Measuring Awareness in Cross-Team Collaborations – Distance Matters

Zia Ur Rehman Kiani
Blekinge Institute of Technology
Karlskrona, Sweden
zuki09@student.bth.se

Darja Šmite
Blekinge Institute of Technology
Karlskrona, Sweden
darja.smite@bth.se

Aamer Riaz
Blekinge Institute of Technology
Karlskrona, Sweden
aari10@student.bth.se

Abstract— **Developing and maintaining team awareness within and across teams working in the same project helps team members in aligning their activities and facilitates implicit coordination. This requires both task and presence awareness. In this paper, we share our findings from a survey in which we measured the level of team awareness in cross-team collaborations with varying degree of separation. To measure the levels of awareness we asked questions like who is who, who knows what, who is on a vacation, who depends on whom and alike. Results from surveying 17 pairs of teams from 15 organizations indicate that level of awareness in cross-team collaborations is generally lower than that within the teams. We also found that task and presence awareness levels are independent and can vary. In addition to distance, we identified a few other factors with potential positive and negative influence on team awareness.**

Keywords— ***Distributed Software Development, Awareness; Task awareness; Presence awareness***

I. INTRODUCTION

High employee retention costs and a lack of skilled resources at a single location [1] led to an increasing globalization of software companies. As a result, distributed and global software development (GSD) became a usual business, and so did large-scale development with developers and teams working jointly in the same software project across geographical, temporal and socio-cultural boundaries [1, 2].

Software development activities are usually carried out in a collaborative fashion [3]. Although individual knowledge is essential in a software development team for performing a task efficiently, it is insufficient when it comes to the task of large-scale distributed software projects [4, 5]. Due to complex dependencies [6], such teams often need to coordinate their work for the successful integration of knowledge from multiple functional and technical domains [4]. Coordination in global software development implies that the individuals working on a common project have common views about the product being built, its functionality, expected outcome, project organization, definitions and shared knowledge about team, task and activities [7]. However, software development projects have a complicated nature due to their characteristics of scalability, uncertainty, interdependency and communication [7]. This is why in reality global software projects are often challenged [8]. Various centrifugal forces, such as dispersion, coordination breakdown, loss of "communication richness", loss of "teamness" and cultural differences [9], aggravate cross-site collaboration.

One emerging area with a potential of positive impact on the practice of global and distributed software development is support of coordination and communication through awareness [3]. Communication generates the commitments that are managed by coordination, which, in turn, arranges the tasks for cooperation [10]. Coordination theory defines the coordination as the *"management of dependencies among task activities"* [4]. Studies of distributed teams reveal that knowledge about the task activities (both group and individual) and information sharing are essential for the successful coordination [4]. Information about all these factors contributes towards the concept, which is referred to as awareness [11].

In this paper we use existing research on the various concepts related to team awareness (see Section II) to create a survey that measures the level of intra-team and inter-team awareness (the survey and research approach are described in Section III). We further study 13 different pairs of teams and use the survey instrument to explore the impact of distance on the levels of awareness (see Section IV). Based on the cross-case analysis and links to related literature we then discuss factors that could have influenced the results (see Section V). Finally, we conclude the paper with a summary of our findings (see Section VI).

II. BACKGROUND AND RELATED WORK

There are several team cognition mechanisms, including shared mental model [12], collective mind [13], transactive memory [14] and team situation awareness [15], which are conceptually different but based on some form of the shared knowledge. Regardless of the task domain, awareness is essential to coordinate the group activities.

Awareness and familiarity are similar but conceptually different constructs. They have acute differences that categorize them into separate groups. Familiarity is a permanent knowledge acquired through learning and experience whereas awareness is situational and provides the contextual knowledge about the current task activities, which is discarded soon after the accomplishment of the task [4]. For instance, working on a task of a software project, developer

gets experience, which adds to his profile and is permanent, but the deadline of that task during the work is situational. It might be critically important at one stage but after meeting the deadline it gets discarded and has no use. Furthermore, familiarity helps to perform the task competently and efficiently in a GSD setup [16, 17] whereas awareness facilitates it by effectively aligning the task activities with the desired goal [4, 11].

Team cognition research suggests that when the team members interact with each other to perform their task activities, with the passage of time, they develop a common understanding about the shared work, which is referred as team knowledge. Team members coordinate their work implicitly through team knowledge as they understand the activities of the other team members [4, 18]. This type of coordination is referred to as the synchronization of the activities of the team members [19].

Developing and supporting the awareness to enhance coordination in distributed software development has been a central point of interest for the last two decades [20, 21]. A generally accepted definition of awareness provided by Dourish and Bellotti is "*an understanding of the activities of others, which provides a context for your own activity*" [11]. Another, more specific definition about geographically dispersed synchronous collaborative group, provided by Gutwin and Greenberg, is:

"*The up-to-the-minute knowledge of other people's activities that is required for an individual to coordinate and complete their part of a group task. Group awareness is maintained by keeping track of information such as other participants' locations in the shared space (where are they working?), their actions (what are they doing?), the interaction history (what have they already done?), and their intentions (what are they going to do next?)*" [22].

Gutwin and Greenberg also proposed a model incorporating different elements of workspace awareness [23], which can be moulded and extended to explore the awareness (see Table I).

