Check for updates

Initiating scientific collaborations across career levels and disciplines – a network analysis on behavioral data

Julia Eberle[1,2] · Karsten Stegmann[1] · Alain Barrat[3,4] · Frank Fischer[1] ·
Kristine Lund[5]

Abstract

Collaborations are essential in research, especially in answering increasingly complex questions that require integrating knowledge from different disciplines and that engage multiple stakeholders. Fostering such collaboration between newcomers and established researchers helps keep scientific communities alive while opening the way to innovation. But this is a challenge for scientific communities, especially as little is known about the onset of such collaborations. Prior social network research suggests that face-to-face interaction at scientific events as well as both network-driven selection patterns (reciprocity and transitivity) and patterns of active selection of specific others (homophily / heterophily) may be important. Learning science research implies, moreover, that selecting appropriate collaboration partners may require group awareness. In a field study at two scientific events on technology-enhanced learning (Alpine Rendez-Vous 2011 and 2013) including $N = 5736$ relations between 287 researchers, we investigated how researchers selected future collaboration partners, looking specifically at the role of career level, disciplinary background, and selection patterns. Face-to-face contact was measured using RFID devices. Additionally, a group awareness intervention was experimentally varied. Data was analyzed using RSiena and meta-analyses. The results showed that transitivity, reciprocity and contact duration are relevant for the identification of new potential collaboration partners. PhD students were less often chosen as new potential collaboration partners, and researchers with a background in Information Technology selected fewer new potential collaboration partners. However, group awareness support balanced this disciplinary difference. Theoretical, methodological, and practical implications of these findings are discussed.

✉ Julia Eberle
    Julia.eberle@rub.de

[1] Department of Psychology, Ludwig-Maximilians-Universität München, Leopoldstr. 13, 80802 München, Germany

[2] Faculty of Philosophy and Educational Research, Ruhr University Bochum, Universitätsstr. 150, 44780 Bochum, Germany

[3] CNRS & Centre de Physique Théoretique, Turing Center for Living Systems, Aix Marseille Université & Université de Toulon, Campus de Luminy, Case 907, 13288 Marseille, France

[4] Tokyo Tech World Research Hub Initiative (WRHI), Tokyo Institute of Technology, Tokyo, Japan

[5] Laboratoire ICAR, ENS de Lyon, CNRS, Université Lumière Lyon 2, 5, parvis René Descartes, Lyon 69007, France

Collaboration and integration of newcomers in scientific communities of practice

While much research has investigated how to engage learners in collaborative scientific inquiry, scientific literature about how real researchers do so is quite scarce. And although a significant body of research is dedicated to the investigation of scientific communities, a good part of it centres on the outcomes of scientific collaboration by providing bibliometric analyses that focus on co-authorship or citations in conference proceedings or journal papers (e.g. Lee et al., 2012). Such research can only give a distal and static perspective on collaboration in a scientific community. Other work however, examines whether collaboration patterns change across disciplines (Clemente-Gallardo et al., 2019), investigates building a new research community based on integration expertise (Bammer, 2013), or even considers the extent to which humanities and the sciences can be integrated into a new shared consilience framework (Slingerland & Collard, 2012). These dynamic perspectives on collaboration give a more comprehensive view of collaborative processes, and work on why (e.g. Maienschein, 1993) and when (Birnholtz, 2007) researchers collaborate contributes to understanding researcher motivations. On the other hand, a better understanding of how to foster new collaborations is needed, especially between newcomers and oldtimers with different disciplinary expertise.

Indeed, the relative lack of knowledge on how researcher collaborations begin is quite surprising for several reasons. First, the collaborative imperative has been getting stronger in many research fields over the recent decades, reflecting the higher impact of collaborative research, the necessity for sharing research resources, and new technological opportunities that simplify collaboration (Bozeman & Boardman, 2014). In addition, complex research questions as well as funding agencies and policy makers specifically demand research collaborations that go beyond an individual discipline, tapping into inter-, trans-, and multidisciplinary collaborations, despite researchers' complaints about these types of collaborations (cf. Rylance, 2015). Therefore, we need knowledge regarding researcher collaboration in general, but we also need knowledge on a first, oft-ignored step: how is such collaboration instigated and supported in informal settings?

Second, scientific communities are by nature constantly evolving and changing. These changes are related to topics, questions, and methods but also to people that form the community, e.g., as newcomers show up and well established oldtimers leave. Both aspects are frequently closely linked, and coping with these continuous changes is vital for scientific communities, requiring additional knowledge regarding the way collaboration between junior and experienced researchers is initiated and can be supported.

The CSCL community is a good example for observing such changes as members make the process quite explicit and participatory. Currently, there is a focus on the content side, for example meta-reviews about the methods that the community has used in the past (Jeong et al., 2014; Xia & Borge, 2019) and an open conversation about directions the community's research may take in the future (e.g. Wise & Schwarz, 2017; Rummel, 2018; Tchounikine, 2019). Several years earlier, Kienle and Wessner (2006) analysed social aspects of the community by investigating participation in CSCL conferences. They found that the community struggled with integrating newcomers and attracting a regular body of members. Mastering these challenges is essential, as a stable membership maintains the scientific community and

insures scientific progress through the continuous contributions of the members. In addition, being part of a relevant scientific community and collaborating with other researchers is crucial for individuals' occupational and personal scientific development. That said, we know only little about how scientific communities are able to handle these challenges effectively and stimulate fruitful collaborations among their diverse members.

In order to understand how scientific communities can foster newcomer integration within the development of collaborations that cross boundaries, we investigate the onset of scientific collaborations and how scientific communities are able to support their members in finding collaboration partners. Specifically, we explore how participants at face-to-face meetings in the interdisciplinary field of technology-enhanced learning select new collaboration partners and what may be influential factors in this process. Face-to-face meetings at conferences and workshops seem to be suitable settings for studying these phenomena. Such meetings provide regular opportunities for newcomers in a given field to become integrated into the scientific community and for members with different disciplinary backgrounds to come together and start new collaborations among all participating researchers (Kienle & Wessner, 2005).

In this study, we focus on four research questions that aim to better understand what aspects are relevant for the onset of scientific collaborations at a scientific event, investigating specifically the process of selecting new collaboration partners and the situation of specific groups within a scientific community.

RQ1: To what extent does face-to-face contact explain the initiation of new research collaborations?

RQ2: To what extent do homophily, prestige, and experience regarding career level and disciplinary background as person-driven selection patterns drive the initiation of new research collaborations?

RQ3: To what extent do reciprocity and transitivity as network-driven selection patterns predict the initiation of new research collaborations?

RQ4: How does group awareness support influence selection patterns in the initiation of new research collaborations?

We will first explore theoretical assumptions and previous findings about scientific communities, the impact of face-to-face interaction, types of selection patterns in social contexts, and the role of group awareness. We will then report on a study that investigates the onset of scientific collaboration during two scientific events for researchers on technology-enhanced learning (the Alpine Rendez-Vous) based on the four research questions, and discuss what can be learnt from our findings.

Potential influential factors on selection processes in scientific communities

According to Kienle and Wessner (2005), a scientific community is a specific form of community of practice that consists of a heterogeneous, often interdisciplinary group of researchers. Its members are usually geographically distributed and sometimes – as in the CSCL community or the broader technology-enhanced learning community – have backgrounds in different disciplines and scientific cultures, resulting in the use of different research methods and scientific theories (Kienle & Wessner, 2006). What brings the members together is a joint field of research interests. Therefore, scientific communities benefit from the integration of new members who bring in complementary ideas and new

perspectives. Literature on communities of practice suggests that the integration of new members is facilitated by legitimate peripheral participation (Lave, 1991) and, in particular, by establishing participation support structures, such as providing opportunities for participation (Eberle, Stegmann, & Fischer, 2014). In interdisciplinary scientific communities, successful collaboration among the community members is another important factor for community cohesion and development. For both newcomers and regular members of a scientific community, participation in scientific events can, therefore, be expected to be beneficial. In the following sections, we will explore mechanisms for this based on previous literature.

Interacting with others: face-to-face contact at scientific events

To understand the onset of interdisciplinary scientific collaboration and newcomer integration in scientific communities, scientific events are a good starting point as they are major opportunities for researchers to get in contact with each other. It is, thus, important to take a closer look at *face-to-face contact between researchers* during scientific events. While research in interactional linguistics illustrates how human interaction is co-constructed through phenomena such as gaze, speech prosody, gestures and general body movement (e.g., Goodwin, 2000) on the micro-level, human interaction is also based on several robust characteristics on a broader level, such as physical and digital proximity, social support, and community belonging. The influence of these characteristics partially depends on the context in which interaction takes place (Isella et al., 2011).

Several studies have investigated face-to-face contact in conference settings. Kibanov et al. (2015) found that researchers who were linked on ResearchGate prior to a conference visit were very likely to interact with each other in the face-to-face situation at the conference. Face-to-face contact between researchers who were not previously linked increased the likelihood for a later link on ResearchGate. However, the length of their face-to-face contact was important: While short contacts had no effect, the likelihood of a link increased with longer face-to-face contact. Taking these findings in relation to the onset of scientific collaborations leads to RQ1.

RQ1: To what extent does face-to-face contact explain the initiation of new research collaborations?

When looking more closely, researchers' face-to-face contact patterns differ between different groups of researchers. Atzmueller et al. (2012) found that face-to-face contact at scientific conferences differs based on participants' career status. The average number of face-to-face contact partners was higher for more experienced researchers than for PhD students and undergraduate students. The weighted eigenvalue centrality, which took numbers of contact partners and the duration of face-to-face contact into account, increased for each career level with professors having the highest values. Furthermore, Barrat et al. (2010) found that researchers' seniority in the field predicted with whom they had face-to-face contact at scientific events. Researchers tended to have the most face-to-face contacts with others whose level of research seniority was similar to their own, meaning that senior researchers stuck together, and junior researchers stuck together.

Another interesting finding on interdisciplinary face-to-face contacts in the study by Atzmueller et al. (2012) was that researchers tended not to stick with researchers of their own core field but used the scientific event also for interactions with researchers of other disciplines or with different research foci than their own. In addition, Kibanov et al. (2015)

found differences in how researchers with different disciplinary backgrounds interacted, specifically that computer scientists tended to socialize more online than in the face-to-face setting compared to non-computer scientists.

Beyond these findings around face-to-face contacts, social network research suggests that the selection of partners for long-term relationships, such as friendship or scientific collaborations, follows several common patterns (Baerveldt et al., 2010). Social selection patterns have been studied in various contexts, and the literature reports a variety of observed patterns. Some selection patterns are based on the opportunities provided by the given social network structure surrounding the selecting person (network-driven patterns), while others focus on personal preferences for specific persons (person-driven patterns). Reciprocity and transitivity as network-driven patterns and homophily/heterophily as a person-driven pattern seem to be among the most important selection patterns (Baerveldt et al., 2010). Note that network-driven patterns and homophily/heterophily effects are not mutually exclusive but can co-exist (Aiello et al., 2010).

