Fostering learning and collaboration in a scientific community – evidence from an experiment using RFID devices to measure collaborative processes

Julia Eberle¹, Karsten Stegmann¹, Kris Lund², Alain Barrat³, Michael Sailer¹, Frank Fischer¹

¹Chair of Education and Educational Psychology, University of Munich, Leopoldstr. 13, 80802 München, Germany
²CNRS, ICAR UMR 5191, Université de Lyon, Université Lumière Lyon 2, ENS de Lyon, 69007 Lyon, France
³CNRS, CPT UMR 7332, Aix Marseille Université, 13288 Marseille, France and Data Science Laboratory, ISI Foundation, 10123 Torino, Italy

Email: julia.eberle@psy.lmu.de, karsten.stegmann@psy.lmu.de, kristine.lund@ens-lyon.fr, michael.sailer@psy.lmu.de, alain.barrat@cpt.univ-mrs.fr, frank.fischer@psy.lmu.de

Abstract: In this study, the integration of new members into a scientific community that comprised to a large extent members from the CSCL community was investigated. New members usually lack the necessary knowledge to interact successfully with more experienced members of a scientific community and to find collaboration partners. We investigated how the level of community participation and support for community knowledge were related to the building of new collaborative relationships during a scientific conference. Participants’ interaction behavior was tracked using RFID devices; social network questionnaires and a bibliographic analysis provided additional data. We found that newcomers do not interact less with other participants than experienced members, but develop fewer collaborative relationships. The chances that newcomers’ interactions lead to the building of new collaborative relationships were increased by access to explicit relevant community knowledge. Making such knowledge explicit seems to be a useful means for supporting newcomers in scientific communities.

Integrating new members into the CSCL community

The CSCL research community has already been the object of research in the past. We have learned that it is a broad interdisciplinary community comprising researchers from more than 11 disciplines who are distributed all over the world but with a majority in North America and Europe and a growing number of members in Asia (Hoadley, 2005; Kienle & Wessner, 2006; Stahl, Spada, Miyake, & Law, 2011). We know that there is an ongoing discussion about what CSCL is, what it comprises, and what a common and shared theory could be; also perspectives on the community differ between members of different geographical locations. Nevertheless, Hoadley (2005) found that the CSCL community consists of a stable core of leading persons. Also a stabilizing trend of CSCL conference participants was observed in 2006 (Kienle & Wessner) which might be seen as a sign for a maturing community.

A healthy scientific community must constantly integrate new members to secure its existence and to bring new ideas in. However, it was observed at previous CSCL conferences that a large percentage of newcomers participated only once and did not participate in later conferences (Kienle & Wessner, 2006). This might be considered an alarming sign and the community might wish to take actions to change this situation and focus more on the integration of new members. So far, research on scientific communities is mostly based on bibliometric analyses focusing on co-authorship or citation analysis of conference proceedings or journal papers (see for example Hoadley, 2005; Newmann, 2004; Lee, Ye, & Recker, 2012). However, papers are artifacts that become only available to the scientific community with a large timely distance to their creation and many factors mediate between the publication of a paper and the beginning of a collaboration between members of the scientific community. To identify those factors that are directly related to the integration of newcomers, such bibliographic analyses are therefore not the optimal choice. For this reason, we conducted a study using a new approach to measure collaboration: RFID devices that all participants of a conference wore and which tracked their interaction with other participants. The aim of this study was to investigate what happens during scientific meetings and how the integration of newcomers can be fostered at such occasions.

Scientific Communities and New Members

The CSCL community can be seen as a scientific community, which is a special form of a community of practice (Kienle & Wessner, 2006). Kienle and Wessner (2005) collected essential characteristics of scientific communities: They consist of a heterogeneous group of members who are usually involved in several scientific communities and are therefore used to switching roles, from expert in one scientific community to less experienced in another scientific community. Members of scientific communities are often geographically
distributed and belong to different organizations; in the CSCL community, members are even located across the whole world, although most members can be found in North America and Western Europe (Hoadley, 2005; Kienle & Wessner, 2006). In many scientific communities, members have backgrounds in different disciplines and scientific cultures, resulting in the use of different methods and theories. What brings them together is a joint field of research interests. Communication and interaction mostly takes place via written artifacts like journal publications and using computer-mediated channels, but there are also regular opportunities for face-to-face meetings, usually organized in the form of conferences and workshops.

