, acting in a manner that is consistent with their social role, in accordance with social norms. In fact, human tutors make extensive use of social intelligence when they motivate and support learners (Lepper et al., 1993). Thus, social intelligence in pedagogical agents may be important not just to gain user acceptance, but also to increase the effectiveness of the agent as a pedagogical intervention.

To test this hypothesis, a model of motivational tutorial tactics was developed, based upon politeness theory (Brown and Levinson, 1987; Johnson et al., 2004). A preliminary series of Wizard-of-Oz studies was conducted in which subjects either received polite tutorial feedback that promotes learner face and mitigates face threat, or received direct feedback that disregards learner face. The polite version led to improvements in learning outcomes, and the effect was amplified in learners who expressed a preference for indirect feedback. We also observed effects on learner attitudes and motivation (Wang et al., 2005).

We term the effect demonstrated here the Politeness Effect. Our results suggest that pedagogical agent research should perhaps place less emphasis on the Persona Effect in animated pedagogical agents, and focus more on the Politeness Effect and related means by which pedagogical agents can exhibit social intelligence in their interactions with learners.

2. Related work

In recent years, the recognition of the importance of affect and motivation in learning has led increasingly to the development of socially-aware pedagogical agents as reflected in the work of Del Soldato and du Boulay (1995) and De Vicente and Pain (2002). Heylen et al. (2003) highlight the importance of these factors in tutors, and examine the interpersonal factors that should be taken into account when creating socially intelligent computer tutors. Cooper (2003) has shown that profound empathy in teaching relationships is important because it stimulates positive emotions and interactions that favour learning. Baylor and Ebbers (2003) have conducted experiments in which learners interact with multiple pedagogical agents, one of which seeks to motivate the learner. Other researchers such as Kort et al. (Aist et al., 2002; Kort et al., 2001), and Zhou and Conati (2003) have been addressing the problem of detecting learner affect and motivation, and influencing it. User interface and agent researchers are also beginning to apply the Brown and Levinson model of politeness to human-computer interaction in other contexts (Andre et al., 1998; Cassell and Bickmore, 2003; Miller, 2002).

Porayska-Pomsta (2004) used the Brown and Levinson model of politeness to analyse teacher communications in classroom settings. Although there are similarities between her approach and the approach described here, her model makes relatively less use of face threat mitigating strategies. This may be due to the differences in the social contexts being modelled.

2.1. Politeness Theory

Brown and Levinson (1987) have devised a cross-cultural theory of politeness, according to which everybody has a positive and negative "face". Negative face is the want to be unimpeded by others (autonomy), while positive face is the want to be desirable to others (approval). Some communicative acts, such as requests and offers, can threaten the hearer's negative face, positive face, or both, and therefore are referred to as Face Threatening Acts (FTAs). Consider a critique of the learner such as "You did not save your factory parameters. Save them now." There are two types of face threat in this example: the criticism of the learner's action is a threat to positive face, and the instruction of what to do is a threat to negative face.

In the Brown and Levinson model, evaluation of face threat depends upon several factors. First, the relative weightiness of different face threats is culturally dependent. The weightiness of a face-threatening act also depends upon the relative power between the speaker and the listener.

Tutors generally have power relative to learners, so we would generally expect tutors to make use of weaker politeness strategies when speaking to learners than the learners use in reverse. Finally, the weightiness of a face threat depends upon the social distance between the two parties. As two people interact over time their social distance often decreases, reducing the severity of face threatening acts and increasing the likelihood that bald-on-record strategies (i.e., direct requests that lack face-saving features) will be used.

Speakers use various politeness strategies to mitigate face threats, according to the severity, or "weightiness", of the FTA. In the above case ("You did not save your factory parameters. Save them now."), the tutor could omit the criticism of the learner and focus on the suggested action, i.e., to save the factory parameters. Alternatively the tutor could perform the face-threatening act off record, i.e., so as to avoid assigning responsibility to the hearer. An example of this would be "The factory parameters need saving." The face threat of the instruction can be mitigated using negative politeness tactics, i.e., phrasing that gives the hearer the option of not following the advice, e.g., "Do you want to save the factory now?" Positive politeness strategies can also be employed that emphasizes common ground and cooperation between the tutor and learner, e.g., "How about if we save our factory now?" Other positive politeness strategies include overt expressions of approval, such as, "That is very good".

2.2. Analyzing politeness in tutorial interactions

To investigate the role of politeness in learner-tutor interaction, we videotaped interactions between learners and a human tutor while the students were working with a particular on-line learning environment, the Virtual Factory Teaching System (VFTS) (Dessouky et al., 2001). VFTS is a web-based factory modeling and simulation system. Students read through an on-line tutorial in a Web browser, and carried out actions on the VFTS simulation as indicated by the tutorial.

Learners were supposed to analyse the history of previous factory orders in order to forecast future demand, develop a production plan, and then schedule the processing of jobs within the factory in order to meet the demand. The tutor sat next to the students as they worked, and could interact with them as the student or the tutor felt appropriate. Completing the entire scenario required approximately two hours of work, divided into two sessions of around one hour each. To analyse the interactions, and use them in designing learner-agent dialog, we transcribed them and annotated them using the DISCOUNT scheme (Pilkington, 1999).

The politeness theory of Brown and Levinson (1987) proved to be effective in accounting for the tutorial tactics observed in these dialogs. The following patterns of politeness strategies were associated with each type of tutor support (listed from most to least frequent):

- Suggesting actions:
    - To avoid threatening student negative face, the tutor mostly applied negative politeness strategies, e.g.: “You will probably want to look at the work centres”, or “Want to look at your capacity?”. A negative politeness strategy used quite often by the tutor is “conventional indirectness”: a compromise between the desires to be direct and to be indirect, resulting in a communicative act that has a non-literal meaning based on conventions. Examples from our transcripts are: “They're asking you to go back and maybe change it”, or “What they're telling you is to go and try to get the error terms”. This strategy enables the tutor to deflect to the system or interface the responsibility of requesting the student to perform an action.
    - In other cases the tutor chose a positive politeness strategy, by phrasing suggestions as activities to be performed jointly by the tutor and the learner, e.g.: “So why don’t we go

back to the tutorial factory…", or emphasize common goals between the tutor and the learner, e.g. "Run your factory, that's what I'd do."

- Providing feedback:
    - Negative feedback might threaten the student's positive face, so the tutor mostly used off-record politeness strategies, e.g.: "So the methodology you used for product 1 probably wasn't that good." In some cases, the tutor provided feedback by promoting interest and reflection, as well as affecting face, using "socratic" communicative acts such as: "Well, think about what you did…".

In a preliminary study (Mayer et al., 2006), students were asked to rate 16 tutorial statements on negative politeness (i.e., how much the tutor allows me freedom to make my own decisions") and positive politeness (i.e., how much the tutor was "working with me"). Consistent with Brown and Levinson's (1987) politeness theory, students rated direct commands and commands attributed to machines as lowest in negative and positive politeness, rated guarded suggestions and guarded questions as highest in negative politeness, and rated guarded suggestions and statements expressing a common goal as highest in positive politeness. This sensitivity to the politeness of on-line tutor's statements was stronger for students with low rather than high computing experience. These results have implications for the design of polite conversational agents in educational software, which were directly tested in the present experiment.

2.3. Motivation Theory

Why is it that tutors attend to learner face, and seek to mitigate threats to learner face? It could be a simple matter of common courtesy, as in any polite discourse. We conjecture that there is more to it than that—that tutors employ politeness strategies in order to promote learning. More precisely, we hypothesize that there is a relationship between learner face and learner motivational states, and that politeness strategies can affect the impact of tutorial tactics on motivational states. The motivational state of the learner during learning can in turn have an impact on learning outcomes.

Motivation is derived from the Latin verb *movere* (to move). In Pintrich and Schunk (2002), motivation is defined as the process whereby goal-directed activity is instigated and sustained. Motivation can influence what, when, and how we learn (Schunk, 1991). Students who are motivated to learn about a topic are apt to engage in activities they believe will help them learn, such as attending carefully to the instruction, mentally organizing and rehearsing the material to be learned, taking notes to facilitated subsequent studying, monitoring their learning progress, evaluating how well they're doing and asking for help when they don't understand the material (Zimmerman and Martinez-Pons, 1992). Collectively, these deep learning activities can improve the learning outcome. Also, motivation bears a reciprocal relation to learning: motivation influences learning and what students learn influences their motivation (Schunk 1991).

