Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Dual Analysis of Literature and Existing Tools for Establishment of Acceptable xAPI Profile for Collaborative Learning

Mitsushiro EZOE[ab*], Yasuhisa TAMURA[c] & Ren KOMATSUZAKI[c]
[a]*The University of Electro-Communications, Japan*
[b]*codeTakt, Inc., Japan*
[c]*Sophia University, Japan*
*mitsushiro.ezoe@codetakt.com

Abstract: The growing use of digital tools in collaborative learning necessitates standardized methods for analyzing learner interactions. While xAPI offers a flexible syntax for logging learning activities, a common vocabulary for collaborative learning is lacking. This hinders cross-platform data integration and analysis. This study investigates the feasibility of applying existing xAPI vocabularies to describe and analyze collaborative learning logs. We conducted a dual analysis: a literature review of 72 Computer-Supported Collaborative Learning (CSCL) studies to identify theoretically important interactions, and an empirical analysis of interaction logs from widely used educational tools in Japanese K-12 settings to identify common real-world actions. Our findings reveal a significant gap between the higher-order interactions studied in research (e.g., assessing, tracking) and the more fundamental actions logged by tools (e.g., editing, submitting). This suggests that while existing vocabularies can cover a majority of current tool-based logs, they fall short of capturing the full complexity of collaborative processes. This study clarifies the current state and challenges of standardizing collaborative learning data, proposing that establishing a baseline for fundamental action logging is a crucial first step toward more advanced, cross-platform learning analytics.

Keywords: Collaborative Learning, Learning Analytics, xAPI, Data Standardization

1. Introduction

Learning Analytics (LA) aims to improve education by analyzing data from various learning environments. As LA expands from individual learning to complex, multi-tool environments like Computer-Supported Collaborative Learning (CSCL), a fundamental challenge arises: data interoperability. Diverse platforms log interactions in proprietary formats, making it difficult to integrate data, reproduce analyses, and scale findings. This fragmentation means that a student's collaborative process, which might span a discussion forum, a shared document, and a presentation tool, is broken into isolated data silos. Consequently, educators and researchers cannot obtain a comprehensive understanding of group dynamics or individual contributions, significantly hindering evidence-based support.

The Experience API (xAPI) is a technical standard designed to address this by providing a common format for learning data. However, xAPI only defines the syntax. For meaningful cross-system analysis, a shared vocabulary and structure—defined in an xAPI Profile—are essential. An xAPI Profile specifies semantic rules for a particular context by defining Concepts (verbs, activity types), Statement Templates (structural patterns for statements), and Patterns (sequences of related events). By enforcing consistency, profiles ensure that records from different tools can be meaningfully compared, enabling the development of reusable dashboards and analytical tools. While profiles exist for domains like self-regulated learning, a widely accepted profile for collaborative learning is still missing.

Collaborative learning involves complex, distributed, and multi-modal interactions such as proposing ideas, reaching consensus, and coordinating actions. These interactions are often spread across multiple digital tools (e.g., LMS, shared whiteboards, messaging apps), each with its own data schema. This fragmentation makes it nearly impossible to gain a holistic view of collaborative processes. A standardized approach to describing these interactions is a critical prerequisite for advancing collaborative learning analytics, especially in contexts like Japan's GIGA School Program, which has rapidly deployed devices nationwide and amplified the need for interoperable data.

However, a significant gap often exists between the complex interactions that learning theories emphasize and the simple, fundamental actions that educational tools actually log. Before a comprehensive new profile can be designed, it is crucial to understand the nature and extent of this gap. This study, therefore, does not aim to propose a new, definitive profile. Instead, it takes a necessary prior step: to assess the feasibility of using existing, widely recognized xAPI vocabularies to describe collaborative learning. We explore the alignment and misalignment between theoretically significant interactions and empirically observed behaviors in real-world educational settings. This leads to our research questions:

- RQ1: What types of learner interactions are frequently studied in CSCL literature?
- RQ2: What types of learner interactions are most frequently logged by common collaborative learning tools in practice?
- RQ3: To what extent can existing xAPI vocabularies bridge the gap between theoretical models and practical tool logs?

2. Methods

This study employs a mixed-method approach, structured in three phases. The scope is focused on the Japanese K-12 educational context, which has seen rapid, nationwide adoption of digital learning tools.

