Effect of Nonverbal Behavioral Patterns on the Performance of Small Groups

Umut Avci
Idiap Research Institute
Martigny, Switzerland
umut.avci@idiap.ch

Oya Aran
Idiap Research Institute
Martigny, Switzerland
oya.aran@idiap.ch

ABSTRACT

In this paper, we investigate the links between the interaction among the group members and team performance, as well as the group composition. Towards this goal, we extract nonverbal audio-visual behavioral features from the group interaction and apply a multivariate binary Influence Model to quantify interactions between team members. We compare and contrast the differences between forms of interaction for distinct performance clusters and describe the factors affecting group performance in terms of group members' influences on each other and traits such as dominance. Our results show that different performance clusters have different interaction types and the groups with high and low performance have a structure where the group members are influenced by one or more people, in contrast to one-to-one pairwise influences. The performance of individuals with specific traits, such as leaders and competent people, has a significant correlation with the group performance, with a correlation coefficient of over 0.50.

Categories and Subject Descriptors

H.1.2 [**Information Systems**]: User/Machine Systems — *Human information processing*

Keywords

Small groups; Nonverbal behavior; Performance analysis

1. INTRODUCTION

The ability to work as part of a team is a crucial selection criterion that applicants encounter in almost all job applications today. With such a principle, companies aim at decreasing their expenditures on employee training that teaches how to work as a group and how to hold efficient meetings. Running effective meetings not only prevents companies from spending valuable time and money in vain but also increases productivity. Hence, developing approaches to improve performance of a team is of great importance for organizations.

Social psychologists have long been interested in identifying key factors that affect the efficiency of groups. The importance of leadership and being led by a right person was discussed in [11]. Authors pointed out that a knowledgeable and assertive leader is able to ensure a meeting to continue on its track, keeps members involved in discussions, and uncovers the participants' strengths. To Sundstrom, cohesion is a highly valuable skill that improves team effectiveness [15]. A cohesive team has the ability to resolve conflicts before it damages the relationship between group members, which is achieved by having a team spirit and supporting each other. In [14], Shaw stressed that heterogeneous teams that are composed of participants with variety of knowledge, skills, and abilities could improve the decision performance. Later studies showed that constructing groups based only on the diversity of people was not always successful at enhancing efficiency. Bringing together like-minded individuals and using heterogeneous teams solely for complex tasks were proposed in [12] and [8] respectively. In [2], Bradley defined four main personality types with their subcategories and analyzed the effect of trait compositions on the performance. The relation between the members' engagement in discussion and the team performance was investigated by Haslam [7] and by van Dick [16]. It was found that the group makes better decisions when they work collectively (in the former) and when members are actively interacting, i.e. less social loafing (in the latter). Woolley showed in one of her findings that the collective intelligence of a group depends on factors other than the individual intelligence of its members, suggesting that the group performance may be lower or higher than the performance of its members [17].

Contrary to the approaches in psychology, there are only a few studies focusing on performance analysis of groups in social computing. In [10], Lepri developed an SVM-based technique to classify individual performances based on nonverbal behavioral features in a group performing Mission Survival Task. Dong provided a model on the relationship between the discussion dynamics and the group performance based on a quantitative analysis of a brainstorming task in [3]. It was found that the group performance is proportional to the number of simultaneous speakers, i.e. generated ideas. In later studies, Dong used mixture of Hidden Markov Processes and Markov jump processes to evaluate the effect of nonverbal audio features on the group performance. He observed that better performances were achieved for faster speaker changes, longer clause lengths and less standard

UM3I'14, November 16, 2014, Istanbul, Turkey.

ACM 978-1-4503-0652-2/14/11 ...$15.00.
http://dx.doi.org/10.1145/2666242.2666246.

deviations of pause [6], and for more speaking turns, and balanced participation rate [5]. Latent Dirichlet allocation (LDA) was used in [9] to extract the group level speaking and looking patterns. Jayagopi built relationships between the extracted patterns and meeting constructs such as group composition and performance. It was found that the group performance correlates with the unsuccessful interruptions skew and convergent gaze.

