This is a draft of the paper "*Awareness support in Global Software Development: A Systematic Review Based on the 3C Collaboration Model*" published at CRIWG 2010 by Springer LNCS series. This is a preliminary version and will differ from the final version. **The original publication is available at www.springerlink.com**
(DOI: http://dx.doi.org/10.1007/978-3-642-15714-1_15)

Please reference the published paper:
Steinmacher, I.; Chaves, A. & Gerosa, M. (2010), Awareness Support in Global Software Development: A Systematic Review Based on the 3C Collaboration Model, *in* Gwendolyn Kolfschoten; Thomas Herrmann & Stephan Lukosch, ed., 'Collaboration and Technology 16th CRIWG Conference on Collaboration and Technology, 2010', Springer Berlin / Heidelberg, pp. 185-201.

Awareness support in Global Software Development: A Systematic Review

Abstract. Global Software Development (GSD) has gained popularity lately and challenges and research opportunities appeared. The development team's physical distribution imposes challenges related to awareness support during collaboration. According to the 3C collaboration model, awareness intermediates communication, cooperation, and coordination. In this paper, we present a systematic review of the literature that describes awareness studies aiming to improve GSD scenario, identifying which of the dimensions of the 3C model are supported by these studies. Results indicate that coordination is far the most explored dimension of 3C, while awareness supporting communication is very little studied and also show a very high number of new tools introduced in GSD domain.

Key words: Awareness; 3C Collaboration Model; Communication; Coordination; Cooperation; Global Software Development

1 Introduction

Computer Supporting Cooperative Work (CSCW) aims to support collaborative work using computer technologies. Research evolution in this area allowed collaboration among geographically distributed people. Software Development industry has been using the benefits brought by CSCW in order to get competitive advantages in terms of cost, quality, and qualified professionals [1]. This new approach is based on geographically dispersed teams, working collaboratively in a same software project.

However, GSD brings new challenges such as contextual, cultural, organizational, geographical, temporal, and political differences [2]. With the increasing number of organizations seeking for GSD and the issues it brought, researches and related literature also increased [3, 4]. Within these researches, there is a great number of studies related to awareness support in distributed development environments. This occurs because awareness is essential when teams are distributed and there is a need to collaborate in order to achieve a common goal.

To provide collaboration within a distributed environment, it is necessary to put additional effort on coordination of dispersed members. Without this coordination, the effort spent on communication will be lost during cooperation. Coordination treats interpersonal conflicts that affect cooperation [5]. 3C collaboration model summarizes the three dimensions of collaboration: coordination, cooperation, and communication.

In this paper we report a systematic literature review on awareness support within GSD scenario. Its purpose was to identify awareness studies that brought improvements so far and how these studies are supporting communication, coordination, and cooperation in GSD. The review also aimed to identify aspects upon which researchers have focused more intently until now, thus allowing analysis and identification of current challenges and opportunities for future works.

This paper is organized as following: in Sect. 2 we present the concepts related to awareness and the 3C model; in Sect. 3 we present the systematic review, including its planning, conduction and analysis; Sect. 4 aims on classifying and summarizing the improvements and the opportunities identified on this review; and, Sect. 5 discusses our findings and the limitations of our review.

2 Awareness and the 3C Collaboration Model in GSD environments

Awareness intermediates the 3C dimensions of collaboration [5]. This section discuss how these relations were explored for this study in GSD environments.

2.1 Awareness in collaborative environments

Awareness was defined by Dourish and Belloti [6] as "an understanding of the activities of others, which provides a context for one's own activities." Its objective is to allow a group of people working collaboratively to realize how and which of their contributions are relevant to the group activities.

Awareness is concerned to support activities that involve two or more individuals, resources or services, voluntarily or involuntarily involved in any collaborative activity. On GSD environments, collaborative teams are geographically dispersed. Thus, physical, temporal and cultural distances highlight the difficulties of providing awareness. The participants of collaborative work often do not know other participants in personal, work in different timezones, do not speak the same language and do not share the same culture. These, among other factors, hinder the information sharing [7], increase conflicts possibility [8] and inhibits informal interaction [9] among team members.

The problems addressed by GSD affect communication, cooperation and coordination among team members. This occurs due to the relationship among these elements and awareness. This relationship is depicted by 3C model, detailed in the next section.

2.2 The 3C collaboration model

The 3C collaboration model, proposed by Ellis et al.[10] and adapted by Fuks et al.[11]. The model defines collaboration as the union of communication, cooperation, and coordination, as represented by Fig. 1.
According to Fig. 1, communication generates commitments that are managed by coordination. Moreover, during communication people negotiate and
make decisions. Coordination arranges task for cooperation, helps managing conflicts, and organizes activities to prevent loss of communication and of cooperation efforts. Cooperation is the joint operation of members of the group in a shared space, seeking to execute tasks, and generate and manipulate cooperation objects. To obtain success, cooperation demands new rounds of communication, generating a cycle that indicates the iterative nature of collaboration. Awareness is the element that intermediates each of the 3Cs, offering feedback to users actions and giving them information about other participants of a collaborative work. More details regarding the 3C collaboration model can be found in [5, 11].

