Computer-Supported Collaborative Learning
DOI 10.1007/s11412-013-9182-3

Explaining authors' contribution to pivotal artifacts during mass collaboration in the Wikipedia's knowledge base

Q9 **Iassen Halatchliyski • Johannes Moskaliuk • Joachim Kimmerle • Ulrike Cress**

Received: 29 October 2012 / Accepted: 18 September 2013

Abstract This article discusses the relevance of large-scale mass collaboration for computer-supported collaborative learning (CSCL) research, adhering to a theoretical perspective that views collective knowledge both as substance and as participatory activity. In an empirical study using the German Wikipedia as a data source, we explored collective knowledge as manifested in the structure of artifacts that were created through the collaborative activity of authors with different levels of contribution experience. Wikipedia's interconnected articles were considered at the macro level as a network and analyzed using a network analysis approach. The focus of this investigation was the relation between the authors' experience and their contribution to two types of articles: *central pivotal articles* within the artifact network of a single knowledge domain and *boundary-crossing pivotal articles* within the artifact network of two adjacent knowledge domains. Both types of pivotal articles were identified by measuring the network position of artifacts based on network analysis indices of topological centrality. The results showed that authors with specialized contribution experience in one domain predominantly contributed to central pivotal articles within that domain. Authors with generalized contribution experience in two domains predominantly contributed to boundary-crossing pivotal articles between the knowledge domains. Moreover, article experience (i.e., the number of articles in both domains an author had contributed to) was positively related to the contribution to both types of pivotal articles, regardless of whether an author had specialized or generalized domain experience. We discuss the implications of our findings for future studies in the field of CSCL.

Keywords Artifact · Mass collaboration · Network analysis · Wikipedia

I. Halatchliyski (✉) · U. Cress
Knowledge Construction Lab, KMRC—Knowledge Media Research Center, Schleichstr. 6, 72076 Tuebingen, Germany
e-mail: i.halatchliyski@iwm-kmrc.de

J. Moskaliuk · J. Kimmerle
Department of Applied Cognitive Psychology and Media Psychology, University of Tuebingen, Tuebingen, Germany

Introduction

Computers facilitate connectivity and coordination among large networks of people (Lipponen 2002; Ryberg and Larsen 2008) and enable them to form communities and build digital knowledge bases. Recently, Web 2.0 environments have greatly lowered the barriers to participative activities for all Internet users (Kapur et al. 2007). As a result, so-called *mass collaboration* has become a common phenomenon (Cress 2013; Cress et al. 2013; Tapscott and Williams 2006). With its specific affordances for knowledge-related activities (Lipponen 2002; Pifarré and Kleine Staarman 2011), mass collaboration presents a whole new field of study in computer-supported collaborative learning (CSCL; Scheuer et al. 2010). Its essence resides not only in new technologies and enhanced connectivity but also in the fact that openly accessible knowledge is now increasingly shared by the masses of learners themselves. Large groups of participants interact from different places and different points in time via a shared virtual space, and their interaction revolves around the creation of shared artifacts. These artifacts often represent a digital knowledge base with a network structure. Direct social interaction for reaching common understanding is largely infeasible under these circumstances (Larusson and Alterman 2009).

Mass collaboration bears three implications for CSCL research regarding collective knowledge bases: (1) The focus should incorporate the interplay between knowledge as substance (i.e., artifacts with meaningful content and interrelations) *and* as participatory activity (i.e., interactive contribution processes). (2) A knowledge base must be studied at the macro level, as it emerges in self-organized, long-term, interactive processes distributed across a large number of people. (3) The network perspective provides a multifaceted methodological approach to a knowledge base as a network of artifacts.

In the study reported here, we used data from the German version of the online encyclopedia Wikipedia, an outstanding example of artifact-based mass collaboration on the Web, to explore a collaboratively created knowledge base (for an extensive review of the large body of publications on the subject, see Okoli et al. 2012). It is a dynamic complex system of interconnected articles deliberately co-produced and modified by collaborative activities. With its large amount of data on the history of articles and authors' contributions, it offers a unique field for studying large-scale, open-ended collaborative processes. The contributions of two authors to the same article may take place years and hundreds of other authors' contributions apart. So—although authors often coordinate their work over article talk pages, that is, discussion threads, and over numerous other channels (Pentzold and Seidenglanz 2006)—a substantial part of the work is coordinated through the dynamically changing article itself. The written content mediates shared understanding on a specific topic, amalgamating views and styles of expression of a multitude of authors into a coherent exposition.

Although Wikipedia is not aimed at "inventing" new knowledge, or at providing a learning environment for the contributors, the processes that unfold in the online encyclopedia have been found to be essentially equivalent to scientific progress and knowledge-building discourse (Cress and Kimmerle 2008; Forte and Bruckman 2006; Kimmerle et al. 2011a; Swarts 2009). The choice and argumentative composition of facts and citations from external sources produce an original knowledge artifact with every article. Obviously, Wikipedia is not just a trivial aggregation of external information, and its articles represent more than just links to the original sources. From the perspective of CSCL research, the complex knowledge-related collaborative activities on Wikipedia are interesting along with the developing knowledge base of mediating artifacts (Cress and Kimmerle 2008;

Halatchliyski et al. 2011), which is a novel product for the online community and the general public, irrespective of Wikipedia's "no original research" policy.[1]

In order to tackle the large-scale dimensions, we employ the concept of a network and the approach of network analysis to the set of interconnected artifacts in two adjacent knowledge domains. Our goal is to exemplify the application of network analysis to the structure of a knowledge base of an online community and relate it to the contribution activity of its authors. Focusing on an article's topological position in the artifact network, we differentiate between two types of pivotal articles, that is, articles that are important for the structure of the knowledge base. An article may be pivotal either in the sense of being *central* within a knowledge domain or in the sense of being *boundary-crossing* between two domains. We examine to what extent different types of editing experience within the knowledge base are important explanatory variables for the contribution to pivotal articles (see Halatchliyski et al. 2010; Sosa 2011).

In the following we briefly recap theory trends in the field of CSCL, integrating views on both collective knowledge as substance and as participatory activity. Based on this theoretical foundation, we discuss the opportunities and challenges of studying collective knowledge in the context of the recent phenomena of mass collaboration and knowledge base networks. We then introduce our research approach based on network analysis metrics, in order to deal with these new challenges in CSCL research. Subsequently, we provide findings from an empirical study on pivotal articles and their contributors, within the artifact network of two adjacent knowledge domains in Wikipedia. Finally, we discuss the implications of our findings for future CSCL research.

Perspectives on collective knowledge

Theories on collective processes of intersubjective meaning-making (Dillenbourg et al. 1996; Koschmann 2002) have left behind individual cognition in order to focus on participation in community practices, negotiation of meanings, and building of shared understanding. Following the so-called *participation metaphor* (Sfard 1998), learning and knowing are depicted as socially shared activities that cross the conceptual boundary from one to the other (see also Scardamalia and Bereiter 1994). Knowing then consists of people's activities and practices that correspond with the specific physical and social context of a situation (Lave 1988; Suchman 1987). Accordingly, collaborative learning and knowing have been placed at the level of *group cognition* by Stahl (2006), emphasizing that they cannot be reduced to the level of cognitive representations and discussion statements of single individuals (see also Koschmann 2002).