In this work, we address the problem of defining factors affecting performance of groups. To this end, we create clusters for distinct performance levels. We then learn interactions of group members from their nonverbal features by using a generative probabilistic approach and analyze performance differences between clusters in terms of interaction dynamics and personality traits. Our paper has several contributions. First, we provided explanations for the reasons behind the difference in group performances in terms of the influences of group members over each other and the dominance of members. To model the influences, we used the Influence Model [1] and presented a multivariate binary version of IM as another contribution of this study. Finally, we revealed the relationship of individuals having particular traits with both the performance of a group and the influential people.

2. OUR APPROACH

A group committed to a common purpose works together to come to an optimal decision by taking into account the opinions of each member in the group. Typically, final outcome determined by the group differs from the decisions made by individuals. Quality of the final say, hence the performance of the group, depends not only on the individuals' strengths such as knowledge and experience but also on the structure of the group, e.g. communication intensity between members and their roles. In order to analyze the effect of the latter on the group performance we followed a two-step approach. Groups were first clustered based on their performances. Then, interactions between members were extracted for each group by using a temporal probabilistic approach. We analyzed the differences between forms of interaction for distinct performance clusters and identified the impact of personal roles on decision-making process. We evaluated our approach by using a data set, collected to identify the emergence of leadership in small groups. Details on the data set is presented in Section 3.

2.1 Finding Performance Clusters

The purpose of the clustering step is to determine the intrinsic grouping of teams based on performance so that intra-cluster consistencies and inter-cluster differences can be analyzed. To this aim, we employed Kernel Density Estimation (KDE) where the main goal is to find the dense regions of data points, i.e. performance values of groups. Being a non-parametric method, it does not rely on a fixed probability model of the clusters as in K-means clustering. Instead, it approximates a probability density function at each data point in the data set [18]. Hence, local minima in the final density function become strong candidates to split the data into clusters. For one-dimensional performance values, KDE is given as $f(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i)$ where n is the number of groups, x_i is the performance of group i, K_h is a kernel function controlling the weight given to the observations x_i at each point x, and h is a smoothing parameter determining the size of the neighborhood around x. In our case, bandwidth was chosen automatically and Gaussian kernel was implemented.

2.2 Modeling Influnces

We modeled influences by using the Influence Model [1]. It is a generative approach that enables modeling interactions between multiple agents. The approach inherits the structural and graphical architecture of a coupled Hidden Markov Model (HMM). In other words, it is a set of multiple HMM chains coupled through cross-time and cross-chain conditional probabilities. Influence model prunes the coupled HMM by drastically reducing the number of parameters to be learned, allowing to include higher number of chains without incurring substantial increase in computational complexity. It is achieved by a parametrization strategy in terms of the "influence" each chain has on other chains. More specifically, the conditional probability $P(S_t^i|S_{t-1}^1, \ldots, S_{t-1}^C)$ is simplified only by keeping the transition probability $P(S_t^i|S_{t-1}^j)$, where t is the time stamp and C is the number of chains. Thus, the full conditional distribution is approximated as follows:

$$P(S_t^i|S_{t-1}^1, \ldots, S_{t-1}^C) = \sum_j \tau_{ij} P(S_t^i|S_{t-1}^j).$$

Here, τ's are constant factors referred as "influences" that quantify how much the state transitions of a given chain depend on a given neighbor. Since the model is not a fully-connected graph as a generalized coupled HMM, it only captures interactions between pairs of chains. Even if the joint effect of multiple chains is absent, pairwise interactions meet our requirements.

In its current version, the Influence Model is able to process multinomial data. We introduced an improvement also to consider multivariate binary case. Parameter learning in the influence model is done by maximum likelihood estimation after applying forward-backward algorithm for the latent state inference. The forward and backward variables are calculated based on the observation probability $P(x_t^{(c)}|s)$; where $x_t^{(c)}$ is the observation for chain c at time t (see [4]). For working with multivariate binary data, we made a Naive assumption of independence between observation features and computed the observation probability as follows:

$$P(x_t^{(c)}|s) = \prod_{i=1}^{N} P(x_{t,i}^{(c)}|s)$$

In the end, the probability of seeing N observations together is the product of their probabilities. Parameters are then estimated by using formulations based on forward and backward variables. In our scenario, the Influence Model was run for every group in the data set regardless the cluster. Number of chains in the model was made equal to the number of group members accordingly and the number of iterations for the parameter learning was set to 100.

3. DATA SET AND FEATURES

3.1 Data Set

The Emergent LEAder corpus (ELEA) is designed to collect information about small group performance, decision making, and leadership [13]. It is composed of 40 meetings, 28 of which are four-person and the rest is three-person. Each meeting took around 15 minutes and participants were asked to perform winter survival task during this time.

