JrnlID 11412_ArtID 9072_Proof# 1 - 15/08/2009

Computer-Supported Collaborative Learning
DOI 10.1007/s11412-009-9072-x

An ontology engineering approach to the realization of theory-driven group formation

Seiji Isotani · Akiko Inaba · Mitsuru Ikeda · Riichiro Mizoguchi

Received: 2 July 2008 / Accepted: 12 August 2009

Abstract One of the main difficulties during the design of collaborative learning activities is adequate group formation. In any type of collaboration, group formation plays a critical role in the learners' acceptance of group activities, as well as the success of the collaborative learning process. Nevertheless, to propose both an effective and pedagogically sound group formation is a complex issue due to multiple factors that influence group arrangement. The current (and previous) learner's knowledge and skills, the roles and strategies used by learners to interact among themselves, and the teacher's preferences are some examples of factors to be considered while forming groups. To identify which factors are essential (or desired) in effective group formation, a well-structured and formalized representation of collaborative learning processes, supported by a strong pedagogical basis, is desirable. Thus, the main goal of this paper is to present an ontology that works as a framework based on learning theories that facilitate group formation and collaborative learning design. The ontology provides the necessary formalization to represent collaborative learning and its processes, while learning theories provide support in making pedagogical decisions such as gathering learners in groups and planning the scenario where the collaboration will take place. Although the use of learning theories to support collaborative learning is open for criticism, we identify that they provide important information which can be useful in allowing for more effective learning. To validate the usefulness and effectiveness of this approach, we use this ontology to form and run group activities carried out by four instructors and 20 participants. The experiment was utilized as a proof-of-concept and the results suggest that our ontological framework facilitates the effective design of group activities, and can positively affect the performance of individuals during group learning.

S. Isotani (✉) · A. Inaba · R. Mizoguchi
Q1 Department of Knowledge Systems, The Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan
e-mail: isotani@acm.org

M. Ikeda
Q1 Department of Knowledge Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan

Keywords Group formation · Ontological engineering · Collaborative learning design

Introduction

Collaborative learning (CL) has a long history in Education (Stahl et al. 2006). Nevertheless, with the fast development of technologies that enhance collaboration and communication, recently, this approach has been attracting more attention and becoming a popular method used in classrooms, e-learning environments, and enterprises. According to Soller et al. (2005), over the past decade the number of technologies that enable people to learn collaboratively have increased considerably. Although these technologies have stimulated the use of group activities to support learning, many researchers have noted problems with a lack of tools and a more systematic approach (computer-understandable approach) to support pedagogically sound group formation, the adequate design of CL scenarios and the intelligent support for students to collaborate more effectively (Inaba et al. 2000; Strijbos et al. 2004; Harrer et al. 2006; Ounnas et al. 2008).

In CL, group formation plays a critical role that affects the acceptance of group activities and the success of the learning process. Some researchers claim that inadequate group formation has been the main reason for many unsuccessful applications that rely on CL (Fiechtner and Davis 1985; Graf and Bekele 2006). Nevertheless, the work of Wessner and Pfister (2001) shows that only a few CSCL systems provide functionality for group formation. The majority focuses either on techniques for sharing resources, or on improvements of group performance.[1]

In this paper we focus our discussion on the necessity of sophisticated group formation to set roles, goals, and activities for each learner before a CL session starts. To propose effective group formation, it is helpful to have a clear and conveyable understanding of many learning theories and their features. However, it is difficult for users (e.g., instructors) to have such a common understanding. Our approach calls upon techniques of ontological engineering to build ontologies that represent, explicitly and formally, the main concepts of each theory which are obtained by our interpretation of theories from group formation perspectives. We then proposed a method for adequately using those concepts. Such an approach does not intend to neglect the existence of other effective methods for group formation. Instead, our approach can (and should) be used jointly with other approaches to increase the benefits of group learning by offering structured and well-linked information that facilitates pedagogically sound CL sessions in a variety of contexts.

The method for group formation using ontologies proposed in this paper consists of, first, understand students' needs (individual goals) and then select a theory (and also group goals) to form a group and design activities that satisfy the needs of all students within a group. Our hypothesis is that if we know beforehand more about students' needs, it is possible to increase the benefits of collaboration by grouping students who can support one another (win-win approach) and propose more personalized CL activities that help the members of a particular group to achieve their goals as individuals and as a group. The proposed method is the opposite of conventional methods where instructors initially design collaborative activities and then assign real learners to the various roles and groups. To demonstrate the feasibility of our method, we run an experiment as proof-of-concept where instructors designed and deployed group activities to support development of participants.

[1] An improvement of group performance does not guarantee an improvement of learning (Dillenbourg 2002).

Each method (conventional and proposed) has its pros and cons, therefore, depending on the situation, an instructor could opt for one of them. Although in this paper we address mainly the support of our proposed approach, the ontology developed in this work can support instructors to use a more “traditional” approach and decide in which conditions he/she should switch to other methods for group formation.

The structure of this paper intends to, first, introduce the current state of the art of group formation. Next, it gives an overview of our theory-driven group formation concept as it is developed to date. Following, it presents the CL ontology and a method for group formation. Finally, in order to validate the usefulness of this ontology, it presents the results of an experiment performed with four instructors that have used our ontologies to form groups with the intent to sharpen the communication skills of 20 participants in an ill-structured environment.

Related work in group formation

In the literature, gathering learners into learning groups has different names, such as group/team formation or group/team composition. However, the meaning of these terms is basically the same, which is to identify concepts that serve as a basis for forming a more effective group. The term *effective* is explored differently among researchers in the field, but often is used as a synonym for the adequate allocation (and/or optimal sharing) of resources to maximize the chances of learning. These resources can be tangible, such as learning materials and tools to support collaboration, or intangible, such as knowledge and skills to be learned. In the following paragraphs, we will show some related work in group formation which combine two or more resources (parameters) to form groups.

Usually, the allocation of resources is based on decisions regarding learner’s profiles, technologies, and predetermined tasks (CL techniques or CL best practices). The use of learner’s profiles helps instructors adequately deliver content adapted to satisfy the necessities of the group and its members. For example, the work of Alfonseca et al. (2006) shows the benefits of using learning styles to gather students with similar styles in order to adapt the content for groups working in adaptive hypermedia environments. Thus, it is possible to increase the heterogeneity of a group according to gender, culture, expertise, and other variables, without adapting the content for each member of the group. Different approaches using information extracted from the learner’s profile to form groups (e.g., knowledge about the content, personality, attributes, and programming styles) are discussed by Greer et al. (1998), Graf and Bekele (2006), Faria et al. (2006) and Ounnas et al. (2008).

Another interesting approach for group formation is to include inputs from the environment, such as the availability of specific tools and learning materials, or emotional parameters of learners, and form groups considering these restrictions. Muhlenbrock (2005), and Wessner and Pfister (2001) point out that the use of special technologies (e.g., PDAs and ubiquitous sensors) to obtain variables from the environment can provide an additional source of information, helping to identify the context (or collaboration context) where the collaboration will take place, and thus, improving the quality of the grouping.

Finally, one of the most used approaches for group formation is the use of CL techniques (also known as CL best practices). Usually, a thoughtless group formation (e.g., random selection of learners) and non-structured CL activities (e.g., free interaction) result in inequitable participation, off-task behavior, resistance to group work, and learners in the same group working at different paces (Dillenbourg 2002; Barkley et al. 2005). Thus, the use of CL techniques aims to ensure better individual accountability and positive group interdependence. Some of the benefits of CL techniques presented by Barkley et al. (2005) are (a) a better explanation of the activity, thus providing learners with a basic overview of the whole picture of the collaboration process; (b) clarification of macro-objectives of the group task which helps learners to understand the benefits of the activity; and (c) outlining the task procedures and describing more precisely what learners should do and how they should behave (assignment of roles), thereby minimizing confusion during the activities. In this context, the available supporting systems for group formation are tied with one CL technique. For example, the work of Soh et al. (2008) proposes an algorithm for automating group formation based on one of the well-known CL techniques called Jigsaw (Aronson and Patnoe 1997). In another related work, Deibel (2005) describes a method to support group formation based on Jigsaw for computer science in-class group work. Both works show interesting advantages in using group formation to foster learner participation, to promote peer teaching, and to motivate critical thinking.

Although the benefits of the group formation approaches are presented in this section, a critical review made by Strijbos et al. (2004, 2007) and Resta and Laferriere (2007) reveals that there is limited research on this topic which makes the design of groups based only on learner's profiles, technologies, and tasks insufficient for proposing well thought out CL sessions. To fully support CL, group formation methods should consider critical elements that affect learner's interaction while taking care to design specific formations with CL activities that elicit expected interaction processes. Furthermore, the impossibility of justifying either theoretically or pedagogically the selection of participants to compose a group is one of the main weaknesses of the available methods, and a strong reason for teachers' hesitation in deploying systems with group formation capabilities.

To improve previous achievements and fill the gaps presented in the previous paragraph, our work aims to provide theoretical knowledge, extracted from learning theories that support CL (e.g., Cognitive Apprenticeship), which can be understood by both humans and computers and be used to further increase the benefits offered by others approaches. Such knowledge provides the theoretical justifications to form groups and offers the fundamental setting for an effective CL design and the essential conditions to predict the impact of interactions in the learning process.

In this theoretical approach, the term *effective* is used differently from other approaches. It is not concerned explicitly with adequate allocation and/or optimal sharing of resources, although it provides support for it. This term refers to the creation of collaborative scenarios that can be theoretically and/or pedagogically justified. In this case, a group formation is *effective* if it can be justified by one or more learning theories. Therefore, group goals, individual goals, learning strategies, learners' conditions, CL activities, and other variables present in a CL scenario should be in agreement with a specific learning theory to validate the effectiveness of the group formation. In this situation, it is possible to enjoy the benefits that learning theories provide, such as the rationale for the design of CL activities, the possibility to predict educational benefits, and finally, well-succeeded (effective) learning.

As shown in Fig. 1, the theoretical approach can be thought of as a higher-level policy that gives pedagogical foundations and better structure to CL, and eventually, can be used jointly with other higher-level policies such as CL techniques (e.g., JIGSAW) cited in this section. These higher-level policies have a common lower-level policy (bottom of Fig. 1). Thus, the higher-level policies are used together with lower-level policies to improve collaboration. An example of lower-level policy is to augment the heterogeneity of participants in a group.

The possibility of using CL techniques harmoniously together with theoretical approaches is highly desired because each gains benefits from the other. On one hand, the activities described in CL techniques can be supported and better explained through the

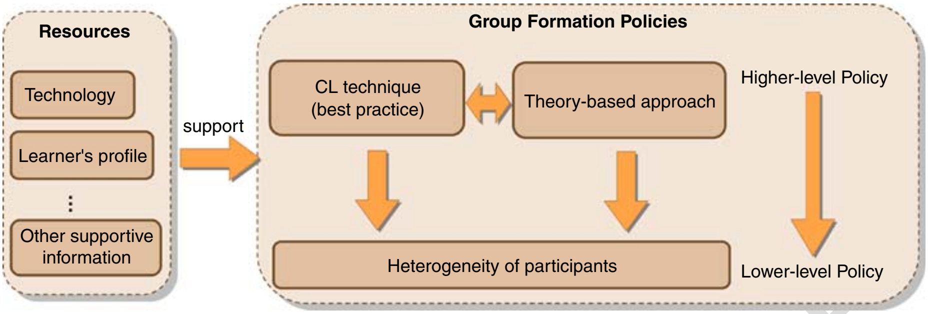


Fig. 1 Proposed theoretical approach that can be used together with other approaches to further increase the benefits of collaborative learning. The "resources" box shows some elements that can be used to support group formation

use of learning theory. On the other hand, descriptions in learning theories can be more easily carried out through the use of concrete activities from CL techniques. The "symbiosis" between CL techniques and a theoretical approach is possible because both have the same goal, which is to create better conditions for learners to learn collaboratively. One of these conditions, often cited by researchers, is the heterogeneity of participants in a group. The heterogeneity can be thought of in terms of different characteristics such as interests, abilities, academic grades, attitudes, knowledge, and others (Graf and Bekele 2006; Resta and Laferriere 2007). The intention of forming groups with heterogeneous participants is what we call lower-level policy. Random selection of learners and learner self-selection are two examples of non-desirable policies commonly used in classrooms to increase heterogeneity. Such approaches provide many unsuccessful collaborative learning sessions (Fiechtner and Davis 1985; Barkley et al. 2005). Thus, to ensure the **adequate** heterogeneity of participants and a better CL session, the research community has been developing a variety of technologies and using different supportive information (left of Fig. 1) obtained from the environment and from learners to increase the use and success of higher-level policies in proposing better CL experiences.