Stahl's (2006) model integrates these levels of individual learning and collective knowledge into an activity system consisting of artifacts, utterances, and interactions as focal points. The sequence of referencing and defining interactions of the individual participants in a particular context continually produces and modifies a network of shared interconnected meanings for the group. Meaning is grounded in the relative positions in this network of mutual references and is not statically attached to physical artifacts or even words. Nevertheless, the meaning-making process is supported by the use of artifacts and words, which have predefined meanings from past discourse activities and which may again become subject to recurrent negotiation by the group participants. Thus, collaboration involves participative interaction along with the creation and reuse of meaningful artifacts,

[1] http://en.wikipedia.org/wiki/Wikipedia:No_original_research

which may often have a physical representation, or may be the focus of collaboration, as argued in the next section. Collective knowledge should then be defined not only as activity (i.e., knowing), but also as substance (i.e., shared artifacts). Knowledge as substance generally manifests itself in the emergent pattern of shared interconnected meanings. Analogously, Wikipedia consists of the collaborative activities and practices of its authors *and* of the networked knowledge base, which can be thought of as snapshots of meaningful structure in constant development. The network structure of referenced artifacts can also be studied with attention to its dynamics.

Both participant interaction and the dynamics of developing artifact shape and content are complementary aspects of the meaning-making process (Hakkarainen et al. 2011; Paavola and Hakkarainen 2009). Two research endeavors explicitly acknowledge the relevance of collaborative creation, use, and transformation of artifacts as epistemic objects (see also Knorr-Cetina 2001) in CSCL: The *metaphor of knowledge creation* (Paavola et al. 2002, 2004) designates artifacts as the goal and the product of collaborative learning. The *co-evolution model of cognitive and social systems* (Cress and Kimmerle 2008) shows how collective knowledge develops with the changing artifact content in the context of a wiki (see also Kimmerle et al. 2011b; Moskaliuk et al. 2009, 2012). The development is presented as successive co-evolution cycles of internalization (i.e., individual learning) and externalization (i.e., creation of collective knowledge; Kimmerle et al. 2010a).

In the present work, our aim is to advance the perspective that—in contrast to the analysis of interaction sequences—artifacts and their meaningful interconnected structure offer a unique way of operationalizing knowledge-related processes in collectives. Maintaining the research focus at the intersubjective level, we extend the concept of collective knowledge to long-term processes and large-scale network structures.

Artifact-based mass collaboration

In line with the participation metaphor of situated learning and knowing, the predominant methodological approach in CSCL has been to study small groups of students in a neatly arranged situation: The students engage in synchronous discourse around a problem-solving task, and the sequence of their interactions represents a major research interest. Lipponen (2002), however, contested the popular definition of collaboration as “a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem” (Roschelle and Teasley 1995, p. 70), because it puts narrow constraints on the object of study. Suthers (2006) also stated that small groups do not deliver an exhaustive picture of collective knowledge processes. Jones et al. (2006) argued for broadening the research focus on collaborative learning to include aspects of networked learning enabled by large-scale technological infrastructures on the Web. In fact, complex knowledge phenomena involve longer periods of time, larger and changing numbers of people, and fuzzy-structured settings (Kapur et al. 2007). In this spirit, any human achievement can be seen as a collaborative accomplishment—in terms of the metaphorical *dwarfs standing on the shoulders of giants*. Extending the view beyond problem-solving small groups enables a macro approach to the complexity of knowledge development across space, time, and collectives of people. This global level of human learning and knowledge creation has rarely been addressed by CSCL research (see Kafai and Peppler 2011).

This large-scale perspective brings to the foreground the connecting role of artifacts in the collaborative process. Bearing in mind that most of the individuals among a vast number of participants cannot interact directly or do not even know each other, intersubjective

JrnlID 11412_ArtID 9182_Proof# 1 - 25/09/2013

understanding and coordinated activities are facilitated by artifacts. This is even more so when the individuals follow a common goal, as in the case of Wikipedia. Each individual must take account of the perspective of the others to contribute by building on the accomplishments of others. Collaborative artifacts represent crystallized knowledge that is preserved from past interactive situations, and that can be built on in the future, giving rise to phenomena like scientific understanding, social practices, and rules. This mechanism of indirect interaction is also referred to as *stigmergy*, where the artifacts created or modified by some individuals stimulate the subsequent activity of other individuals (Susi and Ziemke 2001). Knowledge-related practices in Web 2.0 contexts fall under the participation metaphor of learning and additionally accentuate the creation of knowledge artifacts (Dohn 2009). This view suggests integrating the two perspectives of artifacts as both means and ends of collaboration (see e.g., Kafai and Resnick 2000) and also suggests studying the interplay between knowledge as substance and as participatory activity. In sum, artifact-based mass collaboration develops as a self-organized process around and through the created content, which reduces the need for direct coordinating interactions between the participants.

Networks of knowledge

The study of a mass collaboration knowledge base presupposes an approach that can encompass its macro structure and other large-scale and long-term characteristics. At the same time, it is desirable not to leave the level of artifacts, individuals, and small groups out of focus. According to the actor network theory (Latour 2005), the analysis of social phenomena, of which mass collaboration is an example, should focus on the patterns of mutual influence in the network of actants (i.e., humans as well as artifacts endowed with equal agency). The fundamentals of such a multifaceted approach are provided by the *network* concept.

A network can be defined as a set of dynamically connected nodes that represent units of the same kind, such as persons or knowledge artifacts. The concept has already been used to describe knowledge organization at different levels. The semantic memory of individuals has classically been portrayed as a network of associated knowledge representations (e.g., Collins and Loftus 1975). Stahl (2006) has advanced the idea of networks of references to explain how collective knowledge is created through group discourse activities. In the context of mass collaboration environments like Wikipedia, knowledge is organized in a network of interlinked artifacts (Voß 2005).

Computer technology directly promotes the creation of networked knowledge in a number of ways. The Web itself represents a technological network that maintains hyperlinked information of various kinds. Due to the flexibility of hypertext the recipient can "jump" in multiple directions through the content and combine relevant aspects from different contexts, discerning new meanings (Moskaliuk and Kimmerle 2009). The increased interactivity afforded by Web 2.0 applications also makes network structures and user-generated content prominent. Correspondingly, an increasing number of hyper-structured knowledge bases have emerged from the collaborative activity of a mass of individuals.

