Educational Research Review 34 (2021) 100403

Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev

Review

A systematic review of visual representations for analyzing collaborative discourse

Liru Hu *, Gaowei Chen

Faculty of Education, The University of Hong Kong, 999077, Hong Kong

ARTICLE INFO

Keywords:
Visual representation
Collaborative discourse
Visual analytics
Systematic review

ABSTRACT

Visual analytics combines automated data analysis and human intelligence through visualisation techniques to address the complexity of current real-world problems. This review uses the lens of visual analytics to examine four dimensions of visual representations for analysing collaborative discourse: goals, data sources, visualisation designs, and analytical techniques based on 89 studies. We found visual analysis approaches to be suitable and advantageous for decomposing the temporality of collaborative discourse. However, it has been challenging for current research to simultaneously consider learning theories and follow visualisation design principles when adopting visualisations to analyse collaborative discourse. At the same time, existing visual analysis approaches have mainly targeted learners or researchers in online contexts and mainly focused on *mirroring* collaborative discourse rather than providing advanced affordances such as *alerting* or *advising*. Informed by these findings, we propose a possible future research agenda and offer suggestions for the features of successful collaboration to guide the design of advanced affordances.

1. Introduction

1.1. The complexity of analysing collaborative discourse

Collaboration does not always outperform individual learning (Barron, 2003; Chi & Menekse, 2015). Previous meta-analytic research has revealed a low to medium effect size for an individual's learning in collaborative groups compared with a non-group condition (Cohen's $d = 0.17$ to 0.66; Chen et al., 2018; Johnson & Johnson, 1992; Lou et al., 1996). Researchers have also found that the effectiveness of collaboration is often due to the quality of collaborative discourse (Gillies, 2019). Analysing collaborative discourse could therefore be a meaningful way to understand the complexity of collaborative processes and outcomes. In this review, *collaborative discourse* refers to verbal or written communication in a situation in which a group of individuals with equivalent rights attempt to solve a problem or learn something together by sharing their understanding and negotiating ideas (Dillenbourg, 1999; Hesse et al., 2015).

However, the dynamic and historical features of collaborative discourse (Mercer, 2008; Wise & Chiu, 2011) make it challenging to uncover the complexity of collaborative processes. Human interactions are historically located in specific institutional and cultural contexts and can invoke memories of interactive experiences or shared knowledge. Human interactions also emerge dynamically,

* Corresponding author.
E-mail address: liruhu@connect.hku.hk (L. Hu).

https://doi.org/10.1016/j.edurev.2021.100403
Received 1 March 2020; Received in revised form 22 June 2021; Accepted 24 July 2021
Available online 27 July 2021
1747-938X/

rather than in a planned manner. Interaction must be temporarily and contextually understood (Hmelo-Silver, 2003). This has led to methodological criticism of the pure coding-and-counting approach in analysing collaborative discourse (Csanadi et al., 2018; Swiecki et al., 2020). Tensions have also arisen over the use of existing theories and frameworks to guide the design of computer support for collaboration and harmonise interpretative and computational methodologies (Wise & Schwarz, 2017).

From a practical perspective, productive peer talk seldom happens spontaneously in classrooms. Therefore, students need explicit guidance on how to use language effectively and regulate group interactions (Belland et al., 2013; Cohen & Lotan, 2014; King, 2008; Miller & Hadwin, 2015; Scheuer et al., 2010). There is also the need to support teachers in monitoring multiple groups and adopting dialogic strategies (Kazemitabar et al., 2016; van de Pol et al., 2019; van Leeuwen et al., 2015). Therefore, understanding the nature of collaboration is a challenge for researchers, whereas monitoring the collaborative process and making adjustments as needed is a challenge for students and teachers.

1.2. Visual analytics

The complexity of collaborative discourse can seldom be understood solely via statistical or data mining techniques, a fact that emphasises the indispensability of human interpretation and judgement. Visual analytics has shown great potential in meeting this challenge, through a combination of automated data analysis performed by a computer and interactive visual reasoning performed by a human.

Thomas and Cook (2005, p. 4) defined visual analytics as 'the science of analytical reasoning facilitated by interactive visual interfaces'. This definition assumes that the complexity of current real-world problems makes it impossible to achieve the best solution through purely automated data analysis. Therefore, visual analytics makes use of automated analysis while also involving humans for data interpretation and decision making through interactive graphical interfaces. The original visual analytics process suggested by Keim et al. (2009) featured interactions between visual and automatic methods that created opportunities to gain knowledge from heterogeneous data sources (see Fig. 1). People can apply either visual or automatic analysis to obtain the desired knowledge, but it is more likely that they will need to adopt both visual data exploration and information mining to continuously refine and verify preliminary results. Visualisations in visual analytics not only serve to efficiently and effectively communicate research results but also facilitate confirmatory and exploratory data analysis (Keim et al., 2008).

Many studies have adopted visual analytics approaches in collaborative discourse analysis. Increasing efforts have been made to develop visual analytics tools to help researchers uncover the temporal patterns of collaborative discourse (e.g., Lämsä et al., 2018; Shaffer & Ruis, 2017) and to support group work and teacher guidance (e.g., Resendes et al., 2015; van Leeuwen et al., 2014; Zhang et al., 2018). Some scholars also generally term the application of visual analytics in education *visual learning analytics*, which is an emerging research area (Chen, 2019, 2020; Chen, Chan, et al., 2020; Vieira et al., 2018).

1.3. Research structure and questions

There have been several review articles on visual analysis tools in education, such as on student-facing learning analytics dashboards and educational recommender systems by Bodily and Verbert (2017), on visual learning analytics tools that act as solutions or interventions in education by Vieira et al. (2018), and on team performance visualisation tools by Swiecki and Shaffer (2018). However, to the best of our knowledge, no systematic review has specifically focused on Visual Representations for analysing Collaborative Discourse (VRCD). Therefore, we attempted to systematically examine visual analysis approaches to studying collaborative discourse and to identify possible future research goals for this area.

In this study, we adopted a visual analytics lens to locate studies that adopt visual representations as an analytical approach, rather

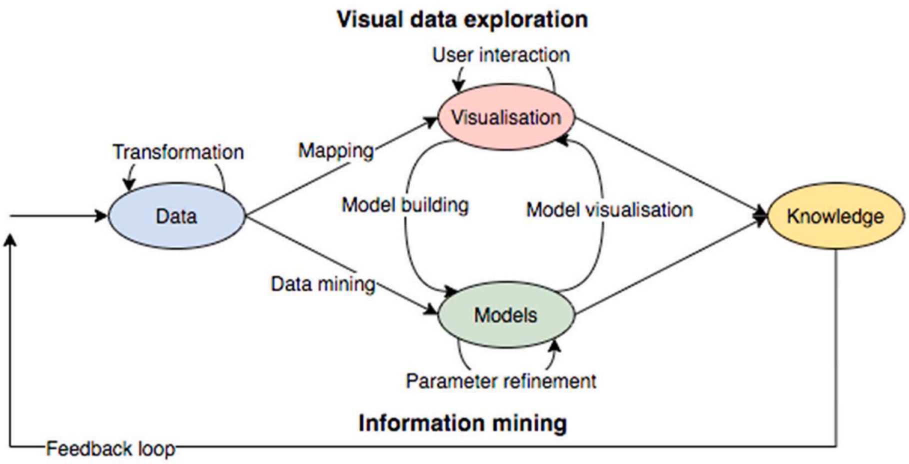

Fig. 1. The visual analytics process proposed by Keim et al. (2009).

than simply as a communication medium. Specifically, the process model of visual analytics proposed by Keim et al. (2009) (see Fig. 1) helped to structure the present review. Although Keim's model has been extended in multiple studies (e.g., Andrienko et al., 2018; Ribarsky & Fisher, 2016; Sacha et al., 2014), data sources, visualisation design, analytic techniques, and knowledge remain the core elements of visual analytics. Consequently, these four elements form the basic structure of our review. We also made the adaptation of changing *knowledge* to *goal* to better incorporate the goal-relevant considerations of visual analysis approaches in a learning context, such as target users, target problems, and theoretical background (e.g., Hillaire et al., 2016; Vieira et al., 2018). Theoretical considerations largely inform and justify the knowledge that users desire to gain from visual analytics approaches. We therefore examined whether VRCD were guided by relevant learning theories to address target problems and support target users.

In brief, we subdivided the review into four major dimensions through a lens of visual analytics: goals, data sources, visualisation designs, and analytic techniques. There were four research questions, accordingly.

1) What are the goals of VRCD in the literature?
2) What are the data sources for these VRCD?
3) What are the design attributes of these VRCD, such as display formats, dynamism, interactivity, and quality?
4) What are the analytic techniques underlying these VRCD?

2. Method

2.1. Strategy for the literature search

We followed the procedures for the Preferred Reporting of Items for Systematic Reviews and Meta-Analyses to locate relevant literature in English (Moher et al., 2009). In our initial searches, we mainly used the following searching expressions and their variants to restrict titles, subjects, or abstracts: (talk OR discussion OR conversation OR dialogue* OR argu* OR (collaborat* AND discourse)) AND visual* NOT 'visually impaired' NOT 'visual dialogue*'. The target databases were ProQuest, Web of Science, the IEEE Xplore Digital Library, and the ACM DL Digital Library. This literature search also covered relevant and high-quality conference proceedings from the most recent five-year period (2014–2018), namely, the International Learning Analytics and Knowledge Conference, the International Conference on Educational Data Mining, the International Conference of the Learning Sciences, the International Conference on Computer-Supported Collaborative Learning (CSCL), the IEEE Conference on Information Visualisation, and the IEEE Conference on Visual Analytics Science and Technology. Finally, a snowballing strategy was adopted to locate other studies that were relevant but might not include the focus keywords. No start time limit was placed on searching, and the search process was carried out between August and October 2019.

2.2. Inclusion/exclusion criteria

In this review, we aimed to synthesise all possible visual representations that have been created to analyse collaborative discourse as either an intervention or research approach. As discussed in the introduction, visualisations in visual analytics should assist either confirmatory or exploratory analysis, rather than simply illustrate research results. Visual analytics should also involve an automatic analysis component. Therefore, to be included in the final review, a study had to meet six inclusion and exclusion criteria according to the lens of visual analytics.

1. A study needed to include at least one figure depicting the visualisation design.
2. The data source of visualisations in a study needed to incorporate collaborative discourse.
3. The visualisations in a study needed to involve at least one automatic analysis technique such as descriptive statistics, inferential statistics, and text analytics.
4. A study was excluded if it incorporated visualisations simply to communicate the research results.
5. A study was excluded if it focused on interactions among individuals with non-equivalent rights, such as teacher–student interactions, parent–child interactions, or human–machine interactions.
6. Posters or two-page articles in conference proceedings were omitted, due to their limited scope.

When VRCD were replicated in multiple studies, or multiple distinct VRCD were present in one study, the following rules were used to choose appropriate analysis units.

1. When multiple articles described/applied the same VRCD (e.g., one author submitted multiple types of work describing the design of one tool, or multiple articles adopted the same VRCD), only the most detailed and/or recent article was retained as the representative publication.
2. When one article contained multiple VRCD, it was split into multiple analysis units only if each of these VRCD was a complete and independent tool (e.g., if an article introduced two distinct VRCD tools, it was split into two analysis units).

2.3. Data coding and analysis

All of the included VRCD were coded using the framework presented in Table 1. Basic information on the article itself was captured, such as the year of publication, type of publication, publication source, institute of the first author, and community to which the article belonged. This allowed us to conduct an overview of features of existing scholarship on this topic. Other major codes corresponding to the four research questions were *goal*, *data source*, *visualisation design*, and *analytic technique*.

2.3.1. Goal

We considered target user, target underlying learning process, timeliness, affordance and learning theory for the goal of VRCD. As to target user, we first categorised it according to the review by Swiecki and Shaffer (2018) and then refined the codes based on our sample. The goal of VRCD has been to serve either *researchers*, enabling them to answer research problems and communicate research results, or *teams* or *educators*, facilitating effective collaboration.

