TIMES AND TURNS IN STIMULATING MEETINGS

Ronald Böck

Otto von Guericke University Magdeburg
ronald.boeck@ovgu.de

Abstract: For human beings, it is a common habit to gather with other people in various situations. Regarding especially business meetings, usually a specific goal is intended which should be reflected in the meeting's outcomes. Therefore, it seems to be assumedly easy to assess the success of a meeting, which is often linked to the effectiveness of an interaction. In contrast, Niebuhr et al. argue that meetings can be also of success for the interaction partners if they are perceived as being "stimulating". This aspect is considered in the current manuscript, analysing the Parking Lot Corpus which comprises 70 group discussions with a distinct goal, allowing to apply various performance measures for the assessment of meetings. We focussed on the (perceived) mean meeting effectiveness, relating it to the total speech time and number of turns provided by the participants. Regarding the individual results of either the group leader or the remaining group members, we see a large variation in the meeting effectiveness, being independent from the meeting length. More detailed analyses of these relations support indications towards stimulating meetings in the sense of Niebuhr et al.

1 Introduction

Gathering is a common habit of human beings (cf. e.g. [1, 2]). The purpose and situation where people came together can be either colloquial, like in families or groups of friends, or (quite) formal, such as business meetings. Therefore, the way of interaction as well as the style of discussion vary a lot and also the particular outcome of the gatherings are different by intention. Regarding especially business meetings, the interaction is driven by specific goals, influencing the outcome and success of the meeting (cf. e.g. [3]). In [3] and [4], it is discussed that usually the success of a meeting is linked to the (perceived) effectiveness of the interaction that might be also shared amongst group members. This aspect relates to the common assumption that short meetings should be considered effective (cf. e.g. [5, 6]).

In contrast, Niebuhr et al. [7] argue on the possibility of "stimulating meetings", being perceived as effective in terms of outcomes and the way of interaction, but are not necessarily short in the sense of measured absolute meeting time. Such meetings are considered as interactions where "the communication partners and their interaction are propelled (by each other) such that the entire group is able to perform better" [7]. Especially the gain (even if only perceived or felt) in the performance, which might manifest later in increased group cohesion or global group performance, is the concept's core aspect. Therefore, the idea of Niebuhr et al. covers a "bet" in a future group performance (cf. e.g. [3, 4]), accepting a currently longer lasting meeting.

However, since in [7] the concept of stimulating meetings and mainly prosodic-phonologic analyses are presented, only a preliminary discussion of trends regarding the meeting duration and the mean effectiveness is given. The current paper extended this perspective to analyses of 1) all groups of the corpus (cf. section 2), being suitable for investigations, and to 2) the relation of acoustic samples of the group leader as well as the remaining group members. For

this, we utilised material of the same corpus (cf. Section 2). Given such focus of investigation, the following research questions contribute to how a stimulating meeting can be (generally) assessed:

RQ 1: Do we find group clusters in the meeting effectiveness measure in relation to representations of meeting time (cf. Section 3.1)?

RQ 2: Do we find clusters within the representations of meeting time (cf. Section 3.1)?

RQ 3: How many groups are arranged to the same cluster comparing the results of leader and remaining group members?

RQ 4: Do we find indications that longer lasting meetings are less effective, which contrasts the concept of stimulating meetings (cf. [7])?

In the following, we present the corpus comprising the analysed group interactions (cf. Section 2). In Section 3, the utilised approaches for time representation and clustering are introduced, followed by the presentation and discussion of results (cf. Section 4).

2 The Parking Lot Corpus

Our investigations are based on the Parking Lot Corpus (PLC) [7] comprising 70 meetings recorded at a public Midwestern United States university, where the group size ranges from three to six participants (cf. Figure 1). For this, the psychology department recruited, in total, 245 undergraduate students. For the sake of comparison, the meeting length was controlled by the experimenter, resulting in an overall time of 20 minutes per interaction. We highlight that this is not reflecting the time really used for discussions. How this issue is being handled is explained in Section 3.1.