The network concept suitably highlights the emergent character of knowledge. According to the theory of conceptual integration and blending (Fauconnier and Turner 2002), the creation of new meanings and knowledge can be thought of as recombination of different existing ideas. Knowledge essentially emerges from the specific way in which various meanings are connected, like nodes in a complex network that can build an infinite number of interconnection patterns. Although the network concept connotes a structural approach, it does not imply a static view on knowledge. Networks are constantly changing as neither their nodes nor their links are enduring entities. Large-scale collective dynamics lead to the bottom-up development of patterns typical of complex systems (Kapur et al. 2007). These patterns then have a top-down impact on the local relations and interactions among individuals and knowledge artifacts.

Based on the network concept, network analysis (see Newman 2010) offers methodological tools to begin dealing with the complex large-scale and long-term patterns in the knowledge base of a mass collaboration environment.

The network analysis approach

Network analysis is a multidisciplinary research approach for examining relational patterns among physical and digital, human and non-human entities. It includes a variety of methodological concepts and instruments to identify, describe, analyze, and visualize positions, relations, clusters of elements, and global network properties. The approach was greatly advanced by sociologists who studied networks of people under the term *social network analysis* (SNA; Wassermann and Faust 1994), but their concepts and methods largely represent mathematical abstractions and are applicable to other kinds of networks. Some of the major applications are: detecting important actors, subgroups, and the actors bridging them; characterizing the position of different artifacts within a network; measuring information paths and flows.

SNA has become an increasingly common method in CSCL research (e.g., Aviv et al. 2003; Cho et al. 2002; Goggins et al. 2012; de Laat et al. 2007; Kimmerle et al. 2013; Palonen and Hakkarainen 2000; Reffay and Chanier 2002; Ryberg and Larsen 2008). Analyses of online social networks are usually based on the logged collaborative interaction between learners that is mediated through a shared digital environment. For example, a network link between two people may mean that the one has read or responded to a contribution of the other, but more indirect relations like the co-presence in a discussion are also possible. Such analyses may yield information on the cohesiveness of learning groups and on the position of individual students relative to the others, at different points in time and overall.

As argued in the previous sections, in addition to the knowledge-related activities of the participants, CSCL research should also incorporate the body of collaborative knowledge artifacts into the analysis. This is especially relevant for a mass collaboration environment, such as Wikipedia, that is directed at creating a knowledge base. The patterns in such a networked body of knowledge artifacts can be appropriately studied with network analysis methods in analogy to bibliometric citation analysis of scientific work (see Glänzel 2003). Mass collaboration manifests itself in knowledge artifacts linked by hyperlinks, similar to scientific papers connected through citations. The emerging learning analytics discipline (Fournier et al. 2011; Siemens 2012) might be a promising field for adapting borrowed bibliometric approaches to networked learning and mass collaboration (see for example Author 2013). Q5

Only a few CSCL studies have analyzed networks of collaboratively created artifacts with content relations. Both Sha et al. (2010) and Oshima et al. (2007) applied automatic algorithms for the identification of semantic relations between the content of artifacts in order to define a so-called semantic network of contributions, to calculate general indices of the network and to cluster the topics of the contributions. Kimmerle et al. (2010b) investigated the development of clusters in a network of Wikipedia articles related to the topic of schizophrenia over a period of five years. They found evidence for co-evolution of the artifact network and the contribution interest of authors over time. Halatchliyski et al. (2010) examined an article network of two adjacent knowledge domains in Wikipedia and identified a group of experienced, boundary-spanning authors who influenced domain integration. The present study extends this approach by relating the concept of pivotal artifacts in a knowledge base to the activity characteristics of the contributing authors. Keegan et al. (2012) also used authors' editing experience to study the collaboration patterns on Wikipedia articles about breaking news.

In sum, the macro perspective on knowledge networks reveals a unique and largely unexplored field within CSCL research. Correspondingly, we argue that network analysis is an appropriate methodological approach when taking this perspective. In the following sections, we present a study with Wikipedia data in which we employ two types of measures of topological centrality to identify pivotal articles in artifact networks: the one captures well-connected artifacts that have important positions within a single knowledge domain; the other accents boundary-crossing artifacts that have an interconnecting position between two knowledge domains. Based on these indices we examine the relation between the authors' editing experience and their contribution to pivotal articles in the knowledge base.

Empirical study

Focusing on two adjacent knowledge domains in Wikipedia, the following study seeks to explain the contribution to pivotal articles in the artifact network of a knowledge base through the editing experience of its authors. Experience in this sense does not designate some scientific or professional expertise but simply the count of an author's content contributions to the investigated knowledge domains. *Pivotal* articles were those with a *central* network position within a single knowledge domain or those with a *boundary-crossing* network position between two knowledge domains. The study includes two levels of analysis: At the level of artifacts, we perform a network analysis on hyperlinked articles, which are categorized a priori in two adjacent knowledge domains. We test our hypotheses at the level of authors by relating their editing experience to their contribution to articles with pivotal network positions.

Level of artifacts The body of knowledge artifacts in a mass collaboration environment may be divided into knowledge domains. The relevant artifacts in the current study were Wikipedia articles, and a *knowledge domain* was a set of articles that had been assigned to the same Wikipedia category, corresponding to a scientific discipline. Hence, our approach bears similarities to scientometric research on the scientific work in neighboring disciplines. The Wikipedia category system is a collaboratively created taxonomy with a nearly hierarchical structure of supra- and sub-categories. Any author can change what category is assigned to an article or a sub-category, and articles are often annotated with multiple categories (Kittur et al. 2009). Accordingly, article categorization is an emergent characteristic of the mass collaboration environment and reflects the logic of the represented knowledge. It is independent of the article network structure and the authorship of articles. Based on the a priori Wikipedia categorization, we chose two knowledge domains for our study. We then distinguished *specialized articles*, which belonged to only one of the two knowledge domains, and *intersection articles*, categorized under both knowledge domains.

Exploring the network structure of a knowledge base at the macro level of knowledge domains, we focused on identifying articles with pivotal network positions. We distinguished between pivotal articles that are central within one knowledge domain and pivotal articles that cross the boundary between two knowledge domains. In our reasoning, both types of articles may be important, supporting on the one hand the internal knowledge organization within a domain and, on the other hand, the interdisciplinary connections between domains.

Therefore, we defined two *separate networks* that corresponded to the hyperlinked specialized and intersection articles in each of the two domains. We also defined a *combined network,* including all the articles and their hyperlinks in both domains taken together. Network nodes represented articles, and network edges represented the hyperlinks not accounting for the direction (as we aimed at examining the relatedness of the articles and not browsing behavior).

Pivotal articles within a knowledge domain were operationalized by applying the eigenvector centrality index (Bonacich 1972) to the articles in the separate networks. This measure characterizes the connectedness of an article relative to all the others in the network: Articles with more direct connections to other well-connected articles obtain higher values. These central articles contain knowledge that is highly significant for a domain. A similar measure is employed by the PageRank algorithm of the Google search engine for ranking the importance of web pages (Page et al. 1998).

