JOURNAL OF LEARNING ANALYTICS

SOLAR SOCIETY for LEARNING ANALYTICS RESEARCH

Early Access Articles, 1–16. https://doi.org/10.18608/jla.2024.8367

Edulyze: Learning Analytics for Real-World Classrooms at Scale

Prasoon Patidar[1], Tricia J. Ngoon[2], Neeharika Vogety[3], Nikhil Behari[4], Chris Harrison[5], John Zimmerman[6], Amy Ogan[7], Yuvraj Agarwal[8]

Abstract

Classroom sensing systems can capture data on teacher-student behaviours and interactions at a scale far greater than human observers can. These data, translated to multi-modal analytics, can provide meaningful insights to educational stakeholders. However, complex data can be difficult to make sense of. In addition, analyses done on these data are often limited by the organization of the underlying sensing system, and translating sensing data into meaningful insights often requires custom analyses across different modalities. We present Edulyze, an analytics engine that processes complex, multi-modal sensing data and translates them into a unified schema that is agnostic to the underlying sensing system or classroom configuration. We evaluate Edulyze's performance by integrating three sensing systems (Edusense, ClassGaze, and Moodoo) and then present data analyses of five case studies of relevant pedagogical research questions across these sensing systems. We demonstrate how Edulyze's flexibility and customizability allow us to answer a broad range of research questions made possible by Edulyze's translation of a breadth of raw sensing data from different sensing systems into relevant classroom analytics.

Notes for Practice

- Classroom sensing technologies are instrumental in systematically capturing the dynamics between instructors and students. Each system has a distinct data structure, making integrating and extending new sensing systems complex.
- Edulyze introduces an analytics platform that processes multi-modal sensing data and translates it into a unified schema that is agnostic to the underlying sensing system.
- Edulyze serves as a foundational tool for pedagogy researchers and instructors. It allows them to bridge various data streams and explore correlations among different metrics of teaching practices, student engagement, and learning outcomes.

Keywords

Large-scale analytics, classroom sensing, research tools, learning pedagogy research

Submitted: 09/02/2024 — **Accepted:** 14/07/2024 — **Published:** 07/08/2024

*Corresponding author[1] Email: prasoonpatidar@cmu.edu, Address: School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, Pennsylvania, 15213, USA. ORCID: https:// orcid.org/ 0000-0003-0034-2767

[2] Email: tngoon@andrew.cmu.edu, Address: School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, Pennsylvania, 15213, USA. ORCID: https:// orcid.org/ 0000-0002-3749-0897

[3] Email: nvogety@gmail.com, Address: School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, Pennsylvania, 15213, USA.

[4] Email: nikhilbehari@gmail.com, Address: MIT Media Lab, 75 Amherst St., Cambridge, Massachusetts, 02139, USA.

[5] Email: chris.harrison@cs.cmu.edu, Address: School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, Pennsylvania, 15213, USA. ORCID: https:// orcid.org/ 0000-0001-5312-3619

[6] Email: johnz@cs.cmu.edu, Address: School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, Pennsylvania, 15213, USA. ORCID: https:// orcid.org/ 0000-0001-5299-8157

[7] Email: aeo@cs.cmu.edu, Address: School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, Pennsylvania, 15213, USA. ORCID: https:// orcid.org/ 0000-0003-2671-6149

[8] Email: yuvraj@cs.cmu.edu, Address: School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, Pennsylvania, 15213, USA. ORCID: https:// orcid.org/ 0000-0001-9304-6080

1. Introduction

Instructor behaviours, both verbal and non-verbal, contribute to student engagement and learning. When students feel noticed and cared for, they pay more attention and try harder at learning (Woolfolk & Brooks, 1985; Fredricks et al., 2004). Instructors often know that both their verbal and non-verbal behaviours impact students; however, they have no easy way to maintain an awareness of these behaviours while they are teaching, especially because they are ephemeral and difficult to recall.

