



**HAL**  
open science

## Seeing the face and observing the actions: the effects of nonverbal cues on mediated tutoring dialogue

Federico Tajariol, Jean-Michel Adam, Michel Dubois

► **To cite this version:**

Federico Tajariol, Jean-Michel Adam, Michel Dubois. Seeing the face and observing the actions: the effects of nonverbal cues on mediated tutoring dialogue. Intelligent Tutoring Systems, Jun 2008, Montreal, France. pp.10. hal-00344510

**HAL Id: hal-00344510**

**<https://hal.science/hal-00344510>**

Submitted on 16 Mar 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Seeing the face and observing the actions: the effects of nonverbal cues on mediated tutoring dialogue

Federico Tajariol<sup>1</sup>, Jean-Michel Adam<sup>2</sup>, Michel Dubois<sup>3</sup>

<sup>1</sup> R & D Orange Labs, France Telecom Groups, Lannion, France. [tajariol@gmail.com](mailto:tajariol@gmail.com)

<sup>2</sup> Laboratory Informatics Grenoble, University of Grenoble, France. [adam@imag.fr](mailto:adam@imag.fr)

<sup>3</sup> Laboratory of Social Psychology, Univ. of Grenoble, France. [Michel.Dubois@upmf-grenoble.fr](mailto:Michel.Dubois@upmf-grenoble.fr)

**Abstract.** Mediated communication technologies, conveying verbal and nonverbal cues, are more and more employed in learning activities. Nevertheless, their effects on teacher-student interaction have been not clearly stated yet. Through two experimental studies, we investigated on the effects of nonverbal communication cues (kinesic and ostensive-inferential) on synchronous mediated tutoring dialogue, in which a tutor and a student communicate through audio-video communication tools. The outcomes show that kinesic cues lead tutor to monitor more carefully learner's ongoing task and to encourage much more them, while ostensive-inferential cues improve learner' task performance and lead both tutor and student to focus better on tutoring speech acts.

**Keywords:** audio-video mediated communication, nonverbal communication, kinesic cues, ostensive-inferential cues, tutoring dialogue.

## 1 Introduction

Mediated communication technologies are more and more used in several human activities, also in teaching environments. These technologies allow distant participants, such a tutor and students, to see one another during the interaction, to work together on any shared documents [1], or just to observe the actions they make in their own working environment [2]. Even if many organizations and pedagogical institutions have already introduced these tools in their education services, the effects on the cognitive dimensions of the communication between teachers and students have not been clearly established yet [3].

The aim of this paper is to understand the effect of audio-video mediated technologies on a specific type of communication, which is the mediated tutoring dialogue. In the first section, we define tutoring dialogue as a joint communication activity and we show the role of nonverbal cues. We then describe our general method and we present two experimentations we conducted on the affordances of two non-verbal cues (kinesic and ostensive-inferential) on synchronous mediated tutoring dialogue.

## **2 Tutoring dialogue and non verbal cues**

As human beings, we use both verbal and nonverbal languages to perform activities, as creating, teaching, etc. When we communicate with our partners, we cannot directly know their thoughts, feelings and intentions; we are just able to infer them by interpreting their utterances [4] and nonverbal behavior [5]. Nevertheless, communication is not a mere sending-receiving messages activity, but it is rather a cooperative action [6]. For example, in tutoring dialogue, both tutor and student coordinate their turns to ground on a mutual understanding, so that tutors may enable students to contextualize their own problem statements and to improve their knowledge [7]. Concerning the nonverbal cues, we may distinguish kinesic cues (e.g., facial expressions, postures) and ostensive-inferential cues (e.g., actions and deictic gestures) [8]. Kinesic cues ensure the conversational floor between tutor and student, and moreover, they inform each participant about feelings and intentions of the other person [8]. Ostensive-inferential cues facilitate the verbal referring process, helping participants to coordinate their actions and to anticipate the other's needs [27]. In face-to-face tutoring dialogue setting, nonverbal cues are immediately available to both tutor and student: for instance, by observing students' facial expression tutors infer when to help them without disturbing needlessly [7]. In a video-mediated setting, according to the social presence theory [10, 11], it would be sufficient to put all nonverbal cues at tutor's and student's disposal to allow a suitable and efficient tutoring dialogue. Nevertheless, this does not seem the best solution. In fact, on the one hand, kinesic cues help distant participants to establish a mutual understanding [12, 13, 14] and ostensive-inferential cues support participants to perform a mediated activity more quickly [15]; on the other hand, nonverbal cues do not always help participants to better perform their activities [16, 17, 18]. However, amongst mediated-communication studies, just a few specifically concern mediated tutoring dialogue. This lack of studies would need priority status in the agenda research on mediated tutoring environments.