Pivotal articles that cross the boundary between two knowledge domains were operationalized by applying the betweenness centrality index (Freeman 1979) to all the articles in the combined network.[2] This measure characterizes the bridging position of an article among the other articles in both domains: Articles that are repeatedly part of the shortest connection between pairs of other articles obtain higher values. These boundary-crossing articles link the two domains and enable knowledge transfer and integration across their boundaries.

Level of authors This level was not a part of the network analysis, which only pertained to the articles and their hyperlinks. Based on the history of contributions to the articles included in the first level of the analysis, we determined the relevant authors and their experience. We used two aspects of experience—*article experience* (i.e., the count of individual articles in both domains an author had contributed to) and *domain experience*. Regarding domain experience, authors were classified into groups of *specialists* (i.e., authors who contributed to one of the domains but not to the other) and *generalists* (i.e., authors who contributed to both domains). As we investigated two domains, there were also two groups of specialists. Generalists were grouped into *intersection generalists* (i.e., authors who have contributed to at least one intersection article, which appeared in both domains) and *non-intersection generalists* (i.e., authors who have contributed to specialized articles in each of the domains but to none of the intersection articles). For purposes of illustration, Table 1 incorporates our definitions into an example. The rows represent articles which belong either to knowledge domain A, to knowledge domain B, or to both domains A and B (intersection articles). According to the definition of domain experience, author 1 is a specialist in A, author 2 is a specialist in B, author 3 is an intersection generalist, and author 4 is a non-intersection generalist. The last row in the table shows the article experience of each of the four authors as the count of articles an author has contributed to.

[2] Both centrality indices were originally developed in SNA research and also used in various other networks (see Leydesdorff 2007)

Table 1 Grouping of authors according to contribution activity and articles' categorization

Articles' a priori categorization	Author 1: specialist in A	Author 2: specialist in B	Author 3: intersection generalist	Author 4: non-intersection generalist
A	x			
A	x			x
A	x			
A & B			x	
A & B			x	
B		x		
B		x		x
Article experience of an author	3	2	2	2

At the level of authors, we determined the relation between authors' experience and their contribution to pivotal articles by building on the measures of pivotal articles in the networks. We calculated author-level aggregate measures of the average centralities—once for eigenvector centrality and once for betweenness centrality—of the articles an author had contributed to. So, an author inherited the averaged centrality of the articles she or he had co-authored. We did this for the combined network as well as for each of the two separate networks independently. The *important* authors within a knowledge domain are those that have the highest aggregated eigenvector centrality, based on the articles they have contributed to. Correspondingly, the *boundary spanners* (Tushman and Scanlan 1981) between two knowledge domains are those authors that have the highest aggregated betweenness centrality based on the articles they have contributed to. They act as gatekeepers at the boundary between two knowledge domains, driving knowledge exchange and integration.

Hypotheses

The goal of the study was to simultaneously investigate the partial effect of authors' *article experience* and their *domain experience* as explanatory variables of their contribution to pivotal articles within and between knowledge domains. Our hypotheses therefore concerned the author level of analysis.

While boundary-spanning contributors might not necessarily have a prominent role within the domains, by definition they should be experienced in both domains (Levina and Vaast 2005). Consequently, we derive the following hypotheses:

Hypothesis 1a Specialists contribute on average to more central, better-connected articles in each of the knowledge domains than generalists. Thus, specialists have a high aggregated *eigenvector* centrality derived from the *separate* domain networks compared with generalists.

Hypothesis 1b Generalists act as boundary spanners and contribute on average to more boundary-crossing articles between both domains than specialists. So, generalists have a high aggregated *betweenness* centrality derived from the *combined* domain network compared with specialists.

Besides domain experience, we expect that authors' article experience (i.e., the count of articles in both domains an author has contributed to) is also a significant explanatory variable of the contribution to pivotal articles. According to the concept of legitimate peripheral participation (Lave and Wenger 1991), experienced authors are expected to have a significant influence in a mass collaboration environment by contributing to pivotal articles within and between knowledge domains:

Hypothesis 2a Authors' article experience is a significant predictor of the contribution to central, well-connected articles, so it is positively related to the aggregated *eigenvector* centrality of authors derived from the *separate* network of each of the knowledge domains.

Hypothesis 2b Authors' article experience is a significant predictor of the contribution to boundary-crossing articles, so it is positively related to the aggregated *betweenness* centrality of authors derived from the *combined* network of both knowledge domains.

In order to estimate the partial effects of article and domain experience, Hypothesis 1a and Hypothesis 1b were simultaneously tested with one model for each of the two knowledge domains. Accordingly, Hypothesis 2a and Hypothesis 2b were simultaneously tested with one model for both domains taken together.

Data and method

We studied the contribution to central pivotal articles within and boundary-crossing pivotal articles between the two a priori delimited knowledge domains *psychology* and *education*, using the categorization system of Wikipedia. Our data was sourced from an official dump file of the German Wikipedia (http://dumps.wikimedia.org/dewiki), containing a snapshot of its state as of January 16, 2012. We chose to study all articles categorized as topics of psychology (German: "Psychologie") or education ("Pädagogik"), as well as all their subcategories. The sample represented two knowledge domains with a similar number of articles and obvious content relations.

Level of artifacts We considered three types of articles in the analysis: 5,085 specialized psychology articles, 4,696 specialized education articles, and 731 intersection articles (i.e., those categorized under both domains). Using eigenvector centrality we measured how well-connected and thus central an article was within each of the two separate domain networks (a total of 5,816 articles in the psychology network and 5,427 articles in the education network). The extent to which an article was boundary-crossing between both domains was measured with its betweenness centrality in the combined network (10,512 articles in total). The higher the eigenvector or betweenness centrality value of an article, the more pivotal was the position of the article within one of the domains or between the two domains. The network analysis measures were calculated with the *igraph* package for R (Csárdi and Nepusz 2006).

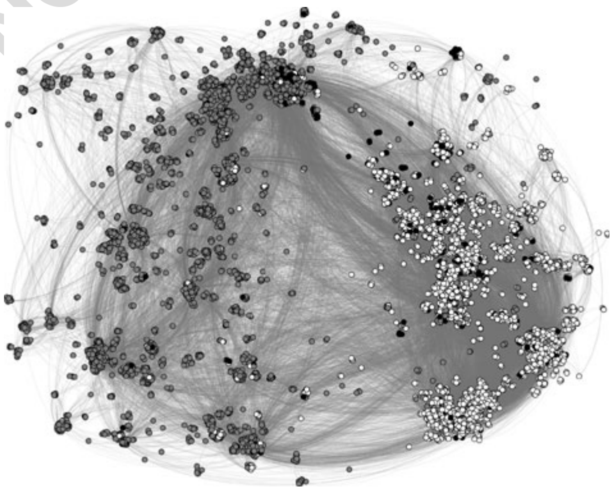
Level of authors We first excluded from the analysis contributions marked as minor, or made by anonymous authors or bots, deletions, reverts to a previous state of the articles, as well as contributions shorter than 150 characters. We also excluded the contributions of administrators and reviewers. Although they contribute a lot of content, their choice of articles and mode of contribution are different and depend on their Wikipedia control tasks. The remaining contributions were made by a total of 8,040 signed-in authors writing in one or both the domains. According to our taxonomy of author groups (see Table 1) we identified 3,980 psychology specialists, 2,762 education specialists, 1,002 intersection generalists, and 296 non-intersection generalists.