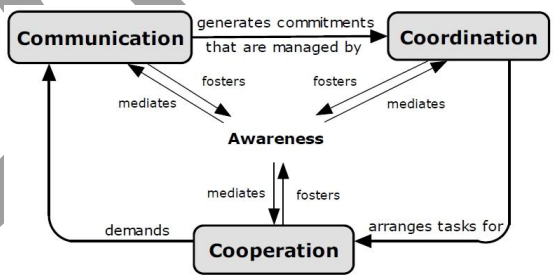

Fig. 1. 3C collaboration model proposed by [10] and adapted by [11]

The relationship among the 3Cs may be used as a guidance to analyze a groupware application domain. A chat, for example, which is a communication tool, requires communication (exchange of messages), coordination (access policies) and cooperation (logging and sharing). Therefore, despite their separation for analysis, there is a constant interplay between them. Fig. 2 (proposed in [12]) presents some applications positioned in the triangle formed by the 3 dimensions.

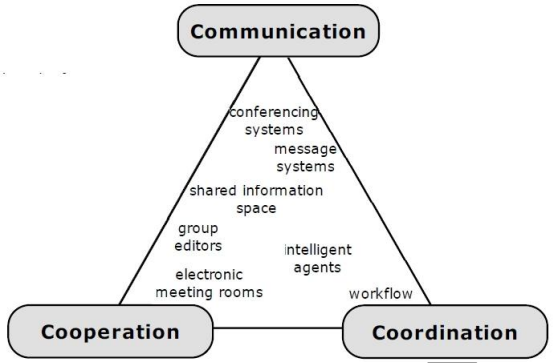

Fig. 2. Applications spread in the triangle formed by the 3C collaboration model [12]

The 3C collaboration model has been often used to classify collaborative tools [12–14]. Organizing collaborative solutions according to this model facilitates the development of collaborative tools since it allows one to resolve problems related to each dimension separately, to compose a complete solution. Moreover, as they are interrelated concepts, deal with them separately may reveal hidden factors if collaboration was treated as an atomic concept.

This paper aims to present awareness studies that improve GSD and classify them according to which of the 3Cs they support. Thus, to make it clear, in this paper we will use the following boundaries to classifying the studies:

- **Communication**: when the study brings improvements to the way messages and information are exchanged among people, reducing gaps, ambiguity and effort needed to understand, establish, and continue a conversation;
- **Coordination**: when the study presented focuses on supporting people to manage themselves by checking and alerting them about the activities, resources and tasks performed by other people that can influence their work;
- **Cooperation**: when the study aims to bring improvements to the shared space or the way users interact with shared artifacts synchronous or asynchronously.

3 Systematic Review

Kitchenham [15] summarizes the stages in a systematic review into three main phases: Planning the Review, Conducting the Review, and Reporting the Review. In this section we present these three steps for our systematic review.

3.1 Review Planning

Review planning includes the identification of the review objective and the development of a protocol. The definition of a review protocol specifies the methods that will be used to undertake a systematic review and aims to reduce the possibility researcher bias [15]. This section summarizes our review protocol.

Research questions: Formulating the research question is the most important activity during protocol definition [15]. The research questions guide the systematic review. In our systematic review the research questions proposed were:
– **Q1**: What are the awareness studies carried out in order to improve GSD scenario?
– **Q2**: Which of the 3Cs are these studies supporting?

Keywords and Query Strings: The keywords were defined based on terms related to GSD; and related to awareness. Table 1 outlines the keywords defined for the questions.

Category C1 has more keywords and reflects the fact that GSD area is maturing, and there are many variations of the same term [3]. The three dimensions of 3C collaboration model (cooperation, coordination and communication) were not included in the query string, because there are studies related to awareness that might not explicitly present one of these words, however they can be classified according to them. The query string was defined as a combination of C1 and C2 using the logical connectors “AND” and “OR”, as presented below:

(Awareness) AND (“Distributed software development” OR “Global software development” OR “Collaborative software development” OR “Global software engineering” OR “Globally distributed work” OR “Collaborative software engineering” OR “Distributed development” OR “Distributed teams” OR “Global software teams” OR “Globally distributed development” OR “Geographically distributed software development” OR “Offshore software development” OR “Dispersed teams” OR “virtual teams”)

Table 1. Keywords defined based on research questions

Reference	Category	Keywords
C1	Global Software Development	“Distributed software development”, “Global software development”, “Collaborative software development”,“Global software engineering”, “Globally distributed work”, “Collaborative software engineering”, “Distributed development”, “Distributed teams”, “Global software teams”, “Globally distributed development”, “Geographically distributed software development”, “Offshore software development”, “Dispersed teams”, “Virtual teams”
C2	Awareness	Awareness

The query string defined was used to retrieve the candidate studies. The following search sources were used to obtain them:

– Science@Direct (http://www.sciencedirect.com);
– El Compendex (http://www.engineeringvillage.com);
– IEEE Digital Library (http://ieeexplore.ieee.org/); and
– ACM Digital Library (http://portal.acm.org).

Studies selection After obtaining the studies by running the query string on the selected sources, papers were analyzed to check their relevance to this systematic review. The analysis was made in order to check if the study deals with awareness on GSD domain, and which dimensions of 3C model are supported by them. It is worth to notice that only studies written in english and with online full paper available were considered.

The process used to include or exclude a study was based on [15, 16, 3] and followed the steps below. The three first steps were performed by two researchers, independently. When at least one of them included a paper as relevant, it was classified as a relevant study. All the steps were reviewed by a third – more experienced – researcher, responsible by checking the information generated.