In the present review, we mainly examined the support by existing VRCD for four dimensions of an implicit learning process: *cognitive*, *social*, *socio-cognitive*, and *emotional*. Collaborative discourse targets multiple dimensions of the underlying learning process (Ludvigsen, 2016). Previous studies have mainly focused on the *cognitive* process that reveals the knowledge construction of group members (e.g., Schnaubert & Bodemer, 2019) and the *social* process that depicts the functioning of the group, including participation, perspective-taking, and social regulations (e.g., Shah & Lewis, 2019). Certain studies have focused on the *socio-cognitive* dimension, which concerns group cognition, mutual understanding, and socially shared task regulation (e.g. Zhang et al., 2007). The *emotional* dimension of a collaborative process, which focuses on the emotions of group members and whether members can monitor emotions and provide emotional support to one another, has also attracted increasing attention in existing research (e.g., Näykki et al., 2014).

The goal of VRCD was further categorised from a temporal perspective as *real-time* or *post-hoc*, based on the target timeliness. The specific affordances of existing visual approaches were categorised into three types—*mirroring* (graphically representing interaction), *alerting* (cueing desired states or important events to pay attention to), and *advising* (suggesting strategies or remedial actions)—according to how they enable different underlying regulative processes (Rodríguez-Triana et al., 2017). This was informed by previous relevant studies of the categorisation of CSCL tools (Jermann & Dillenbourg, 2008; Rummel, 2018; van Leeuwen et al., 2019). For example, Jermann and Dillenbourg (2008) differentiated three types of regulation support: mirroring tools (which assist data collection by providing graphical feedback), metacognitive tools (which assist diagnoses of interaction by providing standards of productive interactions), and guiding tools (which provide remedial actions based on a computational assessment).

The development of learning tools must be guided by specific learning theories (Hillaire et al., 2016; Shaffer & Ruis, 2017; Wise & Schaffer, 2015). Therefore, we further examined whether the goal of VRCD involved any theoretical considerations.

2.3.2. Data source

The data sources of VRCD were first characterised according to basic contextual information, such as group size and communication medium. They were further categorised according to target discourse features. Human conversations are hierarchical, containing nested discourse units, and are composed of basic linguistic units such as words, tone, and pauses. A talk *move* has a functional relation to the conversation of which it is a part (Goffman, 1981). A *turn* refers to a time during which only one speaker holds the floor. Two or more adjacent and functionally related turns form a *sequence*. Informed by the hierarchy of human conversation, we adopted an

Table 1
Coding framework for included studies.

Code	Sub-code	Interpretation
Basic sample statistics	Year of publication	–
	Publication type	*Journal*, *proceeding*, *book chapter*, or *thesis*.
	Institute location	Location of the first author's institute
	Publication place	Name of the journal, conference proceeding, or book in which the study was published or university at which it was produced
	Community	*Learning* oriented or *visualisation* oriented, inferred from the institute location and publication place
Goal	Target user	Three dichotomous variables: *learner*, *educator*, and *researcher*
	Target underlying learning process	Four dichotomous variables: *cognitive*, *social*, *socio-cognitive*, and *emotional*
	Timeliness	Whether the VRCD was provided in *real-time* or *post-hoc* to facilitate the interaction process
	Affordance	Three dichotomous variables: *mirror*, *alert*, and *advise*
	Learning theory	Whether the article mentioned any learning theory background for the production of VRCD
Data source	Group size	*Dyadic*, *small* (usually 3–8 people), or *large*
	Communication medium	*Online*, *face-to-face*, and *audio*
	Target discourse feature	Seven dichotomous variables: *linguistics (pitch, tone, volume, speed)*, *move*, *turn*, *sequence*, *semantics*, *assessment*, and *other*
Visualisation design	Name	The name of VRCD, if provided by the author(s)
	Timeline	Whether the visualisation contained a time axis
	Dynamism	Two dichotomous variables: *static* and *dynamic*
	Interactivity	Whether the visualisation supported human interaction, such as the setting of parameters
	Display format	Fourteen dichotomous variables: *network*, *tree*, *bubble*, *bar*, *step*, *line*, *pie*, *radar*, *raw text*, *novel*, *spiral timeline*, *word cloud*, *heat map*, and *other*
	Design principle	Whether the article mentioned any visualisation design considerations
Analytic technique		The technique for any automated statistical process before data reporting.

open-coding strategy and categorised target discourse features by existing VRCD into the following units: *linguistics*, *move*, *turn-taking*, *sequence*, *semantics*, and *assessment*. *Semantics* refers to the semantic characteristics of collaborative discourse, and *assessment* refers to the post-hoc assessment on the quality of collaborative discourse.

2.3.3. Visualisation design

We considered visualisation categorisation, display formats and design quality when examining the design of VRCD. In this article, we categorised visualisation designs from three dimensions: timeline, dynamism and interactivity. Visualisations are temporal or cumulative, depending on whether they include a timeline. Temporal visualisations may illustrate how discourse unfolds over time, whereas cumulative visualisations may reveal aggregated characteristics or structures of collaborative discourse. Visualisations are *dynamic* or *static* depending on whether they include static elements, such as images, photos, diagrams, and graphs, or dynamic elements, such as animation, video, or simulation. They may also be interactive or not depending on whether human could manipulate design parameters.

The included VRCD involved various display formats to visualise collaborative discourse. The current review mainly employed a common online visualisation classification system (Visual Vocabulary, n.d.). Designs not in this visual vocabulary, except for *raw text*, were coded as *novel* designs. This review also examined whether the selected articles explicitly mentioned any visualisation design considerations since high-quality visualisation design is essential for visual analytics (Keim et al., 2008, 2009).

2.3.4. Analytic technique

Existing VRCD involve a wide range of automatic analytic techniques, including *descriptive statistics* and advanced analytic techniques such as *inferential statistics*, *text analytics*, *network analysis* and *process mining*. The microgenetic analysis of individual utterances is very complicated and laborious (Chiu & Khoo, 2005), and thus restraining the timeliness of feedback on talk quality. A lot of well-established tools and methods in the text analytics community such as topic detection approaches (Teplovs, 2015) have been contextualized to learning (Buckingham Shum & Crick, 2016) to solve these issues.

The technique of *network analysis* could be detailed into social network analysis, socio-semantic network analysis, and epistemic network analysis etc. Social network analysis involves the analysis of social structures by studying the links between nodes (Borgatti et al., 2009; Wasserman & Faust, 1994), and excels at rapidly detecting isolated individuals, influential leaders, or subgroups. Many quantitative indices, such as the index of centrality, exist for describing the connectivity of participants at different levels. Socio-semantic networks analyse the interaction of particular concepts according to their co-occurrences in a turn (Oshima et al., 2018). Unlike other network types, epistemic network analysis utilises a fixed coordination system and supports multiple-level statistical network comparisons (Shaffer & Ruis, 2017).

There are also many *process mining* techniques that could be used to support VRCD. For example, lag sequential analysis calculates

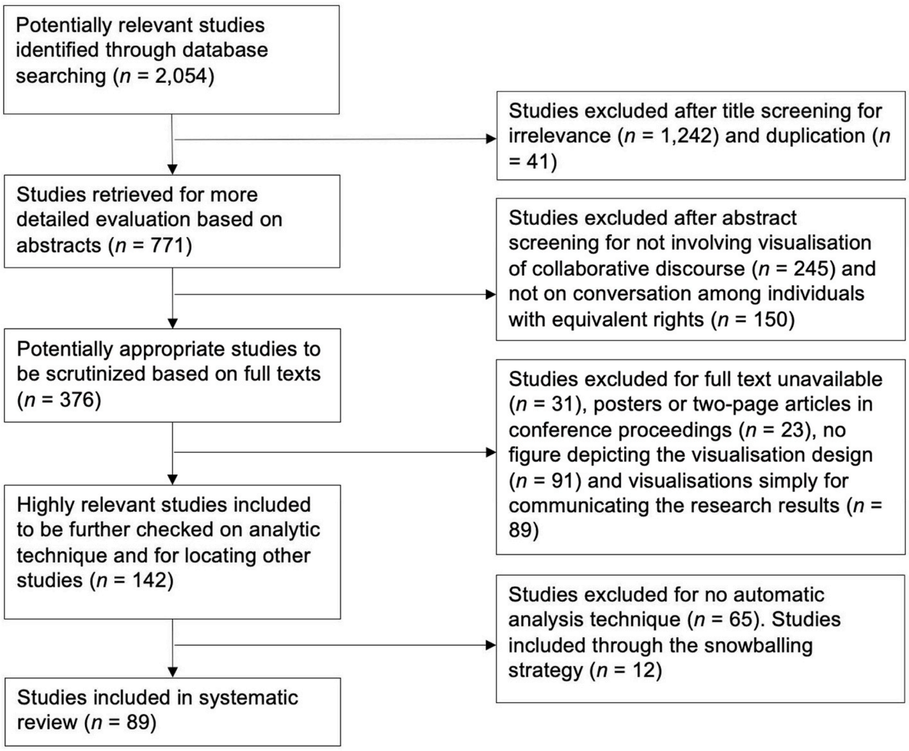

Fig. 2. Flow diagram of included and excluded studies.

Table 2
Basic sample statistics.

No.	Authors and year of publication	Publication type	Institute location	Community	Name
1	Adachi et al. (2014)	Proceeding	Japan	Visualisation	
2	Adeniran et al. (2019)	Proceeding	UK	Learning	
3	Adraoui et al. (2018)	Proceeding	Morocco	Visualisation	
4	Ahn et al. (2012)	Proceeding	USA	Other	TempoVis
5	Allaymoun (2015)	Proceeding	Romania	Learning	
6	Angus et al. (2012)	Journal Article	Australia	Visualisation	Conceptual Recurrence Plot
7	Atapattu et al. (2016)	Proceeding	Australia	Learning	Topic Visualisation Dashboard
8	Bachour et al. (2010)	Journal Article	Switzerland	Learning	Reflect
9	Boroujeni et al. (2017)	Proceeding	Switzerland	Learning	
10	Chen and Zhang (2016)	Journal Article	USA	Learning	Promising Ideas Tool
11	Chen and Zhang (2016)	Journal Article	USA	Learning	Epistemic Discourse Moves Tool
12	Chen (2015)	Proceeding	USA	Visualisation	
13	Chin and Chignell (2006)	Proceeding	Canada	Visualisation	
14	Chinn et al. (2000)	Journal Article	USA	Learning	
15	Critchlow (2006)	Thesis	USA	Visualisation	BuddySquares
16	Dave et al. (2004)	Proceeding	USA	Learning	Forum Reader
17	Deng et al. (2019)	Journal Article	China	Learning	Discussion Analysis Tree (DATree)
18	DiMicco and Bender (2007)	Proceeding	USA	Visualisation	Second Messenger
19	Donath et al. (1999)	Journal Article	USA	Visualisation	Chat Circles, Conversation Landscape
20	Donath et al. (1999)	Journal Article	USA	Visualisation	Loom
21	Dörk et al. (2010)	Journal Article	USA	Visualisation	Visual Back channel
22	Erickson and Kellogg (2003)	Book Chapter	USA	Visualisation	Babble
23	Fu et al. (2017)	Proceeding	China (HK)	Visualisation	iForum
24	Herring et al. (2005)	Proceeding	USA	Visualisation	
25	Hoch et al. (2015)	Journal Article	USA	Other	
26	Hoque and Carenini (2014)	Journal Article	Canada	Visualisation	ConVis
27	Imtiyazi et al. (2016)	Proceeding	Indonesia	Other	
28	Indratmo et al. (2008)	Proceeding	Canada	Visualisation	iBlogVis
29	Janssen et al. (2007)	Journal Article	Netherlands	Learning	Participation Tool
30	Jermann and Dillenbourg (2008)	Journal Article	Switzerland	Learning	
31	Jin (2017)	Journal Article	Korea	Learning	
32	Karahalios (2004)	Thesis	USA	Visualisation	Visiphone
33	Karahalios and Bergstrom (2009)	Journal Article	USA	Visualisation	Conversation Clock
34	Karahalios and Bergstrom (2009)	Journal Article	USA	Visualisation	Conversation Vote
35	Karahalios and Bergstrom (2009)	Journal Article	USA	Visualisation	Conversation Cluster
36	Kazemitabar et al. (2016)	Proceeding	Canada	Learning	Helping Others With Argumentation and Reasoning Dashboard (HOWARD)
37	Kim and Lee (2012)	Journal Article	Korea	Learning	Multidimensional Interaction Analysis Tool (MIAT)
38	Kwon et al. (2016)	Journal Article	Germany	Visualisation	Visualising Online Health Community (VisOHC)
39	Lagatie et al. (2011)	Proceeding	UK	Visualisation	
40	Lämsä et al. (2018)	Journal Article	Finland	Learning	
41	Lee and Tan (2017)	Proceeding	Singapore	Learning	
42	Lee et al. (2009)	Proceeding	Korea	Visualisation	Telescope for Responding comments for Internet Blogs (TRIB)
43	Li et al. (2013)	Proceeding	China	Visualisation	
44	Lund et al. (2017)	Journal Article	France	Learning	FRIEZE
45	Martinez-Maldonado et al. (2012)	Proceeding	Australia	Learning	
46	Mathur and Karahalios (2009)	Proceeding	USA	Visualisation	Voice Space
47	McCormick (2013)	Journal Article	USA	Learning	Social Network Adapting Pedagogical Practice (SNAPP)
48	Ming and Ming (2013)	Proceeding	USA	Learning	
49	Oshima et al. (2012)	Journal Article	Japan	Learning	Knowledge Building Discourse Explorer (KBDeX)
50	Oyama et al. (2014)	Proceeding	Japan	Learning	
51	Pascual-Cid and Kaltenbrunner (2009)	Proceeding	Spain	Visualisation	
52	Pupyrev and Tikhonov (2010)	Proceeding	Russia	Visualisation	
53	Resendes et al. (2015)	Journal Article	Canada	Learning	
54	Sack (2000)	Proceeding	USA	Visualisation	Conversation Map
55	Scardamalia (2004)	Book Chapter	Canada	Learning	Knowledge Forum
56	Schwarz and Asterhan (2011)	Journal Article	Palestine	Learning	Argunaut System
57	Sedrakyan et al. (2020)	Journal Article	Belgium	Learning	
58	Sha et al. (2010)	Proceeding	China (HK)	Learning	Knowledge Space Visualiser (KSV)
59	Shaffer and Ruis (2017)	Book Chapter	USA	Learning	Epistemic Network Analysis (ENA)
60	Shahid et al. (2017)	Journal Article	Pakistan	Visualisation	
61	Shankar et al. (2000)	Proceeding	USA	Visualisation	Fugue
62	Shapiro et al. (2017)	Journal Article	USA	Learning	Mondrian Transcription
63	Siegel et al. (2004)	Proceeding	USA	Visualisation	
64	Sirbu et al. (2019)	Proceeding	Romania	Learning	

