6-22-202

Discovery of Small Group Interactions and Performance from Project Emails

Sasha Ivkovic

MEHMOOD A. CHADHAR

Taiwo Oseni

Sally Firmin

Discovery of Small Group Interactions and Performance from Project Emails

Completed Research Paper

Sasha Ivkovic
Federation University Australia
Ballarat, Australia
s.ivkovic@federation.edu.au

Mehmood Ahmad Chadhar
Federation University Australia
Ballarat, Australia
m.chadhar@federation.edu.au

Taiwo Oseni
Federation University Australia
Ballarat, Australia
t.oseni@federation.edu.au

Sally Firmin
Federation University Australia
Ballarat, Australia
s.firmin@federation.edu.au

Abstract

Despite latest advances in small group research, discovery of group interactions and performance from analysis of small group communication, such as project emails, is still minimally represented. This paper presents a novel approach of studying small groups through analysis of the participants' emails sent to the project manager. We examined 1,105 email messages from managers' email in-boxes across five distinct ICT projects from the personal, social, collaborative, and engaging perspective of the email senders and link the findings to group performance. The study provides theoretical evidence that analysis of incoming communication from project managers' email in-box can be used to measure a group's success. For project managers the approach has the potential to be highly beneficial for monitoring of indicators for the state of project health.

Keywords: small group research, group performance, email analysis, IT projects.

Introduction

In the nineties a new paradigm for Small Group Research (SGR) was introduced that addresses the physical, temporal, and social context within which those groups are embedded. Studies of small groups include social interaction or conversation by people who are acting as group (team) members, having a common group goal, being assigned one or more group tasks to which need to be completed for some common purpose. In addition, they produce some outcome for which members bear collective responsibility and for which acceptability is potentially assessable. According to McGrath (1997) the IPO (input-process-outcome) framework can be used for studying team performance. Inputs describe antecedent factors that enable and constrain members' interactions. These include individual team member characteristics (e.g., competencies, personalities), team-level factors (e.g., task structure, external leader influences), and organizational and contextual factors (e.g., organizational design features, environmental complexity). These various antecedents combine to drive team processes, which describe members' interactions directed toward task accomplishment.

Historically, team processes were categorized as either "taskwork" or "teamwork" (Mathieu et al, 2008). At its essence, taskwork describes functions that individuals must perform to accomplish the team's task, whereas teamwork describes the interaction between team members. Processes such as taskwork and teamwork are important because they describe how team inputs are transformed into outcomes. Outcomes are results and by-products of team activity that are valued by one or more constituencies (Mathieu et al, 2000). Broadly speaking, these may include performance (e.g., quality and quantity) and members' affective reactions (e.g., satisfaction, commitment, viability). According to Mathieu (2008) strategies in this regard may include analysis of small group communication using qualitative research, time-sampling or diary-style investigations, and clever archival approaches. For example, teams who interact through virtual means leave traces of their interactions (e.g., threaded discussion lists, e-mails, video conferences), which, with the proper ethical consideration, can be used to index member interaction in a far more detailed way than would be available for traditional face-to-face teams.

Related literature on analysis of small group communication using qualitative research supports motivation for our study given the assumptions that a group's processes can be obtained from traces of their interactions, namely project related emails. There are several studies that support our motivation. For example, Biffl et al. (2010) demonstrate that mining project related emails can provide better understanding of a project group's performance. Snider (2017) also demonstrate the potential to understand the group activities in project by through the study of the transmission and content of email communication. However, group interactions and performance from analysis of small group communication, such as project emails, is still minimally represented. The purpose of this study is therefore to provide a comprehensive approach which considers discovery of group interactions and performance from project emails. This objective will be achieved using the following research question: "*How does analysis of groups' communications via email explain group performance?"*

This paper is structured as follows: we firstly summarise related work on analyses of email content at a project level for the purposes of measuring a group performance, and in so doing, identify the research issues. The methodology describes data collection from five distinct ICT projects and research design. Findings include discussions of the results obtained a snapshot communication from five projects (cross-sectional analysis) and examine how group communications via email can be used to explain group performance. Finally, we conclude the paper and identify further work.

Related work

Teams exist to perform tasks and performance is the most widely studied criterion variable in the organizational behaviour and human resource management literatures (Mathieu, 2008; Chadhar and Daneshgar,2018). Outputs of the communications generated as part of the project are2 fundamentally related to performance and that analysis of their content can provide understanding, insights and predictions about the project (Hicks 2013). According to Shi et al. (2015) project communication data could contain multiple indicators that can be used to evaluate and represent project performance from different perspectives. For example Wasiak et al. (2011) uses textual classification of engineering project emails to understand i) what topics e-mail discusses, ii) why e-mail is sent, and iii) and how the content is expressed. Classifying email content by the categories of management, information, and problem-solving transactions revealed signatures that align with project phases and, more importantly, problems encountered. Hicks (2013) also demonstrated that identifying different types of email could be useful for revealing problems encountered during project life-cycle indicating the potential value of email content for project management purpose. Work by Hicks (2013) and Wasiak et al. (2010) is advanced by Jones et al. (2015) where the authors uncover the evidence of patterns and trends relating to the email topics that correspond to significant events or issues within the engineering project studied. Snider et al. (2017) claims that such project-specific information extends the ability of managers to understand the activity within their specific engineering project scenario. Work by Ivkovic et al (2018) also highlighted the potential value

