

Telephone conferences for fun: experimentation in people's homes

Andrew F. Monk¹ and Darren J. Reed²,

¹ Centre for Usable Home Technology (CUHTec),
University of York, UK

² Department of Sociology, University of York, UK

Abstract. The paper illustrates how communication experiments may be carried out in a domestic recreational context. Participants situated in their homes were connected into group telephone conversations and simply asked to "chat and enjoy yourselves". Following the conversation, participants provided Likert scale ratings of the experience. In addition, the conversations were recorded and analysed. A total of 211 participants took part in two experiments. Telephone groups had an average size of five people, each speaking in the same conversation from their individual homes. Comments from the participants and Likert rating scales indicated that it was a positive recreational experience. The primary manipulation in each experiment was intended to encourage spontaneous co-involvement of all of the group. In Experiment 1 this was done by changing the way participants were introduced to the group, in Experiment 2 the group was made more salient by providing each member with a list of names. Open ended responses obtained from telephone interviews with participants in Experiment 1 were used to construct a questionnaire for Experiment 2 to measure presence, involvement and communication efficacy. The measures extracted from transcripts included the average length of utterances and equality of contribution as well as a new measure, the number of lines to the first "flow episode" in the transcript. While neither of the manipulations produced significant effects on Likert scale ratings made after the sessions or the measures extracted from transcripts, the paper is able to recommend the measures used and to provide practical advice for other investigators seeking to run communication experiments in a recreational context.

Please use the following format when citing this chapter:

Monk, A. F., Reed, D. J., 2007, in IFIP International Federation for Information Processing, Volume 241, Home Informatics and Telematics: ICT for the Next Billion, eds. Venkatesh, A., Gonsalves, T., Monk, A., Buckner, K., (Boston: Springer), pp. 201-214.

1 Introduction

1.1 Telephone conferencing for recreational purposes

In the home, communication technologies have an important role in the maintenance of social relationships. One-to-one phone calls, for example, are a common form of socialising. This paper explores the use of many-to-many phone connections for similar, purely recreational and social purposes. The paper also illustrates how field studies and ethnographic analyses can be used to inform the design of experimental studies of possible interventions expected to change the user experience of people socialising in telephone conferences.

People are so used to using the telephone for one-to-one conversations that they can find the notion of several people being connected into the same telephone conversation difficult to imagine. Telephone conferencing, which does just this, is widely used for business meetings but rather more rarely for recreation. There are exceptions to this rule. For example, the RNIB has equipment (telephone switches) that it uses to connect groups of people with visual disabilities for purely social purposes. There are also, of course, the premium rate chat lines that became notorious in the late 1980s when teenagers ran up massive bills on their parents' telephones.

A field study by the present authors [1, 2] looked at a scheme run by the Community Resource Team in Hackney, London. Their befriending scheme supports elderly individuals through recreational telephone conferences and weekly one-to-one telephone contact. The telephone conferencing switch and operator service for this scheme was provided by Community Network, a charity offering telephone conferencing facilities to other charities for business and social purposes. The telephone conferences consisted of four to eight older people with a volunteer facilitator. Most of the volunteers were also older people and were trained by Community Network. These bi-weekly link-ups were greatly valued by the group members. One person interviewed described the scheme as a "Godsend". Another reported that before each call she made herself comfortable in a particular chair with a cup of tea: "it's like waiting for someone to visit"[2].

In order to understand the nature of these multi-person recreational telephone conversations, tapes were obtained with the participants' permission from groups who had been meeting this way for some time. These were transcribed by the second author using Directed Conversational Analysis to describe the conversational structures that could be seen within the detailed transcripts [1]. In particular, one could contrast: (i) the rotating two-person conversations seen predominantly at the start of the transcripts, with (ii) incidents of spontaneous co-involvement, or flow, where the whole group of five or six people took part. The latter were of a quite different character, containing shorter, often overlapping, utterances.

The objectives of this work were: (i) to understand what makes for a good recreational telephone conference, and (ii) to make suggestions about how the technology or the procedures within which the technology is used may encourage such an experience. The results from [1] suggested spontaneous co-involvement as a

behavioural indicator of a good recreational telephone conference, as well as various possible manipulations that might encourage it. Many studies would have stopped at this point. However, we were interested to take it one step further and to try out two manipulations designed to encourage spontaneous co-involvement in quantitative experiments. To do this we had to devise a quantitative measure reflecting the behavioural changes we were aiming for. We also needed to develop an experimental procedure that would maintain as many of the important contextual features of mediated domestic socialising as possible while affording the experimental control needed to obtain meaningful results.

