

# Participatory status in electronically mediated collaborative work

Andrew Monk

Department of Psychology  
University of York  
York, YO10 5DD, UK  
A.Monk@psych.york.ac.uk

## Abstract

Overhearing has been observed to be an important conversational resource in a number of cooperative work contexts. Participatory status describes a particular kind of overhearing where two or more primary participants are actively involved in some cooperative task with one or more peripheral participants who are not actively involved, but nevertheless have a legitimate reason for listening in. Our interest in participatory status is focussed on how the way an audio video link is configured could affect a peripheral participant. Clark's theory of language use is successfully used to model participatory status and to make predictions about these effects.

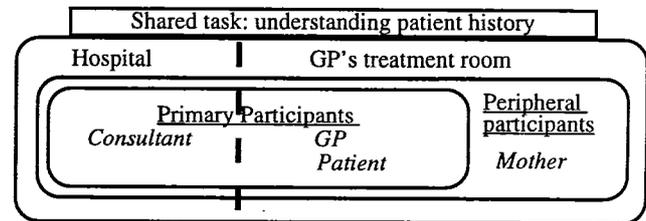
## Introduction

This paper is concerned with the application of psychological theory to improve systems for multimedia communication. Much of the theoretical and empirical work on communication focuses on dyadic rather than multi-person communication. This emphasis on two person groups is also seen in many of the technologies proposed for office communication.

At York we have been examining situations where groups larger than two need to communicate electronically. The work started with some field studies of telemedical consultation. This is an interesting context as it involves: (i) sharing images of the work as well as images of each other, and (ii) three- and four-person groups working together. One of the conclusions drawn from our analysis of this application of communication technology was the value given to overhearing. This might be overhearing by people actively engaged in the current shared work task and overhearing by people who are less actively engaged but who nevertheless get communicative value from knowing what is going on. These field studies were followed up by experiments exploring this phenomenon and the effect of changes in the visibility of participants on the status of overhearers.

This paper is not to report the results of the field studies or the experiments rather to suggest how Clark's theory of

language use may help understand issues we have been exploring surrounding overhearing and participatory status.



**Figure 1.** Primary and peripheral participants at some point in time during a telemedical consultation. The dashed line signifies that the consultant is in a different location to the other participants.

## Primary and peripheral participants

Ethnographic studies of computer supported cooperative work (CSCW) have commonly identified overhearing, or more generally the monitoring of other people's behaviour, as important resources for coordinating work. Heath & Luff, (1996) for example, describe how the operators responsible for line control and passenger information in a London Underground control room coordinate their work by monitoring each other's communicative behaviour with other people. For example, the announcer was observed to make a passenger announcement about a delay that had been inferred from overhearing the controller's conversation with a driver without any explicit communication between the two of them. They cite this example of coordination as an example of a "convergent activity". Other examples they cite of behaviours monitored to coordinate convergent activities are, looking at some information source and picking up the phone. A similar point is made by (Hutchins, 1994) with his concept of distributed cognition. Hutchins, observed work contexts such as aircraft cockpits, ship's bridges, and control rooms. He sees information sources available to all the individuals in a team as cognitive resources. These resources include not only instruments and displays but also the utterances and behaviour of team members.

Participatory status is a particular example of the phenomenon described above. It applies to situations where some of the participants in a cooperative task are more actively involved than others. The distinction is between *primary participants*, for whom resources for coordination are of prime importance, and *peripheral participants* whose work is less directly intertwined with that of the primary participants but who nevertheless can benefit from monitoring the talk and behaviour of those primary participants. By locating this concept within an analysis of the components of the work to be completed, this can be seen as a link between the group oriented perspective of CSCW research and the more individual perspective of conventional task analysis (Watts & Monk, 1998).