TABLE I.   ELEMENTS OF WORKSPACE AWARENESS [23]

Category	Element	Specific Question
Who	Presence	Is anyone in the workspace?
	Identity	Who is participating? Who is that?
	Authorship	Who is doing what?
What	Action	What are they doing?
	Intention	What goal is that action part of?
	Artifact	What object are they working on?
Where	Location	Where are they working?
	Gaze	Where are they looking?
	View	Where can they see?
	Reach	Where can they reach?
How	Action history	How did that operation happen?
	Artifact history	How did this artifact come to be in this state?
When	Event history	When did that event happen?
Who (past)	Presence history	Who was here, and when?
Where (past)	Location history	Where has a person been?
What (past)	Action history	What has a person been doing?

Several tools and techniques [21, 22, 24, 25] have been developed to improve the consciousness among team members by providing the awareness about individual and group activities. *Computer supported cooperative work (CSCW)* applications use different approaches to disseminate actively generated awareness information [11]. Mainly these tools support the group members to achieve the implicit coordination [25, 26] in a multi-team environment, including geographically separated teams.

Literature reveals that CSCW tools, which support the development activities by introducing awareness in GSD environment, can be generally categorized into two groups. Traditional coordination tools provide the basic level of information, e.g. keyboard strokes and mouse clicks, whereas configuration management tools provide information at the individual file or method level [5]. These tools are categorized based on the method they use to provide the awareness information about the activities. Some provide awareness of activities at repository level (BSCW [27], CVS-Watch [28], COOP/Orm [29]) while others provide real time support (JAZZ [30], NightWatch [31], Palantir [32]).

The rationale behind the development of such tools is to create awareness among developers, which helps them to coordinate their work and to identify the potential problems earlier in the phase of development and avoid those during integration [4, 5, 33]. However, recent research empirically demonstrated that teams collaborating across the distance, although affected by the coordination and communication challenges, did not use the available tools and techniques for building awareness as expected [34, 35]. The results indicated that presence of people from different ethnic backgrounds in a collaborative environment produces the need to manage cultural and behavioral issues along with the technical solutions to support awareness [34]. The cross-team coordination challenges are also escalated due to the lack of personal contact and proximity [36, 37]. Existing research shows that increase in the distance between the collaborative teams gradually limits the opportunities of rich collaboration [38] and makes it progressively more difficult for the teams to coordinate their task activities and maintain the presence information about the counterparts [36]. Based on a survey of 775 Microsoft software engineers Begel et al. [37] found that though coordination is difficult, almost all engineers are required to coordinate with others to get their work done. They learned that intra-team and inter-team coordination communication modes varied greatly, due to geographic and personal separation. As a result of a lack of communication, almost fourth of respondents complained that work items they depended on have changed without any notification [37]. While awareness tools have a potential to alleviate these challenges, it is also evident that practices applied locally might not always work globally. This is why a better understanding of the creation of awareness in a multi-team environment is needed.

In this study, we are interested to understand the variation of levels of awareness between the teams collaborating across varying distances. Distance makes the coordination more difficult for large-scale software development teams [38]. Research shows that increase in the distance between the

collaborative teams, gradually limits the opportunities of rich collaboration [39] and consequently makes it progressively more difficult for the teams to coordinate their task activities and maintain the presence information about the counterparts [38]. We therefore would expect that levels of awareness would vary from the high to the low with an increase in separation. Nonetheless, distributed teams might be better equipped with specific collaboration tools, and therefore the answer to the question of the variance in levels of awareness might not be as straightforward.

III. RESEARCH METHODOLOGY

Based on the analysis of related work we identified the following research questions:

RQ1. *What is the level of team awareness (i.e. task awareness and presence awareness) in collocated versus distributed teams of varying distance?*

RQ2. *Which factors affect the team awareness negatively?*

To address the research questions, we performed an empirical survey and interviews with practitioners from 15 software companies, and built our conclusion based on a mixture of qualitative and qualitative data analysis [23]. More details are described in the following sub-sections.

A. Conceptualising the Study Setup

While some studies suggest that coordinating work activities in collocated environment is easy [38], others argue that collaboration declines if the collaborators are more than 30 meters away or they have to climb up the stairs to see the counterparts [39]. Therefore, we started by separating the types of collaboration under study into four different levels, including the separation within the same building for our study:

- **Group-1** Same building, same floor, distance less than 30 meters;
- **Group-2** Same building, different floor or the same floor, but distance more than 30 meters;
- **Group-3** Different locations, but the same time zone;
- **Group-4** Different locations in different time zones.

Next we defined the method of inquiry. Since the focus of investigation was cross-team collaboration, we considered pairs of teams as targeted participants for the survey. A three steps approach (see Fig. 1) was used to perform the survey:

1. In the first step, researchers asked the questions about the interviewee (T1-M1 is the member of Team-1) and inquired about his/her current task and availability. We consider this information as the baseline information about the team member, which is used to measure the awareness in the comparison analysis.

2. Then, we inquired about his/her team members within the same location (T1-M2 and T1-M3 are the members of Team 1) and measure the level of awareness within the team.

3. Finally, we asked questions about all the members of the remote team working on the same project (T2-M1, T2-

M2 and T2- M3 are the members of Team 2). This gave us the insight into cross-team awareness.

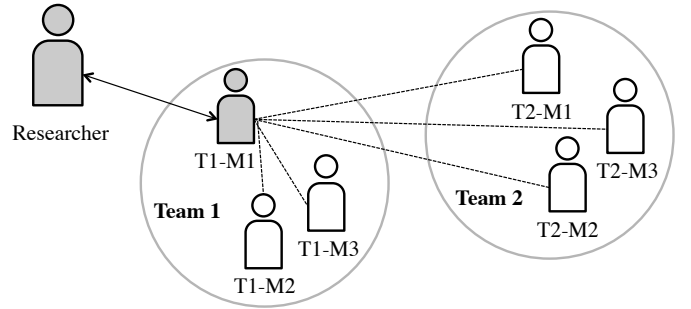The process was repeated for each member of both teams.