Selecting who is accessible: reciprocity and transitivity as network driven selection patterns

Network-driven selection patterns are based on the assumption that the way in which persons choose relations is strongly influenced by the (local) structure of their social network (Baerveldt et al., 2010). This means that if a person is seeking a new relation with someone (e.g., as a friend, for receiving advice, or as a collaboration partner), persons who are easily accessible in one's network are more likely to be chosen compared to persons who are harder to access. Persons who are easily accessible in one's network can be reached through patterns of reciprocity and transitivity.

Reciprocity is one of the simplest network-driven selection patterns: If there is already a connection between two individuals, in which person A has chosen person B (e.g., as a friend or as someone to ask for advice) while person B did *not* chose person A in the past, it is very likely that they will balance their relationship in the future. This can either mean that person B also chooses person A or that their relation will dissolve so that reciprocity is reached. In the context of a face-to-face meeting of a scientific community, researcher A may have read a paper or attended a talk of researcher B, while researcher B is not yet aware of researcher A's work. If researcher A approaches researcher B during the event and they discuss their shared interest, researcher B may also become interested in researcher A's work, and this may be the beginning of a collaboration.

Transitivity is also driven by proximity in the network: Two individuals A and B, who both are connected to a third person C, are likely to build a relation as well. In social network terms, they are building a "transitive triple" or "closing a triangle". Reasons for such transitivity patterns are manifold, e.g., person C can easily introduce A and B, or A and B may share a common interest or activity that was the initial reason for their connection with C. C might also be interested in setting up a connection between A and B to stabilize the relation to both (Baerveldt et al., 2010). This can happen at scientific events, when e.g. professor A and their PhD student B have a coffee between talks and researcher C comes along. As professor A is aware that both B and C share common research interests, A may introduce the two. Such behaviour has indeed been documented for face-to-face contacts at conferences, where certain researchers took the roles of bridges and ambassadors between other researchers (Atzmueller et al., 2012).

Strong network-driven selection patterns at a scientific event are most beneficial for participants with many existing contacts that provide vast opportunities for finding new collaboration partners. Researchers with few previous contacts, as it is typical for newcomers, can benefit much less from network-driven selection patterns. Consequently, we therefore formulate RQ2 accordingly.

RQ2: To what extent do reciprocity and transitivity as network-driven selection patterns predict the initiation of new research collaborations?

Selecting specific others: homophily and heterophily

In contrast to selection patterns that are driven by the social setting around people, homophily is a broadly studied pattern that describes the selection of specific others beyond their mere availability (Baerveldt et al., 2010). Homophily describes the tendency of people to seek contact with others who are similar to them (McPherson et al., 2001), which stands in contrast to heterophilous selection, which is the tendency to select dissimilar others. These patterns have been studied in regard to very different types of similarity/dissimilarity, such as gender, religion, age, education, occupation, social class, behaviour, attitudes, beliefs, abilities, and aspirations (McPherson et al., 2001).

In general, a greater tendency for homophilous selection patterns has been observed as relationships between two similar persons have several advantages, e.g., they require relatively little effort, expectations of each other are comparably clear, and the interaction is easily perceived as rewarding due to shared values (Coleman, 1988; Lin, 2001; Byrne & Clore, 1970; Lott & Lott, 1974). For the identification of scientific collaboration partners, homophilous selection could mean that researcher A prefers to work with researcher B, who has the same disciplinary background as they share similar views on how research is done and what is important to look at. Barrat et al. (2010) have found homophilous tendencies for face-to-face contacts at conferences based on similarity in career levels.

Heterophilous selection, however, has other advantages that may outweigh its effort requirement, especially in the context of scientific collaboration. According to social capital theory, heterophilous selection allows for gaining new resources (Coleman, 1988; Lin, 2001), such as access to new perspectives and complementary expertise. Interdisciplinary scientific communities seem to be built around the assumption that heterophilous collaboration between researchers with different backgrounds is desirable. Heterophilous tendencies have also been observed in relation to prestige effects, when someone seeks a relation to a person with higher status (Lin, 2001). This may also be the case in scientific communities. For example, well-known and distinguished professors are more generally favored as collaboration partners than unknown PhD students, and researchers who have the desired complementary competence for an interdisciplinary project are favored by project leaders.

Knowing about others: group awareness

All so far described face-to-face contact patterns and selection patterns take a macrolook at human behaviour without taking psychological aspects of the acting individuals into account. However, psychological aspects underlie each of these patterns at the individual level. For example, the choice an individual researcher makes in selecting their collaboration partners in a homophilous or heterophilous way tends to depend on whether the researcher perceives another researcher as similar/dissimilar or prestigious. This perception, in turn, depends on the knowledge the researcher has about the other person.

Some knowledge about others is easy to acquire as relevant information is simply visible in face-to-face situations, e.g., the clothes a person wears may indicate a certain life style, status, or even disciplinary background, while other information is invisible and harder to access (Baerveldt et al., 2010, p.286). It is mostly invisible information that is needed to select a scientific collaboration partner, specifically a person's specific expertise and knowledge, as these are crucial for scientific collaboration (Cannon-Bowers & Salas, 2001; Wegner, 1987). While a huge body of research, e.g., on shared / team mental models, group awareness, transactive memory, and grounding, has looked into important factors for successful collaboration in existing small groups that work or learn collaboratively in specific situations, relatively little research has focussed on aspects relevant for the selection of collaboration partners within a broader community before actual collaboration starts.

To select the right collaboration partner, a researcher needs awareness and knowledge about the other researchers in the scientific community, specifically about their personal and social resources (Lin, 2001). A researcher's personal resources may include their area of expertise, current research interests, and knowledge. This means that cognitive group awareness (Janssen & Bodemer, 2013) of the scientific community's members is necessary. Cognitive group awareness refers to individual cognition about the group's or group members' situational and / or stable characteristics, the latter including interest, expertise, or beliefs (Bodemer et al., 2018). A researcher's social resources may include their professional network, access they have to other experts, and their standing in the scientific community, requiring what could be seen as a form of social group awareness (Janssen & Bodemer, 2013).

Newcomers to a scientific community can be expected to possess little knowledge about the new community. Acquiring more of this knowledge is necessary for them to contribute in a meaningful and accepted way (Levine & Moreland, 2013). Although knowledge about individual researchers is in general easy to acquire as most of them present their bios and publications openly on their websites, it is hard to retrieve information about the structure of the community and identify the 'important' people in the community or those who could be specifically relevant for one's own research within the crowd of participants, especially in a conference setting. As Baerveldt et al. (2010) have shown in friendship networks, students with little information about their peers are less active in initiating new friendships. They seem also more likely to use rather passive selection strategies, such as transitivity (forming new friendships with friends of their friends), instead of initiating new friendships with peers who might be a good fit to them, for example regarding norms and values.

Consequently, newcomers in scientific communities may be disadvantaged in finding new collaboration partners as they may not only use fewer active strategies for initiating new collaborations with other researchers, but they may also benefit less from network-based selection patterns as they are not yet linked to many other researchers. As researchers tend to be involved in several scientific communities at the same time, newcomers to scientific communities can be found on all career levels. Therefore, researchers are used to switching roles, often from expert in one scientific community to newcomer in another scientific community (Kienle & Wessner, 2005). However, PhD students are the most common form of newcomers in scientific communities, and they are probably in the most disadvantaged position for finding collaboration partners. In contrast to more experienced researchers, PhD students not only lack group awareness of the specific new scientific community but also knowledge regarding scientific collaboration and how to initiate it in general. They are, moreover, often seen as less prestigious to interact with, as they are likely to have only few personal and social resources relevant for scientific collaboration. This makes PhD students also less attractive collaboration partners from the perspective

of other researchers. However, even experienced members of a scientific community cannot be assumed to be totally aware of all other participants' expertise and research during a face-to-face event, so lack of group awareness within a scientific community is very likely a general problem, although it is more severe for some people than for others.

Summing up the previous argumentation, we still know very little about person-driven selection patterns that may underlie the initiation of collaborations between researchers during scientific face-to-face meetings and how subgroups within an interdisciplinary scientific community may differ in their selection patterns. This leads to RQ3.

RQ3: To what extent do homophily, prestige, and experience regarding career level and disciplinary background as person-driven selection patterns drive the initiation of new research collaborations?

Going a step further, it seems promising to support researchers in gaining a better group awareness of the other participants at a scientific event to help them make informed decisions with whom to get in touch instead of relying on passive network-driven patterns. Tools to support different types of group awareness have already been explored. However, this research has so far focussed on supporting small groups in their coordination and collaboration in computer-supported work or learning settings (Schmidt, 2002; Janssen & Bodemer, 2013). These groups worked on a concrete collaborative task, often mediated by computers, and therefore most group awareness tools targeted (socio-) behavioural, i.e., situational, aspects of the group members, such as their knowledge of the specific task or their participation throughout the collaboration process (Janssen & Bodemer, 2013; Bodemer et al., 2018). However, if the context is the onset of collaborations in a large scientific community in a face-to-face setting with nonspecific outcomes, different support is needed. For example, support can target cognitive group awareness on more stable characteristics of researchers that are relevant in the scientific community, such as others' general area of expertise and interests, and social group awareness of stable aspects within the scientific community, such as others' co-authors or PhD advisors. This need has already been identified, and approaches have been made to address it. Alshareef et al. (2018) have proposed an algorithm that can be used to recommend researchers to each other as potential collaboration partners based on their similarities as evidenced by previous publications. However, as we have pointed out before, this approach targets homophilous collaboration only and does not provide support for heterophilous collaborations in interdisciplinary scientific communities. Another group of researchers has gone a step further and designed a prototype of a group awareness tool for conference participants that included, next to other features, self-reported research interests of participants, recommended persons to interact with, and even helped to initiate the first contact via an online interface (Windhager et al., 2014; Zenk et al., 2014). Their qualitative evaluation of the "Conference Explorer" tool showed that participants were highly satisfied. This finding seems to indicate that there is not only a theoretical and practical need for group awareness support at scientific events, but that participants also use such a support instrument in a way that they find valuable. However, so far, the behavioral consequences of group awareness support and its effects on the identification of collaboration partners has not yet been explored. We will, therefore, investigate experimentally RQ4.

RQ4: How does group awareness support influence selection patterns in the initiation of new research collaborations?

Method

Study context and participants

The study was conducted at the 2011 and 2013 Alpine Rendez-Vous, two scientific events that aimed at bringing together researchers from multiple disciplines working on technology-enhanced learning (TEL) to foster community building and scientific progress in the field. The events were part of a series of four meetings, first organized by the KALEIDO-SCOPE Network of Excellence (1) and later by the STELLAR Network of Excellence (2), two initiatives funded by the European Union to foster European TEL research. The Alpine Rendez-Vous 2013 was promoted by TELEARC and EATEL, two associations that resulted from the two previous Networks of Excellence. Both, Alpine Rendez-Vous 2011 and 2013, lead to a shared outcome of many participants – two books on grand challenge problems that TEL research was supposed to tackle in the upcoming years (3).