This distinction between primary and peripheral participants arose from some field studies. Three sites were studied where a GP or a nurse practitioner at one location consulted a specialist in a hospital over a video link (Watts and Monk, 1997; Watts and Monk, 1998). At all the sites studied, conversations were rarely simple two-party interactions. When a patient is talking to a remote consultant, for example, it may be in the presence of her GP, a relative and a nurse. Accordingly, the task analysis notation we developed to analyse the communications requirements of these sites, the CUD (Comms Usage Diagram, see Monk and Watts 1997; Watts and Monk, 1999) requires the analyst to identify the parties present and their level of participation. At any point in time there will be "primary participants" actively engaged in some shared task and "peripheral participants" who can legitimately monitor the conversation but are not actively engaged in it. For example, at a particular moment in time the consultant may be discussing the history of a patient's problem over the video link with the patient and their GP, who is acting as a kind medical advocate. These three then are the primary participants. Commonly, additional people would be in the treatment room with the GP and patient. These might include nursing staff, a radiographer, friend or relation of the patient or a technician. These additional people are described as peripheral participants (see Figure 1). They would not have a part to play in the current shared task, but they could overhear and see what was going on, and have a close interest in it. Peripheral participants stand to gain from overhearing the primary participants. A radiographer listening in on a telemedical consultation stands to learn something more about interpreting X-rays. A nurse might later be charged with the care of the patient and, by virtue of having witnessed the consultation, already have a good insight into the patient's situation. A relative overhearing some factual inaccuracy may introduce themselves into the conversation in order to contradict what was said.

Further, the field studies suggest that the precise configuration of communications equipment could have a potentially large effect on: (i) the salience of peripheral participants to the primary participants, and (ii) the salience of the activities of the primary participants to the peripheral participants. For example, the quality of multi-party sound

provided by a speakerphone arrangement was often poor forcing the primary participants to resort to telephone handsets. Peripheral participants would then only be able to hear one side of the conversation. More subtle changes may result from the positioning of the camera. This often meant that primary participants could not see peripheral participants, making their presence very much less salient.

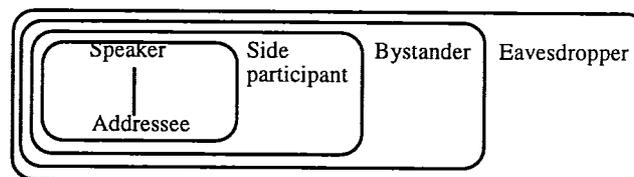


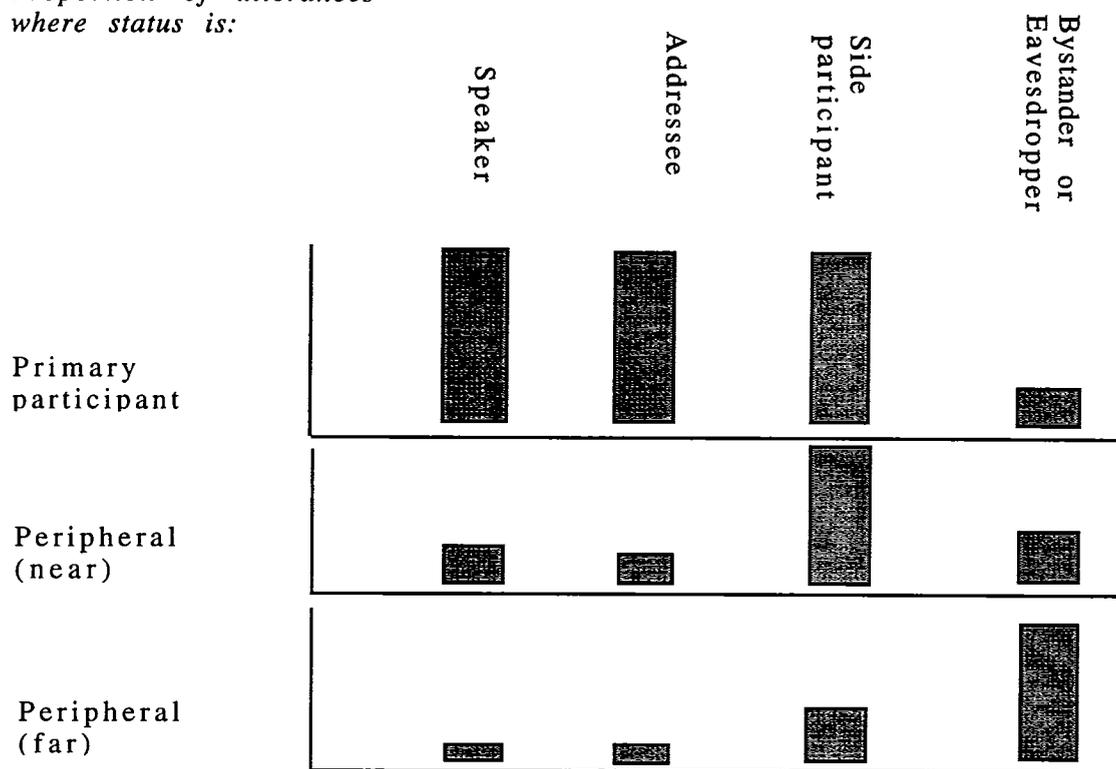
Figure 2. Levels of participation, adapted from Clark 1996.