For the sake of being less intrusive, the data were recorded with an infrastructure that allows freedom of movement. A Microcone, a commercial microphone array, was used to capture audio signal. The device is capable of segmenting speakers automatically and provides audio for prosodic cue extraction. For video recordings, two webcameras were used. Video recordings are only available for 27 meetings.

Winter survival task is a fictional scenario in which the aim is to prioritize 12 items in rank order to survive an airplane crash in winter. The ranking is performed individually at the beginning of the task and as a team after the group discussion to account for the interactions among the participants and to observe the most influential people for factors as dominance and leadership. These rankings are used to compute the Absolute Individual Scores (AIS) and Absolute Group Scores (AGS) based on the absolute difference with respect to the rankings of survival experts.

Participants also filled questionnaires about themselves and their perception on group members before and after the meeting respectively. A five-point scale was used to evaluate the first 16 statements, in which four variables were considered. Perceived Leadership (PLead) characterizes a person who directs the group, imposes her/his opinion, and is involved. Perceived Dominance (PDom) is associated with a person who dominates, is in a position of power, and asserts him- or herself. A person who is competent, intelligent, and has a lot of experience links to Perceived Competence (PComp). The one who is kind, friendly, and sympathetic matches with Perceived Liking (PLike). In the last statement, participants were asked to rank group members based on their dominance to find Ranking of Dominance (RDom). Each participant assigns 1 to the person who is perceived as the most dominant, and 3 or 4 (based on the number of group members) to the one who is perceived as the least dominant. As a result, for each participant, there are 3 or 4 questionnaire outputs reflecting the participants' perception of other group members. Each perceived variable is averaged over all participants per group, and the group ranking is normalized according to the group size. We used averaged and normalized variables in our experiments.

3.2 Nonverbal Features

Audio nonverbal features are obtained by processing the output of the Microcone's automatic speaker segmentation. This information is used to create a binary segmentation for each participant, where status 1 represents speech and status 0 represents non-speech. We processed the segmentation to extract the following audio features for each participant:

Speaking Status ($\mathbf{SS}_i$): The binary speech segmentation for participant i.

Unsuccessful Interruptions-Audio ($\mathbf{UIA}_i$): A binary feature where status 1 represents time instants with Unsuccessful Interruptions and status 0 otherwise, given that *Participant i interrupts participant j if i starts talking when j is speaking, and i finishes her/his turn before j does.*

Visual nonverbal features are composed of two parts: head activity and body activity. A binary segmentation is created for each motion type, where status 1 represents movement of head (body) and status 0 represents stillness (Please refer to [13] for details). By using the segmentation, the following visual features are extracted for each participant.

Head Motion Status ($\mathbf{HMS}_i$): The binary segmentation for head motion for participant i.

Body Motion Status ($\mathbf{BMS}_i$): The binary segmentation for body motion for participant i.

Unsuccessful Interruptions-Head ($\mathbf{UIH}_i$) and Unsuccessful Interruptions-Body ($\mathbf{UIB}_i$): The binary features where status 1 represents time instants with Unsuccessful Interruptions and status 0 otherwise, given that *Participant i interrupts participant j if i starts acting when j is moving, and i finishes her/his turn before j does*, respectively for head and body motion.

A discussion about the selected features and their usage is given in detail in Section 4.

Table 1: Clustering results based on group performances

	Low performance	Average performance	High performance
Cluster cardinality	20	15	5
Performance Mean	54.3	42.93	27.8
Performance Std.Dev.	3.06	3.85	5.76

4. EXPERIMENTS AND RESULTS

In this section, we first shed light on the diversity in group performances via clustering and influence modeling. Then, we provide a discussion on the relation between the individual performances in a group and the performance of the group. Finally, we present an analysis for the association of individuals having specific traits with the group performance and the influential people.

To represent the group performance, we used the AGS values (see Section 3). Since group performances show variability, different performance levels need to be compared to explore the underlying reasons of change. Clustering step enables us to create sets of groups with distinct performance ranges. Kernel Density Estimation takes as an input the Absolute Group Scores of groups and automatically generates three sets corresponding to low, average, and high performance. It is seen from Table 1 that groups are unevenly distributed to the clusters, i.e. there are only 5 groups in the high performance level while there are 20 in the low one. The mean and standard deviation of groups' performances in clusters are also depicted in the table. Note that, the lower the score the higher the performance.