Both Alpine Rendez-Vous meetings were organized in a similar way. The events were deliberately located in a big hotel at a remote place in the French Alps to avoid external distractions and to provide many opportunities for networking among the participants. Both events consisted of a set of independent parallel workshops on specific topics under the scope of TEL, each workshop lasting one and a half days. Half of the workshops took place in the first part of the event, followed by a community event for all event participants in the evening of the second day. The second group of workshops took place after the community event. While each workshop had been selected in a competitive process and was organized independently, there was a general schedule for all workshops at the event to synchronize starting time, breaks, and end time. Between the workshop time slots, all present participants had breakfast, lunch, dinner, and coffee breaks together, as well as a long afternoon break, which allowed for social activities and networking within and across the workshops.

The Alpine Rendez-Vous 2011 (4) consisted of four workshops in the first half and four workshops in the second half of the event. The Alpine Rendez-Vous 2013 comprised five workshops in the first half and five workshops in the second half of the event (5).

Altogether, 136 persons participated in workshops at the Alpine Rendez-Vous 2011 and 152 persons participated in workshops at the Alpine Rendez-Vous 2013, leading to a total of 288 participants. The majority of participants were from European countries, followed by North American and Asian countries. Although classical test theory implies excluding participants who appear twice in the dataset to ensure data independence, we decided to analyse the data of those who participated in several workshops, as excluding them would substantially affect the reliability of social network results within workshops. The same participants are included several times in analyses only on a meta-analysis level. There were 19 participants of the Alpine Rendez-Vous 2011 and 14 participants of Alpine Rendez-Vous 2013 who participated in two workshops during the event. 24 researchers participated in workshops at both events in 2011 and 2013, and of those, four persons participated in three workshops altogether and two persons participated in four workshops altogether. This procedure lead to 321 workshop participants in the dataset and a sample of N = 5736 relations between workshop co-attending participants.

Study Design and data collection procedure

Study design and data collection procedure were the same in both Alpine Rendez-Vous meetings. In an experimental design, the factor *group awareness support* (with vs. without) was varied across different workshops in a randomized way. Additional quasi-experimental variables *career level* (PhD student vs. experienced researcher) and *disciplinary background* (Information Technology vs. Social Sciences) varied naturally among participants within the workshops.

Participants in all workshops were informed about the study when they arrived at the event and signed a form of consent during conference registration. All participants who agreed to participate were then equipped with an RFID device, which immediately started tracking their face-to-face contacts with other participants while in pre-defined public areas. Tracking was deactivated when participants checked out of the hotel and returned their RFID device. Four participants opted out of wearing the RFID device. Additionally, a social network questionnaire was handed out to each participant at the end of each workshop. Participants who had to leave earlier were asked to fill in an online version of the questionnaire. Altogether, only 15 questionnaires were not filled in and the missing values appeared to be randomly distributed over all workshops. Personal data about the participants (career level and disciplinary background) were collected with the social network questionnaire and within the registration form for the event. Table 1 gives an overview of all data collected.

Group awareness support, career level, and disciplinary background

Group awareness support was varied between the two conditions, i.e., between a subset of workshops at each Alpine Rendez-Vous. In the experimental condition, workshop participants received a brochure that aimed at supporting cognitive and social group awareness. The brochure contained profiles of all workshop participants. We opted for a paper-based brochure as it is easy to use (e.g., no need to install an app or log in, in contrast to a digital version) and we expected its physical presence to keep participants aware that group awareness information was available, and consequently support the use of the information during the workshop. The profiles were compiled based on information taken from participants' personal websites to spare participants the burden of providing the information to us. Each profile included basic information about each person (name, picture, and contact information) to help the participants to connect names and faces, to look up forgotten names or allow for arranging personal meetings. We expected this information to reduce barriers for one-on-one communication. The main feature of the brochure was information about

Table 1 Overview of collected data

Instrument	Data
RFID devices	■ Duration of face-to-face contact between participants during the event
Personal questions (registration form / with social network questionnaire)	■ Career level (PhD student vs. experienced researcher) ■ Disciplinary background (Information Technology vs. Social Sciences)
Social network questionnaire	■ Collaboration partners within the workshop before the event ■ Potential future collaboration partners within the workshop after the event

personal resources (research interests and exemplary publications) and information pointing to their social resources (affiliations and background). This information was included to help the participants identify who had matching or complimentary interests and expertise (research interests and topics of exemplary publications), shared or desired contacts (co-authors, supervisors), and relations to scientific communities (journals or conferences in which they had published, research group / university of graduation). Based on the still scarce literature on the mechanisms of group awareness tools (Bodemer et al., 2018), we expected this information to help participants to build a more accurate awareness of the group in the workshop that could then be used to compare with personal desires and interests for future research and collaborations, leading eventually to a better identification of relevant partners for one-on-one communication than without the brochure. Figure 1 provides a schema of the profiles. The brochure was given to the participants in the experimental group at the beginning of the workshop without further instructions. The control group did not receive a brochure.

Data on both career level and disciplinary background were extracted from the registration form for Alpine Rendez-Vous 2011 and from the questionnaire at the end of the workshop for Alpine Rendez-Vous 2013. While participants were originally asked to assign themselves to one of three different *career levels* (PhD student, Early/Mid career (postdoc), and full professor), we collapsed the two latter categories for the analyses. This resulted in a variable distinguishing only between PhD students and experienced researchers (Early/Mid career and full professors). This procedure allowed us to include career level as a single dummy variable into the complex statistical models, leading to results that are interpretable in an easier way. Consequently, the dataset of all workshop participants included 77 PhD students and 244 experienced researchers.

Information about the *disciplinary background* of the participants was handled similarly, resulting in a variable that distinguishes only between researchers with a background in Information Technology from those with a background in the Social Sciences (e.g., psychology, education, learning sciences, etc.). Persons who provided background information that could not clearly be assigned to one of the categories, e.g.,

Fig. 1 Schema of a participant profile in the brochure that served as group awareness support

First name, last name (Organizer)

Affiliations: ...
E-Mail: ...
Background: ...
Research Interests: ...

Exemplary Publications:

• ...

• ...

interdisciplinary backgrounds such as "learning technology", were individually checked and then put into the category that seemed to fit best. This classification was chosen because of the specific goal of the Alpine Rendez-Vous to bring researchers from those two areas together and to foster their collaboration. Consequently, the dataset of all workshop participants included 124 persons with a background in Information Technology and 197 workshop participants with a background in the social sciences.

Data sources

The *social network questionnaire* that was handed out at the end of each workshop is the central data source of this study. Each participant received an individualized questionnaire to make sure that the answers were attributed to the right person. Furthermore, the questionnaire was individually adapted to each workshop and contained a list of all workshop participants' names. Participants were asked to indicate with whom they had already collaborated before the event (and with whom they did not) and with whom they had found potential for future collaboration (and with whom they did not). From these data, we extracted a directed network of previous collaborations within each workshop (social network before the workshop) and a directed network of potential collaborations after the workshop (social network after the workshop). It must be noted that the questionnaire used at the Alpine Rendez-Vous 2011 turned out to be quite complicated to fill out for the participants and several of the participants failed to make explicit statements about each of the other workshop participants. For the Alpine Rendez-Vous 2013, the design of the questionnaire was thus simplified, while the content of the questions remained the same. Nevertheless, also with this new design, several participants did not explicitly indicate their relations to all workshop participants. In such cases, we coded missing indications as no relation. In some cases, we also had to make assumptions about the previous relations based on individual patterns in the way participants had filled in their questionnaires.

RFID devices were used for tracking face-to-face contact between participants. RFID devices developed by the SocioPatterns collaboration (http://www.sociopatterns.org) were integrated into participants' name badges. 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 are face-to-face with each other at close range (about 1 to 1.5 m). A base station (a central computer system that uses a local area network) captured the information sent by the RFID devices, pre-processing data in a way that close face-to-face proximity between two individuals with a temporal resolution of 20 seconds is stored and time-stamped (see Cattuto et al., 2010 for a detailed description of the infrastructure). In order to exclude noise and very brief, insignificant contacts, measurement starts only after the first 20 seconds of contact. The RFID devices only tracked face-to-face contacts within the range of the base stations, which were located in public spaces only (dining hall, main hall of the hotel, and workshop rooms). The data collected this way gives access to the amount of time that two participants spent together. For this study, the data collected by the RFID devices was aggregated to a single variable, *duration of interaction,* for each dyad of workshop participants and included all interactions between these two persons during the whole event (including times within and outside of the workshops).

Analyses

Descriptive results of the data were either directly extracted from the database or were computed in R using Python2.7 and the networkx package. Descriptive quantities were computed separately for each workshop and then averaged over workshops, distinguishing between workshops in the control and in the experimental condition. As the network of potential collaborations after the workshop largely includes previous collaboration, we report specifically within these potential future collaborations the *new* potential future collaborations (i.e., the ones that were not declared as previous collaborations).

For the *investigation of selection patterns*, we used the RSiena package version 1.1-232 in R that allowed to do simulation investigations for longitudinal empirical networks based on a stochastic actor-oriented model (Ripley et al., 2014). In general, RSiena analyses how a social network that is measured at different points in time changes between time points and what factors play a role for this change, requiring at least two social network data sets and optionally data sets with attributes of involved actors and / or with attributes of the relations between actors. An RSiena analysis reveals to what extent the factors that are assumed to drive the change from the first to the last network, actually have the expected effect. To do so, the first social network is taken as the starting point and from there, turn-taking between the actors in the network is simulated until the last social network is reached. In the present study, we simulated for each workshop the creation of new potential collaborations (social network after the workshop) based on the network of previous collaborations (social network before the workshop). Figure 2 gives an example of the dataset of workshop I-1. For the turn-taking process between actors, the creation of a new tie was considered, i.e., the actors "decided" whether they wanted to create a new tie and with whom or to do nothing when it was their turn.

The assumed factors driving the turn-taking process are implemented as "effects" in the so called prediction function. A prediction function can consist of different types of effects. For our study, we included three types of creation effects that represented the research questions: 1) duration of face-to-face contact of a dyad during the workshop to help answer RQ 1 (implemented as a main covariate effect in RSiena); 2) network-driven effects to help answer RQ 2, namely reciprocity (recip effect in RSiena) and transitivity (transTrip effect in RSiena, which was chosen among several options to implement transitivity as it is the most commonly used one); 3) effects of career level and discipline split up into three different effects in RSiena to help answer RQ 3: ego effect (egoX effect in RSiena; taking into account the career level / discipline of the person whose turn it is to select a new collaboration partner); alter effect (alterX effect in RSiena; taking into account the career level / discipline of the person who is chosen as a potential new collaboration partner); and homophily effect (sameX effect in RSiena; taking into account whether the two persons who are involved in a new potential collaboration share the same career level / same discipline). Table 2 gives an overview and description of the analysed effects. Variables were not centred to allow for easier interpretation of the results.