## **3 Research problem**

These considerations on the effects of kinesic and ostensive-inferential cues lead us to investigate on the following issues: which type of nonverbal cue improves the tutor-student learning interaction? Would it be better that the tutor and the student see each other (kinesic cues)? Would it be better that the tutor observes the student's actions to improve (her) his learning (ostensive-inferential cues)? To answer these questions, we conducted two experimental studies to understand the effects of kinesic cues (study 1) and ostensive-inferential cues (study 2) on tutoring dialogue.

## 4 General Method

The two experiments we will report followed a same experimental task, same apparatus, same procedure and same dependent measures.

### 4.1 Task, Apparatus and Procedure

*Task & Apparatus.* Tutors and students were involved in a sort of practical pedagogical work. The tutor had to help two students to learn basic commands of HTML language and to create an easy web page. The two students had to edit some sentences of a text in bold and italic, as well as building an internal link. The tutor had to help each student spontaneously or when a student asked for help, yet (s)he could help only one student at a time, as in a dyadic tutoring situation: while (s)he was communicating to a student, the other student could not hear their dialogue. The tutor and the two students settled down in three separate rooms, each one was equipped with a personal computer (central unit and monitor) and with a monitor for audio-video communication. Each computer was set with a web browser and a simple text editor. Anytime the tutor wanted to communicate with a student, (s)he chose the student by pressing a button on an *ad hoc* interface. Both students could ask the tutor to help them anytime, sending him/her a standard message by means of chat software.

*Procedure.* Each experimental session lasted nearly an hour, and it included three main phases. In the preparing phase (15 minutes), the tutor met the two students and the researcher explained them the aims of the experimentation. Then, the tutor and each student reached their own working room. Each subject filled in a consent form and was briefed on the main functions of the apparatus (e.g., the button to use to start a call, the folder containing the html files, etc.). Next, each student answered a pre-test to evaluate her/his HTML knowledge. Then the experimental phase started (35 minutes). Each student received a four-page HTML manual, which contained some HTML basic commands. Of course, the students needed their tutor's help to design the web page. When time ended, each subject was asked to complete a post-test questionnaire (same questions of the pre-test), then (s)he was debriefed and dismissed.

### 4.2 Measures

We transcribed verbatim all experimental sessions that we had videotaped. Based on studies concerning the grounding processes in communication [6], affordances of visual information in mediated communication [1, 9] and tutoring dialogues [19], we created the following coding scheme: *a) the tutor's proactive behavior*: we distinguished the tutor's spontaneous interventions towards the student as *proactive interventions* from the tutor's *reactive interventions* when s(he) replied to a student's call; *b) the mutual understanding*: we categorized all verbal markers that students and tutors had used during their dialogue to ground their mutual understanding (Table 1); *c) the tutoring intrinsic speech acts*: we categorized the tutors' and students' speech acts related to the intrinsic nature of the tutoring dialogue (Table 1). Three trained researchers coded the transcribed

dialogues. We checked the reliability of coding by means of the reproducibility test [29], obtaining an average value for students' acts (K=.58) and a high value for tutors' speech acts (K=.82).  
d) *Students' task performance*: we scored the web page that students had realized.