The corpus' setting relates to discussions aiming for recommendations to improve the university's parking situation. For this, each group was instructed before the interaction and a kind of a meeting agenda was provided, including a list of questions, which might be discussed. Additionally, a group leader was determined by rolling a die. After the meeting the participants filled several questionnaires (cf. [7]). Furthermore, the meetings were assessed by trained annotators and/or self-annotations by the participants, regarding different performance measures. The full list of annotated measures can be found in [7], spanning the range from assessments of meeting effectiveness to participant's satisfaction to number of highly valuable recommendations. For the current investigation, we particularly focussed on the meeting effectiveness (ME) after Leach et al. [5], averaged across all group members. In particular, each participant rated six questions (cf. [5]) related to the achievements of the current meeting using a 5-point Likert scale (1 . . . extremely ineffective to 5 . . . extremely effective). Therefore, the mean meeting effectiveness is also in the range of one to five.

Given this data collection, we currently analysed those groups that provide suitable material[1] for both, acoustic-based and visual-based investigations, allowing multimodal observations on the groups and their respective performances in future work. In this paper, we present only insights on relations of meeting time, being assessed on acoustics, and meeting effectiveness. Since we considered acoustic utterances of either the group leader or the remaining group members, this results therefore in analyses of 68 leaders and 62 "remaining groups".

[1] This means that especially visual investigations are possible, avoiding for instance occlusions.

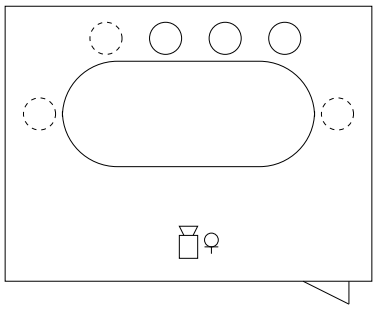


Figure 1 – Schematic visualisation of the setting presenting the relative position of the participants as well as the camera and microphone. Dashed seats represent optional participants in larger groups. The figure is taken from [7].

3 Methods

This section introduces the main concepts for estimating the total "time" used for spoken interations of either the leader or the remaining group members. Further, the cluster approach and the respective parameters are presented.

3.1 Representations of Meeting Time

As mentioned in Section 2, the total length of the meetings was controlled by the experimenter, thus resulting lengths around 20 minutes per session. Therefore, the total duration of interactions could not be used for analysing the relations of meeting time and meeting effectiveness. We decided for using auxiliary measures for the communication or interaction time.

Total Time of Acoustics: The first measure is based on the participants acoustics, including verbal and non-verbal expressions. For this, we separated the acoustic samples of the leader and the remaining group members and calculated the accumulated time of all samples. This results in an (estimated) total time of spoken interaction. It is to be noticed that also the sample times of all remaining group members were cumulated, neglecting the circumstance of different group sizes (cf. Section 2). This approach 1) reflected the distinguished position of the leader against the other group members and 2) accommodated the aspect that the meeting effectiveness is considered as an average across all participants of the interaction. Further, it allowed a direct comparison of interaction parts of the leader and the group.
Regarding the way we estimated the total time or interaction time, this limited the investigations in the sense that meetings without any interaction would mapped to short meetings. Fortunately, such interactions were not found in the PLC (cf. Figures 2(a) and 3(a)) and further, this aspect can be considered as unlikely for our investigations. However, to prevent any biased interpretations, we also considered the number of turns per participant.

Number of Turns: The number of turns indicated the (individual) contribution to the interaction, which levels also the total time of acoustics. Again, we analysed separately the group leader and the remaining group members; their turns were accumulated. Possible interpretations are: Few turns per interaction may reflect a rather tedious interaction; in contrast, many turns represent a rather lively discussion.

3.2 Clustering Approach

The clustering is based on the measures introduced in Section 3.1, representing the (estimated) meeting time. For this, the acoustic statements of the groups were manually segmented and allocated to either the group leader or the remaining group members. We obtained individual assessments of each sample in relation to length, which would allow further prosodic and phonologic investigations[2] in relation to [7]. In the current study, the measures' values were obtained as explained in Section 3.1.