of analysing email content for project management purpose but used "communication metrics". However, the number of investigations of group characteristics extracted from project communication are still scarce, especially investigations in regard to *task production, group well-being,* and*, member support* group functions. According to McGrath (1993, page 2) groups can be characterised in terms of "set of activities and the outcomes that are generated by those activities of a particular set of members using a particular set of tools for a particular set of purposes in a particular set of physical, temporal, and sociocultural context. These activities and outcomes are manifest in three generic group functions – *task production, group well-being,* and*, member support"*.

In project environment the work within each scheduled task is coordinated by communication between managers and team members where the participants maintain project related issues. It often involves the project manager initialising a communication session by sending email message to one or more team members (and/or other interested parties) regarding project or task related issues and receiving response/s. As a result the email responses are in the form of written feedback which often contain the description of the respondents' activities and their experiences while performing these activities. Consequently, project manager's email in-box provides data and information rich repository from which the psychology, communication, and organizational behaviour of small groups can be extracted. In this context there is a need for advancing the analysis of the stored emails that would be useful for measuring group performance with a focus on identifying areas that work or are critical to the project progress. With a growing interest in research for analysis of small group communication via email using qualitative research appear important, although still rare.

These shortcomings provide motivation for this paper. In order to contribute to project management from Information Technology (IT) sector methodology and approach of this paper is designed to answer the research question: "*How does analysis of groups' communications via email explain group performance?"*

Research method and design

This study employed a uses a qualitative method of coding of emails for each group as described in Ivkovic et al. (2018). Team members' communication was discerned by using the longitudinal approach (observations through an extended period during the projects lifetime), while differences and characteristics of incoming communication across five distinct projects was identified via cross-sectional study. Cross-sectional research involves using different groups of people who differ in the variable of interest but share other characteristics such as in this example working on ICT related projects. The terms "ICT project", "project feedback", "performer/requester", "incoming communication" and "project tasks" provided the parameters, or boundaries of selection for projects. Purposeful sampling was adopted for recruitment of participating projects. Two out of five projects were software development projects while the other three projects involved implementation of IT systems as described in Table 1.

Table 1: List of Case Studies with details

#	Project Detail	Group size	Data Source
1	Freelance IT professionals from around the world working on an ongoing Free and Open Source Software (FOSS) development project. Team members work on-line in a virtual team environment.	10-15	120 email messages for a period of 11 months
2	Large Australian State Government organisation implementing a data mining IT project with internal team and external subcontractors from Europe.	20-25	429 email messages for a period of 16 months

3	ICT centre operating as a commercial part of a regional University implementing Geographic Information System (GIS) project with internal team and external subcontractors.	1-5	338 email messages for a period of 22 months
4	Department of a regional University implementing room access e-card project with internal team and external subcontractors.	10-15	80 email messages for a period of 11 months
5	Large Australian Power organisation implementing a GIS IT project with internal team and external subcontractors.	20-25	138 email messages for a period of 13 months

Data collection

The data set consisted of 1,105 email messages from five distinct ICT projects from four Australian organisations and one Free and Open Source (FOSS) software development organisation. In all five projects email was used as a primary communication tool as organisational teams worked either virtually or with external contractors from Australia and overseas. For one project, data was collected from publicly available mailing list containing communication between the Open Source Software developers and the project maintainer (project manager) working as virtual teams. In other cases team members' email messages were collected from the project manager's in-box. In order to ensure richness of available data project managers were asked to provide all email messages related to a single project which had recently finished and related to projects in which they were the project manager. The messages as communication records were not formed in isolation, rather were shaped and modified by event sequences and interactions between the team members and their project managers that unfolded over time.

Data analysis

Due to the researcher's domain knowledge in ICT, concepts and vocabulary from this industry were easily understood and interpreted when email content was analysed. This knowledge was essential for understanding the technical content of emails. The study of incoming communication required both qualitative and quantitative approaches for data collection and analysis. The development of the corresponding coding scheme for discovery of problems and expectations was developed by the researcher with assistance of project managers as described in detail in work (Ivkovic et al. 2018). Coding schema for discovery of emotions was based on work by Scherer (2005). Coding schema for discovery of lexical patterns was based on 11 grammatical pattern categories discussed by (Osherenko 2008) as "Grammatical lexical patterns from linguistic studies" which are also used to identify emotional meaning of texts. When coding was completed quantitative and qualitative results, based on multi-feature analysis were used. For each feature and their values, we provided the meaning and frequency count of the values. Initial results were obtained by applying chi-square test of independence for values of each feature in order to test if the relationship between a feature and projects was significant. The test was based on significance level (p-value) of 0.05. In cases where the null hypothesis was rejected, we applied residual analysis identifying feature values for which project's incoming communication differed significantly. These findings provided information about feature values that were commonly found across projects which in turn describe key group's characteristics through analysis of communication records.