T – Team, M – Team member

Fig. 1. Generic setup for the data collection.

B. Questionnaire Building

Based on the conceptual setup, the questionnaire was built to assist with the measurement of the levels of awareness. In order to satisfy our requirement, we used the elements of workspace awareness (11 out of 16) proposed by Gutwin and Greenberg [22] (See Table I). We did not use two elements (artifact and artifact history) due to the privacy concern of the participating companies. We also did not use the gaze, view and reach elements as they were not relevant to our study. Further, we not only used the present and past elements to define the factors but also defined the future form to explore how much participating teams know about the future events and intentions of each other. Moreover, we defined several factors against each element of the workspace. For example, against the element location we defined the factors such as team's and team members' location in office, seating topology and location of supporting documents. The data was inquired to gain an in-depth knowledge about the situation under study.

As suggested by Gutwin and Greenberg, the elements were used to explore the awareness in different scenarios of virtual collaboration [22]. The elements were extended and moulded using related literature on geographically distributed software development, which was surveyed using snowballing search [36]. After that, narrative interviews were conducted with six software developers who volunteered to share their experiences, that we used to validate the awareness factors that emerged from related literature analysis. During the interviews we asked the volunteers to explain their routines – such as task activities, task and team coordination needs, cross-team interaction processes, problems faced, methods and techniques used to carry out the tasks, acquiring information about other team members (locally as well as globally) and supporting tools used. We analyzed the qualitative data gathered from the volunteers by using thematic coding [40]. The findings confirmed the relevance of the factors identified from the literature. Surprisingly, no new information was retrieved. The list of the factors with their implication on within-team or cross-team collaboration, and awareness information they support is given in Table II.

TABLE II. FACTORS USED TO MEASURE THE LEVEL OF TEAM AWARENESS

Questions	Factors studied	Explanation of the factors	Implication	Awareness
Who is who?	Team members	Knows the team members by their names	T, CT	TE
	Team structure	Knows the composition of the team	T, CT	TE
	Roles	Knows the role of each person in the team	T, CT	TE
	Responsibilities	Knows who is responsible for which work	T, CT	TE
	Task dependency	Knows who has a task dependency with his task	T, CT	TA
Who knows what?	Expertise	Knows the expertise of the team members	T, CT	TE
	Team size	Knows the size of the team	T, CT	TE
	Current task	Knows the current task of the team members	T, CT	TA
	Deadline of task	Knows the deadline of the tasks in the team	T, CT	TA
	Schedule	Knows the current status regarding the schedule	T, CT	TA
	Task nature	Knows the nature and complexity of the task	T, CT	TA
	Facing problem	Knows who is facing a problem (technical, development or dependency) regarding the task	T, CT	TA
	Problem solving	Knows who can solve the problem (technical, development or dependency) regarding the task	T, CT	TA
	Current similar task	Knows who is currently working on the same or a similar task	T, CT	TA
	Previous similar task	Knows who has previously worked on the same or a similar task	T, CT	TA
	Dependent tasks	Knows which task has a dependency with his task	T, CT	TA
	Requirements	Knows about the requirement change process	T, CT	TA
	Task importance	Knows the severity of the tasks in the team	T, CT	TA
	Task understanding	Knows the level of understanding about the tasks in the team	T, CT	TA
	Leaving team	Knows who is leaving the team in coming 3 weeks (permanently or going on a mission)	T, CT	PR
	Joining team	Knows who is joining the team in next one week (new or coming back from a mission)	T, CT	PR
	Going on vacations	Knows who is going on vacation in next 6 weeks	T, CT	PR
	Distance	Geographical distance between the teams	CT	PR
What is where?	Supporting documents or work manuals	Knows where the task guidance manuals are placed in the shared space	T, CT	PR
	Presence in office	Knows the momentary presence of the team members	T, CT	PR
	Location in office	Knows the sitting location in the office	T, CT	PR
	Sitting beside	Knows who sits with whom in the team	T, CT	PR
	Teams location	Knows the other team's location	CT	PR
	Seating topology	Knows the seating arrangement/topology in the team	CT	PR
What is what?	Team type	Knows the team type	T, CT	TE
	Communication	Knows the level of communication	T, CT	TE
	Mutual understanding	Knows the level of mutual understanding	T, CT	TE
	Time zone	Knows the time-zone difference	CT	PR
	Holidays in different time zone country	Knows the situation (holidays' duration and availability of people during that)	CT	PR

Implication: T – within the team, CT– cross teams.   *Awareness:* TA – task awareness factor, PR – presence awareness factor, TE – team awareness factor

By using all those factors, we constructed our survey questionnaire by using the Kitchenham guidelines [41]. This includes survey design (cross-sectional), sample selection, question construction, questionnaire evaluation, data analysis etc. After formulating the survey questions we discussed and refined those for the clarity. Then we asked two volunteers to fill the survey form and give feedback on the clarity and understandability of the questions. We iterated the survey questionnaire for refinement several times. Quantitative data gathered through this pilot pointed out a few shortcomings of semantic data analysis, which were updated before the actual survey study.