Regarding the meeting effectiveness, this performance measure is obtained by reports of the participants using the questionnaire of Leach et al. [5]. Since the particular meeting is observed from different participant but we are mainly interested in the overall group performance, the average meeting effectiveness is calculated (cf. Section 2).

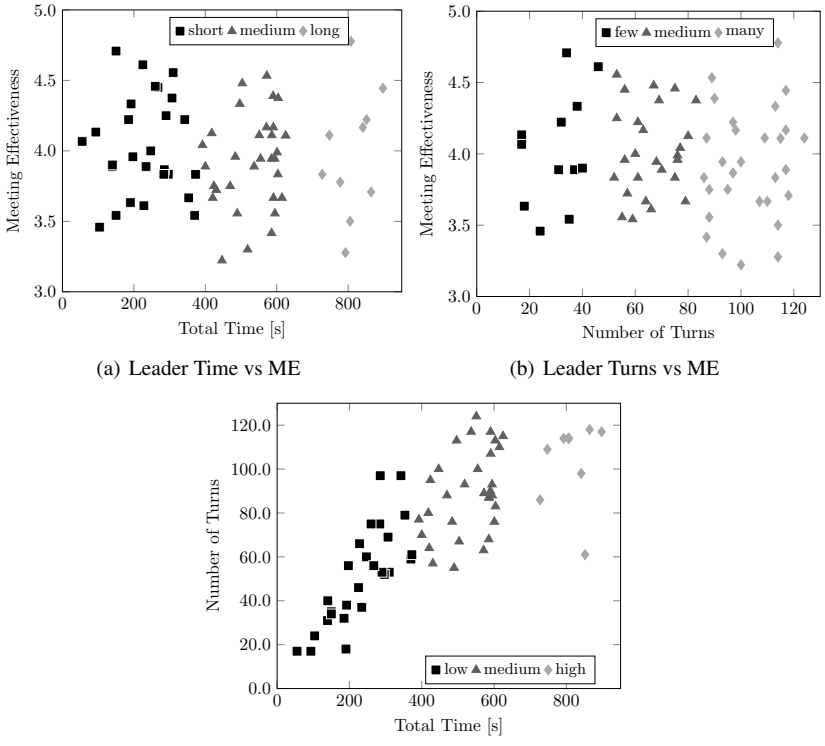
For clustering, we applied a k-means approach, distinguishing the meetings. For this, each group meeting is represented by the total time and the number of turns, separated per leaders and remaining group members, respectively. These values were investigated against the meeting effectiveness (called *Time vs ME* and *Turns vs ME*) as well as against each other (called *Time vs Turns*). As analysing tool, we used the R software (cf. [8]), applying the internal packages, 1) to analyse the optimal number of clusters and 2) to obtain the respective clustering. To estimate the optimal number of clusters the function `fviz_nbclust` (cf. [9]) was utilised, applying the "within sum of square" method. Identifying the (optimal) number of clusters as $k = 3$, we calculated the particular clustering, respectively. The results are visualised in the scatter plots in Figures 2 and 3. To ensure readability, we used both, colours and markers, depicting the same content as particularly explained in the respective legends.

4 Results and Discussion

In the following, we present and discuss the obtained cluster results in relation to the research questions stated in Section 1.