### Side participants, bystanders and eavesdroppers

We can relate this distinction between peripheral and primary participants to Clark's theory of language use and his analysis of overhearing (Schober and Clark 1989). Our analysis is concerned with a level of description to do with work and shared tasks within that work. Theories of language use are concerned with a finer level of granularity, a level of description that considers individual utterances. We will describe the distinction between primary and peripheral participants as "participatory status" and oppose it to language theory distinction that we will call "conversational status". Clark (1996) defines four categories of conversational status in everyday conversation (see Figure 2). A speaker addresses an utterance to an addressee. A side participant is a "ratified participant" (Goffman 1976) i.e., a participant recognised by speaker and addressee as a full member of the conversation. A bystander is not a ratified participant but the speaker is aware they can overhear. Finally, an eavesdropper is an overhearer that the speaker is not aware of.

In Clark's analysis the speaker is engaging in different joint actions with each of these listeners. The utterance is designed for the addressee and the speaker has a strong obligation to monitor understanding and repair any evidence of trouble that may arise. The addressee has similar obligations to try to understand the speaker and to signal when this is and is not possible. The speaker has similar but slightly weaker obligations to a side participant. The side participant is a part of the three-way conversation and so must be able to understand what is said in order to address the other participants appropriately at some later point. However, their rights and obligations are less than those of the addressee. The speaker's obligations to a bystander are further reduced. Finally, speakers can have

*Proportion of utterances  
where status is:*



**Figure 3.** The relationship between participatory status and Clark's categories of overhearer (conversational status)

no obligations to an eavesdropper as by definition they are oblivious to their listening in.

To map between our notion of participatory status and Clark's conversational status one needs to assess the proportion of time (or proportion of utterances) for which each participant has a particular conversational status. Consider a three-party conversation. If for some part of the work someone is a primary participant then one would expect them to be speaker, addressee and side participant for roughly equivalent proportions during that period. Also, they should only rarely, if at all, be a bystander or an eavesdropper. This is depicted in the first row of Figure 3. If on the other hand they are a peripheral participant, they should be most often a side participant, bystander or eavesdropper. Further, this analysis leads to a definition of degrees of peripherality. If a person is mostly a side participant then they may be considered to be less peripheral than if they are mostly a bystanders or eavesdropper. This is depicted in the second and third row of Figure 3.

Degree of peripherality is important as it may affect the mobility of participatory status. In the snap shot of peripheral participatory status provided by Figure 1, the mother is peripheral. At some later point the situation may change. The mother may change her status to be actively involved in the current cooperative task as primary participant. The patient may change his or her status from

primary to peripheral. Mobility of participatory status would seem to be a good thing. One would expect that contexts that allow mobility of participatory status to result in more effective coordination and communication. If the communication facilities or task demands make a peripheral participant more peripheral then it would be expected they would find it harder to become a primary participant.

