JOURNAL OF LEARNING ANALYTICS

Volume 12(1), 338–366. https://doi.org/10.18608/jla.2025.8523

Charting the Development of Collaboration Skills through Collaborative Learning Analytics Systems

Xiaomeng Huang[1*] and Xavier Ochoa[2]

Abstract

Collaboration skills are fundamental to effective collaborative learning, career success, and responsible citizenship. Collaborative learning analytics (CLA) systems hold significant potential in helping students develop these skills by automatically collecting group interaction data, analyzing skill levels, and providing actionable feedback so students can reflect, practise, and improve. Previously, most collaborative feedback systems have focused on improving collaborative processes rather than serving as instructional systems for developing collaboration skills over time. To identify what is needed to navigate toward this new type of tool, our paper proposes an interdisciplinary framework that serves as a guiding compass for designing and evaluating such systems. Through an extensive literature review, we evaluate 15 selected systems through the lens of each element of this framework. We map out the current state of the field and identify four major gaps that need to be addressed to transition from systems that support collaboration to systems that support the development of collaboration skills. These gaps are unexplored collaboration skills, lack of validated indicators, limited modelling techniques, and pedagogical feedback design. Finally, we propose a set of corresponding research agendas to bridge these gaps, providing a forward-looking roadmap for designing effective and actionable CLA systems for collaboration skills development.

Notes for Practice

- Most current collaborative learning analytics (CLA) tools focus on providing temporary support for collaborative processes to improve the outcome of collaboration, while few tools focus on developing students' collaboration skills.
- This paper introduces an interdisciplinary framework integrating theory, pedagogy, and analytics for evaluating and designing CLA tools aimed at developing collaboration skills. Using this framework, we conduct a semi-systematic review to examine the current state of CLA tools for enhancing collaborative behaviours. We identify four major gaps: unexplored skills, lack of validation in construct operationalization, limited use of modelling techniques, and insufficient incorporation of pedagogy.
- Based on the identified gaps, this paper proposes four actionable research agendas. These call for more comprehensive, theory-grounded measurement and pedagogically valid feedback design to guide the future development of CLA systems.

Keywords

Collaboration skills, collaborative learning analytics, feedback, multimodal learning analytics.

Submitted: 14/06/2024 — **Accepted:** 20/01/2025 — **Published:** 18/03/2025

[1]*Corresponding author *Email: xiaomeng.huang@nyu.edu Address: Steinhardt School of Culture, Education and Human Development, New York University, 370 Jay Street, Brooklyn, New York, 11201, USA. ORCID iD: https://orcid.org/0000-0002-6992-061X*

[2]*Email: xavier.ochoa@nyu.edu Address: Steinhardt School of Culture, Education and Human Development, New York University, 370 Jay Street, Brooklyn, New York, 11201, USA. ORCID iD: https://orcid.org/0000-0002-4371-7701*

1. Introduction

Collaboration skills are key to academic success, workplace productivity, and responsible citizenship (Lai et al., 2017). Given the critical importance of these skills for students' academic and professional lives, their development should be a priority in educational and professional settings. However, despite their significant value, a "skills gap" exists between the collaboration skills employers expect and those students actually acquire (Singh Dubey et al., 2022; Orona et al., 2023). This gap arises mainly due to the complexities of assessing and training these skills, including selecting specific skills to focus on, determining optimal times for assessment, accurately evaluating these skills, and integrating both assessments and training into the existing curriculum (Liu et al., 2021). Additionally, educators often struggle to gain sufficient insights into all students' collaborative behaviours in classrooms due to time constraints and the difficulty of observing individual dynamics (Worsley & Ochoa, 2020).

In response to these challenges, collaborative learning analytics (CLA) tools, an emerging area in learning analytics, offer a promising approach (Wise et al., 2021). Unlike systems that simply capture interaction data and apply analytics techniques to analyze collaboration, CLA tools not only automate the capture and analysis of data but also provide the results back to students and educators. This feedback enables them to reflect on and improve their behaviours. Additionally, as Wise and colleagues (2021) argue, for CLA tools to be effective, they need to support students and educators in interpreting and reflecting on their actions and serve as temporary scaffolds that fade as learning is internalized by individuals. In the context of our focus—developing collaboration skills in classroom settings (i.e., collocated environments) across educational levels from K–12 to higher education—CLA tools that leverage multimodal learning analytics (MmLA) offer a promising approach. By capturing rich multimodal data from students' collaborative behaviours, these tools can perform a more comprehensive assessment of their skills, given the rich data collected. These tools provide automated feedback to help individuals develop their collaboration skills, thereby relieving educators of some burden and providing timely, accurate, and actionable feedback that is hard to obtain in traditional settings.

Over the past two decades, considerable efforts by the communities of learning analytics, computer-supported collaborative learning (CSCL), and computer-supported cooperative work (CSCW) have aimed to create analytics-supported systems for improving collaboration in both collocated and remote settings. Many of these systems focus on improving collaborative effectiveness by prompting behavioural modifications to enhance collaboration outcomes. For example, some systems support monitoring joint attention and coordinating awareness through gaze sharing (D'Angelo & Gergle, 2016), some visualize speaking time to monitor group members' participation (Ochoa et al., 2023), and some trace and synthesize group ideas (Zhu et al., 2023). These systems are designed for temporary group-level behavioural modifications rather than for developing long-term, individually acquired collaboration skills. However, they share most of the functionality needed for CLA tools that could be used for collaboration skills development. Therefore, building on previous research, we seek to answer the following questions: (1) What key features should CLA tools include to enhance collaboration skills? (2) How do existing tools align with these features? (3) What steps are necessary to begin designing and developing such tools?

To answer these questions, our paper makes three contributions: First, we present an interdisciplinary framework that integrates theories of collaboration skills, learning analytics, and pedagogical feedback design. This framework serves as our compass, offering a systematic method for guiding the design and evaluation of CLA tools tailored for collaboration skills development. Second, applying our framework, we analyze existing systems to explore the current landscape of relevant research. This exploration highlights the achievements of existing tools while also revealing areas that need further exploration. Finally, from this analysis, we propose a research agenda for further CLA tool development. This agenda, serving as our roadmap, provides guidance for researchers and designers to close the current gaps and enhance the effectiveness of CLA tools in improving collaboration skills.

2. The Compass: TAP Framework

To navigate the landscape of CLA tools aimed at developing collaboration skills, we began by reviewing prior frameworks for evaluating these tools from different perspectives. While previous works have been valuable and inspiring, we found that none of them fully serve our specific purpose. Therefore, we propose an interdisciplinary framework to guide us in evaluating existing tools and designing future ones.

2.1 Review of Prior Frameworks

Our review of frameworks starts with a search for the ones that span the key components of these CLA tools: MmLA, feedback, and collaboration. We specifically focused on frameworks that address at least two of these critical aspects to ensure a comprehensive understanding of the tool's design and application. This exploration revealed that existing frameworks predominantly align with three distinct but interconnected strands: MmLA for feedback, MmLA for collaboration analytics, and the integration of MmLA feedback and collaboration. Each strand contributes uniquely to the broader discourse on how CLA tools can support the development of effective collaboration skills.

The first strand of research focuses on the mechanisms of MmLA for feedback. This strand addresses the question of how MmLA can be used to generate feedback for stakeholders. An example of this approach is the dual-phase MmLA process proposed by Ochoa (2022). In this model, the MmLA process is decomposed into two phases: the mapping phase and the execution phase. The mapping phase begins by selecting learning constructs, which are then operationalized into two components: behavioural markers and analytics. In other words, this phase determines how the constructs are quantified through observable behavioural indicators and identifies the sources for data collection. The execution phase reverses the steps of the mapping phase. It starts with multimodal recording to capture raw data, followed by feature extraction and analysis to detect behaviours and estimate constructs. Finally, feedback is generated for the participants. Similarly, Di Mitri and colleagues

(2018) introduced the MLeAM model, which distinguishes the processes occurring within the physical and digital spaces. They also identified three mediating models: the learner model, the learning task model, and the feedback model. The MLeAM model emphasizes the need for a more comprehensive understanding of the learner and the context. These frameworks propose systematic approaches to transforming multimodal data into actionable feedback across different learning contexts.

The second strand of research focuses on applying MmLA in collaboration, ranging from analyzing to improving collaboration processes. For example, Schneider and colleagues (2021) propose a framework for connecting metrics of multimodal data traces to collaboration outcomes. Moving beyond analysis, Boothe and colleagues (2022) propose an overarching framework for automated collaboration feedback, integrating theory, technology, and pedagogy. This framework emphasizes creating analytics that not only assess but also foster effective collaboration and learning outcomes. Recognizing the importance of context and stakeholders, Echeverria Barzola (2020) adopted a human-centred design (HCD) and an inductive, theory-driven methodology to merge empirical insights from students and teachers with theoretical constructs. These frameworks illustrate the applications of MmLA in collaboration, from analyzing collaboration processes for research purposes to generating feedback that can be more effective for students and teachers.

The third strand of research focuses on using automated feedback for collaboration. The main question is how automated feedback can be used to enhance collaboration. For example, Soller and colleagues (2005) proposed the collaboration management cycle framework, which includes four phases: collecting interaction data, constructing a computational model, enabling participants to compare the current collaboration state with a desired one, and offering advice to bridge the gaps. Within this framework, the authors proposed three types of tools: mirroring tools, which represent interaction data as it is; meta-cognitive tools, which visualize the gaps between the current state and the desired state for participants to reflect on; and guiding systems, which offer actionable feedback for participants to act on. This framework provides an initial step toward understanding how automated feedback systems can be useful in collaboration.

Our review of these frameworks reveals three common themes. First, the emphasis on a theory-driven process in the deployment of MmLA is evident across the studies by Ochoa (2022), Boothe and colleagues (2022), and Schneider and colleagues (2021). These frameworks underscore the importance of a theory-driven approach, advocating for data analytics to be informed by underlying theoretical constructs. Second, the incorporation of pedagogical principles and contextual factors, as highlighted by both Boothe and colleagues (2022) and Echeverria Barzola (2020), reflects a consensus on the necessity of integrating cognitive, metacognitive, and affective considerations into system design to facilitate meaningful behavioural changes in learners. Lastly, the complexity of feedback, as categorized by Soller and colleagues (2005) and Di Mitri and colleagues (2018), reveals varying levels of support. This ranges from increasing students' awareness of their collaborative behaviours to offering guidance that mediates collaborative processes, taking into account the mediating factors of learner, task, and context. Despite these valuable insights, we found that these frameworks are not specifically designed for evaluating and developing CLA tools aimed at enhancing collaboration skills. First, as these frameworks are not focused on skills development, they do not include the necessary pedagogical considerations, such as assessment, training, and adaptive feedback over time to capture learners' progress. Second, while feedback mechanisms are frequently discussed, there is a lack of detailed exploration into pedagogical feedback design that aligns with current research in this area. Third, the reciprocal relationship between theory and empirical insights, where theory both guides the analysis and is enriched by empirical data (Wise & Shaffer, 2015; Giannakos & Cukurova, 2023), is not captured by these frameworks. To bridge these gaps, in the next section, we propose the TAP framework—an interdisciplinary framework for evaluating and designing automated feedback systems for collaboration skills development.

2.2 TAP Framework: An Interdisciplinary Approach

To bridge the gaps in existing frameworks, we propose the TAP (theory, analytics, and pedagogy) framework, illustrated in Figure 1. These three elements have been recognized as important components in previous frameworks: theory guides the operationalization of constructs, analytics generates insights, and pedagogy drives feedback design. By integrating these elements, we aim to cover all phases of CLA tool design, with a special focus on collaboration skills development. Specifically, expanding upon the dual-phase processes proposed by Ochoa (2022) and Boothe and colleagues (2022), our framework integrates the operationalization of constructs and the analytics execution phase with principles of pedagogical feedback design. It emphasizes the importance of dimensioning students' progress to provide adaptive and meaningful feedback over time. Furthermore, the framework proposes a reciprocal process, ensuring that it is informed by theory while simultaneously contributing to theoretical developments through empirical insights. Acting as a compass, the TAP framework offers a comprehensive approach for evaluating current systems and guiding the design of future tools. The following sections will detail the three dimensions of our framework: theory, analytics, and pedagogy. Each dimension is tailored to equip researchers and system designers with a structured approach for evaluating or designing such systems.