Over the years, there has been disagreement about the nature of motivation and the operation of motivational process. However, researchers agree on the behavioral indexes that indicate the presence of motivation. The commonly used indexes are choice of tasks, effort and persistence. Choice of task is a natural indicator of motivation. When students have a choice, what they choose to do indicates where their motivation is. Despite the intuitive appeal of choice of tasks, choice often is not a useful index of motivation in schools or experiment settings because in many classroom and experiments, students typically have few if any choices (Brophy, 1983). Effort becomes the second index since learning often is not easy. Students motivated to learn are apt to expend effort to succeed. Persistence, or time spent on a task, is also commonly used by researchers as a measure of motivation. Students who are motivated to learn tend to persist at the task, especially when they encounter obstacles. Students who choose to engage in a task, expend effort, and persist are likely to achieve at a higher level (Pintrich and Schrauben, 1992; Schunk, 1991).

Knowing the behavioral indication of motivation, tutors increasingly are exploring how to structure cognitive and social factors to optimize learner motivation. The work of Lepper on highly effective human tutors (Lepper et al. 1997; Lepper and Woolverton 2002) shows that up to 50% of the human tutor's interactions with their students are focused on affective elements. This is because learning, especially deep learning, tends to be affiliated with negative emotions, in the face of cognitive disequilibrium (Piaget, 1952). Cognitive disequilibrium has a high likelihood of activating conscious, effortful, cognitive deliberation, questions, and inquiry that aims to restore cognitive equilibrium. The affective states of confusion, and perhaps frustration, are likely to occur during cognitive disequilibrium (Kort, Reilly and Picard, 2001). Moreover, Lepper and Hodell (1989) identified several factors that contribute to learner motivation: curiosity, confidence, challenge, and control. We seek to create pedagogical agents that promote learners' self-confidence and control, in particular, by influencing the learner's sense of autonomy and approval.

2.3.1. Sense of Autonomy

Sense of autonomy is a feeling that your action is unimpeded by others. Sense of autonomy does not directly translate to "control" in educational psychology literature. The notion "control" has two aspects of meaning. The first aspect refers to freedom of choice. Students are in control if they are free to choose what they want to learn, set learning goals, make and follow plans to achieve their goals. Even in a classroom or experiment settings, when a learning task is given, students often are free to choose how they want to carry out the learning task instead of just following the teacher’s instruction. The second aspect refers to “locus of control”, which means the learner’s beliefs about the extent to which their behaviors influence outcomes. Students who believe they have control over whether they succeed or fail would be more motivated to engage in academic tasks, expend effort, and persist on difficult material than students who believe their actions have little effect on outcomes. In turn, these motivational effects should improve learning. Sense of autonomy is more in accordance with the first aspect of the notion “control”. Respect of learners’ need for autonomy promotes self-motivated and self-regulated learning. Providing learners with choices and a sense of control over their learning outcomes may enhance intrinsic motivation (de Charms, 1968; Deci, 1980).

deCharms (1968) used the concept of perceived locus of causality to emphasize that, when people feel autonomous, they experience the initiation of their behavior to be within themselves and they become more intrinsically motivated. On the other hand, when the perceived locus of causality is external to themselves, people tend to lose intrinsic motivation because their need for autonomy is not satisfied. Thus any factor that induces an external perceived locus of causality is predicted to diminish intrinsic motivation, and any factor that fosters an internal perceived locus of causality is predicted to enhance intrinsic motivation.

Numerous studies have found that positive performance feedback enhances intrinsic motivation because it supports their perceived competence (Deci, 1971), while negative performance feedback undermines intrinsic motivation because it diminishes their perceived competence. Thus, telling people they did well at an activity tends to increase their interest in the activity, but telling them they did badly tends to diminish their interest (Vallerand and Reid, 1984). However, subsequent studies by Fisher (1978) and Ryan (1982) indicated than perceived competence enhances intrinsic motivation only if it is accompanied by supports for autonomy. Kast and Connor (1988) did a comparable experiment in which they provided positive feedback with either a controlling or non-controlling style and found that controlling positive feedback undermined intrinsic motivation, whereas positive feedback that was clearly non-controlling enhanced it. This experimental result is consistent with Fisher's (1978) and Ryan's (1982) findings about the importance of giving positive feedback in a way that also supports autonomy.

2.3.2. Sense of Approval

As described in the previous sections, a polite tutor respects student's need for approval. This in turn can help boost and maintain the student's self-efficacy. Compared to self-confidence, self-efficacy is subject-specify confidence in one's ability to perform a task. As Bandura (1986, p. 391) defines it, "self-efficacy is people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances". Students with low self-efficacy tend to have stronger need for approval about their actions in order to maintain their self-confidence. A polite tutor, not by simply flattering the student, but offering feedback in a way that meets student's need for approval, helps increase student's self-efficacy. While student's self-efficacy increases, a polite tutor can help maintain the self-efficacy. Self-efficacy affects choice of activities, effort, and persistence. People holding low self-efficacy for accomplishing a task may avoid it; those who believe they are capable are likely to participate. Especially when they encounter difficulties, efficacious students work harder and persist longer than those with low self efficacy.

Self-efficacy is strongly related to two of the motivation indexes: effort and task persistence (Bandura and Cervone, 1983, 1986: Schunk, 1995). Learners with high self-efficacy beliefs are likely to exert effort in the face of difficulty and persist at a task when they have the requisite skills. Self efficacy has been linked to not just the quantity of effort, but the quality of effort in terms of the use of deeper processing strategies and general cognitive engagement of learning (Pintrich and Schrauben, 1992). Pintrich and De Groot (1990) found that junior high students high in efficacy were more likely to report using various cognitive and self-regulatory learning strategies. In a series of experimental studies, Schunk (1982, 1983a, 1932b, 1983c, 1983d, 1984, 1996) found that students with stronger self-efficacy mastered various academic tasks better than did students with weaker self-efficacy. In addition, these studies showed that efficacy was a significant predictor of learning and achievement, even after prior achievement and cognitive skills were taken into consideration.

This pattern could be explained by the interdependent relationship between self-efficacy beliefs and self-regulated learning strategies (Gaskill and Hoy 2002). Self-regulated learners have been described as individuals who set goals, plan, organize activities, monitor their progress, and evaluate how they are doing (Azevedo and Cromley, 2004; Zimmerman and Schunk, 2001). Level of self-efficacy predicts student's use of cognitive strategies and self-regulation. Use of these strategies then predicts academic achievements (Zimmerman, 1995). As students increase their use of learning strategies and their academic performance improves, their academic self-efficacy increases. Both self-regulated learning and self-efficacy judgments require a similar series of cognitive and metacognitive process, including self-observation, self-judgments, and self-reaction. In judging the discrepancy between "where I am" and "where I want to be", the student ascertains whether current efforts have fallen short of the goal. If so, self-regulated learners can exert more effort or even try another strategy. If self-regulated learners see their progress toward a goal as acceptable, not only do they anticipate the satisfaction of reaching the goal, they also feel enhanced self-efficacy and motivation.

2.4. Politeness and Motivation

Although politeness theory and motivation theory come out of distinct literatures, their predictions regarding the choice of tutorial interaction tactics are broadly consistent. This is not surprising, since the wants described by politeness theory have a clear motivational aspect; negative face corresponds to autonomy, and positive face corresponds somewhat to approval and self-esteem in educational settings. Therefore, we are led to think that tutors may use politeness strategies not only for minimizing the weightiness of face threatening acts, but also for indirectly supporting the student's motivation. For instance, the tutor may use positive politeness for promoting the student's positive face (e.g. his desire to be recognized as a successful learner), and negative politeness for supporting the student's negative face (e.g. his desire to make his own choices).

3. Hypothesis

In conclusion, based on Brown and Levinson's politeness theory, we hypothesize that an intelligent tutor that adopts appropriate politeness strategies can promote learner's motivation and in turn improve learning results. Therefore we predict that students who learn to use the Virtual Factory Training System simulation with an on-screen agent who uses polite requests will learn better than students who learn with an on-screen agent who uses direct requests. More specifically, we hypothesize:

H1: students who learn to use the Virtual Factory Training System simulation with an on-screen agent who uses polite requests will do better than students who learn with an on-screen agent who uses direct requests on post learning test.