Theory-driven group formation

Many learning theories contribute to in-depth understanding and support of CL (e.g., LPP (Lave and Wenger 1991)). By selecting an adequate theory, we can provide the rationale justifying that the suggested group formation can help learners achieve learning goals. One could disagree that it is possible to support or enhance effective group formation by using learning theories. The authors are aware that theories have flaws and are not "*watertight.*" However, learning theories can provide some essential conditions in which learners are able to learn more effectively. By explaining the learning process, besides trying to explain what happens inside of a learner, a learning theory also gives, either explicitly or implicitly, the ***context*** in which ***learning activities*** have been taking place, the target ***knowledge/skill*** that has been tackled, and the ***roles*** played by learners. An example of such a claim can be observed in sentences quoted from Collins (1991) when explaining the theory of Cognitive Apprenticeship: "*...cognitive apprenticeship employ the paradigm of apprenticeship, but with emphasis on cognitive, rather than physical skills ...*" In this theory, the author applies "*... the notion of learning knowledge and skills in context that reflect the way the knowledge will be useful in real life. It is the sine qua non of apprenticeship; but it should*

JrnlID 11412_ArtID 9072_Proof# 1 - 15/08/2009

be thought of in the most general way. In the context of math skills, they might be taught in contexts ranging from running a bank or shopping in a grocery store to inventing new theorems or finding new proofs."

In these quotes, it is possible to grasp some basic ideas described in Cognitive Apprenticeship theory. First, this theory can somewhat support the development of skills, more precisely, cognitive skills; and second, this theory requires that the context of learning activities should incorporate situations from everyday life, more precisely, situations that are familiar to those who are using the activities and which reflect the real-world uses of the skills.

Another possible point of disagreement is that the use of learning theories to adopt some regulations[2] could harm the CL process. However, according to Dillenbourg (2002) and Strijbos and Fischer (2007), effectiveness of CL relies on how well we understand the multiple factors that influence group interactions and use such understanding to prescribe appropriated learning groups and scenarios that facilitate meaningful interactions among learners. From such an observation, the use of theories as ***guidelines*** can increase the effectiveness of CL.

There are many benefits in deploying learning theories to support CL. However, to select an appropriate theory for a specific situation is a difficult and time-consuming task. One of the reasons is the difficulty in understanding the theories because of their complexity and ambiguity. Each theory has different points of view, levels of aggregation, perspective, and emphasis. Furthermore, they are often written in natural language and there is no common vocabulary to describe their characteristics. This difficulty is well observed by Hayashi et al. (2006) in their work to build a framework[3] to support the adequate use of instructional and learning theories for individual learning. Therefore, to allow the rational use of theories to support CL, we must establish a common conceptual infrastructure on which we can clarify, at least partially, what CL is and how learning theories can facilitate the identification of a well thought out group structure.

Ontologies have shown significant results in representing educational theories and using them effectively (Psyche et al. 2005; Hayashi et al. 2006). In CSCL, one of the pioneering works in using ontologies to establish a system of concepts that models CL, with theoretical support, was presented by Inaba et al. (2000). This ontology is referred to as Collaborative Learning Ontology (CL ontology). Since this initial work, many steps have been taken to improve this ontology and facilitate its use to support the development of ontology-aware systems for CL (Inaba et al. 2003, 2004; Isotani and Mizoguchi 2006). Q2

An analysis of the CL ontology presented in the book written by Devedzic (2006) indicates that it can be quite useful to support CL in Semantic Web-based educational systems by offering: (a) a general framework and vocabulary to describe CL scenarios based on theories; (b) standard vocabulary and knowledge that can be used by pedagogical agents facilitating the communication and negotiation among them; (c) clarification of the behavior and roles for learners; and (d) specification of conditions to be met so intelligent systems can shift from individual learning to collaborative learning in the appropriate time and/or situation, besides assigning adequate roles to each learner based on learners' information. Some interesting examples that show the usefulness of this ontology are presented by Barros et al. (2002), Inaba et al. (2002), and Isotani and Mizoguchi (2007). Nevertheless, previous achievements have some room for improvement. It is especially difficult to propose group formation in compliance with theories. To overcome such a limitation, we have been working to clarify the concepts extracted from theories and to promote the adequate use of these concepts. In the

[2] Such a scheme should be understood as a suggestion to improve the quality of CL and not as imposed rules.

[3] More information about this framework can be found in http://edont.qee.jp/omnibus/

next session, we present some of these concepts and explain how they can be used for effective group formation.

Toward an ontology-aware cscl system with theoretical support

Our work uses ontologies as a common conceptual infrastructure in which learning theories and CL are described explicitly and formally. As discussed previously, we aim to enable theory-driven group formation that offers guiding principles that link the design of CL activities with the analysis of interaction processes. This approach allows us to identify intended goals, roles, and strategies for a group and its members during the design process. Then, we can more easily analyze individuals' and groups' interactions to identify whether the proposed interactions were carried out successfully or not and whether learners attained the expected benefits or not. Finally, with a strong analysis of interactions, it is possible to acquire knowledge about learners and propose a better group formation afterward (Fig. 2).

The framework proposed in Fig. 2 is the ideal flow to offer a better learner-centered and theoretically valid CL session. Usually, related research initiates the design of CL activities before selecting learners or forming groups. This approach facilitates the work of designing general CL activities that can be applied in different situations with different learners. Nevertheless, our approach using ontologies already contains the theoretical knowledge that offers the basis for CL design. Thus, in our framework it is possible to focus on forming groups and then use different information (left of Fig. 1) to design specific CL activities to support CL scenarios which can help a group of learners achieve their goals. Furthermore, each component of this framework (group formation, CL design, interaction analysis, and use of meaning results) can be developed fairly independently of each other. This flexibility comes from the fact that all components share the same ontology, and therefore, can follow

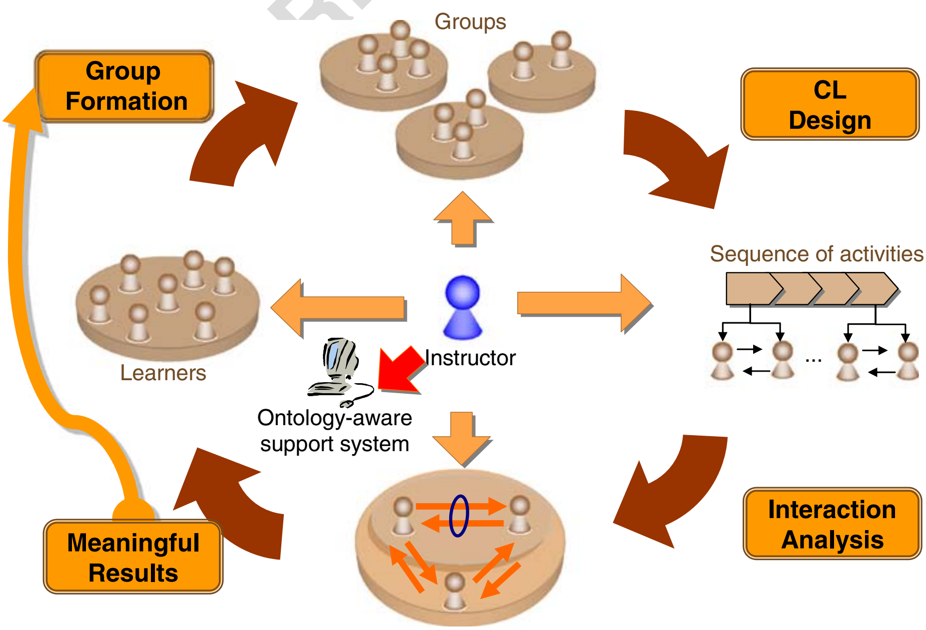


Fig. 2 A full view of the total system of the theory-based group formation and analysis

the same structure of variables, inputs, and outputs. Before the establishment of this framework, previous works in our research group had developed various systems to support CL using ontologies as shown in Inaba et al. (2000, 2002).

Ontology for group formation

To identify the concepts to develop an ontology that supports CL, we need to select appropriate information to propose a principled group formation that creates favorable conditions for learners to perform CL activities and helps instructors to more easily estimate the benefits of a CL session. To accomplish that, we interpreted learning theories from a purpose-specific viewpoint to extract useful information which enables (a) group formation with role assignment, and (b) the specification of interaction flows to facilitate the design of collaborative learning activities. Note that we did not try to do a generic representation of the theories. To prioritize the concepts that should be represented in our ontology, we focused on concepts related to the designing of learning scenarios that have higher impact on changes in the learner's stage of learning. Specifically, we focus on such concepts that allow for a system to answer the following questions based on theoretical support:

– ***What learners can/should participate in the collaboration?***

Some theories require from learners a high degree of knowledge or skill to accomplish some tasks (e.g., distributed cognition). Other theories are specialized to help less-knowledgeable learners. To identify which learners have the potential to get more benefits from a specific theory, our ontology represents the stages of learners' development in terms of knowledge and skills and connect these stages with other concepts such as roles and interactions. More detailed discussion about how we represented knowledge/skills is presented in section "*Main Concepts for Group Formation.*"

- *What goals they have?*

Our ontology intends to be domain independent. Therefore, individual learning goals are represented as changes in the learning stages rather than understanding particular domain concepts. Also group goals are domain independent as well.

- *What roles they play?*

Each theory we analyzed describes roles for learners in a specific CL scenario. Some theories name each role (e.g., Peer Tutoring) and others do not (e.g., Cognitive Flexibility). In our ontology, we named each role and we try to extract from the theories the prerequisites (in terms of knowledge/skills) necessary to play the role and the benefits for playing the role to each player.

- *What tools they can use?*

The learning materials (or learning objects) are especially important to select adequate activities for learners and support CSCL activities. This concept is presented in our ontology and is linked with the interaction processes.

- *What actions/interactions/activities they can/should do?*

One of the main components of CL is the interaction and interaction processes. Each theory proposes different interaction processes to achieve a determined learning goal. Then, our ontology tries to capture such differences and represent them in concepts such as learning strategies and interaction patterns.

Other concepts (e.g., students' behavior and learning styles) were also defined in our CL Ontology. However, this ontology is not complete. Some of the concepts that we still need to represent are: (a) concepts related to learning assessment within a theory, (b) concepts related to the external environment (a CL session can be conducted anywhere), and (c) concepts related to teachers' behavior and strategies to support CL.

Overview of the CL ontology

The CL Ontology is a complex ontology aimed at building a sophisticated system of concepts through a survey of existing learning theories (Inaba et al. 2000). In this initial overview, we explain some concepts of the CL ontology developed to date. In the following subsection, we concentrate on giving more details about three concepts that are essential for group formation.

Collaborative Learning has proven an effective learning method, and sometimes it offers more benefits and advantages than individual learning (Barkley et al. 2005). In CL sessions, learners are encouraged to interact by asking questions, explaining and justifying their opinions, articulating their reasoning, and elaborating and reflecting upon their knowledge, besides many other forms of social interaction (Soller 2001). In fact, educational benefits attained by learners during the CL process depend mainly on interactions. Suthers et al. (2007) emphasize that learning is an "interactional process of change." Learners interact in an attempt to make sense of a situation, and thus, learn (meaning making). In a collaborative environment, learners rely on interactions that are strongly influenced by the characteristics of the learning groups. Therefore, how the gathering of learners takes place is critical to ensure educational benefit. In other words, to attain a learning goal, learners need to interact in a certain way. As we discussed previously, learning theories describe, sometimes implicitly, this *way* of interaction and its expected benefits when performed in an adequate scenario.

The CL ontology offers a framework to describe the concepts extracted from theories that are essential for a successful interaction among learners. To describe these concepts, let us introduce some concepts and their specific terminologies used in the CL ontology:

I: Person in focus.

You or **Y**: Any participant of the group expected to interact with **I.**

I-goal: Individual goal. It represents what a learner is expected to acquire, described as a change of a learner's knowledge/cognitive state. It is good to note that it is not necessary to have a You-goal, because when we focus on the *You* participant, he/she became the new person in focus (*I*).

I-role: Role played by the person in focus.

You-role: Role played by the participant who is interaction with the person in focus.

Y<=I-goal: Learning Strategy. It represents the strategy used by *I* (learner in focus) to interact with *You* (another learner) in order to achieve the *I-goal*.

W(L)-goal: Common learning goal for members of the group (group goal).