In this paper the provide results from analysis of "communication content" based on linguistic and semantic information extracted from phrases within email message texts. By

using natural language parsing technique (through linguistic parser) we search for evidence of problems, emotions, expectations and lexical patters (PEEL) in each sentence/phrase extracted from email texts. For each PEEL element we also identified a source associated with the phrase by labelling it with one of the following options: W=WRITER, R=READER, 3P=THIRD-PARTY, IT=USED FOR OBJECTS. We define the "source" as being the origin of the PEEL element. The writer corresponds to message sender (team member), the reader corresponds to message recipient, third party corresponds to a team member from an external organisation, while 'IT' corresponds to an object such as server, database or code. This additional feature provides more detailed information about each PEEL element. For example, in sentence "*I'm just afraid of our product being blamed for security holes introduced by misconfiguration of misunderstood features.*", we identified the "*I'm just afraid* " phrase as text with emotional content where source of the emotion was the "WRITER" as the origin of the emotion is associated with subject (I). GALC and GI parsers associated the verb afraid with "GALC affective state = FEAR" and "GI polarity = NEG (Anxiety,dread)". The fear in this instance is not real as "afraid" is used in a different context however, in the context of team member's "concern" this entry was recorded in the findings.

Results

In this section we report on findings based on email analysis discussed in detail by Ivkovic et al. (2018). These results are investigations of PEEL analysis extracted from project communication from all projects (1105 emails) in the context of *task production, group well-being,* and*, member support* group functions (McGrath, 1993). Through interpretation of the results we answer the proposed research question "*How does analysis of groups' communications via email explain group performance?"*

In the conclusion we summarise the paper and offer actionable advice to academics and project professionals in regard to understanding group characteristics through complexity of interpersonal communication in projects.

Results on analysis of "communication content" from all projects

The purpose of the PEEL analysis is to bring attention to problems and issues, emotions and affects linked to negative polarity, team members' expectations, as well as expressions of sentiment discovered from hidden texts such as upper-case characters, emphasis, exclamations and emoticons. In regard to groups' characteristics by (McGrath 1993, page 2) and the three generic group functions *task production, group well-being,* and*, member support"* problems are related to *task production*, emotions and affects are related to *group well-being* while team members' expectations are related to *member support*.

Problems: Through analysis of text from 1105 emails across all project groups, 210 problems were found in 161 email messages (14.6%) with ten "problem" categories identified. Each problem category might appear in more than one email message and each email message might contain one or more problem categories. Figure 1 illustrates project related distribution for each problem category. The legend shows how is each project associated with a colour coded bar. For example, in the group of bars for problem id 4 (Knowledge), the height of the red bar (Project 1), shows that more than 60% of all problems identified in Project 1 were related to knowledge (id number 4). This problem category is associated with the "Inability to deliver (complete a task) due to lack of KNOWLEDGE". Project 1 team members were more likely to report knowledge related problems and issues than team members from other projects. After residual analysis was applied, high standard residual of 2.75 for project 1 indicated that the actual frequency count of knowledge related problems was higher than expected. In addition, we found that for Project 5 (the black bar) knowledge related problems (id number 4) were mostly reported as "*I don't know*" or "*I am not sure*" my message senders. This finding

suggests that perhaps some team members could need additional training. These findings demonstrate how identification of problems can be used to measure *task production.*

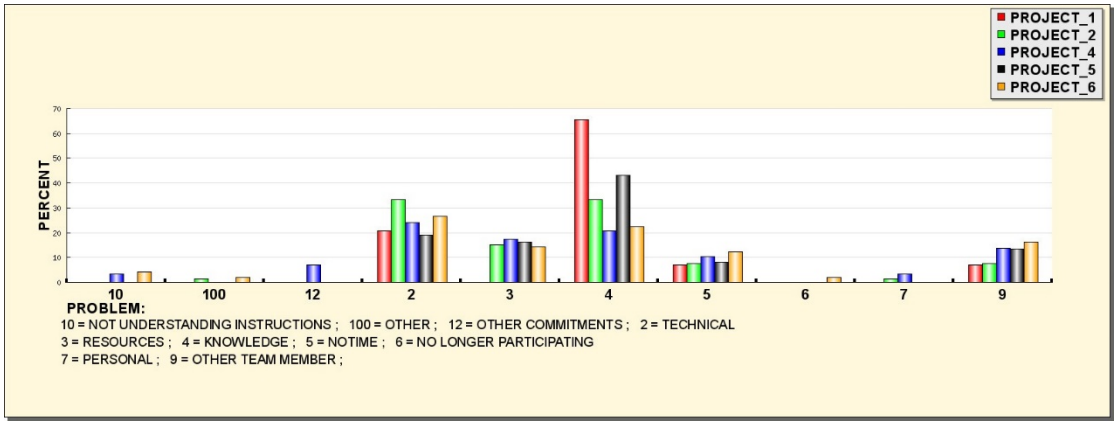

Figure 1: Percentage of reported problems in projects

In Table 2 we show sumary of findings for common PROBLEM categories (column 2) with given project ids shown in the first row, columns 3-7. For example, a high proportion of Project 5 messages were associated with problems (labelled as **High**). For each project the proportions are calculated by dividing number of messages associated with a problem category by the total number of messages. The problem categories are listed in column 2, with "how often" shown in columns 3-7 for each project respectively. For example, in Table 2 under the first row (problems) the group of values in column 6 indicate that Project 5 team members often reported "knowledge" related problems, shown in intersection of row with problem id 4 and column 6 (Proj 5). Furthermore, we report that source of the knowledge reported problems for Project 5 was the team member as message writer (source=W). The source indicates that most common subject in phrases associated with knowledge related problems was grammatical first person; either singular (I) or plural (we).