1. The first analysis was made by reading papers title, excluding those that were considered clearly irrelevant to the research questions.
2. The included studies were then analyzed based on the reading of papers abstract and keywords, considering research questions.
3. Studies included on previous step were further analyzed based on the reading of introduction, conclusion, and specific parts related to the contributions
4. All studies selected so far were read by the researchers and documented to data extraction and quality evaluation on a proper form. Those studies which, despite addressing awareness issues, did not contribute to any significant improvement on GSD domain, were dismissed. We also discarded studies related to a same tool or environment, keeping just the most recent one. Papers included after this step were considered our primary studies.

Studies classification The process of information extraction was based on obtaining information concerning the main contribution of the studies, thus establishing a categorization between objective and subjective results. All papers were categorized based on the classification used in [4]. The categories used on our review are:

(i) case studies;
(ii) theoretical studies (also including conceptual/theoretical frameworks); (iii) experiments;
(iv) tools (also including frameworks and architectures); (v) literature reviews.

Additionally, studies were categorized according to which of 3C dimension is the awareness study presented supporting. This categorization was made by identifying the dimensions supported and weighting them from 0 to 3, according to the level of support the study presents:

3: Mainly supports (main focus of awareness study is on that dimension);
2: Also supports (the dimension is not the main focus, but it is also supported);
1: Indirectly supports (no focus, but brings indirect improvement);
0: Does not support (when no support or improvement is presented).

3.2 Review Conduction

The review was conducted according to the plan presented on previous section. After executing the selection process defined on Section 3.1 a total of 42 primary studies were selected, as illustrated in Table 2.

As it can be observed in Table 2, and was raised by [3], the lack of standard terminology in GSD resulted in a large number of papers to start with, but only a few were selected, resulting in a high sensitivity and low precision search. A high number of papers of unrelated areas (like computer networks, ubiquitous computing, e-learning, and psychology) also contributed to the large number of studies discarded at the beginning. After the title analysis, 143 papers were selected; during second analysis (based on abstract and keywords) 38 papers were dismissed and 105 were further selected; then, after third selection (introduction and conclusion reading), we dismissed other 26 studies. During these three first steps we have discarded 22 more papers repeated or presenting a same tool or environment (in this case only the latest study was considered). Thus, we came to a number of 57 studies selected for an in-depth analysis.

In the deeper analysis, performed by reading the full papers, 15 papers were considered not relevant to the review. The main reason of dismissing papers at this step was that they were not presenting their contributions to the GSD domain area. In the next section, we present the results based on the data collected from this 42 papers selected as primary studies.

Table 2. Included Studies

Source	Papers Found	Excluded per analysis step				Relevant Studies	Primary Studies Selected
		Title	Abstract, Keywords	Introduction, Conclusion	Repeated, Duplicated		
IEEE	37	11	4	6	2	14	10
ACM	325	256	20	9	55	35	26
Science Direct	229	203	9	10	2	5	3
El Com-pendex	86	64	5	1	13	3	3
Total	**677**	**534**	**38**	**26**	**22**	**57**	**42**

3.3 Review Report

Fig. 3 shows the number of relevant studies by year. The first conclusion is that the subject of awareness in GSD evidently is an area which was not widely studied until a few years ago, and that it is only recently that a greater number of publications have appeared. As one can notice, 2007 is the year in which by far the greatest number of studies was published and last three years present the most part of papers of our sample. It is important to highlight that the search was performed early 2010, so it is possible that there can be more studies published in 2009 that were not indexed so far.

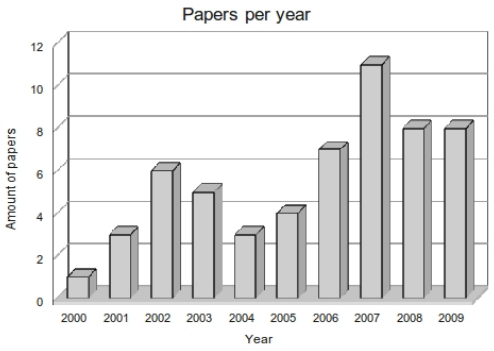

Fig. 3. Amount of relevant studies per year

Fig. 4 presents the distribution of studies according to its categories. All studies were classified in, at least, one category. In this figure, it may be noted the large number of tools introduced. Out of a total of 42 primary studies, 33 studies (79%) are presenting a new tool. It is worth to point out that 21 studies (50%) only present a tool without any experimental analysis of it, presenting a proof of concept or a brief example

on how to use the tool. The other 12 tool related papers present a tool combined with some kind of experimental study. Special attention should be given to one study [17] that presents a tool based on a theoretical study presented there and was further evaluated by an experiment.

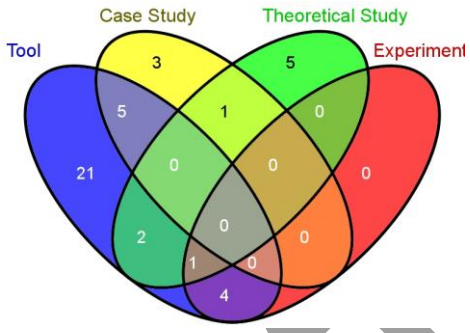

Fig. 4. Venn diagram for types of primary studies found

The other 9 papers (not classified as tool) are categorized as case studies and/or theoretical studies. Five of them are classified only as theoretical studies: three [18–20] present conceptual frameworks for awareness support; one study presents computer support interaction patterns for dispersed team members [21]; and the other one [22] brings an awareness analysis for Open Source communities.