W(A)-goal[4]: Goal of the rational arrangement of the group's activity used to achieve the W(L)-goals and I-goals. It characterizes the CL process according to a specific theory.

[4] W(A)-goal: *W* stands for the **W**hole-group and *A* stands for **A**rrangement.

Using this terminology, the CL ontology describes for a specific situation the reason of the interactions among learners in terms of individuals and group goals as shown in Fig. 3(a) (Inaba and Mizoguchi 2004). This figure represents the learning goals of a group with three learners LA, LB, and LC. Each of these learners has an individual goal (*I-goal*) described as *I-goal(LA)*, *I-goal(LB),* and *I-goal(LC)*, respectively. Concerning interactions among learners, from the point of view of LA, he/she will play a role to interact with LB using the strategy *Y<=I-goal(LB<=LA)* in order to attain his/her *I-goal(LA)*. From the point of view of LB he/she will play a role to interact with LA using the strategy *Y<=I-goal(LA<=LB)* to attain his/her *I-goal(LB)*. There also the point of view of LC when he/she interacts with LA or LB, and so on. Besides the representation of individual goals, there are the group goals. The goal of the whole group is represented by *W(L)-goal(LA, LB, LC)* and *W(A)-goal(LA, LB, LC)*. Furthermore, it is useful to represent the goals of a specific cluster of learners who belongs to a bigger group (a small group inside of bigger group). In Fig. 3(a), the group goals of a small group that contains the learners LA and LB as *W(L)-goal (LA, LB)* and *W(A)-goal(LA,LB) are represented* . Figure 3(b) shows a simple example of the instantiation of the presented concepts to describe a group based on two different theories: Cognitive apprenticeship by Collins (1991) and Observational theory (Bandura 1971).

Figure 3 tries to provide a succinct and comprehensive illustration of some concepts included in the CL ontology. Note that *W(A)-goal* cannot be illustrated as a simple sentence as the other concepts, because this concept is the rational composition of other concepts. Besides that, the *Y<=I-goal* is simply a label without much meaning in the figure. The real semantics and relations of the concepts are represented in our ontology as a system of concepts. In the next section, we detail these concepts and how they are organized and represented in the CL ontology to realize theory-driven group formation.

Main concepts for group formation

This section presents three key concepts, extracted from theories, necessary to understand how groups are formed using our ontology: *learning goal* (individual and group goals), *role,* and *instructional-learning event.*

Our working hypothesis for building a comprehensive ontology and defining individual learning goals is that every theory rests somehow on a common basis to explain learning (and instruction). While the assumed mechanism of developing knowledge/skills is different from each paradigm or theory (e.g., behaviorism, cognitivism, and constructivism), the idea of states and stages in the learning process

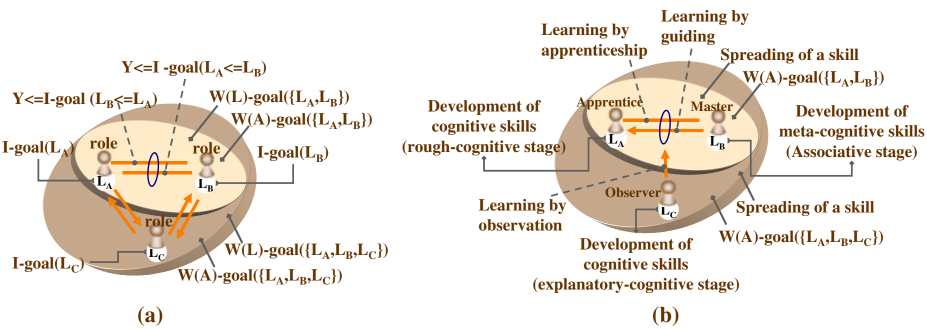


Fig. 3 In **a** we present some concepts and terminologies used in the CL ontology and in **b** an example of the instantiation of these concepts

is common. According to Ertmer and Newby (1993), although instructional/learning theories have unique features and different points of view, they describe the same phenomena of "learning." Thus, it is possible to have an engineering approximation of the states/stages where we can conceptualize "Learning" as changes in the learner's state/stage of development (Hayashi et al. 2006). These changes can occur in an individual learning environment or in more social environment (group learning).

Following such an observation, the authors adopted the theory of knowledge acquisition proposed by Rumelhart and Norman (1978) and the theory of skills development proposed by Anderson (1982) to describe individual learning goals that are domain independent. Both theories are used to give a common background to describe learning as changes in the learner's stage, regardless of whether these changes occur in individual or social environment. According to Rumelhart and Norman (1978), Anderson (1982), and Inaba et al. (2000), although there is a variety of learning goals, the process of a learner's growth can be described in terms of the stages of knowledge acquisition and skill development (Table 1). Thus, concerning individual goals, the CL ontology succinctly describes the learner's knowledge acquisition process and skill development process by adopting the stages and vocabulary used by these theories.

The process of acquiring specific knowledge includes three stages of learning: accretion, tuning, and restructuring (Rumelhart and Norman 1978). **Accretion** is adding and interpreting new information in terms of pre-existing knowledge. **Tuning** is understanding knowledge through its application in a specific situation. **Restructuring** is considering the relationships of acquired knowledge and rebuilding the existing knowledge structure.

Considering the development of skills, there are also three stages of development: the cognitive stage (**rough** and **explanatory**), the associative stage, and the autonomous stage (Anderson 1982). The cognitive stage involves an initial encoding of a target skill that allows the learner to present the desired behavior or, at least, some crude approximation. The **associative** stage is the improvement of the desired skill through practice. In this stage, mistakes presented initially are gradually detected and eliminated. The **autonomous** stage is the gradual and continued improvement of the skill. In this stage, the learner can accurately and quickly perform the desired behavior.

Further, *s(x,y)* is the simplified form that represents the actual stage of the learner: *x* represents the current stage of skill development and *y* represents the current stage of knowledge acquisition. For instance, *s(0,1)* illustrates that the stage of skill development is *nothing* and the stage of knowledge acquisition is *accretion*.

Concerning the description of group goals in the CL ontology (*W(L)-goal*), there are four types: knowledge sharing, creating a solution, spreading of a skill, and knowledge building (or knowledge transmission). These goals were extracted from some of the theories we have

Table 1 Stages of learning development (Inaba et al. 2000)

Individual goals (I-goal)	Stages of development	Abbreviation	Sources
Acquisition of Content-Specific Knowledge	Nothing	s(x, 0), x=0..4	(Rumelhart & Norman, 1978)
	Accretion	s(x, 1) , x=1..4	
	Tuning	s(x, 2), x=1..4	
	Restructuring	s(x, 3), x=1..4	
Development of Skill			(Anderson 1982)
Some Types	Nothing	s(0, y), y=0..3	
- Cognitive Skills - Meta-cognitive Skills - Skill for Self-expression …	Rough-Cognitive	s(1, y), y=0..3	
	Explanatory-Cognitive	s(2, y), y=0..3	
	Associative	s(3, y), y=0..3	
	Autonomous	s(4, y), y=0..3	

analyzed. For example, the Cognitive Flexibility theory supports the sharing of knowledge, and the Cognitive Apprenticeship theory supports the spread of skills.

One of the main factors that affect learners' interactions and, consequently, the achievement of individual/group goals is the ***role*** played by learners. A role provides pedagogical support stating functions, goals, duties, and responsibilities that guide learner's behavior and tend to increase group stability, satisfaction, and communication (Strijbos and Fischer 2007). For example, the role of *"Tutor"* offers educational benefits for a learner who has knowledge about the content, but does not have much experience in using such knowledge. It is because this learner has to explain the content using his or her own words in order to teach and, consequently, obtain a better understanding about it. However, the same role does not bring as much benefit for a learner who already understands the content well and teaches it often. Therefore, we need to know what roles a learner can play in order to support effective group formation. To identify who can play a role and who is appropriate for it, the CL ontology defines the learner's behavior needed to collaborate and two types of prerequisites: necessary conditions and desired conditions. As the names suggest, the necessary conditions are those essential for role play. In other words, if a learner does not fulfill these conditions, he/she cannot play the role, and the desired conditions are those that a learner should satisfy to obtain the full benefits. In other words, if a learner does not fulfill these conditions, he/she can play the role, but the expected educational benefits might not be obtained. Currently, the CL ontology represents 13 roles, their behavior, pre-requisites, and possible benefits for the player, extracted from eight different theories as shown in Table 2 (Inaba and Mizoguchi 2004). In the Column "pre-requisites" the sentences starting with "*" are the necessary conditions for playing a role, and the sentences starting with "-" are the desired conditions to play the role.

To play a role satisfactorily, a learner needs the adequate context. In this research, the context is extracted from each analyzed theory which defines the foundations for effective interaction among learners. From the group formation perspective, this work concentrates on explaining two aspects of such theory-based context: the learning strategy and the CL process, with emphasis on interaction patterns. To express these concepts and their relation with the concepts presented in the previous section, in Fig. 4 we show an updated version of our ontological structure developed to date.

The CL session concept, which is a CL session with theoretical support, consists of two main parts: the learning strategy and the CL process. As we discussed briefly in the previous section, a learning strategy (*Y<=I-goal*) is the form used by a learner to interact with other learners to obtain the desired benefit. Because of that, this concept is intrinsically dependent on the roles played by learners during collaboration and the desired goals of the learner who uses the strategy. Figure 4(a) shows the ontological definition of *Y<=I-goal*. In this figure, *I-role* is played by the main learner (the one who uses the strategy); *You-role* is played by a supporter learner who interacts with the main learner; *I-goal(I)* is an individual goal that can be attained by the main learner through the use of the strategy; and finally, the term *Role Holder*[5] in our ontology refers to a set of learners who can play the specific role in the context determined by *Y<=I-goal*.

Using the structure of Y<=I-goal we can represent, for instance, a configuration of a CL session based on the Cognitive Apprenticeship theory where a learner interacts with other learners to guide them during the resolution of a problem. As shown in Fig. 3(b), from the point of view of the learner who guides, he/she is using the learning strategy (*Y<=I-goal*) called "*learning by guiding*"; his role (*I-role*) is known as the "*master role,*" the role of the

[5] The Role Holder concept is a very deep concept to treat roles adequately in ontologies. Further information about the definition of this concept can be found in (Mizoguchi et al. 2007).

learner who receives the guidance (*You-role*) is known as an "*apprentice role,*" and his/her individual goals (*I-goal*) are to acquire cognitive and metacognitive skills at the autonomous level. Furthermore, using the relation between learners and roles, shown in Fig. 4(b), it is possible to check whether learners have the necessary and desired conditions to play the role, and thus, to identify who has better chances to play the role successfully. In Fig. 5, we show an excerpt of the CL ontology representing part of the configuration of the CL session based on the Cognitive Apprenticeship theory. Note that the points of view of those who are guiding and for those who are apprenticing are represented.

The next concept needing further explanation is the CL process (*W(A)-goal*). This concept specifies the goals of the group activity (*W(L)-goal*) and the rational sequence of interactions (*interaction pattern*) provided by theories that support the achievement of individual and group goals. Previously, we presented the concept of W(L)-goal and its types (knowledge sharing, creating a solution, spreading of a skill, and knowledge building). Following are the details of the concept of interaction patterns.

The essence of CL is the interactions among learners. Recently, the CL community has been putting great efforts to offer support for meaningful interactions. The development of CSCL scripts is one of these efforts. These scripts are guidelines to give structure to CL activities that previously were performed freely producing deficient interactions (Dillenbourg 2002; Miao et al. 2005). To use and to share these scripts adequately, it is necessary to have a common vocabulary. Furthermore, to create a script based on pedagogical/theoretical models, instructors and teachers must be aware of the characteristics of theories/pedagogies and able to represent those characteristics explicitly in terms of the vocabulary.

To support both, common vocabulary and explicit representation of interactions, the CL ontology provides the interaction patterns (Inaba et al. 2002; Isotani and Mizoguchi 2006). These patterns formally describe the flow of the interactions, which specify how group interactions should occur according to a specific learning theory. For example, in Cognitive Apprenticeship theory, initially a master interacts with an apprentice to show the context of a problem; following, the master interacts to demonstrate how to solve the problem in the specific context; and finally, by monitoring and coaching, the master supports the development of the apprentice. This portion of the interaction pattern of Cognitive Apprenticeship theory is justified in the work of Collins (1991), which explains that learning has more chances to occur if it is presented in a specific context, providing in-context examples of correct solutions (or behaviors), and with support of a more knowledgeable partner. It is worth noting that a learning theory does not necessarily have one interaction pattern. Rather, a theory can have many patterns according to different authors which emphasize different aspects of the theory to achieve different (or even the same) educational benefits. In such cases, ontologies provide the common vocabulary and the framework to describe these patterns explicitly and without ambiguity.