Table 2: Overall proportion of PROBLEMS per PROJECT

#	PEEL Feature	Proj 1	Proj 2	Proj 4	Proj 5	Proj 6
	PROBLEMS - Proportion of Messages with Issues/problems reported by project team members	**Moderate**	**Low**	**Low**	**High**	**High**
2	TECHNICAL; Problems due to technical constraints (e.g power failure, network unreachable, server down, IT servicesunavailable)	Sometimes (source W)	Often (source IT)	Sometimes (source IT,W)	Sometimes (source IT)	Often (source IT,R,W,3P)
3	RESOURCES; Problems due to lack of resources (e.g. no templates, no documentation, no written instructions)	Never	Sometimes (source W)	Sometimes (source W)	Sometimes (source W)	Sometimes (source IT,W)
4	KNOWLEDGE; Problems due to knowledge constraints (e.g. does not understand what is required or does not know how to complete the task)	Very Often (source W)	Sometimes (source W)	Rarely (source W)	Often (source W)	Rarely (source W)
5	NO-TIME; No time assigned for the task. Running out of time to complete the task (e.g. other commitments/constraints or lack of time management)	Rarely (source W)	Rarely (source W)	Rarely (source W)	Rarely (source W,3P)	Sometimes (source W)
9	OTHER TEAM MEMBER; Problems due to other team members (e.g. waiting for their task completion)	Rarely (source 3P)	Rarely (source 3P)	Rarely (source 3P)	Rarely (source 3P)	Sometimes (source 3P)

Emotions articulated by team members: This category provides information about emotional content of the email in relation to effective states. For each phrase associated with emotional content a polarity was flagged as either positive or negative. Across all projects, in 86 out of 1105 (7.8%) email messages, we identified 106 sentences, each containing a verb or adjective associated with an effective state. The most common verb was "agree" (positive polarity), followed by "apologise", "happy", and "delighted". Verbs associated with negative polarity, such as "disagree", "nervous", and "frustrated" were not frequent. This approach allows us to compare polarity across projects as illustrated

in Figure 2. The legend shows list of projects each associated with a colour coded bar. Colour coded group of bars represent polarity frequency counts for each project for a given polarity. For example, in the group of bars for positive polarity labelled as "POS", the height of the red bar representing Project 1 shows that 28 (out of 47) verbs or adjectives associated with emotional content for Project 1 were associated with positive polarity. Textual phrases identified from those email messages mostly contained subject "I" and verb "agree'. This finding indicates that Project 1 team members were likely to agree with the project manager in their discussions which can be linked to positive *group well-being.* In regard to negative polarity labelled as "NEG" for Project 1, further analysis of team members' emails included verbs such as "unsure", "disagree", "confused", "worried" and "unhappy". By combining verb of each phrase with subject and object we were able to provide more detailed explanation as follows: "I *unhappy* solution", "I *worried* time-line", "I *surprised* common usage cases changed". This finding demonstrates how negative polarity can be used to measure negative *group well-being.*

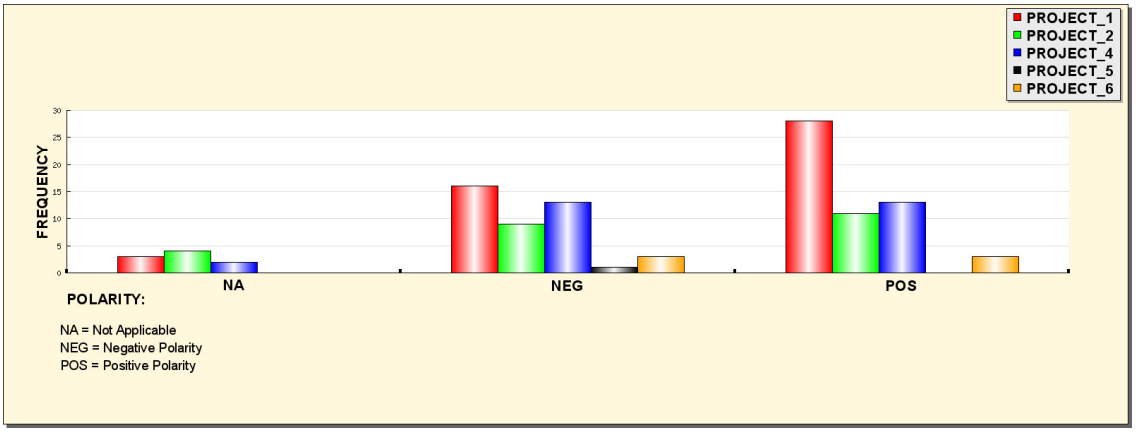

Figure 2: Polarity of emotional content in projects (frequency)

In Table 3 we show sumary of findings for Polarity of emotional content in project (column 2) with given project ids shown in the first row, columns 3-7. For example, the table shows that amongst all projects, Project 1 messages were more likely to be associated with emotional content.