Social network before the workshop					Social network after the workshop					Actor attributes (career level and disciplinary background – both binary)				Relation attributes (duration of interaction in seconds)			
	1049	1055	1201	1311		1049	1055	1201	1311	id	career_phd	discipline_tech			1049	1055	1201	1311
1049	NA	1	0	1	1049	NA	1	1	1	1049	0	0	1049	0	360	340	20
1055	1	NA	0	0	1055	1	NA	1	1	1055	1	1	1055	360	0	480	0
1201	0	0	NA	0	1201	0	0	NA	0	1201	0	0	1201	340	480	0	20
1311	1	1	0	NA	1311	1	1	1	NA	1311	1	1	1311	20	0	20	0

Fig. 2 Example of datasets used for the RSiena analyses (snippets of 4 participants in workshop I-1)

Table 2 Overview and description of the effects on the development of intentions to collaborate analysed in RSIENA and in subsequent meta-analyses

Effect	Description
Reciprocity	Propensity to see mutual potential for collaboration
Transitivity	Propensity to see collaboration potential in the collaboration partner of a previous collaboration partner
ego career level (PhD student)	Effect of being a PhD student (in contrast to an experienced researcher) on identifying potential collaboration partners
alter career level (PhD student)	Effect of being a PhD student (in contrast to an experienced researcher) on being identified as a potential collaboration partner
homophily career level	Propensity that researchers at the same career level identify each other as potential collaboration partners
ego discipline (Information Technology)	Effect of having a background in Information Technology (in contrast to a background in the Social Sciences) on identifying collaboration partners
alter discipline (Information Technology)	Effect of having a background in Information Technology (in contrast to a background in the Social Sciences) on being identified as a potential collaboration partner
homophily discipline	Propensity that researchers with the same disciplinary background identify each other as potential collaboration partners
contact duration	Duration of face-to-face contact during the event

The output of the RSiena analyses consists of estimates for each computed effect, given in log odds ratios (Ripley et al., 2014). These estimates can be interpreted as $e^{estimate}$ being the increase in likelihood that a researcher identifies the potential for a new collaboration with another researcher if the effect is one unit higher for this person compared to a person with a unit lower in this effect. For example, $e^{estimate\ for\ ego\ career\ level}$ provides the difference in likelihood that an experienced researcher identifies potential for a new collaboration compared to a PhD student.

After performing RSiena analyses for each workshop separately, we carried out meta-analyses across the workshops using the metafor package in R (Viechtbauer, 2010). First, we performed separate meta-analysis for each effect across all workshops in order to see if it was a significant predictor for the change between the social network before the workshop to the social network after the workshop, i.e., to answer RQs 1, 2, and 3. Together with these meta-analyses, we performed Cochran's Q-tests to test if there was a significant homogeneity between the workshops regarding the effect. A significant Q-test indicates that the overall effect may not be reliable due to heterogeneity between the studies, i.e., there may be subgroups of workshops in the sample that have not yet been considered (see Hoaglin, 2016 for a detailed discussion). In the present study, a significant Q-test may be an indicator that the effect may be influenced by the group awareness support (RQ 4).

In a second step, we attempted to answer RQ 4 by performing meta-analyses on each effect that included the experimental variation of group awareness support as a moderator. In addition, we computed the mean value of each effect for both groups (workshops that received group-awareness support vs. workshops that did not receive group awareness support) and tested for their significance. As recommended by Ripley et al. (2020) based on Viechtenbauer's (2005) extensive study on different approaches for performing a meta-analysis, we applied Restricted Maximum Likelihood estimation (REML) for each meta-analysis.

Results

In this section, we will provide the results for the four research questions about which aspects drive the initiation of potentially new research collaborations. For each research question, we will first provide descriptive social network results, mostly on the workshop level. After providing the descriptive results, we report on the results of the RSiena analysis.

RQ 1: To what extent does face-to-face contact explain the initiation of new research collaborations?

Altogether, the occurrence of 1,436 face-to-face contacts between participants had been captured during the workshops. The *descriptive results* (a detailed table with descriptive values for each workshop can be found in appendix 1) show that participants had on average contact with 9.07 other participants of their workshop. There is no significant correlation between the number of workshop participants and the average degree of face-to-face contacts (r = .267; *p* = .284), indicating that participants in larger workshops did not have contact with more persons than participants in smaller workshops. However, there is noticeable variance in the number of face-to-face contact partners between the workshops as the average degree ranges from 4.18 to 14.35 face-to-face contact partners. Accordingly, the density of the face-to-face contact network varies, i.e., the portion of measured face-to-face contacts between workshop participants in relation to possible contacts. The average density of all workshops is 0.581, indicating that on average about half of all possible face-to-face contacts between participants had actually occurred. In some of the workshops, only about 30% of all possible face-to-face contacts were realized, while in other workshops, more than 80% of possible contacts were realized. The average clustering coefficients tend to be rather on the high end with an average of 0.698, indicating a good amount of connectedness of the participants in the face-to-face contact network. Comparing these values between the control and experimental condition, we see that the descriptive values for workshops in the experimental condition are slightly higher than for workshops in the control condition.

Looking at the cumulated duration of the face-to-face contacts between workshop participants, we find durations between 20 seconds (the minimum necessary for a face-to-face contact to be captured) and 4 hours and 25 minutes (265 minutes in workshop I-2) and a median of 2.98 minutes across all workshops. The medians of face-to-face contacts differ noticeably between the workshops, with a median of 1.00 minute as the lowest to a median of 7.83 minutes as the highest. On average, the median and the maximum duration of face-to-face contacts were higher in the experimental condition compared to the control condition. Considerable differences between the mean and median of face-to-face contact durations within the workshops indicates that contact duration is not normally distributed. For a normal distribution, face-to-face contact durations would have centred around a given mean with lower and higher values being increasingly rarer the more extreme they get. In contrast to this assumption, Fig. 3 shows the typical distribution pattern that has been found for face-to-face contact duration measured with RFID badges in previous studies: The cumulated duration of face-to-face contacts follows a heavy-tailed distribution and is neither normally nor exponentially distributed. Instead, shorter face-to-face contacts are more common with 20 second durations being the most observed face-to-face contact durations, and longer contacts being increasingly rarer.
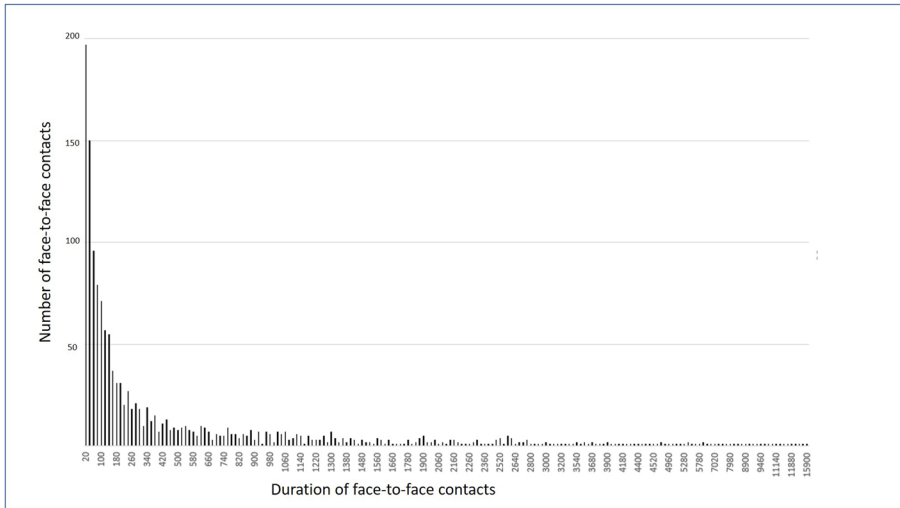
Fig. 3 Duration of face-to-face contacts

The *meta-analysis* reveals that the duration of face-to-face contact during the workshop has a positive significant effect of 0.0002 (SE = 0.0001; p <.001, k = 12) on the initiation of new potential collaborations (see Table 4). In short, the longer two persons had face-to-face contact, the more likely it is that a researcher selects the other as a potential new collaboration partner after the workshop. The likelihood for finding potential collaboration is 1.0012 times higher if there was a face-to-face contact for 60 seconds, 1.127 times for 10 minutes, and 2.054 times for an hour of face-to-face contact. The Q-test is significant, indicating high variability between the workshops that may be to some extent due to the group-awareness intervention that may have helped participants to identify interesting persons to talk with and engage in longer face-to-face contacts with them, instead of using their time for small talk on finding out about potential shared interests.

RQ2: To what extent do reciprocity and transitivity as network-driven selection patterns predict the initiation of new research collaborations?

To understand the role of reciprocity and transitivity for the initiation of new research collaborations, we first look at the descriptive results for the social composition of the workshops regarding previous collaborations and potential collaborations after the workshop, measured by the social network questionnaires (a detailed table with descriptive values for each workshop can be found in appendix 2). The *network statistics for collaborations before the workshop* across all workshops show that each participant had on average 4.74 previous collaboration partners among the workshop participants (SD = 2.71; symmetrized degree, i.e., mutually reported previous collaboration partners). However, the average reciprocity value indicates that participants disagree quite a lot in their perception of whether they had collaborated with each other before (r = 0.544, SD = 0.121); only about half of the previous collaborations that were reported by one participant were confirmed by the indicated previous collaboration partner.

The average clustering coefficient of *previous collaborations* in the workshops shows that there is already a good amount of transitivity (C = 0.501, SD = 0.194). Network statistics for previous collaborations vary between individual workshops but there is relatively little variation between the two experimental conditions.

Network statistics for potential collaborations after the workshops show with an average of 10.57 potential collaboration partners (SD = 3.11; symmetrized degree) a substantial increase compared to the average degree of collaboration partners before the workshop. The degree of reciprocity in indicating potential collaboration partners is on a similar level as for previous collaboration partners (r = 0.483, SD = 0.089). The average clustering coefficient in the workshops (C = 0.764, SD = 0.089), however, increased for potential collaboration partners, meaning that not only are specific participants part of a transitive, well connected network around them, but also that on average most participants have a more transitive, well connected network of potential future collaboration partners than of collaboration partners before the workshop. Again, the descriptive values show no substantial differences between workshops in the control and experimental condition regarding these network statistics, while the individual workshops varied.

The *meta-analyses* of the effects of reciprocity and transitivity show that both are significant predictors for the initiation of new potential collaborations (see Table 5). For reciprocity, the mean effect size is 0.562 (SE = 0.562; p = .023, k = 12). For transitivity, the mean effect size is 0.264 (SE = 0.040; p < .001; k = 12). In short, the likelihood that a researcher selects someone as a potential collaboration partner who was an unreciprocated collaboration partner before the workshop is 1.754 times higher than that the same researcher selects someone with whom no previous collaboration was indicated before the workshop. Additionally, we see that the likelihood that a researcher selects someone who is a collaboration partner of one of the researcher's previous collaboration partners is 1.302 times higher than for selecting someone without a previous relation to their own collaboration partners. The two network-driven selection patterns significantly explain the initiation of new research collaborations. However, the Q-test is significant for both effects (reciprocity: Q (11) = 27.718, p = .004; transitivity: Q (11) = 29.096; p = .002), indicating high variability across the workshops that may to some extent be explained by the group-awareness intervention. Having the brochure at hand may, for example, have helped participants to identify relevant persons right away instead of being dependent on network-driven structures such as being introduced to each other by a common previous collaboration partner.

RQ 3: To what extent do homophily, prestige, and experience regarding career level and disciplinary background as person-driven selection patterns drive the initiation of new research collaborations?