H2: students who learn to use the Virtual Factory Training System simulation with an on-screen agent who uses polite requests will report higher motivational state than students who learn with an on-screen agent who uses direct requests.

4. The Wizard-of-Oz Experiment System

To evaluate the intervention tactics, we created a Wizard-of-Oz experiment system aimed at teaching students how to use the Virtual Factory Teaching System (VFTS). The system includes student's and experimenter's interfaces. The Wizard-of-Oz interface, as part of experimenter's interface, enables a human tutor to use the politeness model to generate the tutorial dialog for those tactics. Behind the scenes, Plan Recognition and Focus of Attention modules monitor the student's progress and help the experimenter analyze student behavior. From the Wizard-of-Oz interface, the tutor selects an item (e.g., "copy_factory") then chooses from among a set of communicative acts associated with the current pedagogical goal (e.g., "indicate action & explain reason" or "tell how to perform action") and generates an intervention. The intervention is sent to the agent window on the student interface. An animation engine (Shaw et al., 2004) produces the gestures and a text-to-speech synthesizer synthesizes speech from the text.

4.1. The student's and experimenter's interface

The student's interface is shown in Fig. 1. The Virtual Factory Teaching System (VFTS) (Dessouky et al., 2001) is displayed on the left. The VFTS is a web-based factory modeling and simulation system developed for industrial systems engineering students for a product inventory and management class. Students model factories by specifying properties of machines and products, forecasting product demand, planning product release, and simulating product production for their factory.

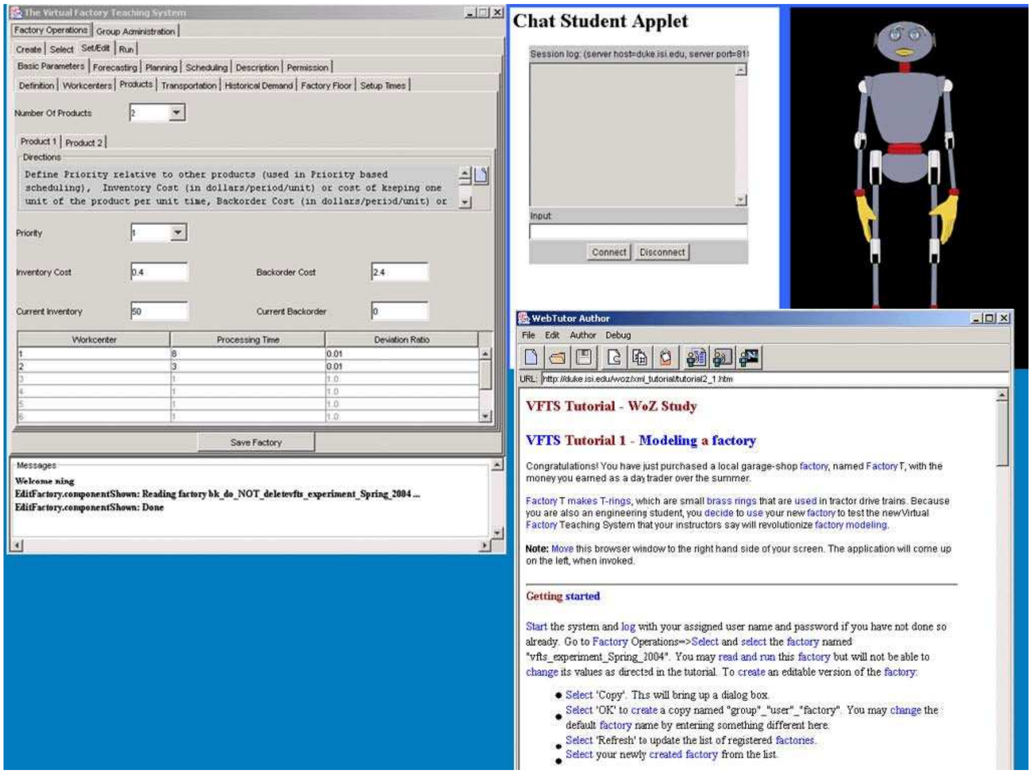

Figure 1. Student's screen during the Wizard-of-Oz experiment.

At the top right of the interface is the *Agent Window*, which contains a chat window for communicating with the agent (or human tutor during the Wizard-of-Oz experiment) and an animated character that generates speech and gestures. On the bottom right is a browser containing the tutorial. Students follow the tutorial on the VFTS using a browser that supports inline questioning, with all questions directed to the Wizard-of-Oz tutor. The tutorial teaches the concepts and skills needed to understand and use the VFTS.

All student keyboard and mouse inputs are sent back to the server for the Plan Recognizer to analyze. The Plan Recognizer compares the student's action with the expected action, categorizes the action as 1) progress toward the goal, 2) an error or inappropriate actions, or 3) a step performed in wrong order, and indicates what the next step should be.

A web camera is placed on top of the monitor to track learner's gaze. This, combined with keyboard and mouse information, is used by a Focus of Attention model to infer which window is the learner's current focus. The focus of attention information helps the system to determine the learner's intended goal.

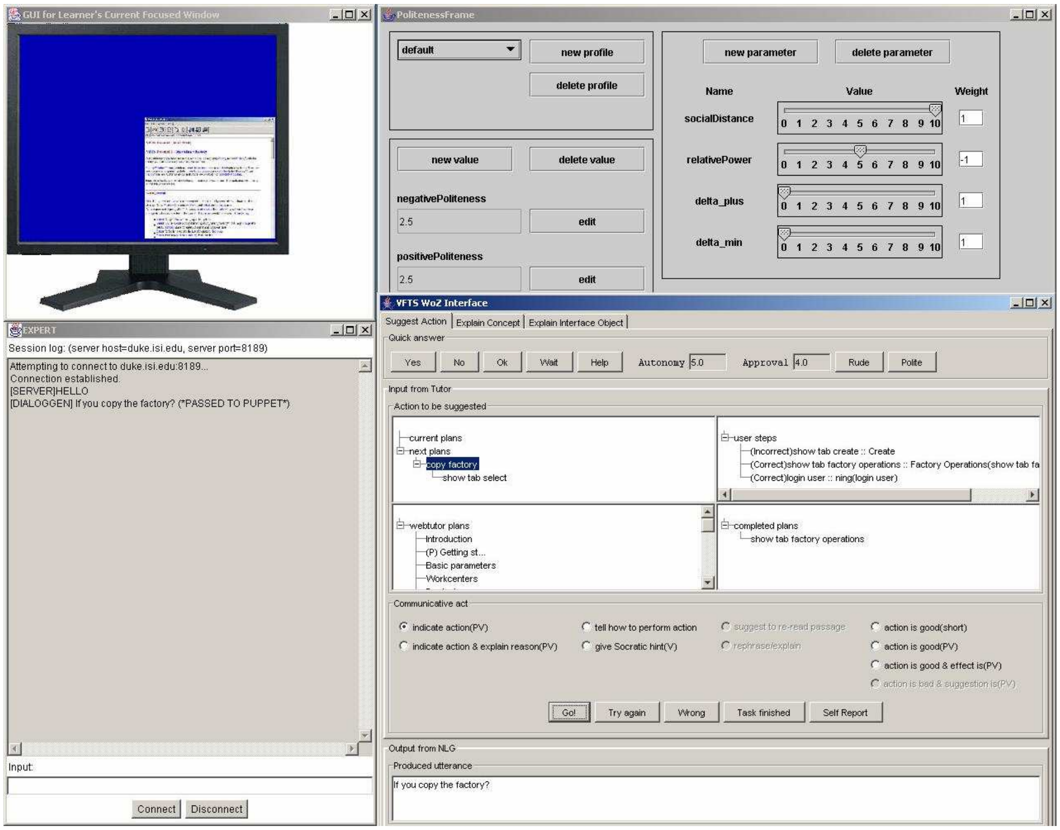

Figure 2. The tutor's screen during the Wizard-of-Oz experiment.