(*continued on next page*)

Table 2 (*continued*)

No.	Authors and year of publication	Publication type	Institute location	Community	Name
65	Smith and Fiore (2001)	Proceeding	USA	Visualisation	Netscan Dashboard
66	Su and Boydell (2016)	Proceeding	Ireland	Visualisation	TopicListener
67	Sundaram et al. (2012)	Journal Article	USA	Visualisation	
68	Tat and Carpendale (2002)	Proceeding	Canada	Visualisation	Bubba Talk
69	Tat and Carpendale (2006)	Proceeding	Canada	Visualisation	CrystalChat
70	Teo et al. (2013)	Proceeding	USA	Learning	
71	Trausan-Matu et al. (2014)	Journal Article	Romania	Learning	Polyphonic Conversation Analysis and Feedback Generation (PolyCAFe)
72	Twitchell and Nunamaker (2004)	Proceeding	USA	Visualisation	
73	van Aalst et al. (2012)	Proceeding	China (HK)	Learning	Knowledge Connections Analyser (KCA)
74	van Leeuwen et al. (2019)	Journal Article	Netherlands	Learning	
75	van Leeuwen et al. (2014)	Journal Article	Netherlands	Learning	
76	van Leeuwen et al. (2015)	Journal Article	Netherlands	Learning	Concept Trail (CT), Progress Statistics (PS)
77	Viégas and Smith (2004)	Proceeding	USA	Visualisation	Newsgroup Crowds
78	Viégas and Smith (2004)	Proceeding	USA	Visualisation	Author Lines
79	Vivian et al. (2015)	Proceeding	Australia	Other	
80	Voyiatzaki and Avouris (2014)	Journal Article	Greece	Learning	Synergo Supervisor
81	Wise et al. (2017)	Proceeding	USA	Learning	
82	Xing et al. (2015)	Journal Article	USA	Learning	
83	Xiong et al. (2012)	Proceeding	China	Learning	
84	Xiong and Donath (1999)	Proceeding	USA	Visualisation	People Garden, People Flower
85	Xu et al. (2013)	Journal Article	China (HK)	Visualisation	
86	Zhang et al. (2018)	Journal Article	USA	Learning	Idea Thread Mapper (ITM)
87	Zhao et al. (2014)	Proceeding	Canada	Visualisation	Personal Emotion Analysis, Reasoning, and Learning (PEARL)
88	Zheng et al. (2018)	Journal Article	China	Learning	
89	Zumbach et al. (2004)	Book Chapter	Germany	Learning	Easy Discussing

Note. Details of this sample are available at: http://italkisee.com/data/Review/VRCD/SampleDetails.xlsx.

the conditional probabilities from one event to another, thereby detecting any significant conditional probability (Bakeman & Gottman, 1997). Frequent sequence mining has been used to uncover frequent sub-sequences that achieve a self-defined frequency and interval distance (e.g., Chen et al., 2017). Markov models are another set of approaches commonly used in temporal learning analytics. The simplest Markov model is a Markov chain, which depicts the transitional probabilities among a set of states. In addition, there are many complex Markov models that may better model collaborative discourse, such as the extended model Markov chain with memory that helps capture the local temporal context of one turn, and the hidden Markov model that facilitates the detection of hidden states underlying observed codes (Reimann, 2009). In addition, statistical discourse analysis is flexible in addressing various challenges in analysing collaborative discourse, such as adjusting for differences across individuals and groups in temporal analysis (e.g., Chen et al., 2012; Chen, Lo, & Hu, 2020; Wise & Chiu, 2011).

2.3.5. Data coding

Two trained coders independently performed all of the coding. Cohen's k was run to determine the degree of inter-coder agreement on each sub-code. The original average agreement between the two coders for the categorical variables on basic sample statistics was 0.93; that for variables on goal was 0.50; that for variables on data sources was 0.51; and that for variables on visualisation design was 0.57. The low average agreement for goal was mainly due to the *mirror* code that existed in almost all of the analysis units and therefore had an extremely low inter-coder agreement ($k = -0.043$). For data sources, the average inter-coder agreement was lowered by codes of *linguistics* ($k = 0.35$), *sequence* ($k = 0.37$), and *socio-cognitive* ($k = 0.22$), due to their small number of occurrences. *Tree* ($k = 0.39$), *step* ($k = 0.29$), and *static* ($k = 0.32$) were also imbalanced codes that decreased the average level of agreement for visualisation design. All disagreements were resolved through further discussion.

3. Results

3.1. Basic sample statistics

Initial search from targeted data sources returned 2054 results. There were 89 VRCD that met the selection criteria (see Fig. 2). Inter-rater agreement on the selection of studies achieved a satisfactory level ($k = 0.87$) (Landis & Koch, 1977). According to Table 2, approximately 65% of included VRCD were produced during or after 2010 and 36% during or after 2015, indicating that this research branch was relatively young. The United States produced the most studies (40%), and the remaining representations were derived from a range of countries, such as Canada (10%, $n = 9$), China (9.0%, $n = 8$), and Australia (4.5%, $n = 4$) and Netherlands (4.5%, $n = 4$).

The included articles were largely from conference proceedings (51%) and journals (43%) and derived from a variety of source publications. Most were from the Hawaii International Conference on System Sciences (HICSS) (10%, $n = 9$), the Proceedings of the International Conference on Computer-Supported Collaborative Learning (CSCL) (5.6%, $n = 5$), IEEE Transactions on Visualisation and

Computer Graphics (5.6%, $n = 5$), the *International Journal of CSCL* (4.5%, $n = 4$), *Computers and Education* (4.5%, $n = 4$), and *Educational Technology and Society* (4.5%, $n = 4$). Publications in the learning-oriented (49%) community were slightly more than those in the visualisation-oriented community (44%). The learning-oriented community mainly published in the International Conference of CSCL, *International Journal of CSCL*, *Computers and Education*, and *Educational Technology and Society*. The primary publication organs for the visualisation-oriented community were HICSS, IEEE Transactions on Visualisation and Computer Graphics, IEEE Computer Graphics and Applications, and the International Conference on Information Visualisation.

3.2. Goal

3.2.1. Target user

Our analysis showed that most VRCD aimed to serve practice by supporting learners (57%), either as a whole group or as individuals within a group, or by supporting educators (30%). There were also a substantial number of VRCD that targeted researchers (33%). This indicated the visualisation's dual role in analysing collaborative discourse; that is, visualisation is a powerful tool to aid the interpretation and improvement of collaborative processes, and can also function as an innovative research method in collaborative discourse analysis, thereby supplementing the original socio-cognitive or interpretive tradition (Dyke et al., 2012; Fu et al., 2016).

3.2.2. Target underlying learning process

The results showed that most of the existing VRCD involved the cognitive dimension of collaborative discourse (57%). In addition, many (54%) of the included VRCD aimed at social relationships among individuals. For example, the Argunaut system provides teachers with group awareness information, such as social networks, cumulative contribution types, and temporal individual contribution records, to facilitate their monitoring of synchronous group discussions (Schwarz & Asterhan, 2011). Those focusing on integrated socio-cognitive processes accounted for only approximately 24%. For example, Lagatie et al. (2011) compared contributed topics across speakers by creating a two-mode network that combined the speaker's identity and topics. The emotional aspect of collaborative discourse was rarely examined (11%). In one example, Chen (2015) developed a multi-level visual analytical system to support the exploration of online discussion forums. The study differentiated negative and positive posts to help understand a community's overall attitude towards a specific topic and identify relevant supporters and opponents.

3.2.3. Timeliness and affordance

This review showed that a majority of the included VRCD (62%) were developed to provide adaptive real-time feedback in the process of group discussion such as group awareness tools that provided students with dynamic information on individual cognition, emotion, or behaviour (e.g., Janssen & Bodemer, 2013; Jin, 2017; Schnaubert & Bodemer, 2019). There were only around 30% of included VRCD that were intended to facilitate post-hoc reflection by providing feedback on group discussion such as the Knowledge Connections Analyser (van Aalst et al., 2012) and Polyphonic Conversation Analysis and Feedback Generation (Trausan-Matu et al., 2014). In addition, we found 7.9% of included VRCD provided both real-time and post-hoc feedback such as Second Messenger (DiMicco & Bender, 2007) and Idea Thread Mapper (Zhang et al., 2018).

Our results also showed that all of the included visual approaches mirrored some features of collaborative discourse, especially after group discussion. Such mirrored features were primarily intended to make individuals aware of their performance in collaboration, thereby improving their future collaborative efforts, or to help them to monitor details of their collaborative process. In contrast, there were very few of included VRCD that aided interpretation by alerting the user to specific features (19%) or advising

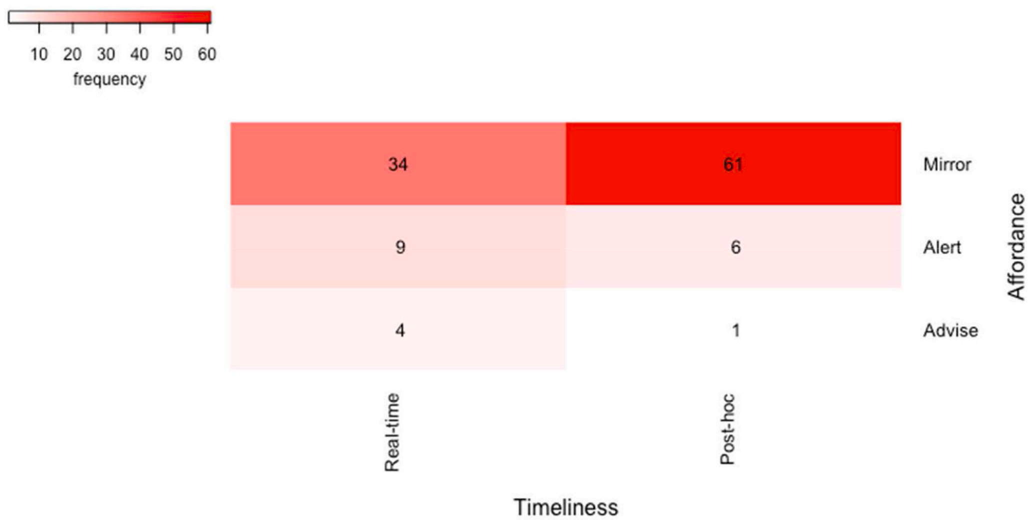

Fig. 3. A heatmap of timeliness and affordance.

desired actions (7%). Explicit affordances of alerting or advising were aimed at supporting an individual's recognition of their problematic status or further selection of proper strategies to improve their current collaboration status. We further found that included VRCD provide alerting and advising mainly during collaboration (65%) while mirroring mainly after collaboration (64%) (see Fig. 3).