Three papers were classified only as case studies: one [23] studies people work rhythm within a company; and the other two [24, 7] study consequences of awareness gaps in broken code builds and communication, respectively. Just one study [25] was classified as both theoretical and case study. And, as one can notice, for this review no study was classified as literature review.

As the number of studies classified as tools is pretty large, we decided to further analyze these studies. Thus, we decided to check what kind of features or improvements these tools provide, and their support to 3C dimension. This classification will be presented in Section 4.4.

Fig. 5 shows the classification of primary studies regarding 3C collaboration model in two different analysis. The first analysis, depicted by Fig. 5(a), is based on a Venn diagram representing the number of awareness studies that present support (in any level) to each dimension of 3C and their intersections. So, a first and clear conclusion that can be made is that communication is very little studied, presenting just 9 related studies and just 2 focusing only on communication. We can also see that great focus is given to coordination and cooperation, as 40 out of 42 studies (95%) present some support to one of these dimensions, and 21 (50%) support both dimensions concurrently.

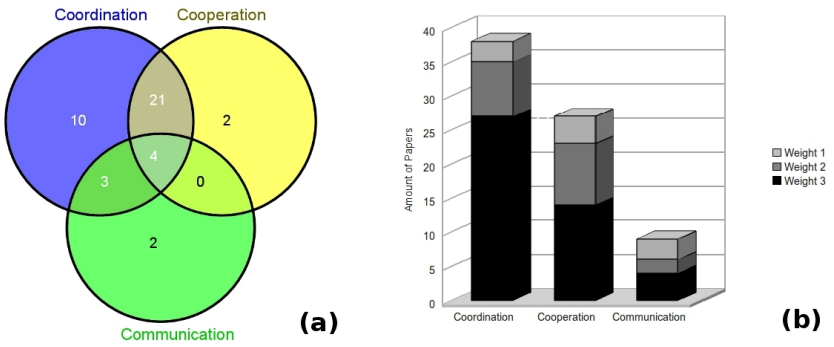

Fig. 5. Venn diagram for types of primary studies found

Fig. 5(b) presents the distribution of studies according to 3C model and to the level of awareness support provided. It is clear that coordination is far the main focus of awareness studies on GSD domain due to the number of studies that mainly supports coordination (weighted with 3 according to the scale presented on Sect. 2.2): a total of 27 studies out of 42 (64%). When we verify studies that support coordination and weighted with 3 or 2, this number grows to 35 (83%). On the other side we have

communication, mainly supported by only 4 studies (9%) and with a total of only 9 studies supporting it.

In summary, Fig. 5 shows that coordination is far the most explored dimension of 3C, while awareness supporting communication seems to be very poorly studied. Some more details about the studies classified according to 3C model dimensions can be found in the next section.

4 Studies discussion and classification

This section aims on classifying and summarizing the improvements proposed and the opportunities identified by performing this systematic review.

4.1 Communication

As it is possible to observe in Figure 5, awareness supporting communication is very little explored within GSD domain. Although many researchers use communication issues as motivation, only 4 awareness initiatives were classified as mainly supporting communication.

Three of these studies focus on providing the users some kind of context in which conversation occurs. Cheng et al. [26] present a tool called Jazz, a collaborative development environment that includes communication facilities, such as chat threading by subject. Jacovi et al.[27] present a tool that allows people to know what are the subjects being discussed on chats. And Fitzpatrick [28] introduces a tickertape tool responsible by bringing CVS commit messages to members of a project, allowing them to start a private or a group chat within the context of the CVS message.

On a study presente by Calefato et al. [9] Jazz environment was improved by presenting its integration with FriendFeed (a social network system), bringing personal interests to work place and offering informal and social communication by using microblogs and forums inside their development environment.

Some studied papers present means to allow communication on GSD environments (e.g. chats and forums), but they are only using existing technologies, without any initiative to make communication easier or more effective. From the viewpoint of software development this is not enough, as, according to [4], the software life cycle requires a great deal of communication using different tools and formats without following communication standards, causing misunderstandings and high response times.

In order to avoid these problems and improve communication, awareness initiatives are needed to avoid ambiguity and misunderstandings, as cultural differences imply different vocabulary which may cause mistakes. Awareness studies focused on generating a common understanding and reducing language ambiguity are desired. Using contextualized information and semantics extraction to improve communication can also be fruitful research areas.

Once communication is highly important and communication tools become more powerful, privacy and security issues regarding access to sensitive user data become critical [29]. So, privacy and security of communication is also a topic that should appear on awareness studies.

4.2 Coordination

Within GSD environment, awareness is seen as a means by which team members can become aware of the work of others that is interdependent with their current tasks, therefore enabling better coordination of teams [24]. This maybe is the justification for 38 out of 42 (90%) awareness studies present some level of support for coordination and 27 (64%) mainly supporting it. Cataldo et al. [20] study the coordination and propose a framework based on product features in order to support coordination within distributed development environments, providing information about members activities and their relations with product features.

In order to provide awareness of coordination, some studies present information regarding the sequence of activities of an ongoing project. One example is the tool called TeamSCOPE [30], that presents a calendar of activities and a log relating artifacts and activities, allowing people to coordinate themselves. Godart et al. [31] introduces the tool ToxicFarm, that offers a workflow view, allowing one to be aware of the activities, their owners and their dependencies. Biuk-Aghai [32] presents visualization approach that aims to support users in gaining a greater understanding of structural and behavioral aspects of virtual collaboration, leading to increased awareness of the activities of the virtual team.