Fig. 2 shows the experimenter's screen. On the bottom right is the Wizard-of-Oz interface. This semi-automated interface enables a human tutor to select tactics and use the politeness model to generate the tutorial dialog for those tactics. The main panel, with its four student activity windows, is on the lower right hand side. The windows display the student's (a) completed plans (a group of actions that together achieve a specific goal), (b) current action, and (c) the inferred next action from the VFTS Plan Recognizer. A fourth window displays the paragraph of the tutorial that is currently visible, to help the experimenter infer the learner's goal.

To communicate with the student, the tutor selects an item in the student activity window (e.g., “copy_factory”) then chooses from among a set of communicative acts associated with the current pedagogical goal (e.g., “indicate action & explain reason” or “tell how to perform action”). The Dialogue Generator then automatically generates a feedback intervention based on tutor’s selection (e.g, “How about we copy the factory?”). The intervention is sent to the Agent Window on the student interface. An animation engine (Shaw et al., 2004) produces the gestures and a text-to-speech synthesizer synthesizes speech from the text.

The window at the top right is the interface for setting the parameters of the politeness model. The parameters are initialized at the beginning of a tutorial session and are not manually modified during the session. The window at the top left shows which area of the screen the learner is focusing on, inferred by the Focus of Attention model. This enables the tutor to tell whether the learner is currently reading the tutorial, working on the VFTS or reading/typing in the Agent Window.

4.2. Dialogue Generator

The Dialogue Generator is used to generate feedback interventions based on tutor’s selection of communicative act and politeness strategies. The politeness module and dialog generator are part of an overall pedagogical agent architecture, which includes other modules to detect learner focus of attention, including eye gaze tracking, and to recognize plans that the learner is carrying out, as described by Qu et al. (2004). This information can be provided to a remote tutor, operating a Wizard-of-Oz interface, so that the tutor can decide when it is appropriate for the agent to interact with the learner.

Based upon analyses of the videotaped study, Sander Kole and Wauter Bosma developed a natural language generator for producing appropriate interaction tactics. The generator takes as input a set of language elements – short noun phrases and short verb phrases in the target domain – and predicates describing the desired dialog move. It chooses an utterance pattern that matches the dialog move predicates most closely, instantiates it with the language elements, and synthesizes an utterance, which is then passed to the guidebot persona for uttering using text-to-speech synthesis.

```
<move>
    <!-- 7S P1 47:11 T -->
    <!-- So number 2, the number of seasons may not be 2 then.-->
    <predicate role="initiating" move="all" name="action1"/>
    <predicate role="initiating" move="all" name="noun1"/>
    <predicate role="initiating" move="hint" name="suggest"/>
    <predicate role="initiating" move="inform" name="identify"/>
    <predicate role="initiating" move="reason" name="explain"/>
    <template>
            So
            <nounphrase case="object" type="parameter" name="noun1.nounphrase1"/>
            may not
            <verbphrase type="parameter" name="action1.action1" form="infinite"/>
            .
    </template>
</move>
```

Figure 3. An example dialog move template

The underlying generation scheme utilizes DISCOUNT as a way of classifying dialog moves. It operates on a set of move templates, each of which includes a set of DISCOUNT predicates and a template for expressing the move in natural language. The templates specify slots for language elements, which are filled from the language elements supplied to the generator. The move templates and language elements are specified using XML syntax and all defined in one language definition file. Fig. 3 shows an example move from the language definition file. The moves are based upon utterances found in the dialog transcripts; the comments at the top of the move template show the original utterance and the transcript and time code where it was found. The move template may classify the move in multiple ways, reflecting the fact that the same utterance may have multiple communicative roles, and different coders may code the same utterance differently.

Using this generation framework, it is possible to present the same tutorial comment in different ways. For example, a suggestion to perform an action, such as saving the current factory description, can be stated either directly (e.g., “Save the factory now”), as a hint, (“Do you want to save the factory now?”), as a suggestion of what the tutor would do (“I would save the factory now”), as a suggestion of a joint action (“Why don’t we save our factory now?”), and etc.

4.3. Politeness Model - Extend Brown and Levinson’s Politeness Model

Having defined this set of dialog moves and having implemented a generator that can produce them, the next challenge is to determine which tactic to employ in which circumstances. How does the choice of interaction tactic depend upon the learner, the topic being discussed, and the state of the social interaction between the learner and the tutor?

The politeness theory of Brown and Levinson (1987) helps provide a rationale for these decisions. Although the Brown and Levinson model is not specifically aimed at modeling tutorial dialog, it provides a good means of accounting for variability in tutorial dialog. The interaction tactics observed in the recorded dialogs, when other than *bald-on-record* statements, have the effect of mitigating face threats. Since offers of advice and requests to perform actions are face threatening acts, the theory predicts that tutors will employ face mitigation strategies for these kinds of interactions, but not for other dialog moves such as explanatory comments. This is consistent with the observed data. The theory predicts that the incidence of face threat mitigation strategies will decrease as tutor and learner interact for longer periods of time. This trend is also observable in the data that we have collected; the incidence of bald-on-record tactics was greater in the follow-on tutorial sessions than in the initial sessions.

We extend the Brown and Levinson model in certain respects. First, whereas Brown and Levinson's politeness model assigns a single numeric value to each face threat, we extend their model to consider positive face threat and negative face threat separately. This enables us to select a redressive strategy that is appropriate to the type of face threat. For example, if an FTA threatens negative face but not positive face, then the politeness model should choose a redressive strategy that mitigates negative face threat; in contrast the basic Brown and Levinson politeness model would consider a redressive strategy aimed at positive face to be equally appropriate. Second, we allow for the possibility that the tutor might wish to explicitly enhance the learner's face, beyond what is required to mitigate immediate face threats. For example, if the tutor judges that the learner needs to feel more in control, he or she will make greater use of redressive strategies that augment negative face.

Altogether, the amount of face threat redress is determined by the following formulas, which are slight elaborations of the weightiness formulas proposed by Brown and Levinson:

$$W_{x+} = D(T,S) - P(T,S) + R_{x+} + \Delta_{+}$$

$$W_{x-} = D(T,S) - P(T,S) + R_{x-} + \Delta_{-}$$

Here W_{x+} and W_{x-} are the amounts of positive and negative face threat redress, respectively, T represents the tutor and S represents the student. D(T,S) is the social distance between the tutor and the student, and P(T,S) is the amount of social power that the tutor has over the student. R_{x+} is the inherent positive face threat of the communicative act (e.g., advising, critiquing, etc.,), R_{x-} is the inherent negative face threat of the act, Δ_{+} is the amount of augmentation of positive face desired by the tutor, and Δ_{-} is the desired augmentation of learner negative face.

Additional factors clearly need to be taken into account besides politeness theory in order fully account for the influence of interaction tactics on learner motivation. For example, politeness theory per se does not explain the relative infrequency in our data of comments aimed solely at positive face, e.g., expressions of praise. In our analyzed dialogs, positive praise is confined to the ends of VFTS sessions, when the learner has completed the assigned tasks. One way to account for this is to note that learners are motivated not just by positive face, i.e., to be approved of by the tutor, but by a desire for self-efficacy, i.e., to approve of their own performance. Since VFTS tasks take some time to complete, it is difficult to tell whether the learner is doing well until after the learner has worked on the task for a long of time. If a learner recognizes this, then frequent praise from the tutor might be regarded as insincere. This is an account that needs to be tested in other domains, where there are more frequent opportunities to evaluate learner work.

4.4. Politeness Model - Implementation

Mattijs Ghijsen and Herwin van Welbergen have developed a politeness module that implements the politeness / motivation model described above, and interfaces to the natural language generator. The combined dialog generator takes as input the desired utterance type, language elements, and a set of parameters governing face threat mitigation (social distance, social power, and motivational support) and generates an utterance with the appropriate degree of face threat redress.

The utterance types are classified in accordance with Bloom’s (1956; Anderson et al., 2001) taxonomy of educational goals. Bloom categorizes instructional actions into three groups: cognitive, concerning the development of intellectual abilities and skills; affective, comprising interests, attitudes, and values; and psycho-motor, regarding the manipulative or motor- skill area. The objectives most relevant to the VFTS are from the cognitive category: Knowledge, i.e. the recall of specifics, universals, methods and processes – such as mastering the concept of forecast, or the planning process; Application, i.e. the use of abstractions in concrete situations – for instance the application of a specific forecasting method to the simulated factory; and Synthesis, i.e. the putting together of elements and parts so as to form a whole –such as producing a plan of operations to perform on the VFTS interface.