3.2.4. Learning theory

We found that approximately 67% of the included VRCD mentioned theoretical considerations, such as knowledge building (e.g., Oshima et al., 2012; Zhang et al., 2018), argumentation (e.g., Chinn et al., 2000), and regulation (e.g., Janssen et al., 2007; Jermann & Dillenbourg, 2008). Further analysis revealed that publications in the learning-oriented community mentioned learning theories (93%) significantly more often than those in the visualisation-oriented community (41%), χ^2 (1, $n = 85$) = 26.19, $p < .001$.

3.3. Data source

This analysis revealed that around 74% of VRCD were contextualized in an online environment, whereas only 18% were tailored to face-to-face talk. There were also 3.4% of representations ($n = 4$) that focused on illustrating dyadic audio conversations. In terms of group size, approximately 47% of these VRCD focused on large groups, whereas 35% focused on small groups, and approximately 11% focused on dyadic groups. Furthermore, almost all of the large groups communicated online, as did approximately 65% of the small groups (see Fig. 4). Finally, only 35% of the small groups communicated face to face, and the dyadic groups primarily communicated via audio or face to face.

In terms of target discourse features, the included VRCD largely focused on turn-taking (71%) and semantics (47%). Turn-taking was usually adopted to reflect relational social space in group interactions (see Fig. 5). It was also frequently analysed to determine socio-cognitive structures. In contrast, semantics were typically investigated to explore individual or group cognition, especially the knowledge construction process. Talk moves were another frequently examined discourse feature (32%); they reflect an individual's cognitive process, especially their thinking trajectory. Therefore, talk moves could also be combined with turn-taking patterns to reveal an individual's socio-cognitive process. In addition, a few studies considered linguistic characteristics of peer talk (e.g., turn length, speaking speed, pitch, tone, volume; 15%), the post-hoc assessment on the quality of collaborative discourse (9%, $n = 8$), and sequence (e.g., phases of collaborative problem solving; 3%, $n = 3$). Linguistic features, in particular turn length, were mainly adopted to indicate the amount of participation of group members. One study also captured the usage of capitals and exclamation marks to indicate the speakers' sentiment (Tat & Carpendale, 2002). Talk sequence reflects a group's underlying cognitive structure. It was mainly analysed to inform the development of group cognition. Post-hoc assessment was used to analyse various aspects of the underlying learning process.

3.4. Visualisation design

3.4.1. Visualisation categorisation

The results of this review revealed that 62% of the included VRCD involved a time axis. Most of the included VRCD incorporated static graphs (73%), whereas very few used dynamic elements (29%). In addition, only approximately 55% of the representations allowed users to define/select visual components to display.

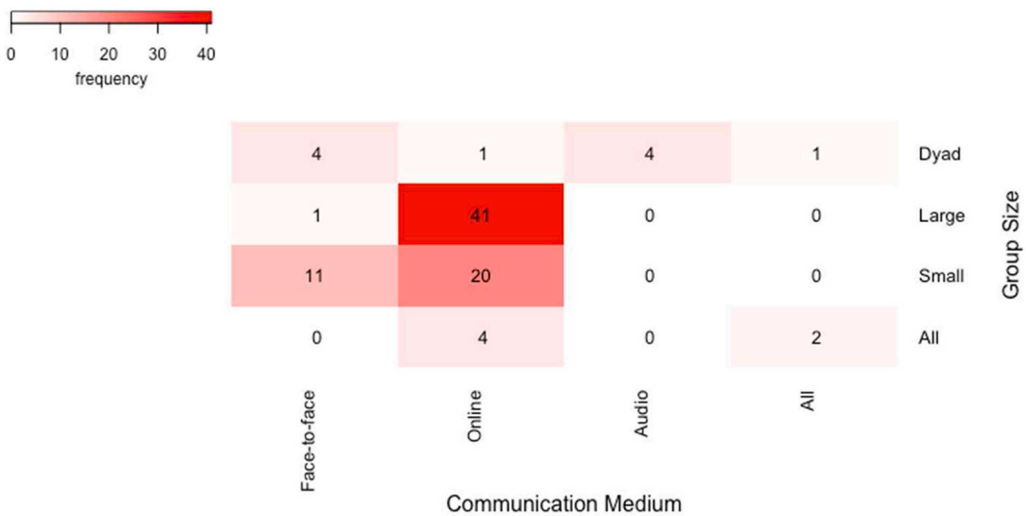

Fig. 4. A heatmap of communication medium and group size.

Fig. 5. A heatmap of target discourse feature and target underlying learning process.

3.4.2. Display format

The results showed that the most used formats were novel designs (42%), networks (36%), raw text (35%), bar plots (27%), bubble plots (24%), line graphs (19%), word clouds (12%), and pie plots (11%). Radar plots, heat maps, and tree graphs (all smaller than 10%) were seldom used.

A substantial proportion of VRCD used novel elements to illustrate collaborative discourse in specific contexts. These novel visualisations were largely designed to illustrate the process of collaboration. For example, Fu et al. (2017) proposed a visualisation called Thread River to show the temporal and hierarchical structure of complex and lengthy online posts. The Mondrian Transcription method was used to illustrate a conversation process with dynamic physical positions (Shapiro et al., 2017). There were also various innovative and engaging VRCD used to show turn-taking processes, such as Visiphone (Karahalios, 2004), Conversation Clock (Karahalios & Bergstrom, 2009), Second Messenger (DiMicco & Bender, 2007), and Reflect (Bachour et al., 2010). In addition, novel interfaces were designed to support synchronous online discussion, such as Chat Circles (Donath et al., 1999) and Fugue (Shankar et al., 2000).

Collaborative discourse consists of many momentary utterances. Our results showed that a network is the most widely used pictorial format to capture the cumulative effects of these moments. Existing VRCD have involved various types of networks, including social networks, socio-semantic networks, epistemic networks, and some other self-defined networks. For example, social networks have been extensively used to illustrate the participatory structure of a conversation (e.g., Adraoui et al., 2018; Jin, 2017). Social networks enable people to see connections among group members and whether there is a sub-group, isolate, or leader. Semantic networks help delineate the semantic structure underlying collaborative discourse (e.g., Sha et al., 2010).

Raw text was also found to be commonly used in the existing VRCD. Interactive visualisations tended to include original collaborative discourse to contextualise extracted patterns or features more often than non-interactive ones (χ^2 (1, n = 89) = 23.91, p < .001). More than half (57%) interactive VRCD incorporated raw text. For example, Concept Trail not only depicted dynamic occurrences of target concepts but also highlighted these concepts in raw conversations (van Leeuwen et al., 2014). Similarly, Knowledge Building Discourse Explorer (KBDeX) contained raw text and highlighted keywords (Oshima et al., 2012, 2018). Raw text provided the original context to aid users' interpretations of the abstracted result, and, thus, the inclusion of raw text echoed the fundamental goal of visual analytics approaches: to combine human intelligence with the computational advantage of technology.

The results also suggested that basic bar graphs or pie plots have been used to communicate purely cumulative features of collaborative discourse. For example, the Knowledge Connections Analyser mainly used familiar bar graphs and pie plots instead of complicated networks to facilitate young students' reflection on the number of collaborators and the percentage of specific types of notes (van Aalst et al., 2012).

The temporal dimension of peer talk is usually represented by turn location, or, sometimes, the actual time. Bubble plots and line graphs have been widely used by various researchers to illustrate how one aspect or feature of peer talk unfolds over time, such as the distribution of certain concepts in group talk (van Leeuwen et al., 2014), the evolution of participation inequity (Lämsä et al., 2018), the state of consensus building (Xiong et al., 2012), and the evolution of collaboration among group members (Trausan-Matu et al., 2014). These features are theoretically assumed to be meaningful for explaining the temporality of peer talk and are typically extracted at different time points to underscore this temporality.

3.4.3. Quality of visualisation design

The results revealed that approximately 55% of VRCD did not mention any visualisation design principles. Given this result, features of VRCD that considered design principles were further explored in terms of publication sources, attributes of visualisations, and target users. We found that publications from the visualisation-oriented community were substantially more likely to follow visualisation design principles (68%) than those from the learning-oriented community (25%), χ^2 (1, $n = 85$) = 16.02, $p < .001$. We also observed that the learning-oriented community was significantly less likely to create innovative designs, χ^2 (1, $n = 85$) = 23.84, $p < .001$. In contrast, the visualisation-oriented community was more likely to follow strict procedures to justify the choice of display formats and visual interface design, which helped generate diverse and innovative visualisations.

The nature of the target users did not significantly influence the consideration of visualisation design principles. If multiple target users were addressed as a single group, we observed that visualisation principles were most incorporated into learner-facing visualisations (53%) and least into those intended for researchers (13%). However, such differences did not reach a significant level, χ^2 (3, $n = 89$) = 5.94, $p = .12$.

To determine the effect of visualisation attributes, the complexity of visualisation was quantified through the addition of the dichotomous variables 'dynamic' and 'interactivity' and fourteen dichotomous variables depicting the display format. We found that visualisations that incorporated visualisation design principles ($M = 3.40$, $SD = 1.65$) were not significantly more complex than those that did not incorporate them ($M = 3.00$, $SD = 1.65$), $t\,(87) = -1.14$, $p = .26$. Neither were interactive visualisations significantly more likely to follow design principles (53%) than non-interactive visualisations (35%), χ^2 (1, $n = 89$) = 2.90, $p = .089$. However, we found that representations incorporating novel elements were substantially more likely to explicitly mention visual design principles (68%) than were those consisting of traditional graphs (29%), χ^2 (1, $n = 89$) = 13.10, $p < .001$.

3.5. Analytic technique

In this review, we found that around 73% of the included VRCD utilised advanced analytical techniques. The most widely used analytical technique, text analytics (39%), was typically adopted to analyse emergent topics in collaborative discourse via topic-modelling methods. Other common applications were automatic coding through supervised machine learning, clustering discourse units based on semantic similarity, and sentiment analysis. Network analysis was also widely used by existing VRCD (27%). A few VRCD incorporated speech processing (7%, $n = 6$) and process mining techniques (2%, $n = 2$).

3.5.1. Text analytics

Our results indicated that a major application of text analytics is to enable the automatic extraction of particular features of discourse, such as underlying topics/concepts and language usage. We also found that text analytics have been used for the automated labelling of discourse units, to provide prompt feedback to groups and prevent the need for tedious, labour-intensive coding tasks. For example, van Leeuwen et al. (2014) adopted text analytics to automatically identify whether a group was in agreement or disagreement and facilitated teachers' in-time diagnosis and intervention when guiding multiple collaborating groups.

3.5.2. Network analysis

We found that the primary network analysis technique was social network analysis. For example, some included VRCD used different centrality metrics to describe the dynamism of a social network from different perspectives (e.g., Boroujeni et al., 2017; Herring et al., 2005). Some VRCD involved socio-semantic network analysis. For example, KBDeX, developed by Oshima et al. (2012), is such a typical visual analysis tool. KBDeX also provides dynamic monitoring of different network indices to support in-depth discourse analysis. This review also indicated that epistemic network analysis has been increasingly used because of its flexibility in text segmenting and node definition, its support of paired network comparison, and its strong theoretical foundation (Shaffer & Ruis, 2017).

3.5.3. Speech processing and process mining

We found that speech processing techniques have been mainly used to support the real-time mirroring of small-group talk (Bachour et al., 2010; DiMicco & Bender, 2007; Karahalios, 2004; Karahalios & Bergstrom, 2009). The included VRCD mainly adopted speech processing techniques to automatically detect the participant speaking at each point and the volume of speech. They also applied various rules to address the issue of overlapping speech, by either registering the loudest voice (Bachour et al., 2010) or visualising the authentic turn-negotiating phase (DiMicco & Bender, 2007; Karahalios, 2004). These visual tools are intended mainly to strengthen social presence in a remote collaboration or increase members' awareness of their participation equality.