Wenger, McDermott, and Snyder (2002) distinguish different levels of community participation in communities of practice located in and across organizations: outsiders, peripheral members, active members, and core members; in addition, some communities of practice have a coordinator. For scientific communities, these participation levels were adapted (Kienle & Wessner, 2005). Outsiders are person who do not intend to contribute to the scientific community, but benefit from its work; for example by reading single papers and maybe exporting ideas and results from them to another scientific community. Peripheral members are, according to the original model, persons who contribute only sporadically to the community and often lack the abilities and knowledge to contribute to more complex tasks. Therefore, newcomers to a community usually start as peripheral members. In a scientific community, peripheral members could be graduate students preparing their first papers but also more experienced researchers who explore a research field that is new to them. While focussing only on observable contributions to conferences, Kienle and Wessner (2005) suggested viewing only passive conference participants as peripheral members. Following that rule, they defined all paper authors as active members. This is, however, only one possible way to conceptualize this participation level as the transition between peripheral and active membership is smooth and graduate students who wrote and presented their paper could still be seen as peripheral members in the first learning stage. Active members usually are defined as those persons who regularly contribute to a community and have the necessary knowledge and skills to do so. Core members are those active members who additionally take over substantial responsibility for the whole community and make efforts to influence its directions. In a scientific community, this can include conference program organizers, journal editors, or scientific board members.

Communication and Collaboration in Scientific Communities

In scientific communities, especially in interdisciplinary ones, successful collaboration can be understood as one of the most desired results of researchers’ learning and of the scientific community’s cohesion. Successful interdisciplinary collaboration in a scientific community requires an integration of the contributing disciplines on some level, for example the mutual integration of concepts, theories, methodologies, and epistemological principles (van den Besselaar & Heimeriks, 2001). The development of mutual understanding and the building of shared representations are important for fruitful communication between experts of multiple domains (Fischer, 2000).

To enable an individual researcher to benefit from and collaborate in a scientific community, this person must to some extent be integrated and has to acquire several types of knowledge that are shared in the community. Successful collaboration requires shared knowledge, including several different types of relevant knowledge, for example about contents or methods, but also about attitudes in the community as well as about the individual members (Cannon-Bowers & Salas, 2001). However, peripheral members usually have only little knowledge about the new community and need to acquire it first to become able to contribute more and in a proper way (Levine & Moreland, 1999). Compared to mono-disciplinary communities, this might be even more complicated in multi- and interdisciplinary communities because of the variety of research lines. The CSCL community still is divided on several questions, no underlying theory or methodology can be found on which all members would agree as being the basis of CSCL. Attitudes and beliefs of a community are even harder to grasp as they are usually not made explicit. Knowledge about individual community members seems comparably easy to acquire as most CSCL researchers present their bios and publications on their website. But for a newcomer, it is hard to identify the ‘important’ persons in a community or those who could be relevant for their own work. This makes it hard for newcomers to gain relevant knowledge about content, methods, community attitudes, and members within CSCL on their own.

A usual way to learn about a scientific community is reading papers. But to read only some of them can give a peripheral community member a very biased idea of the community. In this respect, face-to-face meetings are of high relevance to scientific communities; among other things, they provide possibilities for peripheral members to gain community knowledge and interact with other members. Such meetings make communication easier, especially in scientific communities like CSCL which consist of members with different native languages complicating the distribution of results and effective communication (Kienle & Wessner, 2005). Workshops and conferences are used to foster researchers’ communication and learning about the findings and approaches of others, but also to integrate newcomers. Such events bring participants together and allow them to focus on learning activities and on community building, and can be called encapsulation. Although encapsulation is a widely used strategy in different contexts (Levine & Moreland, 1991), it can be organized in different ways:
workshops usually allow for more one-on-one interaction, while (larger) conferences usually focus on other types of communication. However, it is unclear how one-on-one interaction is related to researchers’ learning in the scientific community. Access to community knowledge, especially to knowledge about other members, seems in particular to be also very relevant and it might be helpful to foster it during encapsulation events.