**Table 1 – Tutoring dialogue coding scheme**

Mutual understanding		
Role	Category	Examples
Tutor	To accept student's utterance	T: "yes" "ok, right"
	To check student's understanding	T: "is it clear now?", "is it ok?"
Student	To accept tutor's utterance	S: "yes", "ok"
	To check tutor's utterance	S: "could you repeat, please?", "what?"
Tutor's and Student's tutoring intrinsic speech acts		
Role	Category	Examples
Tutor	To find out student's ongoing task	T: "Did you try moving it on the red icon?"
	To help student	T: "Close the window and open the other file"
	To encourage student	T: "That's good you've nearly finished"
Student	To give tutor information about ongoing task	S: "I still have to finish this part of the exercise"
	To ask tutor's help	S: "Is the I tag in the HEAD part of the text?"

## 5 Experiment 1: the effects of kinesic cues on tutoring dialogue

### 5.1 Hypotheses

H1) Given the great difficulty of the task for a HTML beginner, we made the hypothesis that if the tutor could observe the students' faces, (s)he would infer their difficulties during the practical work and then (s)he would be *more proactive* in helping them without waiting for their call.

H2) We supposed that mutual understanding would be easier if the students and the tutor could see each other: we expected that the number of verbal markers for the common ground process would be fewer when the participants could see each other than when they could not.

H3) Consequently, we expected that their dialogue would be grounded on intrinsic tutoring speech acts and that intrinsic tutoring speech acts would be more numerous when participants could see each other than when they could not.

H4) We expected that students' performance would improve when tutor could see their faces while they were performing the task, because tutor could help them in a suitable manner.

### 5.2 Participants

We recruited 12 tutors, half men and half women (age M=30.9, S.D.=7.7). All of them were computer scientists owning good skills in HTML programming. We also recruited 48 undergraduate students in psychology (36 women and 12 men; age M=24.1, S.D.=6.2), all of them unskilled in HTML programming.

### 5.3 Experimental conditions

We set the following conditions: c1) *audio only*: tutor and students could only talk to each other through the audio channel; c2) *audio & human face*: the tutor and each of the two students could see each other's face and upper torso on a personal monitor, and they could talk to each other by means of the audio channel.

Each tutor performed in both conditions (within-participants experimental design), whereas half students were assigned to one of the two conditions (between-participants design). We controlled that the age of the students involved in the two conditions were equivalent ( $t(46)=1.318$ ; n.s.) and so was the HTML knowledge ( $t(46)=0.09$ ; n.s.). To guard against order effects, we counterbalanced the tutors' performing order setting.

### 5.4 Main Outcomes

We conducted independent *t-test* and we calculated the value of the effect size *d*.

H1) Tutors' proactive interventions were more in the *audio & human face* than in the *audio only* condition ( $M = 4.6$  (2) vs.  $1.8$  (0.9),  $t(22) = 4.27$ ;  $p < .001$ ,  $d = 1.8$ ). H1 was confirmed.

H2) Concerning the mutual understanding, tutor had significantly produced fewer verbal markers to check students' understanding in *audio & human face* than in *audio only* condition ( $M=3.9$  (3.6) vs.  $10$  (6.4),  $t(22)=2.85$ ;  $p < .01$ ,  $d=1.17$ ). About other students' and tutors' verbal markers, we did not find any statistically significant differences between the two conditions. H2 was partially confirmed.

H3) Concerning the speech acts referred to tutoring dialogue, we present the outcomes about tutors' speech acts and then the outcomes about students' speech acts.

*i*) Tutors produced significantly more speech acts to encourage students in *audio & human face* than in *audio only* condition ( $M=3.2$ (1.3) vs.  $1.2$ (1.2),  $t(22)=2.9$ ;  $p < .01$ ,  $d=1.14$ ). Moreover, the number of tutors' speech acts oriented to know students' ongoing task was significantly higher in *audio & human face* than in *audio only* condition ( $M=60.7$  (33.2) vs.  $35.3$  (8.1),  $t(22)=4.27$ ;  $p < .02$ ,  $d=1.05$ ).