2.1 Theoretical Analysis (RQ1)

We conducted a literature review to identify interaction types central to collaborative learning research. We searched Google Scholar for empirical CSCL studies, resulting in a final set of 72 peer-reviewed articles. For each article, we identified the dataset used and categorized the described learner interactions using verbs from the existing TLA (Teacher-Led Learning Analytics) Profile. This profile was chosen as a reference because it is a well-established, research-oriented vocabulary designed to capture pedagogical interactions. Its focus on describing learning activities from an analytical perspective provided a suitable lens for classifying the objectives of prior empirical research. The process involved mapping the descriptions of datasets and analytical methods in each paper to the most appropriate verbs within the TLA Profile, thereby creating a quantitative overview of research trends in the CSCL field. This process, facilitated by GPT-4o for initial screening and data extraction, allowed us to quantify which collaborative behaviors are most frequently targeted in academic research. A complete list of references cited in this study is provided in the Appendix.

2.2 Empirical Analysis (RQ2)

To understand real-world learner behavior, we analyzed anonymized interaction logs from schoolTakt, a major classroom support system used in Japanese elementary and junior high schools, alongside a functional comparison with two other popular tools, LoiLoNote School and Miraiseed. These tools were selected due to their high market share and representative functionalities in the Japanese K-12 context. The GIGA School Program has led to the widespread adoption of these specific systems, making their logs a valuable source of authentic, large-scale data on student interactions. Our analytical framework (Figure 1) was designed to abstract the specific UI elements of each tool into generalizable action types, enabling a more fundamental comparison of their core functionalities and the types of data

they generate. As illustrated in Figure 1, we classified user actions based on a three-level framework: (a) actions on the content of the learning item, (b) actions on the learning item itself, and (c) actions directed toward others' items. This framework allowed us to identify the most common, fundamental user operations. Figure 2 provides a concrete example of a typical activity flow within schoolTakt, illustrating the sequence of these fundamental actions from the start of a lesson to its completion.

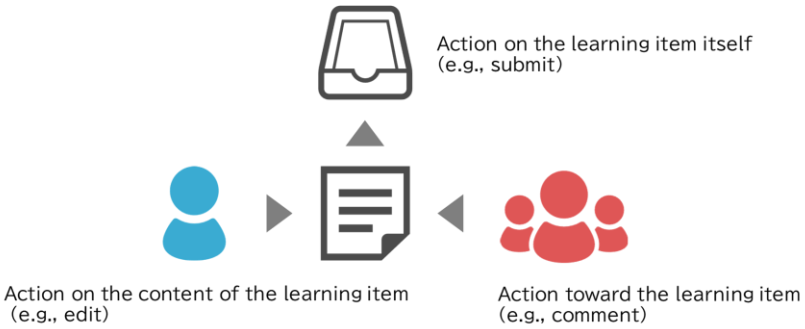

Figure 1. Classification of Actions Related to Learning Items

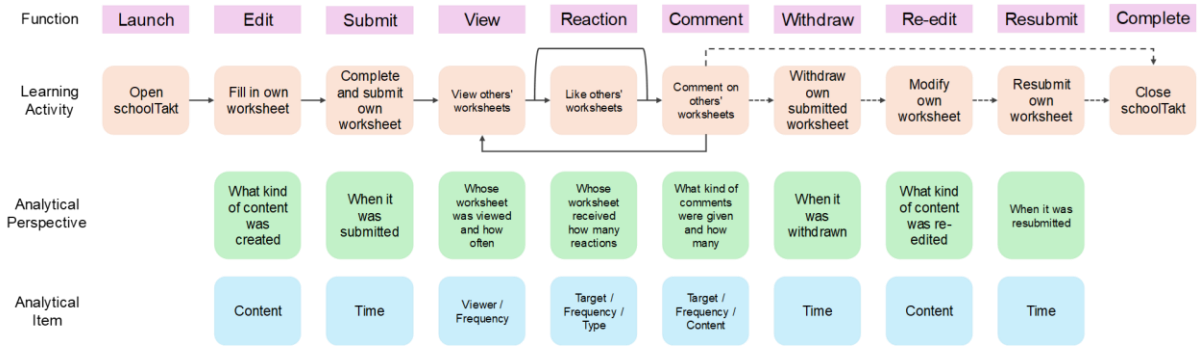

Figure 2. Example of Collaborative Learning Activities Using schoolTakt

2.3 Comparative Analysis (RQ3)

Finally, we compared the findings from the theoretical and empirical analyses. We evaluated the extent to which the interaction types identified in the literature (RQ1) align with the actions logged by the tools (RQ2). This comparative analysis aimed to determine the coverage and limitations of applying a standardized, existing vocabulary to describe both theoretical constructs and practical behaviors.

