

What Can Be Learned About Computer-Supported Collaborative Learning From a Bibliographic Coupling Analysis?

Heisawn Jeong, Hallym University, heis@hallym.ac.kr
Sujong Seo, Hallym University, westwaterbell@gmail.com
Ji Yeon Jeong, Hallym University, qkdua45@gmail.com
Cindy E. Hmelo-Silver, Indiana University, chmelosi@indiana.edu
Sebastian Grauwin, University of Lyon, sebgrauwin@gmail.com

Abstract: Bibliographic coupling (BC) of 869 CSCL papers published between 2005 and 2014 showed a map of CSCL research organized around shared references. The CSCL map consists of ten BC research clusters with varying size and topics. A subset of CSCL were not well-connected, suggesting a degree of fragmentation in CSCL research. Bibliometric analysis also revealed hidden or implicit research practices.

Introduction

With advances in technology, there has been much progress in the field of Computer-Supported Collaborative Learning (CSCL). In spite of its success in producing positive learning gains (Chen, Wang, Kirschner, & Tsai, 2018), there is awareness that CSCL is at an inflection point (Wise & Schwarz, 2017). A number of recent studies have examined CSCL research practices and showed that, while sharing the goal of supporting collaborative learning with technology, CSCL research is practiced and studied in a number of different ways and differing views exist as to the critical aspects of collaborative learning (Jeong, Hmelo-Silver, & Yu, 2014). Tang, Tsai, and Lin (2014) examined pairs of CSCL publications cited together and, by additionally applying exploratory factor analysis and social network analysis, identified six major research themes and 15 core publications. These field-wide reflections on CSCL research helped us understand the major research topics along with major theoretical and methodological frameworks.

The goal of this paper is to understand the current landscape of CSCL research between 2005 and 2014 using a bibliographic coupling analysis. Bibliographic coupling (BC), like co-citation analysis, use bibliometric information about publications (e.g., authors, references), but differs in that it uses shared citations between publications and identify clusters of research that share references. It can help us understand how CSCL research is grouped around major reference bases and what are the nature and characteristics of these references. The features of the clusters in terms of keywords, journal outlets, institutional backgrounds of the authors, for example, can also help us understand CSCL research practices around CSCL publications.

Method

A CSCL corpus constructed in earlier examinations of CSCL research practices and outcomes was used (Jeong et al., 2014; Hmelo-Silver et al., under review). The corpus was constructed by journal-based searches as well as keyword-based searches on the Web of Sciences and ERIC databases. It contained a total of 869 papers published between 2005 and 2014. Metadata extraction includes information about the authors, title of the publications, keywords, publication source, institutions, country affiliated with authors, references, for example.

Bibliographic Coupling (BC) links were formed between publications when they share references (Kessler, 1963). A community detection algorithm based on modularity optimization (an implementation of the Louvain algorithm) was then applied to partition networks of linked papers into clusters. These clusters are represented in a map, in which a node represents a cluster with its size being proportional to the number of papers within the clusters. A frequency analysis was carried out to the papers within each cluster and provided results for (1) top 20 author keywords (2) top 10 publication sources, (6) top 10 countries, (7) top 10 references, (8) top 10 references sources, and (9) top 10 representative papers of the clusters. Representative papers here refer to papers that are most closely aligned with the papers within the cluster. In technical terms, they refer to the papers with the highest in-degree, where the in-degree of a paper is defined as the number of papers with the topic it is connected to by the shared references. Note that even when items such as author keywords or references appear in a given cluster with high frequency, the high frequency appearance may not be unique to that cluster, that is, the same keyword might appear frequently in other clusters too. A sigma value was used to identify whether the information is significant or significant, but note that it only indicates whether the item was more or less unique to the clusters and should not be likened to a statistical test.

Results

Bibliometric characteristics of the corpus and BC map

The most cited reference of the whole corpus was Vygotsky (1978) followed by Dillenbourg (1999) and Kirschner, Jochems, Dillenbourg, and Kansellaar (2002). Vygotsky was cited in 124 times (14%) in the corpus, twice as much as Dillenbourg (1999). The top three keywords used to describe the CSCL research were collaborative/cooperative learning, computer-mediated communication (CMC), and interactive environments. These keywords appear in many CSCL clusters, indicating the homogeneous nature of the corpus. In terms of publication sources, they were mostly published in *Computers and Education*, *Computers in Human Behavior*, and *Journal of Computer-Assisted Learning*. The *International Journal of CSCL (ijCSCL)* and *Journal of the Learning Sciences (JLS)* ranked fourth and fifth. The largest proportion of CSCL research (38%) were published in *Computers and Education*. It is by no means a CSCL exclusive journal, but publishes a large numbers of papers each (e.g., 12 issues per year, a contrast to 4 issues by *ijCSCL*). The CSCL research base is international with the United States (25%) followed by Taiwan (13%) and the UK (8%) forming the top three countries in which the authors were based.