1.2 A "six-room" experiment where the rooms are people's homes

There is a tradition of communication experiments going back to the 1970s [3, 4] where people participate in some joint experimental task while in separate rooms using some sort of communication technology. These experiments compare different ways of technologically mediating communication and have become known as two-room experiments (for a summary of this and more recent research see[5]). Very little of this work has looked at mediated communication in groups of more than two people (but see[6, 7]). The aim of the experiments described below was to emulate the recreational telephone conferences pioneered by Community Network and others in an experiment. This would involve connecting six individuals, i.e., a six-room experiment.

It was decided that these experiments should be run using a conventional telephone conference with the participants situated in their own homes. The alternative would have been to bring them in to the laboratory.

One of the reasons that most of the research on mediated communication uses only two rooms is practical, i.e., finding the laboratory space. Connecting people in their own homes with a telephone conference offers a solution to this practical limitation. The other reason for not getting people into the laboratory for these experiments is methodological. While it is reasonable to generalise from experiments in the laboratory to work contexts, this is less reasonable with domestic and recreational contexts. Coming to the laboratory and then doing some task assigned by the experimenter is much like work. Conclusions drawn about technologies designed to support tasks that simulate work can then be (cautiously) generalised to parallel tasks in real work contexts. However, the laboratory is not like the home, even if one sets out a room with domestic furniture and decorations, the context will be artificial and unfamiliar. There will not be the same interruptions and distractions. Any conclusions drawn would be much harder to generalise. Thus while connecting people from their own homes will increase the error variance, it will also make the results more realistic.

When selecting the experimental "task" for these experiments, little attempt was made to influence how the conversations proceeded. In Experiment 1 participants were told "we just want you to chat and enjoy yourselves". In Experiment 2 there were some suggested initial topics and a very short warm up game, but again it was emphasised that they could talk about whatever they liked. This lack of control again makes for more error variance from variation in the way different groups approach

the conversation. However, this is natural variance and increases the generality of any conclusions drawn. For another experimental design that successfully confounds context and task with participant in this way see [8].

2. Experiment 1

The Directed Conversational Analysis of the tapes from Community Network (see [1]) identified a possible manipulation in terms of the way participants are introduced to the telephone conference. The effects of this manipulation are examined in this experiment.

2.1 Method

Design

Telephone groups consisted of 4 to 6 individual participants. Each group was randomly assigned to one of two experimental conditions. In the Individual condition participants were introduced to a single individual as they were brought into the telephone conference. In the Group condition they were introduced to the whole group. 92 participants took part in 18 telephone groups run using telephone conferencing switches at Community Network. Participants were recruited by email from the University of York Alumni Database of people willing to be contacted by email about university activities.

Once the instructions for introducing the groups to the conversations had been sent to Community Network we were in the hands of the operators. Examination of the tapes indicate that the assignments we had asked for were not always followed. Table 1 gives the actual allocation of groups by size and condition as confirmed from the tapes. There are two extra groups of size 5 in the Group condition. This equates to a small difference in average size (5.10 vs. 5.13) and so all 18 groups were included in the analysis.

Table 1. Number of groups of each size in the two introduction conditions in Experiment 1.

Size	Four	Five	Six
Individual	2	3	3
Group	2	5	3

Procedure

Starting times for the conference calls were arranged by telephone. In addition, participants were sent instructions indicating that they would be telephoned at the agreed time and date and that they would then be able to talk. To achieve the manipulation that formed the independent variable, the telephone operators at Community Network were asked to follow one of two scripts when introducing participants to a telephone group. In both cases the script involved ringing the first

person on the list provided by us and putting them on hold while the next was contacted. Once the operator had the second person on the list connected on their phone they would introduce them to the first and they could talk. At this point the two scripts deviated. In the Individual condition the process was continued with each new group member being introduced to the first person in. The operator would break into the conversation between the participants already introduced and then connect the new group member, e.g., if Anne was the first person to join the call it might go:

“excuse me Anne”

[Anne replies]

“hello Anne it’s the Operator, I have Peter joining you”

In the Group condition each new group member was introduced to the group as a whole, e.g.,

“excuse me everyone”

[conversation stops]

“hello everyone it’s the Operator, I have Peter joining you”

After approximately 30 minutes the operator broke in and warned the participants that the call would finish in 5 minutes.

