

Toward supporting group dynamics

Fabio Pianesi, Massimo Zancanaro, Vera Falcon, Elena Not
ITC-irst
Via Sommarive, 18
38050 Povo Trento Italy
{pianesi,zancana,falcon,not}@itc.it

Abstract. The complexity of group dynamics occurring in small group interactions often hinders the performance of teams. The availability of rich multimodal information about what is going on in meetings makes it possible to explore ways of providing support to dysfunctional teams from facilitation to training sessions, addressing both the individuals and the group as a whole. A necessary step in this direction is that of capturing and understanding group dynamics. In this paper, we discuss a particular scenario, in which meeting participants receive a multimedia feedback on their relational behavior, as a first step towards increasing self-awareness. We describe the background and the motivation for a coding scheme partially inspired by the Bales' Interaction Process Analysis aimed at identifying suitable observable behavioral sequences and an experimental investigation on the acceptability of such a service.

1 Introduction

Most of the current research work in using multimodality to support group interaction is aimed toward providing easy access to computerized services for the group to efficiently accomplish its tasks [12]. For example, in the CHIL project, most of the services provided are aimed at offering better ways of connecting people (the Connector service) and supporting human memory (the Memory Jog) [32]. The research in the AMI project mostly focuses on off-line multimedia retrieval and multimedia browsing of information obtained from meetings [26]. The DARPA-funded project CALO supports a group in creating a project schedule by automatically interpreting gestures and speech, including the learning of new words [21].

In the field of CSCW where the focus is often in distributed meetings, the social relationships among the participants of a meeting has been recognized as a

Please use the following format when citing this chapter:

Pianesi, Fabio, Zancanaro, Massimo, Falcon, Vera, Not, Elena, 2006, in IFIP International Federation for Information Processing, Volume 204, Artificial Intelligence Applications and Innovations, eds. Maglogiannis, I., Karpouzis, K., Bramer, M., (Boston: Springer), pp. 302–311

fundamental aspect of the meetings' efficacy since the seminal work of Tang [30]. Although in face to face interactions visibility and awareness of others' behaviour is not a problem, the success of a meeting is often hindered by how participants behave. According to a survey in [14] professionals agree that as much as 50% of meeting time is unproductive and that up to 25% of meeting time is spent discussing irrelevant issues. In order to improve meetings performance, external interventions such as facilitators and training experiences are commonly employed. Facilitators are neutral and non-evaluating servants of the meeting group. Their role is to help participants maintain a fair and focused behavior during the meeting, directing and setting the pace of the discussion. They help and solicit the group agreement about content to be analyzed and about the correct process to adopt to reach a viable solution. They protect individuals from personal attacks and are responsible for maintaining an open and balanced conversational flow (preventing monologues and assuring everyone gets a change to express their ideas).

Ideally, systems for CSCW in co-located settings should include virtual facilitators who "observe" and analyze the behavior of meeting participants, by means of audio/visual sensory devices, and intervene to mediate the discussion. Obviously, this is an extremely challenging design- and implementation-task, requiring incremental steps of problem dissection and solving. First and foremost, the acceptability of suggestions and directives based on social observations coming from a virtual facilitator need to be assessed. Would a meeting participant consider as reliable and acceptable suggestions like: "Bob, you're sure doing a lot of talking. Let's hear from some of the other people¹", when coming from a virtual facilitator? Would he be willing to change his behavior after that, therefore impacting on the group dynamics? And, more in general: would he be willing to be observed and analyzed in his social behavior? Indeed, the shift of focus from task-based functionalities commonly supported in CSCW systems to functionalities targeting the social behavior is not only difficult from a technical point of view but also challenging from a design point of view since it can be expected that these types of services trigger complex, and possibly negative, reactions from the users. The problem is too rich of complex facets to be addressed as a whole.

To take a step at a time, our research is currently focusing on the latter two questions above: would meeting participants accept to work in an augmented environment whose task is not merely to provide technological support but also that of deliberately keeping trace of their behavior and understanding their role in the discussion? To better isolate the variables to be evaluated, we decided to shift from the facilitator-scenario, where high interactivity and human-computer dialogue is involved, to a simplified scenario where the meeting participant is confronted with the external, virtual, observer just at the end of the meeting. We therefore investigated the feasibility, the usefulness and the acceptability of a functionality inspired by *coaching*; it consists of a report about the social behaviour of individual participants that is generated from multimodal information, and privately delivered to them. The underlining idea is that the individual, the group(s) they are parts of, and the whole organization might benefit from an increased awareness of participants about their own behavior during meetings.