4.1 Analysis of Internal Dynamics

We analyzed the internal dynamics of groups in terms of the influence of each person in a group on the other group members. For this purpose, we run the Influence Model for each group in our data set and obtained influence matrices that allow quantitative interpretation of the interaction between group members. In the Influence Model, each person corresponds to a chain with two states for each modality: *Silent* and *Speaking* for the audio, *Still* and *Moving* for the video. Observation features of a person for the audio include a binary speaking vector of that person and 2 or 3 (depending on the group cardinality) binary interruption vectors, which are the unsuccessful audio interruptions of other group members on the chain-owner. Observation features of a person for the video are composed of two binary motion vectors, i.e. head and body, of that person and binary unsuccessful visual interruption vectors as it is defined in the previous case. We believe that unsuccessful interruption is a proper feature indicating one's determination on the discussion. In the case of audio, it can be considered as an evidence for the robustness of speaker's argument and for the speaker's ability to defend her/his ideas. On the other hand, for the video, it is an indicator of co-movement such as nodding or tilting of the body to the side that signals agreement or objection. In either modality, the speaker continues her/his speech and motion without being disturbed by the interruption, which is a sign of higher influence.

An influence matrix has a dimension of $D \times D$, where D is calculated as the product of the group cardinality and the number of states. Each cell in the matrix corresponds to influence of a person's state (in the rows) on another person's state (in the columns). The influence values range from 0 (blue) to 1 (red). Our analysis of influence matrices for different clusters shows that there is a relation between the forms of interaction and the performance level. We observed that, for the low performance cluster, a person has influence only over another person. Many people have influence over each other for the average performance cluster. And for the high performance cluster, a person has influence over multiple people. This can be seen in Figure 1, which shows sample influence matrices of three groups, using only audio features, each coming from a distinct performance cluster. The first one, Figure 1a, belongs to a group in the low performance cluster. Here, Person 1 has influence on Person 2 such that when Person 1 is *Silent*, Person 2 tends to be in *Speaking* state and when Person 1 is in *Speaking* state, Person 2 tends to be *Silent*. Note that, Person 1 affects no one except Person 2. The same pattern can be observed for the other group members as well. The influence matrix for a group in the average performance cluster, Figure 1b, shows that every group member except Person 1 has influence on each other. Person 2 has influence on Person 1, Person 3 has influence on Person 2 and 4, Person 4 has influence on Person 1, 2, and 3. In this case, Person 1 and Person 2 are affected by multiple group members. The last one, Figure 1c, depicts influences for a group in the high performance cluster. Person 1 has an effect on Person 3 and Person 3 has influence on both Person 1 and Person 2. What makes it different from the average performance cluster is that neither Person 1 nor Person 2 is affected by any other member than Person 3, which makes her/him the most influencing member in the group. Note that, groups in the figure were selected manually. We observed that these samples characterize the clusters, i.e. majority of the influence matrices in the clusters are similar to those presented. A detailed analysis using the video modality and the multimodal case is shown in the next section.

4.2 Analysis of Individual vs. Group Performances

In this part, the relation between the members' individual performances in a group and the performance of the group will be analyzed in terms of the dominance and the influences of group members.

Due to the page limitations, we will not go through the whole ELEA corpus. Instead, a representative group from each performance level was chosen. The selection was made based on Root Mean Squared Error (RMSE), a measure that quantifies the sample standard deviation of the differences between the group members' performances and the group performance. RMSE was calculated for each group in a cluster and the one with the highest RMSE score was selected as the group characterizing the cluster. Such criterion gives us groups having high variability in individual performances, that intuitively allows us to make inferences on the connection between individual and group performances. RMSE is calculates as $\mathrm{RMSE}_i^c = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y-y_j)^2}$ where RMSE_i^c is the Root Mean Squared Error of group i in cluster c, n is the number of members in group i, y is the performance of group i, and y_j is the individual performance of the jth member in group i.