In the last stage of the analysis at the level of authors we aggregated article measures from the network analysis as an average over the articles an author had contributed to. This procedure resulted in two types of values: the eigenvector centrality of an author in a separate network, measuring how important the total contribution of an author within one domain was; and the betweenness centrality of an author in the combined network, measuring the extent to which an author contributed as a boundary spanner between domains. These aggregate measures enabled us to simultaneously investigate the partial significance of article and domain experience of an author as explanatory variables of his or her contribution to pivotal articles.

Results

Before we present the tests of the hypotheses (which concern the level of authors), we first provide the most relevant results from the analysis at the level of articles. Figure 1 depicts the combined network of articles in both knowledge domains. The grey dots represent education articles, the white ones psychology articles, and the black dots show intersection articles. The curved lines display the hyperlinks between the articles. The visualization was made with the Gephi platform (Bastian et al. 2009) using the OpenOrd layout algorithm that organizes the dots according to their interconnections. Thus, a number of dots that have direct connections to each other are represented as a cluster. Over ten repetitions of the algorithm the produced layouts were very similar.

It is interesting to note in the layout that both adjacent domains are clearly distinguishable as two separate parts in the combined network. The intersection articles are dispersed among both the education and the psychology parts of the network and do not form a homogenous network cluster. Some of the intersection articles have more connections to psychology articles and others are more tightly bound to education articles.



Q1 **Fig. 1** The combined network of Wikipedia articles in education and psychology

We found moderate rank correlations between articles' eigenvector centrality in the education network and betweenness centrality in the combined network (τ=.53, p<.001) as well as between eigenvector centrality in the psychology network and betweenness centrality in the combined network (τ=.43, p<.001). In other words, boundary-crossing articles between the two domains are not necessarily central pivotal articles in either of the domains.

We corroborated this finding with independent-samples unequal-variances t-tests comparing the group of intersection articles with the specialized articles. Both betweenness and eigenvector centrality had distributions strongly skewed to the right, that is, only a few articles had high values, and the majority of them had very low values. We applied a logarithmic transformation to these variables in order to make them better fit the assumptions of the *t*-test. As expected from their definition, intersection articles were shown to be boundary-crossing articles in the combined network, with a significantly higher mean log betweenness centrality than that of specialized articles: M_{int}=7.01, SD=3.36 vs. M_{spec}=5.50, SD=3.95; $t(887.9)$=11.60, p<.001. Thus, a specialized article was less likely to occupy a boundary-crossing position between the domains than an intersection article. In support of our reasoning on the moderate correlations between eigenvector and betweenness centrality, intersection articles were shown to be less important in both separate networks, with a significantly lower mean log eigenvector centrality than that of specialized articles; in the education network: M_{int}=−4.95, SD =1.74 vs. M_{spec}= −4.64, SD=1.44; $t(892.5)$=−4.60, p<.001; in the psychology network: M_{int}=−4.64, SD =1.44 vs. M_{spec}=−4.31, SD=1.37; $t(928.6)$=−5.72, p<.001. Thus, a specialized article was more likely to occupy a central position in its domain than an intersection article.

We now turn to the analysis at the author level and the results of the main hypothesis tests. We excluded authors with article experience of less than 2, in order to enable fair comparisons between the groups of generalists and specialists. We did this because non-intersection generalists by definition have a minimal article experience of 2, as they have contributed to at least one education and one psychology article.[3] Our sample was reduced to 1,663 authors (640 psychology specialists, 292 education specialists, 435 intersection generalists, and 296 non-intersection generalists). We used three ANCOVA models—two for the contribution to pivotal articles within each of the two domains and one for the contribution to pivotal articles between the domains. Both article experience and domain experience of an author were included in the models as predictors of the extent to which the author contributed to pivotal articles. Thus, their incremental predictive value could be simultaneously estimated. Again, we applied a logarithmic transformation to the continuous variables betweenness centrality, eigenvector centrality, and article experience, whose distributions were strongly skewed to the right, in order to make them better fit the preconditions of the ANCOVA. Article experience entered the models as a continuous predictor; domain experience was modeled as intercept dummy variables. The coefficients of these dummy variables directly indicated the differential effect of the generalist groups compared with a specific group of specialists for equal levels of article experience of specialists and generalists. The basic model for the three networks was:

$$W_i = \alpha + \beta X_i + \gamma Y_i + \delta Z_i + \varepsilon_i$$

where

W_i predicted log betweenness or log eigenvector centrality of author i
X_i 1 if author i is an intersection generalist, 0 otherwise

[3] Additional t-tests comparing the excluded intersection generalists and specialists with article experience of 1 corresponded with the results reported in the following.

Y_i 1 if author i is a non-intersection generalist, 0 otherwise
Z_i log article experience of author i
ε_i error term.

Hypothesis 1a assumed that specialists contribute on average to more central articles in each of the domains and thus have a higher aggregated *eigenvector* centrality derived from the *separate* domain networks compared with generalists. This assumption was partially supported for intersection generalists; in the education domain: $\beta=-0.15$, $t(1019)=-1.41$, $p=.159$; in the psychology domain: $\beta=-0.36$, $t(1367)=-5.04$, $p<.001$. It was fully supported for non-intersection generalists; in the education domain: $\gamma=-0.28$, $t(1019)=-2.45$, $p=.015$; in the psychology domain: $\gamma=-0.35$, $t(1367)=-4.35$, $p<.001$. Consequently, the overall effect of domain experience was marginally significant in the education domain ($F(2, 1019)=2.99$, $p=.051$) and significant in the psychology domain ($F(2, 1367)=16.27$, $p<.001$).

Hypothesis 1b assumed that generalists act as boundary spanners (i.e., contribute on average to more boundary-crossing articles) and thus have a high aggregated *betweenness* centrality derived from the *combined* domain network compared with education and psychology specialists taken together. This assumption was supported as well; for intersection generalists: $\beta=0.54$, $t(1659)=4.46$, $p<.001$; for non-intersection generalists: $\gamma=0.31$, $t(1659)=2.29$, $p=.022$; with a significant overall effect of domain experience: $F(2, 1659)=10.45$, $p<.001$.

Hypothesis 2a assumed that article experience is a significant predictor of aggregated eigenvector centrality of the authors derived from the separate domain networks. This was supported for both knowledge domains; in the education domain: $\delta=0.38$, $t(1019)=5.19$, $p<.001$; in the psychology domain: $\delta=0.30$, $t(1367)=5.88$, $p<.001$.