C. Selection of Participants

As, we were not limited to the specific region any software company working in GSD was a potential candidate for our research. Upon an Internet search, 22 software companies working in GSD were identified, by using stratified random sampling method. We contacted those companies by email, but received no response. Consequent personal visits or phone calls resulted in declines. The reasons varied. Some companies

refused, since they believed their experience is not relevant for the study. Others mentioned the lack of time to participate and/or confidentiality concerns.

Finally, convenience sampling was used to employ the participants. Confirmation of interest was received from 34 teams (17 pairs of teams) in 15 software companies. All companies were involved in distributed development of large-scale software systems.

D. Data Collection

Once the questionnaire was ready and companies were selected survey responses were collected from large-scale software development teams working in different domains of software development via individual, structured face-to-face interviews and a web-based survey. Major contribution (52% of responses) came from the telecommunication sector and approximate duration of each face-to-face inquiry session was 45 minutes. The study was conducted in seven countries worldwide. This helped to explore the level of team awareness between teams ranging from co-located to the ones separated by large geographical distance and multiple time zones.

However, dispersion of the teams in seven countries made it impossible to conduct face-to-face interviews with all participating teams. Thus for the teams which were not available for face-to-face meetings, we sent a web-based link to fill the online survey form. Moreover, we provided them with the technical assistance via phone, Skype and emails to avoid any misinterpretation or confusion while filling the survey. Clarification of the questions raised during the process was also performed by the researchers. Since many awareness elements are time dependent, a five-days time limit was set for those filling the online survey to avoid any inconsistency in the context. Finally, in order to ensure the accuracy of the context, we contacted the participants after the completion of each survey form and further discussed the situation with reference to the context of the other participants of that team. A large amount of written notes was taken during this process.

All interviews that we conducted were audio recorded, except one in which we forgot to turn on the audio recorder, and web based survey responses were stored in the survey application.

In total, data collected accounted 133 responses from 34 teams (17 pairs of teams) representing 15 software companies.

E. Data Analysis

Survey data was stored electronically in spreadsheets. After normalization, we selected 107 responses from 26 teams (13 pairs of teams) representing 12 software companies (see Table III). We rejected data collected from eight teams that did not meet one or several of the following criteria for the study:

- Wrong or incompletely filled survey form,
- Participation within the team less than 50%,
- Late online responses (after the five-days deadline),
- Lack of response from the collaborating team in a pair.

TABLE III. INFORMATION ABOUT PARTICIPANTS

Groups	Pairs of teams	Teams	Local participants	Cross-site participants	Total participants
Group-1	4	8	28	18	46
Group-2	4	8	15	17	32
Group-3	2	4	4	4	8
Group-4	3	6	13	8	21
Total:	**13**	**26**	**60**	**47**	**107**

At this stage, the correct and consistent data was converted to a presentable form to be analyzed – both statistically and semantically – to find out the level of team awareness for all pairs of teams. As described in the study setup, we used the individual's data acquired through the first step of the survey as the benchmark for semantic analysis. For instance, personal data provided by each participant about his/her task, presence, vacation plans etc. was considered as the baseline information about that particular team member. Comparison of the personal information with how other team members (within the team and across the team) responded regarding his/her task, presence and vacation plans, helped us to judge whether the other team members are aware of these events.

We assigned the weights and calculated the percentages to measure the level of awareness for each individual. Weight is assigned to keep the balance between the newly hired experienced people and those already working but having less experience. Then, we aggregated it to the team level in order to get the level of team awareness for the whole team.

Finally, we converted the results into logarithmic form to make them more presentable.

IV. RESULTS

In response to our research questions we first present the results regarding the levels of awareness within the team, and then results pertaining to the levels of awareness in cross-team collaborations of varying distances.

A. Awareness within the team

Based on the general perceptions prevalent in software engineering research, we expected to see a very high level of awareness within the team, in other words, in all co-located teams participating in the survey. Our results show that on average, awareness level was measured only as high (see Table IV). Furthermore, average value of presence awareness was medium, which is much lower than initial expectations. Interestingly, when dividing the analysis into groups (can be seen further in Table VI) we find that only teams under Group-2 on average meet the expected level of team awareness, which is very high.

In order to understand why teams under Group-2 achieve the very high level of team awareness, while teams under Group-1, Group-3 and Group-4 fail, we performed a stratified evaluation of the demographic attributes and found that previous history and experience of working together, team size, team type and the level of interaction among the team members might have influenced the results.

B. Cross-Team Awareness

A general perception is that the distance and awareness are inversely related to each other. It means that the greater the distance between the collaborative teams, the lower the level of awareness can be expected. Focus/nimbus model [20, 21] dealing with the large spaces to disseminate awareness information in virtual collaborative environments also suggests the same. This implies that, as we will move from Group-1 to Group-4, the expected level of awareness across collaborating teams will gradually decrease (See Table V).

Similar to the pattern of awareness within the team, the level of awareness for cross-site teams is higher for Group-2 (determined as medium) as compared to Group-1, Group-3 and Group-4, which scored low. We also found that none of the groups meets the expected level of team awareness for cross-site collaboration. Teams under Group-1 and Group-4 were the expected extremes (i.e. very high and very low respectively). However, our results indicate that teams under Group-1 and Group-4 have the same low level of awareness.

To understand which factors elevated the level of team awareness for Group-2 (medium) and why the growing distance did not determine the decreasing level of awareness,

we looked into the demographic attributes and performed another level of analysis by strata. We found that previous history and experience of working together, team size, team type and the level of interaction among team members might have played a more significant role than the geographic separation. These factors are further discussed in the following sub-section.