Table 3: Overall proportion of Emotional Polarity per PROJECT

#	PEEL Feature	Proj 1	Proj 2	Proj 4	Proj 5	Proj 6
	EMOTIONS - Proportion of Messages with Affective States linked to Polarity	**High**	**Low**	**Low**	**Low**	**Low**
	Messages contain verbs/adjectives linked to POSITIVE Polarity	Sometimes (source W)	Rarely (source W)	Rarely (source W)	Never	Rarely (source W)
	Messages contain verbs/adjectives linked to NEGATIVE Polarity	Sometimes (source W)	Rarely (source W)	Rarely (source W, R)	Rarely (source W)	Rarely (source W)

Team members' expectations: Throughout analysis of 1105 email messages from all projects, in 538 messages (48.7%) we identified 14 team members' expectation categories. Figure 3 shows project related distribution for each expectation category. The figure illustrates that across all projects the most common expectation category was "task completion" (id 4). This type of expectation is related to emails where team member as a message sender was expecting message receiver to complete a task which corresponds to requests for member support. Message excerpts "If possible, please add these to the list (both are relatively critical)" and "Can you please have a look at it and reply to all when fixed please?" illustrate "task completion" type of expectation (id 4). The legend shows list of projects with corresponding colours relating to colour coded bars. For example, in the group of bars for expectation category with id 4 (task completion), the height of the blue

bar (Project 4), shows that more than 35% of all team members' expectations identified in Project 4 were related to task completion. The residual analysis shows high standard residual of 2.98 for Project 4 indicating that the actual frequency count for "task completion expectation" was higher than expected. This finding demonstrates how discovery of team members' expectations can be used to measure *member support*.

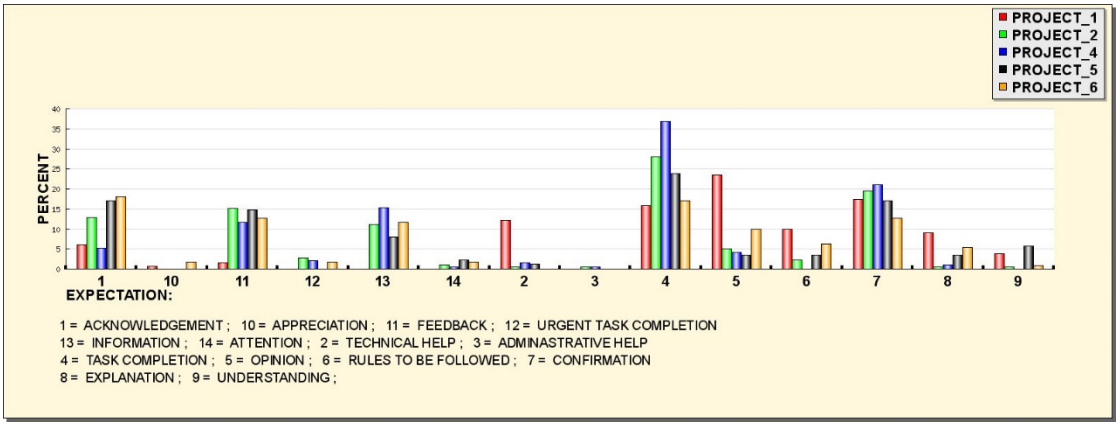

Figure 3: Percentage of team member expectations in projects

In Table 4 we show sumary of findings for common EXPECTATION categories (column 2) with given project ids shown in the first row, columns 3-7. For example, a high proportion of messages from Project 1 and Project 5 messages were associated with higher expectations (labelled as **High**) on the message receiver (denoted as source R). Project 1 team members were more likely to expect "Opinion" (id 5) than team members from other projects. This type