Individual workshops vary visibly in their *composition* regarding participants' career level and disciplinary background (see appendix 2 for details). Looking at the career level, we find workshops with 0 to 8 PhD students while most participants were experienced researchers in a later career phase, with an average of 4.3 PhD students across all workshops. As some of the workshops were more technology-oriented and others focussed more on learning aspects, the number of participants with a background in Information Technology (in contrast to a background in Social Sciences) ranging from 1 to 16, with an average of 6.9 participants with a background in Information Technology per workshop. Comparing the workshop compositions between the control and experimental group, we find that workshops in the experimental condition had a slightly higher number of PhD students (M = 4.6, SD = 1.6) than workshops in the control condition (M = 4.0, SD = 2.2). Workshops in the experimental condition had also a slightly higher number of participants with a background in Information Technology (M = 7.6, SD = 2.8) than the control condition (M = 6.3, SD = 4.4).

Table 3 Percentage of observed relations between different categories of participants compared to all possible relations

	Collaborations before workshop	Contact during workshop	Potential collaborations after workshop
Relations between career levels			
PhD x PhD	17.9%	36.6%	56.9%
Experienced x Experienced	33.9%	36.1%	73.3%
PhD x Experienced	22.3%	34.2%	59.3%
Relations between backgrounds			
Inf. Technology x Inf. Technology	36.3%	44.0%	66.3%
Social Sciences x Social Sciences	25.8%	38.4%	62.8%
Inf. Technology x Social Sciences	23.9%	33.0%	58.6%

Percentages were computed separately for each workshop and then averaged over workshops; only workshops with more than 3 persons of each category are included

Table 3 provides an overview of percentages of the set of relations between different career levels and different disciplinary backgrounds compared to the set of all possible relations, averaged over all workshops. Looking at career level, we find that the fraction of previous collaborations is the lowest in dyads of two PhD students (17.9%) and much higher for dyads of experienced researchers (33.9%). The average number of face-to-face contacts during the workshop, however, does not differ substantially with respect to career levels. For the potential collaborations after the workshop, we see again descriptive differences. Potential collaborations occur more often between two experienced researchers (73.3%) than between two PhD students (56.9 %) or between two researchers of different career levels (59.3%).

Focussing on participants' disciplinary background, we see that dyads of participants with a background in Information Technology show the highest percentage of collaborations before the workshop (36.3%), the highest percentage of face-to-face contacts (44%), and the highest percentage of potential collaborations after the workshop (66.3%) compared to dyads of participants with a background in Social Sciences or mixed dyads.

The *meta-analyses* investigated three effects for career level and for disciplinary background (see Table 4). For career level, we find no significant ego and homophily effects, but a significant negative alter effect of -0.516 (SE = 0.221; p = 0.20, k = 12). This means that there is no tendency that PhD students select their potential collaboration partners differently than experienced researchers. Furthermore, researchers tend not to select their potential collaboration partners based on the aspect that they share the same career level. However, participants tended to select fewer PhD students as potential collaboration partners, indicating their lower prestige compared to more experienced researchers. The likelihood that a PhD student is selected as a potential new collaboration partner is 0.597 times that of an experienced researcher. The Q-test is significant, though, indicating high variability across the workshops that may at least partly be explained by the group awareness intervention (Q(11) = 26.467; p =.006). For example, with group-awareness support, participants may have been able to see relevant research interests of PhD students and get in touch with them instead of relying on stereotypes of a PhD student not being an interesting potential collaboration partner.

Table 4 Results of the meta-analyses of effects on selection of potential new collaboration partners

	Overall model (k = 12)*		Control condition (k = 6)*		Experimental condition (k = 6)*	
	Estimate (sd)	p	Estimate (sd)	P	Estimate (sd)	p
reciprocity	0.561 (0.223)	.012	0.408 (0.303)	.177	0.703 (0.334)	.035
transitivity	0.264 (0.040)	<.001	0.274 (0.064)	<.001	0.236 (0.035)	<.001
ego career level (PhD student)	-0.811 (0.579)	.161	-	-	-	-
alter career level (PhD student)	-0.516 (0.221)	.020	-0.412 (0.176)	.019	-0.668 (0.555)	.229
homophily career level	-0.199 (0.267)	.456	-	-	-	-
ego discipline (Information Technology)	-1.262 (0.548)	.021	-2.266 (0.502)	<.001	-0.414 (0.748)	.580
alter discipline (Information Technology)	-0.057 (0.203)	.779	-	-	-	-
homophily discipline	-0.047 (.0160)	.770	-	-	-	-
interaction duration	0.0002 (0.0001)	<.001	0.0004 (0.0002)	.078	0.0002 (0.0001)	.001

*Based on data of 12 workshops in which the RSiena analyses converged (for details see appendix 1)

For disciplinary background, we find no significant alter and homophily effects but a significant negative ego effect of -1.262 (SE = 0.548; p = .021, k = 12). Consequently, there seems to be no tendency that one of the disciplinary backgrounds is more prestigious and, therefore, more desirable for potential future collaboration than the other. Researchers also tended not to stick to their own discipline. However, participants with a background in Information Technology seem to select significantly fewer other participants for future potential collaboration. A researcher with a background in Information Technology is only 0.283 times as likely to select any new potential collaboration partner than a researcher with a background in the Social Sciences. The Q-test is significant, however, showing high variability across the workshops that may be to some degree due to the group awareness intervention (Q(11) = 23.555; p = .015). The group awareness intervention may have supported participants with different disciplinary backgrounds in distinctive ways, as disciplinary background may be associated with preferences about how to learn about others – either in personal communication or in written form.

RQ 4: How does group awareness support influence selection patterns in the initiation of new research collaborations?

As reported before, the meta-analyses revealed significant effects for reciprocity, transitivity, alter career level, ego disciplinary background, and duration of face-to-face contact (see table 4) and in addition, in all of these cases there was a significant degree of unexplained heterogeneity between the workshops. Consequently, we conducted meta-analyses with group awareness support as a moderator between each of the effects and the selection of new potential collaboration partners to test if the two conditions differed from each other (a forest plot for the effects can be found in appendix 3). In a second step, we looked into the details of the effects within both conditions by conducting separate meta-analyses for workshops in the control condition (without group awareness support) and workshops in the experimental condition (with group awareness support) for each effect (see table 4). This second step allowed us to test whether the effects within each condition were still significant when only studied for this subgroup of workshops.

For *reciprocity*, we find that group awareness support is not a significant moderator of this effect, i.e., control and experimental conditions do not differ (Q(1) = 0.437; p = .508, k = 12). Looking into the details within both conditions based on the post-hoc within-group meta-analyses, the reciprocity effect is positive in both conditions but only significant in the experimental condition. In short, group awareness support has no influence on the effect that researchers preferably select someone as a new potential collaboration partner who was already an unreciprocated collaboration partner before the workshop compared to someone with whom no previous collaboration was indicated before the workshop.

For *transitivity*, group awareness support is also not a significant moderator (Q(1) = 0.0001; p = .994; k = 12). In both, the control and the experimental condition, transitivity has a positive significant effect of a similar magnitude. This means that group awareness support does not influence the general tendency of researchers to select someone as a new potential collaboration partner who was already a collaboration partner of one's own collaboration partner.

Regarding the *alter effect of career level*, there is no significant moderation effect of group awareness support either, i.e., neither condition differs from the other (Q(1) = 0.117, p = .733; k = 12). Looking at the estimates within the conditions, we find the negative alter effect of career level in both, however, it is only significant in the control group and not in the experimental group. This finding can be interpreted in the way that group awareness support has no effect on researchers' preferences to choose experienced researchers over PhD students as new potential collaboration partners.

For the *ego effect of disciplinary background*, there is a significant moderator effect of group awareness support. i.e., control and experimental condition differ significantly from each other (Q(1) = 4.828; p = .028; k = 12). The post-hoc meta-analyses within the conditions reveal that the effect is negative in both conditions; however, it is much higher in the control condition than in the experimental condition. The effect is only significant in the control condition. This means that group awareness support reduces the effect that researchers with a background in Information Technology select fewer new potential collaboration partners than researchers with a background in Social Sciences.

The effect of *face-to-face contact duration* is also not moderated by group awareness support, i.e., no difference between control and experimental condition (Q(1) = 2.398; p = 122; k = 12). Although the effect is positive in both conditions and significant in the experimental condition, it is only marginally significant in the experimental condition. In short, group awareness support does not predict that longer face-to-face interactions between two researchers increase their likelihood for a new potential collaboration after the workshop.

Discussion

The aim of this study was to investigate the onset of scientific collaboration during scientific events. We found that participation in interdisciplinary scientific events was in general associated with the identification of new potential collaboration partners. Network-driven selection patterns were more predictive for the identification of new potential collaboration partners than person-driven selection patterns based on individual characteristics of researchers. That is, whether a researcher was a PhD student or an experienced researcher and researchers' disciplinary backgrounds was less relevant than their previous network of collaborators before the event (although these aspects are likely to be connected to a certain extent). This study found no evidence for homophily effects in this respect. Neither were experienced researchers more likely to select other experienced researchers as collaboration partners rather than PhDs students nor were social scientists more likely to decide in favour of a collaboration with another social scientist instead of an information scientist. The reverse was also true. However, time of interaction matters: The more time researchers spent with each other during the scientific event, the more likely they were to plan a new collaboration with each other. Additional group awareness support during the scientific event did not lead to substantial changes in these selection patterns. One notable exception was that discipline-based differences disappeared with group awareness support.

The results imply that specifically providing opportunities for one-on-one face-to-face contact during scientific events is vital for the onset of scientific collaboration – a challenge for online conferences. In addition, participants coming to a scientific event with no or only a small existing social network and PhD students could benefit from specific support in another form other than group awareness support. Researchers with a background in information technology, in contrast, can benefit from group-awareness support.

The role of face-to-face contact for the initiation of new research collaborations

For face-to-face contacts at the two scientific events, we found a heavy-tailed distribution with many short contacts. There were fewer longer face-to-face interactions between individuals and more shorter face-to-face interactions. This pattern is in line with previous

findings of face-to-face contacts and has been found in face-to-face contacts at conferences (e.g. Barrat et al., 2010) but also in other settings such as among students in schools (e.g. Fournet & Barrat, 2014). Our finding, therefore, supports a stable pattern for face-to-face contacts in communities. Regarding the role of face-to-face contacts for starting scientific collaborations, we found that they are highly relevant. The longer participants had contact with each other at the event, the more likely they were to report an intention to collaborate afterwards. This finding fits well with results of other studies, which found a close connection between face-to-face contact and online activities between researchers before and after scientific events (e.g., Kibanov et al., 2015). Our finding also confirms that activities that conference organizers have pursued for a long time, so far only based on assumed relevance, are indeed important to foster scientific collaboration and the integration of newcomers. Such activities include poster presentations that allow for direct face-to-face interaction in contrast to formal talks, but also to explicitly make room for informal conversations in the program during lunch, dinner, or receptions. Given the challenge for the upcoming years when researchers will be less able to travel and meet face-to-face in order to prevent the spread of diseases and due to sustainability considerations, the question arises to what extent we can transfer this study's results to pure on-line and hybrid interactions. We will need to find ways to translate opportunities for informal one-on-one or small group contact into online formats when face-to-face scientific events cannot be the standard anymore. For example, the Science of Team Science conference was held on-line in 2020 and used a slack channel as well as a zoom sign-up sheet to create informal and spontaneous spaces for "coffee-chat" discussion. How does the influence of such innovations compare with direct face-to-face interaction?