3. Results

3.1 Theoretical Analysis (RQ1)

The theoretical analysis covered 72 CSCL studies, which together utilized 338 distinct data instances for their analyses. The frequency of TLA Profile verbs used to describe these instances showed a clear and uneven distribution, as summarized in Figure 3. A small number of verbs related to data collection and evaluation were used frequently, while most of the 38 available verbs in the profile were rare. Specifically, analytical verbs like captured (used in 154 instances), tracked (135), and assessed (98) dominated the findings. This strong focus on monitoring and evaluation verbs suggests that the data models in the literature often prioritize

the act of analysis itself over the semantic description of specific collaborative processes. In other words, the research tends to describe the data as something that was "captured for analysis" rather than describing the rich, underlying collaborative action (e.g., "negotiating meaning" or "building consensus") that the data represents.

	All	(%)	Measured Data	(%)	Analyzed Data	(%)
データの個数	338		230		108	
captured	154	45.5621	137	59.5652	17	15.7407
tracked	135	39.9408	105	45.6522	30	27.7778
assessed	98	28.9941	56	24.3478	42	38.8889
evaluated	73	21.5976	39	16.9565	34	31.4815
explored	51	15.0888	29	12.6087	22	20.3704
contextualized	51	15.0888	25	10.8696	26	24.0741
surveyed	47	13.9053	43	18.6957	4	3.7037
socialized	26	7.6923	17	7.3913	9	8.3333
asserted	15	4.4379	15	6.5217	0	0.0000
organized	14	4.1420	10	4.3478	4	3.7037
clarified	10	2.9586	5	2.1739	5	4.6296
located	9	2.6627	9	3.9130	0	0.0000
directed	7	2.0710	2	0.8696	5	4.6296
selected	7	2.0710	2	0.8696	5	4.6296
inferred	6	1.7751	1	0.4348	5	4.6296
screened	5	1.4793	0	0.0000	5	4.6296
validated	3	0.8876	1	0.4348	2	1.8519

Figure 3. Verb Frequency in CSCL Literature

3.2 Empirical Analysis (RQ2)

The analysis of tool logs revealed a focus on more fundamental, content-creation actions. Table 1 provides a functional classification of the three major tools. While core functions like Edit and Submit are common, the implementation of peer-interaction functions like View, Comment, and Like/Stamp varies significantly. For example, while all tools support content creation (Edit), the mechanisms for peer interaction (View, Comment) are not standardized, posing a direct obstacle to cross-platform analysis of collaborative behaviors. The functions are primarily concrete, individual contributions to a shared space, as visualized in the activity flow in Figure 2.

Table 1. *Common Functions in Collaborative Learning Tools*

Action Category	Action Type	LoiLoNote	Okulink Plus	schoolTakt
Actions on the Content of the Learning Item	Edit	○	○	○
	Submit	○	○	○
	Send	○	○	–
	Receive	○	○	–
Actions on the Learning Item Itself	Withdraw	–	–	○
	Delete	○	○	*1
	Duplicate	○	○	○
Actions Toward the Learning Item	View	*2	*2	○
	Like	–	*3	○
	Stamp	–	○	–
	Comment	–	○	○

*1: LoiLoNote and Okulink Plus permit students to delete learning items from their own screens.
*2: In LoiLoNote and Okulink Plus, students must first receive an item and then import it into their own board for viewing.
*3: Okulink Plus offers several reaction stamps, including applause and flower stamps, commonly used in Japanese classrooms to indicate high praise.

3.3 Comparative Analysis (RQ3)

A clear discrepancy emerged between the theoretical and empirical findings. While research emphasizes analytical verbs like assessed and tracked, the tools primarily log foundational actions like edit, submit, and comment. This indicates a significant gap: existing tools are effective at capturing what students produce, but not necessarily how they collaborate in terms of higher-order cognitive and social processes like negotiation or metacognitive monitoring. An existing vocabulary like the TLA profile can describe many of the tool-based actions (e.g., posted for submitting, interacted for commenting). However, it lacks the specific granularity to differentiate the nuanced interactions that CSCL research deems important.

4. Discussion

This study investigated the feasibility of applying a standard xAPI vocabulary to collaborative learning logs, revealing a significant gap between the focus of academic research and the realities of tool-based data collection. This discussion explores the implications of this gap and proposes a pragmatic path forward.

4.1 The Gap Between Collaborative Learning Theory and Practice

Our findings show that while existing educational tools in the Japanese K-12 context log a variety of user actions, these are primarily primitive, task-oriented behaviors (e.g., editing, submitting). In contrast, CSCL research focuses on higher-order processes like meaning negotiation and metacognitive monitoring. This discrepancy is a key challenge for data-driven collaborative learning analytics. Current tools do not explicitly capture the complex cognitive and social interactions that are theoretically most valuable. This may be because tools are designed to support basic classroom management and content delivery, where simple, observable actions are easier to implement and interpret for teachers. Capturing higher-order processes often requires more sophisticated logging mechanisms or even manual coding, which is not practical for everyday classroom use. This suggests that many higher-order collaborative processes may occur implicitly or go unlogged due to these platform limitations.