These cognitive goals, applied to the set of interface objects in the VFTS interface, and to the concepts and tasks described in the tutorial materials for the VFTS, determine the set of possible communicative acts that the dialog generator needs to generate. The repertoire of utterance patterns and language elements was extended as needed in order to cover this set.

To choose the appropriate interaction tactic, the politeness generator first computes target positive and negative politeness values for the desired utterance. The positive and negative politeness values are computed in order to counteract the weightiness of the face threat, as well as to achieve additional motivational influence, as indicated in the formulas in the previous section. Social distance, social power, and motivational influence are all parameters that are supplied to the politeness generator and are potentially adjustable.

Once the target politeness values are chosen, the generator chooses from a library of natural language templates one that matches the target politeness values most closely. Each template, as in Fig. 1, is assigned a positive and negative politeness value. A template is chosen that minimizes the sum of the distances between desired and chosen politeness values, for both positive and negative politeness. When multiple templates have an appropriate politeness value one is chosen that matches the greatest number of move predicates.

Table 1. Face threat redress strategies for different utterance types

Utterance Type	Politeness strategies
Suggest action	Bald on record, conventional indirectness, joint goal, student goal, question, suggestion, tutor goal
Explain concept	Bald on record, positive politeness, attend to hearer, student’s goal, impersonalize, off record
Explain tutorial	Bald on record, tutor goal, joint goal, suggestion
Suggest interface operation	Bald on record, conventional indirectness
Explain interface object	Bald on record
Socratic hint	Socratic hint
Action feedback	Bald on record, positive politeness

To apply this politeness module, it is necessary to assign politeness values to each template in the library. To assign these values, we grouped templates according to the politeness strategies that they exhibit, as shown in Table 1. Examples of each politeness strategies used are shown in Table 2. These categories were derived from analysis of the background dialog transcripts, and then mapped onto the general strategies identified by Brown and Levinson such as bald-on-record. We then assumed that all templates employing a given strategy should be assigned the same level of politeness—not strictly true, since perceived politeness depends upon context, but adequate as a first approximation. We then created a questionnaire containing examples of different politeness strategies, and had subjects evaluate each example in terms of negative and positive politeness. The mean negative and positive politeness scores were then assigned to the templates in the library.

Table 2. Face threat redress strategies used and examples

Politeness strategies	Example
Attend to hearer	As you might already know, a stochastic model randomizes the planning methodology.
Bald on record	Read the paragraph more carefully.

Conventional indirectness	The system is asking you to set the parameter.
Impersonalize	It might be useful to read the paragraph.
Joint goal	Why don't we read the paragraph more carefully.
Off record	It's important to read the paragraph.
Positive politeness	It is usually hard to understand the paragraph.
Question	Do you want to set the parameter?
Socratic hint	Take a look at the paragraph and see what you think.
Student goal	You might want to read the paragraph.
Suggestion	I suggest that you read the paragraph more carefully.
Tutor goal	I would read the paragraph more carefully.

5. Experiment Method

5.1. Participants and Design

Fifty-one students participated in the study, including 17 students from the University of Southern California (USC) and 34 students from the University of California, Santa Barbara (UCSB). The students from USC were either engineering graduate or undergraduate students, and all were male. The students from UCSB were mostly undergraduate students from introductory psychology classes. Among subjects from UCSB, 17 of them are female and 17 are male. Five students from USC participated in a pilot study, which allowed us to test the experiment set-up. Four other students were excluded due to technical difficulties during the experiment or extremely high ability to complete the task without tutoring. In the main experiment, 20 students served in the polite group and 17 served in the direct group.

The experiment system was configured so that the on-screen agent produced polite suggestions (Polite version) or direct suggestions (Direct version) in response to learner queries. In the Polite version, positive and negative politeness values varied randomly in a moderate to high range, causing the tutor to use politeness in a variety of ways both in giving hints and in providing feedback. In the Direct version, positive and negative politeness values were fixed at minimum values, forcing the tutor to communicate directly and not allowing for mitigation of face threat. In all other respects the two versions were identical.

5.2. Measures

Pre-questionnaire packet Pre-questionnaire packet includes a background questionnaire and a personality questionnaire. The background questionnaire solicited basic demographic information, including the learner's age, gender, major, learning style, experience with computers, and background in engineering. The personality questionnaire consisted of 20 questions from the International Personality Item Pool (2003), intended to measure self-esteem, need for cognition, extroversion, and optimism.

Post-questionnaire packet Post-questionnaire packet includes a motivation questionnaire and a learning outcome test. The motivation questionnaire evaluated the learner's motivation and perception of the Wizard-of-Oz tutor. The learning outcome test tested how much students have learnt from the Virtual Factory Training System.

Computer skills, engineering background, preference for indirect help All three scales were measured in the pre-questionnaire. Computer skills was measured using a single item "My level of SKILL in using a computer is: " with a 5 point metric (1 = Very low; 5 = Very high). Students were asked about their engineering background by answering "Yes" or "No" to the statement "I study/work in an engineering discipline." To measure students' preference for indirect help, we asked them to evaluate the statement "I prefer to be told about my mistakes in a DIRECT MANNER or I prefer to be told about my mistakes in a POLITE / INDIRECT MANNER (e.g.

direct: ‘This is wrong’, polite: ‘Perhaps you could try this…’)”. This scale was measured using a 5 point metric (1 = Strongly prefer direct; 5 = Strongly prefer indirect).

Self-esteem, optimism, need for cognition, extroversion These four scales are measured using items from the International Personality Item Pool (2003). These scale were measured with a 5 point metric (1 = Very Inaccurate; 5 = Very Accurate). These items were issued in the pre-questionnaire packet.

Tutor helpfulness scale. We constructed a 8-item tutor helpfulness scale (coefficient alpha = .73), included in the post-questionnaire. This scale was measured with a 5 point metric (1 = Strongly Disagree; 5 = Strongly Agree). Sample items include: “The tutor made it easier for me to follow each step.” and “The tutor helped me to identify my mistakes.”

Self-efficacy scale. Self-efficacy was measured using a 3-item scale (coefficient alpha = .71) we constructed, presented in the post-questionnaire. This scale was measured with a 5 point metric (1 = Strongly Disagree; 5 = Strongly Agree). Sample items include: “I feel confident of my ability to complete another VFTS problem of the SAME LEVEL.”

Sense-of-Control scale. We constructed a 5-item sense-of-control scale (coefficient alpha = .63), included in the post-questionnaire. This scale was measured with a 5 point metric (1 = Strongly Disagree; 5 = Strongly Agree). Sample items include: “I felt I was given total freedom in making decisions at every step.”

Interest, liking of tutor, wiliness to work with tutor again. We indexed the students’ interest using the item “I think my interest increased as the tutorial progressed.” Liking of tutor was index using item “I like the tutor.” Work with tutor again was indexed using item “I would like to work with the tutor again.” These items were presented to the subjects in the post-questionnaire. All three scales were measured with a 5 point metric (1 = Strongly Disagree; 5 = Strongly Agree).

Learning scale Students were asked to take a test which contains 35 questions about the Virtual Factory Training System. Sample items include: "What does planning function do?" Within the 35 questions, 17 were easy questions and 18 were difficult questions. These questions were included in the post-questionnaire packet. Students were presented with the test after filling out the motivation questionnaire fro the post-questionnaire packet. All questions are multiple choice questions. Each question is graded for 1 point. There is a total of 35 points.

Frequency of tutor intervention Frequency of tutor intervention was measured by counting how many times tutor has intervened, then divide it by the length of the session.

5.3. Apparatus

The apparatus consisted of a Dell desktop computer with a 20 inch color monitor at the student end and a Dell desktop computer with dual 20 inch color monitors at the tutor end.

5.4. Procedure

Students were randomly assigned to either a polite treatment or a direct treatment. Students were seated at a desk with a computer and were tested individually. Two pre-tests were administered: the background questionnaire and the personality questionnaire. To provide experience with the VFTS, the experimenter provided a brief introduction to the game and asked the learner to solve a basic assembly-line design problem. The learner then interacted with the VFTS – using either the polite or direct tutor – for approximately 35 minutes. Then, two post-test questionnaires were administered: the motivation questionnaire and the learning outcome test.