TABLE IV. LEVEL OF WITHIN THE TEAM AWARENESS

Team type	Expected level of awareness	Level of awareness found			
		Team Awareness factor	Task Awareness factor	Presence Awareness factor	Team awareness
Within the team for all groups	Very high	High	High	Medium	High

TABLE V. LEVEL OF CROSS-TEAM AWARENESS

Team type	Expected level of awareness	Level of awareness found			
		Team Awareness factor	Task Awareness factor	Presence Awareness factor	Team awareness
Group-1	Very high	Low	Low	Low	Low
Group-2	↓	Medium	Low	Medium	Medium
Group-3		Low	Low	Medium	Low
Group-4	Very low	Very low	Low	Very low	Low

Group 1: Same building, same floor, distance less than 30 meters
Group 2: Same building, different floor or the same floor, but distance more than 30 meters
Group 3: Different locations, but the same time zone
Group 4: Different locations in different time zones

C. Factors affecting team awareness

In order to find the explanations for our results, we analyzed the awareness factors from Table II for a possible positive or negative association with the level of team awareness. Moreover, the demographic information of the teams was also considered to know the circumstances in more detail. We consulted the data and through stratified analysis of the demographic data we identified factors that might have influenced the level of team awareness in our study (see Table VI). We discuss them in detail in the following sub-sections for both within the team and across the team.

1) Experience: We found that the lack of experience of the team members affected the level of task awareness, presence awareness and also the factors of team awareness, which eventually lowered the overall team awareness of Group-1 for both within and across the teams. We analyzed the different facets of individual experience and concluded that team awareness for pairs under Group-1 were affected by the following factors:

a) Lack of prior work experience (fresh graduates),

b) Lack of prior company experience (new hires),

c) Lack of prior team experience (new team member).

Our findings were further supported when we studied the reasons of comparatively better level of team awareness in Group-2 for both within and across the teams. Further, a better level of team awareness for cross-site teams in Group- 4 also suggests that experience may have positively affected the level of team awareness. It shows that the effect of experience is stronger even for the teams, which are distributed across the geographical, temporal and socio-cultural boundaries.

We also observed that developers with higher experience of working together in the same team and in the relevant domain are more likely to support and maintain the task awareness and presence awareness, which consequently increases the level of team awareness. This finding also extends prior research on team cognition, which suggests that team members with similar task experience by working on similar projects have more organized shared knowledge [18], and compliments research advocating the recency of projects as a factor affecting awareness [42].

In order to consolidate our confidence on the results even further, we considered the team members with less experience as outliers and removed them from analysis but still we found no considerable change in our results. We argue that this is because, the lack of experience of the team members is not the only factor, which decreases the level of team awareness. Demographic data analysis revealed that other factors such as team size, team type and level of cross-site interaction might have also contributed to the level of team awareness, as described in continuation.

2) Team size: Another important factor that potentially impacted the team awareness negatively was the team size. We found that teams under Group-1 are larger in size (for both onsite and offsite teams) as compared to the other groups (see Table VI). The teams under Group-1 were almost double in size than the teams in the other groups. Our observations also support the existing research, which advocates that the increase in the size of team exponentially increases the dependent links within the team and brings substantial project management and coordination overhead [16]. Consequently, it results in a decrease in the level of team awareness by negatively influencing the effective comprehension of the task and presence awareness.

The main reason for a comparatively better team awareness under Group-3 (for both within and across the teams) was the smaller size of the teams in that group, which, we assume, made it possible to maintain the necessary level of awareness in coordination of their task activities across different locations. Other than Group-3, we also found that the

teams with smaller size possessed higher level of team awareness as compared to the larger teams (see Table VI).

3) Team type: Team type emerged as another candidate impediment in achieving the high level of team awareness. Research shows that the agile teams spend the major part of their working hours in discussions and learning about each other's work activities and resolving problems [43]. It supports development of a common vision about the project and alignment of team member activities. These processes have an obviously positive influence on task and presence awareness.

Analysis of the demographic data shows that teams under Group-1 expected to demonstrate very high level of team awareness both within and across the teams (as they are situated in the same building). Our findings revealed that Group-1 did not meet the required level of team awareness and this effect was stronger for cross-site teams. We found that three out of four pairs of teams under Group-1 did not follow agile methods. Thus we believe that this might have been the reason for low level of task and presence awareness.

As discussed earlier, Group-3 (for both onsite and offsite teams) was composed of small teams working from the same

time zone but dispersed locations. Both factors (small size and temporal proximity) might mean that team members in this group were well aware of each other. Although the teams in that group were not declared agile, when we contacted project managers they informed us of similar practices of informal discussion each morning followed in these projects. As a result, we observe a comparatively better level of team awareness in Group-3 teams despite their physical distribution across different geographical locations.

4) Interaction: Another potential reason for the low level of team awareness we relate with the lack of cross-site interaction. We derived that the teams, which maintained a high level of cross-site communication, demonstrated a higher level of team awareness. This finding also complements the prior research emphasizing the importance of interaction between the dispersed teams [7].

We also found that teams with more frequent communication were usually either agile or smaller in size. Furthermore, we also found that the members of such teams were comparatively more experienced. So consistent with prior research, our findings also propose that small and agile teams [43, 44] and teams composed of experienced members [18] are more likely to have a higher level of team awareness.