The included VRCD in this review adopted process mining techniques to dissect the peer communication process (e.g., Sedrakyan et al., 2020). For example, some VRCD used lag sequential analysis to analyse sequential patterns of temporal progress and reveal the dynamic features of human collaborative activities (e.g., Chang et al., 2017; Lefstein et al., 2015).

4. Discussion and future research agenda

4.1. Visual analysis approaches could help uncover temporal patterns

We found that most of VRCD under study involved the temporal dimension of collaborative discourse, presented as networks, novel designs, raw text, bubble plots, or line graphs. We also found that text analytics, network analysis, speech processing techniques, and process mining were the main advanced analytical techniques underlying existing VRCD, as these are powerful tools to support the detection of temporal patterns. These findings indicate that visual analysis approaches are appropriate and advantageous for dissecting the temporality of collaborative discourse.

The temporal analysis of interactions has also been increasingly emphasised in recent years (e.g., Csanadi et al., 2018; Knight et al., 2017; Swiecki et al., 2020), as the temporal development of peer dialogue must be known for educational sense-making to be performed (Mercer, 2008) and to further scaffold promotive interdependencies of collaboration.

Interdependencies of peer talk exist at multiple time scales (Wise & Chiu, 2011). Previous visual analysis approaches have largely focused on the micro-time context and analysed collaborative discourse at the turn or move level. However, it is challenging to determine the number of recent turns in the micro-time context; therefore, one turn has typically been taken as the time unit in existing visual analysis approaches on semantic analysis, especially in socio-semantic network analysis (e.g., Oshima et al., 2018).

Nevertheless, many studies have shown that the length of the micro-time context for different discourse features may vary and even extend beyond the previous turn (Chen et al., 2012; Chiu & Khoo, 2005; Molenaar & Chiu, 2014). This problem is addressed by another common technique known as the sliding window, which allows for a flexible definition of the micro-time context. The sliding window involves the segmentation of discourse into multiple overlapping stanzas, according to a flexible window length (e.g., 10 s, three lines of talk, or three turns). However, determining an appropriate window length for specific contexts and target features remains a challenge, although there have been attempts to find an automated solution to this challenge (Ruis et al., 2018, 2019).

4.2. *Learning theories and visualisation design could be further integrated*

Visual analytics is an emerging multidisciplinary research area, but our review of the current visual analysis research on collaborative discourse revealed a gap between the visualisation-oriented and learning-oriented communities. This was particularly apparent in the differences in these communities' theoretical considerations and visualisation designs, suggesting that the greater the emphasis on learning theory in VRCD, the less the consideration given to visualisation design. According to the literature we reviewed, more than 90% of VRCD in the learning-oriented community set goals that were informed by theoretical considerations, but only 25% of them justified their visualisation design decisions. In contrast, around 40% of VRCD in the visualisation-oriented community referred to learning theories, but around 70% of them mentioned visualisation design principles. Our findings are in line with a previous review on the overall development of visual learning analytics in education (Vieira et al., 2018), which concluded that information visualisation experts focused on visualisation innovations and seldom referred to educational theories, whereas education researchers were guided by education theories but always used traditional display formats and ignored visualisation design principles.

It remains challenging for these two communities to incorporate each other's good practices and thereby construct a common design framework to guide the development of VRCD. A particular challenge is the integration of learning theories in the process model of visual analytics (see Fig. 1). The visualisation-oriented community generally begins with a domain challenge or problem, follows with a strict procedure to elaborate and justify the design process, and iteratively improves the design based on users' and relevant domain experts' feedback (e.g., Fu et al., 2017; Kwon et al., 2016). It usually considers domain theories by consulting domain experts, but this seldom explicitly justifies the theoretical considerations (e.g., Stenliden et al., 2017). In contrast, the learning-oriented community proposes new visualisation designs to address issues in learning research (e.g., Shapiro et al., 2017), but it seldom reports design considerations as in the visualisation community.

We therefore call for a higher level of collaboration between these two communities. Innovative approaches have started to achieve this goal. For example, Hillaire et al. (2016) proposed a six-step model to guide the development of visual learning analytics tools. In this model, the first step is to define an educational goal informed by educational theories, and subsequent steps involve the definition of the target users, an interdisciplinary paper prototyping process, a formative evaluation, mock data, and implementation.

It is important to note that the observed gap between the two communities concerning the usage of learning and visualisation theories was largely based on the explicit reports of the authors. Different publication practices in the two communities might make the authors less likely to report theories or design principles that might be unfamiliar to their audiences. For example, the IEEE Visualisation Conference, a major venue for visualisation research, holds the following annual conferences: Visual Analytics Science and Technology, Information Visualisation, and Science Visualisation. All three conferences require submitted papers to elaborate and justify design decisions when developing a visualisation system to solve a target problem. Further research could use more comprehensive data to confirm the finding on the gap between communities.

4.3. *Teachers could be better supported to facilitate group discussion*

Existing VRCD have largely targeted learners and researchers, with insufficient attention being paid to educators. A large body of research has emphasised the role of teachers in guiding peer talk (Gillies et al., 2008; Gillies & Khan, 2008). Teachers can facilitate productive collaborative talk by providing adaptive interventions in the process of student collaboration (Webb et al., 2009), but it is very challenging for teachers to decide which type of guidance to give to student groups and when. The need for such guidance also tends to overload teachers when they are required to monitor multiple groups in a typical classroom setting.

Consequently, visual analysis approaches have been adopted to support teachers' facilitation of student collaboration. For example, the Argunaut system helps facilitate teachers' moderation of multiple synchronous group discussions by providing them with structured real-time group awareness information (Schwarz & Asterhan, 2011). However, this is currently a small body of research that has yet to examine how VRCD could support teacher guidance in student collaboration (van Leeuwen & Rummel, 2019).

Another wide body of knowledge informs how teachers can facilitate student collaborative communication, which could be used to guide the future development of VRCD. For example, we found that many studies suggested that teachers could intervene in group work when students reach a collaborative stalemate and fail to make progress or when one student dominates the discussion and limits the development of authentic dialogue (e.g., Cohen, 1994). Reznitskaya and colleagues (Reznitskaya et al., 2007, 2009) highlighted the importance of teacher prompting in facilitating student argumentative skills. Continuous scaffolding provided by teachers can also create an inclusive and supportive environment, in which students are encouraged to be critical and constructive thinkers, and to use talk to build high-quality interaction with others. At the same time, teachers should exercise caution in how much assistance they supply to student groups, as task-related help may be detrimental to group problem solving performance (Chiu, 2004; Gillies, 2004). Rather, researchers have shown that it is beneficial for teachers to guide students' attention to important issues, exhort students to deepen their thinking, and provide expanded elaborations (Chiu, 2004; Hogan et al., 1999).

4.4. More attention could be given to the face-to-face context

Previous VRCD have focused on the online context, largely neglecting the role of face-to-face contexts. This may be attributable to data accessibility, as, in contrast to an online context, preparing face-to-face data for further analysis is onerous. Therefore, it is challenging to provide real-time or in-time post hoc feedback concerning face-to-face collaborative talk, which limits the application of VRCD to analysing this mode of interaction. The popularity and strengths of online discussion may also be factors, as these are more flexible in terms of time and structure. An online discussion environment also makes it easier to embed VRCD.

Notwithstanding these challenges, group talk is very common in typical classroom settings. Students in the face-to-face context also need to be aware of their talk quality and should be provided with explicit guidance on how to use language effectively and how to regulate group interactions (e.g., King, 2008; Näykki et al., 2017). Some existing VRCD have adopted speech processing techniques to provide timely feedback on small-group talk (Bachour et al., 2010; DiMicco & Bender, 2007; Karahalios, 2004). However, the feedback focused mainly on basic turn-taking structure. Hence, researchers could make more of an effort to design timely VRCD on semantics and moves in group talk.

Face-to-face real-time talk is distinct from online asynchronous communications in multiple aspects (Asterhan & Schwarz, 2010). For example, a face-to-face environment tends to elicit instantaneous reactions, whereas an online environment, especially if it is asynchronous, allows students more thinking time. In addition, face-to-face talk is augmented by audio-visual stimuli such as pitch, tone, gaze, facial expressions, and body gestures (Bless & Greifeneder, 2017) and therefore tends to establish a stronger socio-emotional presence than online discussion (Garrison, 2007). How these contextual differences might lead to different findings concerning collaborative discourse remains underexplored, especially with respect to those features that typify high-quality collaborative discourse.

4.5. More efforts could be made to provide advanced support

We found that mirroring was the dominating affordance in existing VRCD; there were few instances of advanced affordances, such as alerting or advising. This is attributable to the complexity of collaboration as well as the diversity of contextual elements. It is challenging for VRCD to provide users with advanced support such as that which indicates a desired state, guides attention to problematic events, or suggests desired actions. Human knowledge and experience are therefore necessary to interpret VRCD, reflect the collaborative process, and take further action, which is in line with the objective of visual analytics approaches.

However, VRCD that mirror complex collaborative discourse also tend to be complicated. An effective visual design should balance functional and aesthetic considerations (Simoff et al., 2008). With respect to the functional aspects, a successful graph should make the patterns, trends, or comparisons of the presented data easily and immediately comprehensible. In terms of aesthetics, the visual appeal of a graph should not obscure its message (Kosslyn, 2006). Going even further, Kosslyn (2006) specified eight psychological principles as handy guidelines to ensure that a graphic design connects with an audience, directs and holds their attention, and promotes their understanding of the data. Therefore, a complex VRCD could include advanced affordances, such as alerting or advising to guide attention, limit cognitive load, and aid in high-level reflection.

There is a large body of literature on high-quality collaborative discourse that could inform the design of these advanced affordances. For example, successful collaborative work requires individuals to build positive interdependence (Wang, 2009) and dynamically regulate collaborative processes (Borge et al., 2018; Borge & White, 2016; Järvela et al., 2013). Individuals should also establish mutual goals, collaboration principles, a collective timeline, and dynamic monitoring and adjusting practices in their collaborative processes (Stanton & Fairfax, 2007). Social loafing or diffusion of responsibility may jeopardise individual or group outcomes (Webb et al., 2009). Individuals in effective learning teams actively press for explanations and justifications from their peers (Gillies, 2019). Thus, VRCD might appropriately alert users to or advise them on the occurrence of severe cognitive divergence, social conflict, constant isolated contributions, consistent isolates/subgroups, or the neglect of specific talk moves.

In addition, advanced analytics could be employed to provide further insights for designing the advanced affordances or to help provide timely feedback. For example, epistemic network analysis could be integrated with social network analysis to allow more insights to be gained into group work (Gašević et al., 2018). Natural language processing methods could be adopted to capture the interdependency of dialogue data and provide users with real-time feedback (Li et al., 2007; Sullivan & Keith, 2019). Neural network approaches have also been extensively explored for automatic labelling, as these approaches can outperform traditional machine learning methods (Meng et al., 2018).

5. Conclusions

Visual analytics gives learners, educators, and researchers an opportunity to interpret and improve collaborative communication in a more systematic and engaging manner. We conducted an overview of the scholarly landscape of current applications of visual analysis approaches in collaborative discourse based on four dimensions: goals, data sources, visualisation designs, and analytic techniques. We found that visual analysis approaches are suitable and advantageous to decompose the temporality of collaborative discourse. However, visual analytics is a multidisciplinary research area, and good practices in the visualisation-oriented and learning-oriented communities have yet to be fully integrated; that is, few existing VRCD have considered both learning theories and visualisation design principles. Constructing a VRCD design framework that will deepen the integration of visual analytics and learning analytics is therefore a crucial and challenging task for future researchers in this promising field.

We also found that more attention should be devoted to supporting teachers and face-to-face group discussion. Consequently, another challenge for future VRCD will be to go beyond simply mirroring discussion processes to providing advanced affordances such as alerting or advising, thereby further guiding users' attention and facilitating their decision making. The well-established features of high-quality collaborative discourse will undoubtedly spur and inspire the efforts of future researchers to strengthen the support for VRCD.

Acknowledgement

This work was supported by the Research Grants Council, University Grants Committee (Grant No. 17608318) and by the Public Policy Research Funding Scheme from the Policy Innovation and Co-ordination Office of the Government of the HKSAR (Grant No. 2017.A8.073.18C). The authors would also like to thank Mr. Pengfei Li and Mr. Alben Jinjian Yu for their kind help in data coding, and anonymous reviewers for their constructive feedback in revisions of this manuscript.