*ii*) Consequently, students' speech acts oriented to give tutors any details about the ongoing task was significantly higher in *audio & human face* than in *audio only* condition ( $M=63.6$  (28.4) vs.  $43.8$  (14.4),  $t(22)=2.15$ ;  $p < .05$ ,  $d=0.88$ ).

No statistically significant differences between the two conditions were found about other speech acts. H3 was partially confirmed.

H4) Concerning the learning score, no significant differences were found between the two conditions (*audio only*  $M=4.2$  (1.4) vs. *audio & video-person*  $M=4.6$  (1.5),  $t(22)=0.83$ ; n.s.,  $d = 1.1$ ). H4 was rejected.

### 5.5 Discussions

The aim of this first experimentation was to measure the effects of kinesic cues on several measures, such as the tutor's proactive behavior, the mutual understanding between the tutor and the students, the intrinsic tutoring speech acts and the student's task performance. Concerning the tutor's proactive behavior, the

tutor assisted more spontaneously the students without waiting for their help request when tutor and students could see each other. This outcome confirms that, as in face-to-face setting [7], tutors take the floor before students explicitly produce an aid request. Moreover, tutors could easily check the students' understanding level by just observing their faces. This outcome corroborates that the addressees' face would help to be aware of their understanding [12]. Even if, however, the students' verbal markers are not significantly different between the two conditions, we must consider that tutors have generally the floor in tutoring dialogue and that students play the role of a reactive addressee rather than the leading role [7]. On the other hand, it could be possible that, given the difficulty of the task, they chose to check tutor's utterances by using verbal markers in order to be sure of the content of tutor's aid. The outcomes concerning the students' task performance imply that human faces do not improve the performance in procedural tasks [3, 17, 18].

## **6 Experiment 2: the effects of ostensive-inferential cues**

The experimental method of this experiment was the same than for the first one.

### **6.1 Hypotheses**

H1) We expected the tutor would be more proactive when (s)he could observe the students' actions, because (s)he was always aware of their difficulties and (s)he could take the floor before they asked for help.

H2) We expected that the number of verbal markers for mutual understanding would be fewer if tutors could observe the students' actions.

H3) We expected that the number of intrinsic tutoring speech acts would be higher when the tutors could observe the students' actions rather than when they could not.

H4) We also predicted the students' performance would improve when the tutor could observe their actions during the practical work session, because (s)he could choose the most suitable help to their needs.

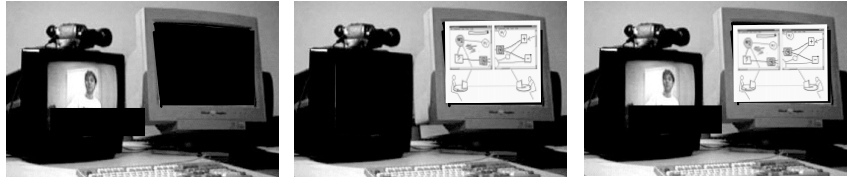
### **6.2 Participants**

We recruited the same tutors ( $N = 12$ ) as for the first study, and 72 undergraduate students in psychology (50 women and 22 men; age  $M=23.8$ ,  $S.D.=5.1$ ), all of them unskilled in HTML programming.

### **6.3 Experimental conditions**

We set the three following conditions: c1) *audio & human face*: the tutor and the students could see each other's face and upper torso on the screen and they talked to each other through the audio channel; c2) *audio & student's actions*: the tutor and the students could talk through the audio channel, and the tutor could observe both students' computer screens (by means of VNC™ software). So, the tutor

could observe the actions that each student was performing on their own computer desktop and (s)he could make the decision to help them; c3) *audio & human face & student's actions*: the tutor could observe both students' face and upper torso and their computer screens, while (s)he could talk to each other by means of the audio channel.