Table 4: Overall proportion of EXPECTATIONS per PROJECT

#	PEEL Feature	Proj 1	Proj 2	Proj 4	Proj 5	Proj 6
	EXPECTATIONS - Proportion of Messages with Team members' expectations	**High**	**Low**	**Moderate**	**High**	**Moderate**
1	ACKNOWLEDGEMENT; expects only a simple acknowledgement from the message recipient	Rarely (source R)	Sometimes (source R)	Rarely (source R)	Often (source R)	Often (source R)
4	TASK COMPLETION; has expectations of the message recipient to directly or indirectly take some action for a task completion	Sometimes (source R)	Very Often (source R)	Often (source R)	Sometimes (source R)	Sometimes (source R)
5	OPINION; expects to hear message recipient's opinion about items mentioned in the message	Very Often (source R)	Rarely (source R)	Rarely (source R)	Rarely (source R)	Sometimes (source R)
6	RULES TO BE FOLLOWED; has expectations of the message recipients to follow some rules or standards	Sometimes (source R,3P)	Rarely (source R)	Never	Rarely (source R,3P)	Rarely (source R,3P)
7	CONFIRMATION; has expectations of the message recipient to validate, verify or give assurance	Often (source R)	Often (source R)	Often (source R)	Sometimes (source R)	Rarely (source R)
11	FEEDBACK; has expectations of the message recipient to provide feedback with comments and instructions or to further elaborate on items mentioned in the message	Never	Sometimes (source R)	Rarely (source R)	Sometimes (source R)	Rarely (source R)
13	INFORMATION; has expectations of the message recipient to provide information on items mentioned in the message	Never	Sometimes (source R)	Often (source R)	Rarely (source R)	Sometimes (source R)

of expectation is related to emails where team member, as message writer expects a reply based on subjective discussions about propositions and new ideas mentioned in the message. Message excerpt "*What do you think about it*?" illustrates this type of expectation.

Hidden meaning in email texts expressed as Lexical Patterns: From 1105 messages in 80 (7.2%) we identified 92 phrases containing lexical patterns linked to emotional meaning of texts. Out of 11 lexical pattern categories we found three most commonly used lexical pattern categories in team member's emails as as illustrated in Figure 4. "Emphasis" (id 6) are used as

special weight placed on something considered important, "Exclamations" (id 2) are used for dramatic effect, while "Interjections" (id 1) are used to express hesitation, realisation and doubt. For example, in the group of bars in Figure 4, for "lexical pattern" category id 6 (Emphasis), the height of the green bar (17 phrases) shows that Project 2 team members were more likely to use emphasis in their email text than other team members. Our analysis of email revealed two common writing methods that contained hidden meanings in team members' text: a) usage of upper case characters, emoticons, and exclamation marks to emphasize feelings, emotions and moods, and b) usage of question marks, filled pauses and hesitation devices such as "Uuhm", "uh", "ah" and "oh" to express hesitation, realization and doubt. For example, Project 1 team members were more likely to use interjections in their email texts than other team members. For example in this message excerpt "*Uuhm, I've had a bad day and mixed up some stuff, I guess.*", the Uuhm interjection is implying hesitation by the team member. In our analysis we interpreted this phrase as "*I guess I made a mistake*". The findings demonstrate how identification of hidden meaning in email texts can be used to measure *group well-being* (usage of upper case characters, emoticons, and exclamation marks to emphasize feelings, emotions and moods) and to measure *task production* through identification od hesitations, realizations and doubts.

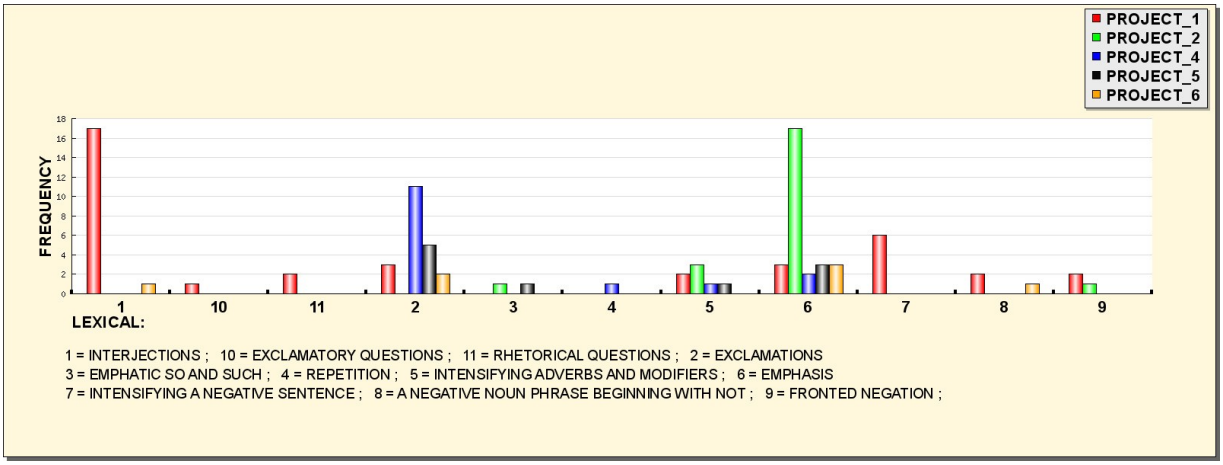

Figure 3: Lexical patterns in project emails (frequency)

In Table 5 we show sumary of findings for common Lexical Pattern categories (column 2) with given project ids shown in the first row, columns 3-7. For example, a high proportion of messages from Project 1 messages were associated with lexical patterns (labelled as **High**).