References

Studies included in the review were marked with an asterisk.

(*) Adachi, H., Myojin, S., & Shimada, N. (2014). Tablet system for sensing and visualizing statistical profiles of multi-party conversation. In *2014 IEEE 3rd global conference on consumer electronics (GCCE)* (pp. 407–411). IEEE.

(*) Adeniran, A., Masthoff, J., & Beacham, N. (2019). Model-based characterization of text discourse content to evaluate online group collaboration. In *International conference on artificial intelligence in education* (pp. 3–8). Cham: Springer.

(*) Adraoui, M., Retbi, A., Idrissi, M. K., & Bennani, S. (2018). Network visualization algorithms to evaluate students in online discussion forums: A simulation study. In *2018 International conference on intelligent systems and computer vision* (pp. 1–6). IEEE.

(*) Ahn, J. W., Hammock, J., Parr, C., Preece, J., Shneiderman, B., Schulz, K., Hansen, D., Rotman, D., & He, Y. (2012). Visually exploring social participation in Encyclopedia of Life. In *2012 International conference on social Informatics* (pp. 149–156). IEEE.

(*) Allaymoun, M. H. (2015). Graphical visualization of rhetorical structures in chats. In *Proceedings of the 11th International scientific conference eLearning and software for education* (pp. 407–414). Carol I National Defence University Publishing House.

Andrienko, N., Lammarsch, T., Andrienko, G., Fuchs, G., Keim, D., Miksch, S., & Rind, A. (2018). Viewing visual analytics as model building. *Computer Graphics Forum, 37*(6), 275–299.

(*) Angus, D., Smith, A., & Wiles, J. (2012). Conceptual recurrence plots: Revealing patterns in human discourse. *IEEE Transactions on Visualization and Computer Graphics, 18*(6), 988–997.

Asterhan, C. S., & Schwarz, B. B. (2010). Online moderation of synchronous e-argumentation. *International Journal of Computer-Supported Collaborative Learning, 5*(3), 259–282.

(*) Atapattu, T., Falkner, K., & Tarmazdi, H. (2016). Topic-wise classification of MOOC discussions: A visual analytics approach. In *Proceedings of the 9th International conference on educational data mining.* Raleigh, NC, USA: (EDM).

(*) Bachour, K., Kaplan, F., & Dillenbourg, P. (2010). An interactive table for supporting participation balance in face-to-face collaborative learning. *IEEE Transactions on Learning Technologies, 3*(3), 203–213.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. New York: Cambridge University Press.

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences, 12*(3), 307–359. https://doi.org/10.1207/S15327809JLS1203_1

Belland, B. R., Kim, C. M., & Hannafin, M. J. (2013). A framework for designing scaffolds that improve motivation and cognition. *Educational Psychologist, 48*(4), 243–270. https://doi.org/10.1080/00461520.2013.838920

Bless, H., & Greifeneder, R. (2017). Introduction: What is social cognition research about? In R. Greifeneder, H. Bless, & K. Fiedler (Eds.), *Social cognition: How individuals construct social reality* (2nd ed., pp. 1–15). London: Psychology Press. https://doi.org/10.4324/9781315648156.

Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies, 10*(4), 405–418. https://doi.org/10.1109/TLT.2017.2740172

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science, 323*(5916), 892–895. https://doi.org/10.1126/science.1165821

Borge, M., Ong, Y. S., & Rosé, C. P. (2018). Learning to monitor and regulate collective thinking processes. *International Journal of Computer-Supported Collaborative Learning, 13*(1), 61–92.

Borge, M., & White, B. (2016). Toward the development of socio-metacognitive expertise: An approach to developing collaborative competence. *Cognition and Instruction, 34*(4), 323–360. https://doi.org/10.1080/07370008.2016.1215722

(*) Boroujeni, M. S., Hoppe, H. U., Hecking, T., & Dillenbourg, P. (2017). Dynamics of MOOC discussion forums. *ACM international conference proceeding series*, 128–137. https://doi.org/10.1145/3027385.3027391.

Buckingham Shum, S., & Crick, R. D. (2016). Learning analytics for 21st century competencies. *Journal of Learning Analytics, 3*(2), 6–21.

Chang, C. J., Chang, M. H., Chiu, B. C., Liu, C. C., Chiang, S. H. F., Wen, C. T., Hwang, F. K., Wu, Y. T., Chao, P. Y., Lai, C. H., Wu, S. W., Chang, C. K., & Chen, W. (2017). An analysis of student collaborative problem solving activities mediated by collaborative simulations. *Computers & Education, 114*, 222–235, 2017.

(*) Chen, Y. (2015). Visual opinion analysis of threaded discussions. In *2015 IEEE International conference on data mining workshop* (pp. 646–651). IEEE.

Chen, G. (2019). Visual learning analytics to support classroom discourse analysis for teacher professional learning and development. In N. Mercer, R. Wegerif, & L. Major (Eds.), *International handbook of research on dialogic education* (pp. 167–181). Abingdon, UK: Routledge.

Chen, G. (2020). A visual learning analytics (VLA) approach to video-based teacher professional development: Impact on teachers' beliefs, self-efficacy, and classroom talk practice. *Computers & Education, 144*, 103670.

Chen, G., Chan, C. K. K., Chan, K. K. H., Clarke, S. N., & Resnick, L. B. (2020a). Efficacy of video-based teacher professional development for increasing classroom discourse and student learning. *The Journal of the Learning Sciences, 29*(4–5), 642–680.

Chen, G., Lo, C. K., & Hu, L. (2020b). Sustaining online academic discussions: Identifying the characteristics of messages that receive responses. *Computers & Education, 156*, 103938.

Chen, G., Chiu, M. M., & Wang, Z. (2012). Predicting social cues during online discussions: Effects of evaluations and knowledge content. *Computers in Human Behavior, 28*(4), 1497–1509. https://doi.org/10.1016/j.chb.2012.03.017

Chen, B., Resendes, M., Chai, C. S., & Hong, H. Y. (2017). Two tales of time: Uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interactive Learning Environments, 25*(2), 162–175. https://doi.org/10.1080/10494820.2016.1276081

Chen, J., Wang, M., Kirschner, P. A., & Tsai, 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.

(*) Chen, B., & Zhang, J. (2016). Analytics for knowledge creation: Towards epistemic agency and design-mode thinking. *Journal of Learning Analytics, 3*(2), 139–163. https://doi.org/10.18608/jla.2016.32.7.

Chi, M. T., & Menekse, M. (2015). Dialogue patterns in peer collaboration that promote learning. In L. B. Resnick, C. S. C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 263–274). Washington, DC: American Educational Research Association.

(*) Chin, A., & Chignell, M. (2006). A social hypertext model for finding community in blogs. In *Proceedings of the seventeenth conference on Hypertext and hypermedia* (pp. 11–22). New York, NY, USA: ACM Press.

(*) Chinn, C. A., O'Donnell, A. M., & Jinks, T. S. (2000). The structure of discourse in collaborative learning. *The Journal of Experimental Education, 69*(1), 77–97. https://doi.org/10.1080/00220970009600650.

Chiu, M. M. (2004). Adapting teacher interventions to student needs during cooperative learning: How to improve student problem solving and time on-task. *American Educational Research Journal, 41*(2), 365–399. https://doi.org/10.3102/00028312041002365

Chiu, M. M., & Khoo, L. (2005). A new method for analyzing sequential processes: Dynamic multilevel analysis. *Small Group Research, 36*(5), 600–631. https://doi.org/10.1177/1046496405279309

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64*(1), 1–35.

Cohen, E. G., & Lotan, R. A. (2014). *Designing group work: Strategies for the heterogeneous classroom* (3rd ed.). New York: Teachers College Press.

(*) Critchlow, M. D. (2006). *Conveying social presence Data on the Desktop with BuddySquares* (Doctoral dissertation). University of Illinois at Urbana-Champaign.

Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). When coding-and-counting is not enough: Using epistemic network analysis (ENA) to analyze verbal data in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 13*(4), 419–438.

(*) Dave, K., Wattenberg, M., & Muller, M. (2004). Flash forums and forumReader: Navigating a new kind of large-scale online discussion. In *Proceedings of the 2004 ACM conference on computer supported cooperative work* (pp. 232–241). ACM.

(*) Deng, S., Zhou, Y., Zhang, P., & Abbasi, A. (2019). Using discussion logic in analyzing online group discussions: A text mining approach. *Information & Management, 56*(4), 536–551.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–19). Oxford: Elsevier.

(*) DiMicco, J. M., & Bender, W. (2007). Group reactions to visual feedback tools. In *International conference on persuasive technology* (pp. 132–143). Berlin, Heidelberg: Springer.

(*) Donath, J., Karahalios, K., & Viégas, F. (1999). Visualizing conversation. *Journal of Computer-Mediated Communication, 4*(4), JCMC442.

(*) Dörk, M., Gruen, D., Williamson, C., & Carpendale, S. (2010). A visual backchannel for large-scale events. *IEEE Transactions on Visualization and Computer Graphics, 16*(6), 1129–1138.

Dyke, G., Kumar, R., Ai, H., & Rosé, C. P. (2012). Challenging assumptions: Using sliding window visualizations to reveal time-based irregularities in CSCL processes. In *Proceedings of the 10th International conference of the learning sciences* (Vol. 1, pp. 363–370). Sydney: Australia.

(*) Erickson, T., & Kellogg, W. A. (2003). Social translucence: Using minimalist visualisations of social activity to support collective interaction. In K. Höök, D. Benyon, & A. Munroe (Eds.), *Designing information spaces: The social navigation approach* (pp. 17–41). London: Springer-Verlag.

Fu, E. L. F., van Aalst, J., & Chan, C. K. K. (2016). Toward a classification of discourse patterns in asynchronous online discussions. *International Journal of Computer-Supported Collaborative Learning, 11*(4), 441–478.

(*) Fu, S., Zhao, J., Cui, W., & Qu, H. (2017). Visual analysis of MOOC forums with iForum. *IEEE Transactions on Visualization and Computer Graphics, 23*(1), 201–210. https://doi.org/10.1109/TVCG.2016.2598444.

Garrison, D. R. (2007). Online community of inquiry review: Social, cognitive, and teaching presence issues. *Journal of Asynchronous Learning Networks, 11*(1), 61–72.

Gašević, D., Joksimović, S., Eagan, B. R., & Shaffer, D. W. (2018). Sens: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior, 92*, 562–577. https://doi.org/10.1016/j.chb.2018.07.003, 2019.

Gillies, R. M. (2004). The effects of communication training on teachers' and students' verbal behaviours during cooperative learning. *International Journal of Educational Research, 41*(3), 257–279. https://doi.org/10.1016/j.ijer.2005.07.004

Gillies, R. M. (2019). Promoting academically productive student dialogue during collaborative learning. *International Journal of Educational Research, 97*, 200–209. https://doi.org/10.1016/j.ijer.2017.07.014

Gillies, R. M., Ashman, A., & Terwel, J. (Eds.). (2008). *The teacher's role in implementing cooperative learning in the classroom*. Boston, MA: Springer.

Gillies, R. M., & Khan, A. (2008). The effects of teacher discourse on students' discourse, problem-solving and reasoning during cooperative learning. *International Journal of Educational Research, 47*(6), 323–340. https://doi.org/10.1016/j.ijer.2008.06.001

Goffman, E. (1981). *Forms of talk*. Philadelphia, PA: University of Pennsylvania Press.

(*) Herring, S. C., Kouper, I., Paolillo, J. C., Scheidt, L. A., Tyworth, M., Welsch, P., & Yu, N. (2005). Conversations in the blogosphere: An analysis "from the bottom up". In *Proceedings of the thirty-eighth Hawai'i International conference on system sciences (HICSS-38)* (p. 107). Los Alamitos: IEEE Press.

Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin, & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 37–56). Dordrecht, Netherlands: Springer.

Hillaire, G., Rappolt-Schlichtmann, G., & Ducharme, K. (2016). Prototyping visual learning analytics guided by an educational theory informed goal. *Journal of Learning Analytics, 3*(3), 115–142. https://doi.org/10.18608/jla.2016.33.7

Hmelo-Silver, C. E. (2003). Analyzing collaborative knowledge construction: Multiple methods for integrated understanding. *Computers and Education, 41*(4), 397–420.