TABLE VI. LOGARITHMIC RESULTS FOR ALL PAIRS OF TEAMS

Team type	Pairs	Task awareness		Presence awareness		Team awareness (onsite)		Team awareness (offsite)		Experience		Team size		Team type	Cross-site interaction
		Within team	Cross-team	Within team	Cross-team	Within team	Cross-team	Within team	Cross-team	Within team	Cross-team	Within team	Cross-team		
Group-1	Pair-01	1.92	1.41	1.67	1.13	1.85	1.41	1.82	1.30	1.00	1.11	0.60	0.48	Not Agile	Medium
	Pair-02	1.90	1.54	1.66	1.38	1.85	1.49	1.82	1.45	0.78	0.90	1.08	0.70	Not Agile	Medium
	Pair-03	1.82	1.36	1.71	1.18	1.84	1.53	1.73	1.36	0.70	0.48	0.78	0.30	Not Agile	Low
	Pair-04	1.95	1.62	1.92	1.75	1.93	1.71	1.94	1.75	1.11	1.04	0.78	0.90	Agile	High
Group-2	Pair-05	1.93	1.57	1.78	1.82	1.88	1.81	1.83	1.70	0.70	1.00	0.30	0.48	Not Agile	Medium
	Pair-06	1.92	1.53	1.90	1.62	1.91	1.69	1.90	1.51	0.85	1.18	0.70	0.60	Not Agile	High
	Pair-07	1.83	1.53	1.79	1.49	1.77	1.49	1.81	1.43	0.90	1.00	0.60	0.48	Agile	Medium
	Pair-08	1.93	1.45	1.94	1.71	1.93	1.66	1.94	1.64	1.15	1.18	0.60	0.85	Agile	High
Group-3	Pair-09	1.93	1.38	1.72	1.76	1.92	1.48	1.86	1.45	1.08	0.85	0.30	0.30	Not Agile	High
	Pair-10	1.88	1.47	1.82	1.41	1.67	1.52	1.92	1.51	0.85	0.95	0.30	0.30	Not Agile	Low
Group-4	Pair-11	1.91	1.54	1.77	0.85	1.90	1.20	1.81	1.26	0.95	0.85	0.90	0.60	Agile	Low
	Pair-12	1.90	1.56	1.69	1.38	1.85	1.48	1.83	1.54	0.90	0.90	0.48	0.30	Not Agile	Medium
	Pair-13	1.81	1.58	1.79	1.45	1.79	1.51	1.80	1.45	0.90	0.90	0.30	0.30	Not Agile	High

V. DISCUSSION

Our study presents the level of team awareness between the teams of varying distance. Related literature suggests that GSD teams face a wide variety of coordination problems as the temporal, socio-cultural and geographic distance increases between them [1, 9, 39]. Such distances also cause the lack of awareness between the teams as people at distant sites share less context and common grounds [10]. Moreover, as the distance grows it augment the problems faced by software development teams by decreasing the awareness even further [20, 21].

It motivated us to divide the distance into slabs of varying sizes and study how it interacts with the awareness on each slab. In order to validate this phenomenon, we tested it on a spectrum of distance (ranging from collocated teams to the teams which are distributed across different time zones and

geographical boundaries) to find the gradual decrease with increase in distance. Our expectations were that within the team, awareness remains similar for all the teams at all the places and it decreases in cross-site collaborations. We developed these expectations because teams at a single location do not have the geographical and temporal distances. Further, even if such teams are multi-cultural and people have different backgrounds, they face very few problems because they sit in the same place and they can easily sort the things out in a face-to-face meeting.

Against the expectations our findings revealed slightly different results (see Table IV and V). We found that the team awareness was not gradually decreasing with the increase of distance and also within the team the level of awareness was not as high as expected, and varied for some of the teams. However, we did find that cross-team awareness was generally lower than awareness within the teams. We further associated

these results with other factors that have influenced awareness beside the distance. These are: prior experience of working together, team size, team type and frequency of interaction.

Although we were eager to see positive results from testing our initial hypotheses, we believe our results have useful implications for practice. In fact, the identified factors can be used to improve distributed development by implementing structures and employing practices that can positively affect team awareness even and especially when teams are separated.

Our results also demonstrate that the levels of task and presence awareness in the same group of separation varied. In particular, pairs of teams in the same group (level of distribution) could have presence awareness relatively higher than task awareness. Interestingly, task awareness in all groups scored low, while presence awareness varied. Following our expectations maintaining presence awareness between temporary separated teams appeared to be challenging, as the presence awareness level in that group scored very low. However, presence awareness for neighbouring teams scored low, which we associated with the large size of teams. For teams separated by more than 30 meters within the same building until separation within the same time zone was medium, which is higher for that of task awareness in the same groups. This might mean that presence awareness information is more important for the team members in order to coordinate the task related knowledge through ad hoc inquiries, whenever the necessity arises. Varying awareness needs have also been observed by de Souza and Redmiles [42], who found that the awareness networks of a software developers are fluid and change during the course of a project. Hence, in the future we suggest to use awareness measurement in combination with the awareness needs, since although high awareness levels alleviate coordination, the opposite does not necessarily have a devastating effect.

A. Study Limitations

Our study has several limitations. First of all, varying number of responses for each group might have influenced the results. However, we believe that convergence of the emerging levels from pairs of teams in the same groups increase the reliability of the results.

Secondly, a more detailed survey could have potentially helped to obtain more subtle details, however it might have also influenced the reliability of responses, since longer questionnaires would bore the respondents. For future work we suggest to look at awareness factors (see Table III) independently with a possibility to extend the questionnaire.