4.2 Implications for Data Standardization

This limitation, however, does not diminish the importance of standardization. On the contrary, it highlights the need for a foundational layer of interoperability. Before we can analyze complex interactions, we must be able to reliably compare basic actions across different platforms. For educational service providers, agreeing on a common way to describe fundamental actions like "submitting a card" or "commenting on a peer's idea" is a necessary and pragmatic first step. This study demonstrates that existing xAPI vocabularies provide a viable, albeit incomplete, starting point for this endeavor. This directly addresses the practical challenge that even primitive analysis is difficult in the current fragmented data landscape. Without a shared understanding of what submit means across platforms, any analysis of higher-order processes that follow a submission is built on unstable ground.

4.3 A Pragmatic Path Forward

The findings suggest a clear, pragmatic path for advancing collaborative learning analytics. Instead of immediately attempting to create a single, all-encompassing xAPI profile that captures every nuance of collaboration, the community—including researchers, developers, and educators—should first focus on standardizing a "core" set of fundamental, observable actions. This core vocabulary, derived from the common functions identified in our empirical analysis (e.g., edited, submitted, viewed, commented), would establish a baseline of interoperability. Once this foundation is in place, more complex, context-specific layers can be

added to capture higher-order processes. This layered approach ensures that basic cross-platform analysis becomes immediately possible, while providing a stable framework upon which more sophisticated analytics can be built in the future. This approach allows for both immediate practical benefits and long-term research advancement.

4.4 Limitations and Future Work

This study has several limitations that suggest future research directions. First, its focus on Japanese K-12 educational tools limits the generalizability of the findings to other contexts where different pedagogical approaches and tools may be used. Second, by relying on existing log data and pre-defined verbs, we could not capture the full richness of higher-order collaborative behaviors that were not explicitly logged. The use of pre-defined verbs may have constrained the analysis, preventing the identification of more nuanced, context-specific actions.

Future work should follow the pragmatic path outlined above, including broader, multi-national, cross-platform studies to validate and expand the proposed “core” vocabulary. Research should also explore learning activity designs and tool features that make higher-order processes more observable and loggable, such as structured debates or peer-review workflows. Examining teacher dashboards or AI-driven feedback systems using this standardized core data would be valuable, potentially demonstrating the practical benefits of this foundational approach.

5. Conclusion

This study examined the alignment between collaborative learning theory and practice through the lens of data standardization. We found a clear gap between the higher-order interactions valued in research and the fundamental actions logged by widely used educational tools. While a comprehensive analysis of collaboration is not yet feasible with current logs alone, establishing a standardized vocabulary for foundational actions is a critical first step. Providing a common data language enables more robust, cross-platform analytics. This pragmatic approach will improve the utility of learning analytics for educators and pave the way for future research to capture more complex dynamics. Bridging the gap between theory and practice requires building from a solid, interoperable foundation—one action at a time.

References

Benesse Corporation. (2025). Miraiseed. https://bso.benesse.ne.jp/miraiseed/index.html

codeTakt Inc. (2025). schoolTakt. https://schooltakt.com/

Lang, C., Siemens, G., Wise, A., Gasevic, D., & Merceron, A. (Eds.). (2022). Handbook of learning analytics (2nd edition). SOLAR, Society for Learning Analytics and Research New York, NY, USA.

LoiLo Inc. (2025). LoiLoNote School. https://n.loilo.tv/ja/

Manso-Vázquez, M., Llamas-Nistal, M., & Mikic-Fonte, F. A. (2018). An xAPI Application Profile to Monitor Self-Regulated Learning Strategies.

Rosé, C. P., Fischer, F., Järvelä, S., & Weinberger, A. (Eds.). (2021). The international handbook of computer-supported collaborative learning. Springer.

Sawyer, R. K. (Ed.). (2014). The Cambridge handbook of the learning sciences (2nd ed.). Cambridge University Press.

Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjón, B. (2016). Applying standards to systematize learning analytics in serious games. Computer Standards & Interfaces.

Appendix

You can view the bibliographic information of the final set of 72 peer-reviewed articles for Theoretical Analysis (RQ1) at the following URL.
https://drive.google.com/file/d/1v2-fL1AGeRodU24KkJJhDAvF8Q8SMQng/view?usp=sharing