Hypothesis 2b assumed that article experience is a significant predictor of aggregated betweenness centrality of the authors derived from the combined network of both knowledge domains taken together. This assumption was also supported ($\delta=0.60$, $t(1659)=6.75$, $p<.001$).

In sum, our hypotheses were largely confirmed except for a non-significant difference in the expected direction between education specialists and intersection generalists in the education domain (Hypothesis 1a). We found no significant interaction effects between article and domain experience, that is, there was no difference in the impact of article experience among the four groups of generalists and specialists. The reported results were confirmed by testing conservative ANCOVA models, using ranks (ordinal transformation) instead of the log transformed article experience, betweenness and eigenvector centrality.

Discussion

In the empirical study reported here our aim was to explain the authors' contribution to pivotal articles in the artifact network of two Wikipedia knowledge domains in relation to domain experience and article experience of the collaborating authors. Specialists (i.e., authors with contribution experience in only one of the domains) were expected to contribute on average to more central pivotal articles in each of the separate domains than generalists (i.e., authors with contribution experience in both domains). Generalists were expected to act as boundary spanners by contributing on average to more boundary-crossing pivotal articles between both domains than specialists. We further expected that article experience (i.e., the total number of articles an author has contributed to) was positively related both to the contribution to central articles within each of the two knowledge domains, and to the contribution to boundary-crossing articles between both knowledge domains.

The hypotheses of the study were supported by the empirical results. We found that both domain experience and article experience of an author are significantly related to the contribution to pivotal articles in the artifact network. Even the single non-significant result tended to be consistent with the hypothesis that education specialists would contribute to more central pivotal articles in the education domain than intersection generalists. Intersection generalists were defined as authors with at least one contribution to an intersection article. In this respect, we found that intersection articles were boundary-crossing articles between domains and were responsible at least to some extent for the integration of knowledge across the domain boundaries. However, they were not so central and important within each of the two particular domains. Thus, the non-significant difference must be the consequence of other very central specialized articles in the education domain to which the intersection generalists had contributed. Even so, education specialists contributed on average to more central articles in the education domain than intersection generalists. Furthermore, as intersection articles turned out to be boundary-crossing articles, it is unsurprising that intersection generalists proved to be boundary spanners between the domains. However, non-intersection generalists also proved to be boundary spanners, confirming the significance of experience in both domains for the contribution to boundary-crossing articles.

Thus, our results suggest several principles of contribution to pivotal articles at domain level in a knowledge base: As we distinguished between pivotal articles that are central within a single knowledge domain and those that cross the boundaries between two domains, a difference between the authors who contributed to these two types of pivotal articles became evident. This division of roles in the mass collaboration process is related to the domain experience of the authors. Specialized experience in one domain goes together with contributions to central pivotal articles in that domain. Generalized experience in two domains goes together with contributions to pivotal articles that cross the boundaries between the domains. At the same time, the article experience of an author, regardless of the domain experience, is positively related to the contribution to both types of pivotal articles.

The reported results built on and enhanced our previous investigations (Halatchliyski et al. 2010) into knowledge construction in the context of a different pair of domains in Wikipedia. By differentiating two types of authors' experience we can now show that authors with experience in only one domain are not peripheral. These specialists play an important role in a mass collaboration environment, as their contribution is central within that knowledge domain. By isolating the relative significance of the explanatory variables domain experience and article experience, our understanding of the contribution to pivotal artifacts is now more differentiated. Generalists tend to contribute to boundary-crossing articles between domains but they are just as likely to contribute to very central articles within each of the domains, if they have a high article experience, that is, if these generalists contribute to a large number of articles. Accordingly, specialists tend to contribute to central articles within their domain but they might also act as boundary spanners and contribute to boundary-crossing articles between domains, if they have a high article experience.

While drawn from the limiting perspective of two of the knowledge domains in Wikipedia, psychology and education, these results indicate a division of labor between generalists and specialists and a broad significance of the contribution experience of the collaborators. From a design point of view this speaks for the need of a general participation encouragement and empowerment of the long tail in networked environments. As a great number of the participants typically make few and isolated contributions, it is vital for the mass collaboration process to attract repeated contributions and commitment to pivotal artifacts. This can be facilitated at many levels of the design of an environment, from lowering the usability threshold of active participation to developing incentive systems to stimulate voluntary contributions.

Conclusion

This paper conveys a two-fold contribution to CSCL research. It provides evidence for the significant relation between authors' experience and their contribution to pivotal artifacts at the level of knowledge domains in Wikipedia. It also provides an example of an integrative theoretical perspective within CSCL that views collective knowledge both as substance (i.e., collaborative artifacts) and as participatory activity (i.e., collaborative contributions). In accordance with this perspective, we took a multi-layered approach incorporating analysis at the level of artifacts and at the level of authors. Our approach is appropriate for the self-organized, long-term and large-scale process of mass collaboration that produces a dynamic networked knowledge base of artifacts and their interconnections. Besides wikis, other multi-user virtual environments for learning, such as massive open online courses (MOOCs), or for gaming represent promising research contexts where our approach may be applicable. The condition is to identify a network of collaborative artifacts that is open to further interactive development by the participants. Such contexts may be different from formal education as learners self-regulate their motivation to participate and to achieve goals.

Considering that knowledge building in small-group settings also manifests itself in the creation of shared artifacts (Paavola and Hakkarainen 2009; Scardamalia and Bereiter 1994), it could be worthwhile to extend our approach to integrate the results of small-group and mass collaboration research into a general theoretical framework. Surely, this would suggest combining the structural approach of network analysis, which is useful to discern abstract patterns, with content and interaction analysis techniques, which can supply richer interpretation of the observed patterns.

Another direction for future research would be to augment our approach with temporal aspects of knowledge development by analyzing an artifact network at different points in time. A dynamic network analysis has been shown to yield further insights into the essentially temporal collaborative process (see e.g., Halatchliyski et al. 2012, 2013). Therefore, it would be interesting to examine the longitudinal aspects of the knowledge contained in pivotal articles in a knowledge base. As a structural backbone, such pivotal knowledge may be an important factor directing the development of new knowledge.

In line with the suggestions for further research and extension of the presented approach, we reassert our view that CSCL research should take a detailed account of the recent phenomenon of mass collaboration. The CSCL research community needs consider the increasing impact of mass collaboration on learning and knowledge creation. In our opinion, CSCL research would benefit from treating a collaborative artifact not only as a means of interaction support in small groups but also as a goal of the creation process within self-organized communities and networks. With the adoption of a network perspective, large-scale structures and long-term processes of knowledge development become accessible for investigation.

References

Aviv, R., Erlich, Z., Ravid, G., & Geva, A. (2003). Network analysis of knowledge construction in asynchronous learning networks. *Journal for Asynchronous Learning Networks, 7*, 1–23.

Bastian, M., Heymann, S., & Jacomy, M. (2009). *Gephi: An open source software for exploring and manipulating networks. Proceedings of the Third International Conference on Weblogs and Social Media* (pp. 361–362). Menlo Park: AAAI Press.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology, 2*, 113–120.