Table 2: Statistics of sample groups from each performance cluster

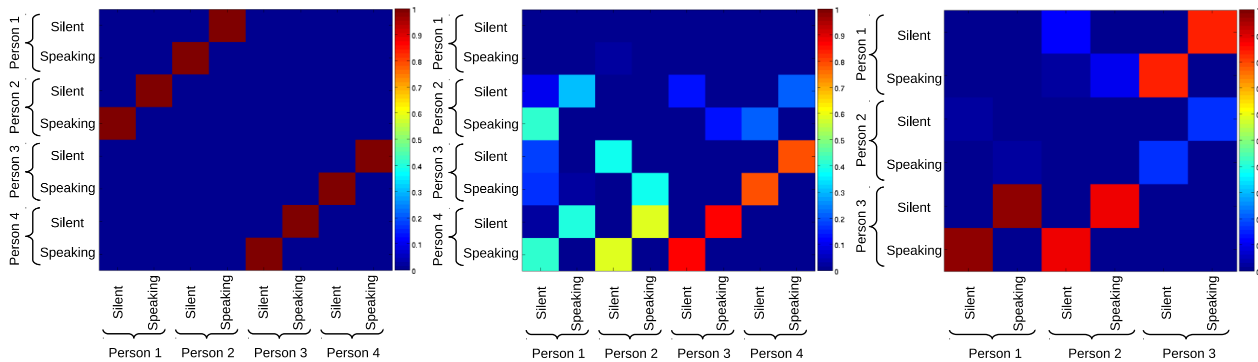
	Low Group 40	Average Group 29	High Group 5
RMSE	19.44	10.30	28.96
Mean RMSE	8.27	9.34	20.67
Group cardinality	4	4	3
Individual performances	30,56,40,38	48,48,50,54	58,30,56
Group performance	58	40	22
Dominant member	2nd	4th	3rd

Table 2 presents the results of RMSE selection process. Group 40, 29, and 5 are the selected teams representing the groups in low, average, and high performance clusters respectively. RMSE and Mean RMSE correspond to a score of a single group and to an average score of groups in clusters. As can be seen, individual performances are not uniform within the groups. The table also shows the rank of the most dominant person in each group. The fact that a dominant person is in a position of power and has ability to state her/his ideas confidently channeled us to consider dominance as an evidence of influence. That's why, we investigate the effect of the most dominant person in determining the performance of the group. We used the RDom variable from the questionnaires to find the most dominant person in each group.

For multimodal analysis, we calculated separate influence matrices for the audio and visual modalities and took their average to characterize the multimodal influence matrix. Note that the video recordings in the ELEA corpus are available only for a subset of meetings and that multimodal analysis was performed only for meetings for which the video recordings were available.

Figure 2 depicts the influence matrix of Group 5 in the high performance cluster (see Table 2). Since the video recording is not available for this team, only the audio modality is shown. Person 1 has no influence over other group members while Person 2 affects all. Person 1 and 3 tend to listen Person 2 while (s)he is speaking, and tend to speak when (s)he is silent. Person 2 is only influenced by Person 3. From Table 2, we see that decisions made by group members are worse than the one made as a group. More importantly, the performance of the most dominant person in the group, i.e. third one with the performance of 56, is much lower than the group performance of 22. One would expect that (s)he worsen the performance of the group as s(he) is dominant with low performance. However, being the most influential group member, Person 2 increases the group performance as (s)he has the highest performance of all, i.e. 30.

A similar situation is observed for Group 29 in Table 2, corresponding to the average performance cluster. Individual performances fall behind the group performance and the performance of the most dominant member is the worst of all. The multimodal influence matrix indicates two important interactions: one for Person 2 and another for Person



(a) Low performance cluster (b) Average performance cluster (c) High performance cluster