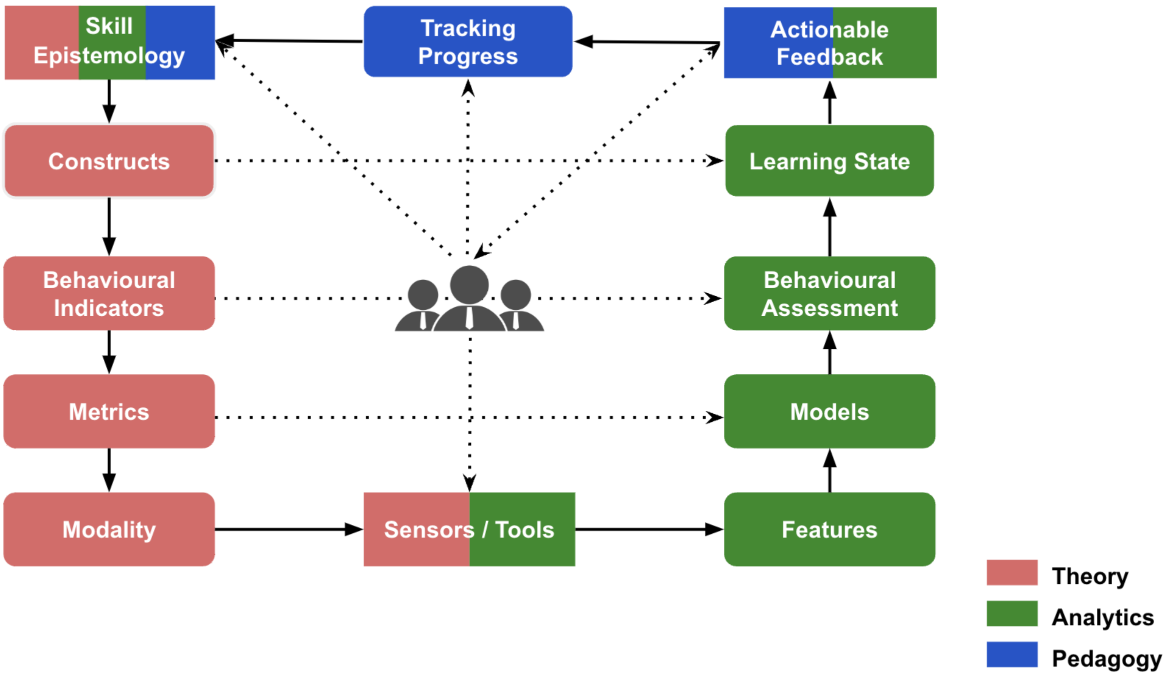

Figure 1. TAP framework: evaluating and designing CLA tools for effective collaboration skills development.

2.2.1 Skill Epistemology

To develop collaboration skills with CLA tools, skill epistemology serves as both the starting and ending point of the framework. Skill epistemology addresses fundamental questions, such as: What are collaboration skills? How can we measure and teach them? The answers to these questions inform the theoretical operationalization of the skills, the methods for measuring skill levels, and the design of feedback to support student learning. The framework begins by considering the theories of collaboration and collaboration skills, the pedagogy for teaching these skills, and the specific context of interest. Decisions about which skills to teach directly influence the operationalization of constructs, while the chosen pedagogy shapes the design of feedback. The framework ends with skill epistemology as well; the insights gained from the entire cycle—implementing automated feedback to support students' collaboration skills—will expand our understanding of collaboration skills and improve how we teach them.

2.2.2 Theory Dimension

Once the skill epistemology has been selected, the next step involves identifying relevant constructs. These constructs are abstract concepts that manifest as observable learning behaviours (Ochoa, 2022) and must be relevant to the intended application context. The selection of constructs should consider the learners, the tasks they will engage in, and the specific contexts in which these tasks occur. This decision-making process involves understanding whom the tool will benefit, the nature of the tasks it will support, and the environment in which it will be used. Once constructs are selected, a literature review is essential to decide what behavioural indicators can reflect these constructs and how to quantify those indicators. Based on the identified behavioural indicators and metrics, one can specify the modalities of the required data traces. This decision guides the selection or development of the appropriate sensors or tools needed to capture the data traces.

For example, one could start by pinpointing the key sub-skills outlined in existing collaboration skills research, such as the knowledge, skills, and ability (KSA) for teamwork proposed by Stevens and Campion (1994) or the Pearson's collaboration skills framework (Lai et al., 2017). Based on these inventories of skills, the selection of sub-skills should align with the context of the intended application. For example, in classroom discussions where the issue of some students being overlooked by their peers is common, active listening might emerge as a critical construct in communication skills. To operationalize active listening, conducting a detailed literature review is necessary. For instance, the active listening observation scale (ALOS) (Worthington & Bodie, 2017), with its robust indicators covering verbal and nonverbal behaviours and general behavioural perceptions, could serve as an effective operationalization tool. This scale helps identify essential data traces—like verbal communication, eye gaze, body posture, and facial expressions—indicating that audio and video modalities are necessary. Equipped with this insight, the next step involves selecting or designing the appropriate sensors or systems to capture these modalities and then moving to the analytics execution phase. This approach ensures a grounded, context-aware, and theory-informed pathway from theory phase to analytics phase. By prioritizing construct validity (Cronbach & Meehl, 1955), our framework ensures that constructs are operationalized based on established theories and validated methodologies. This process ensures that our approach is grounded in rigorous research, rather than relying on readily available data.

2.2.3 Analytics Dimension

Transitioning to the analytics dimension, our process starts with the sensors or tools we designed or selected. Using these sensors, we collect multimodal data relevant to our constructs of interest. Next, we extract features from this data, guided by the behavioural indicators identified during the theory phase. With these features, we construct computational models informed by the metrics established in the theory dimension. Recognizing the limitations of relying solely on unimodal data (Worsley & Ochoa, 2020; Ochoa, 2022), we emphasize data fusion from multiple sources. This integration provides a comprehensive view of a learner's skill level.

For instance, consider developing active listening skills. Our chosen sensors begin by capturing verbal communication from discussions and nonverbal data like eye movements and body posture. We focus on eye gaze directions to infer where a learner is looking, and we analyze head and body positions to identify engagement behaviours like nodding or leaning forward. These features directly correspond to "expressing understanding non-verbally," a behavioural indicator of active listening identified in ALOS (Bodie et al., 2015). Using these extracted features, we build computational models to determine if a learner exhibits non-verbal understanding. Each feature contributes to assessing active listening facets. However, a comprehensive evaluation of active listening cannot rely on a single indicator. Therefore, by merging insights from various behaviours and data sources, we gain a more comprehensive understanding of learners' active listening skills. This structured approach, from theoretical underpinnings to practical application, ensures construct validity and addresses the complex nature of collaboration skills. Equipped with these comprehensive analyses, we are well prepared to provide learners with targeted and actionable feedback to help them develop collaboration skills.

2.2.4 Pedagogy Dimension

To ensure that feedback is truly actionable, we strongly emphasize the role of pedagogy within our framework. The pedagogy dimension is structured around three key elements: actionable feedback, tracking progress, and centring students. These elements are visually represented as an ongoing and iterative feedback loop in our framework (see Figure 1), which highlights the dynamic and continuous nature of the learning process. Actionable feedback, the cornerstone of our pedagogy dimension, is multifaceted. Traditionally, feedback has been viewed as a one-way flow of information from teachers to students (Hattie & Timperley, 2007). However, constructivist perspectives have redefined feedback as a dialogic process, where students are active participants, using feedback to enhance their learning (Lipnevich et al., 2016; Lim et al., 2019). Our framework synthesizes these perspectives by first focusing on content-centred feedback, using classifications from Narciss and Huth (2004) and Panadero and Lipnevich (2022) that classify feedback design into its content, function, and presentation. We then extend our focus to the students' engagement with the feedback—how they cognitively and emotionally interact with and apply the feedback to their learning processes. Therefore, although represented as a single block in the visual model, actionable feedback includes two interrelated layers: the structural components of the feedback (content, function, and presentation) and the interactive processes involving student engagement and application. This dual-layer approach ensures that feedback is truly actionable for students. Following the actionable feedback, our framework introduces the last piece of the puzzle: the process of monitoring progress over time. By tracking the progress, our framework incorporates a longitudinal perspective to monitor and support the ongoing development of collaboration skills.

For example, in enhancing active listening skills, if the analysis indicates that a student seldom responds to their peers, it is crucial to design feedback that considers its content, function, and presentation. First, the content of the feedback should go beyond simply setting clear goals like "respond more to others." It should specifically identify current performance gaps—such as a lack of verbal and non-verbal acknowledgements during discussions—and offer targeted strategies for improvement. For instance, the feedback could illustrate instances where the student failed to nod or make eye contact while others spoke and suggest specific adjustments, such as "Try nodding in agreement or offering verbal affirmations like 'I see' or 'That makes sense' to show you are engaged." In terms of function, the feedback should prioritize the learning process over merely completing tasks. It should guide students on how to interact effectively with others, encouraging them to self-monitor and regulate their behaviours. This approach helps move away from simplistic labels like "good" or "bad" listener to more nuanced self-assessments, fostering a deeper understanding of active listening as a skill. For the presentation of feedback, it is important to customize the timing and modality to fit both individual and group needs. Utilizing the capabilities of an automated feedback system, students can access multimodal feedback, such as audio and video recordings of key moments during collaboration. This allows them to see concrete examples of their behaviours and directly links feedback to their actions. Moreover, integrating students' engagement with feedback is essential. Understanding students' prior knowledge about active listening, their motivation to improve, and their willingness to engage with feedback is also important for designing effective interventions. Finally, the system needs to monitor students' progress to ensure that feedback is adaptive to their evolving learning state. When students demonstrate improvement in providing feedback to their peers, the intensity and frequency of the feedback can be gradually reduced. This fading of guidance encourages students to rely more on their internal judgments and skills, thus supporting the transition from externally guided practice to self-regulated behaviour. The entire cycle—from selecting relevant theories and employing analytics to generating actionable feedback on active listening skills—should contribute back to enhancing the skill epistemology itself. These practices could deepen our understanding of how active listening skills can be developed effectively and refine the pedagogical practices used to teach these skills.

2.3 Conclusion

In summary, our interdisciplinary TAP framework, consisting of the theory, analytics, and pedagogy dimensions, offers a comprehensive approach to designing and evaluating automated feedback systems for collaboration skills development. It aims to establish standards and guidelines for developing feedback systems that are deeply rooted in theory, centred on student engagement, grounded in robust analytics, and informed by sound pedagogical principles. Using this framework as a "compass," the following section will examine existing feedback systems in collocated collaboration environments. This exploration aims to assess the current state of the field, identify gaps and areas for improvement, and propose directions for future research and system development based on the guidelines and insights provided by our framework. By applying our framework to review existing systems, we hope to chart a path forward, highlighting how future feedback systems can more effectively support the development of collaboration skills.

3. Current Landscape Exploration: Review of Existing Systems

In this section, we explore the current landscape of research focusing on the development of collaboration skills among students through analytics-supported feedback systems. By adopting a semi-systematic review methodology, we identified and analyzed 15 systems from the past two decades. We detail the data source, analysis procedures, and gaps identified in the following subsections.

3.1 Data Source

To better understand the collaboration skills or behaviours targeted by each system, we reviewed definitions of collaboration skills from several well-regarded frameworks across different fields, including management studies (Stevens & Campion, 1994), CSCL (Järvelä & Hadwin, 2013), and educational measurement (OECD, 2017; Ofstedal & Dahlberg, 2009). Unlike collaborative problem-solving (CPS) skills (OECD, 2017), which primarily focus on cognitive skills within CPS contexts, collaboration skills encompass a broader spectrum. These skills include both cognitive and social-emotional abilities critical for effective interaction in various academic and workplace settings, such as conflict resolution, communication, negotiation, and relationship building (Stevens & Campion, 1994; Ofstedal & Dahlberg, 2009; Janssen & Bodemer, 2013; Lyndgaard & Kanfer, 2024). Our aim was to identify the specific skills needed for effective collaboration. By comparing the definitions of skills listed in these frameworks, we found considerable overlap across these fields. Therefore, we categorized these skills into two types: intrapersonal skills and interpersonal skills. This classification serves as a reference to identify which collaboration skills existing systems focus on and which areas need further exploration. While not comprehensive, it offers guidance in identifying gaps in the development of collaboration skills through these systems.