¹ This sample sentence has been taken from [14], page 101.

The main goal of this paper is to discuss the users' response to such a multimodal system. In order to investigate the grounds of our approach, we also present a coding scheme for annotating observed group dynamics as a prerequisite for the actual realization of social-aware services (of which the relational report is one example).

2 Supporting Group Efficacy by Relational Reports

In social psychology, the “Ringelmann effect” or “social loafing” postulates that the individuals' performances decrease with the size of the group [23; 20]. Indeed, there is evidence that this effect has a lesser impact when the amount of contribution of each participant can be compared to the contribution of the others [18]. The usefulness of an external feedback has also been considered in the multimodal coaching field [7; 25], based on the support provided to users by embodied conversational agents.

In discussing the role of collaboration for teachers, Andersen [1] suggests that peer coaching sessions among teachers provide a scheduled opportunity to step out of the reflexive mode and think reflectively, and that the coaching process allows the externalization of both tough contents and processes that are normally internal, making them available to examination. Reflective thinking is fundamental in the development of meta-strategic knowledge [24]. That is, the kind of knowledge that experts possess about their own thinking and that includes information about the effectiveness and applicability of various strategies. There are three stages in the reflective process [10]: (i) the return to experience (what happened?); (ii) attending to feelings (how did I feel, why did I act or react this way?); and (iii) the re-evaluation of the experience (what does it mean?). Reflective thinking, as noted by Andersen [2], can be effectively fostered in coaching relationship. The coach is a “different observer” than the coachee; by bringing a different perspective to the relationship, the coach can see both circumstances and possibilities that the coachee can't [9].

In our work we propose a multimodal system that monitors the group behavior and generates individual reports about the participants' behavior. The system observes the meeting as a coach would do, and not as a recorder. This means that the system does not keep trace of exactly what people said and what they did as the meeting goes by. The generated reports are not minutes, but represent a more qualitative, meta-level interpretation of what happened in the social dynamics of the group. They do not contain information like “in the first part of the meeting you have talked for ten minutes about machine learning techniques useful to solve the problem” but rather “in the first part of the meeting you have provided the group with background information” or “you have prevented others from intervening in the discussion”. The reports are delivered privately to each participant after the meeting, with the purpose of informing them about their behavior rather than evaluate it. Hence, the system acts as a coach for the individual group participants. Since the coaching relationship is based upon trust and permission [13], we may expect problems regarding acceptability to arise.

2.1 A Sample Multimedia Relational Report

The technological plausibility and feasibility of the automatic generation of these *relational reports* is supported by the fast growing results gained in the field of multimodal multiparty meeting processing [25], which make more reliable the multimodal perception and understanding of a wide array of information sources (speech, voice tone and prosody, gestures, handwriting, sketches and other manual activity, body and head posture, eye gaze, ...).

The relational reports are generated by a rule-based system starting from a “social” annotation of the meeting according to the coding scheme described in the next section. The report is built according to the task and socio-emotional roles simultaneously assumed by the speaker during the interaction. Each final relational report has the form of a multimedia presentation where a talking head describes the behavior of the participant in an informative rather than normative way. The presentation is enriched with short audio-video clips from the actual meeting that exemplify the information presented. An example of the relational reports we are considering is the following: “You have actively contributed to the meeting,” while showing a neutral expression. Then, the talking head changes into an happy expression, and continues “You’ve helped to focus the discussion on the relevant topics and have provided useful information and opinions to clarify some of the issues. At the beginning and at the end of the meeting, you have kept records of various aspects of group progress.” Finally, moving to a sad expression she adds: “You have profitably cooperated with your colleagues, even though a contrast with R.Z. emerged during the first part of the meeting.”

3 Observing Group Behavior

A coding scheme for annotating group behaviour should, in the first place, be usable by human annotators. This is needed both for exploratory research aiming at understanding relevant social phenomena and assessing the coding scheme, and to provide data bases that can be used to train systems. Then, the categories of coding scheme must be capable of mapping onto constellations of low-level patterns that can be detected through vision and speech.