All this being said, the potential of our face-to-face contact data may not have been fully exploited within this paper by reducing it to duration as a single, aggregated variable. As Malik (2018) pointed out, there may be other relevant aspects such as repetition of contacts during the event or the specific time point on which face-to-face contacts occurred. In addition, the timing of contacts (e.g., during workshop activities vs. during social events such as lunch or coffee breaks) could give interesting insights. These aspects need to be explored in the future as including them was out of the scope of this paper.

On the methodological level, the measurement of face-to-face contacts using sociometric badges is a quite innovative technique for the learning sciences, although it has been used in other research areas for some time. Sociometric badges come with certain benefits but also a few disadvantages that need to be considered for the interpretation of study results. There are only a few methods that capture face-to-face interactions besides RFID badges, namely self-reports, observations, and video-recordings of social situations. Regarding objectivity, RFID badges were found to be superior to methods that rely on participants' memory, such as diaries and surveys (Mastrandea et al., 2015), and this is very likely true also for external ratings of observed or video-recorded social interactions, such as RFID badges, which eliminate human errors and biases during data collection and analysis. In addition, RFID badges allow the investigation of new research questions that have previously not been easy to study in a quantitative manner, specifically in field settings and when the object of study is face-to-face interactions between large numbers of people that roam between different places. However, measurement with RFID badges is not free from errors: There are still human-related issues that harmed accurate data collection, specifically that study participants forgot to wear their badges at all times, they placed the badges at unfavourable spots on their bodies (e.g., on trouser pockets), or badges simply flipped

when people moved and therefore did not capture every face-to-face interaction. There are also technical difficulties related to the badges as Elmer et al. (2019) report, namely that a significant number of face-to-face contacts are not captured by the sociometric badges. False positives (i.e., RFID measurement of contacts that did not actually take place) are not a problem with RFID badges. To make up for "lost" contacts and to increase measurement validity, Elmer and colleagues recommend the use of aggregated face-to-face contact duration, as we did in this study. As the badges fail to capture face-to-face contact data perfectly and we probably missed a significant amount of time researchers spent in face-to-face contacts, we conclude that it is more likely that our analysis results underestimate the importance of face-to-face contact for the initiation of collaboration intentions and the integration of newcomers than overestimating them. Nevertheless, in future studies, a combination of different forms of measurement of face-to-face interaction could be a helpful improvement to increase validity.

The role of reciprocity and transitivity as network-driven selection patterns for the initiation of new research collaborations

Reciprocity and transitivity are essential network-driven selection patterns that were previously found to explain the building of relationships in many other contexts. Our results show that this is also the case for building new collaboration intentions among researchers. The reciprocity effect in this study can be understood as a convergence in dyads' shared understanding that a collaboration would be worthwhile to establish between a researcher who had not been aware that there was a shared research interest and a researcher who was already aware of the shared research interest because they remembered some kind of previous collaboration. An interesting finding is that the reciprocity effect is not positive in all workshops (see Appendix 3), meaning that in some workshops there is an opposite effect, and researchers tended to specifically not reciprocate one-sided previous collaborations. It is unclear why this is the case as these workshops seem not to differ from other workshops. However, a similar question has been discussed regarding unreciprocated friendships in adolescents. This discussion, unfortunately, has led to inconclusive results and no satisfying answer as well: Although an obvious explanation is that the not reciprocating person has a higher status in a social network and more opportunities for friendships – or in our case collaborations – outside of the investigated social network that would allow this person to be more selective, empirical evidence suggests that this is not the case (cf. Block, 2015). Consequently, we assume that we have found a similar phenomenon in our data that requires more research on the nature of reciprocity in social networks in general.

In addition, it seems noteworthy that differences in where researchers set their threshold of what counts as a collaboration may have added some noise to the data as they quite often disagreed whether there had been a previous collaboration between a given dyad before the workshop. We tried to counteract this problem by defining "collaboration" on the questionnaires but still some participants may have thought of formal collaborations in a joint project or intensive work on shared publication only, while others may also have included informal collaboration such as previous discussions about a topic of interest or shared co-authorships on papers with a large number of co-authors that did not work closely together. In social network terms, some researchers may have interpreted the question as asking for weak ties and others as asking for strong ties (Granovetter, 1973). We, therefore, assume

that the reciprocity effect in forming new collaboration intentions, which we found in the meta-analysis, may to some extent be moderated by researchers' understanding of the term "collaboration". Researchers with a more formalized understanding of collaboration, came to the conclusion during the workshop that a formalized collaboration (a strong tie) would be worthwhile to establish with a person with whom they only had informally collaborated before (i.e. recognized and discussed shared research interests).

The transitivity effect appears in all workshops as well as in the overall meta-analysis as a positive effect, pointing out the importance of initial contacts for the establishment of collaboration intentions with other researchers. This is in line with previous research (e.g., Baerveldt et al., 2010) and with social capital theory (Bourdieu, 1986) as we can conclude that researchers, who attend a scientific event where many of their previous collaboration partners are, have access to more opportunities for building new collaboration intentions than researchers who are not well connected when attending a scientific event.

The role of homophily, prestige, and experience regarding career level and disciplinary background as person-driven selection patterns for the initiation of new research collaborations

Previous literature had found that homophily is a widely common person-driven selection pattern, but heterophily would be desirable for the selection of new potential collaboration partners in an interdisciplinary research field. We found no evidence for homophily based on career level and disciplinary background in this study, indicating that the onset of inter-disciplinary and career-level spanning collaboration in the field of technology-enhanced learning works quite well from this perspective.

When looking at differences between PhD students and experienced researchers, however, our findings complement previous findings on researchers' face-to-face contacts at conferences that showed different contact patterns for these groups (Atzmueller et al., 2012). Our assumption that PhD students might lack experience in how to select potential new collaboration partners and would behave differently in their selection behaviour, in contrast, was not confirmed. PhD students, consequently, seem not to need any specific support on how to identify potential new collaboration partners. What we found, however, was a prestige effect, favouring experienced researchers as potential collaboration partners. This finding may to some extent be explained by other researchers' assumptions about PhD students' limited availability for collaborations in the near future, e.g., because they and their resources are consumed by their PhD project and they cannot yet serve as Co-PIs on larger research projects. Increasing the visibility of their research at a scientific event, however, may be helpful, enabling others to more easily identify their potential for future collaboration based on their research, either with a more long-term perspective or for smaller collaborative projects such as joint special issues. In particular, PhD students who cannot benefit from transitivity effects of a mentor or other gatekeepers to the community may benefit from support measures that make other participants aware of them.

For disciplinary backgrounds, we found the opposite effect. No discipline was more prestigious than the other, but researchers seem to show different selection patterns for their future potential collaboration partners. In line with previous findings that showed that researchers with a background in Information Technology tend to be engaged in fewer face-to-face interactions during scientific events (Kibanov et al., 2015), they also select fewer potential collaboration partners. In comparison to the career level prestige effect, this effect of disciplinary background is much higher. Exploring further disciplinary differences

and reasons for such differences seems worthwhile, especially as researchers with a background in Information Technology have been the main sample of most studies that looked at contact patterns during conferences and other disciplines have received less attention.

The effect of group awareness support on selection patterns in the initiation of new research collaborations

Based on previous findings that information about group members leads to more active forms of selection (Baerveldt et al., 2010), we had assumed that group awareness support influences how researchers select potential future collaboration partners. Our findings show no such effect for most of the previously described selection patterns except for one: The ego effect of disciplinary background, for which we found that researchers with a background in Information Technology select fewer potential future collaboration partners than social scientists. For the workshops that received group awareness support, the difference in the selection of new potential collaboration partners between participants with different disciplinary backgrounds disappeared. Looking at the intended mechanism of providing support for group awareness, namely, to help in perceiving and understanding information about others that can then tacitly guide collaborative processes (Bodemer et al., 2018), this makes sense. By increasing researchers' awareness of the other participants' invisible attributes, researchers can adapt their behaviour in accordance with the new information and identify more persons' potential for future collaboration. The other effects we found to drive researchers' selection behaviour, in contrast, are less accessible on the cognitive level of an individual person. To change these patterns, other forms of support may be needed. For example, mentoring or tandem programs may be helpful to reduce disadvantages of participants with few prior contacts in the scientific community.

All this being said, our findings on the effect of group awareness support rely on a quite limited sample, contrasting only six groups per condition. Further replicative studies are necessary to validate our findings of group awareness support on the community level. The findings, though, are in line with subjective positive assessment of a group awareness tool for conference participants (Windhager et al., 2014; Zenk et al., 2014). It seems, therefore, worthwhile to further develop and investigate the mechanisms of group awareness tools to support not only small group collaboration but also to foster the identification for collaboration potential within scientific communities. It seems especially relevant to continue to consider the question regarding what information should be shown to participants as we can assume that the available information will have an impact on what selection patterns are emphasized. Specific recommendations of relevant researchers based on previous publications as suggested by Alshareef et al., (2018) may support homophilous selection. Specific recommendations based on current research interests as implemented by Windhager et al., (2014) may, in contrast, allow for more diverse selection and be especially helpful to participants with a very clear focus on what future collaboration they want to pursue. Very broad information, such as used in the present study, may also allow for diverse and focused selection and also help spur situational interest in specific persons while browsing their profiles, but be less helpful in settings with a very large number of participants as browsing so many profiles is time consuming. So far, this can only be speculated upon and needs further research.

Limitations of the study

In addition to previously discussed limitations, it is likely that this study suffers from the common problem that some relevant variables have not been taken in to account. Especially the amount of unexplained variance between the workshops indicates that, indeed, some relevant aspects have not been included. This may include variables related to personality, e.g. extraversion of researchers may be relevant, prestige of certain workgroups, or workshop design, which may allow for more or less interaction and insights about the other participants. To understand the role of these variables demands further investigation that was beyond the scope of the current study.

Due to the study's design, we also need to be careful with causal interpretations. Most of the investigated effects and variables were not experimentally induced but either resulted from the measured social networks (reciprocity and transitivity) or were of quasi-experimental nature and given by the nature of the participants or their behaviour during the workshops (career level, disciplinary background, and face-to-face contact duration). Consequently, we cannot be sure whether these variables actually caused the initiation of future collaboration intentions. Group awareness support was the only experimental varied variable, being qualified to be causally interpreted with the given limitations discussed previously.

Finally, as we aimed at understanding the onset of scientific collaborations, looking only at the identified potential future collaborations is a big limitation. In a next step, data that indicates if the identified potential was actually put into practice needs to be looked at. In addition, we did not take into account what role social media may have played for the initiation of new collaborations among researchers in this study. As the use of social media has increased in the last couple of years, it may not only be relevant for the initiation of distant collaborations but also affect the initiation of research collaborations at face-to-face events, pointing at a need for replication of our findings with specific focus on the role of social media.