(*) Hoch, C., Zellner, M., Milz, D., Radinsky, J., & Lyons, L. (2015). Seeing is not believing: Cognitive bias and modelling in collaborative planning. *Planning Theory & Practice, 16*(3), 319–335.

Hogan, K., Nastasi, B. K., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17*(4), 379–432. https://doi.org/10.1207/S1532690XCI1704_2

(*) Hoque, E., & Carenini, G. (2014). Convis: A visual text analytic system for exploring blog conversations. *Computer Graphics Forum, 33*(3), 221–230.

(*) Imtiyazi, M. A., Alamsyah, A., Junaedi, D., & Pradana, J. A. (2016). Word association network approach for summarizing Twitter conversation about public election. In *2016 4th International conference on information and communication technology (ICoICT)* (pp. 1–4). IEEE.

(*) Indratmo, J., Vassileva, J., & Gutwin, C. (2008). Exploring blog archives with interactive visualization. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 39–46). ACM.

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

(*) Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior, 23*(3), 1105–1125. https://doi.org/10.1016/j.chb.2006.10.005.

Järvela, S., Järvenoja, H., Malmber, J., & Hadwin, A. (2013). Exploring socially shared regulation in the context of collaboration. *Journal of Cognitive Education and Psychology, 12*(3), 267–286. https://doi.org/10.1891/1945-8959.12.3.267

(*) Jermann, P., & Dillenbourg, P. (2008). Group mirrors to support interaction regulation in collaborative problem solving. *Computers and Education, 51*(1), 279–296. https://doi.org/10.1016/j.compedu.2007.05.012.

(*) Jin, S. H. (2017). Using visualization to motivate student participation in collaborative online learning environments. *Educational Technology & Society, 20*(2), 51–62.

Johnson, D. W., & Johnson, R. T. (1992). Positive interdependence: Key to effective cooperation. In R. Hertz- Lazarowitz, & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 174–199). Cambridge, UK: Cambridge University Press.
(*) Karahalios, K. G. (2004). *Social catalysts: Enhancing communication in mediated spaces* (Master's thesis). ACM Digital Library (AAI0807517).
(*) Karahalios, K. G., & Bergstrom, T. (2009). Social mirrors as social signals: Transforming audio into graphics. *IEEE Computer Graphics and Applications, 29*(5), 22–32.
(*) Kazemitabar, M. A., Bodnar, S., Hogaboam, P., Chen, Y., Sarmiento, J. P., Lajoie, S. P., & Chan, L. (2016). Creating instructor dashboards to foster collaborative learning in on-line medical problem-based learning situations. In *International conference on learning and collaboration technologies* (pp. 36–47). Cham: Springer.
Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008). Visual analytics: Scope and challenges. In S. J. Simoff, M. H. Böhlen, & A. Mazeika (Eds.), *Visual data mining. Lecture notes in computer science* (Vol. 4404, pp. 76–90). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-71080-6_6.
Keim, D. A., Mansmann, F., & Thomas, J. (2009). Visual analytics: How much visualization and how much analytics? *ACM SIGKDD Explorations Newsletter, 11*(2), 5–8. https://doi.org/10.1145/1809400.1809403
(*) Kim, M., & Lee, E. (2012). A multidimensional analysis tool for visualizing online interactions. *Journal of Educational Technology & Society, 15*(3), 89–102.
King, A. (2008). Structuring peer interaction to promote higher-order thinking and complex learning in cooperating groups. In R. M. Gillies, A. F. Ashman, & J. Terwel (Eds.), *The teacher's role in implementing cooperative learning in the classroom* (pp. 73–91). Boston, MA: Springer US.
Knight, S., Wise, A. F., & Chen, B. (2017). Time for change: Why learning analytics needs temporal analysis. *Journal of Learning Analytics, 4*(3), 7–17. https://doi.org/10.18608/jla.2017.43.2
Kosslyn, S. M. (2006). *Graph design for the eye and the mind.* New York, NY: Oxford University Press.
(*) Kwon, B. C., Kim, S. H., Lee, S., Choo, J., Huh, J., & Yi, J. S. (2016). VisOHC: Designing visual analytics for online health communities. *IEEE Transactions on Visualization and Computer Graphics, 22*(1), 71–80. https://doi.org/10.1109/TVCG.2015.2467555.
(*) Lagatie, R., Wild, F., De Causmaecker, P., & Scott, P. (2011). Exposing knowledge in speech: Monitoring conceptual development in spoken conversation. In *Proceedings of 2011 IST-Africa conference* (pp. 1–8). IEEE.
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.
(*) Lämsä, J., Hämäläinen, R., Koskinen, P., & Viiri, J. (2018). Visualising the temporal aspects of collaborative inquiry-based learning processes in technology-enhanced physics learning. *International Journal of Science Education, 40*(14), 1697–1717. https://doi.org/10.1080/09500693.2018.1506594.
(*) Lee, Y. J., Bae, M. J., Woo, G., & Cho, H. G. (2009). A personalized visualizing and filtering system for a large set of responding messages on internet discussion forums. In *2009 ninth IEEE International conference on computer and information technology* (Vol. 2, pp. 160–165). IEEE.
(*) Lee, A. V. Y., & Tan, S. C. (2017). Temporal analytics with discourse analysis: Tracing ideas and impact on communal discourse. In *Proceedings of the seventh International learning analytics & knowledge conference* (pp. 120–127). ACM.
Lefstein, A., Snell, J., & Israeli, M. (2015). From moves to sequences: Expanding the unit of analysis in the study of classroom discourse. *British Educational Research Journal, 41*(5), 866–885. https://doi.org/10.1002/berj.3164
Li, Y., Wang, J., Liao, J., Zhao, D., & Huang, R. (2007). Assessing collaborative process in CSCL with an intelligent content analysis toolkit. In *Seventh IEEE International conference on advanced learning technologies* (pp. 257–261). IEEE.
(*) Li, J., Zhang, D., & Zhang, P. (2013). Supporting dynamic situation awareness in online group discussion: A visualization approach. In *2013 46th Hawaii International conference on system sciences* (pp. 470–479). IEEE.
Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research, 66*(4), 423–458. https://doi.org/10.3102/00346543066004423
Ludvigsen, S. (2016). CSCL: Connecting the social, emotional and cognitive dimensions. *International Journal of Computer-Supported Collaborative Learning, 11*(2), 115–121.
(*) Lund, K., Quignard, M., & Shaffer, D. W. (2017). Gaining insight by transforming between temporal representations of human interaction. *Journal of Learning Analytics, 4*(3), 102–122.
(*) Martinez-Maldonado, R., Kay, J., Yacef, K., & Schwendimann, B. (2012). An interactive teacher's dashboard for monitoring groups in a multi-tabletop learning environment. In *International conference on intelligent tutoring systems* (pp. 482–492). Berlin, Heidelberg: Springer.
(*) Mathur, P., & Karahalios, K. (2009). Visualizing remote voice conversations. In *CHI'09 extended abstracts on human factors in computing systems* (pp. 4675–4680). Boston, Massachusetts, USA: ACM.
(*) McCormick, J. (2013). Visualizing interaction: Pilot investigation of a discourse analytics tool for online discussion. *Bulletin of the IEEE Technical Committee on Learning Technology, 15*(2), 10–13.
Meng, Y., Rumshisky, A., & Sullivan, F. (2018). Automatic labeling of problem-solving dialogues for computational microgenetic learning analytics. In *Proceedings of the eleventh International conference on language resources and evaluation* (pp. 4056–4060). Miyazaki, Japan: LREC 2018.
Mercer, N. (2008). The seeds of time: Why classroom dialogue needs a temporal analysis. *The Journal of the Learning Sciences, 17*(1), 33–59. https://doi.org/10.1080/10508400701793182
Miller, M., & Hadwin, A. (2015). Scripting and awareness tools for regulating collaborative learning: Changing the landscape of support in CSCL. *Computers in Human Behavior, 52*, 573–588. https://doi.org/10.1016/j.chb.2015.01.050
(*) Ming, N. C., & Ming, V. L. (2013). Visualizing topics, time, and grades in online class discussions. In *Proceedings of the tenth computer-supported collaborative learning conference* (Vol. 2, pp. 105–108). Madison, USA: ISLS.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Reprint-preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Physical Therapy, 89*(9), 873–880.
Molenaar, I., & Chiu, M. M. (2014). Dissecting sequences of regulation and cognition: Statistical discourse analysis of primary school children's collaborative learning. *Metacognition and Learning, 9*(2), 137–160. https://doi.org/10.1007/s11409-013-9105-8
Näykki, P., Isohätälä, J., Järvelä, S., Pöysä-Tarhonen, J., & Häkkinen, P. (2017). Facilitating socio-cognitive and socio-emotional monitoring in collaborative learning with a regulation macro script–an exploratory study. *International Journal of Computer-Supported Collaborative Learning, 12*(3), 251–279.
Näykki, P., Järvelä, S., Kirschner, P. A., & Järvenoja, H. (2014). Socio-emotional conflict in collaborative learning—a process-oriented case study in a higher education context. *International Journal of Educational Research, 68*, 1–14.
Oshima, J., Oshima, R., & Fujita, W. (2018). A mixed-methods approach to analyze shared epistemic agency in jigsaw instruction at multiple scales of temporality. *Journal of Learning Analytics, 5*(1), 10–24.
(*) Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge building discourse explorer: A social network analysis application for knowledge building discourse. *Educational Technology Research & Development, 60*(5), 903–921. https://doi.org/10.1007/s11423-012-9265-2.
(*) Oyama, K., Watanabe, H., & Takeuchi, A. (2014). Visualization of turn-taking and mental workload in collaborative working environment. In *Proceedings of the 2014 IEEE 18th International conference on computer supported cooperative work in design (CSCWD)* (pp. 140–145). IEEE.
(*) Pascual-Cid, V., & Kaltenbrunner, A. (2009). Exploring asynchronous online discussions through hierarchical visualisation. In *2009 13th International conference information visualisation* (pp. 191–196). IEEE.
(*) Pupyrev, S., & Tikhonov, A. (2010). Analyzing conversations with dynamic graph visualization. In *2010 10th International conference on intelligent systems design and applications* (pp. 748–753). IEEE.
Reimann, P. (2009). Time is precious: Variable-and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 4*(3), 239–257.
(*) Resendes, M., Scardamalia, M., Bereiter, C., Chen, B., & Halewood, C. (2015). Group-level formative feedback and metadiscourse. *International Journal of Computer-Supported Collaborative Learning, 10*(3), 309–336.
Reznitskaya, A., Anderson, R. C., & Kuo, L. J. (2007). Teaching and learning argumentation. *The Elementary School Journal, 107*(5), 449–472. https://doi.org/10.1086/518623
Reznitskaya, A., Kuo, L. J., Clark, A. M., Miller, B., Jadallah, M., Anderson, R. C., & Nguyen- Jahiel, K. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education, 39*(1), 29–48. https://doi.org/10.1080/03057640802701952

Ribarsky, W., & Fisher, B. (2016). The human-computer system: Towards an operational model for problem solving. In *Proceedings of the annual Hawaii International conference on system sciences* (pp. 1446–1455). IEEE.

Rodríguez-Triana, M. J., Prieto, L. P., Vozniuk, A., Boroujeni, M. S., Schwendimann, B. A., Holzer, A., & Gillet, D. (2017). Monitoring, awareness and reflection in blended technology enhanced learning: A systematic review. *International Journal of Technology Enhanced Learning, 9*(2–3), 126–150. https://doi.org/10.1504/IJTEL.2017.084489

Ruis, A. R., Siebert-Evenstone, A. L., Pozen, R., Eagan, B. R., & Shaffer, D. W. (2018). A method for determining the extent of recent temporal context in analyses of complex, collaborative thinking. In *Proceedings of International conference of the learning sciences (ICLS 2018)* (pp. 1625–1626). London, UK: ISLS.

Ruis, A. R., Siebert-Evenstone, A. L., Pozen, R., Eagan, B. R., & Shaffer, D. W. (2019). Finding common ground: A method for measuring recent temporal context in analyses of complex, collaborative thinking. In *Proceedings of the computer supported collaborative learning conference (CSCL 2019)* (pp. 136–143). Lyon, France: ISLS.