Another approach used to provide awareness to support coordination is gathering information on source code version management repository to make users aware of changes on artifacts that affects their work. Cook et al. [33] presents CAISE, a tool that notifies developers regarding dependent codes, user

dependence and impacts of code changes, based on commits. FASTDash [34] is a visualization tool that seeks to improve activity awareness using a representation of the shared code (extracted from SVN/CVS) that highlights team members' current activities. Many other studies revealed that code repositories are a rich source of information for awareness generation [35, 36, 26, 37, 38, 17].

Also using information from source code version management repository, but studying social network analysis techniques, De Souza et al.[39] introduces Ariadne, a tool that extracts information from code repositories, analyzing sociotechnical dependencies, thus helping to find coordination problems using social network visualization. Tesseract [40] and SmallBlue [41] are other tools that also present socio-technical network analysis as an alternative to improve awareness.

Expertise search is another mechanism studied to provides awareness to support coordination. Expertise browser [42] is a tool to assist users in identifying experts for specific artifacts or tasks, making them aware of how experienced they are and the amount of experts. SmallBlue [41] is an expertise search system that can be used to identify experts, see dynamic profile information and get information about the social distance to the experts; it supports someone to get the right people to work within a given context (task, expertise or area of interest). Minto [43] presents Emergent Expertise Locator that uses emergent team information extracted from source code repositories to propose experts

The studies presented with focus on coordination are bringing improvements, however, the development process must also be adapted to provide the team members with a better awareness of the project status. Most part of the studies presents something based on historical information extracted from source code repositories, but, based on [17] it is also necessary to get recent information, once, according to [34], key information items used to gain awareness are the items that change on a daily, hourly or minute-by-minute basis. Another open research topic is to focus on dynamics of social networks coordination raised by Damian et al. [24], trying to maintain awareness on the emerging and unplanned interactions that should appear during the development cycle.

4.3 Cooperation

Activities within GSD environment requires awareness information to help distributed developers to edit and discuss shared artifacts, reducing negative impacts of distribution. In this sense two studies [44, 45] use code annotation to present changes being concurrently made by other developers in a shared artifact. Dekel and Herbsleb [46] also use code annotation to provide awareness on a tool called eMoose, that allows developers to write informal comments in the code, stores them in a central database and spreads them to other developers using annotated module.

A well explored kind of awareness that supports cooperation is the warning and prevention of interaction conflicts on cooperative artifacts. Lighthouse [38]also provides support to coordination by capturing change events directly on developers workspace to avoid conflicts keeping a shared and up-to-date UML design representation of the actual code. Estublier and Garcia [47] present a study based on cooperative policies to control concurrent engineering in order to avoid conflicts and propose awareness support considering different concurrent models. Palant`ır [8] is another tool that supports cooperation by making users aware of (direct or indirect) conflicts on source code and helping them to come to a solution. Holmes and Walker [45] and Ignat [44] are two other studies that present techniques to avoid conflicts during cooperative handling of code artifacts.

A different awareness study is presented by Everitt et al. [48], that propose Designers' Outpost, a tool that allows users writing on Post-it notes and adding them to an electronic whiteboard, and organize information by physically moving Post-its around on the board; the tool provides synchronous remote cooperation and introduced awareness regarding changes on the electronic whiteboard.

Despite the amount of studies presenting support to cooperation, it is an area that deserves more attention. One example is that there is no study presenting ways to suggest pieces of code to complete a given function or method based on similar codes extracted from other developers' code. Another possible opportunity is to focus on providing awareness support to cooperation in development phases other than coding; for example, present support for clients and analysts cooperate during requirement extraction and specification phases.

4.4 Summary

This subsection aims to summarize and classify the awareness features that the studied tools present to support. This classification was made to (i) further organize the high number of tools found on this review

and (ii) provide a quick reference to GSD environment developers and researchers regarding which awareness features are desirable when designing their environments.

The information can be found on table 3 and presents the references for the studies classified according to which 3C model dimensions they support. The table does not include frameworks or architectures [49–53].

In addition to the analysis of opportunities made on each subsection of this section, and to the Table 3, it is possible to highlight other possible research topics that are not explored on any 3C dimension. Firstly, we have not found studies dealing with awareness to overcome issues related to cultural, political, geographical differences, although this is frequently presented as motivation on GSD studies. Another research opportunity is the definition of policies to provide awareness within a GSD environment, maintaining privacy of team members and organization as well as information security. Another opportunity that was too little explored on all three dimensions is providing awareness not only to coding phase, but also to other software engineering phases.

Table 3. Features to support 3C dimensions

	Coordination	Cooperation	Communication
Conflict indication	[34, 38, 47, 8, 36, 35, 44, 33]	[47, 44]	
Artifact change indication	[31, 30, 34, 26, 38, 37, 36, 45]	[31, 37, 47, 8, 44]	
Activity control (workflow, logs, agenda, worklist)	[31, 30, 54, 28, 32, 35]		
Presence/status indication	[31, 30, 55, 26, 56, 54, 48]		
Context/sub ject-aware message exchange			[30, 26, 27, 54, 28]
Historical log		[30, 17, 47]	
Historic based expert search/recommendation	[17, 43, 42, 41]		
Social/socio-technical network	[39, 40, 35, 41]	[17, 39]	
Source code annotation		[46, 45, 44]	
Collaborative artifact synchronous handling		[38, 54, 48]	
Screen Sharing	[55, 26]	[26, 36]	
Informal/social communication			[9]

5 Conclusion

In this paper we have presented a systematic review of the literature on the use of awareness to support GSD projects. This study was motivated by the absence of systematic reviews about awareness supporting GSD domain, and the lack of studies organizing groupware researches according to 3C collaboration model. The focus of this review was to identify how was the evolution of awareness and what are the emergent challenges related to awareness on GSD. It can be considered a starting point to establish the issues upon which subsequent research will be focused.