While classroom observation and self-reports are most widely used in research to study the impact of instructor behaviours on student engagement (Andersen et al., 1979), these labour-intensive and often inaccurate approaches do not scale. Ubiquitous computing offers one path forward. It can serve as extra sets of eyes and ears, capturing multi-modal data on instructor and student actions, facial expressions, gestures, gaze, and speech at a scale and precision far beyond the capability of human observers (Ahuja et al., 2019; Martinez-Maldonado, Mangaroska, et al., 2020; Martinez-Maldonado, Echeverria, Schulte, et al., 2020; Sümer et al., 2023). Furthermore, using video capture tools for professional development allows educators to revisit and reflect on specific teaching moments, enhancing their self-efficacy and thus enabling a more deliberate practice of effective teaching strategies (Fong et al., 2019). Ubiquitous computing can capture the interplay between instructors' actions and student reactions, providing rich data for instructor self-reflection to share with a community of practice and providing insights for further pedagogical research (Alzoubi et al., 2021; Li et al., 2010; Oviatt, 2018). Some existing systems track instructor and student physiological data (Gao, Marschall, et al., 2022), interaction with classroom materials (Saquib et al., 2018), and instructor gaze and movement around the classroom (Ahuja et al., 2019; Martinez-Maldonado, Echeverria, Schulte, et al., 2020).

While incorporating these powerful and ubiquitous sensing systems in the classroom presents a great opportunity for understanding how teacher-student interactions can influence learning outcomes, there are two major challenges in deriving value from these types of instructor-student sensing systems that researchers could address. The first is gleaning insights from low-level sensing system data. While multi-modal sensing systems are ubiquitous and can collect a breadth of data types, making sense of data from multiple sources is challenging for instructors (Oviatt, 2018; Yan et al., 2022). The second challenge is organizing the diverse data and building around the context of classroom interactions. Data collection and analysis from sensing systems remains complex (Cukurova et al., 2020; Oviatt, 2018; Ochoa et al., 2016). Typically, sensing systems collect data over single interaction modalities (e.g., Moodoo (Martinez-Maldonado, Echeverria, Schulte, et al., 2020) collects location data, and n-Gage (Gao et al., 2020) captures physiological data like heart rate). These sensing systems can provide a richer set of insights if combined, since engagement and class activity involve multiple modalities interacting with each other; e.g., combining data from Moodoo (instructor's location) and n-Gage (students' physiology) can help in understanding how an instructor's use of classroom space impacts students' engagement. However, each sensing system employs a unique data structure, preventing different systems from integrating data and easily adding new sensors and sensor stream data (Foster et al., 2024). These challenges motivate our research questions:

RQ1: How might we organize diverse and heterogeneous sensing system data across different systems and modalities?

RQ2: How might researchers use this aggregated data to answer relevant research questions about instructor-student behaviours and engagement?

In this paper, we present Edulyze[1], which provides a unified schema to structure processed classroom data across various behaviour modalities. Motivated by prior work in instructors' data interests and relevant data in understanding student engagement, we developed a set of case studies that show how Edulyze's data and visualizations might be used to answer questions about *instructor behaviours*, *student behaviours*, and the *interaction between instructor and student behaviours*. These analytics are built into Edulyze, which is configurable, modular, and extensible by design and allows researchers to add additional behavioural modules or extend the output schema without requiring customization of the underlying Edulyze architecture. To demonstrate Edulyze's performance and usability, we collected and analyzed data from three classroom sensing systems, Edusense (Ahuja et al., 2019), ClassGaze (Ahuja et al., 2021), and Moodoo (Martinez-Maldonado, Echeverria, Schulte, et al., 2020), to demonstrate (a) the breadth and diversity of the features and data created by Edulyze and (b) the simplicity by which these data can be analyzed and visualized into insights in Edulyze to help answer these questions and pose new ones. These analytics allow educational stakeholders, such as researchers, to investigate pedagogically relevant questions.

Notably, Edulyze is agnostic to underlying sensing frameworks and customizable without the need to develop custom scripts for analyzing various modalities of data. This independence from sensing frameworks helps in two ways. It allows systems researchers to focus on building better classroom sensing frameworks without worrying about the modality of data presented to instructors, and it lets pedagogy researchers use off-the-shelf sensing systems and extend Edulyze to push the boundaries of learning sciences. Our goal with the design of Edulyze is to enable the concept of the quantified self or personal informatics

[1] Edulyze Repository: https://github.com/edusense/edulyze

ISSN 1929-7750 (online).

into teaching, which would enable instructors, researchers, and other stakeholders to answer ever-increasing questions about the effect of teaching behaviours and practices on student learning and engagement. We make the following contributions:

- We present Edulyze, a system that combines data from various classroom sensing systems to generate analytical insights about classroom behaviours and presents a unified schema to structure processed classroom data.
- We implemented the Edulyze system and evaluated its performance with three state-of-the-art classroom sensing systems (Edusense, ClassGaze, and Moodoo) to show that Edulyze can efficiently process heterogeneous data across a variety of sensing modalities.
- Finally, we provide open-source resources and present five case studies to show how researchers might use Edulyze to analyze and visualize multi-modal classroom analytics to further learning science.