Among the available notions of group members’ roles [22], that defining them in terms of behaviour enacted in a particular context was of particular interest to us. Indeed, that notion moves away from a strictly organizational perspective in which roles are defined by the social positions within the group, and it differs from approaches defining roles according to the social expectation associated with a given position [22]. In other words, functional roles allow exploiting information about what actually happened in the course of the interaction, while reducing the necessity for knowledge about the group’ structure, history, position in the organization, etc.

Benne and Sheats [6] provided a list of “functional roles” recognizable in working groups, according to which interest is predominant in the behaviour: task-oriented, maintenance-oriented or individual-oriented. The first two kinds of roles are directed toward the group’s need: *task-oriented roles* provide facilitation and

coordination, while *maintenance roles* contribute to structure and preserve interpersonal relations, in order to reduce tensions and maintain smooth group functioning. The third type of roles, the *individual roles*, is “individual-centred” and performed in order to reach individual need rather than group’s goals. During the interaction, each person can enact more than one role. Even if a definite answer to the quest for “optimal” group productivity is not provided, the authors emphasise the threat due to a strong presence of individual roles, and the importance of achieving the task and maintaining the relationship between members.

Similarly, Bales [4] proposed the Interaction Process Analysis, a framework to study small group interaction, by classifying roles in face-to-face interaction in a two-dimensional space.

Building on Benne and Sheats’s functional roles and on Bales’ two dimensional approach, and drawing on observations performed on a set of face-to-face meetings, a coding scheme was produced consisting of five labels for the Task Area and five labels for the Socio Emotional Area. The *Task Area* includes functional roles related to facilitation and coordination tasks as well as to technical experience of members, while *Socio Emotional Area* concerns the relationships between group members and the functional roles oriented toward the functioning of the group.

The Task Area roles are the following.

Orienteer: is the person orienting the group. S/He introduces the items on the agenda, defining the position of the group relative to the goals and helps keeping the group focused and on track. S/He summarizes the main ideas of the group, recording the most important arguments in the discussion, the minutes, and the group decisions. S/He spells out suggestions in terms of examples or develops meanings, offers a rationale for suggestions previously made and tries to deduce how an idea would work out if adopted by the group.

Giver (Information, Opinion, Suggestion): is the person providing factual information and answering to questions: s/he often has expertise on a given topic. S/He states his/her beliefs and attitudes about an idea: expresses personal values as opposed to factual information; shows or clarifies the relationship or linkage among various ideas and suggestions, trying to pull them together.

Seeker (Information, Opinion, Suggestion): S/He requests suggestions and information to promote effective group decisions. S/He can also ask for clarification of values and opinions expressed by other members of the group.

Procedural technician is the person who does something for the group. S/He uses the material resources available to the group, managing them for the sake of the group.

Followers: S/He listens, does not participate actively to the interaction.

The roles pertaining to the Socio Emotional Area are the following.

Attacker: S/He may work in many ways – deflating the status of others, expressing disapproval of the values, acts or feelings of others, attacking the group or the problem it is working on, joking aggressively, showing envy toward another’s contribution by trying to take credit for it. S/He consistently reacts negatively to other’s ideas: makes very critical comments, usually indirectly using humor.

Gate-keeper: S/He is the moderator within the group, who mediates the communicative relations: s/he attempts to keep communication channels open by encouraging or facilitating the participation or by proposing regulation of the flow of

communication. S/He mediates the differences between other members, attempts to reconcile disagreements, relieves tension in conflict situations.

Protagonist: S/He takes the floor without need to be consulted driving the conversation. S/he assumes a personal perspective asserting his/her authority or superiority because of his/her status or because of the particular task she/he is performing.

Supporter: S/He shows a cooperative attitude indicating understanding, attention and acceptance as well as providing technical and relational support to other members of the group. S/He also keeps a collaborative climate sharing the common objects and trying to make them available to each member.

Neutral: Going along with the group, s/he passively accepts the idea of others, serving as an audience in group discussion.