In this study we investigate factors which influence the integration and learning of newcomers in a scientific community. We adopt a social network perspective on learning and integration by focusing on the building of collaborative relationships between community participants as the visible and desired consequence of integration and learning. The social network approach offers two different ways to look at the building of collaborative relationships. First, we can look at individual persons and how successful they were in building new relationships; second, we can also look at all individual relations between two community members and what factors influence the probability of a random relation to become a collaborative relationship:

1. Persons-related RQ: To what extent can support for community knowledge and participants’ level of participation predict a participant’s number of newly built collaborative relationships?
2. Relation-related RQ: To what extent can time spent interacting, support for community knowledge, and a persons’ level of community participation predict the development of a new collaborative relation?

Method

Study Design, Context and Participants

The study was planned with a quasi-experimental design in which the factor support for community knowledge was varied across different workshops and the factor community participation level varied naturally among participants. This design was implemented at a small conference organized by a European research community which is closely related to the CSCL community. The aim of the conference was to bring together researchers form multiple disciplines who worked in the field of technology-enhanced learning. The conference was organized in a non-standard way and consisted of 8 workshops and a doctoral consortium. The workshops were organized in two series of 4 workshops taking place in parallel. The number of participants for each workshop varied between 14 and 22 persons. Each workshop lasted one and a half days while the doctoral consortium lasted for the whole 4 days. The conference took place at a hotel in a remote place and all workshop organizers and participants also lived in this hotel during the time. Participants were selected based on a review process of papers they had submitted.

All together, 152 persons participated in the conference. The majority of them came from European countries, but there were also participants from many other countries. For this study, only persons who had participated in one of the workshops were taken into account. Persons who were only involved in the doctoral consortium were left out because of their special conditions during the conference. Also the data of participants who had missing values or whose answer patterns made their credibility questionable was left out. For the analysis of the two research questions, further constraints (explained below) resulted in different sample sizes. For RQ 1, the sample consisted of 89 participants. For RQ 2, the sample consisted of 742 dyadic relations in which 125 persons were involved.

Data Collection and Instruments

Data about the participants’ interaction during the conference, their collaborations with other participants beforehand, and their intended collaborations after participation to a workshop were measured using two different tools: RFID devices and social network questionnaires. Further information about the participants was taken from the application form for the conference. Additionally, a bibliographic analysis of co-authorships listed in Google Scholar was performed.

Tracking face-to-face proximity with RFID devices

During the conference, each participant wore an RFID device, developed by the SocioPatterns collaboration (http://www.sociopatterns.org) that was integrated into the name badge. The devices engage in bidirectional low-power radio communication. As the human body acts as a shield for the used radio frequency, and as the badges are worn on the chest, badges can exchange radio packets only when the individuals wearing them face each other at close range (about 1 to 1.5 m). The measuring infrastructure can capture that there was a close face-to-face proximity between two individuals with a temporal resolution of 20 seconds, and gives therefore access to the amount of time that two participants spent together (see Cattuto et al. (2010) for a detailed description of the infrastructure). Only two participants of the conference declined to wear the devices.

Social network questionnaires

After each workshop, participants were asked to fill in a social network questionnaire about their relations within the workshop: they were given a list of all workshop participants’ names and were asked to indicate with whom they had collaborated already before the conference and with whom they had found potential for future
collaboration. As some conference attendees participated in two workshops, 160 questionnaires were handed out, from which 150 were returned.

**Bibliographic analysis on publications listed in Google Scholar**

About 1.5 years after the conference had taken place, we performed a Google Scholar search to have an indicator if the subjective indications in the social network questionnaire had lead to an objective measurable collaborative outcome.

To identify if two participants had collaborated with each other before but forgot to indicate that in the social network questionnaire, we performed a search for joint publications before the conference. We restricted this search to papers published in 2010 or earlier (query term: “as_yhi=2010”). For each author’s name, at least two variations were included (query term: “(author: Doe J OR author: Doe John”)”. For each possible pair of workshop participants, a separate search was conducted by combining them, for example the query term “(author: Doe J OR author: Doe John) AND (author: Smith S OR author: Smith Samantha)”.