6. Results

Since the experiment materials and the procedures were identical, we combined the data collected from the experiments carried out in the summer of 2004 at USC and in the fall of 2004 at

UCSB. A two-way analysis of variance (ANOVA) using condition (polite vs. direct) and experiment location (USC vs. UCSB) as between-subject factors showed that there was no significant interaction between condition and experiment location, $F(1, 33)=0.003$, $p=0.957$. Therefore, we focused on comparing the polite and direct conditions using two-tailed t-tests on the combined data (with alpha at $p < .05$).

We grouped 37 students into two groups: 20 students in the polite and 17 in the direct group, based on the treatment they received. For each group, we calculated the average score of the learning outcomes tests and applied Student's t-test to analyze the differences between the groups.

6.1. Do Students Learn Better with a Polite Tutor?

Overall, students who received the polite treatment scored better ($M_{polite}=19.45$, $SD_{polite}=5.61$) than students who received the direct treatment ($M_{direct}=15.65$, $SD_{direct}=5.15$). This is consistent with what we found in our previous pilot study (Wang et al., 2005). A t-test revealed that the difference between the groups was statistically significant, $t(35)= 2.135$, $p=0.040$, $d=0.73$. The scores range from 7 to 31 with an average of 17.70.

Is Politeness Particularly Effective for Certain Types of Students?

Even though the politeness strategies had a positive impact on student learning overall, we are interested in whether certain subgroups of students were particularly influenced by politeness strategies. To investigate this issue we partitioned students into subgroups based on their responses to the background and personality questionnaires, and then compared the mean score on the learning posttest by polite and direct groups for each subgroup. In particular, we partitioned students based on their self-reported computer skills, engineering background, preference for direct help, and various personality traits.

Is Politeness Particularly Effective for Students with Lower Computer Skills?

From students' self-ratings of their computer skills, we found that almost all students rated their computer skills either average or above average. We then grouped students into 2 groups, 19 with average computer skills and 17 with above average computer skills (one student with below average computer skill was not included due to technical difficulties). Overall, students with above average computer skills performed better than students with average computer skills. This may because our test-bed, VFTS, is a relatively complicated computer-based teaching system. Better computer skills help students understand the basic concepts of operations in VFTS. But for students with average computer skills, those who received polite treatment (M_{polite}=18.42, SD_{polite}=5.02) performed marginally better than those who received direct treatment (M_{direct}=14.14, SD_{direct}=3.39), $t(17)= 1.993$, $p=0.063$. We did not observe this difference within students with above average computer skills. In this case the tutor, either polite or direct, has less impact on students learning performance. On the other hand, students with poorer computer skills relied more on tutor to help them through the learning task.

Is Politeness Particularly Effective for Students without Engineering Background?

We asked the students whether they work or study in an engineering discipline. Within the students with no engineering background (28 students), we found a major difference between the polite (M_{polite}=18.80, SD_{polite}=5.80) and direct groups (M_{direct}=14.08, SD_{direct}=4.37), $t(26)=2.403$, $p=0.024$. We did not find much difference within engineering students (9 students). VFTS is a system built for Industrial Engineering students. For students who do not work/study in an engineering discipline, such as psychology students, performing tasks in the VFTS could be much more challenging. This is consistent with our hypothesis that students with better ability to perform the task relied less on the tutor.

Is Politeness Particularly Effective for Students who Prefer Indirect Help?

Direct help consists of tutor feedback that is devoid of any politeness strategy, while Indirect help consists of feedback that is phrased using politeness strategies. Based on students' preference of direct or indirect help, we grouped them into 3 groups: 15 preferred direct help, 13 preferred indirect and 9 had no preference. For students who preferred direct help or did not have any preference, we did not observe any difference made by the polite tutor. In contrast, for students who specifically reported their preference for indirect help, the polite tutor had a significant effect on their learning performance (M_{polite}=20.43, SD_{polite}=5.74, M_{direct}=13.00, SD_{direct}=4.56), $t(11)= 2.550$, $p=0.027$.

Is Politeness More Effective for Students with Particular Personality Characteristics?

We measured 4 personality traits: self-esteem, optimism, need for cognition and extroversion. On self-esteem and optimism, we found our sample distribution was skewed – most subjects had a high self-esteem and were optimistic. We grouped students based on their level of need for cognition and extroversion. On overall learning results, we did not find interactions between these two personality traits and use of politeness strategies. However, on students' performance on learning difficult concepts, there were some interesting differences between the polite and direct groups.

For the 20 students who scored high on extroversion, we found out that the polite tutor helped students to learn difficult concepts more than the direct tutor did (M_{polite}=10.46, SD_{polite}=2.07, M_{direct}=8.56, SD_{direct}=1.59, $t(18)= 2.259$, $p=0.037$). We found the same difference for 22 students who scored high on need for cognition (M_{polite}=10.00, SD_{polite}=1.48, M_{direct}=8.18, SD_{direct}=2.52, $t(20)= 2.061$, $p=0.053$). Students with high need for cognition are probably more motivated to learn difficult concepts. Students with high extroversion are more open to discuss their problems with the tutor when trying to understand difficult concepts. This leads us to conclude that, when learning materials are relatively challenging, students with either high extroversion or need for cognition were more likely to be influenced by politeness strategies.

Do Students Like the Polite Tutor More than the Direct Tutor?

On the post-questionnaire, students were asked whether or not they liked the tutor. We grouped students into 2 groups based on their answers: 20 students liked the tutor and 17 did not or had no preference. We did not find statistical significance between the polite and direct group among students who did not like the tutor or did not have a preference. However, among students who reported that they liked the tutor, we found that students who worked with the polite tutor performed better than students who worked with the direct tutor (M_{polite}=20.33, SD_{polite}=5.26, M_{direct}=15.50, SD_{direct}=4.96, $t(18)$=2.058, p=0.054), especially on learning difficult concepts (M_{polite}=11.08, SD_{polite}=2.61, M_{direct}=8.38, SD_{direct}=1.77, $t(18)$=2.559, p=0.020). However, whether students like the tutor or not is not as accurate a predictor of learning performance as preference for direct help.

Do Students Want to Work with the Tutor Again?

We also asked students in the post-questionnaire whether or not they would like to work with the tutor again. We grouped students into 2 groups based on their answers: 22 students indicated they would like to work with tutor again and 15 did not or had no preference. We did not find statistical significance between polite and direct group within students who would not like to work with the tutor again or did not have a preference. However, within students who reported a desire to work with the tutor again, we found that students who worked with the polite tutor performed better on learning difficult concepts than students who worked with the direct tutor (M_{polite}=10.92, SD_{polite}=2.75, M_{direct}=8.50, SD_{direct}=1.51, $t(20)$= 2.482, p=0.022).

Does Frequency of Tutor Intervention Moderate the Politeness Effect?

Tutor attentiveness could be a factor that affected students' learning outcomes. During the experiment, tutor attentiveness was balanced under both experimental conditions. However, how many times of tutor gave feedback to the students depended on the students' need. We grouped students into two groups based on the amount of tutor feedback: 11 students in low and 26 students in average-to-high groups. On average students spent 36 minutes on the VFTS. We considered a low interaction as less than 20 feedbacks during the experiments, while average to high is 20 or more feedbacks. Our hypothesis is that when the number of tutor interventions is low, politeness would have less effect on students' learning. The result confirmed our hypothesis. We found that when the tutor's interventions were low, the polite tutor did not affect students learning as much. However, when the tutor's interventions were average to high, the polite tutor made a big difference (M_{polite}=18.21, SD_{polite}=5.60, M_{direct}=13.83, SD_{direct}=3.35, $t(24)$= 2.365, p=0.026).

6.2. Motivation Results

Does Politeness Increase Students' Self-Efficacy or sense-of-control?

We compared students' answer to the post-questionnaire and found no significant difference on self-report of self-efficacy (M_{polite}=2.03, SD_{polite}=0.76, M_{direct}=2.13, SD_{direct}=0.61, $t(35)$= –0.45, p=0.28) and sense-of -control (M_{polite}=2.93, SD_{polite}=0.54, M_{direct}=3.04, SD_{direct}=0.81, $t(35)$= -0.47, p=0.11) between Polite and Direct condition.