2. Related Work

The research challenge in measuring student engagement is that verbal and non-verbal behaviours are fleeting. Prior research on the impact of teaching behaviours on student engagement traditionally relied on human observations and self-reports (Andersen et al., 1979). However, human observation is time-consuming and reliant on the availability of observers, and self-reports can be inaccurate and subject to bias (Martinez-Maldonado, Schulte, et al., 2020; Pryor et al., 1977). Classroom sensing systems can not only increase the temporal fidelity of observation of classroom behaviours but also capture a wide variety of relevant behavioural metrics. Wearable sensing systems use inexpensive, lightweight sensors worn in the classroom to measure physiological metrics or movement. For example, wearable devices such as wristbands and headsets can measure heart rate and other physiological proxies for student engagement (Gao et al., 2020) or predict teacher activity and classroom orchestration (Prieto et al., 2016). Some wearable sensing systems, such as Moodoo (Martinez-Maldonado, Echeverria, Schulte, et al., 2020) and Sensei (Saquib et al., 2018), can track individual movements to provide information about how instructors move about their classrooms and interact with students. Audio and camera-based systems have the benefit of wide-area sensing without the need for wearing sensors. For instance, EduSense (Ahuja et al., 2019) uses a pair of cameras affixed to opposite walls to collect various multidimensional classroom data. Such data can be used to monitor classroom behaviours such as gaze (Ahuja et al., 2021; Mangaroska et al., 2018), posture (Karen Chen & Gerritsen, 2021), students' hand raises (Ahuja et al., 2021), and facial expressions (Sümer et al., 2023) in static, inflexible classrooms such as lecture halls.

These multi-modal classroom analytics from sensing systems can provide instructors with a nuanced picture of fleeting student and instructor behaviours for further reflection and improvement (Martinez-Maldonado, Mangaroska, et al., 2020; Rodríguez-Triana et al., 2016; Martinez-Maldonado, Schulte, et al., 2020; Cukurova et al., 2020). However, numerous challenges remain in practice. Specifically, many multi-modal analytics systems are not sustainable or scalable because they require a large amount of reconfiguration in order to develop usable analytics (Yan et al., 2022). In addition, unifying data from multiple sensing systems is challenging due to the need for synchrony across data sources (Foster et al., 2024). As a result, these multi-modal analytics systems are learning-design specific (e.g., Ochoa et al., 2018, for small-group presentations) and not adaptable to new contexts (Yan et al., 2022). In particular, the complexity of low-level multi-modal data can make it difficult to synthesize data into meaningful insights (Cukurova et al., 2020). While multi-modal analytics may be most useful when captured by a variety of sensing systems and sources, data collection and analysis are often built upon custom code or for specific contexts (Oviatt, 2018), leading to challenges in unifying and synchronizing data across different data sources (Foster et al., 2024). Additionally, a recent deployment study of multi-modal learning analytics (MMLA) in the classroom reported that teachers suggested a detachable architecture and a desire for minimal technical support in maintaining MMLA systems (Martinez-Maldonado et al., 2023), which shows a need for more flexible systems for ease of use and analysis. Edulyze presents a breadth of analytics data that relates to both spatial pedagogy and non-verbal behaviours that can help instructors and other stakeholders gather insights about the relationship between these behaviours and student engagement.