Assessing the coding scheme

In order to assess our coding scheme, two independent judges have manually annotated four hours and half of video recorded meetings. In particular, three hours and half have been considered for the Socio Emotional Area and one hour for the Task Area.

Video snippets are annotated with t-uples like <task; p; o; start:1.796; end:170.066; duration:168.269> where the various elements in the tuple respectively indicate: (i) the type of occurring role (whether in the Task Area or Socio Emotional Area); (ii) an id code identifying the observed participant; (iii) an id code identifying the observed role (e.g. “o” for observer); (iv) the start time of the observed role; (v) the end time of the observed role; (vi) its duration. The annotators’ agreement has been assessed using the K of coefficient [11], considering the number of seconds as measure. The mean score was 0.88 for the Task Area and 0.8 for the Socio Emotional Area. The coding scheme can therefore be considered reliable [see Falcon et al. 2005].

The assessment of a reliable coding scheme is a vital step before proceeding to define procedures for the automatic perception, fusion, interpretation and annotation of the multimodal input coming from sensory devices which monitor the meeting. Within CHIL, we are currently investigating which verbal and non-verbal multimodal patterns can be used to automatically identify the social roles.

3.1 From coding to reports

The example presented in section 2.1 was manually built from the annotations of a real meeting. For each participant, a table was compiled showing the distribution of his/her roles. For each role, we took into consideration the total number of seconds the relevant participant played that role, as well as the role’s distribution during the meeting (for the sake of simplicity, we considered the meeting as divided into quarters).

A number of patterns were then elicited and mapped to verbal statements that describe the behaviour. The pattern “high Protagonist+Orienteer held for XY part of the meeting” is mapped onto the sentence “You have actively contributed to the XY

part of the meeting". Similar patterns were then used to build the rest of the example above.

More elaborate strategies involve reasoning about the behaviour of various participants at a time. For example, should the participant for whom the report is prepared have maintained high percentages of "orienteer" role especially at the beginning and at the end of the meeting with the others being silent or neutral, the report could include a statement like "at the beginning of the meeting you have helped define the agenda and initiate the discussion, summing up the outcome of the meeting at the end". At the opposite, should the considered participant have maintained significant percentages of "orienteer" and "seeker" roles for the most part of the meeting, a statement as the following could be included: "During the meeting you have played a leading role, defining discussion topics and soliciting your colleagues' participation." In case sequences of seeker-(recorder)-attacker are observed, the report could be complemented with: "in some cases, however, you have displayed a critical and aggressive behaviour as a response to your colleagues' contribution".

Within CHIL, there is an on-going effort aimed at the automatic generation of relational reports starting from meeting annotations. The system is based on an existing strategic text planner which accesses a repository of declaratively defined discourse schemata based on the pattern rules described above. The linguistic realization of sentences is currently template-based. Notwithstanding the limitations of the current, preliminary, implementation of the generator, a positive response to the technological feasibility of the system has emerged. An example of brief report actually automatically generated from real data is the following: "You have actively contributed to the meeting, supporting the discussion with your knowledge, helping here and there to focus the discussion on the relevant topics, taking notes. You've had a particular involvement especially at the end of the discussion."

4 An initial evaluation of the relational reports

In this section we present and discuss a study we conducted to understand the way people would react to the very idea of the relational report. In particular, we were interested in investigating four different dimensions: the perceived usefulness of such a service; its reliability — that is, whether people think that an automatic system can reliably provide a report on such a delicate matter as individual behaviour in group situations; its intrusiveness — that is, the perceived degree of intrusiveness of a service that monitors group and individual behaviour to provide reports on their relational behaviour; and its acceptability — what affects the acceptance of the report by addressees?

4.2 The experiment

Eleven groups of four people each were requested to enter in a structured discussion of about half an hour following the Mission Survival Task².

All the groups were video-recorded using four fixed omni-directional cameras, closed-talk microphones and seven T-shaped microphone arrays, each consisting of four omni directional microphones. There was no attempt to hide the recording devices since one of the purposes of the experiment was to evaluate the acceptability of being recorded. All the participants were non-technical persons and they had been told that the purpose of the experiment was data collection for the development of a multimodal system.