The BC map of CSCL research is presented in Figure 1. The map consists of 10 clusters of varying size. Cluster labels were automatically selected from the most significant keywords used by the papers in the clusters. Note that not all papers in the corpus were included in the BC map. Out of the 869 papers in the corpus, a subset of the papers (n=122) did not share any references with other papers in the corpus and were not included in the cluster map. There were also a set of small clusters, consisting of two or three papers, that were unconnected to the rest of the clusters. Presence of such unconnected papers and clusters can be due to the inconsistencies in the data because of slight variations and inconsistencies in reference formatting (e.g., presence or absence of middle names, subtitles, etc.) that can make the same reference be treated differently by the BC algorithm.

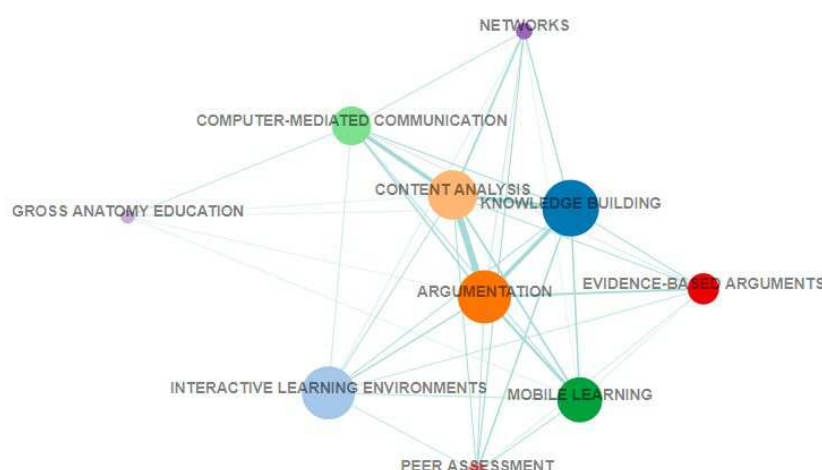


Figure 1. BC Map of CSCL research clusters.

Major CSCL research clusters

The major clusters refer to the five biggest clusters. They represent major areas of CSCL research in which many papers were published. They were: *knowledge building* (145), *argumentation* (127), *interactive learning environment* (127), *content analysis* (109), and *mobile learning* (90). Most of the research topics are concerned with issues related to developing CSCL environments but with different foci. While the *interactive learning environments* cluster is about generic interactive environment without strong association with specific technology or tools, the *knowledge building* is closely tied with Knowledge Forum as well as accompanying learning theories and pedagogy. The *argumentation* cluster reflects one of the major pedagogical foci in CSCL. The presence of a separate *mobile learning* cluster suggests a strong interest in these technologies by themselves. Although CSCL can occur both online and offline, much of the online collaboration is mediated by computers. The *content analysis* cluster reflects interests in method needed to analyze data often produced in computer-mediated discourses.

In addition to research topics or focus, these clusters also differ in the kinds of references they cited. A few references such as Vygotsky (1978) appear in multiple clusters, but highly cited references differ across clusters. This happened even when the clusters appear to study similar topics. For example, both the *knowledge building* and *interactive learning environments* clusters are interested in building collaborative environments. In the *knowledge building* cluster, theoretical papers about knowledge building communities and methodological papers about design experiments were one of the among the most cited references. On the other hand, in the

interactive learning environments cluster, publications on self-efficacy and statistical power analysis were highly cited (see Table 1). These differences in references show that the same issue of building a collaborative environment are designed and researched from quite different theoretical and methodological perspectives.

Clusters also differ in terms of the contexts in which the research was conducted, that is, the country in which authors are based in. As noted in the previous section, US-based researchers authored the majority of the papers throughout the corpus, but European-based researchers were more visible in the *argumentation* cluster producing 74% of the papers in this cluster. Researchers in Asia, especially from Taiwan, were prominent in the *interactive learning environment* and *mobile learning* clusters. Cultural emphasis on different styles of discourse and/or technology might have played a role, but more exploration is needed to better understand the causes of these differences.

Table 1: Top three references (authors and titles) of the major CSCL clusters (number of publications)

Clusters	Top Three References
Knowledge building (145)	1. Vygotsky (1980). Mind in society: The development of higher psychological processes. 2. Brown (1992). Design experiments. 3. Scardamalia, & Bereiter (1994). Computer support for knowledge-building communities.
Argumentation (127)	1. Webb (1989). Peer interaction and learning in small groups. 2. Cohen (1994). Restructuring the classroom: Conditions for productive small groups. 3. Clark & Brennan (1991). Grounding in communication.
Interactive learning environment (127)	1. Bandura & Wessels (1997). Self-efficacy 2. Cohen (1988). Statistical power analysis for the behavioural sciences. 3. Vygotsky (1980). Mind in society: The development of higher psychological processes.
Content analysis (109)	1. Hara et al. (2000). Content analysis of online discussion in an applied ed psy course. 2. Rourke et al. (2001). Methodological issues in the content analysis of computer conference transcripts. 3. Gunawardena et al. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing.
Mobile learning (90)	1. Zurita, & Nussbaum (2004). CSCL using wirelessly interconnected handheld computers. 2. Dillenbourg (1999). Collaborative learning: Cognitive and computational approaches. 3. Vygotsky (1980). Mind in society: The development of higher psychological processes.