- **Intrapersonal skills:** These include skills like setting personal goals, meeting role expectations, and monitoring one's own progress and contributions.
- **Interpersonal skills:** This category refers to the skills necessary for effective collaboration, including setting group goals, co-planning and coordination, monitoring group progress, group evaluation, reflection and adaptation of group strategies, building relationships, interacting with others, negotiation and argumentation, and resolving conflicts.

Over the past two decades, interdisciplinary efforts within learning analytics, CSCL, and CSCW have led to the development of tools aimed at supporting collaboration. While the importance of collaboration skills is well recognized, our review indicates that direct targeting of these skills by existing systems remains rare. Most studies focus on enhancing collaboration by providing tools that facilitate shared information, such as concept maps or gaze sharing, rather than directly developing collaboration skills. Systems that attempt to modify collaborative behaviours typically offer temporary support rather than aiming for long-term skill development. Despite this, we have included these systems in our review due to their alignment with our objectives and their potential for skill development support. With appropriate modifications, these systems could be adapted to more effectively focus on long-term collaboration skill enhancement.

Due to the varied interpretations and conceptualizations of these systems—ranging from tools designed to improve awareness to those fostering participation, rapport, and contribution—we have opted for a semi-systematic review as our methodological approach. This choice aligns with the guidance from Snyder (2019), who notes that semi-systematic reviews are suitable for topics conceptualized differently across disciplines. Typically, this approach employs qualitative methods to identify, analyze, and synthesize emergent themes across various research traditions. It is often used to provide a comprehensive overview of the current state of the field and to outline a future research agenda. To conduct our semi-systematic review, we selected three systematic reviews that align with the three principal dimensions of our framework: the operationalization of theory, the implementation of analytics, and the design of analytics-supported feedback, all within the context of collaboration settings. Published within the past five years and adhering to PRISMA guidelines, these reviews provide comprehensive examinations of systems that serve as valuable starting points, covering a broad spectrum of systems relevant to our interests. The selected reviews are

- **Praharaj and colleagues (2021)**, which focuses on modelling collocated collaboration through MmLA, emphasizing behavioural indicators and collaboration indexes such as equality and synchrony, and offering insights into the dynamics of face-to-face collaborative interactions;
- **Schneider and colleagues (2021)**, which examines the empirical links between multimodal metrics and collaborative constructs, shedding light on the quantitative measures that can inform collaboration quality and effectiveness; and
- **Hu and Chen (2021)**, which explores the use of visual representations in analyzing collaborative discourse, highlighting how visual feedback can support understanding and improvement in collaborative settings.

After thoroughly reviewing all studies included in these three systematic reviews, we applied our inclusion and exclusion criteria to identify the most relevant studies for our topic. Additionally, we expanded our search through the reference lists of the initially identified studies. The search and filtering process is detailed in the following section. To identify systems designed to support collaboration skills development or regulate collaborative behaviours, we established the following selection criteria:

1. **Collocated environments.** We focused on studies that involved face-to-face, physical space collaboration, excluding those on remote collaboration. This decision was made not only because the data sources collected in collocated environments are multimodal in nature and usually offer a richer set of features, but also because these environments remain the primary learning settings in schools.
2. **Collaboration for learning.** We excluded studies that focused on individual learning instead of collaborative learning.
3. **Student-centric systems.** Our interest was in tools that directly enhance or regulate students' collaboration skills, not in tools mainly for teachers to orchestrate classroom activities.
4. **Feedback systems.** We looked for systems providing students with direct feedback on their collaboration, not just those advancing theoretical knowledge about collaboration or creating metrics.

We employed a binary coding scheme (0 for criteria not met, 1 for criteria met) to systematically evaluate each study. Studies received scores from 0 to 4 based on how many of our criteria they satisfied, with only those achieving a score of 4 included in our final analysis. Starting with 295 papers cited in the three systematic reviews, we first removed 32 duplicates, leaving us with 263 unique studies for initial evaluation based on titles and abstracts. During this initial screening, we found that some studies met more than one exclusion criterion. Specifically, 110 papers were primarily focused on remote collaboration settings and thus excluded; 36 studies did not directly address collaborative learning; 45 studies were focused on systems designed not for student use; and 204 studies did not include direct feedback to students on their collaborative behaviours. After adjusting for overlaps—where a single study could be excluded for multiple reasons—we were left with 18 papers that initially seemed to meet all four criteria. The coding process followed an iterative approach, with regular meetings between the two authors to refine criteria and develop coding strategies. One author served as the primary coder, conducting the first round of coding across all papers and taking detailed notes to inform discussions during the meetings. The second author independently reviewed the coding results to ensure accuracy and consistency. Any discrepancies or ambiguities were discussed and resolved collaboratively during these meetings until consensus was reached. While inter-rater reliability was not formally calculated, the collaborative and iterative nature of this process ensured the rigour and validity of the coding. A more detailed review of these 18 full-text papers resulted in the exclusion of an additional five papers that, upon closer examination, were used for neither supporting collaboration skills development nor regulating collaboration behaviours. Moreover, by exploring the reference lists of the selected papers, we identified two additional studies relevant to our criteria. This selection process is illustrated in Figure 2. As a result, we identified 15 collaboration systems, as summarized in Table 1. The system is sorted by the last names of the lead authors, in alphabetical order.

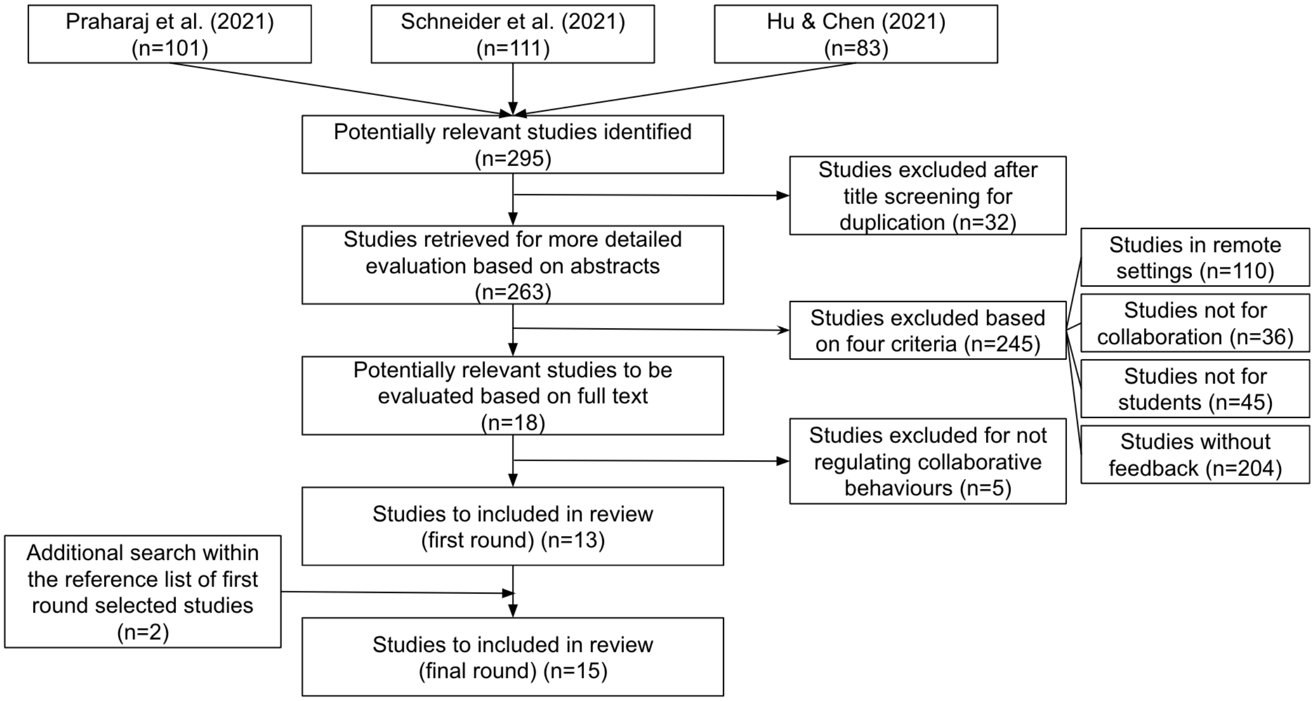

Figure 2. Flow diagram of study selection process.

In the next section, we use our proposed framework to analyze these systems. Our key questions are as follows: What has been achieved so far in collocated CLA tools for developing collaboration skills or supporting collaborative behaviours? What existing achievements can we build upon? What should be done next, and how should it be done?

3.2 Data Analysis

To comprehensively understand the current landscape of tools supporting collaboration skills development, we applied our integrated framework to each identified system. Our analysis focused on 18 dimensions derived from the 12 components of our framework. For the "actionable feedback" component, as detailed in Section 2.2.4, we further unpacked it into six specific dimensions: feedback content, feedback function, feedback presentation, cognitive engagement, emotional engagement, and applying feedback. We detail our criteria in Table 2. Each system was coded across these 18 dimensions, and the initial findings were documented in a spreadsheet, with each column corresponding to one of the aspects listed above. This structured approach allowed us to systematically evaluate how each system aligns with our framework and to identify areas of strength and gaps in the use of theory, analytics operations, and pedagogical feedback design in the context of collaboration skills development.

Building upon these analyses, through the lens of our framework, we have identified 14 trends, illustrated in Figure 3, and summarized them into four major gaps in the next sections. Our intention is not to critique any individual research, as each study made valuable contributions within its scope, but to illuminate opportunities for enhancing the design and implementation of an analytics-supported feedback system to develop collaboration skills.

3.3 Gap 1: Unexplored and Insufficiently Supported Collaboration Skills

The first gap we identified is in the collaboration skills targeted by the reviewed systems. While most systems aim to support collaborative behaviours in specific contexts, only two studies explicitly define the skills they intend to improve. This reflects a broader trend of focusing on immediate behavioural modification rather than long-term skill development. We evaluated these systems by comparing with the collaboration skills inventory we previously identified, and we summarized the findings in Table 3 for intrapersonal skills and Table 4 for interpersonal skills. In these tables, we mark the addressed skills and leave blank those that are as yet unexplored. Our analysis reveals the following trends:

Table 1. Overview of systems reviewed.

No.	Authors	System
1	Bachour et al. (2010)	*Reflect:* An interactive tabletop that displays individuals' speaking time in real time
2	Balaam et al. (2011)	An ambient display that visualizes non-verbal behaviour synchrony in real time
3	Bergstrom and Karahalios (2007)	*Conversation Clock:* Projected concentric rings that visualize individuals' speaking time, volume, and simultaneous contribution in real time
4	DiMicco et al. (2004)	*Second Messenger:* A shared display visualizing the amount of participants' speaking time
5	Echeverria et al. (2017)	*DBCollab:* A collaborative database design tool with a shared interactive tabletop, displaying a dashboard of individual and group speaking time and touch input after a collaborative design activity
6	Karahalios and Bergstrom (2009)	*Conversation Vote:* A tabletop that visualizes participants' contribution and votes for ideas
7	Karahalios and Bergstrom (2009)	*Conversation Cluster:* A tabletop that visualizes the classified content created within the group over time
8	Kim et al. (2008)	*Meeting Mediator:* A portable system that visualizes speaking time and interactivity in the group in real time
9	Koh et al. (2020)	A web-based tool showing a radar chart of self and peer ratings, with prompts to help students make personal collaboration goals, follow steps, and monitor progress
10	Kulyk et al. (2005) Terken and Sturm (2010)	A shared display that visualizes individual's speaking time, duration of current turn, and visual attention from speakers and listeners in real time
11	Praharaj et al. (2018)	A shared display that visualizes individuals' speaking time and turn taking in real time
12	Starr et al. (2018)	A shared display that visualizes individuals' speaking time
13	Tausch et al. (2014)	*Group Garden:* A shared display that visualizes the number of ideas that each participant contributed during brainstorming sessions
14	Tausch et al. (2016)	A shared display that visualizes the number of ideas that each participant contributed and group contribution during brainstorming sessions
15	Wong-Villacrés et al. (2016)	A report showing students' self-assessment and observers' and peers' rating in contribution, communication, and respect

Table 2. Analysis dimensions and criteria.