Table 5: Overall proportion of Lexical Patterns per PROJECT

#	PEEL Feature	Proj 1	Proj 2	Proj 4	Proj 5	Proj 6
	LEXICAL PATTERNS - Proportion of Messages with Hidden Meaning in Email Texts expressed as *Lexical Pattern*	**High**	**Low**	**Low**	**Low**	**Low**
1	Messages contain Interjections; expresses hesitation, realisation and doubt	Sometimes (source W)	Never	Never	Never	Rarely (source W)
2	Messages contain Exclamations; used for dramatic effect	Rarely (source W)	Never	Sometimes (source W)	Rarely (source W)	Rarely (source W, R)
6	Messages contain Emphasis; putting emphasise on something considered important	Rarely (source W, IT, R)	Sometimes (source W)	Rarely (source W)	Rarely (source W)	Rarely (source W)

Discussion

In order to answer the research question: "*How does analysis of groups' communications via email explain group performance?"* we focus on McGrath (1993, page 2) statement that groups

can be characterised in terms of "set of activities and the outcomes that are generated by those activities of a particular set of members using a particular set of tools for a particular set of purposes in a particular set of physical, temporal, and sociocultural context. These activities and outcomes are manifest in three generic group functions – *task production, group well-being,* and*, member support"*.

Problems reported by team members provide insights into project issues which are directly linked to *task production* process. Analysis of textual content of emails related to problems revealed that team members' were often not able to complete a project task due to insufficient knowledge. Further analysis of these sentences showed that the source of the knowledge problems was often the 'WRITER'. For example, in this message excerpt "I was not sure what is the right place to add this." the subject "I" indicates that the writer reports about insufficient knowledge of himself. The finding indicates that across all projects team members as message writers reported on knowledge problems commonly related to themselves. While this finding is to some degree expected, it is important to note that discovery of problems reported in project emails have different meaning when the "source" attribute is taken into account. In this instance results from the analysis can identify members and areas of their insufficient knowledge. In order to improve group's performance the management can provide additional information, organise workshops, additional training, or assign another person to assist in this matter. The second most common problem category was due to technical constraints (e.g. power failure, network unreachable, server down, IT services unavailable) and problems due to lack of resources (e.g. no templates, no documentation, no written instructions). From the performance point of view those messages were associated with project tasks and sub-tasks which were not progressing well.

Identification of email content for "expectations" enabled us to measure groups' functions where *member support* was required. For example discovery of messages where urgent reply was expected from a group member, expectations of reading and replying to work related emails in a timely manner, and being responsive is seen as being sensitive and caring.

We demonstrated that analysis of emails associated with emotional content can be used to identify members' frustrations or dissatisfaction and as such can be used to measure *group well-being*. Throughout all projects only a small proportion of email messages contained emotional sentences. However, discovery of emotions provided valuable feedback as they could be linked to negative polarity, identifying team members' reactions to their negative experiences while performing project activities.

In addition to discovery of verbs linked to emotional content we applied lexical pattern analysis. The use of emotion-denoting words sentences from email texts were tested for occurrences of emoticons, abbreviations, acronyms, interjections, "?" and "!" marks, repeated punctuation and capital letters. We demonstrated that analysis of message text can provide window into email senders' emotive and cognitive worlds. Through lexical pattern analysis *group well-being* can measured through discovery of upper case characters, emoticons, and exclamation marks that emphasize feelings, emotions and moods while *task production* can measure through identification od hesitations, realizations and doubts.

Conclusion and outlook

The main purpose of this paper was to explain group performance from group communications via email. Historically, team processes were categorized as either "taskwork" or "teamwork" (Mathieu et al, 2008). In this paper we focused on teamwork which describes the interaction between team members via "communication". Our investigation through analysis of PEEL features are based on analysis of individual phrases extracted from message text sentences.The contribution provides conceptual understanding of the role that "communication" plays in the different project environments as well as practical implementation of linking email analysis to

group performance. In order to answer the research question and explain group performance from group communications via email we outline the following claim by Mathieu (2008) "Teams exist to perform tasks and performance is the most widely studied criterion variable in the organizational behaviour and human resource management literatures. In part, this can be attributed to measurement issues". In project-based settings, individuals are often part of a resource pool that is drawn from according to some combination of their knowledge, skills, and abilities (KSAs) and the needs of the project or team task. Whereas member cognitive ability may be valuable for teams in general, task-related knowledge levels are likely to be even more important to teams that perform a task over time (Mathieu 2008). However, we believe that beside KSAs group performance can be predicted by their engagement in communication. Engagement is directly linked to team members' performance which in turn determines project progress. Project managers' email in-boxes provided references to actual project events and led to the position that the in-box provided an information rich repository of written records related to team members' engagement in projects. Understanding engagement involves "decoding" of incoming emails through implementation of multi-variant observations including:

1. Evidence of psychological engagement; through team members' use of emails to express their experiences, thoughts, ideas, opinions, reflections and feelings to their project manager (and other team members).
2. Evidence of physical engagement; which in this instance acts as a self-driven mechanism for expressing in words reflections and experiences about their physical engagement related to act of "doing" (completing project tasks).
3. Evidence of communication engagement; which is linked to physical and psychological engagement therefore it acts as a relationship between the sender and the receiver, as well as a vehicle for the transfer of written content such as thoughts and status updates.