Does Politeness Increase Students' Interest?

Again, we didn't find statistically significant difference between students received polite and direct treatment on self-reported interest (M_{polite}=3.40, SD_{polite}=1.19, M_{direct}=3.35, SD_{direct}=0.86, $t(35)$=0.14, p=0.89).

Does Politeness Increase Students' Perception of Tutor Helpfulness?

From the post-questionnaire, we found no statistically significant difference between students from polite and direct group on tutor helpfulness (M_{polite}=2.05, SD_{polite}=0.47, M_{direct}=2.45, SD_{direct}=0.56, $t(35)=-2.25$, p=0.49). This indicates that politeness didn't increase students' perception of tutor's helpfulness.

6.3. Discussion on Motivation Result

In our study, we used a self-report questionnaire to measure students' motivation change. Self-report questionnaires are easy to administer and score, but problems often arise when inferences must be drawn about students' responses. The general validity of the questionnaire can be called into question. In addition, there are other concerns about self-reports; for example, students could be giving socially acceptable answers that do not match their beliefs. In some of our other studies about personalized tutor, we find that students rate a personalized and non-personalized lesson as equal in interest to them but they do better on a cognitive test with the personalized version. Perhaps people are just not very good at gauging their level of motivation or interest on a questionnaire. Behavioral measures are probably more valid than questionnaires for assessing people's level of motivation.

7. Discussion

In this paper, we reported on the effect of politeness strategies on students' learning performance, which we call the politeness effect. Across all students, a polite agent, compared to a direct agent, had a positive impact on students' learning outcomes. Richer interaction amplified this effect. In particular, for students with need for indirect help or who had lower ability for the task, the polite agent was much more effective than the direct agent. For students with high extroversion or who were more open to communication with the agent, the polite agent helped them better understand difficult concepts. Making students like the agent appeared to help students learn. It was not the physical appearance of the agent, but rather the helpfulness and feedback style adopted by the agent that created the effect.

Tutorial advice and feedback are certainly not the only places to apply politeness strategies. In our study, we artificially restricted the use of politeness in tutorial interaction to ensure that the polite condition and the direct condition were as similar as possible. In real human-human interaction, people employ a range of additional strategies to build rapport and react empathetically. These strategies have been modelled in other learning domains (Bickmore, 2003; Johnson et al. 2004), and could complement the strategies studied here. We did not include them in this particular study because it would have increased the frequency of tutorial interaction, making it harder to tell whether the politeness effect was really a consequence of the frequency of interaction rather than the politeness strategies themselves. The study presented here shows the politeness effect on learning after a short-term of interaction with. However, the politeness effect goes beyond short-term tutoring. Long-term tutoring should have the same if not stronger effect.

The politeness effect goes beyond the engineering training system we demonstrated here. Other studies we have conducted showed that politeness strategies do occur pervasively in other domains such as second language learning (Johnson et al., 2004). However, more research will have to be done to study their effects on learning outcomes in other domains.

We recommend that developers of intelligent tutors and pedagogical agents examine the tutorial messages that their tutors are generating from a politeness perspective, as politeness may have an impact on the tutors' effectiveness. Meanwhile, more research needs to be done to study how the politeness effect applies in other learning contexts, and investigate other aspects of social actor modelling that go beyond the tactics studied here.

8. Acknowledgement

We would like to thank the various people who have contributed to the Social Intelligence Project, including Wauter Bosma, Maged Dessouky, Mattijs Ghijsen, Sander Kole, Kate LaBore, Hyokeong Lee, Helen Pain, Lei Qu, Sahiba Sandhu, and Herwin van Welbergen. This work was supported in part by the National Science Foundation under Grant No. 0121330, and in part by a grant from Microsoft Research. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or any other funder.

9. References

Aist, G., Kort, B., Reilly, R., Mostow, J., & Picard, R., 2002. Adding human-provided emotional scaffolding to an automated reading tutor that listens increases student persistence. In S. A. Cerri, G. Gouardères, & F. Paraguaçu (Eds.), Proceedings of the 2002 International Conference on Intelligent Tutoring Systems (pp. 320-329). Berlin: Springer.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C., 2001. A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York: Longman.

André, E., Rist, T., Müller, J., 1998. Guiding the user through dynamically generated hypermedia presentations with a life-like character. In Proceedings of the 1998 International Conference on Intelligent User Interfaces (pp. 21-28). New York: ACM Press.

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

Azevedo, R., & Cromley, J.G., 2004. Does training on self-regulated learning facilitates students' learning with hypermedia. Journal of Educational Psychology, 96, 523-535

Bandura, A., 1986. Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. & Cervone, D., 1983. Self-evaluative and self-efficacy mechanisms governing the motivational efforts of goal systems. Journal of Personality and Social Psychology, 45, 1017-1028

Bandura. A. & Cervone, D., 1986. Differential engagement of self-reactive influences in cognitive motivation. Organizational Behavior and Human Decision Processes, 38, 92-133.

Baylor, A.L., & Ebbers, S., 2003. Evidence that multiple agents facilitate greater learning. In U. Hoppe, F. Verdejo, & J. Kay (Eds.), Proceedings of 2003 International Conference on Artificial Intelligence in Education. Amsterdam: IOS Press.

Bickmore, T., 2003. Relational Agents: Effecting Change through Human-Computer Relationships. Unpublished PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

Bloom, B.S. (ed.), 1956, Taxonomy of educational objectives: The classification of educational goals. Handbook I, cognitive domain. London: Longman.

Brophy, J., 1983. Conceptualizing student motivation. Educational Psychologist, 18, 200-215

Brown, P., Levinson, S.C., 1987. Politeness: Some universals in language use. New York: Cambridge University Press.

Cassell, J., & Bickmore, T., 2003. Negotiated collusion: Modeling social language and its relationship effects in intelligent agents. User Modeling and User-Adapted Interaction, 13, 89–132.

Cooper B., 2003. Care – Making the affective leap: More than a concerned interest in a learner's cognitive abilities. International Journal of Artificial Intelligence in Education, 13, 3-9.

De Vicente, A., & Pain, H., 2002. Informing the detection of the students' motivational state: An empirical study. In S.A. Cerri, G. Gouardères, & F. Paraguaçu, (Eds.), Proceedings of the 2002 International Conference on Intelligent Tutoring Systems (pp. 933-943). Berlin: Springer.

deCharms, R., 1968. Personal causation: The internal affective determinants of behavior. New York: Academic Press

Deci, E.L., 1971. Effects of externally mediated rewards on intrinsic motivation. Journal of Personality and Social Psychology, 18, 105-115.

Del Soldato, T., & du Boulay, B., 1995. Implementation of motivational tactics in tutoring systems. International Journal of Artificial Intelligence in Education, 6, 337-378.

Dessouky, M.M., Verma, S., Bailey, D., & Rickel, J., 2001. A methodology for developing a Web-based factory simulator for manufacturing education. IEEE Transactions, 33, 167-180.

Fisher, C. D., 1978. The effects of personal control, competence, and extrinsic reward systems on intrinsic motivation. Organizational Behavior and Human Performance, 21, 273-288.

Gaskill, R.J., & Hoy, A.W., 2002. Self-efficacy and self-regulated learning: The dynamic duo in school performance. In J. Aronson (Ed.), Improving academic achievement: Impact of psychological factors on education (pp. 185-208). New York: Academic Press

Graesser, A. C., Moreno, K., Marineau, J., Adcock, A., Olney, A., & Person, N., 2003. AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head? In U. Hoppe, F. Verdejo, and J. Kay (Eds.), Proceedings of International Conference on Artificial Intelligence in Education (pp. 47-54). Amsterdam: IOS Press.

Heylen, D., Nijholt, A., op den Akker, R., Vissers, M., 2003. Socially intelligent tutor agents. In Rist, T., Aylett, R., Ballin, D., Rickel, J. (Eds.), Proceedings of Social Intelligence Design Workshop (pp. 341-347). Berlin: Springer.

International Personality Item Pool, 2003. http://ipip.ori.org/ipip/

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.

Johnson, W.L., Rizzo, P., Bosma W., Kole S., Ghijsen M., & Welbergen H., 2004. Generating socially appropriate tutorial dialog. In E. André, L. Dybkjær, W. Minker. & P. Heisterkam. (Eds.): Proceedings of 2004 Affective Dialogue Systems, Tutorial and Research Workshop (pp. 254-264). Berlin: Springer.