### Clark's levels of joint action applied to overhearers

Clark's notion of levels of joint action, is a potential basis with which to model the effects of manipulations of communication facilities on the degree of peripherality experienced by a peripheral participant. Central to this account is the concept of a joint action and his extension of Grice's (Grice 1957) notion of cooperation by describing communication in terms of four levels of joint action (see Table 1). At level 1 a speaker and an addressee have to cooperate to the extent that the speaker intends to execute some behaviour and the addressee to attend perceptually to it. At level 2 the speaker presents a signal to the addressee who cooperates by intending to identify it. At level 3 the speaker is intending to signal that something is the case and the addressee is recognising that this is so from the signal

and its context. At level 4 the speaker is proposing a "joint project" and an addressee is considering it. The joint actions at each level in the action ladder have coincident starting and finishing points and upward causality (e.g., level 1 joint action is completed in order to complete a level 2 joint action). Upward causality implies downward evidence. Thus evidence that a level 2 joint action has been completed implies that the accompanying level 1 joint action was also completed.

Speaker A's actions	Addressee B's actions
4 A is proposing a joint project w to B	B is considering A's proposal of w
3 A is signalling that p for B	B is recognising that p from A
2 A is presenting signal s to B	B is identifying signal s from A
1 A is executing behaviour t for B	B is attending to behaviour t from A

**Table 1.** Levels of joint action (after Clark, 1996, p. 389)

Clark's recursive definition of a joint task asserts that all participants know that all the other participants know they are engaged in the joint task. Thus at each level of joint action speaker and addressee need to be confident that the other is indeed treating this as a joint task. Table 2 suggests how this might be achieved. At level 4, evidence that the other partner is treating this as a joint task comes from the content of the joint project proposed by A and from the history of B's responses to joint proposals. At level 3 there is again the possibility that the content can provide evidence, but there is also the possibility of verbal and visual back channels. There are no obvious categories of evidence at level 2 so reassurance must come by downward evidence from level 3 or level 4. At level 1 there is the possibility of monitoring visual attention.

Evidence leading speaker A to consider addressee B	Evidence leading addressee B to consider speaker A
4 B has responded appropriately to previous proposals (H)	A's proposal is relevant to B (H)
3 B has responded appropriately to previous signals (H); A can hear verbal back channels (H); A can see visual back channels (S)	A's signal is directed at B (H); A's signal refers to common ground specific to B (H)
2 Only by downward evidence	Only by downward evidence
1 A can see B is attending (S)	B can see A's behaviour is directed at B (S)

**Table 2.** Evidence that the other person is taking part in the joint task, speaker and addressee. (H) = must be able to hear other; (S) = must be able to see other.

We can generalise this scheme to each conversational status discussed above. Table 3 presents the same analysis for a side participant. The actions are equivalent except that now the speaker is taking joint actions with the addressee B and a side participant C. What distinguishes the situation of two addressees from the situation of one addressee and one side participant is that there is no joint action at level 4. Joint actions are required at levels 2 and 3 in order to build up common ground for when C becomes the attendee. So C may be recognising the signal that p and A signalling that p to both B and C but there is only one addressee for the proposed joint project. Table 4 presents the analysis of where speaker and side participant get reassurance that the other is indeed taking part in these joint actions. Apart from the lack of a joint action at level 4, this is the same as Table 2.

Speaker A's actions	Side participant C's actions
4 No joint action	No joint action
3 A is signalling that p for B and C	C is recognising that p from A
2 A is presenting signal s to B and C	C is identifying signal s from A
1 A is executing behaviour t for B and C	C is attending to behaviour t from A

**Table 3.** Levels of joint action with a side participant.

Evidence leading speaker A to consider side participant C	Evidence leading side participant C to consider speaker A
4 No joint action	No joint action
3 C has responded appropriately to previous signals (H); A can hear verbal back channels from C (H); A can see visual back channels from C (S)	A's signal is directed at B and C (H); A's signal refers to common ground specific to C (H)
2 Only by downward evidence	Only by downward evidence
1 A can see C is attending (S)	B can see A's behaviour is directed at B and C (S)

**Table 4.** Evidence that the other person is taking part in the joint task, speaker and side participant. (H) = must be able to hear other; (S) = must be able to see other.

Table 5 presents the very weak levels of joint action between a speaker and a bystander. The speaker is primarily engaging in a joint action with attendee and side participant, however he or she knows that the bystander can overhear and so may make allowances for this. Note that what the bystander D may take from this overhearing may be different from what was intended by the other participants. Hence if p is what B and C are intended to recognise then p' is what D recognises.