No.	Dimensions	Criteria
1	Skill epistemology	The collaboration skills or collaborative behaviours targeted
2	Constructs	The constructs being operationalized
3	Behavioural indicators	Observable behaviours indicating collaboration skills
4	Metrics	Quantitative measures used to assess these indicators
5	Modality	Modalities of data (e.g., audio, video) captured
6	Sensors/tools	Specific systems used to capture these modalities
7	Features	Data features extracted from the collected modalities
8	Models	Computational models or algorithms applied to the features for analysis
9	Behavioural assessment	Evaluation of collaboration skills based on the outcomes
10	Learning state estimation	Inference of a learner's skill level from behavioural assessments
11	Feedback content	Information provided to learners about their collaboration
12	Feedback function	The function of feedback provided (e.g., feedback on task or process)
13	Feedback presentation	How feedback is delivered (e.g., timing, schedule, visualization)
14	Cognitive engagement	How the system facilitates learners' understanding of the feedback
15	Emotional engagement	How the system facilitates learners' emotional responses when receiving the feedback
16	Apply feedback	How learners use feedback to improve collaboration skills or collaborative behaviours
17	Tracking progress	Whether the system monitor learners' development over time
18	Inform theory	Whether the empirical findings contribute back to theories

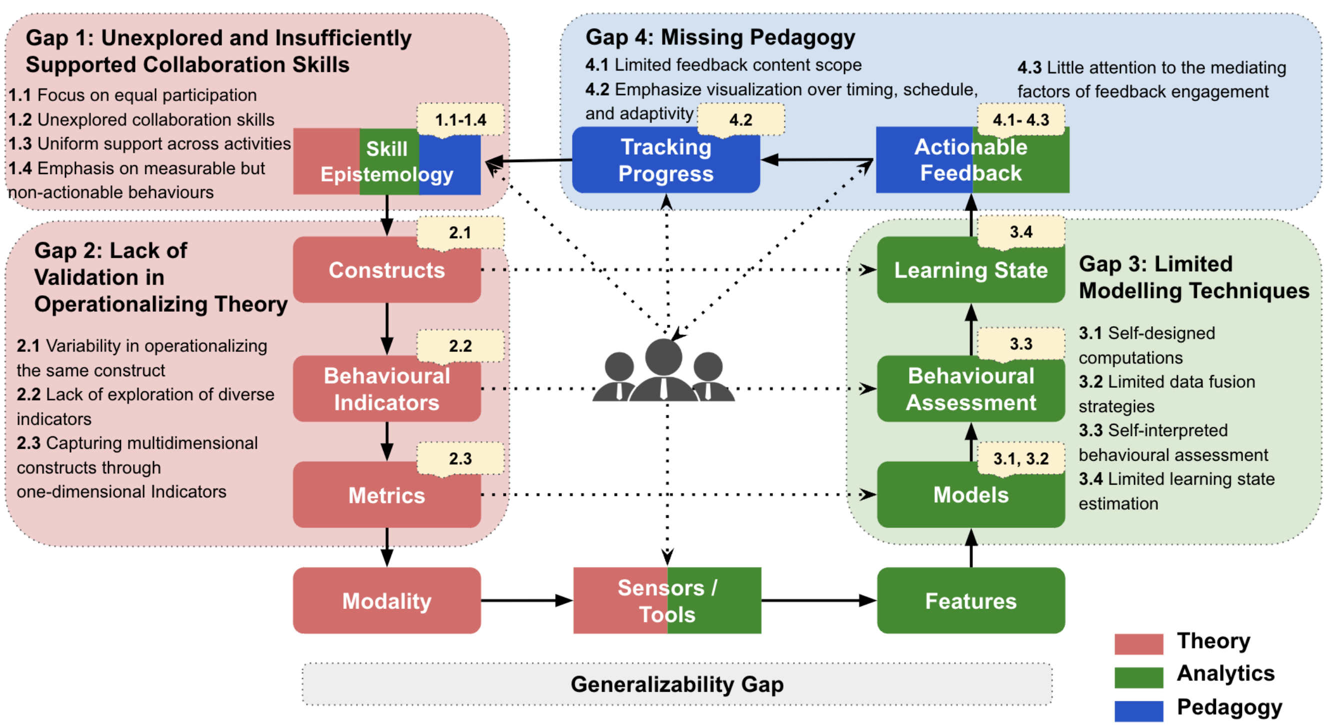

Figure 3. Overview of the research gaps identified in current CLA tools for supporting collaboration skills development.

Table 3. Intrapersonal skills targeted by the systems reviewed.

No.	Papers	Construct	Individual Goal Setting	Filling Role Expectation	Monitor Individual Progress and Contribution
1	Bachour et al. (2010)	Equal participation			**X**
2	Balaam et al. (2011)	Rapport			
3	Bergstrom and Karahalios (2007)	Equal participation			**X**
4	DiMicco et al. (2004)	Equal participation			**X**
5	Echeverria et al. (2017)	Equal participation			**X**
6	Karahalios and Bergstrom (2009)—Conversation vote	Equal participation and contribution			**X**
7	Karahalios and Bergstrom (2009)—Conversation cluster	Equal participation and contribution			**X**
8	Kim et al. (2008)	Interactivity, dominance			**X**
9	Koh et al. (2020)	Coordination, mutual performance monitoring, constructive conflict, and team emotional support			
10	Kulyk et al. (2005) and Terken and Sturm (2010)	Equal participation and equal attention			**X**
11	Praharaj et al. (2018)	Equal participation			**X**
12	Starr et al. (2018)	Equal participation			**X**
13	Tausch et al. (2014)	Equal contribution			**X**
14	Tausch et al. (2016)	Individual contribution, group progress, and equal contribution			**X**
15	Wong-Villacrés et al. (2016)	Contribution, communication, and respect			**X**

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License (CC BY 4.0)

Table 4. Interpersonal skills targeted by the systems reviewed

No.	Papers	Construct	Group Goal Setting	Co-Planning and Coordination	Monitor Group Progress	Group Evaluation and Adaptation of Strategies	Relationship Building	Interaction with Others	Negotiation and Argumentation	Conflict Resolution
1	Bachour et al. (2010)	Equal participation			X	X				
2	Balaam et al. (2011)	Rapport						X		
3	Bergstrom and Karahalios (2007)	Equal participation			X	X				
4	DiMicco et al. (2004)	Equal participation			X	X				
5	Echeverria et al. (2017)	Equal participation			X	X				
6	Karahalios and Bergstrom (2009)—Conversation vote	Equal participation and contribution			X	X				
7	Karahalios and Bergstrom (2009)—Conversation cluster	Equal participation and contribution			X	X				
8	Kim et al. (2008)	Interactivity and dominance						X		
9	Koh et al. (2020)	Coordination, mutual performance monitoring, constructive conflict, and team emotional support		X	X		X	X		X
10	Kulyk et al. (2005) and Terken and Sturm (2010)	Equal participation and equal attention			X	X		X		
11	Praharaj et al. (2018)	Equal participation			X	X				
12	Starr et al. (2018)	Equal participation				X				
13	Tausch et al. (2014)	Equal participation			X					
14	Tausch et al. (2016)	Individual contribution, group progress, and equal contribution			X					

Continued on next page

No.	Papers	Construct	Group Goal Setting	Co-Planning and Coordination	Monitor Group Progress	Group Evaluation, Reflection, and Adaptation of Group Strategies	Relationship Building	Interaction with Others	Negotiation and Argumentation	Conflict Resolution
15	Wong-Villacrés et al. (2016)	Contribution, communication, and respect			**X**		**X**			

1. **Focus on equal participation.** First, the systems we reviewed have a significant focus on equal participation. Out of the 15 systems analyzed, 12 primarily focus on equal participation as a key construct. We marked this construct as the skill "monitoring individual and group progress" in Table 4. However, we acknowledge that the skill of monitoring individual and group progress is multifaceted and can be operationalized in various ways. Equal participation is only one of the relevant constructs.

2. **Unexplored collaboration skills.** Second, we noticed that several key skills have not been targeted. Specifically, skills such as individual goal setting and role expectations in the intrapersonal category, as well as group goal setting, negotiation, argumentation, and conflict resolution in the interpersonal category, are not being supported by the systems we reviewed.

3. **Uniform support across activities.** Third, the support provided by these systems across different collaborative processes appears consistent, regardless of the team's specific activity—whether familiarizing with the task, understanding the task, or discussing solutions. This uniformity suggests a lack of adaptation in the tools to the dynamic nature of group activities.

4. **Emphasis on measurable but non-actionable behaviours.** Lastly, some systems emphasize measuring behaviours that, while quantifiable, do not necessarily support effective collaboration directly. For example, non-verbal synchrony can be measured and is an indicator of collaboration quality (Balaam et al., 2011), but it is a byproduct of effective collaboration rather than a behaviour that can be directly supported through external interventions.

The underlying reasons for these trends are likely rooted in different research paradigms and conceptualizations of collaboration. Researchers who view collaboration primarily as a method for learning often focus on quantifiable data, such as speaking time or speech volume, as proxies for participation. They believe that joining the discussion and speaking for a fair amount of time is a prerequisite for effective learning. However, if the goal is to develop collaboration skills, the focus needs to shift to the skills themselves and the temporal dimension. This includes identifying which collaboration skills are targeted, understanding how different skills are used in various phases of collaboration, and determining how these skills can be developed over time. These considerations are crucial for providing pedagogical feedback that helps students develop and refine their skills. The different goals and conceptualizations of collaboration contribute to the observed gap. For CLA tools aimed at developing collaboration skills, it is essential to set clearer goals in system design that align with the actual development of particular skills.

3.4 Gap 2: Lack of Validation in Operationalizing Theory

Moving on to the later part of the theory dimension of our framework, the second gap we identified involves operationalizing the selected construct. This involves examining the behavioural indicators used as proxies to measure that construct. The selection of behavioural indicators is important as they determine what is measured, how well the construct is assessed, and how potentially accurate and useful the feedback could be. We summarized the behavioural indicators used in these studies in Table 5 and identified three major trends:

1. **Variability in operationalizing the same construct.** We observed that the same construct has been operationalized in varied ways. For example, equal participation has been quantified using seven different behavioural indicators: speaking time, speaking volume, overlapping speech, turn-taking, silence, amount of gesturing, and self- and peer-perceived dominance. Among all the indicators, speaking time appears 10 times, while all other indicators appear twice or once.

2. **Lack of exploration of diverse indicators.** There is a noticeable disparity in the number of indicators used across different constructs, which indicates a significant disparity in the research focus. Among these, "equal participation" is the most extensively studied, with multiple indicators dedicated to its assessment. In contrast, other important constructs—such as respect, coordination, mutual performance monitoring, constructive conflict, and team emotion support—are often evaluated using only a single indicator, typically self- or peer-reported surveys. This imbalance reflects a broader research gap: while equal participation receives considerable attention, other critical constructs remain underexplored.

3. **Capturing multi-dimensional constructs through one-dimensional indicators.** Some indicators do not fully capture their intended multifaceted constructs. For example, inclusion in collaboration is a complex construct involving behaviours such as active listening, acknowledging others, and valuing diverse perspectives. In the study by Terken and Sturm (2010), visualizations of speakers' and listeners' eye gaze were used as indicators of inclusion and rapport. Although eye gaze might suggest attention, it alone is insufficient to quantify inclusion.

Table 5. Summary of constructs and behavioural indicators.