Conclusion and future research

In this paper, we explored factors that influence the onset of scientific collaborations at two scientific events for research on technology-enhanced learning. We aimed at understanding how scientific communities can support their members in finding new collaboration partners and in that way including newcomers to the scientific community. In addition, we found that the overall duration of face-to-face contact during the scientific event is highly predictive for establishing potential for future collaborations. Important mechanisms are transitivity and reciprocity as they shape researchers' selection patterns. Both basically rely on the benefits of a previously existing network before attending a scientific event. We found differences between subgroups of participants, specifically that PhD students are less prestigious future collaboration partners than experienced researchers and that researchers with a background in Information Technology select fewer future collaboration partners. However, we found no homophily effects on career level or disciplinary background. This finding indicates that subgroups of scientific communities may benefit from different types

of support during scientific events. Group awareness support was experimentally varied in this study and we found that researchers with a background in Information Technology changed their selection patterns and identified more potential future collaboration partners with the support than without. In addition, this study gives a first example of how using RFID badges to measure face-to-face interactions can be used to investigate "collaboration in the wild" and provide new types of data for research on collaborative learning.

These findings show a need for further research in several directions. On the one hand, a more detailed investigation of face-to-face interaction time during scientific events may be fruitful as face-to-face interaction in formal or informal settings of a scientific event may have different effects. On the other hand, it seems worthwhile to further explore how and to what extent researchers realize their collaboration intentions, e.g., by looking at joint publications and other scientific artefacts co-created after the event. Of specific interest for understanding the realization of collaboration intentions may be the role of researchers' early face-to-face interaction patterns during the scientific event, the role of technologies for maintaining longstanding collaborations after the event (e.g. social networking tools), and the identification of potentially different types of research collaborations that could be followed up in more longitudinal studies. Also, further basic research on additional factors influencing the onset of collaboration is needed, such as factors that explain when people will not be successful in initiating new collaborations and what role researchers' personality plays in this process. Another rather practical research direction within and beyond the CSCL community is the question how digital and hybrid environments for scientific events can be designed to allow for "face-to-face-like" interaction between participants as well as for interactions that allow the powerful transitivity effects to happen. Also unique benefits of digital and hybrid environments need to be explored that can even improve interaction during (partially) technology mediated scientific events, such as sophisticated group awareness tools.

Notes

1. For more information on KALEIDOSCOPE see here: https://cordis.europa.eu/project/rcn/71192_en.html
2. For more information on STELLAR see here: https://cordis.europa.eu/project/rcn/89455_en.html
3. Alpine Rendez-Vous 2011's outcome is the book Grand Challenges in Technology Enhanced Learning (Fischer et al., 2014); The outcome of the Alpine Rendez-Vous 2013 is Grand Challenge Problems in Technology-Enhanced Learning II: MOOCs and Beyond (Eberle et al., 2016).
4. For more information about the Alpine Rendez-Vous 2011, please see here: http://web.archive.org/web/20110513105838/http://www.stellarnet.eu:80/programme/wp3/rendez-vous
5. For more information about the Alpine Rendez-Vous 2013, please see here: http://web.archive.org/web/20141022002451/https://metah.imag.fr/alpine-rendez-vous/home/?lang=fr

Appendix 1

Table 5 Descriptive statistics for face-to-face contacts

workshop	N	average degree*	density*	average clustering*	mean duration (in min)	median duration (in min)	maximum duration (in min)
Control Condition							
I-1**	15	8.93	0.638	0.762	6.92	2.33	76.33
I-3	17	5.18	0.324	0.439	6.22	1.67	80.00
I-4	24	6.48	0.324	0.529	5.65	1.67	41.00
I-6**	17	6.25	0.417	0.513	2.70	1.00	21.00
I-7	21	8.10	0.426	0.482	6.63	2.33	86.00
II-2	21	10.19	0.510	0.695	17.65	3.33	198.00
II-4**	10	5.20	0.578	0.715	14.50	2.67	73.00
II-6	22	14.00	0.667	0.753	9.35	2.67	193.00
II-8	17	12.35	0.772	0.880	13.37	5.67	167.67
II-10**	9	6.00	0.750	0.852	13.75	4.00	118.67
Overall Mean	17.3	8.27	0.541	0.662	9.87	2.73	106.47
Experimental Condition							
I-2	23	7.09	0.338	0.553	28.37	2.33	265.00
I-5**	13	4.18	0.418	0.717	4.80	2.00	21.67
I-8	25	7.92	0.344	0.511	3.80	1.33	25.67
II-1**	17	12.71	0.794	0.828	15.73	5.00	113.33
II-3	18	14.35	0.897	0.911	21.07	7.83	149.00
II-5	16	10.67	0.762	0.793	13.53	4.00	171.33
II-7	19	11.22	0.660	0.781	10.18	2.00	198.67
II-9	17	12.50	0.833	0.847	6.82	2.00	74.67
Overall Mean	18.5	10.08	0.631	0.743	13.03	3.32	127.42
All Workshops							
Mean	17.8	9.07	0.581	0.698	11.28	2.98	115.78

Workshops are labelled from "I-1" to "II-10" with "I" referring to workshops at the Alpine Rendez-Vous 2011 and "II" referring to workshops at the Alpine Rendez-Vous 2013. Within the events, workshops were randomly numbered from 1 to 8 and 1 to 10, respectively

Average degree within a workshop is computed as number of observed face-to-face contacts between individual participants within a workshop. *Density* of the workshops' face-to-face contact network is computed as the ratio of observed relations to possible relations in the workshop. *Average clustering* of each workshop indicates whether the participants of a workshop were all evenly connected or rather split in sub-groups; **For these workshops, RSiena analyses did not converge. Consequently, these workshops were not included in the meta-analyses
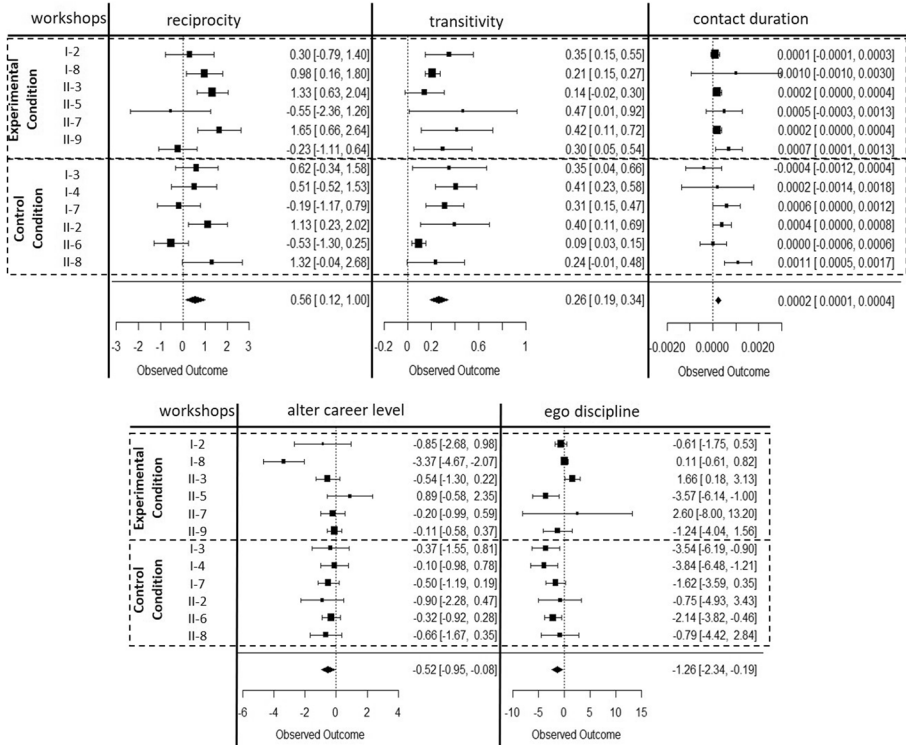
Appendix 2

See Table 6

Table 6 Workshop compositions and network statistics

Workshop	Network composition				Collaborations before workshop				Potential Collaborations after workshop			
	N participants	N relations	N PhD students	N technology background	average degree*	density*	average clustering*	reciprocity	average degree*	Density*	average clustering*	reciprocity
Control Condition												
I-1**	15	210	3	5	2.53	0.181	0.689	0.684	12.40	0.886	0.914	0.581
I-3	17	272	2	4	2.24	0.140	0.329	0.368	7.88	0.493	0.660	0.388
I-4	24	552	6	6	6.50	0.283	0.512	0.538	13.92	0.605	0.730	0.557
I-6**	17	272	8	16	6.24	0.390	0.795	0.509	11.29	0.706	0.770	0.479
I-7	21	420	5	3	6.76	0.338	0.511	0.507	14.95	0.748	0.833	0.490
II-2	21	420	3	13	2.48	0.124	0.370	0.731	7.05	0.352	0.622	0.392
II-4**	10	90	3	6	0.60	0.067	0.000	0.333	5.40	0.600	0.807	0.519
II-6	22	462	6	4	11.73	0.558	0.793	0.597	16.91	0.805	0.878	0.457
II-8	17	272	4	1	3.41	0.213	0.479	0.655	8.12	0.507	0.638	0.521
II-10**	9	72	0	5	4.00	0.500	0.538	0.500	5.11	0.639	0.774	0.478
Overall Mean	17.3	304.2	4.0	6.3	4.65	0.279	0.502	0.542	10.30	0.634	0.762	0.486
Experimental Condition												
I-2	23	506	4	8	4.61	0.209	0.455	0.358	11.91	0.542	0.689	0.277
I-5**	13	156	3	11	5.23	0.436	0.589	0.647	10.46	0.872	0.916	0.647
I-8	25	600	4	8	8.32	0.347	0.640	0.567	13.28	0.553	0.748	0.470
II-1**	17	272	8	8	7.29	0.456	0.726	0.484	10.47	0.654	0.780	0.416
II-3	18	306	5	9	3.22	0.190	0.388	0.483	8.44	0.497	0.630	0.500
II-5	16	240	2	2	4.00	0.267	0.579	0.781	10.75	0.717	0.798	0.628
II-7	19	342	5	4	3.16	0.175	0.349	0.567	9.68	0.538	0.731	0.511
II-9	17	272	6	11	2.24	0.140	0.280	0.474	12.12	0.757	0.821	0.379
Overall Mean	18.5	336.8	4.6	7.6	4.87	0.277	0.501	0.545	10.89	0.641	0.764	0.479

Table 6 (continued)

Workshop	Network composition				Collaborations before workshop				Potential Collaborations after workshop			
	N participants	N relations	N PhD students	N technology background	average degree*	density*	average clustering*	reciprocity	average degree*	Density*	average clustering*	reciprocity
All Workshops												
Mean	17.8	318.7	4.3	6.9	4.74	0.279	0.501	0.544	10,56	0,637	0,763	0,483

Workshops are labelled from "I-1" to "II-10" with "I" referring to workshops at the Alpine Rendez-Vous 2011 and "II" referring to workshops at the Alpine Rendez-Vous 2013. Within the events, workshops were randomly numbered from 1 to 8 and 1 to 10, respectively

* Networks were symmetrized for computing these parameters; *average degree* within a workshop is computed as number of observed relations between individual participants within a workshop. *density* of the workshops' social networks is computed as the ratio of observed relations to possible relations in the workshop. *average clustering* of each workshop indicates if the participants of a workshop were all evenly connected or rather split in sub-groups. *reciprocity* in each workshop is computed as the extent to which participants reciprocally nominated each other; **For these workshops, RSiena analyses did not converge. Consequently, these workshops were not included in the meta-analyses

Appendix 3

See Fig. 4.