Cluster Results: In the current study, we investigated the relation of (estimated) interaction times and the mean effectiveness of the meeting (cf. research question RQ 1) as well the representations of interaction times against each other (cf. research question RQ 2). The time is represented by the aggregated time of samples in seconds as well as the number of turns provided by the respective participants (cf. Section 3.1). Using the k-means approach, we obtained the results visualised in the scatter plots in Figures 2 and 3, where each data point represents a particular group meeting. In the clustering, we distinguished the leader of the group (cf. Figure 2) and the remaining group members (cf. Figure 3), observing the influence on the effectiveness separately.
Regarding the scatter plots related to meeting effectiveness, we saw that interestingly not the meeting effectiveness is an indicator for distinguishing the meetings, but rather the total time and the number of turns. In this sense, the common argumentation (cf. e.g. [2, 4, 7]) seems to be proven. Regarding the number of turns in the leaders (cf. Figure 3(b)), also a trend towards the common interpretation of meeting effectiveness can be seen. However, in contrast, analysing the ME values even in long meetings, we observed highly efficient meetings. Further, in the *short* and *few* clusters in Figure 2(a)/(b) and Figure 3(a)/(b), respectively, the meeting effectiveness also varied a lot. So, there is no clear indication that short meetings are necessarily effective or longer meetings are considered as ineffective (cf. research question RQ 4). This is in line with the discussion in [7] that longer lasting meetings might be of success and,

[2] This is not matter of the current manuscript.



(a) Leader Time vs ME

(b) Leader Turns vs ME

(c) Leader Time vs Turns

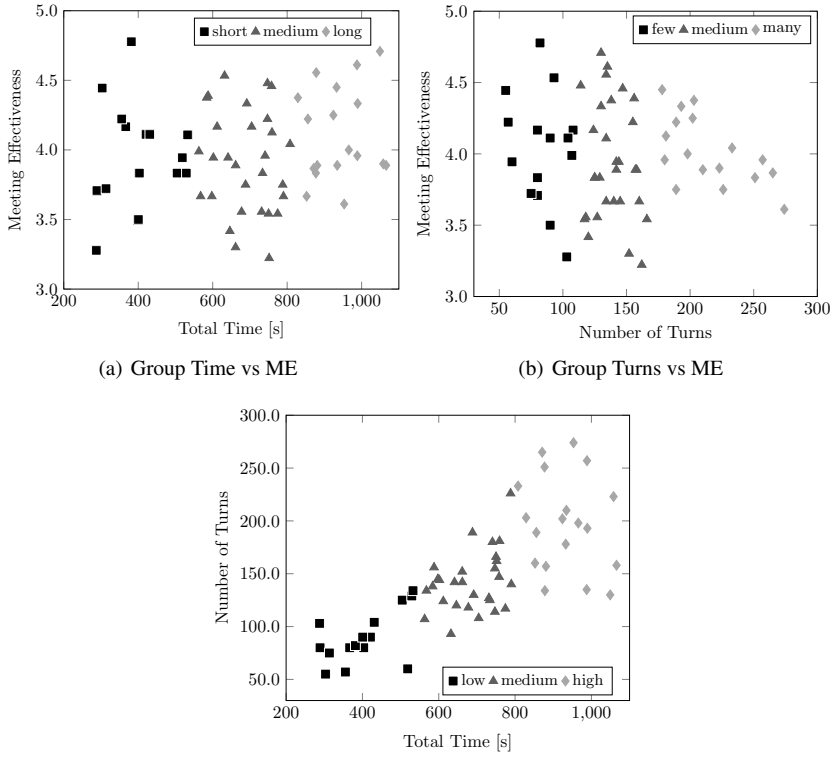
Figure 2 – Scatter plots visualising the clustering results for the leader investigations. The combination of marker shapes and colours is used to increase the readability.

especially, can be "stimulating".

In research question RQ 2, we asked for possible relations between total time and number of turns. This is linked to the idea: A high number of turns (perhaps a lively discussion) might not directly result in long meetings. Given the observations from the respective clusters in Figures 2(c) and 3(c), this hypothesis could not be confirmed. In fact, especially in the remaining group members clustering (cf. Figure 3(c)), we see that short meetings relate to a low number of turns. This is an interesting aspect which leads to either an immediate agreement on proposals or a tight leadership style. However, the other perspective is related to the aspect that the group member do not know each other beforehand. Therefore, it can be assumed that the longer meetings with larger numbers of turns are used for familiarisation (cf. also paragraph Indication for Stimulating Meetings). Both aspects are matter of further investigations and need collaborative work with social scientists.

Shared Clusters across Measures: To answer research question RQ 3, we analysed the obtained clusters and compared them regarding the corresponding groups.