Finally, we believe that our findings are generalizable only for distributed software development teams working on a shared project. While local awareness might be influenced by the variety of factors, as observed in our case, our conclusions regarding the challenges with awareness in cross-site collaborations versus co-located development should be also relevant for others.

VI. CONCLUSIONS

In this study, we have explored the level of team awareness between the teams on a varying distance. Several studies have discussed the effect of the distance on coordination and awareness in software development teams [4, 45] but the cross-site awareness still seems immature. Our result extend existing research by focusing on task awareness and presence awareness separately for teams separated by varying distance.

Our results are based on the findings from surveying 26 teams (13 pairs of teams) in 12 companies worldwide. We considered the pairs of teams working jointly on a shared project and collaborating to perform their task and team activities. Although the common perception and our initial expectation was that distance is the main hindrance between the teams which reduces the team awareness. Our findings showed that 1) task awareness levels differed from presence awareness levels, and 2) frequent interaction within and across the teams, small team size, prior working experience and agile practices fostering interaction all positively influenced the level of awareness.

Although, collocated development is often associated with high levels of awareness while distance is looked at as the enemy, factors identified in our study may help alleviating coordination in distributed teams through task and presence awareness and shall be thus studied in more detail in the future.

REFERENCES

[1] J. D. Herbsleb and D. Moitra, "Global software development," *IEEE Software*, vol. 18, no. 2, pp. 16–20, Apr. 2001.

[2] H. Holmstrom, E. Conchuir, P. Agerfalk, and B. Fitzgerald, "Global software development challenges: A case study on temporal, geographical and socio-cultural distance," Proceedings of the IEEE International Conference on Global Software Engineering, pp. 3–11, 2006.

[3] C. Gutwin, K. Schneider, D. Paquette, and R. Penner, "Supporting group awareness in distributed software development," in *Proceedings of the International Conference on Engineering Human Computer Interaction and Interactive Systems*, Berlin, Heidelberg, 2005, pp. 383–397.

[4] J. Espinosa, S. Slaughter, R. Kraut, and J. Herbsleb, "Team Knowledge and Coordination in Geographically Distributed Software Development," *Journal of Management Information Systems*, vol. 24, no. 1, pp. 135–169, Jul. 2007.

[5] A. Sarma and A. van der Hoek, "Towards Awareness in the Large," in *International Conference on Global Software Engineering*, 2006, pp. 127 –131.

[6] K. Crowston and E. E. Kammerer, "Coordination and collective mind in software requirements development," *IBM Systems Journal*, vol. 37, no. 2, pp. 227 –245, 1998.

[7] R. E. Kraut and L. A. Streeter, "Coordination in software development," *Commun. ACM*, vol. 38, no. 3, pp. 69–81, Mar. 1995.

[8] E. Ó Conchúir, H. Holmström Olsson, P. J. Ågerfalk, and B. Fitzgerald, "Benefits of global software development: exploring the unexplored," *Software Process: Improvement and Practice*, vol. 14, no. 4, pp. 201–212, 2009.

[9] E. Carmel, *Global software teams: collaborating across borders and time zones*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1999.

[10] J. D. Herbsleb, "Global Software Engineering: The Future of Socio-technical Coordination," in *Future of Software Engineering*, Washington, DC, USA, 2007, pp. 188–198.

[11] P. Dourish and V. Bellotti, "Awareness and coordination in shared workspaces," in *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, NY, USA, 1992, pp. 107–114.

[12] J. A. Cannon-Bowers and E. Salas, "Reflections on shared cognition," *Journal of Organizational Behavior*, vol. 22, no. 2, pp. 195–202, 2001.

[13] K. E. Weick and K. H. Roberts, "Collective Mind in Organizations: Heedful Interrelating on Flight Decks," *Administrative Science Quarterly*, vol. 38, no. 3, p. 357, Sep. 1993.

[14] D. M. Wegner, "A Computer Network Model of Human Transactive Memory," *Social Cognition*, vol. 13, no. 3, pp. 319–339, Sep. 1995.

[15] M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, Mar. 1995.

[16] J. A. Espinosa, S. A. Slaughter, R. E. Kraut, and J. D. Herbsleb, "Familiarity, Complexity, and Team Performance in Geographically Distributed Software Development," *Organization Science*, vol. 18, no. 4, pp. 613–630, Jul. 2007.

[17] R. S. Huckman, B. R. Staats, and D. M. Upton, "Team Familiarity, Role Experience, and Performance: Evidence from Indian Software Services," *Management Science*, vol. 55, no. 1, pp. 85–100, Jan. 2009.

[18] J. Espinosa, R. Kraut, S. Slaughter, J. Lerch, J. Herbsleb, and A. Mockus, "Shared Mental Models, Familiarity and Coordination: A Mulit-Method Study of Distributed Software Teams," *Human-Computer Interaction Institute*, Jan. 2001.

[19] G. M. Wittenbaum and G. Stasser, "Management of information in small groups," in *What's social about social cognition? Research on socially shared cognition in small groups*, J. L. Nye and A. M. Brower, Eds. Thousand Oaks, CA, US: Sage Publications, Inc, 1996, pp. 3–28.

[20] S. Benford, J. Bowers, L. E. Fahlén, C. Greenhalgh, and D. Snowdon, "User embodiment in collaborative virtual environments," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 1995, pp. 242–249.