After the call the participants were telephoned one at a time by the second author and interviewed about their experience of the conversation. The interview was structured around four statements and the participants’ reactions to them:

"It really felt like I was with the other people";

"I was very involved with the conversation";

"This is a great way to get to know people better";

"I was able to speak as much as I wanted".

Each participant was asked for a five point rating on the statement and then to elaborate on their answer. Finally, they were asked for "one good thing and one bad thing about the call". Both the conversation and the interview were recorded, with the participant's permission.

2.2 Results

Ratings and interview data

The sampling unit for the analyses presented in both experiments is the telephone group. The dependent variables derived from the analyses of conversation provide one score per group. While each participant provides a separate score for the dependent variables derived from ratings, these are not independent as, within a telephone group, they are based on the participants' experience of the same conversation. For this reason, ratings from participants were averaged to give a mean

rating per group before further statistical analysis in all of the analyses presented in the results sections of both experiments. For this experiment rating data was obtained for all but one participant (Group condition, group size 6).

Table 2. Mean (and standard deviation) ratings on a scale 1 to 5 where 5 is "strongly agree" in Experiment 1.

	Individual (n = 8)	Group (n =10)
1. It really felt like I was with the other people;	3.00 (.71)	3.00 (.35)
2. I was very involved with the conversation;	3.81 (.45)	3.55 (.33)
3. This is a great way to get to know people better;	3.72 (.72)	3.59 (.50)

Table 2 presents mean ratings for the four scales. The predicted advantage for the Group condition is not evident in any of the scales. A 2-way factorial analysis of variance was carried out for each of these mean ratings with condition as one between subjects variable and size of group as the other. There was no significant main effect of condition. There was a significant main effect of group size for statement 2, "I was very involved with the conversation" ($F(2,12) = 8.050, p = .006$; means, 4.063, 3.700, 3.360 for group sizes 4-6 respectively). This generally higher degree of involvement might have been expected in the smaller groups. There were no significant interactions.

Analyses of the conversations

The subjective experiences of our participants, as elicited in the post conversation interviews described above, represents their recollection and reconstruction of what happened and how they felt. While this is valuable evidence about the effects of the manipulation, we were particularly keen to assess the quality of the experience through behavioural measures obtained directly from transcripts of the conversations themselves. To this end, the first part of the conversations (about 15 minutes) were transcribed by trained transcribers naive to the experimental manipulation. Each transcript was checked by a second paid transcriber to ensure no utterances had been omitted and that the utterances were accurately attributed to participants. The transcription followed the normal convention in Conversational Analysis of continuing a long utterance on a new line, thus a turn could be more than one line long. All the quantitative analyses that follow use line, rather than turn, as the count indicating progress through the transcript. The length of a line is approximately 60 characters.

Line length as a measure of fluidity: Short mean utterance length has been suggested as an indication of fluency [9]. Here utterance length was measured as the average line length in characters and the percentage of continued lines. These are presented in Table 3. There is clearly no effect of the manipulation on any of these statistics. The same 2-way analyses performed on the ratings revealed no significant main effects or interactions for mean length for all lines, or percentage of continuation lines.

Table 3. Mean statistics (with standard deviation) extracted from the transcripts for the two experimental conditions in Experiment 1.

	Individual (n= 8)	Group (n=10)
Mean line length for all lines	36.2 (4.5)	37.1 (6.6)
Mean percentage of continuation lines	29 (7)	33 (14)
Equality of contribution	.963 (.026)	.942 (.047)
Mean lines to first flow episode	266 (237)	428 (309)

Equality of contribution: One would expect the condition where participants were introduced to the group to result in more equal contributions. Equality was measured using the information theory entropy equation $-\sum p \log_2(p)$, where p is the proportion of lines contributed by each participant. This has a maximum value when the proportions are all equal. These maxima depend on the group's size so each entropy was then divided by the maximum entropy attainable for that groups size to give a number between 0 and 1 (see [10]; also [11] for a discussion of similar measures). This statistic is also presented in Table 3. The difference between means was small and in the opposite direction to that expected. The same 2-way analyses performed on the ratings revealed no significant main effects or interactions for this variable.