In CL ontology, interaction patterns are composed of necessary and complementary interaction activities as shown in Fig. 4(d). The interaction activities are represented by *influential I_L events*, which is the abbreviation for influential instructional-learning event. A similar structure for individual learning was presented by Hayashi et al. (2006). Each *I_L event* is composed of both an instructional event and a learning event. These two events are composed by an actor of an action, the action, and the benefits of this action to the actor (Fig. 4e). The actor of an action is the Role Holder, which means that an actor can only be a learner who plays a specific role in the CL process (e.g., master and apprentice). An actor can act as an instructor (learner doing an instructional action) or as a learner (learner doing a learning action), and through the interaction among actors, the attainment of educational benefits occurs. This formalization in the CL ontology allows explicit representation of the t2.1 **Table 2** Roles, their prerequisites and expected benefits

	Learning theory	Role	Pre-requisite (condition)	Expected effect
t2.3	Anchored Instruction (CTGV 1992)	Anchored instructor	* having the knowledge	• Acquisition of content specific knowledge (tuning)
t2.4			* knowing how to diagnose others	
t2.5			- not having experience in diagnosing others	• Development of cognitive skill (associative stage)
t2.6		Problem holder	* having a problem	• Acquisition of Content Specific Knowledge (tuning)
t2.7			- having the knowledge	
t2.8	Cognitive Apprenticeship (Collins 1991)	Master	* knowing how to use the cognitive skill	• Development of cognitive and/ or meta-cognitive skill (autonomous stage)
t2.9			* having experience in using the cognitive skill	
t2.10			* knowing how to use meta-cognitive skill	
t2.11			* having experience in using the meta-cognitive skill	
t2.12		Apprentice	*Nothing*	• Development of cognitive and/or meta-cognitive skill (cognitive stage & associative stage)
t2.13	Cognitive Flexibility (Spiro et al. 1988)	Panelist	* knowing how to use a skill for self-expression	• Development of skill for self-expression (associative stage)
t2.14			* having his/her own opinion	
t2.15			- not having experience in using the skill for self-expression	
t2.16		Audience	* having the knowledge	• Acquisition of content specific knowledge (restructuring)
t2.17			* having experience in using the knowledge	
t2.18			* having related knowledge in the domain	
t2.19	Distributed Cognition (Salomon 1993)	Full participant	* having the knowledge	• Acquisition of content specific knowledge (restructuring)
t2.20			* having experience in using the knowledge	
t2.21			* having related knowledge in the domain	

t2.22			* knowing how to use the cognitive skill	• Development of cognitive skill (autonomous stage)
t2.23			* having experience in using the cognitive skill	• Development of meta-cognitive skill (autonomous stage)
t2.24			* knowing how to use the meta-cognitive skill	
t2.25			* having experience in using the meta-cognitive skill	
t2.26	LPP (Lave and Wenger 1991)	Full Participant	(Similar as full participant in Distributed Cognition)	(Similar as full participant in Distributed Cognition)
t2.27		Peripheral participant	* knowing how to use the cognitive skill	
t2.28			* knowing how to use the meta-cognitive skill	• Development of cognitive skill (associative stage)
t2.29			- not having experience in using the cognitive skill	• Development of meta-cognitive skill (associative stage)
t2.30			- not having experience in using the meta-cognitive skill	
t2.31	Peer Tutoring (Endlsey 1980)	Peer tutor	* having the target knowledge	• Acquisition of Content Specific Knowledge (tuning)
t2.32			- not having experience in using the knowledge	
t2.33			- misunderstanding the knowledge	
t2.34		Peer tutee	- not having the knowledge	• Acquisition of Content Specific Knowledge (accretion)
t2.35	Observational Theory (Bandura 1971)	Observer	Nothing	depending on what to observe
t2.36	Socio-Cultural theory (Vygotsky 1978)	Diagnoser	* knowing how to use cognitive skill	• Development of cognitive skill (associative stage)
t2.37		Client	* knowing how to use meta-cognitive skill	• Development of meta-cognitive skill (associative stage)

Q3 The sentences in this table are created to be context/domain independent

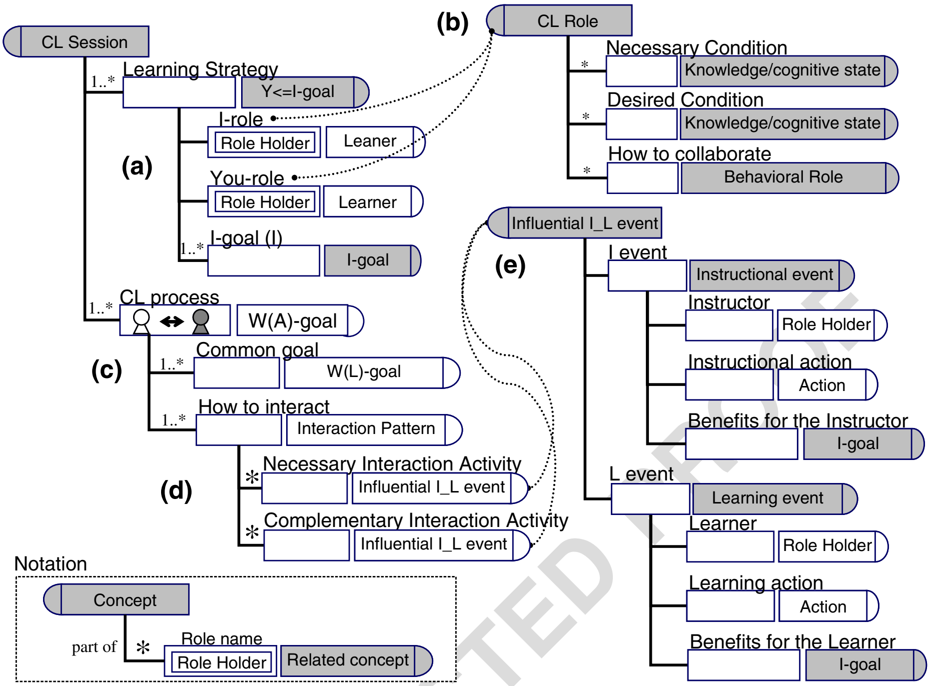


Fig. 4 Part of the Ontological Structure used for group formation

interaction and its benefits from both points of view: for those who do the action and for those who receive the action. Furthermore, it also provides a macro-view of the CL process, in terms of flow of interactions and sequence of activities, and a micro-view of the CL

Q4

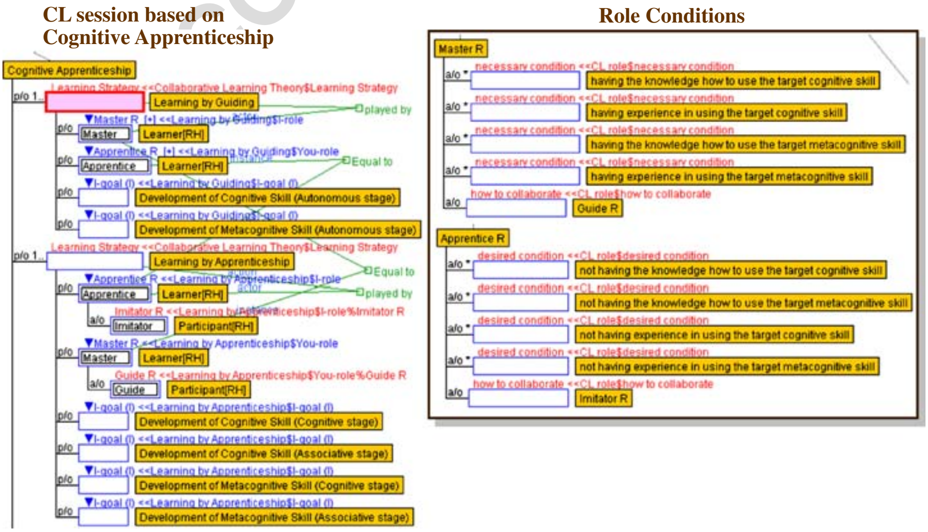


Fig. 5 The use of the CL ontology to represent a CL session based on the Cognitive Apprenticeship Theory. In the figure a/o means *attribute-of* relation and p/o means *part-of* relation

process, in terms of actions and reactions among learners which facilitate the educational benefits of each action.

Note that the *I_L events* are fundamental to link group formation, CL design, and interaction analysis as shown in Fig. 2. Once a group is formed according to a theory, we can use the *I_L events* (interaction patterns) to identify the best sequence of activities for the group following the same theory; and finally, we can analyze the real actions of each learner and compare them with the expected actions defined in the *I_L events*. Thus, if learning does not occur as expected, it is possible to pinpoint the deficient interactions and propose a solution to solve it.

A group formation method

Subsequently, the question becomes how to use the CL ontology developed to date to form groups. First of all, the ontology is used as a common vocabulary to set up the CL session. After that, we use the relationship among concepts to identify the best group formation that satisfies the session requirements.

A conventional method for group formation is, first, select a group goal and a basic structure (based on learning theories, best practices, or CSCL scripts) to design collaborative tasks/scenarios and then assign real learners to the various roles and groups. This practice can be easily used in face-to-face classrooms and it is often used together with CSCL scripts (Kobbe et al. 2007). To support this method, the CL ontology explicitly shows for each analyzed theory the common goal of individuals within a group (group goal). With the linked information presented in the ontology, it is also possible to utilize interaction patterns to help instructors in designing goal-oriented collaborative activities. Finally, if the instructor can gather information of students, he/she can use the ontology to assign roles for each learner in a systematic manner.

Although this method has been used effectively in a variety of contexts, other interesting approaches for group formation which use users' information more efficiently can be sought. Therefore, an alternative method that we want to explore in this paper consists of, first, understand students' needs (individual goals); then select a theory (and also group goals) to form a group; and finally design activities that satisfy the needs of all students in a group. Our main hypothesis is that by having students' information beforehand, we can better understand students' needs. We can thereby increase the benefits of collaboration by grouping students who can support one another and propose more personalized CL activities, which help them to achieve their goals as individuals and as a group.

These two methods complement each other. Therefore, depending on the situation, an instructor can opt for one of them. The first method that will be referred to as *group-individual* orientation method does not require prior knowledge of students and helps instructors to adopt and implement group learning in classroom and e-learning environments. The other method will be referred to as the *individual-group* orientation method. It requires students' information beforehand to adapt group formation and CL activities. With more personalized activities for each group/student, we can facilitate the achievement of individual goals as well as group goals. The CL Ontology can be used to help instructors to decide which method is the best for a specific situation. Table 3 shows some pros and cons of each method.

To propose a group formation using the individual-group orientation method, it is necessary to understand students' needs and then use this information adequately to form groups. Therefore, using the concepts in the CL ontology, we divide the process of group formation in two phases: planning (getting information) and grouping (forming groups).

Table 3 Pros and Cons of two methods for group formation

Group-individual orientation	Individual-group orientation
Pros	Pros
- Does not require prior knowledge of participants.	- Groups and activities are personalized to fulfill the needs of each participant.
- Easy to be adopted and implemented in classroom or e-learning environments.	- Appropriate roles are assigned according to participants' conditions.
- Any learner can join a group and roles can be assigned by participants or teachers.	- Group goals are defined according to its members and not "imposed" based on preconceptions.
- Well-known approach can be used in CSCL scripts.	- Only learners who can potentially contribute to the others (and vice versa) can join a group.
- Easy to apply the same activity for all groups and participants.	- Thanks to the above, a convincing interaction specification appropriate for the learning goal can be specified for each learner and hence, an appropriate group can be formed.
Cons	Cons
- Group formation and role assignment are not adopted to consider the conditions of each learner.	- To adequately assign roles, it requires prior knowledge of participants' behaviors and stage of knowledge/skills.
- Group goals are defined prior to group formation. Therefore, these goals may not be appropriate for all groups and learners.	- Learners who are not suitable to play any role in a specific scenario cannot join the CL process.
- Learners who may harm collaborative learning processes of other learners are not treated adequately.	- CL activities might be different for each formed group requiring the use of semantically enabled environments to track students' interactions within a context.
After collaboration, learners may not achieve their individual goals because collaborative learning activities were not design to support them.	