The evidence providing mutual reinforcement for these weak joint actions is described in Table 6. For the speaker this comes from the speaker seeing that the bystander is attending to what is being said. For the bystander evidence comes from attending to, identifying and recognising previous signals from A.

Speaker A's actions	Bystander D's actions
4 No action	No action
3 A is signalling that p (for B and C) and realises that D will recognise that p'	D is recognising that p' from A
2 A is presenting signal s (to B and C) and realises that D will identify signal s'	D is identifying signal s' from A
1 A is executing behaviour t (for B and C) and knows D is attending	D is attending to behaviour t from A

**Table 5.** Levels of joint action with a bystander.

Evidence leading speaker A to consider bystander D	Evidence leading bystander D to consider Speaker A
4 None	None
3 None	D has recognised previous signals from A (H)
2 None	D has identified previous signals from A (H)
1 A knows D can hear; A can see D is attending (S)	D can hear A behaving (H); D can see A behaving (S)

**Table 6.** Evidence that the other person is taking part in the joint task, speaker and bystander. (H) = must be able to hear other; (S) = must be able to see other.

By definition the speaker is not even aware of an eavesdropper thus there can be no joint tasks. Tables 7 is described as "levels of action" rather than "levels of joint action" as in tables 1, 3 and 5. Table 8 records the evidence that may encourages the eavesdropper to listen to the speaker.

Speaker A's actions	Eavesdropper E's actions
4 No action	No action
3 No action	E is recognising that p from A
2 No action	E is identifying signal s from A
1 No action	E is attending to behaviour t from A

**Table 7.** Levels of action of an eavesdropper.

Experience leading eavesdropper E to monitor Speaker A
4 None
3 E has recognised previous signals from A (H)
2 E has identified previous signals from A (H)
1 E can hear A behaving (H); E can see A behaving (S)

**Table 8.** Experience leading an eavesdropper to listen to a speaker. (H) = must be able to hear other; (S) = must be able to see other.

## Predicting the effects of audio-visual configurations on degree of peripherality

The above analysis can be used to generate hypotheses concerning the effects of changing the information available about a peripheral participant to a primary participant and vice versa. Consider the configuration used in one of our experiments. Three students are involved in a role play in which one of them is a student faced with disciplinary proceeding on account of failing three courses. This student role has to decide which of the three courses to appeal against in a discussion, over a video link, with an administrator role. The peripheral participant is a tutor role. The student is in the tutor's room and the task is so organised that the tutor is highly involved, though he or she cannot take part in the discussion.

Two audio video configurations were used in our experiment. The configuration used in the "high visibility" condition is described in Table 9. The administrator can see the student on a video monitor. They are sitting shoulder to shoulder at a table (see Figure 4). The tutor is strongly discouraged from trying to join in the discussion by the task instructions. The student wears a boom microphone that is the administrator's only source of sound from the remote office, and so cannot hear the tutor anyway.

The student and tutor sit in front of the video monitor providing sound and an image of the administrator who they can both see and hear. The student and tutor can hear each other but as the tutor is discouraged from saying anything by the task instructions, in practice, the student does not hear the tutor. In table 9 the tutor's audio availability is thus recorded in parentheses indicating that in practice no audible signals from the tutor are available to the student, even though they are in the same room. Similarly, while the student and tutor are in the same room and so can see each other but only by turning their heads away from the video monitor in front of them, our data shows that the task demands are such that the student spend nearly all the time looking at the monitor. Thus the tutor's audio availability is also recorded in parentheses indicating that in practice no visible signals from the tutor are available to the student.

Role	Status	Can see	Can hear
Admin.	Primary	Student, Tutor*	Student
Student	Primary	Admin., (Tutor)	Admin., (Tutor)
Tutor	Peripheral	Admin., Student	Admin., Student

**Table 9.** Who could see and hear who in the high visibility condition in the experiment. \* in the low visibility condition the tutor was not visible to the administrator. Parentheses indicate that in practice the availability was minimal (see text).