Regarding the confusion matrices in Table 1, we saw that in most combinations similar clusters were shared across the measures. Especially considering the remaining group clusters, a more focussed grouping could be achieved, for instance, suggesting that short meetings (in relation to meeting effectiveness) might also be clustered to few turns (in relation to meeting effectiveness) as visualised in Table 1(b). In contrast, for the leaders, the collection is either clearly focussed (cf. Table 1(b)) or rather losely linked (cf. Tables 1(a) and (c)). From these results, we argue



(a) Group Time vs ME

(b) Group Turns vs ME

(c) Group Time vs Turns

Figure 3 – Scatter plots visualising the clustering results for the investigations observing the remaining group members. The combination of marker shapes and colours is used to increase the readability.

that for a group of participants the meeting effectiveness and the representation of time could be estimated easier than for individual persons like the leader. For them, the individual perception of what an effective meeting is and means may vary easily. However, given the PLC, the individual differences could not be assessed from the data. Therefore, the findings indicate only possible trends and show options for further, more detailed investigations in this sense.
Additionally, we also analysed the assignments of clusters across "experiments", which means: Do we see same assignments in the leader clustering and the remaining group members clustering? Regarding the three possible clusters, the least accordance can be seen for *Turns vs ME* where only two groups share the *medium* cluster and one group the *many* cluster. For the other two clusters only the *medium* one is shared across the groups, respectively; in each case, seven groups could be identified. Again, this provides some evidence to the individual perception of the meeting effectiveness. Nevertheless, more social related analyses need to be conducted to explain the current findings, considering social, communicational, and psychological aspects.

Indication for Stimulating Meetings: Research question RQ 4 emphasises the differences between short, concentrated meetings and stimulating meetings which may last longer but being nevertheless perceived as effective. As already stated in paragraph Cluster Results, the meeting effectiveness is decoupled from the particular length of the meeting (cf. Figure 2(a) and Figure 3(a)) as well as from the number of turns (cf. Figure 2(b) and Figure 3(b)). For the clusters focussing on the group leader, the variation of meeting effectiveness within each cluster is more pronounced compared to the remaining group members. Given these results, we could not find

Table 1 – Distribution of clustering combinations in form of confusion matrices. The tables indicate how many groups show similar clusters across the measures (individual clusters can be found in Figures 2 and 3). Leader represents the clustering obtained regarding only the leader, where Group indicates the investigations on the remaining group members.

(a) Leader Time vs ME & Turns vs ME

Time/Turns	few	medium	many
short	12	13	2
medium	0	12	19
long	0	1	9

(b) Group Time vs ME & Turns vs ME

Time/Turns	few	medium	many
short	12	3	0
medium	3	21	5
long	0	6	12

(c) Leader Time vs ME & Time vs Turns

Time/TimeTurns	low	medium	high
short	27	0	0
medium	0	31	0
long	0	0	10

(d) Group Time vs ME & Time vs Turns

Time/TimeTurns	low	medium	high
short	15	0	0
medium	0	28	1
long	0	0	18

(e) Leader Turns vs ME & Time vs Turns

Turns/TimeTurns	low	medium	high
few	12	0	0
medium	13	12	1
many	2	19	9

(f) Group Turns vs ME & Time vs Turns

Turns/TimeTurns	low	medium	high
few	12	3	0
medium	3	21	6
many	0	4	13

clear evidence for the aspect that longer meeting are perceived comparably less effective. At first, this contrasts the common interpretation (cf. e.g. [2, 4]), given the PLC data. In addition, these findings support to some extend the argumentation of [7] towards stimulating meetings. In the PLC's setting the group members usually do not know each other before the meeting. Therefore, the stimulating aspects is also linked to issues of creating a kind of group conversion or group coherence (cf. [10]). This is related to the mentioned "bet", where the currently longer meeting is used for familiarisation, allowing (perhaps) afterwards short meetings of established groups. Although we see some evidence for stimulating meetings, the current study is also limited in a way that especially the corpus lacks control groups definitely knowing each other, which should be investigated in future research.