[21] T. Rodden, "Populating the application: a model of awareness for cooperative applications," in *Proceedings of the ACM conference on Computer supported cooperative work*, NY, USA, 1996, pp. 87–96.

[22] C. Gutwin and S. Greenberg, "A Descriptive Framework of Workspace Awareness for Real-Time Groupware," Computer Supported Cooperative Work (CSCW), vol. 11, no. 3–4, pp. 411–446, Sep. 2002.

[23] J. W. Creswell, Research design: qualitative, quantitative, and mixed methods approaches, 2nd ed. London: Sage, 2003.

[24] A. J. Espinosa, J. F. Lerch, and R. E. Kraut, "Explicit versus implicit coordination mechanisms and task dependencies: One size does not fit all," in Team cognition: Understanding the factors that drive process and performance, pp. 107–129, APA, 2004.

[25] C. R. B. de Souza and D. F. Redmiles. "An empirical study of software developers' management of dependencies and changes". In: Proceedings of the International Conference on Software Engineering. ACM, New York, NY, USA, 2008, pp. 241-250.

[26] G.-J. de Vreede, N. Jones, and R. J. S. Mgaya, "Exploring the Application and Acceptance of Group Support Systems in Africa.," J. of Management Information Systems, vol. 15, no. 3, pp. 197–220, 1999

[27] W. Appelt, "WWW Based Collaboration with the BSCW System," in *SOFSEM: Theory and Practice of Informatics*, J. Pavelka, G. Tel, and M. Bartošek, Eds. Springer Berlin Heidelberg, 1999, pp. 66–78.

[28] B. Berliner and others, "CVS II: Parallelizing software development," in *Proceedings of the USENIX Winter 1990 Technical Conference*, 1990, vol. 341, p. 352.

[29] B. Magnusson and U. Asklund, "Fine grained version control of configurations in COOP/Orm," *Software Configuration Management*, pp. 31–48, 1996.

[30] L. T. Cheng, C. R. B. de Souza, S. Hupfer, J. Patterson, and S. Ross, "Building collaboration into IDEs," *Queue*, vol. 1, no. 9, p. 40, 2003.

[31] C. O'Reilly, P. Morrow, and D. Bustard, "Improving conflict detection in optimistic concurrency control models," in *Software Configuration Management*, vol. 2649, B. Westfechtel and A. VanderHoek, Eds. Berlin: Springer-Verlag Berlin, 2003, pp. 191-205.

[32] A. Sarma, Z. Noroozi, and A. Van Der Hoek, "Palantír: raising awareness among configuration management workspaces," in *Software Engineering, 2003. Proceedings. 25th International Conference on*, 2003, pp. 444–454.

[33] I. Steinmacher, A. P. Chaves, and M. A. Gerosa, "Awareness Support in Global Software Development: A Systematic Review Based on the 3C Collaboration Model," in *Collaboration and Technology*, G. Kolfschoten, T. Herrmann, and S. Lukosch, Eds. Springer Berlin Heidelberg, 2010, pp. 185–201.

[34] M. Cataldo, M. Bass, J. D. Herbsleb, and L. Bass, "On Coordination Mechanisms in Global Software Development," in *IEEE International Conference on Global Software Engineering*, 2007, pp. 71 –80.

[35] D. Damian, L. Izquierdo, J. Singer, and I. Kwan, "Awareness in the Wild: Why Communication Breakdowns Occur," in *IEEE International Conference on Global Software Engineering*, 2007, pp. 81 –90.

[36] L. A. Goodman, "Snowball Sampling," *Ann. Math. Statist.*, vol. 32, no. 1, pp. 148–170, Mar. 1961.

[37] A. Begel, N. Nagappan, C. Poile, and L. Layman. "Coordination in large-scale software teams". In: Proceedings of the ICSE Workshop on Cooperative and Human Aspects on Software Engineering (CHASE '09). IEEE Computer Society, Washington, DC, USA, 2009, pp. 1-7.

[38] J. D. Herbsleb and R. E. Grinter, "Architectures, coordination, and distance: Conway's law and beyond.," IEEE Software, vol. 16, no. 5, pp. 63–70, 1999.

[39] G. Olson and J. Olson, "Distance matters," Human-Computer Interaction, vol. 15, no. 2/3, pp. 139–178, 2000.

[40] V. Braun and V. Clarke, "Using thematic analysis in psychology," Qualitative Research in Psychology, vol. 3, pp. 77–101, Jan. 2006.

[41] Kitchenham, B., Pfleeger, S.L., 2001–2002. Principles of Survey Research, Parts 1 to 6. Software Engineering Notes.

[42] C.R.B. de Souza and D.F. Redmiles, "The Awareness Network, To Whom Should I Display My Actions? And, Whose Actions Should I Monitor?," *IEEE Transactions on Software Engineering*, 2011, 37(3), pp.325-340.

[43] V. G. Stray, N. B. Moe, and A. Aurum, "Investigating Daily Team Meetings in Agile Software Projects," in *EUROMICRO Conference on Software Engineering and Advanced Applications*, 2012, pp. 274 –281.

[44] E. Bradner, G. Mark, and T. D. Hertel, "Effects of team size on participation, awareness, and technology choice in geographically distributed teams," in *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, 2003, p. 10 pp.

[45] J. D. Herbsleb, A. Mockus, T. A. Finholt, and R. E. Grinter, "An empirical study of global software development: distance and speed," in *Proceedings of the 23rd International Conference on Software Engineering*, Washington, DC, USA, 2001, pp. 81–90