Figure 1: Influence Matrices for performance clusters using audio features

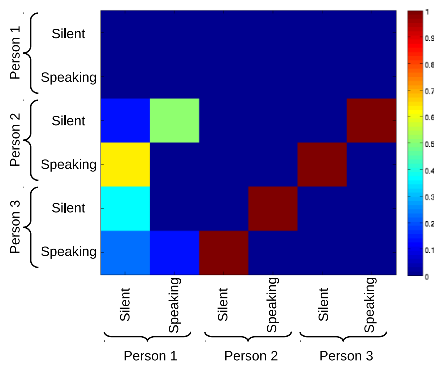
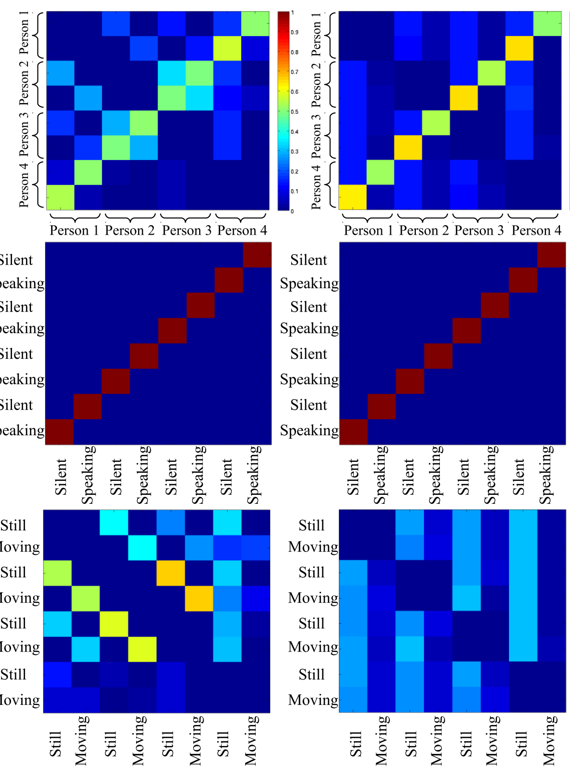


Figure 2: Influence matrix of a group in the high performance cluster

3 (see Figure 3a). They both have effects on multiple people. To investigate their impact more clearly, the audio and visual modalities need to be examined separately. The influence matrix for the audio shows a one-to-one structure, meaning there is a pairwise interaction between members, e.g. Person 4 listens Person 1, Person 3 listens Person 2 and so on. Hence, influences of members to each other are equal for the audio case. When we analyze the influence matrix for the video, we see that the first and the third member tend to move when Person 2 is moving. The influence of Person 3 on Person 1 is smaller when compared with the two significant influences of Person 2. Evaluation of the visual modality reveals that Person 2 is more influential than Person 3. By being the most influential member with relatively high performance, Person 2 boosts the group performance.

Contrary to the Group 29 and 5, all individual performances in Group 40 are better than the group performance. Here, the lowest individual performance belongs to the most dominant person and other members have significantly higher performances. Therefore, we do not expect to see any influential member in the group but Person 2, who is the most dominant one. Indeed, when we look at the influence matrices for Group 40 (see Figure 3b), we do not observe an influential person. A one-to-one interaction is noted for the audio case. For the video case, there is no interaction between the members because they have a tendency to not to move regardless the state of people. Since there is not an influential person in the group to rise the performance, the performance of the most dominant person prevails, causing a drop in the group performance. Semantic analysis of interactions was omitted as we focused on using nonverbal behavioral cues.



(a) (b)

Figure 3: Influence Matrices: Average Performance (a) and Low Performance (b) clusters for Multimodal (top), Audio (middle), and Video (bottom)

Using complementary modalities, i.e. audio and visual, allows us to assess the interactions between group members more accurately. For cases in which one of the modalities provides information and the other one does not, relying only on a single modality may lead to misinterpretation of results. Our example in Figure 3 is the epitome of such a situation. If we had analyzed the influence matrices by only looking at the audio modality, we would have missed the interactions included in the visual modality.

4.3 Group Composition and Performance

In this section, we investigate how individuals with specific traits (such as leaders or dominant people) affect the overall group performance. For this purpose, we used the PLead, PDom, PComp, PLike, and RDom variables that

Table 3: Pearson Correlation Values between individual performances for each perceived variable and group performances. Significance values $*: p < 0.01$

	PLead	**PDom**	**PComp**	**PLike**	**RDom**
Correlation	0.60*	0.32*	0.53*	0.20	0.46*

Table 4: Influential People ratio for Perceived Variables

	PLead	**PDom**	**PComp**	**PLike**	**RDom**
Ratio	0.45	0.45	0.33	0.20	0.45

were obtained from the questionnaires. For each variable, we selected the person in each group that has the highest score for that variable and calculated the correlation between their individual performances (AIS) and the group performances (AGS). We computed the Pearson correlations at 1% significance level (i.e., $p < 0.01$). Table 3 shows the Pearson correlation values. The value for RDom, for instance, is the correlation between the individual performance of the most dominant people in the groups and the performance of the groups. These results suggest that a moderate correlation exists for perceived variables except PLike. Highest correlation is with PLead, with a correlation of 0.60.