Cho, H., Stefanone, M., & Gay, G. (2002). Social network analysis of information sharing networks in a CSCL community. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of the Computer-Supported Collaborative Learning Conference* (pp. 43–50). Mahway: Lawrence Erlbaum Associates.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Cress, U. (2013). Mass collaboration and learning. In R. Luckin, P. Goodyear, B. Grabowski, S. Puntambekar, J. Underwood, & N. Winters (Eds.), *Handbook on design in educational technology* (pp. 416–425). London: Taylor and Francis.

Cress, U., & Kimmerle, J. (2008). A systemic and cognitive view on collaborative knowledge building with wikis. *International Journal of Computer-Supported Collaborative Learning, 3*, 105–122.

Cress, U., Barron, B., Halatchliyski, I., Oeberst, A., Forte, A., Resnick, M., & Collins, A. (2013). Mass collaboration—an emerging field for CSCL research. In N. Rummel, M. Kapur, N. Nathan, & S. Puntambekar (Eds.), *To see the world and a grain of sand: Learning across levels of space, time and scale: CSCL 2013 Proceedings* (pp. 557–563). Madison: International Society of the Learning Sciences.

Csárdi, C., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems, 1695*.

de Laat, M., Lally, V., Lipponen, L., & Simons, R.-J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis. *International Journal of Computer-Supported Collaborative Learning, 2*, 87–103.

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In H. Spada & P. Reiman (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning science* (pp. 189–211). Oxford: Elsevier.

Dohn, N. (2009). Web 2.0: Inherent tensions and evident challenges for education. *International Journal of Computer-Supported Collaborative Learning, 4*, 343–363.

Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.

Forte, A., & Bruckman, A. (2006). From Wikipedia to the classroom: Exploring online publication and learning. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of the 7th International Conference Of The Learning Sciences* (pp. 182–188). Mahwah: Erlbaum.

Fournier, H., Kop, R., & Sitla, H. (2011). *The value of learning analytics to networked learning on a personal learning environment. Proceedings of the 1st International Conference on Learning Analytics and Knowledge LAK'11* (pp. 104–109). New York: ACM Press.

Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks, 1*, 215–239.

Glänzel, W. (2003). Bibliometrics as a research field. A course on the theory an application of bilbiometric indicators. Budapest: Magyar Tudományos Akadémia Könyvtára. [http://www.citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.5311&rep=rep1&type=pdf].

Goggins, S., Valetto, G., Mascaro, C., & Blincoe, K. (2012). Creating a model of the dynamics of socio-technical groups. *User Modeling and User-Adapted Interaction, 22*, 1–35.

Granovetter, M. S. (1973). The strength of weak ties. *The American Journal of Sociology, 78*, 1360–1380. Q7

Hakkarainen, K., Ritella, G., & Seitamaa-Hakkarainen, P. (2011). Epistemic mediation, chronotope, and expansive knowledge practices. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Connecting computer-supported collaborative learning to policy and practice: CSCL2011 conference proceedings* (Vol. II, pp. 948–949). Hong Kong: International Society of the Learning Sciences.

Halatchliyski, I., Moskaliuk, J., Kimmerle, J., & Cress, U. (2010). *Who integrates the networks of knowledge in Wikipedia? Proceedings of the 6th International Symposium on Wikis and Open Collaboration (Article No. 1)*. New York: ACM Press.

Halatchliyski, I., Kimmerle, J., & Cress, U. (2011). Divergent and convergent knowledge processes on Wikipedia. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Connecting computer-supported collaborative learning to policy and practice: CSCL2011 conference proceedings* (Vol. II, pp. 566–570). Hong Kong: International Society of the Learning Sciences.

Halatchliyski, I., Oeberst, A., Bientzle, M., Bokhorst, F., & van Aalst, J. (2012). Unraveling idea development in discourse trajectories. 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* (Vol. II, pp. 162–166). Sydney: International Society of the Learning Sciences.

Halatchliyski, I., Hecking, T., Göhnert, T., & Hoppe, H. U. (2013). Analyzing the flow of ideas and profiles of contributors in an open learning community. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.), *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge, LAK '13* (pp. 66–74). New York: ACM.

Jones, C., Dirckinck-Holmfeld, L., & Lindstrom, B. (2006). A relational, indirect, meso-level approach to CSCL design in the next decade. *International Journal of Computer-Supported Collaborative Learning, 1*, 35–56.

Kafai, Y. B., & Peppler, K. A. (2011). Beyond small groups: New opportunities for research in computer-supported collective learning. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Connecting computer-supported collaborative learning to policy and practice: CSCL2011 conference proceedings* (Vol. I, pp. 17–24). Hong Kong: International Society of the Learning Sciences.

Kafai, Y., & Resnick, M. (2000). *Constructionism in practice: Designing, thinking, and learning in a digital world*. Mahwah: Erlbaum.

Kapur, M., Hung, D., Jacobson, M., Voiklis, J., Kinzer, C., & Chen, D.-T. (2007). Emergence of learning in computer-supported, large-scale collective dynamics: A research agenda. In C. A. Clark, G. Erkens, & S. Puntambekar (Eds.), *Proceedings of the International Conference of Computer-supported Collaborative Learning* (pp. 323–332). Mahwah: Erlbaum.

Keegan, B., Gergle, D., & Contractor, N. (2012). *Do editors or articles drive collaboration? Multilevel statistical network analysis of Wikipedia coauthorship. Proceedings of the 2012 ACM Conference on Computer-Supported Cooperative Work* (pp. 427–436). New York: ACM Press.

Kimmerle, J., Cress, U., & Held, C. (2010a). The interplay between individual and collective knowledge: Technologies for organisational learning and knowledge building. *Knowledge Management Research and Practice, 8*, 33–44.

Kimmerle, J., Moskaliuk, J., Harrer, A., & Cress, U. (2010b). Visualizing co-evolution of individual and collective knowledge. *Information, Communication and Society, 13*, 1099–1121.

Kimmerle, J., Moskaliuk, J., & Cress, U. (2011a). Using wikis for learning and knowledge building: Results of an experimental study. *Educational Technology & Society, 14*, 138–148.

Kimmerle, J., Moskaliuk, J., Cress, U., & Thiel, A. (2011b). A systems theoretical approach to online knowledge building. *AI & Society: Journal of Knowledge, Culture and Communication, 26*, 49–60.

Kimmerle, J., Thiel, A., Gerbing, K.-K., Bientzle, M., Halatchliyski, I., & Cress, U. (2013). Knowledge construction in an outsider community: Extending the communities of practice concept. *Computers in Human Behavior, 29*, 1078–1090.

Kittur, A., Chi, E. H., & Suh, B. (2009). What's in Wikipedia? Mapping topics and conflict using socially annotated category structure. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems* (p. 1509). New York: ACM Press.