To identify joint papers after the workshop, we performed a second Google Scholar search similar to the first one, but restricted to papers published in 2011 or later (query term: as_ylo=2011).

**Dependent and Independent Variables**

**Level of prior community participation**

Each participant was allocated to a community participation category, either as a peripheral community member or as an active community member. The allocation was based on the participant’s previous participation in the scientific community (similar to the allocation criteria used by Kienle and Wessner, 2006). Those participants who fulfilled at least one of the following criteria were assigned as active members: they attended the previous conference 1.5 years ago, they were organizers of one of the workshops at the present conference, or they were members in one of the boards of the scientific community. Peripheral members fulfilled none of these criteria.

**Access to community knowledge**

In 3 of the 8 workshops, support for community knowledge was given with the aim to foster collaboration between participants. Support for community knowledge was implemented as knowledge about the individual community members. A brochure with the following information about all participants was compiled: their name, picture, contact information, affiliations, background, research interests, and exemplary publications. This brochure was handed out to the participants at the beginning of the workshop without further instructions.

**Number of newly built relationships**

To answer RQ1, we computed for each person the number of new collaboration partners. Participants with whom the person had already collaborated before the conference were not included as new collaborative relationships; this was either indicated by the person, the partner, or both of them in the questionnaire or by previous collaborative publication between the person and the partner found in the bibliographic analysis. Three different types of newly built relationships were computed based on different measures: the number of new interaction relationships, the number of new subjective collaborative relationships, and the number of new objective relationships.

**Number of new interactive relationships.** The number of new interaction partners was computed for all persons from their face-to-face time with other workshop participants (recorded by the RFID infrastructure). All participants of the workshop with whom a person had interacted during the workshop (but not collaborated before the workshop) were counted.

**Number of new subjective collaborative relationships.** This outcome variable was computed from the social network questionnaire. We counted the number of participants with whom the particular person had indicated to have identified potential for future collaboration, but only if the respective participant had also indicated to have identified potential for future collaboration with this person.

The sample for RQ 1, in which the number of newly built relationships was the outcome variable, included only a sub-sample of the conference participants. Only those participants were taken into account who had indicated in the questionnaire to have not collaborated before with at least 10 of the other workshop participants. This constraint was made because of statistical reasons: A person who had only the chance to build a new relationship with 2 other participants of the workshop (because he or she had collaborated with all others before) would bias the results because this person might have built more new relationships if possible.

**Interaction time**

Interaction time was computed for each dyad of participants in a workshop. The time was taken from the RFID-based measurements. For the relations within one of the 4 workshops in the first part of the conference, only the interaction time from the beginning of the conference to the end of the workshop was taken into account. For the
relations within one of those 4 workshops that took place in the second part of the conference, only the interaction time from the beginning of the workshop until the end of the conference was taken into account. Observed interaction times between pairs of participants ranged from 0 seconds to 75 minutes.

**Development of a new collaborative relation**

To answer RQ 2, we analyzed all possible new relation between two participants of the same workshop who had not collaborated with each other before. Therefore, relations were exclude if one or both persons in the relation had indicated in the social network questionnaire to have collaborated with each other already before the workshop or if a previous joint publication was identified in the bibliographical analysis. Three different types of possible new relations between two participants were computed based on different measures; all three types were dichotomous variables: the development of an interactive relationship, the development of a new subjective collaborative relationship, and the development of a new objective relationship.

- **Development of an interactive relationship.** If a face-to-face interaction had been recorded with the RFID devices measurement between two persons who formed a possible new relation, this was taken as a newly developed interactive relationship (= 1). If no face-to-face interaction was recorded, the relation was treated as one without a newly developed interactive relationship (= 0).
- **Development of a new subjective collaborative relationship.** If both persons in a relation had indicated in the social network questionnaire to have identified potential for future collaboration after the workshop, this was taken as a newly developed relation (=1). If none or only one of the two persons in the relationship had indicated to have identified potential for future collaboration, this was seen as no relationship (= 0).
- **Development of a new objective relationship.** Taking the results of the Google Scholar search of co-authorships, we looked for each possible pair of two workshop participants, if they had published a joint paper after the conference. If a joint paper was found this was taken as newly developed objective relationship (= 1). If no jointly published paper was found, this was taken as no newly developed objective relationship (= 0).