We have presented our findings in two phases: in the first phase an initial quantitative data was presented, including number of studies per year, types of primary studies and their classification according to the 3C model; in the second stage, we have analyzed and discussed the data extracted from the primary studies, enabling us to show some conclusions in Sect. 4 about the current state of art and practice of the awareness in GSD domain and the contributions and challenges found for cooperation, coordination and communication.

In general, most part the of the primary studies (79%) focuses introducing a new tool with some awareness support to GSD. The main focus is given to studies gathering information from source code version management repositories, used to provide awareness supporting both coordination and cooperation.

During this review we found a lack of studies and tools giving solutions to awareness providing regarding recent (or real time) context. We did not find any study linking awareness in GSD and ubiquitous computing. Merging these areas should be a nice research topic once awareness is already being discussed in ubiquitous computing during some time[57]. We also did not find any clue on how to use awareness related to the physical location of a team member, regarding, for example, how to treat different cultures, national laws or organization restrictions. The closer study we found was a case study presented by Begole et al. [23], aiming on find ways to coordinate teams in different timezones according to their temporal rhythmic activities patterns on a day-by-day and weekly basis.

In terms of the classification according to 3C collaboration model, the main conclusion is that we have to much support for coordination and very little studies focused on communication. We can say that awareness to support communication is an area that demands more attention and can be a fruitful research topic. Coordination, even appearing as the main focus of awareness within GSD area, is still presenting some opened opportunities like those presented in [17] and [24] and summarized in 4.2.

This systematic review is a initial study in order to define awareness needs for a collaborative GSD environment, trying to balance support for communication, coordination and cooperation.

5.1 Limitations

Systematic review is a powerful method to search for primary studies within a given domain. But as any other method, it also presents some limitations. With the increasing number of studies in this area, this review may have missed some papers that address the use of awareness to support GSD. We also did not perform our search into every possible source. Four digital libraries were selected based on previous studies [4, 3] and on the subject under review. We dismissed some sources because they do not bring full papers (case of Google Scholar and Springer Link). But, we are confident that the sample collected give a good indication of the state of the art and practice of awareness on GSD domain.

The papers included in this review have involved two researchers cross checking each paper for inclusion. A third researcher was responsible for reviewing the information generated after each step. However, the findings of this review may have been affected as the classification is a human process and it is based on some criteria could be considered subjective.

References

1. Robinson, M., Kalakota, R.: Offshore Outsourcing: Business Models, ROI and Best Practices. Mivar Press, GA, USA (2004)

2. Ivcek, M., Galinac, T.: Aspects of quality assurance in global software development organization. In: 31st International Convention MIPRO. (2008) 150–155

3. Prikladnicki, R., Damian, D., Audy, J.L.N.: Patterns of evolution in the practice of distributed software development in wholly owned subsidiaries: A preliminary capability model. In: Proc. of the 2008 IEEE ICGSE, IEEE CS (2008) 99–108

4. Jimenez, M., Piattini, M., Vizcan˜o, A.: Challenges and improvements in distributed software development: A systematic review. Advances in Software Engineering (2009)

5. Fuks, H., Raposo, A., Gerosa, M., Lucena, C.: Applying the 3C model to Groupware Development. IJCIS 14(2-3) (2005) 299

6. Dourish, P., Belloti, V.: Awareness and coordination in shared workspaces. In: ACM CSCW'92, ACM Press (1992) 107–114

7. Chang, K.T., Ehrlich, K.: Out of sight but not out of mind?: Informal networks, communication and media use in global software teams. In: CASCON'07, ACM (2007) 86–97

8. Sarma, A., Redmiles, D., van der Hoek, A.: Empirical evidence of the benefits of workspace awareness in software configuration management. In: ACM SIGSOFT, ACM (2008) 113–123

9. Calefato, F., Gendarmi, D., Lanubile, F.: Embedding social networking information into jazz to foster group awareness within distributed teams. In: SoSEA '09: workshop on Social software engineering and applications, ACM (2009) 23–28

10. Ellis, C., Gibbs, S., Rein, G.: Groupware: some issues and experiences. Communications of the ACM 34(1) (1991)

11. Fuks, H., Raposo, A., Gerosa, M., Pimentel, M., Lucena, C.: The 3c collaboration model. In Kock, N., ed.: The enciclopedia of e-Collaboration, USA, Texas A&M International University (2007) 637–644

12. Borghoff, U.M., Schlichter, J.H.: Computer-Supported Cooperative Work: Introduction to Distributed Applications. Springer-Verlag New York, Inc. (2000)

13. Neale, D.C., Carroll, J.M., Rosson, M.B.: Evaluating computer-supported cooperative work: models and frameworks. In: CSCW '04, ACM (2004) 112–121