To create the low visibility condition we simply blanked out the side of the image presented to the administrator that contained the tutor (see right hand side of figure 4). To predict what effect this would have on the degree of peripherality of the peripheral party one needs to relate table 9 to tables 2, 4, 6, and 8. These list the evidence speakers and overhearers have available to encourage them to consider the other people. Each piece of evidence is annotated with an "H" or "S" to indicate whether that evidence depends on being able to hear or see the other person.

### *Student's relationship to tutor and administrator's relationship to student:*

Table 9 shows that the visibility condition has no effect on what the student and tutor, or administrator and student, can see or hear of one another and so we would expect no effect on degree of peripherality in these relationships. This is not true of the relationship between tutor and administrator.

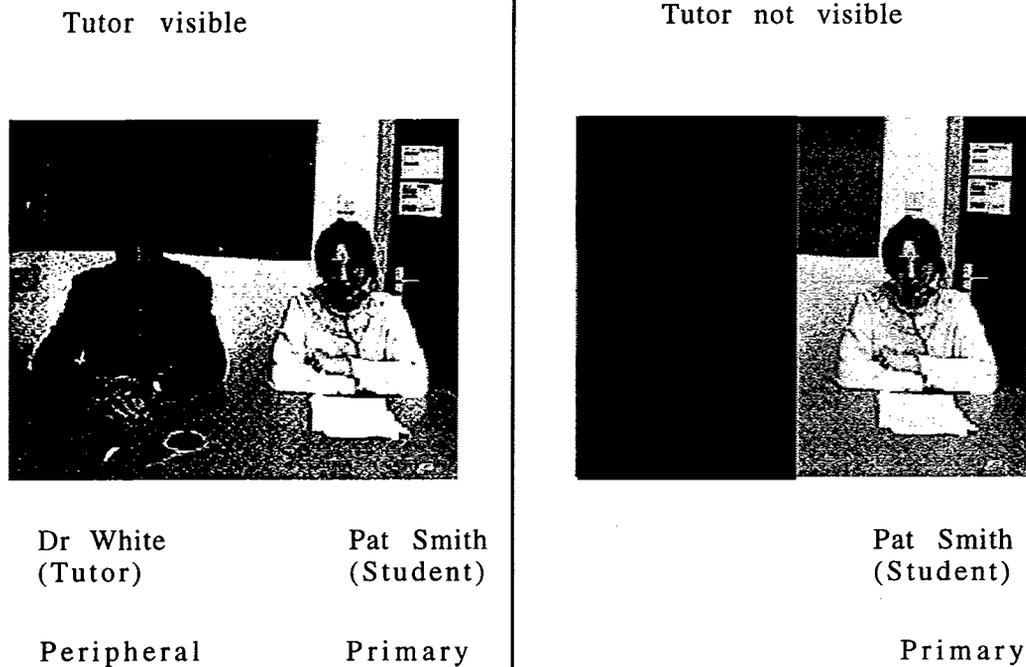
### *Administrator's relationship to tutor:*

The tutors do not speak so they may be an addressee, side participant bystander or eavesdropper. As they cannot reply it seems unlikely they will often be an addressee and so degree of peripherality will depend mainly on how much they are side participants (see Figure 3). Table 4 lists the evidence providing mutual reinforcement for the joint actions needed for the tutor to be a side participant when the administrator is speaker.

The tutor can see and hear the administrator in both conditions and so all this evidence is available to the tutor. The administrator cannot hear the tutor in either and can only see the tutor in the high visibility condition. The evidence leading the administrator as speaker to consider the tutor as a side participant can be through seeing, leading to the prediction that the high visibility condition would encourage the administrator to relate to the tutor as a side participant more in the high visibility condition. This may result in the administrator presenting more evidence to the tutor who will in turn be more likely to behave as a side participant in a pattern of mutual reinforcement.

Role	Status	Can see	Can hear
Admin.	Primary	Student, Tutor*	Student, Tutor
Student	Primary	Admin., (Tutor)	Admin., Tutor
Tutor	Peripheral	Admin., Student	Admin., Student

**Table 10.** Who could see and hear who in the high visibility condition if the audio constraints had been relaxed. \* in the low visibility condition the tutor would not be visible to the administrator.