To set up a CL session (planning phase), first, it is necessary to determine what the target individuals have done in the past (experience) and what they can do now (initial levels of knowledge/skills). In this phase, it is possible to identify, for example, the necessities of individuals and which roles they are able to play. An assessment of the content-worth learning and/or the content needed to be learned should follow. The content should be divided into knowledge to be acquired and skills to be developed. The relationships among knowledge-knowledge, knowledge-skill and skill-skill should also be identified. Finally, select the educational goals expected to be achieved by individuals and/or by the entire group for the specific content. The initial levels of knowledge/skills and the educational goals of each individual should be stated in terms of stages of learning development s(x, y) as indicated in Table 1. A more detailed specification of this process is presented by Isotani and Mizoguchi (2008). Furthermore, each step of the planning phase can be completed, at least partially, by following some instructional design strategies published by many different researchers. Some of these strategies and well-known researchers in the field of instructional design can be found in Romiszowski's book (1981).

By using the CL ontology presented in Fig. 4, the collected information can be used appropriately to form groups (grouping phase). Observe that we have many possibilities to form a group using our ontology. Let us explore one strategy concerning individual goals. In this case, the ontology helps identify conditions where learners can achieve their individual goals by performing CL activities. First, by looking in the I-goal slot (Fig. 4a) of the ontology, we can identify which CL session, supported by a specific theory, can help learners achieve their goals. If we cannot find a session, it means that the theories represented in our ontology cannot help the improvement of the specific learning goal. However, usually there is more than one theory that can help learners achieve their goals. Each theory-based CL session in our ontology provides the settings that the CL activities should conform with. To join a session, a learner needs to satisfy the conditions to play a specific role and to follow a strategy, (Fig. 4b) along with other specific conditions prescribed or described by theories. If a learner does not satisfy the conditions of the session, then he/she cannot get the full benefits prescribed by it, or worse, it could harm the CL process. A session also provides the CL process that clarifies the common goal of a group and the interaction patterns (sequence of CL activities) that can be followed by learners to obtain the desired individual and group goals (Fig. 4c). In previous works, we have shown how to design CL activities using this ontology. A simple pseudo-algorithm to exemplify the use of the CL ontology to form groups considering only individual goals is shown in Table 4. The main goal of this pseudo-algorithm is to use the individual goals available in the learner profiles to find a set of learners that does not violate any necessary condition described in the CL ontology. This means that we try to divide a given set of learners into several groups obtaining a portion of the learners that satisfies a set of conditions. This portion does not necessarily cover all learners, but instead, creates groups where all learners in a given group can attain their individual goals, and the conditions of the groups (e.g., roles) are in agreement with a specific theory. This pseudo-algorithm is just one simple alternative for using the information contained in the CL ontology. More Algorithm using agent technologies, Web services, and other new technologies can be applied to provide a better use of this ontology.

In summary, the pseudo-algorithm showed in Table 4 can be described in narrative as follows:

- ***Planning Phase—Set up a CL session:***

1.1. To determine what the target individuals have done in the past (experience) and what they can do now (initial levels of knowledge/skills). This step aims to identify the needs of individuals and the roles they are able to play.
1.2. Assess the content worth learning and/or the content which needs to be learned. The content should be divided into: knowledge to be acquired, and skills to be developed. The relationships between knowledge-knowledge, knowledge-skill, and skill-skill should also be identified.
1.3. Elect the learning goals which are expected to be achieved by individuals, and/or by the entire group for the specific content.
1.4. State the initial levels of knowledge/skills and the learning goals of each individual in terms of the stages of learning development *s(x,y)* as indicated in Table 1. Each step described above can be completed (at least partially) by following certain instructional design strategies. Some of them can be found in the work of Romiszowski (1981).

- ***Grouping Phase—Forming the Groups:***

There are many possibilities when forming a group. Let us explore one way related to individual goals.

2.1. Match the individuals' goals with a CL session by looking at the *I-goal*. If no match is found, it means that the theories represented in our ontology cannot help to

JrnlID 11412_ArtID 9072_Proof# 1 - 15/08/2009

t4.1 **Table 4** A pseudo-algorithm for group formation considering only individual goals

```
//Goal: use individual goals in the learner profiles to form groups that do not violate any necessary condition
//described in the CL ontology

Setup_CLSession(); //create a list of learners and
                   //initiate the learner profiles (e.g., identifying individual goals and learners' conditions)
                   //setup the environment (e.g., content to be learned, materials available, map the domain
                   //content in the CL ontology, etc.)

Effective_Groups_for_I-goals(learner L, CL_ontology CLO)
  //given a list of learners form groups based of theories that satisfies the individual goals
  //all information of the learner profiles should be accessible from the variable L
  For each learner L do
    For each theory T in the CLontology do
      For each Strategy ST in T do
        L.I-goal    //Individual goal of the learner L obtained from the leaner profile
        ST.I-goal   //Benefits for an individual when using the strategy ST
        If L.I-goal = ST.I-goal then //match individual goals with strategies that help to achieve them
          L.conditions      //actual condition (e.g., knowledge/skill) of a learner
          ST.requirements //necessary requirements to play any role in a given strategy ST
          If canPlayRole(L.conditions, ST.requirements) then
            If sessionsList.existCLSession(T) then  //if there is a CL session using this theory
              CL1 ← returnCLsession(T)
              CL1.add(L, ST) //add learner to an existing CL session
              L.effectiveCLsessionsAdd(CL1)   //track how many sessions can be effective for this learner
            Else
              //create a new session with the specific theory, with one learner playing the role ST
              sessionsList.CreateNewSession(T, L, ST)
  For each session CL in sessionsList do
    For each requirement R in the CLontology do
      If CL does not follow the theory requirements in the CLontology then
        solveRequirement(CL)
          //example
          ;; If a necessary role in CL cannot be played //session cannot be used
              ;; CL.removeAllLearners() ;
              ;; sessionsList.remove(CL);
          ;; If CL.overloadOfLearners(role) then //overload of learners playing the same role
              ;; While(CL.overload(role))
                  ;; Lr ← CL.removeLearnerThatWorstFit();//remove the learners who get less benefits and
                                                        //contribute less for the success of the group work
                  ;; Lr.effectiveCLsessionsRemove(CL);
          //and other restrictions
    CL.setGroupGoal();
    CL.designActivities();
    CL.startSession();

  For each learner L do
    If L.effective CL sessions = Ø then
      L .mode = individual learning //these learners will not obtain many benefits if work in groups
                                    //then start a individual learning mode for these learners
```

Q5 This table is written using the standard guidelines to write pseudo-codes in computer science. More detailed information can be obtained on Cormen et al. (2001)

improve the specific goal. However, usually there is more than one session that can help learners to achieve their goals.

2.2. Check whether learners have the necessary and desired conditions to play a role. Learners who meet all of the conditions are given a high priority to join the group; learners with only the necessary conditions have a low priority; and the other learners cannot join the group, because they could harm the CL process.

2.3. Set the group goal (*common goal*); and design CL activities according to the interaction patterns that are described or prescribed by theories. These patterns can be followed by learners in order to obtain the desired individual and group goals. In previous studies, we have shown how to design CL activities using this ontology.

Note that, unlike other approaches, the method of group formation using ontologies can provide the rationale for group formation. For each choice made to form a group, the ontology provides pedagogical justifications that explain it. For example, we can support instructors by explaining why some learners should collaborate and why others should not; it is also possible to help them set reasonable goals for learners and for the entire group considering the theoretical point of view, the learners' preconditions, and the content to be learned; also, we can ask learners to play specific roles in order to produce a more sophisticated collaboration.

Another interesting way to address the problem of group formation is to utilize ontologies to propose constraints that need to be satisfied. These constraints can be defined as strong constraints (must be satisfied) or weak constraints (an agent can decide whether this constraint should or should not be satisfied). Mapping the ontological representation to constraints that can be solved by existing engines (provided by Semantic Web technologies) is a straightforward and powerful alternative to hardwired algorithms. An example of an algorithm for group formation using constraints can be found in the work of Ounnas et al. (2008).

Experiment

With the objective of obtaining information about the impact of forming groups using theory-driven group formation with our ontologies, we designed an experiment as a **proof-of-concept**. The main goals of the experiment were to gather information and verify (a) whether instructors can use the concepts contained in the ontology adequately, and (b) if the framework of the group formation suggested by the ontology is relevant to the success of the CL session.

The study was carried out with two pairs of qualified instructors, each pair from a different institution, and 20 participants who were expected to develop information sharing and self-expression skills. The participants are from seven different countries from Latin America, pursuing different degrees in Japan (e.g., Medicine, Education, Agronomy), between the ages of 18 and 35 years old. All participants are volunteers in a NGO (Non-Governmental Organization) that supports (a) children's education and (b) international exchange programs that promote cultural understanding. The participants need to learn how to work with people from different countries and with different cultures. Also, they need to improve their skills to present their work concisely and in an understandable manner for a broad audience. We chose such an ill-structured environment for two main reasons: (a) since 2004, these participants have been working together, but have been suffering from many problems in collaborating and sharing information; and (b) in an ill-structured environment, it is easier to identify when a set of changes in the CL settings affects the success of the CL process. We expended about 2 months to complete the whole experiment.

The experiment consists of two phases. The first phase is the planning (set up) of the CL session and the second phase was its actual execution. In the first phase, instructors were asked to deal with the group problem using their own methods. After that, they should find an agreement and select or merge some of the created CL sessions. We specifically asked the instructors to give details about the content to be learned by the participants, their choices to form groups, to define individual and group goals, and to create a sequence of activities (including tools to be used). Next, the same tasks were done using our ontology with methods similar to those proposed in the previous section.

Basically, three different tasks (information sharing tasks) were used by instructors: (a) *construction of mind map*: Each participant has pieces of information (e.g., about their country) and they need to create a complete picture about the situation (e.g., poverty) in each country showing differences and commonalities (e.g., government actions). Finally, they need to come up with a consensus to create a mind map that covers all information discussed by the group; (b) *Cultural exchange:* Each participant is coupled with another participant from different countries and they have to teach about their cultures, and (c) *Exposition*: Each participant gives a small presentation about their own work/study/research and others have to summarize what has been presented. The main goal of these activities was to help participants to acquire knowledge and skills to work in multi-cultural and multi-racial environments where communication skills (not language skills) are essential to exchange information adequately.

The second phase was the application of the proposed sessions. For each CL session, about half of the participants used the scenario proposed by instructors without support of our ontology (control groups), and the other half used the scenario with ontological support (experimental groups). All groups (experimental and controlled) received the support of instructors while the activities were taking place. For each session, different participants were selected to join the experimental groups according to the necessary requirements described in the ontology. All sessions were recorded and evaluated by both instructors and participants who filled out questionnaires after the sessions. The duration of each activity was about 3 to 6 h plus some intervals of 30 min and each CL session was composed by one or more activities. Finally, regarding the conduction of designed collaborative scenarios we did not use any special computer-based support (e.g., CSCL scripts or IMS-LD engines). In our experiment, instructors act as recommender systems, given individual recommendations for each participant before the CL sessions start.

In total, four CL sessions were created. The first one, with the main goal of spreading specific knowledge among participants, was performed in pairs where the more knowledgeable participant should "*teach*" the content to the less knowledgeable one. Four groups followed a Peer Tutoring based CL session (Endlsey 1980), and six control groups did not have any specific guideline. In the second session, the main goal was to improve skills of self-expression. Five groups were created with four members each. Three experimental groups followed a Cognitive Flexibility based CL session (Spiro et al. 1988) where learners had to expose their opinions from different perspectives. For the other two control groups, it was advised that learners should expose their opinion during the task, but no restriction was imposed to ensure it. The third and fourth sessions were engaged in mind map constructions, with the main goal of improving cognitive and metacognitive skills, and the skills for self-expression. Four groups were created with five members for each. One group followed the Cognitive Apprenticeship CL session (Collins 1991) with one teacher and four apprentices; another one followed the LPP CL session (Lave and Wenger 1991) with two full participants and three peripheral participants; and the other two were control groups that received support from instructors, yet their interactions were not restricted in any sense. The group that followed Cognitive Apprenticeship theory had activities such as demonstrations and guided tasks. Although the final goals were the same, the group that followed LPP theory had activities such as discussions and exchanging of ideas. In Table 5, we show some interaction between learners and their educational benefits.