14. Muhammad, A., Enr´ıquez, A.M.M., Decouchant, D.: Awareness and coordination for web cooperative authoring. In: AWIC. (2005) 327–333

15. Kitchenham, B.: Procedures for performing systematic reviews. Technical report, Keele University and NICTA (2004)

16. Travassos, G., J., B.: Revisoes sistematicas aplicadas a engenharia de software. In: XXI SBES Brazilian Symposium on Software Engineering. (2007)

17. Omoronyia, I., Ferguson, J., Roper, M., Wood, M.: Using developer activity data to enhance awareness during collaborative software development. CSCW'09 18(5-6) (2009) 509–558

18. Dustdar, S., Gall, H.: Process awareness for distributed software development in virtual teams. In: Euromicro Conference, 2002. Proc. 28th. (2002) 244–250

19. Storey, M.A.D., C˘ ubrani´c, D., German, D.M.: On the use of visualization to support awareness of human activities in software development: a survey and a framework. In: SIGVIS'05, ACM (2005) 193–202

20. Cataldo, M., Herbsleb, J.D.: End-to-end features as meta-entities for enabling coordination in geographically distributed software development. In: Proc. of the ICSE Workshop on Software Development Governance, IEEE CS (2009) 21–26

21. Schu¨mmer, T., Lukosch, S.: The absent participant: more patterns for group awareness. In: PLoP '06, ACM (2006) 1–14

22. Gutwin, C., Penner, R., Schneider, K.: Group awareness in distributed software development. In: CSCW'04, New York, NY, USA, ACM (2004) 72–81

23. Begole, J., Tang, J., Smith, R., Yankelovich, N.: Work rhythms: analyzing visualizations of awareness histories of distributed groups. In: CSCW'02, ACM (2002) 334–343

24. Damian, D., Izquierdo, L., Singer, J., Kwan, I.: Awareness in the wild: Why communication breakdowns occur. In: ICGSE 2007. (2007) 81–90

25. Cataldo, M., Wagstrom, P.A., Herbsleb, J.D., Carley, K.M.: Identification of coordination requirements: implications for the design of collaboration and awareness tools, New York, NY, USA, ACM (2006) 353–362

26. Cheng, L., Hupfer, S., Ross, S., Patterson, J.: Jazzing up eclipse with collaborative tools. In: OOPSLA workshop on eclipse technology eXchange, ACM (2003) 45–49

27. Jacovi, M., Soroka, V., Ur, S.: Why do we reachout?: functions of a semi-persistent peer support tool. In: ACM SIGGROUP, ACM (2003) 161–169

28. Fitzpatrick, G., Marshall, P., Phillips, A.: Cvs integration with notification and chat: lightweight software team collaboration. In: CSCW'06, ACM (2006) 49–58

29. Godefroid, P., Herbsleb, J., Jagadeesany, L., Li, D.: Ensuring privacy in presence awareness: an automated verification approach. In: CSCW'00, ACM (2000) 59–68

30. Jang, C.Y., Steinfield, C., Pfaff, B.: Virtual team awareness and groupware support: An evaluation of the teamscope system. IJHCS 56(1) (2002) 109–126

31. Godart, C., Molli, P., Oster, G., Perrin, O., Skaf-Molli, H., Ray, P., Rabhi, F.: The toxicfarm integrated cooperation framework for virtual teams. Distributed and Parallel Databases 15(1) (2004) 67–88

32. Biuk-Aghai, R.: Visualizing structural and behavioural aspects of virtual collaboration. In: Proc. of 10th IEEE WETICE. (2001) 279–284

33. Cook, C., Churcher, N., Irwin, W.: Towards synchronous collaborative software engineering. In: Software Engineering Conference. (2004) 230–239

34. Biehl, J.T., Czerwinski, M., Smith, G., Robertson, G.G.: Fastdash: a visual dashboard for fostering awareness in software teams. In: CHI '07: Proc. of the SIGCHI conference on Human factors in computing systems, ACM (2007) 1313–1322

35. Sarma, A., van der Hoek, A.: Towards awareness in the large. In: Global Software Engineering, 2006. ICGSE '06. International Conference on. (Oct. 2006) 127–131

36. O'Reilly, C., Bustard, D., Morrow, P.: The war room command console: shared visualizations for inclusive team

coordination. In: SOFTVIS'05, ACM (2005) 57–65

37. De Lucia, A., Fasano, F., Oliveto, R., Santonicola, D.: Improving context awareness in subversion through fine grained versioning of java code. In: IWPSE'07, ACM (2007) 110–113

38. da Silva, I., Chen, P., der Westhuizen, C.V., Ripley, R., der Hoek, A.V.: Lighthouse: coordination through emerging design. In: OOPSLA'06, ACM (2006) 11–15

39. de Souza, C.R., Quirk, S., Trainer, E., Redmiles, D.F.: Supporting collaborative software development through the visualization of socio-technical dependencies. In: GROUP'07, ACM (2007) 147–156

40. Sarma, A., Maccherone, L., Wagstrom, P., Herbsleb, J.: Tesseract: Interactive visual exploration of socio-technical relationships in software development. In: ICSE '09, IEEE CS (2009) 23–33

41. Ehrlich, K., Lin, C.Y., Griffiths-Fisher, V.: Searching for experts in the enterprise: combining text and social network analysis. In: GROUP'07, ACM (2007) 117–126