**Figure 4.** Administrator role's view in high and low visibility conditions

*What if the tutor had been able to hear the tutor:*

To illustrate further how Clark's levels of joint action leads to detailed predictions consider what would have happened if the administrator had been able to hear the tutor in both visibility conditions and the task instructions had allowed the tutor to speak, all be it infrequently so that they remain a peripheral participant. Table 10 describes this audio visual configuration. Again the difference is in the relationship between administrator and tutor. Consulting table 4 we see that there is now the potential for level 3 evidence for the administrator as speaker that the tutor is a side participant, available by hearing. Downward evidence implies joint activity at level 2 and level 1 and so the availability of visual evidence is going to be less important. The prediction is then that the effect of the visibility condition on the degree of peripherality of the peripheral participants will be less if those peripheral participants are allowed to speak and the remote administrator can hear them do so.

**Conclusions**

This paper has shown that our concept of participatory status can be related to Clark's theory of language use.

Further, the modelling required to do this demonstrates that Clark's theory of language use has sufficient power to make quite detailed predictions in this context. By taking a particular example, it was possible to predict how a relatively subtle manipulations of the communication technology should change the degree of peripherality of a peripheral participant.

The models developed have a similar potential to make predictions about practical design decisions, e.g., when it is potentially important for someone to be able to see someone else and when it is not. Before doing so these concepts need to be operationalised in experimental procedures and practically important effects demonstrated. In fact, when degree of peripherality is operationalised using interpersonal awareness scores the predictions are not born out (see Table 11). In our experiments, the effect of visibility condition is seen in the student's ratings of the tutor, not in the tutor's rating of the administrator and the administrator's ratings of the tutor, as predicted. More analytic and empirical work is needed to understand these effects.

Our next step is to apply the models to other detailed experimental results we already have and to test them further with new experiments and field studies.

Rater	Rated	Mean interpersonal awareness	
		High vis.	Low vis.
Tutor	Admin.	73 (16)	78 (9)
	Student	73 (13)	82 (9)
Admin	Student	81 (10)	86 (10)
	Tutor	25 (16)	20 (20)
Student	Admin	78 (18)	80 (9)
	Tutor	61 (26)	28 (20)

**Table 11.** Mean interpersonal awareness ratings (and standard deviations). Scores are based on 8 analogue scales. There were ten three-person groups in the high visibility condition and a further ten in the low visibility condition (see Monk and Watts, in preparation).

## References

- Clark, H. H. 1996. *Using Language*. Cambridge, CUP.
- Goffman, E. 1976. Replies and responses. *Language in Society* 5: 257 - 313.
- Grice, H. P. (1957). Meaning. *Philosophical Review* 66: 377-388.

Heath, C. C. and Luff P. 1992. Collaboration and control: crisis management and multimedia technology in London underground line control rooms. *CSCW Journal* 1: 69-94.

Monk, A. F. and Watts L. 1997. Task analysis for collaborative work. *Human-Computer Interaction: INTERACT'97*, Sydney, Australia, Chapman & Hall.

Monk, A.F. and Watts, L. (in preparation) Peripheral participation in video-mediated communication.

Schober, M. F. and H. H. Clark 1989. Understanding by addressees and overhearers. *Cognitive Psychology* 21: 211 - 232.

Watts, L. A. and A. F. Monk 1997. Telemedical consultation: task characteristics. *CHI'97 Conference on Human Factors in Computing Systems*, Atlanta, Georgia, USA., ACM Press.

Watts, L. A. and A. F. Monk 1998. Reasoning about tasks, activity and technology to support collaboration. *Ergonomics* 41(11): 1583-1606.

Watts, L. A. and A. F. Monk 1999. Telemedicine: what happens in teleconsultation. *International Journal of Technology Assessment in Health Care* 15(1): 219-234.

**Acknowledgements:** The author would like to thank the UK Economic and Social Research Council for funding this work through their Cognitive Engineering Programme, also Leon Watts for his considerable contribution to the ideas described here.