Few days after, the participants received an individual report elaborated by a social psychologist. Each report described the behaviour of the participant in terms of the functional roles played during the meeting. The psychologist took the roles of the coding scheme as a reference, adopting a descriptive style and without mentioning explicitly the role labels (for example, the "Orienteer/Protagonist" label was paraphrased as "[...] She initiates the discussion by proposing an importance order, justifying it and using a quiet tone of voice [...]"). In writing the reports, the psychologist considered only behavioural aspects of the participants, such as the posture and the tone of voice, and not aspects related to content such as the individual contributions to the discussion.

Half of the participants were told that their report was automatically elaborated by an intelligent system able to monitor the groups' behaviour, while the other half (i.e. the control group) were told that the report was written by a psychologist.

The attitude toward the report was tested by a seven item questionnaire aimed at assessing the perceived usefulness, its reliability, the perceived degree of intrusiveness and its acceptability. A semantic differential aimed at assessing the appropriateness, the completeness and the clarity of the report (the semantic differential was part of the 6-scale questionnaire proposed by Garrison, 2003 with a Cronbach alpha of 0.9482) was also used.

4.1 Results

The answers to the questionnaire were analyzed with a two-tailed multivariate ANOVA ($p=.05$), considering 42 questionnaires: half for the "expert source" of the report and half for the (pretended) "system source" of the report. The independent variable was the source of the report, in order to monitor how it affects the dimensions investigated in the questionnaires items (dependent variables). Generally, there were no statistically significant differences among the questionnaire's responses in the two groups. The subjects' attitudes were more positive toward the system source (though not is a significant way) in regard to: (i) the perceived usefulness of the report for improving their own relational behaviour ($F(1,40)=.366$), (ii) the perceived usefulness for improving interactions in meetings ($F(1,40)=$

² This task consists in having the group discussing how to survive in a disaster scenario, like moon landing or a plane crashing in Canada. This kind of test has often been used in experimental and social psychology to elicit discussions in groups.

.143), and (iii) the willingness to remember the report observations ($F(1,40) = .175$). Vice versa, the subjects' attitudes were more positive toward the human expert in regard to: (i) the usefulness of the report for stimulating reflection on behavioural aspects not considered in the past ($F(1,40) = 2.138$), (ii) the completeness of the report in catching relevant behavioural aspects ($F(1,40) = .293$ for caught and $F(1,40) = .518$ for not caught aspects), and (iii) the analysis ability of the expert ($F(1,40) = 1.675$).

Regarding the subscales of the semantic differential, they were analyzed with a two-tailed multivariate ANOVA with $p = .05$. In this analysis, a more positive attitude toward the human expert emerges, but not at a statistical significant level. The more relevant difference emerged in the appropriateness sub-scale ($F(1,40) = 4.007$, $p = .05$, only marginally significant), less evident in completeness ($F(1,40) = 2.079$) and clarity ($F(1,40) = .901$) sub-scales.

5 Conclusion

In supporting co-located groups, multimodality is generally used to help the group in getting easier access to computerized services. In this paper, we proposed that it can be employed to provide meta-level services that are meant to impact on the group dynamics. This shift of focus from task-based functionalities to functionalities targeting the social behaviour is not only difficult from a technical point of view but also challenging from a design point of view since it can be expected that these types of services trigger complex, and possibly negative, reactions from the users.

As a case study of such a typology of services, we investigate a functionality inspired by coaching: it consists of a report about the social behaviour of individual participants that is generated from multimodal information, and privately delivered to them. We provide a motivation for this functionality by drawing from studies on the so-called "reflective process" and by discussing the benefit of coaching provided by humans to other humans. The underlining idea is that the individual, the group(s) they are parts of, and the whole organization might benefit from an increased awareness of participants about their own behaviors during meetings. Since the coaching relationship is based upon trust, we expected problems regarding acceptability to arise when an automatic system enters into play. Yet, the initial evaluation conducted in a controlled experiment where the experimental group was told that the report was prepared by an automatic system did not show any statistical difference in acceptability with respect to the control group that received a report prepared by a human expert.

References

1. P. R. Cohen, R. Coulston, , and K. Krout, Multimodal Interaction During Multiparty Dialogues: Initial Results. In *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, October 14-16, Pittsburgh, PA, 2002, pp.448-453.