Fig. 4 Forest plots for all significant effects

References

Aiello, L. M., Barrat, A., Cattuto, C., Ruffo, G., & Schifanella, R. (2010). Link Creation and Profile Alignment in the aNobii Social Network. In *2010 IEEE Second International Conference on Social Computing (SocialCom)* (pp. 249–256). https://doi.org/10.1109/SocialCom.2010.42

Alshareef, A. M., Alhamid, M. F., & El Saddik, A. (2018). Recommending Scientific Collaboration Based on Topical, Authors and Venues Similarities. 2018 IEEE International Conference on Information Reuse and Integration (IRI). https://doi.org/10.1109/iri.2018.00016

Atzmueller, M., Doerfel, S., Hotho, A., Mitzlaff, F., & Stumme, G. (2012). Face-to-Face Contacts at a Conference: Dynamics of Communities and Roles. In M. Atzmueller, A. Chin, D. Helic, & A. Hotho (Eds.), *Modeling and Mining Ubiquitous Social Media, 7472*, 21–39. Springer.

Baerveldt, C., van de Bunt, G. G., & Federico de la Rua, AD. (2010). Why and how selection patterns in classroom networks differen between students: The potential influence of network size preferences, level of information, and group membership. *REDES - Revista hispana para el análisis de redes sociales, 19*(12), 273–298.

Bammer, G. (2013). *Disciplining Interdisciplinarity: Integration and implementation sciences for researching complex real-world problems*. ANU E Press.

Barrat, A., Cattuto, C., Szomszor, M., van den Broeck, W., & Alani, H. (2010). Social dynamics in conferences: Analyses of data from the live social semantics application. In P. F. Patel-Schneider, Y. Pan, P. Hitzler, P. Mika, & L. Zhang (Eds.), *ISWC`10: Proceedings of the 9th International Semantic Web Conference on the semantic web, Volume II* (17-33), Berlin: Springer.

Birnholtz, J. P. (2007). When Do Researchers Collaborate? *Toward a Model of Collaboration Propensity Journal of the American Society for Information Science and Technology, 58*(14), 2226–2239.

Block, P. (2015). Reciprocity, transitivity, and the mysterious three-cycle. *Social Networks, 40,* 163–173.

Bodemer, D., Janssen, J., & Schnaubert, L. (2018). Group awareness tools for computer-supported collaborative learning. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (pp. 351–358). Routledge.

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). Greenwood.

Bozeman, B., & Boardman, C. (2014). Research collaboration and team science: A state-of-the-art review and agenda. Springer Briefs in Entrepreneurship and Innovation. Springer International Publishing.

Byrne, D., & Clore, G. L. (1970). A reinforcement model of evaluative responses. *Personality, 1*, 103–128.

Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior, 22*, 195–202.

Cattuto, C., van den Broeck, W., Barrat, A., Colizza, V., Pinton, J.-F., & Vespignani, A. (2010). Dynamics of person-to-person interactions from distributed RFID sensor networks. *PloS ONE, 5*(7), e11596.

Clemente-Gallardo, J., Ferrer, A., Iniguez, D., Rivero, A., Ruiz, G., & Tarancon, A. (2019). Do researchers collaborate in a similar way to publish and to develop projects? *Journal of Informetrics, 13*, 64–77.

Coleman, J. S. (1988). Social capital in the creation of human capital. *The American Journal of Sociology, 94*.

Eberle, J., Lund, K., Tchounikine, P., & Fischer, F. (Eds.). (2016). Grand Challenge Problems in Technology-Enhanced Learning II: MOOCs and Beyond: Perspectives for Research, Practice, and Policy Making Developed at the Alpine Rendez-Vous in Villard-de-Lans (1st ed. 2016). SpringerBriefs in Education. Cham: Springer. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1079147

Eberle, J., Stegmann, K., & Fischer, F. (2014). Legitimate Peripheral Participation in Communities of Practice: Participation Support Structures for Newcomers in Faculty Student Councils. *Journal of the Learning Sciences, 23*(2), 216–244. https://doi.org/10.1080/10508406.2014.883978

Elmer, T., Chaitanya, K., Purwar, P., & Stadtfeld, C. (2019). The validity of RFID badges measuring face-to-face interactions. *Behavioral Research Methods, 51*, 2120–2138. https://doi.org/10.3758/s13428-018-1180-y

Fischer, F., Wild, F., Sutherland, R., & Zirn, L. (2014). *Grand Challenges in Technology Enhanced Learning: Outcomes of the 3rd Alpine Rendez-Vous. SpringerBriefs in Education*. Cham, s.l.: Springer International Publishing. Retrieved from https://doi.org/10.1007/978-3-319-01667-2

Fournet, J. & Barrat, A. (2014). Contact Patterns among High School Students. *PLoS One, 9*(9), e107878. https://doi.org/10.1371/journal.pone.0107878

Goodwin, C. (2000). Action and Embodiment Within Situated Human Interaction. *Journal of Pragmatics, 32*, 1489–1522.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology, 78*(6), 1360–1380.

Hoaglin, D. C. (2016). Misunderstandings about Q and 'Cochran's Q test'in meta-analysis. *Statistics in Medicine, 35*(4), 485–495.

Isella, L., Stehlé, J., Barrat, A., Cattuto, C., Pinton, J. F., & van den Broeck, W. (2011). What's in a crowd? Analysis of face-to-face behavioral networks. *Journal of Theoretical Biology, 271*(1), 166–180.

Janssen, J., & Bodemer, D. (2013). Coordinated Computer-Supported Collaborative Learning: Awareness and Awareness Tools. *Educational Psychologist, 48*(1), 40–55. https://doi.org/10.1080/00461520.2012.749153

Jeong, H., Hmelo-Silver, C. E., & Yu, Y. (2014). An examination of CSCL methodological practices and the influence of theoretical frameworks 2005–2009. *International Journal of Computer-Supported Collaborative Learning, 9*(3), 305–334.

Kienle, A., & Wessner, M. (2005). Principles for Cultivating Scientific Communities of Practice. In P. Besselaar, G. Michelis, J. Preece, & C. Simone (Eds.), *Communities and Technologies 2005: Proceedings of the Second Communities and Technologies Conference, Milano 2005* (pp. 283–299). Dordrecht: Springer.

Kienle, A., & Wessner, M. (2006). The CSCL community in its first decade: Development, continuity, connectivity. *International Journal of Computer-Supported Collaborative Learning, 1*(1), 9–33. https://doi.org/10.1007/s11412-006-6843-5

Kibanov, M., Atzmueller, M., Illig, J., Scholz, C., Barrat, A., Cattuto, C., & Stumme, G. (2015). Is Web Content a Good Proxy for Real-Life Interaction? A Case Study Considering Online and Offline Interactions of Computer Scientists. In *Proceedings of the 2015 ACM/IEEE International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2015*. Paris, France.

Lave, J. (1991). Situating learning in communities of practice. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 63–82). Washington, DC, US: American Psychological Association. https://doi.org/10.1037/10096-003

Lee, V. R., Ye, L., & Recker, M. (2012). What a long strange trip it's been: A comparison of authors, abstracts, and references in the 1991 and 2010 ICLS Proceedings. In J. van Aalst, K. Thompson, M. J. Jacobson, & P. Reimann (Eds.), *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012)* (pp. 172–176). Sydney, NSW, Australia: ISLS.

Levine, J. M., & Moreland, R. L. (2013). Knowledge Transmission in Work Groups: Helping Newcomers to Succeed. In J. M. Levine, L. L. Thompson, & D. M. Messick (Eds.), *Series in Organization and Management. Shared Cognition in Organizations: The Management of Knowledge* (pp. 267–296). Hoboken: Taylor and Francis.

Lin, N. (2001). *Social capital: A theory of social structure and action. Structural analysis in the social sciences: Vol. 19*. Cambridge, New York: Cambridge University Press. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=74320

Lott, A. J., & Lott, B. E. (1974). The role of reward in the formation of positive interpersonal attitudes. In T. L. Huston (Ed.), *Foundations of interpersonal* (pp. 171–192). Academic Press.

Malik, M. M. (2018). *Bias and beyond in digital trace data* (Doctoral dissertation, Carnegie Mellon University).

Mastrandea, R., Fournet, J., & Barrat, A. (2015). Contact patterns in a high school: A comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PLoS ONE, 10*(9), e0136497.

Maienschein, J. (1993).Why Collaborate? *Journal of the History of Biology, 26*(2), (Summer 1993), 167-183.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology, 27*, 415–444.

Ripley, R. M., Snijders, T. A. B., Boda, Z., Vörös, A., & Preciado, P. (2014). *Manual for RSiena version 4.0 (version May 10, 2014)*. Oxford. Retrieved from University of Oxford, Department of Statistics; Nuffield College website: http://www.stats.ox.ac.uk/siena/

Ripley, R. M., Snijders, T. A. B., Boda, Z., Vörös, A., & Preciado, P. (2020). *Manual for RSiena version 4.0 (version January 7, 2020)*. Oxford. Retrieved from University of Oxford, Department of Statistics; Nuffield College website: http://www.stats.ox.ac.uk/siena/

Rummel, N. (2018). One framework to rule them all? Carrying forward the conversation started by Wise and Schwarz. *International Journal of Computer-supported Collaboration, 13*(1), 123–129.

Rylance, R. (2015). Grant giving: Global funders to focus on interdisciplinarity. *Nature, 525*(7569), 313–315.

Schmidt, K. (2002). The problem with "awareness": Introductory remarks on "awareness in CSCW." *Computer Supported Cooperative Work (CSCW), 11*(3–4), 285–298.

Slingerland, E., & Collard, M. (Eds.). (2012). *Creating Consilience: Integrating the Sciences and the Humanities*. Oxford University Press.

Tchounikine, P. (2019). Learners' agency and CSCL technologies: Towards an emancipatory perspective. *International Journal of Computer-supported Collaboration, 14*(2), 237–250.

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics, 30*, 261–293.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). Springer.

Windhager, F., Smuc, M., & Zenk, L. (2014). Conferences, Coffee and Complexity. Supporting Cognition and Communication at Conventions. *Proceedings of the 18th International Conference on Information Visualisation (iv2014).*

Wise, A., & Schwarz, B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-supported Collaboration, 12*(4), 423–467.

Xia, Y. & Borge, M. (2019). A Systematic Review of the Quantification of Qualitative Data in the Proceedings of International Conferences of CSCL from 2005 to 2017. In K. Lund, G.P., Niccolai, E., Lavoué, C., Hmelo-Silver, G., Gweon, & M., Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, Volume 2 (pp. 620-623).* Lyon, France: International Society of the Learning Sciences.

Zenk, L., Smuc, M., & Windhager, F. (2014). Beyond the name tag. Connecting people and knowledge at conferences. In B. Lutz (Ed.), *Wissen nimmt Gestalt an: Beiträge zu den Kremser Wissensmanagement-Tagen 2013* (pp. 215-225). Krems: Edition Donau-Universität Krems.