Note. Publications in the table only appear in the table (not in the reference section) due to space limitations.

Minor CSCL research clusters

Minor clusters refer to the five smallest clusters, which were: *Computer-mediated communication* (n=66), *evidence-based argumentation* (n=41), *peer assessment* (n=13), *networks* (n=10), and *gross anatomy education* (n=7). *Computer-mediated communication* address various issues arise in computer-mediated communication situations (e.g., social presence, dialogue patterns). The *evidence-based argumentation* cluster research online discourse and computer-mediated learning, but with emphasis on structured pedagogical activities including argumentation, PBL, and/or inquiry. The rest of the minor clusters address feedback provided by peers (*peer assessment* cluster), networked learning and social network analysis (*networks* cluster), and education in a particular discipline (*gross anatomy education* cluster).

These clusters were again differentiated by the references they cited. This difference is noticeable when clusters with overlapping research topics were compared. As described earlier, *computer-based communication* (CMC) and *evidence-based argumentation* both share as emphasis in online discourse, but differed in their top references. In the CMC cluster, social and communication theories ranked high, whereas in the *evidence-based argumentation* cluster, references on pedagogies (e.g., PBL or web-based inquiry) were highly cited. We can also compare this cluster to the *argumentation* cluster, one of the major clusters, in which reviews and conceptual papers played a bigger role in grounding the research (see Table 1). These differences represent different traditions of argumentation research in CSCL and indicates how research examining the same topic such as argumentation can be built on quite different intellectual traditions.

Like major clusters, minor clusters varied in the country in which authors are based. While US based works were the majority, European researchers were again more visible in the *evidence-based argumentation* and *networks* cluster. Asian researchers were highly represented in the *peer assessment* cluster.

Summary and conclusion

In this paper, we examined CSCL research based on the references they share and presented a BC map of CSCL research. The map showed that CSCL research is organized into ten clusters of research. They cover different topics, but the majority of the papers addresses topics related to knowledge building, interactive learning environments, and argumentation. There were also clusters that represent a small but distinct research topics such as peer assessment or gross anatomy education. BC analysis helped us to reveal these clusters and organizations that did not receive much attention. Both minor and major clusters are part of CSCL research. We need to be more mindful of integrating findings and lessons from all the clusters when we interpret and assess CSCL research.

Some of the clusters, although addressing similar topics, formed distinct clusters because they cited different references. Cited references show the intellectual backgrounds of the research. Many of them are related to the theoretical and methodological approaches adopted in the research. The diversity in theoretical backgrounds and methodology provide productive tensions from which new knowledge can emerge. At the same time, there is a danger that they may remain as separate body of knowledge bases. They do not need to be merged into one, but we need to make sure there are enough connections and awareness of each other within the field to ensure the coherence of CSCL as a research field.

The findings of the paper also showed that certain topics were published more in some journals over others and that certain topics were more researched by researchers from certain countries and regions, hinting that some of the clustering may not be purely intellectual. Researchers may value certain research questions, theoretical perspectives and/or research methods over others depending on their backgrounds and training. Journals are also likely to establish its identity by the editors, editorial boards, and reviewers. These factors may have been implicit in the past, but BC analysis revealed the influence of some of these factors. It will take time to explore these further and interpret properly, but it a new analytical tool that we can use to understand how the knowledge outcomes produced in research may interact with the process that produce the knowledge.

References

- Chen, J., Wang, M., Kirschner, P. A., & Tsai, C. C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research*, 88(6), 799-843.
- Dillenbourg, P. (1999). *Collaborative learning: Cognitive and computational approaches. advances in learning and instruction series*. New York, NY: Elsevier Science.
- Hmelo-Silver, C. E., Jeong, H., McKeown, J., Hartley, K., Faulkner, R. T. & Andrade, A. (under review). Synergies among the pillars: A meta-synthesis of Computer-Supported Collaborative Learning in STEM domains.
- Kessler, M. M. (1963). Bibliographic coupling between scientific articles. *American Documentation*, 24, 123–131.
- Kirschner, P. A., Jochems, W., Dillenbourg, P., & Kanselaar, G. (2002). *Three worlds of CSCL: Can we support CSCL?* Heerlen: Open University of the Netherlands.
- Jeong, H., Hmelo-Silver, C. E., & Yu (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.
- Tang, K. Y., Tsai, C. C., & Lin, T. C. (2014). Contemporary intellectual structure of CSCL research (2006–2013): A co-citation network analysis with an education focus. *International Journal of Computer-Supported Collaborative Learning*, 9(3), 335-36
- Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, 12(4), 423-467.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*. Cambridge, MA: Harvard University.

Acknowledgments

This work is supported by the National Research Foundation of Korea (Grant# 2016R1D1A1B03935697) and US National Science Foundation (US National Science Foundation (Grant # 1249492).