Constructs	Behavioural Indicators	Number of studies
Equal participation	Speaking time	10
	Speaking volume	2
	Overlapping speech	2
	Turn-taking	2
	Silence	1
	Amount of gesturing	1
	Self- and peer-perceived dominance	1
Contribution	Ideas proposed	3
	Voting	1
	Ideas alignment	1
	Artifacts created	1
Rapport	Eye contact	2
	Simultaneous movement	1
	Tempo similarity	1
	Posture matching	1
	Back channel responses	1
Communication	Self and peer rating	1
	Observer rating	1
Respect	Self and peer rating	1
	Observer rating	1
Coordination	Self and peer rating	1
Mutual performance monitoring	Self and peer rating	1
Constructive conflict	Self and peer rating	1
Team emotion support	Self and peer rating	1

The limited use of theory is not uncommon in MmLA research studies. For example, Schneider and colleagues (2021) and Giannakos and Cukurova (2023) observed that many studies in multimodal collaboration do not explicitly mention the theoretical frameworks employed. The reasons for the absence of guiding theory likely stem from a bottom-up inductive approach. While starting with data can generate interesting insights, it also runs the risk of the "streetlight effect" (Ochoa, 2022). To avoid this issue, we recommend adopting more theory-driven and validated construct operationalization.

3.5 Gap 3: Limited Modelling Techniques

Within the analytics dimension of our framework, we focused on examining the calculation of behavioural indicators, the extraction of features from data, the models employed, and the behavioural assessments given. For cases where no computational model is used, we categorized them into three types: arithmetic calculations, which involve simple addition or subtraction; manual coding, where human coders tag actions; and raw data, which involves presenting unprocessed data directly to participants. We summarized the computational strategies used for each construct in Table 6 and the behavioural assessment in Table 7. From this analysis, we identified four major trends:

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License (CC BY 4.0)

Table 6. Summary of metrics and models.

Constructs	Behavioural Indicators	Metrics	Features	Models	Studies
Equal participation	Speaking time	The difference between perfectly balanced participation (i.e., 25% for a four-person group); average segment length	Duration of speaking time	Arithmetic calculation	Bachour et al. (2010), DiMicco et al. (2004), Echeverria et al. (2017), Karahalios and Bergstrom (2009), Kim et al. (2008), Kulyk et al. (2005), Terken and Sturm (2010), Praharaj et al. (2018), and Starr et al. (2018)
	Speaking volume	Pass a certain threshold; variation in speech volume	Speech volume	Arithmetic calculation	Bergstrom and Karahalios (2007) and Kim et al. (2008)
	Overlap speech	Pass a certain threshold; length of simultaneous speaking divided by total length of discussion	Duration of simultaneous speaking	Arithmetic calculation	Bergstrom and Karahalios (2007) and Kim et al. (2008)
	Turn taking	Total number of turns per individual	Each instance a participant takes over from either conversation or silence	Raw data	Kim et al. (2008) and Praharaj et al. (2018)
	Silence	Duration of silence	Duration of silence	Raw data	Bergstrom and Karahalios (2007)
	Amount of gesturing	Mean and variance of individual's movement	Body movement	Unknown	Kim et al. (2008)
	Self and peer perceived dominance	Mean and standard deviation of rating from all participants; average score of peer rating	N/A	Unknown	Kim et al. (2008)
Contribution	Ideas proposed	Count the number of new ideas proposed	N/A	Arithmetic calculation	Karahalios and Bergstrom (2009) and Tausch et al. (2014, 2016)
	Voting	Count the number of votes one gave and received	N/A	Arithmetic calculation	Karahalios and Bergstrom (2009)
	Ideas alignment	Track similar ideas	Words; themes	Topic modelling	Karahalios and Bergstrom (2009)

Continued on the next page

Table 6 – continued from previous page

Constructs	Behavioural Indicators	Metrics	Features	Models	Studies
	Artifact created	Number and percentage of actions	Position of pen tips and shapes of pen traces	Arithmetic calculation	Wong-Villacrés et al. (2016)
Rapport	Eye contact	Count the number of times it happens	N/A	Manual coding	Balaam et al. (2011), Kulyk et al. (2005), and Terken and Sturm (2010)
	Simultaneous movement				Balaam et al. (2011)
	Tempo similarity in gestures and movements				
	Posture matching				
	Back channel responses				
Communication and respect	Self and peer rating	Average score of peer rating	N/A	Arithmetic calculation	Wong-Villacrés et al. (2016)
	Observer rating				
Coordination, mutual performance monitoring, constructive conflict, team emotional support	Self and peer rating				Koh et al. (2020)

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License (CC BY 4.0)

Table 7. Summary of behavioural assessment.

Constructs	Behavioural Indicators	Behavioural Assessment
Equal participation	Speaking time	Over-/under-participators are those who spoke more/less than the mean participation level plus the standard deviation of participation levels among all participants. Shorter speech segments indicate higher interactivity.
	Speaking volume	Louder volume indicates the confidence of the speaker, desire to lead or attempt to be heard over other speakers; variation in speech volume is the variation in speech energy, which indicates dominance in a group.
	Overlap speech	Simultaneous contributions can indicate negotiation between speakers, back-channels, or a common reaction such as laughing; less overlapping speech indicates people are more likely to collaborate.
	Turn taking	More turns indicate a higher level of interactions.
	Silence	Allow participants to judge the current length of silence.
	Amount of gesturing	An indicator of communication skills
	Self- and peer-perceived dominance	Subjects with values higher than one standard deviation over the mean were considered dominant (self and peer ratings are highly correlated).
Contribution	Ideas proposed	More ideas proposed indicates better performance in brainstorming activity.
	Voting	Voting indicates participation.
	Ideas alignment	None
	Artifact created	Compare with other members in the group.
Rapport	Eye contact	More eye contact indicates attention.
	Simultaneous body movement	More simultaneous body movement indicates more rapport among team members.
Communication, respect, coordination, mutual performance monitoring, constructive conflict, team emotional support	Self, peer, and observer rating	Comparison of their self ratings with the average of ratings from their team members

1. **Self-designed computations.** In line with the variability in operationalizing constructs identified in gap 2, we observed a range of self-designed metrics used to quantify behavioural indicators. Common metrics include calculating the deviation of an individual's speaking time from the desired balanced speaking time. However, other metrics, such as the average length of speaking segments, are also used. These computational strategies often rely on basic arithmetic calculations or displaying raw data. This indicates that the potential for more advanced modelling techniques has not been fully explored. Additionally, the absence of validated scales or thresholds guiding these calculation choices highlights the need for more standardized and validated metrics.

2. **Limited data fusion strategies.** Another observed trend is the limited use of data fusion strategies. Many studies rely on single-modality data derived from a single feature extracted from raw data. While such data can provide initial insights, a more comprehensive understanding typically requires the fusion of multiple data modalities to construct higher-level indicators (Ochoa, 2022). Although researchers often triangulate their findings with qualitative data, such as surveys or interviews, or validate them against human coders' results, there is a lack of studies where multiple data traces are fused to develop a holistic model of students' behaviours. This highlights a gap in leveraging the full potential of MmLA to enhance the depth and reliability of behavioural assessments.

3. **Self-interpreted behavioural assessment.** The third trend we noted is the self-interpreted behavioural assessment, where researchers subjectively interpret computational results. For instance, louder speaking volume has been interpreted as an indication of both confidence and dominance within a group (Bergstrom & Karahalios, 2007). Similarly, variations

in speaking volume have been associated with dominance (Kim et al., 2008). This issue is also related to the lack of theoretical support connecting the indicators and the behavioural assessment, specifically whether and how strongly the indicators provide evidence for the assessment. This leaves room for varied interpretations of behavioural data. Without validated guidelines for interpretation, it becomes challenging to provide accurate and constructive feedback to students.

4. **Lack of learning state estimation.** In our framework, the ending point of the analytics dimension is learning state estimation, which involves synthesizing various analyses to estimate students' levels of collaboration skills based on multiple behavioural assessments. Although none of the studies we reviewed fully achieve this, some have made progress toward this goal. For instance, Koh and colleagues (2020) present a report on four different dimensions of collaboration skills to students. While this report does not provide a comprehensive estimation of skill levels based on assessments in each dimension, it aims to offer students more holistic feedback on their collaboration skills.

These findings align with trends identified by other researchers in the broader field. For example, Schneider and colleagues (2021) observed that most multimodal collaboration analysis studies they reviewed relied on a single modality without fully leveraging the potential of multimodal data to generate deeper insights. Similarly, they found that most studies used simple calculations, such as arithmetic operations, rather than more sophisticated modelling techniques. The underlying reasons for these trends also relate to the lack of theoretical guidance and varied conceptualizations of collaboration and collaboration skills. Additionally, technical challenges and privacy concerns can also make capturing multimodal data difficult. The combination of these factors highlights the need for more integrated and theory-informed approaches to effectively using multimodal data in assessing and developing collaboration skills.

3.6 Gap 4: Missing Pedagogy

The final component of our framework is the pedagogy dimension. We examined feedback in terms of content, function, presentation, cognitive and emotional engagement, how students apply the feedback, and whether the system or study tracked students' progress. For feedback content, function, and presentation, we primarily used definitions developed by Hattie and Timperley (2007) and Narciss (2013). Specifically, following Hattie and Timperley (2007), we assessed whether the feedback included information specifying the goals (feed up), past performance (feed back), and next steps (feed forward). We also specified whether the feedback was addressed to the group or the individual. For feedback function, we evaluated whether it focused on the task (FT), the process (FP), or self-regulation (FR), following Hattie and Timperley (2007)'s framework. Lastly, we described the design of the feedback presentation. Our findings are summarized in Table 8.

Additionally, we summarized the student experience–related design considerations in Table 9. First, we examined whether any design considerations addressed students' sense-making processes, specifically if researchers considered whether students could understand and make sense of the feedback. We summarized the results in the "cognitive engagement" column. If the studies only asked students whether they understood the feedback, we labelled it as "post-survey" or "post-interview." Next, we examined whether affective factors such as emotion and motivation were considered in the design decisions and put the results in the "emotional engagement" column. To better understand the observed changes when students received feedback, we summarized the results in the "apply feedback" column. Lastly, we looked at whether and how the system tracked students' progress, documenting this in the "track progress" column. Through these analyses, we identified three trends:

1. **Limited feedback content scope.** According to Hattie and Timperley (2007), feedback content should include feed back (information about past performance), feed up (the goals they should achieve), and feed forward (the next steps). While all the systems provide "feed back" on students' collaboration performance, only two studies included the goals, and only one study included the next steps. We also examined the level at which the feedback is addressed—whether it is for the individual or for the group. Most studies included feedback that addresses both individual and group levels. Although not explicitly stated, the feedback at the group level is expected to help individuals set goals for themselves by observing others' performance. However, studies have shown that this type of "social mirror" might not work for everyone (Soller et al., 2005). The absence of explicit goals and next actionable steps could potentially make the feedback less actionable.

2. **Emphasize visualization over timing, schedule, and adaptivity.** According to Narciss (2013), the presentation of feedback includes several dimensions, such as timing, schedule, adaptation, coding, and modes. Our analysis found that all the studies focused on the mode of feedback, particularly on how to communicate ideas more efficiently using simple visualizations like pie charts or line charts. Some studies iterated their visualizations based on testing results to improve clarity. However, other crucial aspects of feedback presentation—such as timing, schedule, and adaptivity—are rarely considered. This limited focus on visualization alone may overlook the potential benefits of well-timed and adaptive feedback in enhancing students' learning and collaboration skills.

3. **Little attention to the mediating factors of feedback engagement.** Another trend we observed is the limited consideration of students' cognitive and emotional engagement during the system design phase. Although some

Table 8. Summary of feedback content, level, function, and presentation.