Regarding the assessment process, during the experiment we work together with the four instructors who performed many tasks to evaluate learners. To check the stage of development (knowledge/skills), instructors evaluated learners by giving: pre-tests, post-tests and questionnaires. Also, they analyzed learners' interactions/behaviors during the CL

Table 5 Some Interactions and their benefits for two groups based on different theories

Interaction	Expected benefits (From→To) Role A	Expected benefits (From→To) Role B	Learning Theory
	Master	Apprentice	Cognitive Apprenticeship (Collins, 1991)
Demonstration	s(3, 2)→s(4, 2)	s(0, x)→s(1, x); s(1, x)→s(2, x); x=0,1,2	
Instigating thinking	s(3, 2)→s(4, 2)	s(1, x)→s(2, x); x=0,1,2	
Monitoring/Coaching	s(3, 2)→s(4, 2)	s(1, x)→s(2, x); s(2, x)→s(3, x); x=0,1,2	
	Full Participant	Peripheral participant	LPP (Lave & Wenger, 1991)
Requesting details	s(3, 2)→s(3,3)	s(0,x)→s(1,x); x=0,1,2	
Instigating discussion	s(3, 2)→s(4,3)	s(1,x)→s(3,x); x=0,1,2	
Exchanging information	s(3, 2)→s(4,3)	s(1,x)→s(3,x); x=0,1,2	

sessions and how these interactions affected the final product of the group. Based on these results, they try to determine the stage of development of each learner. For example, if a learner gets a bad score in the pre-test, has a poor performance in the group, and gets a bad score in the post-test, then the instructor could say that such learner does not have any knowledge or skill s(0,0). In another example, if the learner gets a bad score in the pre-test, performs fairly in the group, and gets a better score in the post-test, then the instructor could say that this learner learned some basic concepts and moved from s(0,0) to s(0,1). The explanation of each stage of learning and some strategies to roughly identify them are described in the works of Rumelhart and Norman (1978) and Anderson (1982).

Results and discussion

The interface between instructors and ontologies was mediated by one of the authors. The intention was to capture the necessities of users and to check the usefulness of concepts represented in our ontologies (and not the usefulness of a particular system built using ontologies). With the encouraging feedback and data obtained in the experiment, we believe it is feasible to propose a complete ontology-aware system to support CL as shown in Fig. 1.

Concerning the first phase (planning), all the instructors agreed that the use of the ontology was quite helpful in obtaining insight about the group formation and in designing CL activities. It was discovered that many unconscious choices of instructors, in fact, have been explicitly represented in our ontology. Furthermore, instructors have considered it informative and meaningful that the concepts in our ontology were linked with the relevant theory. Besides this, the theory supports the rationale behind each choice to form a group and to design CL activities; in some cases, the instructors were able to select the theory they felt more comfortable working with. Another benefit pointed out by instructors was the facility to create and to share CL sessions. When each instructor produced their own sessions/scenarios using their own vocabulary, it was quite difficult to discuss the benefits of each one in order to find a common agreement and to merge them. One example of such a problem occurred when producing a CL activity without support (in our case, without ontological support) and then tried to share this activity with another person. In this session, the use of a mind mapping tool was previously established. Then, to identify the problems of spreading information in a determined community and to create a mind map, one pair of instructors proposed the following activity: "*(a) identify specific problems; and (b) cluster these problems into more general problems.*" The other pair of instructors proposed the following: "*(a) examine the main general problems; and (b) break them into small components clarifying their relationships.*" When the pairs exchanged their proposed activities, initially, both pairs classified these activities as different ones. However, after a more careful analysis, they realized the activities had the same goal (***what to achieve***) which was to identify the problems and their subproblems in a given topic and show and identify the correct relationships between them. The main difference between the activities was ***how to achieve*** the goal. The first activity described a bottom-up approach while the second one described a top-down approach.

According to instructors' comments, using CL ontology, the activities and sessions they described were more easily comprehended when they exchanged their created CL activities and sessions. Furthermore, the ontology was used only as a guideline or basic structure to help them propose CL sessions with theoretical justification. Thus, the instructors also had the flexibility of not heavily relying on the theories and adding the characteristics they think the groups needed in order to work effectively. Our research shows that the use of the ontology did not restrict instructors' actions or their creativity. Instead, it helped them to focus on the main problem and to make efforts in parts where their expertise was required the most.

A simple, yet prime, example of the CL ontology usage for group formation is evident from the planning of the first session. In this session, the main goal was to spread knowledge among participants. Using conditions such as the level of knowledge of the participants and the desired goal, our ontology suggests that a peer-tutoring-based CL session could be well applied in this situation. Such a suggestion encouraged instructors to pair participants of the highest level of the content-specific knowledge (restructuring stage), the tutors, with participants of the lowest level of the content-specific knowledge (Nothing), the tutees. These participants correspond to the top and bottom in Table 6, respectively.

However, the ontology suggests that the tutor should ***not*** be those who have the knowledge in restructuring stage. Instead, the tutors should be those who have knowledge in the accretion stage, which means they have the necessary knowledge, but do not have experience in teaching it to others, possibly leading to some misunderstandings (Fig. 6). As we presented in previous sections, there are at least two reasons for this suggestion based on theoretical justifications (Endlsey 1980). First, if the tutors already have knowledge in the

Table 6 Level of content specific knowledge of participants in session 1

Member ID	Knowledge	Member ID	Knowledge
20	Restructuring	13	Tuning
7	Restructuring	8	Accretion
3	Restructuring	10	Accretion
4	Restructuring	14	Accretion
5	Tuning	15	Accretion
6	Tuning	11	Accretion
1	Tuning	17	Nothing
18	Tuning	9	Nothing
19	Tuning	2	Nothing
16	Tuning	12	Nothing

Q4

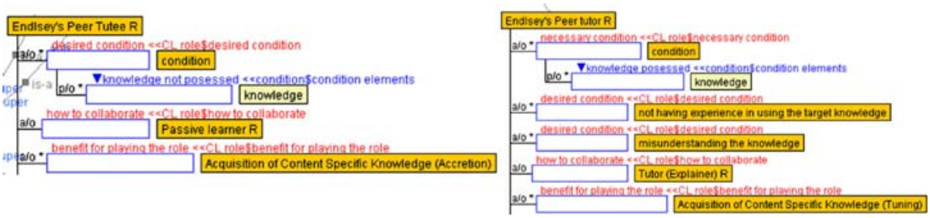


Fig. 6 Conditions for role play in a peer-tutoring-based CL session

restructuring stage, it means that they already understand the content well and have either been using or teaching it many times. Then, in this case, only the participants playing the role of tutees will attain some measurable benefits. The second reason for using participants in the accretion stage as tutors is that they must explain the content to teach or share their knowledge and, consequently, they (a) can obtain a better understanding about it, (b) can be aware of possible misunderstandings in their own knowledge, and (c) can solve some of these misunderstandings by asking for help. Thus, both tutor and tutee can obtain measurable benefits, increasing the successfulness of the CL session. By receiving such pedagogically valid advice, the instructors were quite pleased to change their position when creating groups using the peer-tutoring-based CL session (experimental groups) and groups paired randomly (control groups). As shown in Table 6, four participants did not have the desired knowledge. Thus, instructors proposed four experimental groups and six control groups.

Besides stages of development (knowledge and skills), to form groups, instructors also considered other information such as: the language spoken by participants (to facilitate self-expression), educational background (to increase heterogeneity of thoughts), culture (to increase cultural exchange), previous relationships with other participants (to avoid meaningless interactions), gender (to avoid groups with only men or women), and intrinsic behavior of participants. Table 7 shows some information used to form groups in the first CL session based on Peer Tutoring.

Note that differently from conventional experiments which try to compare individuals who participated only in control groups or experimental groups, our experiment has a different objective which is to identify if any participant at any condition (actual and previous learning history) can join an experimental group and have a better learning

Table 7 Some information used to form groups in a peer-tutoring-based CL session

Member ID	Knowledge	Language	Educ. Background	Country	Gender	Behavior	Group ID
8	Accretion	Spanish/Japanese	Medicine	Paraguay	F	reflective	1
10	Accretion	Spanish	Medicine	Peru	F	Active	2
14	Accretion	Spanish/Portuguese	International trading	Colombia	M	Active	4
15	Accretion	Portuguese	Japanese Drum	Brazil	M	Active	3
17	Nothing	Portuguese/Japanese	Education	Brazil	F	Active	1
9	Nothing	Portuguese	Acupuncture	Brazil	F	Reflective	3
2	Nothing	Spanish/Japanese	Economics	Paraguay	F	Reflective	2
12	Nothing	Portuguese	Architecture	Brazil	F	Active	4

experience if compared with his/her peers in control groups. Also, it is possible to compare his/her performance with previous performances when interacting with other learners in a control group setting. Furthermore, to avoid too much interference between CL sessions, in our experiment, *each CL session is considered a unique event*. Then, each session had its own pre- and post-tests, the interaction analysis considers only the interaction which occurred within a session, students evaluate their partners concerning their participation within a session, and so forth. Roughly the following schema was adopted to form groups and run activities for each session:

Start session A:

1- Learners' knowledge/skills for the particular domain and topic are assessed (pre-test).
2- Then, according to the pre-test, participants are assigned to experimental or control groups according to the algorithm proposed in this paper. Experimental groups are requested to follow a specific guideline according to a selected theory, and control groups can work more freely.
3- During collaboration, instructors assess the evolution of experimental and control groups.
4- At the end of the session, learners have a post-test. Questionnaires are also filled by participants to evaluate their peers.
5- Then, Instructors re-analyze the interactions of each participant (using recorded videos).
6- Finally, an overall evaluation of each group/participant is presented.
7- Groups are "dissolved."

End Session A.

For each session, we fixed learners in experimental groups and control groups. Therefore, it was possible to check the development of each participant within each session. When a new session starts we went back to step 1 above. In the new session, the content, skills, and knowledge tackled were different from the previous session and the pre-test was done again. In Table 8, it is shown the distribution of participants in each session.

After the members of each group were chosen, instructors used the interaction patterns represented in our ontology to properly propose the sequence of CL activities. As we discussed previously, each interaction pattern is a model of typical interaction processes described in one of the learning theories in the CL ontology. In a learning theory, educational benefits obtained by a learner through interactions are either implicitly or explicitly described. Thus, the interaction patterns have been developed to facilitate specific interaction processes that are recommended by theories to achieve specific learning goals (Inaba et al. 2002, 2003). For example, the first session used the interaction pattern for Peer Tutoring (Endlsey 1980). An illustrative visualization of this pattern is shown to the right of Fig. 7. Solid boxes represent necessary interactions, or those that are essential to attain the desired educational benefits; dotted boxes represent complementary interactions, or those that support the achievement of desired benefits but are not essential. Each of these boxes possesses some events related to it. In the case of Peer Tutoring, we have a Tutor event and a Tutee event. The arrow shows desired transitions between interactions. This pattern is richly represented in CL ontology where the actions of each participant, their benefits, and other information are also explicitly and formally described. A small portion of the CL ontology describing the interaction activities within the colored box is shown to the left of Fig. 7.

In the second phase of our experiment, we tried to verify differences between the control groups and the groups formed using our ontology (experimental groups). For each CL

Table 8 Distribution of participants into groups during each session

Session	Type	Groups and Participants	Member ID
1	Experimental	G1.1	8; 17
		G1.2	10; 2
		G1.3	15; 9
		G1.4	14; 12
	Control	G1.5	11; 20
		G1.6	3; 4
		G1.7	5; 13
		G1.8	18; 6
		G1.9	19; 7
		G1.10	16; 1
2	Experimental	G2.1	3; 7; 17; 19
		G2.2	4; 6; 16; 18
		G2.3	10; 13; 14; 20
	Control	G2.4	1; 5; 8; 9
		G2.5	2; 11; 12; 15
3	Experimental	G3.1	3; 13; 15; 19; 20
		G3.2	1; 4; 10; 11; 14
	Control	G3.3	2; 8; 9; 16; 18
		G3.4	5; 6; 7;12;17
4	Experimental	G4.1	2; 4; 7; 8; 9
		G4.2	3; 11; 12; 13; 18
	Control	G4.3	1; 6; 14; 17; 19
		G4.4	5; 10; 15; 16; 20

session, instructors checked how the participants interacted with each other, the groups' achievements, and the benefits obtained by individuals, besides other indicators. Although the number of participants is not statistically significant to make a richer analysis or stronger conclusions, we have found some interesting results.

First, instructors observed that in the control groups more than half of the scheduled time of some sessions was filled with meaningless interaction, instead of performing the necessary activities that would improve the desired skills. Meaningless interactions were defined by instructors as those that interfere with the good "health" of the group and the progress of collaboration among group members. Examples of meaningless interactions are: arguing among members, long discussion without any concrete result, "off-topic" discussion, abrupt interruption while good collaboration is taking place, excessive participation (of one member) or lack of it, besides many others. Furthermore, it was noted that on many occasions, members of experimental groups who had worked well together in previous sessions could not work together in control groups, harming the CL process. One explanation is that in the experimental groups, participants who were chosen adequately (rather than randomly, as it usually happens), had defined roles and could follow well-structured interaction patterns. As many studies have shown, following these regulations can decrease the chances of undesirable interactions (Dillenbourg 2002; Strijbos et al. 2004). In Fig. 8, we show the percentage of meaningful interactions of both experimental and control groups. In this figure, it is possible to observe that the experimental groups spent more time in meaningful interaction than the control groups in all the designed CL sessions.