5 Conclusion

Based on the PLC, we analysed aspects of total times and number of turns in group meetings in relation to the mean meeting effectiveness, annotated and perceived by the participants. We specifically focussed on aspects being related to "stimulating meetings" as discussed in [7]. Clustering the group meetings in relation to the material obtained by either the group leader or the remaining group members showed that even in longer lasting meetings the meeting effectiveness is not necessarily decreasing (cf. Figures 2 and 3). This somehow contradicted common expectations from the literature (cf. e.g. [2, 3, 4]). In fact, it rather strengthens the idea that meetings can be stimulating in the sense that the outcome is considered within the group to be productive. However, we could not find similar indications regarding relations of total times and number of turns. Further, we analysed the number of clusters being shared across clusters within each setting (i.e. only leader or only remaining group members). The respective confusion matrices are presented in Table 1. Additionally, the number of groups sharing same clusters across settings were also assessed. In fact, the results are sobering, just showing a weak tendency in terms of groups being clustered as *medium* (cf. paragraph Shared Clusters across Measures in Section 4). Finally, we discussed our results in relation to an understanding of stim-

ulating meetings. Especially in this sense, more research focussing on individual perceptions of effectiveness are necessary.

Therefore, in future research, prosodic characteristics of both, the leader and the remaining group participants, as well as the speech charisma of the interacting partners should be analysed and related to the current findings. Further, the individual perception on the meeting, its outcomes, and the interaction should be evaluated, considering also social aspects. Therefore, we encourage researchers to use self-ratings of the group members to assess the perceived level of effectiveness but also the level of being stimulated by the interaction. Both measures need to be validated in settings where the group member either know each other or are strangers.

Acknowledgements

We acknowledge support by "Adaptive Strategies in Assistive Technologies in Multi-Person Interaction" (ASAMI) funded by the Federal State of Sachsen-Anhalt, Germany, under the grant number: I 138.

References

[1] MCGRATH, J.: *Groups: Interaction and Performance*. Prentice-Hall, Englewood Cliffs, USA, 1984.

[2] BROWN, R.: *Group Processes – Dynamics within and between Groups*. Blackwell Publishers, Oxford, UK, 2000.

[3] KOZLOWSKI, S. W., S. M. GULLY, E. R. NASON, and E. M. SMITH: *Developing adaptive teams: A theory of compilation and performance across levels and time*, pp. 240–292. Wiley, Hoboken, USA, 1999.

[4] YOERGER, M., J. A. ALLEN, and J. CROWE: *The impact of premeeting talk on group performance*. *Small Group Research*, 49(2), pp. 226–258, 2018.

[5] LEACH, D. J., S. G. ROGELBERG, P. A. WARR, and J. L. BURNFIELD: *Perceived meeting effectiveness: The role of design characteristics*. *Journal of Business and Psychology*, 24(1), pp. 65–76, 2009.

[6] SALAS, E., S. I. TANNENBAUM, S. W. J. KOZLOWSKI, C. A. MILLER, J. E. MATHIEU, and W. B. VESSEY: *Teams in space exploration: A new frontier for the science of team effectiveness*. *Current Directions in Psychological Science*, 24(3), pp. 200–207, 2015.

[7] NIEBUHR, O., R. BÖCK, and J. A. ALLEN: *On the sound of successful meetings: How speech prosody predicts meeting performance*. In *Proceedings of the 23rd ACM International Conference on Multimodal Interaction*, p. s.p. ACM, Montreal, Canada, 2021.

[8] R CORE TEAM: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL `https://www.R-project.org/`.

[9] KASSAMBARA, A. and F. MUNDT: *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*, 2020. URL `https://CRAN.R-project.org/package=factoextra`. R package version 1.0.7.

[10] TSKHAY, K. O. and N. O. RULE: *Accuracy in categorizing perceptually ambiguous groups: A review and meta-analysis*. *Personality and Social Psychology Review*, 17(1), pp. 72–86, 2013.