Flow test: In all the conversations analysed, a pairwise conversational structure was observed to persist, even when all the participants had been connected to the conversation. The pairs would change but the structure is basically pairwise conversation: A talks to B for several turns, then A to C and so on. Reed [1] describes this in terms of Goffman's [12] frame analysis as conversation with a primary frame of seriousness. This primary frame can be broken by activities such as shared and invited laughter. At these points in the conversation the frame changes to one of play and the structure of the conversation becomes fluid and genuinely multi-party. During these states of what Goffman [13] calls spontaneous co-involvement, turns become shorter and there is much overlapping and latched speech. A similar phenomenon is described by Edelsky [14]. The phenomenon might also be thought of as a version of Csikszentmihalyi's flow [15] applied to the group rather than the individual, that is, a period of mutual and strong engagement in the conversation. For this reason we shall refer to these episodes in the transcripts as flow states.

The main hypothesis, derived from [1], was that the groups with introduction to a single individual would be slower to reach flow. Flow is defined as, spontaneous, relaxed and inclusive. It is composed of small turn utterances in close succession ("latched talk"), sometimes overlapping, but when this happens there is no competition. Everybody gets to take part. Topics are light and topic change occurs easily. There are no awkward silences.

To make analysing 18 15-minute transcripts for this first flow event tractable the following procedure was followed. The transcripts were transferred to a spread sheet that was programmed to examine a window of $2n$ lines from the current line where n is the size of the group. If this window contained a line from all n participants the initial line was automatically flagged by the spreadsheet as the start of a potential flow episode. The second author then examined these flagged $2n$ windows wherever there were n consecutive flagged lines in a row. This was a three stage process.

1. A sequence was not counted as flow if any of the lines in the 2n window was one of the following:

- (i) a long silence (10 seconds or more);
- (ii) administrative business, e.g., "who is here", "how many people are expected";
- (iii) off call, out of call or about call business, e.g. "who's that in the background", "can anybody hear that noise";
- (iv) any part of a routinized turn in the round where there is an expectation everyone should reply, e.g., greetings, "where does everybody live" (and playing the guessing game in Experiment 2).

2. A sequence was not counted as flow if the only line attributed to a participant in the 2n window was:

- (i) a non-significant turn utterances or an unrecognisable verbalisations, e.g., a grunt;
- (ii) a laughter token or minimal turn utterance such as "mm" or "aha".

3. The content was examined to see if the participants did indeed seem to be enjoying themselves. If they had been arguing, for example, the sequence would have been discounted. In the event, while there was disagreement, there were no potential flow sequences with content of this kind.

If none of these exemptions applied, the starting line was marked as the start of a flow episode.

The number of lines between the introduction of the last member of the group and the start of the first flow episode was then computed. If the group never reached flow the value entered was the line number of the last line in the transcription. The mean lines to the first flow episode are given in Table 3. While there appears to be a relatively large difference between these means it was in the opposite direction to that predicted and the standard deviations were very large. The same 2-way analyses performed on the ratings revealed no significant main effects or interactions for lines to first flow episode in this experiment.

Conclusions from Experiment 1

Changing the way that participants are introduced had no significant effect on the ratings or on the measures extracted from transcripts. While this is a null result, and hence may be due to a lack of sensitivity in the measures used and the experimental design rather than a real lack of effect, one may tentatively conclude that this manipulation probably does not reflect a practically important design consideration.

3. Experiment 2

Experiment 2 looked at the possibility of encouraging flow episodes with visual aids as an addition to the voice channel. For half the groups, instructions sent prior to the session contained a list of the names of the participants, in an attempt to make more salient the fact that they were talking to a group. By making the group more salient in this way we expected to encourage a quicker movement to spontaneous co-involvement.

3.1 Method

Design

Each telephone group of 4 to 6 individual participants was randomly assigned to one of two experimental conditions. In the With List condition participants had a printed list of the participants in the group with their instructions for the session. In the Without List condition the instructions did not include this list. 119 participants took part in 24 telephone groups. 94 were recruited by email from the alumni database of which 50 had taken part in Experiment 1. Five groups were recruited by letter from a database of older adults in Northumbria. There were two of these latter 5 groups in the With List condition, and 3 in the Without List condition.