**Figure 0 - Experimental conditions study 2 (from left): audio & human face vs. audio & student's action vs. audio & human face &**

As for the first study, each tutor performed all conditions (within-participants design) and we counterbalanced the order of the tutors' performance to guard against order effects. The students were equally assigned to the three experimental conditions. We controlled their age ( $F(2, 69) = 1.73$ ; n.s.) and their HTML knowledge ( $F(2, 69) = 0.16$ ; n.s.) in each of the three groups.

#### 6.4 Main Outcomes

We conducted an analysis of variance (ANOVA) and *a priori* comparisons between conditions.

H1) About the tutors' proactive behaviour, the number of their spontaneous interventions were significantly different between conditions ( $F(2, 33) = 4.671$ ;  $p < .02$ ,  $\eta^2 = 0.22$ ). In fact, the tutors were significantly more proactive in *audio & human face & student's actions* than in *audio & student's actions* condition ( $M = 6.6$  (1.9) vs. 4.6 (2),  $t(33) = 2.366$ ;  $p < .02$ ,  $d = 1.02$ ) and than in *audio & video-person* condition ( $M = 6.6$  (1.9) vs. 4.2 (2.2),  $t(33) = 2.859$ ;  $p < .007$ ,  $d = 1.16$ ). H1 was confirmed.

H2) About mutual understanding, the number of tutors' and students' verbal markers used to ground the understanding are not significantly different between conditions. H2 was rejected.

H3) About tutoring speech acts, we will first present outcomes for tutors (i) and then for students (ii).

i) Tutors. One-way ANOVA showed that the number of tutors' speech acts to know the students' ongoing task was significantly different between conditions ( $F(2,33) = 4.529$ ;  $p < .02$ ,  $\eta^2 = 0.21$ ). Tutors produced fewer speech acts to know the students' task in *audio & student's actions* than in *audio & human face* condition ( $M = 27.4$  (7.6) vs. 35.3 (8.1),  $t(22) = 2.8$ ;  $p < .01$ ,  $d = 1$ ) and fewer in *audio & human face & student's actions* than in *audio & human face* condition ( $M = 26.6$  (6.8) vs. 35.3 (8.1),  $t(22) = 2.6$ ;  $p < .005$ ,  $d = 1.16$ ). Moreover, the tutors' speech acts to help the student were significantly different between conditions ( $F(2, 33) = 15.248$ ;  $p < .000$ ,  $\eta^2 = 0.48$ ). In fact, tutors' speech acts to help students were fewer in *audio & human face* than in *audio & student's actions* ( $M = 14.8$  (10.4) vs. 40.2 (12.9),  $t(33) = 4.99$ ;  $p < .000$ ,  $d = 2.18$ ) and fewer in *audio & human face* than in *audio & human face &*



*student's actions* ( $M=14.8$  (10.4) vs.  $37.9$  (13.8),  $t(33)=4.55$ ;  $p < .000$ ,  $d=1.91$ ). Finally, concerning the tutors' speech acts to encourage students, their number was significantly different between conditions ( $F(2, 33) = 4.813$ ;  $p < .01$ ,  $\eta^2 = 0.22$ ). In fact, tutors more often encouraged the students in *audio & human face* than in *audio & student's actions* condition ( $M=3.2$  (2.3) vs.  $0.8$  (1.1),  $t(33)=2.69$ ;  $p < .01$ ,  $d=1.33$ ) and more often in *audio & human face & student's actions* than in *audio & human face* ( $M=3.2$  (2.8) vs.  $0.8$  (1.1),  $t(33)=2.69$ ;  $p < .01$ ,  $d=1.12$ ).

ii) Students. About the students' speech acts to give tutors information about the ongoing task: the one-way analysis of variance showed significant differences between conditions ( $F(2, 33) = 4.529$ ;  $p < .02$ ,  $\eta^2 = 0.21$ ). Students less often gave their tutors information about their ongoing task in *audio & human face & student's actions* than in *audio & human face* condition ( $M=26.5$  (14.1) vs.  $43.8$  (14.4),  $t(22)=2.5$ ;  $p < .01$ ,  $d=1.21$ ) and less often in *audio & student's actions* than in *audio & human face* condition ( $M=27.7$  (18.4) vs.  $43.8$  (14.4),  $t(22)=2.6$ ;  $p < .005$ ,  $d=0.97$ ). All other students' speech acts did not differ between conditions. H3 was partially confirmed.