Through application of the email analysis from managers' email in-boxes across five distinct ICT projects from the personal, social, collaborative, and engaging perspective of the email senders we were able to linking the findings to group performance. In regard to groups' characteristics by (McGrath 1993, page 2) and the three generic group functions *task production, group well-being,* and*, member support"* we demonstrated how identification of problems can be used to measure *task production*, identification of emotions articulated by team members can be used to measure *group well-being*, and, discovery of team members' expectations can be used to measure *member support*. In addition we demonstrated how identification of hidden meaning in email texts can be used to measure *group well-being* (usage of upper case characters, emoticons, and exclamation marks to emphasize feelings, emotions and moods) and to measure *task production* through identification od hesitations, realizations and doubts. On the basis of team members problems, emotions linked to negative polarity, expectations and other emotional content discovered from hidden texts such as upper case characters, emphasis, exclamations and emoticons we were able to explain groups' performances across six ICT related projects.

In order to provide depth of understanding, and triangulation of the data further research would involve a questionnaire with project managers which would be used to evaluate analysis approach presented in this paper. The purpose of the questionnaire is to determine project manager's views on values of extracted knowledge from email analysis presented in this paper in the context of improving our analysis approach.

Study limitations

This research is limited to written communication via email collected from IT related projects and so is limited to feedback that is in a textual form collected from a central communication point. While applied analysis and recommendations provided potentially useful "calculations of reality" the primary researcher and practitioners needed to commit time to work together on the interpretation of the results. On the

basis of the project manager's domain knowledge it was the managers decision to choose which discovered areas of concern were important for further management actions.

References

Biffl, S., Sunindyo, W.D. and Moser, T. 2010, July. "A Project Monitoring Cockpit Based On Integrating Data Sources in Open Source Software Development," In *SEKE* (pp. 620-627).

Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M., & MacKenzie, S. B. 1995. On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. Personnel Psychology, 48: 587-605.

Hicks, Ben. 2013. "The language of collaborative engineering projects," *In: DS 75-6: Proceedings of the 19th International Conference on Engineering Design (ICED13)*, Design for Harmonies. Vol. 6, pp. 19-22.

Chadhar, M., & Daneshgar, F. (2018). Organizational Learning and ERP Post-implementation Phase: A Situated Learning Perspective. *J. Inf. Technol*, *19*, 138-156.

Ivkovic, S., Firmin, S., Oseni, T., & Chadhar, M. (2018). Communication Metrics Extracted from Project Managers' Email In-Box. In PACIS (p. 103).

Jones, S., Payne, S., Hicks, B., Gopsill, J. and Snider, C. 2015, July. "Subject lines as sensors: co-word analysis of email to support the management of collaborative engineering work," In *International Conference on Engineering Design 2015 (ICED 2015)*. University of Bath.

Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. Journal of management, 34(3), 410-476.

McGrath, J. E., Arrow, H., Gruenfeld, D. H., Hollingshead, A. B., & O'Connor, K. M. (1993). Groups, Tasks, and Technology: The Effects of Experience and Change. Small Group Research, 24(3), 406–420. https://doi.org/10.1177/1046496493243007

McGrath, J. E. (1997). Small group research, that once and future field: An interpretation of the past with an eye to the future. Group dynamics: Theory, research, and practice, 1(1), 7.

McGrath, J. E., Arrow, H., & Berdahl, J. L. (2000). The study of groups: Past, present, and future. Personality and Social Psychology Review, 4(1), 95-105.

Osherenko, A.(2008). Towards semantic affect sensing in sentences. In Proceedings of the AISB 2008 Symposium on Affective Language in Human and Machine (pp. 41-44).

Shi, L., Newnes, L., Culley, S., Gopsill, J., Jones, S. and Snider, C. 2015, July. "Identifying and Visualising KPIs for Collaborative Engineering Projects: a Knowledge Based Approach," In *ICED15, The 20th International Conference on Engineering Design, 2015*. University of Bath.

Scherer, Klaus (2005). "What are emotions? And how can they be measured?" In: Social Science Information 44, pp. 695-729.

Snider, C., Škec, S., Gopsill, J.A. and Hicks, B.J. 2017. "The characterisation of engineering activity through email communication and content dynamics, for support of engineering project management," *Design Science*, *3*.

Snider, C., Gopsill, J. A., Jones, S., Shi, L., & Hicks, B. 2015. "Understanding engineering projects: an integrated vehicle health management approach to engineering project monitoring," *In DS 80-3 Proceedings of the 20th International Conference on Engineering Design (ICED 15)* Vol 3: Organisation and Management, Milan, Italy, 27-30.07. 15 (pp. 065-074).

Tashakkori, A., & Creswell, J. W. 2007. The new era of mixed methods.

Wasiak, J., Hicks, B., Newnes, L., Dong, A. and Burrow, L. 2010. "Understanding engineering email: the development of a taxonomy for identifying and classifying engineering work," *Research in engineering design*, *21*(1), p.43.

Wasiak, J., Hicks, B., Newnes, L., Loftus, C., Dong, A., & Burrow, L. 2011. "Managing by e-mail: what e-mail can do for engineering project management," *IEEE Transactions on engineering management*, 58(3), 445-456.