Again, equipment failures and other misadventures mitigated against efforts to balance the group sizes in each condition. Also there were less groups in the transcript analysis than the ratings analysis as some groups lacked tapes (see table 4). The difference in average size is 4.92 vs. 5.00 for the ratings analyses and 4.67 vs. 5.09 for the transcript analyses.

Table 4. Number of groups of each size in the two instruction conditions in Experiment 2. The reduction in the number of groups providing transcripts is due to equipment failure recording four groups.

Size	Ratings			Transcripts		
	Four	Five	Six	Four	Five	Six
With List	5	3	4	5	2	2
Without List	4	4	4	3	4	4

Questionnaire

A questionnaire was constructed using the recorded interviews data from Experiment 1. These interviews were transcribed and coded using grounded theory analysis [16]. These codes were then grouped into three themes: presence (e.g., how much it felt like one was with the other people); involvement (e.g., how much they had or had not felt immersed in the conversation or how positive an experience it was) and communication efficacy (ways in which mediation affected their ability to communicate, because they could not see each other or recognize each others' voices). Quotes were then selected to illustrate the themes and these quotes converted into agree/disagree statements for use in a questionnaire obtainable on request from the authors.

Procedure

The procedure was similar to that in Experiment 1 (Group condition only). Participants were sent a letter with details of when their call would occur. They were told they could talk about anything but two suggestions were made to get the discussion going: the recent accolades received by the University of York in a Sunday Times poll, and a planned new campus for the university. In order to give both groups a reason to have these instructions in front of them, and thus also a list of participants in the With List condition, each participant was given the name of a

different city to use in a brief guessing game. The letter also enclosed the 19 item questionnaire described above to be filled in and returned after the call. The operator introduced each new participant to the group as in the Group condition in Experiment 1 and again warned them when 5 minutes was left.

3.2 Results

Ratings

Mean ratings were computed for three composite scores corresponding to presence, involvement and communication efficacy. These scores were then averaged for the participants in each telephone group, as in Experiment 1, and this mean score entered into the analyses presented below. Not all participants returned their questionnaires. The return rate was 87% for the With List condition and 89% for the Without List condition.

Reliability coefficients for the questionnaire derived from a content analysis of the discussion following Experiment 1 were computed for the three subscales using the averaged group data that was that the basis of the comparisons carried out below ($N=24$). Presence has an alpha of .64. Involvement has an alpha of .90 and Communication efficacy .55. Together the complete 19 item test has an alpha of .88 and we would therefore recommend it to other investigators needing a measure of conversational experience in mediated communication.

The means of these scores for each condition are given in Table 5. It can be seen that there is very little difference between the two conditions. As in Experiment 1 a 2-way between subjects analysis of variance was carried out with condition and group size as independent variables for each of the three composite scores. There were no significant main effects of condition or group size in any of the three measures.

Both Involvement and Communication Efficacy had significant 2-way interactions ($F(1,18) = 4.483$, $p = .026$; and $F(1,18) = 3.685$, $p = .046$; respectively). This was due to group size 5, showing the expected advantage of making the group salient, whereas the other two groups sizes show the opposite effect.

Table 5. Mean (and standard deviation) questionnaire scores on composite rating scales for Experiment 2.

	With List (n = 12)	Without List (n = 12)
Presence score	2.81 (0.29)	2.83 (0.40)
Involvement score	2.73 (0.80)	2.95 (0.83)
Communication efficacy score	3.08 (0.37)	3.18 (0.30)

Analyses of the conversations

As in Experiment 1 the first part of the conversations were transcribed by trained transcribers naive to the experimental manipulation. The mean length of characters per line and percentage of continuation lines was computed as in Experiment 1 and means for these measures are included in Table 6. The 2-way analyses of these

variables showed no significant main effects or interactions. Finally, the equality of contributions was computed for each group. Here the 2-way analysis showed no significant main effect of condition or interaction, but a significant main effect of group size ($F(2,15) = 4.841$, $p = .024$; means, .987, .960, .953 for group sizes 4-6 respectively). This also accords with expectation and replicates the results of Carletta et al. [10] who noted a similar negative correlation between group size and equality using a slightly different statistic (see also [11]).