H4) The students' web page score was significantly different between conditions ( $F(2, 69) = 5.776$ ;  $p < .005$ ,  $\eta^2 = 0.14$ ). The students performed significantly better in *audio & human face & student's actions* than in *audio & human face* condition ( $M = 5.9$  (1.6) vs.  $4.5$  (1.5),  $t(69)=3.27$ ;  $p < .002$ ,  $d=0.88$ ) and better in *audio & student's actions* than in *audio & human face* condition ( $M=5.5$  (1.2) vs.  $4.5$  (1.5),  $t(69)=2.4$ ;  $p < .02$ ,  $d=0.73$ ). H4 was confirmed.

## 6.5 Discussions

The second study focused on the effects of ostensive-inferential cues referring to student's actions on tutoring dialogue. The overall outcomes showed: i) concerning the tutor's proactive behavior, the effect of ostensive-inferential cues was not as important as we expected. In fact, the number of tutor's proactive interventions produced with ostensive-inferential cues was nearly the same as produced with kinesic cues only. It appeared that ostensive-inferential cues only did not lead tutors to increase their proactive contributions. In fact, the tutors' proactive behavior would increase if both the kinesic and ostensive-inferential cues were available to the tutors. Against our expectations, ostensive-inferential cues did not improve the mutual understanding. Moreover, kinesics and ostensive-inferential cues support in an equivalent manner the mutual understanding. This outcome suggests that even if the tutor could observe both students' faces and actions, (s)he checked their understanding through verbal markers, and so corroborating tutors' behavior in a face-to-face setting [20]. However, in other mediated communication activities [1, 15], ostensive-inferential cues decreases verbal markers used for mutual understanding. This would suggest that the properties of ostensive-inferential cues have different effects on mutual understanding between distant partners, depending on the characteristics of the specific joint activity. Besides, when ostensive-inferential cues are available, tutors' questions to know the students' task status decreased and tutors' speech acts to help students increased. We highlight that the number of tutors' encouraging speech acts was higher with kinesic cues than with ostensive-inferential cues. This suggests that through kinesic

cues the content of tutoring dialogue is oriented towards socio-relational issues rather than task issues. Concerning the students' task performance, ostensive-inferential cues improved students' task performance better than kinesic cues did.

## 7 Conclusion

The aim of these two experiments was to understand the effects of nonverbal cues, such as kinesic and ostensive-inferential cues, on mediated tutoring dialogue. Main outcomes of experiment 1 showed that when the tutor and the students could see each other's face, the tutor more often took the floor spontaneously, (s)he produced fewer verbal markers to check students' understanding and (s)he more willingly encouraged them. A possible interpretation is that kinesic cues, increasing the level of social presence, led tutor to monitor more carefully learners' ongoing task and to encourage them, helping both students and tutors in mutual understanding process. Outcomes from experiment 2 partially corroborated the first ones, and also confirmed some previous researches (e.g., kinesic cues improve proactive behavior) [9]. Moreover, ostensive-inferential cues let the tutor and the student focus on intrinsic tutoring contents, and improved student's performance. Although these outcomes should be considered under some limits (e.g., visual parallaxes sometimes lowered the quality of the interaction between participants), if we are asked to suggest some tips for the design of tools supporting synchronous mediated tutoring dialogue, we argue that: a) if we want the tutor to be proactive, it would be better to let the tutor and the student see each other; b) if we want the student to improve learning, it would be better to let the tutor observe the student's actions. These outcomes show that ostensive-inferential cues improve the learning performance of students involved in procedural task (practical works, training in using software, etc.) It would be necessary to corroborate their effects on declarative content tasks. More researches would be needed to better understand the effects of the interaction between the different nonverbal cues, as well as to allow an efficient tutoring dialogue.