**Data Analysis**

To answer RQ 1, the data was analyzed using hierarchical linear modeling and applying a HLM model (using the software HLM 6.08 by Raudenbush, Bryk, and Congdon, 2004) as each person was nested within a workshop. HLM allowed us to control for random effects caused by differences in the workshop which could not be controlled otherwise. As the dependent variable was Poisson-distributed, a logarithmic link function was used.

For RQ 2, instead of looking at the outcome of individual persons, we looked at the relations between two persons. In this way, data of the same person appeared several times in the dataset and the relation-data was cross-classified within two persons. To control for person-specific effects we used an HCM2 model (using the software HLM 6.08). Each relation appeared twice in the dataset, so each of both persons in a relation was identified as cross-classification variable twice. After carrying out the analysis, we divided the degrees of freedom in half again to deliver appropriate results for the real sample. Additionally, it was necessary to apply a logarithmic link function to the model to account for the binomial (dichotomous) distribution of the outcome variable. Additionally, a $\chi^2$-test was applied to investigate the development of new objective collaborative relationships because the percentage of identified new objective collaborative relationships was too small to apply HCM here as well. For all analyses, the significance level was set to .05.

**Results**

**Person-related RQ1: Influences on a person’s number of new interactive relationships.** No significant effects of community participation level, support for community knowledge or of an interaction of those variables on the participants’ number of interaction partners during the workshops could be identified.

**Person-related RQ1: Influences on a person’s number of new subjective collaborative relationships.** The population-average model to predict a participant’s number of new collaboration partners after the workshop revealed the following results: An average participant who was an active member of the scientific community and had not received support for community knowledge acquired on average 3.4 new collaboration partners ($\gamma = 1.222$, $SE = 0.142$, $t(7) = 8.622$, $p < .01$). Peripheral community members, in contrast, acquired on average only 2.0 new collaboration partners ($\gamma = -0.319$, $SE = 0.153$, $t(86) = -2.087$, $p = .04$). However, peripheral community members who received support for community knowledge were able to find on average 3.6 new collaboration partners ($\gamma = 0.437$, $SE = 0.213$, $t(86) = 2.051$, $p = .04$). No significant effect of support for community knowledge was found on active community member’s number of new subjective collaborative relationships.

**Relation-related RQ 2: Influences on the development of a new interactive relationship.** The HCM model showed that for two random participants of a workshop who did not receive support for community knowledge, the average probability to develop a new interactive relationship was 30.3% ($\theta = 0.832$, $SE = 0.150$, $t(698) = -5.541$, $p < .01$). This probability varied significantly across individual workshop participants. If they
received support for community knowledge in their workshop, their probability to develop a new interactive relationship was significantly reduced by 10.3 % ($\theta = -0.551$, $SE = 0.235$, $t(698) = -2.340$, $p < .01$). No significant effect of community participation level was found on the probability to develop a new interactive relationship.

**Relation-related RQ 2: Influences on the development of a new subjective collaborative relationship.**

We identified the following HCM model to predict the probability of a random relation between two workshop participants to develop a new subjective collaborative relationship after they had participated in the same workshop: If both persons were active community members and they had not interacted face-to-face with each other, their probability to mutually identify potential for future collaboration was on average 27.9 % ($\theta = -0.948$, $SE = 0.162$, $t(697.5) = -5.861$, $p < .01$). However, this varied significantly across individual persons. For every minute two persons spent interacting, the probability for them to develop a new subjective collaborative relationship increased significantly ($\theta$ (in seconds) = 0.001, $SE < 0.001$, $t(697.5) = 2.057$, $p = .03$). For example, a relation in which the two persons had spent 1 minute interacting with each other, had an increased probability to develop into a new subjective collaborative relationship by 1.2 % compared to a relation in which no direct interaction was measured. If one of the two persons was a peripheral member, the probability for the relation to develop into a new subjective collaborative relationship was significantly reduced by 6.7 % ($\theta = -0.361$, $SE = 0.197$, $t(697.5) = -1.836$, $p < .04$). No significant effect was found for support for community knowledge.