No.	Systems	Feedback Content	Feedback Level	Feedback Function	Feedback Presentation
1	Bachour et al. (2010)	Feed back	Individual, group	FR	Colour-coded circles displayed on a tabletop
2	Balaam et al. (2011)	Feed back	Group	FR	Ripples on an ambient display
3	Bergstrom and Karahalios (2007)	Feed back	Individual, group	FR	Projected concentric rings of coloured bars on a table
4	DiMicco et al. (2004)	Feed back, feed up	Individual, group	FR	Coloured histograms with annotations of participating states
5	Echeverria et al. (2017)	Feed back	Individual, group	FR	Pie chart and line graph
6	Karahalios and Bergstrom (2009) [Conversation vote]	Feed back	Individual, group	FT	Coloured rectangles
7	Karahalios and Bergstrom (2009) [Conversation cluster]	Feed back	Group	FT, FP	Word cloud, timeline of topics
8	Kim et al. (2008)	Feed back	Individual	FR	Coloured circles with different postions and line thickness
9	Koh et al. (2020)	Feed back, feed up, feed forward	Individual	FR	Radar chart
10	Kulyk et al. (2005) and Terken and Sturm (2010)	Feed back	Individual, group	FR	Coloured circles on a table
11	Praharaj et al. (2018)	Feed back	Individual, group	FR	Line chart over time
12	Starr et al. (2018)	Feed back	Individual, group	FR	Coloured bar chart
13	Tausch et al. (2014)	Feed back	Individual, group	FR	A garden of flowers
14	Tausch et al. (2016)	Feed back	Individual, group	FR	A garden of flowers
15	Wong-Villacrés et al. (2016)	Feed back	Individual, group	FR	Table and pie chart

studies employed post-surveys or interviews to gauge how students understood and felt about the visualizations, none incorporated these engagement factors into the initial design considerations. Furthermore, we found that only half of the studies examined how students applied the feedback. Of these, five studies reported effectiveness for only a subset of the group or only when the systems were in use. The lack of engagement factors could be a potential reason for this limited effectiveness. Lastly, only two studies used repeated measures to examine students' behavioural change over time, with all other systems being evaluated only once. This suggests a need for more longitudinal studies to understand the long-term impact of feedback systems on students' collaboration skills.

While the systems we reviewed have made significant contributions to offering feedback that facilitates collaborative behaviour modifications, their evaluation results are mixed and do not work for everyone. The reasons for these mixed results can vary. From a pedagogical perspective, our analysis indicates that most feedback offered may not be as actionable as intended. This is particularly crucial for CLA systems designed to develop collaboration skills, as these tools are not merely socio-technical but also instructional systems. Therefore, to advance the design of CLA systems, it is essential to incorporate pedagogically grounded feedback design principles.

Table 9. Summary of Feedback Engagement

No.	Systems	Cognitive Engagement	Emotional Engagement	Apply Feedback	Track Progress
1	Bachour et al. (2010)	Post-survey	Post-survey	A subset of group	None
2	Balaam et al. (2011)	Post-interview	Post-survey	Increase non-verbal synchrony subconciously	None
3	Bergstrom and Karahalios (2007)	Iterated design to make it easy to understand	None	None	The balancing behaviour disappeared when the visualization was removed.
4	DiMicco et al. (2004)	Add annotations	Post-survey	A subset of group	None
5	Echeverria et al. (2017)	None	None	None	None
6	Karahalios and Bergstrom (2009) [Conversation vote]	None	None	A subset of group	None
7	Karahalios and Bergstrom (2009) [Conversation cluster]	None	None	None	None
8	Kim et al. (2008)	None	None	A subset of group	None
9	Koh et al. (2020)	None	None	None	None
10	Kulyk et al. (2005) and Terken and Sturm (2010)	None	None	None	None
11	Praharaj et al. (2018)	None	None	None	None
12	Starr et al. (2018)	None	None	None	None
13	Tausch et al. (2014)	None	None	Effective for all	None
14	Tausch et al. (2016)	None	Post-survey	Feedback on both individuals and groups is more effective.	None
15	Wong-Villacrés et al. (2016)	None	Post-survey	None	Repeated measure over time

3.7 Generalizability Gap

Lastly, a general gap we observed is the scale and setting of evaluation studies: they predominantly consist of small-scale, one-time lab studies, limiting the generalizability of the findings. For instance, the size of participant groups ranged from a single group of four to as many as 144 groups of three or four students, with the majority being university students or adults. Furthermore, the scarcity of studies investigating the long-term effects of these systems is a significant gap. The lab-based nature of most evaluations raises questions about the ecological validity of the findings, suggesting a need for more diverse and longitudinal research designs to understand the impact of these tools in real-world settings. Although our framework does not focus explicitly on user testing, the identified trends underscore significant gaps in the current approach to designing and implementing automated feedback systems. These gaps highlight the need for studies that evaluate not only the immediate effectiveness of these tools but also their capacity to support the long-term development of collaboration skills. Addressing these concerns will be crucial for advancing the field and ensuring that these systems can meaningfully contribute to students' collaboration skills development in diverse and dynamic educational environments.

3.8 Conclusion

By applying our framework to the selected systems, we identified four major gaps, spanning theory, analytics, and pedagogical design dimensions. The first gap concerns the targeted collaboration skills, which consist of various sub-skills, some of which are critical yet unexplored. Second, the approaches used to operationalize the selected constructs are often not theory driven or pedagogically sound. Additionally, we observed a limited use of advanced modelling techniques, which, though not necessary, have the potential to generate more comprehensive insights. Lastly, for a CLA system aimed at developing collaboration skills, the feedback design from a pedagogical perspective should not be neglected. It is essential to consider feedback design from both content and student perspectives to ensure that it is actionable and meaningful for students' skills development. Previous research has made valuable contributions by collecting multimodal data from participants, performing analytics, and presenting feedback on collaborative behaviours to help groups raise awareness and make changes. Building on this foundation, and with the goal of creating more effective skill-development applications, we propose a set of research agendas in the next section. These agendas aim to bridge the identified gaps and advance this area.

4. The Roadmap: Research Agenda

To bridge the four gaps we identified, we propose four corresponding research agendas. These agendas aim to outline a feasible and actionable roadmap, moving from designing CLA tools that support temporary collaborative behavioural changes to developing sustained collaboration skills over time.

4.1 Research Agenda 1: Move toward Contextualized and Nuanced Development of Collaboration Skills

To bridge the identified gap 1, "underexplored and insufficiently supported collaboration skills," we propose the following two research agenda items. The first aims to adopt a developmental perspective to add additional lenses on developing collaboration skills, while the second focuses on building concrete steps to move in that direction.

4.1.1 Adopt a Developmental Perspective of Collaboration Skills Learning

First and foremost, we propose that, in addition to offering temporary support for collaboration, a developmental perspective on collaboration skills learning should be adopted. This addition should occur at two levels: the content level and the student level. The content level shift requires a thorough understanding of the skill itself from both theoretical and pedagogical perspectives: what the skill entails, how its definition may vary across different contexts, how the skill can be learned, and what the different phases and sub-goals of developing this skill are. The content generated by the systems should build toward achieving these pedagogical goals. At the student level, we propose shifting the view of learners from mere information receivers to active participants with varying learning trajectories and ways of engaging with the systems. This requires a thorough understanding of students' motivations, needs, challenges, prior knowledge, and skills. By centring on students' individual learning paths, the system can provide tailored support that addresses their specific needs.These shifts at both the content and student levels ensure that collaboration theory and the complex, dynamic nature of the collaboration process are taken into consideration. This approach ensures that the focus is on what is genuinely important for skill development and avoids providing uniform support at all times.

We have observed some promising design approaches that are making progress toward this direction. For instance, the Networking Performance Skill System employs evidence-centred design (ECD) to design the content delivery based on the evidence from student interactions (Scalise et al., 2007). ECD integrates four process models: the student model, representing students' current knowledge and skills; the task model, defining tasks that allow students to demonstrate their knowledge; the evidence model, which interprets performance using the first two models; and the assembly model, which guides the presentation of learning tasks. This systematic approach adapts to students' learning progression and adjusts the content based on their current skill levels. Another example is the application of HCD in educational settings, where systems are developed collaboratively with both students and teachers to address real-world educational needs. For instance, Echeverria Barzola (2020) implemented an HCD-MmLA approach to deeply understand the nuanced and contextualized needs of teachers and students in MmLA interface design. This method involves three stages: a co-design session that elicits needs and challenges in teaching and learning practices, an inductive mapping stage that links context to learning analytics components such as data and visualizations, and a theory-driven mapping stage that validates constructs identified in previous theories and literature. This ensures that the design is not only theoretically robust but also centred around the actual needs of students and teachers.

These promising design approaches could pave the way for us to shift toward a developmental and dynamic perspective of collaboration skills learning. Before advancing to this design phase, we recommend an initial and actionable step: compiling a comprehensive inventory of collaboration skills. This inventory will serve as a critical foundation for future design and research.

4.1.2 Compile a Contextualized Collaboration Skills Inventory

Collaboration skills have been widely researched in education, communication, and management studies. Although we compiled an initial list of skills in Section 3.1, it is by no means exhaustive or contextualized. To develop a comprehensive inventory, we propose starting with an extensive literature review to identify a wide range of collaboration skills. This should be followed by refining the list through discussion with diverse stakeholders, such as teachers, employers, and students. This collaborative approach ensures that the skills list is not only detailed but also tailored to real-world requirements. Each skill should be defined within its context, with clear descriptions of sub-skills and classification into different skill levels. These definitions and classifications will guide both design and research of automated systems in the future. A notable example is the collaboration skills taxonomy proposed by Lai and colleagues (2017), which classifies collaboration skills and team performance into five levels: non-participant, participant, cooperator, coordinator, and conflict resolver. Each level is characterized by specific behaviours. This taxonomy provides an initial framework for assessing and supporting collaboration skills.

By compiling this inventory prior to the design and research phases, we then ensure that our focus remains on the skills that students genuinely need to develop, rather than merely on data that is easily measurable. This approach also helps address the issue of some critical skills that are currently underexplored.

4.2 Research Agenda 2: Move toward Theory-Driven and Validated Construct Operationalization

To address these issues, we advocate shifting to the theory-driven approach (Wise & Shaffer, 2015) to operationalize the intended construct. This shift is to ensure that the constructs measured are both theoretically robust and empirically validated. Inspired by emerging practices in the field, we propose two actionable research agenda items.

4.2.1 Use Existing Validated Instruments

After identifying a specific collaboration skill, an extensive literature review should be conducted across related fields to determine if there are any existing validated measurement instruments that have already been operationalized with behavioural indicators. For example, active listening, a well-researched construct in communication studies, has various validated behavioural scales (Worthington & Bodie, 2017). While these instruments were initially created for human coders, recent advances in MmLA and large multimodal language models offer the potential to accelerate the measurement process and achieve accuracy levels beyond human capabilities. By leveraging sensors and algorithms to collect and analyze multimodal data traces, guided by validated scales, this approach ensures that our measures are theoretically grounded and robust.

4.2.2 Develop Validated Instruments

When validated scales for behavioural indicators of an intended construct are unavailable, the initial step involves exploring correlations between these indicators and validated scales. This exploration helps determine which data traces are significant and how they relate to the construct within a specific context. For instance, joint visual attention (JVA) is a construct commonly studied in MmLA. By exploring the relationship between JVA and the human coders' results using the validated rating scheme of collaboration quality (Meier et al., 2007), researchers have found that JVA positively correlates with high-quality collaborative interactions (Schneider & Pea, 2013; Schneider et al., 2018). Additionally, Cukurova and colleagues (2018) developed the Nonverbal Indexes of Students' Physical Interactivity (NISPI) framework, which can be used to interpret key parameters of students' CPS skills with students' non-verbal interaction data. They mapped nonverbal behaviours to four validated CPS constructs: synchrony, individual accountability, equality, and intra-individual variability. These constructs cover both individual and group levels of CPS. While these studies primarily focused on measurement rather than feedback delivery, their findings could provide a strong foundation for automating the measurement of student behaviours and provide actionable feedback on their collaboration skills.

4.3 Research Agenda 3: More Comprehensive Assessment Strategy

To address the gap caused by limited modelling techniques, we advocate for the adoption of more comprehensive assessment strategies. Collaboration skills include a complex and multifaceted set of sub-skills, and each sub-skill involves multiple constructs. Additionally, due to the dynamic nature of collaboration processes, adopting a developmental perspective is necessary. Lastly, with the goal of developing collaboration skills, it is important to consider pedagogy—specifically, how the assessment can be used to improve the teaching and learning of such skills. To move in this direction, we propose the following two research agenda items.

4.3.1 Theoretically Comprehensive Measurement

As identified in our analysis and in other MmLA literature reviews as well (Giannakos & Cukurova, 2023; Schneider et al., 2021), the use of theory in MmLA research is limited. There is a clear need for using theories to guide the analysis, assessment, and interpretation of results (Wise & Shaffer, 2015). In the psychometric field, researchers are making progress toward establishing theory-grounded measurements. For instance, Hao and colleagues (2019) proposed a general scoring strategy for assessing CPS. This strategy incorporates three components: collaboration outcome–based measures, collaboration process–based representations, and the mapping between these outcomes and processes. The outcome measures are binary, indicating whether a problem is solved. The process-based representations extract features from three sources: manually annotated communications; linguistic features; and metrics based on the amount, interactivity, and timing of communication. Lastly, statistical or machine learning models are used to map between the outcome measures and the process representations. This top-down approach ensures that the complexity of CPS assessment is not reduced to a simplistic single measurement or score.