Table 6. Mean statistics (with standard deviation) for process measures extracted from the transcripts for the two experimental conditions in Experiment 2

	With List (n = 10)	Without List (n = 11)
Mean line length for all lines	30.7 (4.7)	30.0 (7.2)
Mean percentage of continuation lines	23.2 (6.4)	21.8 (11.3)
Equality of contribution	.969 (.030)	.968 (.020)
Mean lines to first flow episode	254 (219)	300 (181)

Mean lines to flow from the line where the last participant entered the conversation were computed as described in the results of Experiment 1. Here the results are in the right direction (see Table 6) but again the standard deviations are large. The 2-way analysis of variance showed no significant main effect of condition, and no significant interaction but this time there was a significant effect of group size with the size 4 groups reaching flow in less than half the number of lines of the other group sizes ($F(2,15) = 5.054$, $p = .021$; means, 134, 318, 423 for group sizes 4-6 respectively). One would expect a small group to find it easier to meet our criteria for flow and so this result confirms that the measure has some validity and sensitivity to this manipulation in the experiment.

4. General discussion

4.1 Running experiments in people's homes

The studies described above demonstrate the feasibility of running formal quantitative communication experiments in a recreational setting. A degree of realism was achieved by having the participants join in from home and simply asking them to "chat". As in our earlier studies, participants were connected from their own homes and the conversation was largely undirected. Unlike the telephone conferences in the field studies, however, participants had not met before. Nevertheless, the conversational mechanisms observed were very similar, giving some credence to the idea that there will be a primary frame of seriousness that can be broken with bouts of spontaneous co-involvement in all telephone conferences.

One can also argue that the context was a truly recreational one. Despite the fact that they had been asked to meet with a group of strangers, participants generally enjoyed the telephone conferences. In Experiment 2 the questionnaire included a

question to tap into this. The average rating to the statement "I found the call to be an extremely positive experience" was 3.7 in the With List condition and 3.8 in the Without List condition where 1 is strongly disagree and 5 is strongly agree. Given the "extremely" and "strongly" qualifiers, this is a positive endorsement for this recreational experience. Also, as part of the questionnaire used in Experiment 2, participants were invited to make any additional comments they wanted. The large majority of comments were positive about the taking part in the telephone conference. It was 'a most enjoyable exercise' and a 'great experience'.

4.2 The effect of the experimental manipulations

Changing the way that participants are introduced, in Experiment 1, seemed an obvious way of influencing the specific target we had set ourselves, that is to reduce the amount of pairwise conversation and hasten the onset of the first instances of flow. Had the manipulation been successful the results would have had important implications for the running of recreational telephone conferences. In the event it is difficult to know whether the null result obtained is due to the general insensitivity of the experimental design or whether it is really unimportant. The number of groups, 8 in one condition and 10 in the other, should have been sufficient to detect a practically important effect. Tentatively one may conclude that this manipulation has no practically interesting effect on the conversations. The fact that the trend in the results was actually in the opposite direction to that expected leads us to speculate that there may have been two effects working against each other in this experiment. Introducing each new participant to the whole group may have emphasised the group, but introducing each person to the same individual may have made that person a kind of facilitator. If the randomly chosen person to whom the introduction was made facilitated better conversations the result may have been a more rapid move to spontaneous co-involvement. Had they been less effective as facilitators, the other effects may have come through resulting in the large amount of variance observed.

In Experiment 2 the group was made more salient by providing a list of names. The idea was that by being able to see a list of all the group members, each participant would be made continually aware that the current speaker was not the only person there. A number of possible variants on this theme were considered when designing this experiment: making one group anonymous, providing biographies or pictures; they were rejected as all of these variants introduced additional factors and additional practical constraints. The result was a relatively small intervention that appears to have had no effect, although here again the cautions needed when interpreting null results apply. Given different equipment constraints it would be very interesting to explore the use of some of the graphical awareness tools proposed for audio and text conferencing, where the present status of a participant (present/absent, speaking/silent) is graphically represented via avatars.