**Relation-related RQ 2: Influences on the development of a new objective collaborative relationship.**

There was a significant difference between the relations of participants who had received support for community knowledge and those relations in which no support was available regarding the development of new objective collaborative relationships ($\chi^2(1,699) = 21.11$, $p < .01$) favoring those who had received the means of support. No significant differences was identified between peripheral and active members. Also no significant difference was found between relations in which persons had directly interacted with each other and those without direct interaction during the workshop. No difference was found as well between those relations in which a subjective collaborative relationship was reported and those in which no such relationship was reported.

**Conclusion**

Encapsulation events like conferences and workshops are an important means of scientific communities to bring their members together and foster integration of new members. Therefore, peripheral members of scientific communities are an important group of participants. Although peripheral community members seem not to differ from active community participants regarding the number of interactions they have during a conference or the probability for interacting with a random other participant, they seem disadvantaged regarding the outcome of these interactions. Peripheral members’ chances to develop subjective collaborative relationships are lower resulting in a lower number of new subjective relationships after the conference. However, we do not find these results in the more objective measurement on real collaborative outcomes, but this might be due to the fact that measuring joint papers 1.5 years after the conference is still a bit early. The process to plan a joint research project, collect and analyze data, write a paper and successfully publish it takes usually a long time and it seems advisable to rerun a Google Scholar search at a later point in time.

In this study we assumed that the disadvantages of peripheral community members could be based on their lack of community knowledge. Therefore, we supported the participants of some of the workshops with explicit community knowledge. Providing participants with this support reduced the chances that a participant interacted with a random person, but did not reduce the number of their interaction partners during the conference. This can be seen as a hint that participants who had the community knowledge support were able to identify more precisely who would be a relevant interaction partner and enabled participants to use their few time more efficiently by focusing on interactions with those participants. This is in line with the result that the longer two persons spent interacting with each other, the higher their chances to build a new collaborative subjective relationship. However, our community knowledge support was not directly related to the chances for building a new subjective collaborative relationship or the number of newly built subjective collaborative relationships. Looking at the few results of joint papers we identified so far, we see a clear relation between receiving community knowledge support and successful objective collaborative relationships. Taking all these results together, we can assume that community knowledge support enables participants of scientific meetings, especially those who still possess little community knowledge, to identify promising partners for collaboration more efficiently and to focus on longer and more effective interactions with those, which increases the chances of plans for future collaboration to become real and visible collaborations.

Additionally to the limitations of the objective collaborative relations measurement through our Google Scholar search, also some technical problems with the RFID devices have to be reported: we can not claim to really have measured all face-to-face interactions between workshop participants because some participants lost or forgot their RFID devices for some time or the devices run out of battery. Also, the name badges in which the devices were integrated flipped quite often, so the body of the participant wearing the badge shielded the radio signals. Although we surely missed some interactions between participants, the results show that RFID devices
can work as a promising new method to measure collaboration. But researchers who want to use this technique in the future can surely improve its use by taking care of the reported problems.

Aside from technical questions, the results of this study confirm the importance of shared community knowledge for collaboration (Cannon-Bowers & Salas, 2001). The relation between community knowledge has so far been studied on the cognitive level, but our findings confirm them also from a social network perspective with visible outcomes. Acquiring shared community knowledge seems important for the building of new collaborations and the integration of peripheral members into a scientific community. Additionally, our results confirm finding form previous studies about tactics which can be used by communities of practice to integrate and support their newcomers and peripheral members (Eberle, Stegmann, & Fischer, 2012). This study had identified explicit access to community knowledge as an important means to foster the learning of peripheral members and foster their collaboration with more active members. This finding seems to be transferable to scientific communities as a special type of community of practice.

We can also draw some practical implications from this study for future CSCL conference and workshop organizers as well as for their participants: Organizers can support the integration of peripheral participants by providing explicit access to information about the community and its members and by planning for enough time for their participants to interact on a one-on-one basis with each other. Participants, on the other hand, can positively influence their workshop and conference experiences by informing themselves beforehand about the other participants and their backgrounds and by focusing on longer one-on-one interactions with other participants.

References