Knorr-Cetina, K. (2001). Objectual practice. In T. R. Schatzki, K. Knorr-Cetina, & E. von Savigny (Eds.), *The practice turn in contemporary theory* (pp. 175–188). London: Routledge.

Koschmann, T. (2002). Dewey's contribution to the foundations of CSCL research. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of the Computer-supported Collaborative Learning Conference* (pp. 17–22). Mahway: Lawrence Erlbaum Associates.

Larusson, J., & Alterman, R. (2009). Wikis to support the "collaborative" part of collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 4*, 371–402.

Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. New York: Oxford University Press.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life (learning in doing)*. Cambridge: Cambridge University Press.

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

Levina, N., & Vaast, E. (2005). The emergence of boundary spanning competence in practice: Implications for implementation and use of information systems. *MIS Quarterly, 29*, 335–363.

Leydesdorff, L. (2007). "Betweenness centrality" as an indicator of the "interdisciplinarity" of scientific journals. *Journal of the American Society for Information Science and Technology, 58*, 1303–1309.

Lim, W.-Y., So, H.-J., & Tan, S.-C. (2010). eLearning 2.0 and new literacies: Are social practices lagging behind? *Interactive Learning Environments, 18*, 203–218. Q8

Lipponen, L. (2002). Exploring foundations for computer-supported collaborative learning. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of the Computer-supported Collaborative Learning Conference* (pp. 72–81). Mahway: Lawrence Erlbaum Associates.

Moskaliuk, J., & Kimmerle, J. (2009). Using wikis for organizational learning: Functional and psycho-social principles. *Development and Learning in Organizations, 23*, 21–24.

Moskaliuk, J., Kimmerle, J., & Cress, U. (2009). Wiki-supported learning and knowledge building: Effects of incongruity between knowledge and information. *Journal of Computer Assisted Learning, 25*, 549–561.

Moskaliuk, J., Kimmerle, J., & Cress, U. (2012). Collaborative knowledge building with wikis: The impact of redundancy and polarity. *Computers & Education, 58*, 1049–1057.

Newman, M. (2010). *Networks: An introduction*. Oxford: Oxford University Press.

Okoli, C., Mehdi, M., Mesgari, M., Nielsen, F.Å., & Lanamäki, A. (2012). The people's encyclopedia under the gaze of the sages: A systematic review of scholarly research on Wikipedia. Working paper. [http://ssrn.com/paper=2021326].

Oshima, J., Oshima, R., & Knowledge Forum ® Japan Research Group. (2007). Complex network theory approach to the assessment on collective knowledge advancement through scientific discourse in CSCL. In C. A. Clark, G. Erkens, & S. Puntambekar (Eds.), *Proceedings of the International Conference of Computer-Supported Collaborative Learning* (pp. 563–565). Mahwah: Erlbaum.

Paavola, S., & Hakkarainen, K. (2009). From meaning making to joint construction of knowledge practices and artefacts – A trialogical approach to CSCL. In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Computer supported collaborative learning practices: CSCL2009 conference proceedings* (pp. 83–92). Rhodes: International Society of the Learning Sciences.

Paavola, S., Lipponen, L., & Hakkarainen, K. (2002). Epistemological foundations for CSCL: A comparison of three models of innovative knowledge communities. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of the Computer-Supported Collaborative Learning Conference* (pp. 24–32). Mahway: Lawrence Erlbaum Associates.

Paavola, S., Lipponen, L., & Hakkarainen, K. (2004). Models of innovative knowledge communities and three metaphors of learning. *Review of Educational Research, 74*, 557–576.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). The PageRank citation ranking: Bringing order to the eb. Technical report. Stanford University. [http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf].

Palonen, T., & Hakkarainen, K. (2000). Patterns of interaction in computer-supported learning: A social network analysis. In B. Fishman & S. O'Conner-Divelbiss (Eds.), *Proceedings of the Fourth International Conference on the Learning Sciences*. Mahwah: Lawrence Erlbaum.

Pentzold, C., & Seidenglanz, S. (2006). Foucault@Wiki: First steps towards a conceptual framework for the analysis of wiki discourses. In *Proceedings of the 2006 International Symposium on Wikis* (pp. 59–68). New York: ACM Press.

Pifarré, M., & Kleine Staarman, J. (2011). Wiki-supported collaborative learning in primary education: How a dialogic space is created for thinking together. *International Journal of Computer-Supported Collaborative Learning, 6*, 187–205.

Reffay, C., & Chanier, T. (2002) Social network analysis used for modeling collaboration in distance learning groups, In S. A. Cerri, G. Gouardères, & F. Paraguaçu (Eds.), *ITS 2002 Lecture Notes in Computer Science, 2363*, 31–40.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. E. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69–197). Berlin: Springer.

Ryberg, T., & Larsen, M. C. (2008). Networked identities: Understanding relationships between strong and weak ties in networked environments. *Journal of Computer-Assisted Learning, 24*, 103–115.

Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences, 3*, 265–283.

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning, 5*, 43–102.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*, 4–13.

Sha, L., van Aalst, J., & Teplovs, C. (2010). A visualization of group cognition: Semantic network analysis of a CSCL community. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the disciplines: Proceedings of the 9th International Conference of the Learning Sciences* (Volume 1, pp. 929–936). International Society of the Learning Sciences: Chicago, IL.

Siemens, G. (2012). *Learning analytics: Envisioning a research discipline and a domain of practice. Proceedings of the Second International Conference on Learning Analytics and Knowledge LAK'12* (pp. 4–8). New York: ACM Press.

Sosa, M. E. (2011). Where do creative interactions come from? The role of tie content and social networks. *Organization Science, 22*, 1–21.

Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge: MIT Press.

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.

Susi, T., & Ziemke, T. (2001). Social cognition, artifacts, and stigmergy: A comparative analysis of theoretical frameworks for the understanding of artifact-mediated collaborative activity. *Cognitive Systems Research, 2*, 273–290.

Suthers, D. (2006). Technology affordances for intersubjective meaning-making: A research agenda for CSCL. *International Journal of Computers Supported Collaborative Learning, 1*, 315–337.

Swarts, J. (2009). *The collaborative construction of "fact" on Wikipedia. Proceedings of the 27th ACM International Conference on Design of Communication* (pp. 281–288). New York: ACM Press.

Tapscott, D., & Williams, A. D. (2006). *Wikinomics: How mass collaboration changes everything*. New York: Portfolio.

Tushman, M. L., & Scanlan, T. J. (1981). Boundary spanning individuals: Their role in information transfer and their antecedents. *The Academy of Management Journal, 24*, 289–305.

Voß, J. (2005). Measuring Wikipedia. In P. Ingwersen & B. Larsen (Eds.), *Proceedings of 10th International Conference of the International Society for Scientometrics and Informetrics* (pp. 221–231). Stockholm: Karolinska University Press.

Wassermann, S., & Faust, K. (1994). *Social network analysis: Methods and application*. New York: Cambridge University Press.