Johnson, W.L., LaBore, C., & Chiu, Y.-C., 2004. A pedagogical agent for psychosocial intervention on a handheld computer. In Proceedings of AAAI Fall Symposium on Health Dialog System (pp. 64-70). Menlo Park, California: AAAI Press.

Johnson, W.L., Wu, S. & Nouhi, Y., 2004. Socially intelligent pronunciation feedback for second language learning. In Proceedings of International Conference on Intelligent User Interface workshop in Modeling Human Teaching Tactics and Strategies.

Kast, A., & Connor, K., 1988. Sex and age differences in response to informational and controlling feedback. Personality and Social Psychology Bulletin, 14, 514-523.

Kort B., Reilly R., & Picard R.W., 2001. An affective model of interplay between emotions and learning: Reengineering educational pedagogy – building a learning companion. In T. Okamoto, R. Hartley, K. Kinshuk & J. Klus (Eds.), Proceedings of International Conference on Advanced Learning Technologies. Los Alamitos, CA: IEEE Computer Society.

Lepper, M.R., Drake, M., & O’Donnell-Johnson, T.M., 1997. Scaffolding techniques of expert human tutors. In K.H.M. Pressley (Ed.), Scaffolding student learning: Instructional approaches and issues (pp. 108-144). New York, NY: Brookline Books.

Lepper, M.R., & Hodell, M., 1989. Intrinsic motivation in the classroom. In C. Ames & R. Ames (Eds), Research on motivation in education (Vol. 3, pp. 73-105). San Diego, CA: Academic Press.

Lepper, M.R., & Woolverton, M., 2002. The wisdom of practice: Lessons learned from the study of highly effective tutors. In J. Aronson (Ed.), Improving academic achievement: Impact of psychological factors on education (pp. 135-158). Orlando, FL: Academic Press.

Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J., 1993. Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), Computers as cognitive tools (pp. 75-105). Hillsdale, NJ: Erlbaum.

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., Bhogal, R. S., 1997. The persona effect: Affective impact of animated pedagogical agents. In S. Pemberton (Ed.), Proceedings of 1997 Conference on Human Factors in Computing Systems (pp. 3539-366). New York: ACM Press.

Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J., 2004. A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. Journal of Educational Psychology, 96, 389-395.

Mayer, R. E., Johnson, W. L., Shaw, E., & Sandhu, S., 2006. Constructing computer-based tutors that are socially sensitive: Politeness in educational software. International Journal of Human-Computer Studies, 64, 36-42.

Mayer, R. E., Sobko, K., & Mautone, P. D., 2003. Social cues in multimedia learning: Role of speaker's voice. Journal of Educational Psychology, 95, 419-425.

Miller C. (ed.), 2002 Etiquette for Human-Computer Work. AAAI Fall Symposium. AAAI Technical Report FS-02-02. AAAI Press.

Moreno, R. & Mayer, R. E., 2000. Meaningful design for meaningful learning: Applying cognitive theory to multimedia explanations. In Proceedings of 2000 World Conference on Educational Multimedia Hypermedia & Telecommunications (pp. 747-752). Charlottesville, VA: AACE Press.

Moreno, R., & Mayer, R. E., 2004. Personalized messages that promote science learning in virtual environments. Journal of Educational Psychology, 96, 165-173.

Moreno, R., Mayer, R. E., Spires, H., & 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.

Nass, C., Isbister, K., & Lee, E.-J., 2000. Truth is beauty: Researching embodied conversational agents. In Cassell, J., Sullivan, J., Prevost, S. & Churchill, E. (Eds.), Embodied Conversational Agents (pp. 374-402). Cambridge, MA: The MIT Press.

Piaget, J. 1952. The origins of intelligence. New York: International University Press.

Pilkington, R.M., 1999. Analyzing educational discourse: The DISCOUNT scheme. (Technical report 99/2). Leeds, United Kingdom: University of Leeds, Computer-Based Learning Unit.

Pintrich, P.R., & De Groot, E., 1990. Motivational and self-regulated learning components of classroom academic performance. Journal of Educational Psychology, 82, 33-40.

Pintrich, P.R., & Schrauben, B., 1992. Students' motivational beliefs and their cognitive engagement in classroom tasks. In D. Schunk & J. Meece (Ed.), Student perceptions in the classroom: Causes and consequences (pp. 149-183). Hillsdale, NJ: Erlbaum.

Pintrich, P.R., & Schunk, D.H., 2002. Motivation in education: Theory, research, and applications. Upper Saddle River, New Jersey: Merril Prentice Hall.

Porayska-Pomsta, K., 2004. Influence of Situational Context on Language Production. Unpublished Ph.D. thesis, University of Edinburgh, Edinburgh, United Kingdom.

Qu, L., Wang, N., Johnson, W.L., 2004. Pedagogical agents that interact with learners. In Proceedings of Workshop on Balanced Perception and Action in ECAs.

Reeves, B., Nass, C., 1996. The media equation. New York: Cambridge University Press.

Ryan, R.M., 1982. Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. Journal of Personality and Social Psychology, 43, 450-461.

Schunk, D, H., 1982. Effects of effort attributional feedback on children's perceived self-efficacy and achievement. Journal of Educational Psychology, 74, 548-556.

Schunk, D, H., 1983a. Ability versus effort attributional feedback: Differential effects on self-efficacy and achievement. Journal of Educational Psychology, 75, 848-856.

Schunk, D, H., 1983b. Developing children's self-efficacy and skills: The roles of social comparative information and goal setting. Contemporary Educational Psychology, 8, 76-86.

Schunk, D, H., 1983c. Goal difficulty and attainment information: Effects on children's achievement behaviors. Human Learning, 2, 107-117.

Schunk, D. H., 1983d. Reward contingencies and the development of children's skills and self-efficacy. Journal of Educational Psychology, 75, 511-518.

Schunk, D. H., 1984. Sequential attributional feedback and children's achievement behaviors. Journal of Educational Psychology, 76, 1159-1169.

Schunk, D. H., 1991. Self-efficacy and academic motivation. Educational Psychologist, 26, 207-231.

Schunk, D. H., 1995. Self-efficacy and education and instruction. In J.E. Maddux (Ed.), Self-efficacy, adaptation, and adjustment: Theory, research, and application (pp. 281-303). New York: Plenum Press.

Schunk, D. H., 1996. Goal and self-evaluative influences during children's cognitive skill learning. American Educational Research Journal, 33, 356-382.

Shaw, E., LaBore, C., Chui, Y.-C., & Johnson, W.L., 2004. Animating 2D digital puppets with limited autonomy. In Proceedings of the 2004 International Symposium on Smart Graphics. (pp. 1-10). Berlin: Springer.

Vallerand, R. J., & Reid, G., 1984. On the causal effects of perceived competence on intrinsic motivation: A test of cognitive evaluation theory. Journal of Sport Psychology, 6, 94-102.

Wang, N., Johnson, W.L., Rizzo P., Shaw, E. and Mayer, R.E., 2005. Experimental Evaluation of Polite Interaction Tactics for Pedagogical Agents. In Proceedings of the 2005 International Conference on Intelligent User Interfaces (pp. 12 - 19). New York: ACM Press.

Zhou X., & Conati C., 2003. Inferring User Goals from Personality and Behavior in a Causal Model of User Affect. In Proceedings of the 2003 International Conference on Intelligent User Interfaces (pp. 211-218). New York: ACM Press.

Zimmerman, B. J., 1995. Self-efficacy and educational development. In A. Bandura's (Ed.), Self-efficacy in changing societies. New York: Cambridge University Press.

Zimmerman, B.J., & Martinez-Pons, M., 1992. Perception of efficacy and strategy use in the self-regulation of learning. In D.H. Schunk & J.L. Meece (Eds.), Student perceptions in the classroom (pp. 185-207). Hillsdale, NJ: Erlbaum.

Zimmerman, B. J. & Schunk, D. (Eds.), 2001, Self-regulated learning and academic achievement: Theoretical perspectives (pp. 153-189). Mahwah, NJ: Erlbaum.

Tables:

Figures:

Figures can be printed in color online and in grayscale in print.