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

Sacha, D., Stoffel, A., Stoffel, F., Kwon, B. C., Ellis, G., & Keim, D. A. (2014). Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics, 20*(12), 1604–1613. https://doi.org/10.1109/TVCG.2014.2346481

(*) Sack, W. (2000). Discourse diagrams: Interface design for very large-scale conversations. In *Proceedings of the 33rd annual Hawaii International conference on system sciences* (pp. 1–10). IEEE.

(*) Scardamalia, M. (2004). CSILE/Knowledge forum®. In A. Kovalchick, & K. Dawson (Eds.), *Education and technology: An encyclopedia* (pp. 183–192). Santa Barbara, CA: ABC-CLIO, Inc.

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning, 5*(1), 43–102.

Schnaubert, L., & Bodemer, D. (2019). Providing different types of group awareness information to guide collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 14*(1), 7–51.

(*) Schwarz, B. B., & Asterhan, C. S. (2011). E-Moderation of synchronous discussions in educational settings: A nascent practice. *The Journal of the Learning Sciences, 20*(3), 395–442. https://doi.org/10.1080/10508406.2011.553257.

(*) Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior, 107*, 1–15. https://doi.org/10.1016/j.chb.2018.05.004, 2020.

(*) Shaffer, D. W., & Ruis, A. R. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 175–187). New York, NY, USA: Society for Learning Analytics Research (SoLAR).

(*) Shahid, N., Ilyas, M. U., Alowibdi, J. S., & Aljohani, N. R. (2017). Word cloud segmentation for simplified exploration of trending topics on Twitter. *IET Software, 11*(5), 214–220.

Shah, N., & Lewis, C. M. (2019). Amplifying and attenuating inequity in collaborative learning: Toward an analytical framework. *Cognition and Instruction, 37*(4), 423–452.

(*) Shankar, T. R., VanKleek, M., Vicente, A., & Smith, B. K. (2000). Fugue: A computer mediated conversational system that supports turn negotiation. In *Proceedings of the 33rd annual Hawaii International conference on system sciences*. IEEE.

(*) Shapiro, B. R., Hall, R. P., & Owens, D. A. (2017). Developing and using interaction geography in a museum. *International Journal of Computer-Supported Collaborative Learning, 12*(4), 377–399.

(*) Sha, L., Teplovs, C., & van Aalst, J. (2010). A visualization of group cognition: Semantic network analysis of a CSCL community. In *Proceedings of the 9th International conference of the learning sciences* (Vol. 1, pp. 929–936). Chicago, Illinois, USA: ISLS.

(*) Siegel, M. A., Ellis, S. E., & Lewis, M. B. (2004). Designing for deep conversation in a scenarios-based e- learning environment. In *Proceedings of the 37th Hawaii International conference on system sciences*. IEEE.

Simoff, S., Böhlen, M. H., & Mazeika, A. (Eds.). (2008). *Visual data mining: Theory, techniques and tools for visual analytics*. Berlin Heidelberg: Springer.

(*) Sirbu, M. D., Dascalu, M., Crossley, S. A., McNamara, D. S., & Trausan-Matu, S. (2019). Longitudinal analysis and visualization of participation in online courses powered by cohesion network analysis. In *Proceedings of the 13th International conferences on computer-supported collaborative learning* (pp. 640–643). Lyon, France: ISLS.

(*) Smith, M. A., & Fiore, A. T. (2001). Visualization components for persistent conversations. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 136–143). ACM.

Stanton, P., & Fairfax, D. (2007). Establishing individual accountability for learning in an exam-less, group project course. In *Proceedings of the 2007 Middle Atlantic section fall conference of the American society for engineering education* (pp. 1–9).

Stenliden, L., Nissen, J., & Bodén, U. (2017). Innovative didactic designs: Visual analytics and visual literacy in school. *Journal of Visual Literacy, 36*(3–4), 184–201.

(*) Su, J., & Boydell, O. (2016). TopicListener: Observing key topics from multi-channel speech audio streams. In *2016 IEEE second International conference on big data computing service and applications* (pp. 85–94). IEEE.

Sullivan, F. R., & Keith, P. K. (2019). Exploring the potential of natural language processing to support microgenetic analysis of collaborative learning discussions. *British Journal of Educational Technology, 50*(6), 3047–3063. https://doi.org/10.1111/bjet.12875

(*) Sundaram, H., Xie, L., De Choudhury, M., Lin, Y. R., & Natsev, A. (2012). Multimedia semantics: Interactions between content and community. *Proceedings of the IEEE, 100*(9), 2737–2758.

Swiecki, Z., Ruis, A. R., Farrell, C., & Shaffer, D. W. (2020). Assessing individual contributions to collaborative problem solving: A network analysis approach. *Computers in Human Behavior, 104*, 105876.

Swiecki, Z., & Shaffer, D. W. (2018). Toward a taxonomy of team performance visualization tools. In *Proceedings of International conference of learning sciences* (pp. 328–335). London, UK: ISLS.

(*) Tat, A., & Carpendale, S. (2002). Visualising human dialog. In *Proceedings of the sixth International conference on information visualisation* (pp. 16–21). IEEE.

(*) Tat, A., & Carpendale, S. (2006). CrystalChat: Visualizing personal chat history. In *Proceedings of the 39th annual Hawaii International conference on system sciences (HICSS'06)* (Vol. 3, pp. 1–10). IEEE.

(*) Teo, H. J., Johri, A., & Mitra, R. (2013). Visualizing and analyzing productive structures and patterns in online communities using multilevel social network analysis. In *Proceedings of the tenth computer-supported collaborative learning conference* (Vol. 2, pp. 173–176). Madison, USA: ISLS.

Teplovs, C. (2015). Theory-led design of instruments and representations in learning analytics: Developing a novel tool for orchestration of online collaborative learning. *Journal of Learning Analytics, 2*(2), 44–46. https://doi.org/10.18608/jla.2015.22.4

Thomas, J. J., & Cook, K. A. (2005). *Illuminating the path: The research and development agenda for visual analytics*. Piscataway, NJ: IEEE Press.

(*) Trausan-Matu, S., Dascalu, M., & Rebedea, T. (2014). PolyCAFe-automatic support for the polyphonic analysis of CSCL chats. *International Journal of Computer-Supported Collaborative Learning, 9*(2), 127–156. https://doi.org/10.1007/s11412-014-9190-y.

(*) Twitchell, D. P., & Nunamaker, J. F. (2004). Speech act profiling: A probabilistic method for analyzing persistent conversations and their participants. In *Proceedings of the 37th annual Hawaii International conference on system sciences*. Los Alamitos, CA: IEEE Press.

(*) van Aalst, J., Chan, C. K. K., Tian, S. W., Teplovs, C., Chan, Y. Y., & Wan, W. S. (2012). The knowledge connections analyzer. In *Proceedings of the 10th International conference of the learning sciences* (pp. 361–365). Sydney, Australia: ISLS.

van de Pol, J., Mercer, N., & Volman, M. (2019). Scaffolding student understanding in small-group work: Students' uptake of teacher support in subsequent small-group interaction. *The Journal of the Learning Sciences, 28*(2), 206–239. https://doi.org/10.1080/10508406.2018.1522258

(*) van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2015). Teacher regulation of cognitive activities during student collaboration: Effects of learning analytics. *Computers and Education, 90*, 80–94. https://doi.org/10.1016/j.compedu.2015.09.006.

van Leeuwen, A., & Rummel, N. (2019). Orchestration tools to support the teacher during student collaboration: A review. *Unterrichtswissenschaft, 47*(2), 143–158. https://doi.org/10.1007/s42010-019-00052-9

(*) van Leeuwen, A., Rummel, N., & van Gog, T. (2019). What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations?. *International Journal of Computer-Supported Collaborative Learning, 14*(3), 261–289. https://doi.org/10.1007/s11412-019-09299-x.

(*) van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014). Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers and Education, 79*, 28–39. https://doi.org/10.1016/j.compedu.2014.07.007.

(*) Viégas, F. B., & Smith, M. (2004). Newsgroup crowds and authorlines: Visualizing the activity of individuals in conversational cyberspaces. In *Proceedings of the37th annual Hawaii International conference on system sciences* (pp. 1–10). IEEE.

Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers and Education, 122*(March), 119–135. https://doi.org/10.1016/j.compedu.2018.03.018

Visual Vocabulary (n.d.). Visual vocabulary: Designing with data. Retrieved July 24, 2020, from http://ft-interactive.github.io/visual-vocabulary/.

(*) Vivian, R., Tarmazdi, H., Falkner, K., Falkner, N., & Szabo, C. (2015). The development of a dashboard tool for visualising online teamwork discussions. In *2015 IEEE/ACM 37th IEEE International conference on software engineering* (Vol. 2, pp. 380–388). IEEE.

(*) Voyiatzaki, E., & Avouris, N. (2014). Support for the teacher in technology-enhanced collaborative classroom. *Education and Information Technologies, 19*(1), 129–154.

Wang, Q. (2009). Design and evaluation of a collaborative learning environment. *Computers & Education, 53*(4), 1138–1146. https://doi.org/10.1016/j.compedu.2009.05.023

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications (Structural analysis in the social sciences* (Vol. 8). Cambridge; New York: Cambridge University Press.

Webb, N. M., Franke, M. L., De, T., Chan, A. G., Freund, D., Shein, P., & Melkonian, D. K. (2009). "Explain to your partner": Teachers' instructional practices and students' dialogue in small groups. *Cambridge Journal of Education, 39*(1), 49–70. https://doi.org/10.1080/03057640802701986

Wise, A. F., & Chiu, M. M. (2011). Analyzing temporal patterns of knowledge construction in a role-based online discussion. *International Journal of Computer-Supported Collaborative Learning, 6*(3), 445–470. https://doi.org/10.1007/s11412-011-9120-1

(*) Wise, A. F., Cui, Y., & Jin, W. Q. (2017). Honing in on social learning networks in MOOC forums: Examining critical network definition decisions. In *Proceedings of the 7th International learning analytics & knowledge conference* (pp. 383–392). New York, NY, USA: ACM.

Wise, A. F., & Schaffer, D. W. (2015). Why theory matters more than ever in the age of big data. *Journal of Learning Analytics, 2*(2), 5–13. https://doi.org/10.18608/jla.2015.22.2

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. https://doi.org/10.1007/s11412-017-9267-5

(*) Xing, W., Wadholm, R., Petakovic, E., & Goggins, S. (2015). Group learning assessment: Developing a theory- informed analytics. *Journal of Educational Technology & Society, 18*(2), 110–128.

(*) Xiong, C., Chen, C., & Jiang, N. (2012). Research on the visualization in group deliberation environment. In *Proceedings of the 2012 7th International conference on computer science and education* (pp. 146–149). ICCSE.

Xiong, R., & Donath, J. (1999). Peoplegarden: Creating data portraits for users. In *Proceedings of the ACM symposium on user interface software and technology* (pp. 37–44). ACM.

(*) Xu, P., Wu, Y., Wei, E., Peng, T. Q., Liu, S., Zhu, J. J., & Qu, H. (2013). Visual analysis of topic competition on social media. *IEEE Transactions on Visualization and Computer Graphics, 19*(12), 2012–2021.

Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of 9- and 10-year-olds. *Educational Technology Research & Development, 55*(2), 117–145. https://doi.org/10.1007/s11423-006-9019-0

(*) Zhang, J., Tao, D., Chen, M. H., Sun, Y., Judson, D., & Naqvi, S. (2018). Co-organizing the collective journey of inquiry with idea thread mapper. *The Journal of the Learning Sciences, 27*(3), 1–41. https://doi.org/10.1080/10508406.2018.1444992.

(*) Zhao, J., Gou, L., Wang, F., & Zhou, M. (2014). Pearl: An interactive visual analytic tool for understanding personal emotion style derived from social media. In *2014 IEEE conference on visual analytics science and technology* (pp. 203–212). IEEE.

(*) Zheng, Y., Xu, C., Li, Y., & Su, Y. (2018). Measuring and visualizing group knowledge elaboration in online collaborative discussions. *Journal of Educational Technology & Society, 21*(1), 91–103.

(*) Zumbach, J., Hillers, A., & Reimann, P. (2004). Supporting distributed problem-based learning: The use of feedback mechanisms in outline learning. In T. S. Roberts (Ed.), *Online collaborative learning: Theory and practice* (pp. 86–102). Hershey, PA: Idea Group Inc.