2. A. Waibel, H. Steusloff, and R. Stiefelhagen, CHIL: Computer in the Human Interaction Loop. In NIST ICASSP Meeting Recognition Workshop, Montreal, Canada, 2004.
3. I. McCowan, D. Gatica-Perez, S. Bengio, D. Moore and H. Bourlard Towards Computer Understanding of Human Interactions. In: Ambient Intelligence, E. Aarts, R. Collier, E. van Loenen & B. de Ruyter (eds.), Lecture Notes in Computer Science, Springer-Verlag Heidelberg, 2004, pp. 235-251.
4. E. Kaiser, D. Demirdjian, A. Gruenstein, X. Li, J. Niekrasz, M. Wesson, and S. Kumar, A Multimodal Learning Interface for Sketch, Speak and Point Creation of a Schedule Chart. In *Proceedings of the Sixth International Conference on Multimodal Interfaces (ICMI 2004)*, October 14-15, State College, Pennsylvania, USA, 2004, pp.329-330.
5. J.C. Tang, Findings from observational studies of collaborative work, *International Journal of Man-Machine Studies*, **34**, 143-160 (1991).
6. M. Doyle, and D. Straus, *How To Make Meetings Work* (The Berkley Publishing Group, New York, NY. 1993).
7. D.A. Kravitz, and B. Martin,. Ringelmann rediscovered: The original article, *Journal of Personality and Social Psychology*, **50** (May 1986), 936-941 (1986).
8. J.M. Jackson, and S.G. Harkins, Equity in effort: an explanation of the social loafing effect. *Journal of Personality and Social Psychology*, **49**, 1199-1206 (1985).
9. S.G. Harkins and M. Jackson. The Role of Evaluation in Eliminating Social Loafing, *Personality and Social Psychology Bulletin* **11**(4), 457-465 (1985).
10. T. Bickmore, and R. Picard, Subtle Expressivity by Relational Agents. In *Proceedings of the CHI 2003 Workshop on Subtle Expressivity for Characters and Robots* , April 7th, Fort Lauderdale, FL, 2003.
11. Y. Liu, and Y.S. Chee, Intelligent Pedagogical Agents with Multiparty Interaction Support. In *Proceedings of IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT 2004*, Los Alamitos, CA: IEEE Computer Society, 2004, pp. 134-40.
12. C. Andersen, A Theoretical Framework for Examining Peer Collaboration in Preservice Teacher Education. In *Proceedings of the 2000 Annual International Conference of the Association for the Education of Teachers in Science*, January 6-9, Akron, Ohio, 2000.
13. D. Kuhn, M. Garcia-Milà, A. Zohar, and C. Andersen, Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*, **60**(4), 1995.
14. D. Boud, , R. Keogh, and D. Walker, *Reflection: Turning experience into learning*.(Kogan Page, London, 1988).
15. C. Andersen, *A microgenetic study of science reasoning in social context*. Unpublished doctoral dissertation, Columbia University, MO, (1998).
16. G. Bloom, C. Castagna, and B.Warren, More than mentors: Principal coaching. *Leadership*. May/June (2003).
17. A.L. Costa, and R.J. Garmston,. *Cognitive Coaching: A Foundation for Renaissance Schools*. (Christopher-Gordon Publishers, Norwood, MA, 2002).
18. D. Katz, and R.L. Kahn, *The social psychology of organizations* (2nd ed.). (John Wiley, New York, 1978).
19. K.D. Benne, and P. Sheats, Functional Roles of Group Members, *Journal of Social Issues* **4**, 41-49 (1948).
20. R.F. Bales, *Personality and interpersonal behavior*. (Holt, Rinehart and Winston, New York, 1970).
21. V. Falcon, C. Leonardi, F. Pianesi, and M.Zancanaro, Annotation of Group Behaviour: a Proposal for a Coding Scheme. In *Proceedings of Workshop on Multimodal Multiparty Multimodal Processing (MMMP)* at ICMI, October 7, Trento, Italy, 2005, pp. 39-46.
22. Garrison, B. The perceived of Electronic Mail in Newspaper Newsgathering. In *Proceedings of Communication Technology and Policy Division, Association for Educational Journalism and Mass Communication Midwinter Conference*, March 1, Boulder, Colorado, 2003.