Furthermore, we recommend measuring the effectiveness of these systems by adopting a developmental perspective. This approach is in line with the principle of assessment proposed by Scalise and colleagues (2007), which suggests that assessments should reflect the trajectory and progression of student learning. This principle advocates for evaluating the growth and development of students' understanding of knowledge and skills over time, rather than relying on a single measurement. By considering both the dynamics of collaboration processes and students' progression over time, this approach ensures a more comprehensive evaluation and lays the foundation for more targeted and actionable feedback.

4.3.2 Pedagogically Valid Data Fusion Strategies

To leverage the potential of multimodal data sources and gain deeper insights into students' learning trajectories, performing data fusion is essential (Ochoa, 2022). However, it is crucial to employ pedagogically valid strategies when fusing this data. Before proceeding with data fusion, we recommend assessing the importance of each feature within pedagogical contexts. For instance, it is important to differentiate features that are suitable for measurement from those that are appropriate for feedback. Non-verbal synchrony, such as simultaneous body movements or electrodermal activity (EDA) synchrony, may be natural results of effective collaboration but are not necessarily actionable for individuals seeking to improve. Moreover, understanding what is pedagogically important to teach in the given context is crucial. We advocate for designing assessments that clearly inform students about the quality of their skills, which can enhance their understanding and skills development.

4.4 Research Agenda 4: Promote Feedback Actionability through a Student-Centred Approach

To address the gap of "missing pedagogy," we recommend compensating for this shortfall in two key areas: feedback content and feedback engagement. Given that the development of collaboration skills is fundamentally pedagogical, it is crucial to consider the entire spectrum of pedagogical feedback. This includes understanding internal feedback mechanisms, how external feedback interacts with internal feedback, and other mediating factors in the feedback process. Additionally, introducing participatory design and co-design principles from the field of human–computer interaction has the potential to advance our understanding and practices in these areas. Based on these considerations, we propose the following two research agenda items.

4.4.1 Feedback Content: Clear Goals, Direct Paths, and Adaptivity

From the feedback content perspective, the systems we reviewed generally lack clear action goals and plans for students to act upon. Furthermore, none of these systems appear to consider the temporal aspects of collaboration. However, the field of learning analytics is already advancing in this direction. For example, in the CPS coaching system proposed by D'Mello and colleagues (2024), the system was designed based on the CPS framework, which consists of three facets and several subskills. At each stage, students received feedback on their previous performance, learned about a specific facet and its indicators, applied what they learned in the next round of collaboration, and then received feedback again. In this cycle of "collaborate-feedback-learn-collaborate-feedback," students were given explicit goals and specific standards, and the feedback served as scaffolds for them to reflect on and practise their efforts toward the goal. Another example is the analytics-supported feedback tool proposed by Jung and Wise (2024), which not only provides a direct pathway to actions based on the analytics but also aligns the timing of these analytics with different phases of learning activities. These approaches go beyond creating easy-to-understand visualizations or dashboards; they integrate analytics with learning more seamlessly, enhancing both adaptivity and relevance.

4.4.2 Feedback Engagement: Centring Students

Feedback should not be viewed merely as information delivered to students but as a dialogic process between the feedback mechanism and the student. In the learning analytics community, some empirical models have been proposed to better understand how instructors use dashboards. For example, in the model of teacher dashboard usage proposed by Van Leeuwen and colleagues (2021), the process of seeing and acting upon feedback is not linear. Several teacher characteristics, such as pedagogical knowledge, data literacy, professional routines, and age and gender, mediate the sense-making processes of using teacher dashboards. To account for these mediating factors, the recent introduction of participatory design and co-design into learning analytics represents an initial step toward understanding these learner factors. For instance, Sarmiento and colleagues (2020) engaged students as co-designers for a learning analytics dashboard to centre their voices. Moreover, it is crucial to understand the entire cycle of internal and external feedback processes. The interactive tutoring feedback model proposed by Narciss (2013) detailed learner factors that mediate feedback effectiveness, including students' goals, understanding of standards, prior knowledge, and motivation. Before designing feedback, it is essential to consider these factors because the development of collaboration skills requires specific learning conditions. By acknowledging these factors, we aim to design systems that genuinely cater to learners' conditions and needs, thereby improving the actionability of the feedback.

We propose these research agendas to bridge the identified gaps in the current systems of CLA, specifically those aimed at changing students' collaboration behaviours in collocated environments. By suggesting shifts in perspectives and practices and providing examples, we aim for these agendas to serve as a roadmap for designing and developing effective analytics-supported feedback systems that foster the development of collaboration skills.

5. Conclusion

In this paper, we have addressed the development of collaboration skills through analytics-supported feedback in collocated environments. First, we proposed an integrated design and evaluation framework that incorporates theory, analytics, and pedagogy (TAP). We then conducted a semi-systematic review of papers focusing on student-facing feedback systems aimed at improving collaborative behaviours in collocated environments. Using the 15 criteria in our proposed framework, we explored the current landscape of this area. By comparing current practices with our framework, we identified 14 research trends and summarized them into four major gaps: unexplored and insufficiently supported collaboration skills, lack of validation in operationalizing theory, limited modelling techniques, and missing pedagogy.

To bridge these identified gaps, we proposed four research agendas. For addressing collaboration skills, we suggest shifting the focus from temporary collaboration processes to supporting collaboration skills in a more nuanced and contextualized way. To tackle the lack of validation in construct operationalization, we recommend using existing validated instruments or developing new ones to ensure that construct operationalization is theoretically grounded. Regarding the limited modelling techniques, we advocate for more comprehensive, theoretically grounded, and pedagogically valid assessment strategies, recognizing that collaboration skills are multifaceted and inherently pedagogical. Lastly, for feedback design, we propose a student-centred approach that focuses on clearer content and considers students' cognitive and affective engagement to make feedback truly actionable. Together, these research agendas outline a roadmap and initial steps toward designing effective feedback systems for the development of collaboration skills.

5.1 Practical Implications

While the framework and research agenda are primarily aimed at researchers, educators and learning designers can also benefit from integrating elements of the framework into their daily teaching and design practices. For example, if resources allow, feedback design grounded in theoretical and pedagogical principles could be incorporated into the curriculum even before the use of technology. Additionally, by collaborating with researchers and engaging in genuine research–practice partnerships, educators and designers could not only implement systems to support the development of collaboration skills but also gain valuable insights from the research and design process to enhance their own practices.

For researchers, while certain items in the proposed research agenda, such as compiling a collaboration skills inventory, can be accomplished by a small team in a relatively short time, implementing the broader agenda and fully realizing the framework will require interdisciplinary collaboration and long-term effort. Furthermore, ethical considerations must be carefully addressed, including protecting students' data privacy and ensuring the cultural responsiveness of CLA tools. Such considerations are essential to ensure that the framework is equitable and beneficial for all learners, across diverse educational and cultural contexts.

5.2 Limitations

This study should also be understood with its limitations. First, our focus on collocated environments, which are naturally suited for implementing MmLA, limited the scope of our search, resulting in only 15 systems meeting our criteria. While including systems in online environments would likely add more information to the identified gaps and expand the research agendas, the main trends found in this paper are expected to remain valid. The scarcity of research in this field underscores the potential for future studies in this direction. Second, we acknowledge that the systems reviewed may not represent an exhaustive list of all available systems. This is partly due to the reliance on systematic reviews published in 2021, which did not include systems developed after that time. However, since these reviews systematically covered research spanning the past 20 years, and we conducted additional reference-based searches, we believe our analysis effectively captures the major trends and gaps in the field. Our primary goal was to identify overarching themes and directions in the field rather than providing a comprehensive inventory of all systems. Third, regarding the generalizability of this study, we note that most of the systems reviewed were tested with university students or adults. This underscores the need for more research focusing on K–12 settings to ensure broader applicability. Additionally, we did not discuss cultural contexts in our analysis. Future studies should investigate how CLA tools can be made culturally responsive to better support collaboration skills development in diverse international settings. Finally, our paper aims to shed light on the structures and designs of CLA systems for collaboration skills development rather than evaluating their effectiveness. Future studies involving more statistical evaluation would be needed to assess the effectiveness and generalizability of these systems.

Collaboration skills are among the most fundamental skills in the 21st-century workplace, and CLA tools hold significant potential to help students develop these skills. Our study represents an initial step in addressing the development of collaboration skills through learning analytics. We advocate for broader examinations and implementations of the theoretical, analytical, and pedagogical aspects of these CLA tools. By doing so, we can enhance the effectiveness of these tools in developing essential collaboration skills in students, ultimately preparing them for success in their academic, professional, and civic lives.

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors declared no financial support for the research, authorship, and/or publication of this article.

References

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. https://doi.org/10.1109/TLT.2010.18

Balaam, M., Fitzpatrick, G., Good, J., & Harris, E. (2011, May). Enhancing interactional synchrony with an ambient display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (SIGCHI 2011), 7–12 May 2011, Vancouver, British Columbia, Canada (pp. 867–876). ACM. https://doi.org/10.1145/1978942.1979070

Bergstrom, T., & Karahalios, K. (2007). Conversation Clock: Visualizing audio patterns in co-located groups. In R. Sprague (Ed.), *Proceedings of the 2007 40th Annual Hawaii International Conference on System Sciences* (HICSS 2007), 3–6 January, 2007, Waikoloa, Hawaii, USA (p. 78). IEEE. https://doi.org/10.1109/HICSS.2007.151

Bodie, G. D., Vickery, A. J., Cannava, K., & Jones, S. M. (2015). The role of "active listening" in informal helping conversations: Impact on perceptions of listener helpfulness sensitivity and supportiveness and discloser emotional improvement. *Western Journal of Communication*, *79*(2), 151–173. https://doi.org/10.1080/10570314.2014.943429

Boothe, M., Yu, C., Lewis, A., & Ochoa, X. (2022). Towards a pragmatic and theory-driven framework for multimodal collaboration feedback. In *Proceedings of the 12th International Conference on Learning Analytics and Knowledge* (LAK 2022), 21–25 March 2022, online (pp. 507–513). ACM. https://doi.org/10.1145/3506860.3506898

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281. https://doi.org/10.1037/h0040957

Cukurova, M., Luckin, R., Millán, E., & Mavrikis, M. (2018). The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. *Computers & Education*, *116*, 93–109. https://doi.org/10.1016/j.compedu.2017.08.007

D'Angelo, S., & Gergle, D. (2016). Gazed and confused: Understanding and designing shared gaze for remote collaboration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI 2016), 7–12 May 2016, San Jose, California, USA (pp. 2492–2496). ACM. https://doi.org/10.1145/2858036.2858499

Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning*, *34*(4), 338–349. https://doi.org/10.1111/jcal.12288

DiMicco, J. M., Pandolfo, A., & Bender, W. (2004). Influencing group participation with a shared display. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (CSCW 2004), 6–10 November 2004, Chicago, Illinois, USA (pp. 614–623). ACM. https://doi.org/10.1145/1031607.1031713

D'Mello, S. K., Duran, N., Michaels, A., & Stewart, A. E. B. (2024). Improving collaborative problem-solving skills via automated feedback and scaffolding: A quasi-experimental study with CPSCoach 2.0. In J. Masthoff (Ed.), *User modeling and user-adapted interaction* (pp. 1087–1125, Vol. 34). Springer Nature Link. https://doi.org/10.1007/s11257-023-09387-6

Echeverria, V., Martinez-Maldonado, R., Chiluiza, K., & Shum, S. B. (2017). DBCollab: Automated feedback for face-to-face group database design. In *Proceedings of the 25th International Conference on Computers in Education* (ICCE 2017), Main Conference Proceedings, 4–8 December 2017, Christchurch, New Zealand. APSCE. https://library.apsce.net/index.php/ICCE/article/view/2240