Q4

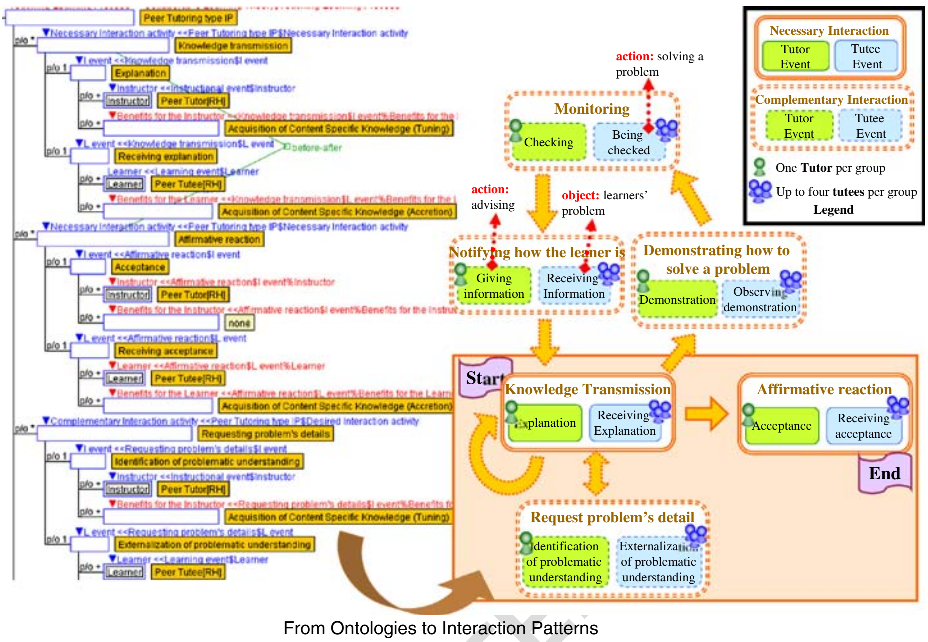
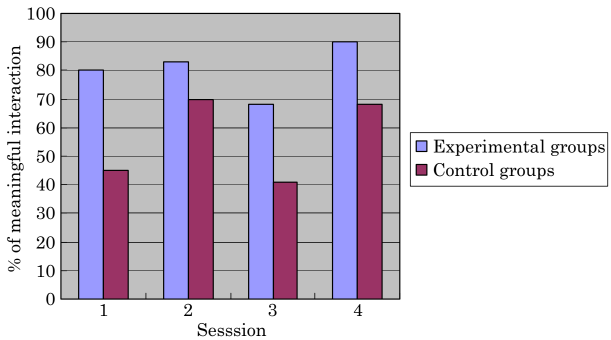


Fig. 7 In the right, an illustration of the interaction pattern for Peer Tutoring. In the left, a small portion of the CL ontology representing the interactions within the colored box

Another interesting result obtained in our experiment was that in most sessions the participants in the experimental groups had more improvement, and the performance of the whole group was better when compared with the control groups. Figure 9 shows the final scores for each participant given by instructors considering both qualitative (e.g., how the interactions were performed) and quantitative (e.g., individual and group test scores) parameters. In the left side of each graph on Fig. 9, we cluster the scores of participants who joined the experimental group for each session. For example, in the first session, eight participants joined experimental groups based on Peer Tutoring, and their grades are presented in the first eight columns of the top-left graph; in the second session, twelve participants joined the experimental groups based on Cognitive Flexibility, and their grades are also presented in the first twelve columns of the top-right graph. The same follows for sessions 3 and 4 (columns 1 to 10). According to the instructors, most of the participants

Fig. 8 Percentage of meaningful interactions during the CL session



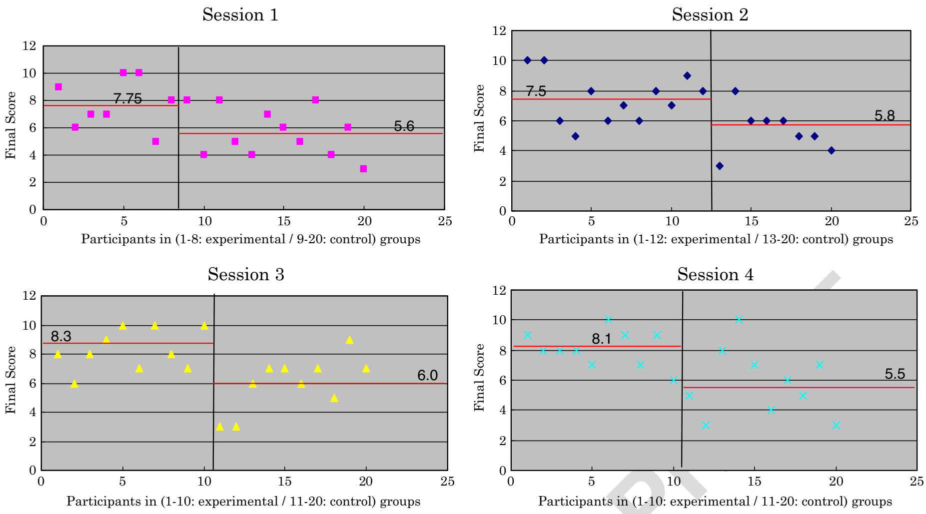


Fig. 9 Scores of Participants after each CL session. On the left of each graph you have participants in experimental groups and on the right participants in control groups. The red lines are the median of participants' scores

who joined the experimental groups achieved their individual goals, and the groups performed more smoothly. As a result, we can observe that the average of the participants in experimental groups had a better score when compared with the average of participants in control groups. In each graph, the median of the scores obtained by participants in experimental and control groups are shown as a red line. It is also worth to observing that in the experimental groups only 2 scores (out of 40), which means 5% of the total, were lower than the borderline (value 6) and in the control groups we have 18 scores below the borderline, which means 45% of the total. Furthermore, in the experimental groups 25% of the scores were equal to or above 9, and in the control groups only 5%. Finally, the average score, considering all sessions in the experimental groups, was 7.9, with standard deviation (σ) equal to 1.38, while in control groups it was 5.7, with standard deviation equal to 1.80. Although the number of subjects is not statistically significant, this experiment can be used as proof-of-concept to demonstrate the feasibility of our framework. In this situation, these results suggest that our group formation methodology might have some good impact on learning development in the group learning context. Furthermore, we identified that the majority of learners obtained good results when interacting in experimental group settings even if they had bad performance in previous sessions when interacting in control group settings. Furthermore, learners who had good results in experimental group settings often had worse performance when in control groups. This result suggests that independently of the previous group learning experiences, it might be possible to give a good support to learners by providing a good group formation and CL scenario where learners can interact more effectively.

Through an interview with instructors and other learners, we have confirmed that the main reasons for such a clear difference of performance between the experimental and control groups are the adequate formation of group members together with the structured design of CL activities. In our experiment, instructors and participants who were in experimental groups had in hand the sequence of interactions (interaction pattern) suggested by our ontology and instantiated by instructors, together with explanations about what the goals of each interaction were and what actions were expected for each participant. Thus, when a participant in a group interacted in a way that did not contribute to the goal of the interaction and/or the expected actions have not been performed, then other participants or the instructor could ask him/her to keep following the script (interaction pattern). The approach of using ontologies to give explicit information to participants empowered them allowing for the group's self-regulation to fit their interactions in the proposed interaction pattern. The possibility of providing support for group-based self-regulation has recently been introduced as a good mechanism to structure and investigate individual's interactions in group contexts (Sassenberg and Karl-Andrew 2008). According to the questionnaires filled out by instructors, having an adequate group formation and interaction pattern seemed to result in a better involvement of participants, which facilitated positive interdependence and individual accountability of participants. According to participants' questionnaires, because everyone had a role in the group and could check how they would contribute to the achievement of the group goal, they felt a sense of partnership with each other. Thus, contributions of each participant were more respected and expected; group consensus was reached more rapidly.

For example, in sessions 3 and 4, the groups had the goal to share the skill for building a mind map, and the experimental groups based on Cognitive Apprenticeship participants were chosen according to pedagogical specifications. These specifications explicitly informed the roles, tasks, and individual goals for each participant. The participants who played the role of "Master" had to increase his or her ability to build a map (which means to develop the specific skill in the autonomous stage) while the "apprentices" had to learn how to build a map adequately (which means to develop the specific skill in the associative stage). Throughout specific tasks, the teacher helped the "apprentices" produce a map by externalizing his or her cognitive processes while building maps and monitoring apprentices. As a result, on one hand, the learners playing the role of "master" acquired the desired individual goal. On the other hand, by observing, imitating, and being monitored, the "apprentices" developed the desired skill effectively and smoothly. In these sessions, the participants in the control groups did not achieve an effective group performance. Although some members of the control groups achieved their individual goals, the groups could not achieve their desired goals. A lack of coordination was observed among participants, which frequently generated strong disagreement among some and caused an increased indisposition of participants to working together. As one participant had pointed out: "*We spent too much time to organize our thoughts and only a little time left to present solutions for the topic. I believe if we had more time before the activity to discuss the topic (informally) our results could have been improved.*" Another problem was that some participants did not contribute toward the group goal. For example, in one of the sessions one participant complained about the behavior of another one: "*one participant didn't interact with the group. She was just listening without doing anything for the group, except when someone asked her opinion. I think if someone wants to work in a group he/she must work for the group.*" In this case, the participant who did not interact with others in the group behaved passively because she did not know what the group was expecting her to do. In such situations, participants were more likely to work alone to develop their own skill than supporting their colleagues in a group learning environment. It was also identified that the indisposition among some participants remained after the CL sessions, indicating that a previous harmful CL session may partially have a negative effect on future CL sessions.

However, according to the participant's opinions, it was somewhat difficult to follow some settings of experimental groups such as appointed roles, strategies, and tasks. One of their arguments was that sometimes they had to neglect their personal behavior to get the task completed as requested. Another comment was that it would be preferable to have more than one sequence of activities, so that they could choose a sequence that suited them better, avoiding or at least decreasing the sense of obligation in completing an unwilling sequence of activities. Those complaints are reasonable and will be taken into consideration to improve our ontology.

In conclusion, the results of this experiment as a proof-of-concept suggest that the framework of group formation presented in our ontology can be used to adequately form effective groups. This verification is essential in order to provide intelligent systems with theoretical knowledge that clarifies how learning theories help instructors to form groups, to design CL activities, and to enhance learning outcomes. The ontology presented in this work aims to represent the knowledge of intelligent educational systems that support CL, while playing a central role in the decision making about ***how***, ***when***, and ***why*** learning theories can (or should) be used to form groups considering the factors that influence the CL process.

Conclusions

The main goal in this work was to demonstrate that, to some extent, it is possible to use ontological engineering to “operationalize” and capture important concepts of learning theories from a purpose-specific perspective and thereby support group formation in CSCL. Our assumption is that each theory has strong and weak points and depending on the situation we can switch from one theory to another. To allow that, concepts of learning theories must be explicitly represented and ontologies are used for this purpose. Thus, an intelligent system can use ontologies to help users to form groups, design CL activities, and realize when a theory is more appropriate than others, considering participants’ conditions, teacher’s preferences, and other resources in the environment. According to our analysis, some of the critical elements presented in learning theories that affect group formation and learners’ interactions are: (a) individual goals, (b) group goals, (c) group arrangement goals, (d) roles, (e) learning strategies, (f) learner’s behavior, (g) interaction patterns, and (h) learners’ stage of knowledge/skill. All these elements and many others are semantically connected and explicitly represented in our CL Ontology.

We also proposed a method for adequately using the concepts on our ontology which consist of (a) understanding students’ needs, and then (b) selecting a theory to support group formation and designing of CL activities that satisfy the needs of all students in a group. Our assumption was that by having students’ information beforehand, we can increase the benefits of collaboration by grouping students who can support one another, assigning roles adequately according to student’s conditions, and proposing more personalized CL activities that help them to achieve their individuals and group goals. This approach does not intend to neglect the existence of other effective methods for group formation, such as the conventional one that first, selects a group goal and a basic structure to design collaborative tasks and then assigns real learners to roles and groups. Instead, our ontology can support a variety of group formation methods and depending on the situation, different methods should be combined and used to increase the benefits of group learning.