4.3 Combining field studies with experiments

Perhaps the most important contribution of this paper has been to demonstrate the value of combining experiments with field studies. The benefits work both ways. For example, spontaneous co-involvement (conversational flow) was identified in the field study as a desirable feature of group conversations. The measure, lines to flow, directly resulted from this characterisation. However, in addition, translating the concept into an operationalised quantitative measure that could be used to compare different implementations required a much more clearly defined version of the idea. The scoring scheme identified in section 2.2 can thus be viewed as a clarification of the field studies, resulting from having to design an experiment. The same is true of the experimental manipulations, which, while unsuccessful, do serve to operationalise and hence clarify the design goal we were pursuing.

This paper demonstrates the feasibility of doing communication experiments in the realistic recreational context of people's own homes. It also takes us some way towards the goal of devising practical and sensitive measures for assessing enjoyment directly from user behaviour. Lines to flow is a quantitative measure derived directly from the participants' conversation, reflecting a potentially important aspect of enjoyment. This work is now being extended in this laboratory to examine other measures of group behaviour [17]. Such measures provide new ways of thinking about user experience in addition to the more commonly collected post-experience subjective ratings.

Acknowledgements

This work was funded by grant L328253006 to the first author from the UK ESRC via the PACCIT programme. We are grateful to other participants in the PACCIT programme and members of the HCI group at York who have helped to shape this research.

References

1. Reed, D.J., Fun on the phone: the situated experience of recreational telephone conferences, in *Funology: from usability to enjoyment*, M.A. Blythe, et al., Editors., Kluwer: Dordrecht, the Netherlands (2003) 67-79
2. Reed, D.J. and A.F. Monk, Using familiar technologies in unfamiliar ways and learning from the old about the new. *Universal Access in the Information Society*, 3 (2004) 114-121
3. Short, J., E. Williams, and B. Christie, *The social psychology of telecommunications*, London: John Wiley and Sons (1976)
4. Chapanis, A., Interactive Human Communication. *Scientific American*, 232 (1975, March) 36 - 42
5. Finn, K.E., A.J. Sellen, and S.B. Wilbur, *Video-mediated communication*, Mahwah, New Jersey: Lawrence Erlbaum Associates (1997)
6. Anderson, A.H., J. Mullin, R. Katsavras, R. McEwan, E. Grattan, P. Brundell, and C. O'Malley, Multimediating multiparty interactions, in *Interact'99*, Edinburgh: Amsterdam: IOS Press (1999)
7. Monk, A.F. and L.A. Watts, Peripheral participation in video-mediated communication. *International Journal of Human-computer Studies*, 52 (2000) 775-960

8. Monk, A.F., E. Fellas, and E. Ley, Hearing only one side of normal and mobile phone conversations. *Behaviour and Information Technology*, 23 (2004) 301-305
9. Daly-Jones, O., A.F. Monk, and L.A. Watts, Some advantages of video conferencing over high-quality audio conferencing: fluency and awareness of attentional focus. *International Journal of Human-Computer Studies*, 1998, 49(1): p. 21 - 59.
10. Carletta, J., S. Garrod, and H. Fraser-Krauss, Communication and Placement of Authority in Workplace Groups: The Consequences for Innovation. *Small Group Research*, 1998, 29(5): p. 531-559.
11. Fay, N., S. Garrod, and J. Carletta, Group discussion as interactive dialogue or as serial monologue: the influence of group size. *Psychological Science*, 2000, 11(6): p. 481-486.
12. Goffman, E., *Frame analysis. an essay on the organisation of experience*. 1974, Boston: North East University Press.
13. Goffman, E., *Fun in games*, in *Encounters: two studies in the sociology of interaction*, E. Goffman, Editor. 1961, Bob Merril: Indianapolis.
14. Edelsky, C., Who's got the floor. *Language in Society*, 1981, 10: p. 383-421.
15. Csikszentmihalyi, M. and K. Rathunde, The measurement of flow in everyday life: towards a theory of emergent motivation. *Nebraska Symposium on Motivation*, 1993, 40: p. 57-97.
16. Strauss, A. and J. Corbin, *Basics of qualitative research: grounded theory procedures and techniques*. 1990, Newbury Park: Sage.
17. Lindley, S.E. Designing interfaces to afford enjoyable social interactions by collocated groups. *Proceedings of CHI 2005, Extended Abstracts*, 2005, 1122-1123. New York: ACM Press.