Echeverria Barzola, V. I. (2020). *Designing feedback for collocated teams using multimodal learning analytics* [PhD]. University of Technology Sydney. https://hdl.handle.net/10453/140936

Giannakos, M., & Cukurova, M. (2023). The role of learning theory in multimodal learning analytics. *British Journal of Educational Technology*, *54*(5), 1246–1267. https://doi.org/10.1111/bjet.13320

Hao, J., Liu, L., Kyllonen, P., Flor, M., & von Davier, A. A. (2019). Psychometric considerations and a general scoring strategy for assessments of collaborative problem solving. *ETS Research Report Series*, *2019*(1), 1–17. https://doi.org/10.1002/ets2.12276

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hu, L., & Chen, G. (2021). A systematic review of visual representations for analyzing collaborative discourse. *Educational Research Review*, *34*, 100403. https://doi.org/10.1016/j.edurev.2021.100403

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

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist*, *48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006

Jung, Y., & Wise, A. F. (2024). Probing actionability in learning analytics: The role of routines, timing, and pathways. In *Proceedings of the 14th International Conference on Learning Analytics and Knowledge* (LAK 2024), 18–22 March 2024, Kyoto, Japan (pp. 871–877). ACM. https://doi.org/10.1145/3636555.3636914

Karahalios, K., & Bergstrom, T. (2009). Social mirrors as social signals: Transforming audio into graphics. *IEEE Computer Graphics and Applications*, *29*(5), 22–32. https://doi.org/10.1109/MCG.2009.85

Kim, T., Chang, A., Holland, L., & Pentland, A. S. (2008). Meeting mediator: Enhancing group collaboration using sociometric feedback. *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (CSCW 2008), 8–12 November 2008, San Diego, California, USA, 457–466. https://doi.org/10.1145/1460563.1460636

Koh, E., Tee, Y.-H., Suresh, D., & Caleon, I. (2020). A digital formative assessment intervention for cultivating teamwork skills. In M. Gresalfi & I. S. Horn (Eds.), *The interdisciplinarity of the learning sciences, 14th international conference of the learning sciences (ICLS) 2020* (pp. 1597–1600, Vol. 3). International Society of the Learning Sciences (ISLS). https://repository.isls.org//handle/1/6375

Kulyk, O., Wang, J., & Terken, J. (2005). Real-time feedback on nonverbal behaviour to enhance social dynamics in small group meetings. In S. Renals & S. Bengio (Eds.), *Machine learning for multimodal interaction. MLMI 2005. Lecture notes in computer science* (Vol. 3869). Springer. https://doi.org/10.1007/11677482_13

Lai, E., DiCerbo, K., & Foltz, P. (2017). *Skills for today: What we know about teaching and assessing collaboration* (tech. rep.). Pearson. https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/efficacy-and-research/skills-for-today/Collaboration-FullReport.pdf

Lim, L., Dawson, S., Joksimovic, S., & Gašević, D. (2019). Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference? In *Proceedings of the Ninth International Conference on Learning Analytics and Knowledge* (LAK 2019), 4–8 March 2019, Tempe, Arizona, USA (pp. 250–259). ACM. https://doi.org/10.1145/3303772.3303804

Lipnevich, A. A., Berg, D. A. G., & Smith, J. K. (2016). Toward a model of student response to feedback. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 169–185). Routledge. https://doi.org/10.4324/9781315749136

Liu, O. L., Ling, G., & Fishtein, D. (2021). Charting the skills for the future: Perspectives from employers and higher education institutions. *Chinese/English Journal of Educational Measurement and Evaluation*, *2*(2). https://doi.org/10.59863/VNWR4406

Lyndgaard, S. F., & Kanfer, R. (2024). Interpersonal, intrapersonal, and cognitive tactics: A thematic analysis of adults' 21st century learning management. *Journal of the Learning Sciences*, *33*(1), 175–211. https://doi.org/10.1080/10508406.2024.2303777

Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning*, *2*(1), 63–86. https://doi.org/10.1007/s11412-006-9005-x

Narciss, S. (2013). Designing and evaluating tutoring feedback strategies for digital learning environments on the basis of the Interactive Tutoring Feedback Model. *Digital Education Review*, *23*, 7–26. https://www.learntechlib.org/p/131614/

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, D. Leutner, & R. Brünken (Eds.), *Instructional design for multimedia learning: Proceedings of the Fifth International Workshop of SIG 6 Instructional Design of the European Association for Research on Learning and Instruction* (EARLI 2002), 27–29 June 2002, Erfurt, Germany (Vol. 181195). Waxmann.

Ochoa, X. (2022). Chapter 6: Multimodal learning analytics: Rationale, process, examples, and direction. In C. Lang, G. Siemens, A. F. Wise, D. Gašević, & A. Merceron (Eds.), *The handbook of learning analytics* (pp. 54–65). SoLAR. https://doi.org/10.18608/hla22.006

Ochoa, X., Echeverria, V., Carrillo, G., Heredia, V., & Chiluiza, K. (2023). Supporting online collaborative work at scale: A mixed-methods study of a learning analytics tool. In *Proceedings of the 10th ACM Conference on Learning @ Scale* (L@S 2023), 20–22 July 2023, Copenhagen, Denmark (pp. 237–247). ACM. https://doi.org/10.1145/3573051.3596165

OECD. (2017, August). *PISA 2015 assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving* (tech. rep.). OECD. https://doi.org/10.1787/9789264281820-en

Ofstedal, K., & Dahlberg, K. (2009). Collaboration in student teaching: Introducing the collaboration self-assessment tool. *Journal of Early Childhood Teacher Education*, *30*(1), 37–48. https://doi.org/10.1080/10901020802668043

Orona, G., Liu, O. L., & Arum, R. (2023). The language of 21st century skills: Next directions for closing the skills gap between employers and postsecondary graduates. *Chinese/English Journal of Educational Measurement and Evaluation*, *4*(2). https://doi.org/10.59863/OIVI3767

Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review*, *35*, 100416. https://doi.org/10.1016/j.edurev.2021.100416

Praharaj, S., Scheffel, M., Drachsler, H., & Specht, M. (2018). Multimodal analytics for real-time feedback in co-located collaboration. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, & M. Scheffel (Eds.), *Lifelong* *technology-enhanced learning. EC-TEL 2018. Lecture notes in computer science* (pp. 187–201, Vol. 11082). Springer. https://doi.org/10.1007/978-3-319-98572-5_15

Praharaj, S., Scheffel, M., Drachsler, H., & Specht, M. (2021). Literature review on co-located collaboration modeling using multimodal learning analytics—can we go the whole nine yards? *IEEE Transactions on Learning Technologies*, *14*(3), 367–385. https://doi.org/10.1109/TLT.2021.3097766

Sarmiento, J. P., Campos, F., & Wise, A. (2020). Engaging Students as Co-Designers of Learning Analytics. In *Companion Proceedings of the 10th International Learning Analytics and Knowledge Conference* (LAK 2020), 23–27 March, Frankfurt, Germany (pp. 29–32). SoLAR. https://www.solaresearch.org/wp-content/uploads/2020/06/LAK20_Companion_Proceedings.pdf

Scalise, K., Bernbaum, D. J., Timms, M., Harrell, S. V., Burmester, K., Kennedy, C. A., & Wilson, M. (2007). Adaptive technology for e-learning: Principles and case studies of an emerging field. *Journal of the American Society for Information Science and Technology*, *58*(14), 2295–2309. https://doi.org/10.1002/asi.20701

Schneider, B., & Pea, R. (2013). Using eye-tracking technology to support visual coordination in collaborative problem-solving groups. In N. Rummel, M. Kapur, M. Nathan, & S. Puntambekar (Eds.), *To see the world and a grain of sand: Learning across levels of space, time, and scale: CSCL 2013 conference proceedings—full papers & symposia* (pp. 406–413, Vol. 1). International Society of the Learning Sciences. https://repository.isls.org/handle/1/1935

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2018). Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning*, *13*(3), 241–261. https://doi.org/10.1007/s11412-018-9281-2

Schneider, B., Sung, G., Chng, E., & Yang, S. (2021). How can high-frequency sensors capture collaboration? A review of the empirical links between multimodal metrics and collaborative constructs. *Sensors*, *21*(24), 8185. https://doi.org/10.3390/s21248185

Singh Dubey, R., Paul, J., & Tewari, V. (2022). The soft skills gap: A bottleneck in the talent supply in emerging economies. *The International Journal of Human Resource Management*, *33*(13), 2630–2661. https://doi.org/10.1080/09585192.2020.1871399

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Soller, A., Martinez, A., Jermann, P., & Muehlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *Journal of Artificial Intelligence in Education*, *15*(4), 261–290. https://dl.acm.org/doi/abs/10.5555/1434935.1434937

Starr, E. L., Reilly, J. M., & Schneider, B. (2018). Toward using multi-modal learning analytics to support and measure collaboration in co-located dyads. In J. Kay & R. Luckin (Eds.), *Rethinking learning in the digital age: Making the learning sciences count, 13th international conference of the learning sciences (ICLS) 2018* (Vol. 1). International Society of the Learning Sciences. https://repository.isls.org//handle/1/888

Stevens, M. J., & Campion, M. A. (1994). The knowledge, skill and ability requirements for teamwork: Implications for human resource management. *Journal of Management*, *20*(2), 503–530. https://doi.org/10.1016/0149-2063(94)90025-6

Tausch, S., Hausen, D., Kosan, I., Raltchev, A., & Hussmann, H. (2014). Groupgarden: Supporting brainstorming through a metaphorical group mirror on table or wall. In *Proceedings of the Eighth Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (NordCHI 2014), 26–30 October 2014, Helsinki, Finland (pp. 541–550). ACM. https://doi.org/10.1145/2639189.2639215

Tausch, S., Ta, S., & Hussmann, H. (2016). A comparison of cooperative and competitive visualizations for co-located collaboration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI 2016), 7–12 May 2016, San Jose, California (pp. 5034–5039). ACM. https://doi.org/10.1145/2858036.2858072

Terken, J., & Sturm, J. (2010). Multimodal support for social dynamics in co-located meetings. *Personal and Ubiquitous Computing*, *14*(8), 703–714. https://doi.org/10.1007/s00779-010-0284-x

Van Leeuwen, A., Knoop-van Campen, C. A. N., Molenaar, I., & Rummel, N. (2021). How teacher characteristics relate to how teachers use dashboards: Results from two case studies in K–12. *Journal of Learning Analytics*, *8*(2), 6–21. https://doi.org/10.18608/jla.2021.7325

Wise, A. F., Knight, S., & Buckingham Shum, S. (2021). Collaborative learning analytics. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning* (pp. 425–443). Springer International Publishing. https://doi.org/10.1007/978-3-030-65291-3_23

Wise, A. F., & Shaffer, 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

Wong-Villacrés, M., Granda, R., Ortiz, M., & Chiluiza, K. (2016). Exploring the impact of a tabletop-generated group work feedback on students' collaborative skills. In R. Martínez Maldonado & D. Hernández Leo (Eds.), *Proceedings* *of the First International Workshop on Learning Analytics across Physical and Digital Spaces* (CrossLAK 2016) *co-located with the Sixth International Conference on Learning Analytics and Knowledge* (LAK 2016), 25–29 April 2016, Edinburgh, UK (pp. 58–64). CEUR. https://ceur-ws.org/Vol-1601/CrossLAK16Paper11.pdf

Worsley, M., & Ochoa, X. (2020). Towards collaboration literacy development through multimodal learning analytics. In *Companion Proceedings of the 10th International Learning Analytics and Knowledge Conference* (LAK 2020), 23–27 March, Frankfurt, Germany (pp. 585–595). SoLAR. https://www.solaresearch.org/wp-content/uploads/2020/06/LAK20_Companion_Proceedings.pdf

Worthington, D. L., & Bodie, G. (Eds.). (2017). *The sourcebook of listening research: Methodology and measures*. John Wiley & Sons, Inc. https://doi.org/10.1002/9781119102991

Zhu, X., Shui, H., & Chen, B. (2023). The synthesis lab: Empowering collaborative learning in higher education through knowledge synthesis. In *CSCW '23 Companion: Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*, 14–18 October 2023, Minneapolis, Minnesota, USA (pp. 245–248). ACM. https://doi.org/10.1145/3584931.3606996