One main difference between our approach and conventional ones is the view about the relationship between group formation and the design of interactions. Usually, conventional approaches separate group formation from design of interactions. However, the authors consider a different approach where these two problems are intrinsically connected. Such a view comes from the fact that learning theories propose guidelines to group learners for specific sequences of interactions. Therefore, to make the realization of theory-driven group formation come true, both group formation and structuring of interactions need to be treated as a single unit. Otherwise we cannot say that a group formation is theoretically sound. Based on our commitment, it would be inconsistent to form groups adequately and let participants interact freely (and vice versa). Such differences in viewpoint could be an interesting issue for future research and study cases on this topic.

To verify the usefulness and effectiveness of our ontology and method for group formation, we conducted an experiment as a proof-of-concept with 4 instructors and 20 participants. The results of the experiment indicate that the concepts in the ontology helped instructors to form groups and design CL activities with theoretical justifications. Additionally, although our results are not statistically guaranteed, they suggest that individuals in experimental groups in which each member was carefully selected and the interactions were partially moderated following the prescriptions in the ontology, performed and learned better than those in the control groups whose members were not selected so rigorously and could interact freely with others. We hope that with the insights of this work other researchers in the educational field will have an interest in proposing many different ways of grouping students, and perhaps combining aspects of different theories to explore possibilities for increasing collaborative learning benefits.

We believe the ontology developed in this work is a step forward in the development of the foundations of an intelligent authoring tool for CL, with well-grounded theoretical knowledge, that supports group formation, facilitates the design of CL activities, and minimizes the load of interaction analysis (Fig. 2). The experiment shows that the CL ontology can provide useful information to support CL processes, and it can be further improved considering the comments of instructors and participants. Furthermore, we have already started to merge concepts in the CL ontology with the OMNIBUS ontology, which represents the theoretical knowledge for individual learning, developed by Hayashi et al. (2006, 2008), and is freely available through the Internet at http://edont.qee.jp/omnibus/. Thus, it will be possible to select the best situations for switching from an individual learning mode to a collaborative learning mode (and vice versa) and to create learning scenarios "*on the fly*" that more adequately support instructors and help learners achieve their goals. Our ultimate goal is the realization of *AAAL: Anytime, Anywhere, Anybody Learning* through the development of theory-aware systems which use ontologies to help instructors and learners, structure learning activities and materials compliant with instructional/learning theories, and guide them to perform individual or collaborative learning. We believe such systems have huge potential for making *AAAL* meaningful to both instructors and learners.

Finally, the current use of ontologies has been suffering from a lack of good interfaces that allow instructors/teachers to easily connect with the formal notation of concepts. In order to develop better ontology-aware systems, there is a strong need for researchers from Human-Computer Interaction and Artificial Intelligence to cooperate to create smart interfaces that completely hide the ontology from end-users and ask or present only the minimum amount of information necessary to do some reasoning (information in the ontology should be automatically extracted from the knowledgebase). Using such smart interfaces, it would be possible to decrease the need of end-users (e.g., teachers) to work with formal notation and minimize the overload of information that they have to deal with to perform a task. In future research, we can utilize the results of this work as the basis to propose more user-friendly ontology-aware systems.

Acknowledgement We would like to thank the reviewers and editors of the *ijCSCL* for their helpful comments and suggestions. We also would like to thank the Nippon Foundation, the Association of Nikkei & Japanese Abroad, JICA (Japan International Cooperation Agency), IBM Research and the Department of Knowledge Systems (MizLab) for their financial and technical support.

References

Alfonseca, E., Carro, R. M., Martín, E., Ortigosa, A., & Paredes, P. (2006). The impact of learning styles on student grouping for collaborative learning: a case study. *User Modeling and User-Adapted Interaction, 16*(3–4), 377–401.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*(4), 369–406.

Aronson, E., & Patnoe, S. (1997). *The jigsaw classroom: Building cooperation in the classroom* (2nd ed.). New York: Addison Wesley Longman.

Bandura, A. (1971). *Social learning theory.* New York: General Learning.

Barkley, E., Cross, K. P., & Major, C. H. (2005). *Collaborative learning techniques: A practical guide to promoting learning in groups*. San Francisco: Jossey Bass.

Barros, B., Verdejo, M. F., Read, T., & Mizoguchi, R. (2002). Applications of a collaborative learning ontology. In *Proceedings of the Mexican International Conference on Artificial Intelligence*, LNCS 2313, 103–118.

Cognition and Technology Group at Vanderbilt. (1992). Anchored instruction in science education. In R. Duschl & R. Hamilton (Eds.), *Philosophy of science, cognitive psychology, and educational theory and practice* (pp. 244–273). Albany: SUNY.

Collins, A. (1991). Cognitive apprenticeship and instructional technology. In L. Idol & B. F. Jones (Eds.), *Educational values and cognitive instruction: Implications for reform* (pp. 121–138). Hillsdale: Erlbaum.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001). *Introduction to algorithms*. Cambridge: The MIT.

Deibel, K. (2005). Team formation methods for increasing interaction during in-class group work. *ACM SIGCSE Bulletin, 37*(3), 291–295.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL?* (pp. 61–91). Heerlen: Open University Nederland.

Devedzic, V. (2006). *Semantic web and education*. New York: Springer.

Endlsey, W. R. (1980). *Peer tutorial instruction*. Englewood Cliffs: Educational Technology.

Ertmer, P. A., & Newby, T. J. (1993). Behaviorism, cognitivism, constructivism: comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 6*(4), 50–70.

Faria, E. S. J., Adán-Coello, J. M., Yamanaka, K. (2006). Forming groups for collaborative learning in introductory computer programming courses based on students' programming styles: An empirical study. In *Procceedings of the ASEE/IEEE Frontiers in Education Conference*, S4E-6–S4E-11.

Fiechtner, S. B., & Davis, E. A. (1985). Why some groups fail: a survey of students' experiences with learning groups. *Organizational Behavior Teaching Review, 9*(4), 75–88.

Graf, S., & Bekele, R. (2006). Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization. In *Proceedings of Intelligent Tutoring Systems*, LNCS 4053, 217–226.

Greer, J., McCalla, G., Cooke, J., Collins, J., Kumar, V., Bishop, A., & Vassileva, J. (1998). The Intelligent Helpdesk: Supporting Peer-Help in a University Course. International Conference on Intelligent Tutoring Systems, LNCS 1452, 494–503.

Harrer, A., McLaren, B. M., Walker, E., Bollen, L., & Sewall, J. (2006). Creating cognitive tutors for collaborative learning: steps toward realization. *User Modeling and User-Adapted Interaction, 16*(3–4), 175–209.

Hayashi, Y., Bourdeau, J., & Mizoguchi, R. (2006). Ontological support for a theory-eclectic approach to instructional and learning design. In *Proceedings of the European Conference on Technology Enhanced Learning*, LNCS 4227, 155–169.

Hayashi, Y., Bourdeau, J., & Mizoguchi, R. (2008). Structurization of learning/instructional design knowledge for theory-aware authoring systems. In *Proceedings of the International Conference on Intelligent Tutoring Systems*, LNCS 5091, 573–582.

Inaba, A., & Mizoguchi, R. (2004). Learners' roles and predictable educational benefits in collaborative learning. In *Proceedings of the International Conference on Intelligent Tutoring Systems*, LNCS 3220, 285–294.

Inaba, A., Supnithi, T., Ikeda, M., & Mizoguchi, R. (2000). How can we form effective collaborative learning groups. In *Proceedings of Intelligent Tutoring Systems*, LNCS 1839, 282–291.

Inaba, A., Ohkubo, R., Ikeda, M., & Mizoguchi, R. (2002). An interaction analysis support system for CSCL. In *Proceedings of the International Conference on Computers in Education*. IEEE Press. 358–362.

Inaba, A., Ikeda, M., & Mizoguchi, R. (2003). What learning patterns are effective for a learner's growth? In *Proceedings of the International Conference on Artificial Intelligence in Education*, 219–226.

Isotani, S., & Mizoguchi, R. (2006). An integrated framework for fine-grained nalysis and design of group learning activities. In *Proceedings of the International Conference on Computers in Education*, IOS Press, v. 151, 193–200.

Isotani, S., & Mizoguchi, R. (2007). Deployment of ontologies for an effective design of collaborative learning scenarios. In *Proceedings of the International Workshop on Groupware*, LNCS 4715, 223–238.

Isotani, S., & Mizoguchi, R. (2008). Adventures in the boundary between domain-independent ontologies and domain content for CSCL. In *Proceedings of the International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*, LNAI 5179, 523–532.

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., & Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*(2–3), 211–224.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Miao, Y., Hoeksema, K., Hoppe, H. U., & Harrer, A. (2005). CSCL scripts: Modelling features and potential use. In *Proceedings of the International Conference on Computer Support for Collaborative Learning*, 423–432.

Mizoguchi, R., Sunagawa, R., Kozaki, K., & Kitamura, Y. (2007). The model of roles within an ontology development tool: Hozo. *Applied Ontology, 2*(2), 159–179.

Muhlenbrock, M. (2005). Formation of learning groups by using learner profiles and context information. In *Proceedings of the International Conference on Artificial Intelligence in Education*, 507–514.

Ounnas, A., Davis, H., & Millard, D. (2008). A framework for semantic group formation. In *Proceedings of the IEEE International Conference on Advanced Learning Technologies*, 34–38.

Psyche, V., Bourdeau, J., Nkambou, R., & Mizoguchi, R. (2005). Making learning Design standards works with an ontology of educational theories. In *Proceedings of the International Conference on Artificial Intelligence in Education*, 539–546.

Resta, P., & Laferrière, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review, 19*(1), 65–83.

Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning, and restructuring: Three modes of learning. In J. W. Cotton & R. Klatzky (Eds.), *Semantic factors in cognition* (pp. 37–53). Hillsdale: Erlbaum.

Romiszowski, A. J. (1981). *Designing instructional systems*. New York: Nichols.

Salomon, G. (1993). *Distributed cognitions*. Cambridge: Cambridge University Press.

Sassenberg, K., & Karl-Andrew, W. (2008). Group-based self-regulation: the effects of regulatory focus. *European Review of Social Psychology., 19*, 126–164.

Soh, L., Khandaker, N., & Jiang, H. (2008). I-MINDS: a multiagent system for intelligent computer-supported collaborative learning and classroom management. *Journal of Artificial Intelligence in Education, 18*(2), 119–151.

Soller, A. (2001). Supporting social interaction in an intelligent collaborative learning system. *International Journal of Artificial Intelligence in Education, 12*(1), 40–62.

Soller, A., Martínez-Monés, A., Jermann, P., & Muehlenbrock, M. (2005). From mirroring to guiding: a review of state of the art technology for supporting collaborative learning. *Journal of Artificial Intelligence in Education, 15*(4), 261–290.

Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In *Proceedings of the Annual Conference of the Cognitive Science Society*. 375–383.

Stahl, G., Koschmann, T., & Suthers, D. (2006). *CSCL: An historical perspective. Cambridge handbook of the learning sciences* (pp. 409–426). Cambridge: Cambridge University Press.

Strijbos, J. W., & Fischer, F. (2007). Methodological challenges for collaborative learning research. *Learning & Instruction, 17*(4), 389–393.

Strijbos, J. W., Martens, R. L., & Jochems, W. M. G. (2004). Designing for interaction: six steps to designing computer-supported group-based learning. *Computers and Education, 42*(4), 403–424.

Suthers, D. D., Dwyer, N., Medina, R., & Vatrapu, R. (2007). A framework for eclectic analysis of collaborative interaction. In *Proceedings of the International Conference on Computer Supported Collaborative Learning (CSCL)*, 694–703.

Vygotsky, L. S. (1978). *Mind in society: The development of the higher psychological processes*. Cambridge: Harvard University Press. (re-publication).

Wessner, M., & Pfister, H. (2001). Group formation in computer-supported collaborative learning. In *Proceedings of ACM CSCW*, 24–31.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES.

Q1. Please check authors' affiliations if captured appropriately.
Q2. "Inaba et al. 2004" was cited here but not found in the Reference list. Please provide complete bibliographic information. Alternatively, delete the citation in the text.
Q3. Footnote 5 was transfered here as table footer. Please check.
Q4. Figures 5, 6, and 7 are poor in quality. Kindly provide new images with better quality, otherwise, please advise if we can proceed with the figures as is.
Q5. Footnote 7 was tranfered here as table footer. Please check.