We also looked for the correspondence between the influential people and people with specific traits. For each group, the most influential person was chosen as the one with the highest overall influence over other members. The most dominant group members, leaders, competents, and the most liked ones were acquired from the questionnaire outputs. We then calculated the correspondence score as the ratio of overlap between the influential people and the people associated with the perceived variables. Table 4 indicates that 45% of the influential people are also dominant and leader, and that around 33% of them are competent. The least correspondence was observed for the liked people, in which only 20% of them are influential.

5. CONCLUSIONS

In this work, we presented a framework to detect the relationship of a group's performance with the interaction between group members and the individuals' personality traits. The results indicate that the interaction between group members is an important factor in determining the group performance. Relationships involving pairwise interactions (one-to-one), multiple interactions (many-to-many), and interactions for a specific person (one-to-many) characterize groups in the low, average, and high performance levels respectively. Besides, the group performance tends to be in parallel with the performance of the most influential person in the group. We also observed that the performances of individuals with personality traits have a moderate correlation with the group performance. We believe that our work provides a solid baseline for future applications such as performance predictions via influence interpretation.

6. ACKNOWLEDGMENTS

This research has been supported by the Swiss National Science Foundation (SNSF) Ambizione fellowship under the SOBE project (PZ00P2_136811).

7. REFERENCES

[1] S. Basu, T. Choudhury, B. Clarkson, and A. Pentland. Towards measuring human interactions in conversational settings. In *CUES*, 2001.

[2] J. H. Bradley and F. J. Hebert. The effect of personality type on team performance. *Management Development*, 16(5):337–353, 1997.

[3] W. Dong, T. Kim, and A. Pentland. A quantitative analysis of the collective creativity in playing 20-questions games. In *C & C*, pages 365–366, 2009.

[4] W. Dong, B. Lepri, A. Cappelletti, A. S. Pentland, F. Pianesi, and M. Zancanaro. Using the influence model to recognize functional roles in meetings. In *ICMI*, pages 271–278, 2007.

[5] W. Dong, B. Lepri, and A. Pentland. Automatic prediction of small group performance in information sharing tasks. In *CoRR*, 2012.

[6] W. Dong and A. Pentland. Quantifying group problem solving using social signal analysis. In *ICMI-MLMI*, pages 40–43, 2010.

[7] A. Haslam. *Psychology in Organizations: The Social Identity Approach.* SAGE, 2004.

[8] M. Higgs, U. Pewina, and J. Ploch. Influence of team composition and task complexity on team performance. *Team Performance Management*, 11(7/8):227–250, 2005.

[9] D. Jayagopi, D. Sanchez-Cortes, K. Otsuka, J. Yamato, and D. Gatica-Perez. Linking speaking and looking behavior patterns with group composition, perception, and performance. In *ICMI*, pages 433–440, 2012.

[10] B. Lepri, N. Mana, A. Cappelletti, and F. Pianesi. Automatic prediction of individual performance from "thin slices" of social behavior. In *MM*, pages 733–736, 2009.

[11] J. E. McGrath. *Groups: Interaction and Performance.* Prentice-Hall, 1984.

[12] M. H. Roy. Small group communication and performance: do cognitive flexibility and context matter? *Management Decision*, 39(4):323–330, 2001.

[13] D. Sanchez-Cortes, O. Aran, M. S. Mast, and D. Gatica-Perez. A Nonverbal Behavior Approach to Identify Emergent Leaders in Small Groups. *IEEE Transactions on Multimedia*, 14(3-2):816–832, 2012.

[14] J. B. Shaw and E. Barrett-Power. The effects of diversity on small work group processes and performance. *Human Relations*, 51(10):1307–1325, 1998.

[15] E. Sundstrom, K. P. de Meuse, and D. Futrell. Work teams: Applications and effectiveness. *American Psychologist*, 45(2):120–133, 1990.

[16] R. van Dick, J. Stellmacher, U. Wagner, G. Lemmer, and P. A. Tissington. Group membership salience and task performance. *Managerial Psychology*, 24(7):609–626, 2009.

[17] A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi, and T. W. Malone. Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330(6004):686–688, 2010.

[18] M. J. Zaki and W. M. Jr. *Data Mining and Analysis: Fundamental Concepts and Algorithms.* Cambridge University Press, 12 2014.