42. Mockus, A., Herbsleb, J.D.: Expertise browser: a quantitative approach to identifying expertise. In: ICSE'02, ACM (2002) 503–512

43. Minto, S., Murphy, G.C.: Recommending emergent teams. In: MSR'07: 4th International Workshop on, IEEE CS (2007)

44. Ignat, C.L.: Annotation of concurrent changes in collaborative software development. In: ICCP International Conference on. (2008) 137–144

45. Holmes, R., Walker, R.J.: Promoting developer-specific awareness. In: CHASE'08, ACM (2008) 61–64

46. Dekel, U., Herbsleb, J.D.: Pushing relevant artifact annotations in collaborative software development. In: CSCW'08, ACM (2008) 1–4

47. Estublier, J., Garcia, S.: Process model and awareness in scm. In: SCM'05: 12th International workshop on, ACM (2005) 59–74

48. Everitt, K.M., Klemmer, S.R., Lee, R., Landay, J.A.: Two worlds apart: bridging the gap between physical and virtual media for distributed design collaboration. In: CHI'03, ACM (2003) 553–560

49. Kirsch-Pinheiro, M., de Lima, J.V., Borges, M.R.S.: A framework for awareness support in groupware systems. Computers in Industry 52(1) (2003) 47–57

50. Mangan, M., Borges, M., Werner, C.: A middleware to increase awareness in distributed software development workspaces. In: WebMedia & LA-Web. (2004) 62–64

51. Nutter, D., Boldyreff, C.: Historical awareness support and its evaluation in collaborative software engineering. In: WETICE 15th Workshop on. (2003) 171–176

52. Farshcian, B.A.: Integrating geographically distributed development teams through increased product awareness. Information Systems 26(3) (2001) 123–141

53. Bouthier, C., Bitcheva, J., Canals, G.: A context-adaptive group awareness tool. In: Saint Workshops 2005. (2005) 238–241

54. Geyer, W., Richter, H., Fuchs, L., Frauenhofer, T., Daijavad, S., Poltrock, S.: A team collaboration space supporting capture and access of virtual meetings. In: GROUP'01, ACM (2001) 188–196

55. Tee, K., Greenberg, S., Gutwin, C.: Artifact awareness through screen sharing for distributed groups. IJHCS 67(9) (2009) 677–702

56. Hansen, K.M., Damm, C.H.: Instant collaboration: using context-aware instant messaging for session management in distributed collaboration tools. In: NordiCHI'02, ACM (2002) 279–282

57. Baldauf, M., Dustdar, S., Rosenberg, F.: A survey on context-aware systems. In: IJAHUC. Volume 2., Inderscience (2007) 263–277

6 Apendix A

Table 4 presents the studies classified as primary and their related information. This table will be published online in case of acceptance. It is available within this paper due to the blind review process.

Table 4. Primary studies selected in the systematic review

Refe-rence	Year	Source	Type of Study	Commu-nication	Coopera-tion	Coordina-tion
[23]	2002	ACM	Case Study	0	0	3
[34]	2007	ACM	Tool, Case Study	0	0	3
[32]	2001	IEEE	Tool	0	0	3
[53]	2005	IEEE	Tool	0	0	3
[9]	2009	ACM	Tool	3	0	0
[25]	2006	ACM	Theoretical Study, Case Study	0	0	3
[20]	2009	ACM	Theoretical Study	0	1	3
[7]	2007	ACM	Case Study	1	0	3
[26]	2003	ACM	Tool	2	1	3
[33]	2004	IEEE	Tool	0	1	3
[24]	2007	IEEE	Case Study	0	0	3
[38]	2006	ACM	Tool	0	3	2
[46]	2008	ACM	Tool	0	3	0
[37]	2007	ACM	Tool	0	2	3
[39]	2007	ACM	Tool	0	1	3
[18]	2002	IEEE	Theoretical Study	0	1	3
[41]	2007	ACM	Tool, Case Study	0	0	3
[47]	2005	ACM	Theoretical Study, Tool	0	3	1
[48]	2003	ACM	Tool	0	3	1
[52]	2001	Science Direct	Tool	0	3	2
[28]	2006	ACM	Tool, Case Study	3	0	2
[54]	2001	ACM	Tool	2	3	2
[31]	2004	El Compendex	Tool	0	1	3
[22]	2004	ACM	Theoretical Study	0	0	3
[56]	2002	ACM	Tool	1	0	3
[45]	2008	ACM	Tool	0	3	2
[44]	2008	IEEE	Tool	0	3	2
[27]	2003	ACM	Tool, Case Study	3	0	0
[30]	2002	El Compendex	Tool, Experiment	1	2	3
[49]	2003	Science Direct	Tool	0	3	2
[50]	2004	IEEE	Tool	0	3	3
[43]	2007	ACM	Tool, Experiment	0	0	3
[42]	2002	IEEE	Tool	0	0	3
[51]	2003	IEEE	Tool	0	3	3
[17]	2009	El Compendex	Theoretical Study, Tool, Experiment	0	3	1
[36]	2005	ACM	Tool, Case Study	0	3	2
[35]	2006	IEEE	Theoretical Study, Tool	0	1	3
[8]	2008	ACM	Tool, Experiment	0	3	2
[40]	2009	ACM	Tool, Experiment	0	1	3
[21]	2006	ACM	Theoretical Study	0	2	3
[19]	2005	ACM	Theoretical Study	1	2	3
[55]	2009	El Compendex	Tool	0	0	3