## 8 References

1. Kraut, R. E., Fussell, S.R., Siegel, J.: Visual information as a conversational resource in collaborative physical tasks. *Human Comp. Inter.* 18, 13--49 (2003).
2. Gergle, D., Kraut, R.E., Fussell, S.R.: Language efficiency and visual technology: minimize effort with visual information. *J. Lang. Soc. Psych.* 23, 491--517 (2004).
3. Masoodian, M., Apperley, M., Frederickson, L.: Video support for shared work-space interaction: an empirical study. *Inter. Comp.* 7, 237--253 (1995).
4. Sperber, D., Wilson, D.: *Relevance*. Blackwell, Oxford (1987).
5. Argyle, M., Graham, J.: The Central Europe Experiment: looking at persons and looking at things. *J. Envir. Psych. Nonv. Behav.* 1, 6--16 (1977).
6. Clark, H.H.: *Using languages*. Cambridge University Press, Cambridge (1996).

7. Fox, B.A.: *The Human Tutorial Dialogue Project: Issues in the Design of Instructional Systems*. Lawrence Erlbaum, Hillsdale NJ (1993).
8. Manusov, V., Patterson, M.L.: *The handbook of nonverbal communication*. Sage Publications (2006).
9. Karsenty, L.: Cooperative work and shared visual context : an empirical study of comprehension problems in side-by-side and remote help dialogues. *Hum. Comp. Inter.* 14, 283--315 (1999)
10. Daft, R.L., Lengel, R.H.: Organisational information requirements, media richness and structural design. *Management Science*, 32, 554-571 (1986).
11. Whittaker, S.: Theories and methods in mediated communication. In Graesser, A. C., Gernsbacher, M. A., Goldman, S.R. (eds.) *Handbook of discourse processes*, pp. 123--164. Lawrence Erlbaum, Hillsdale, NJ (2002).
12. Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S., Bruce, V.: Face-to-face interaction and video mediated communication: a comparison of dialogue structure. *J. Exp. Psych.: Applied* 3, 105--125 (1997).
13. Isaacs, E., Tang, J.C.: What video can and cannot do for collaboration. *Multim. Syst.* 2, 63--73 (1994).
14. Whittaker, S., Geelhoed, E., Robinson, E.: Shared workspaces: how they work and when are they useful? *Inter. J. Man-Mach. Studies* 39, 813--842 (1993).
15. Gergle, D., Kraut, R.E., Fussell, S.R.: Language efficiency and visual technology: minimize effort with visual information. *J. Lang. Soc. Psych.* 23, 491--517 (2004).
16. Anderson, A.H., Newlands, A., Mullin, J., Fleming, A.M., Doherty-Sneddon, G., Van Der Velden, J.: Impact of video-mediated communication on simulated service encounters. *Inter. Comp.* 8, 193-206 (1996).
17. Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S., Bruce, V.: Face-to-face interaction and video mediated communication: a comparison of dialogue structure. *J. Exp. Psych.: Applied* 3, 105--125 (1997).
18. O'Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G., Bruce, V.: Comparison of face-to-face and video-mediated communication. *Inter. Comp.* 8, 177--192 (1996).
19. Chi, M.T.H., De Leeuw, N., Chiu, M.H., Lavancher, C.: Eliciting self-explanations improves understanding. *Cogn. Science* 18, 439--477 (1994).
20. Van Lehn, K., Siler, S., Murray, C., Yamauchi, T., Baggett, W.B.: Why do only some event cause learning during human tutoring? *Cogn. Inst.* 21, 209-249 (2003).