3. System Design

Two key goals of Edulyze are to enable structured processing of raw sensing data received from classroom sessions and to derive higher-order insights about the classroom environment and instructor/student behaviours. To better understand what instructors wished to know about their non-verbal teaching behaviours, we conducted semi-structured interviews with eight college instructors who were assigned to teach in instrumented classrooms. We introduced them to the concepts of classroom proxemics, gaze, and immediacy. Next, we probed them for insight into which non-verbal behaviours they found interesting. Instructors expressed interest in viewing their own and their students' non-verbal behaviours, as well as the correlation between the two. Specifically, they highlighted the importance of student data related to engagement, attentiveness, and focus, as indicated by gaze and body positions. Instructors also wanted detailed insights into their own behaviours, including how much

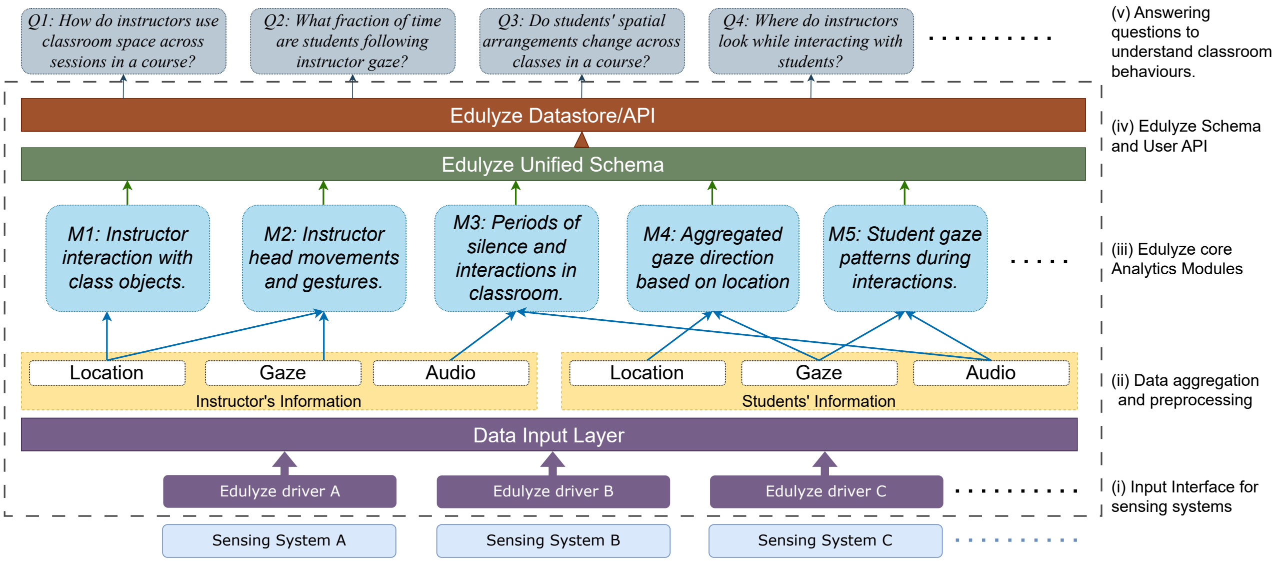

Figure 1. High-level overview of Edulyze, highlighting the key components. (i) Classroom sensing systems provide input to Edulyze, which is (ii) aggregated, preprocessed, and restructured in location, gaze, and audio information on instructor and students. (iii) This information is processed by Edulyze analytics modules to generate classroom analytics and (iv) stored as a unified Edulyze schema. Data from the Edulyze schema can be fetched using Edulyze API and (v) used to answer pedagogically relevant questions by various stakeholders, such as instructors, pedagogy researchers, and university administrators.

time they spent in different classroom areas and activities and how their gaze patterns affected student engagement. Based on our interviews and findings from prior work (Ahuja et al., 2019; Martinez-Maldonado, Echeverria, Schulte, et al., 2020), we conclude that gaze, location, and audio data of instructors and students are important for measuring instructor and student behaviours in classrooms. We also observe that a wide variety of prior classroom sensing systems provide raw sensing data that can be translated to either gaze, location, or audio modality, making them compatible with Edulyze (more details in Section 3.1).

Figure 1 illustrates the high-level architecture of Edulyze. It consists of four key components: (i) *Data Input* layer for interfacing with classroom sensing systems to collect, preprocess, and restructure data; (ii) *Edulyze Analytics* layer, which comprises analytics modules to compute behavioural features using the preprocessed data from the Data Input layer; (iii) *Edulyze Schema* to represent the output from all analytics modules in a unified manner; and (iv) *Edulyze Datastore and API*, which provides a RESTful interface for posting data to the Edulyze datastore, from which it can be queried by various stakeholders to study pedagogy-related questions. In this section, we present each of these components in detail.

3.1 Data Input Layer

Classroom sensing systems use various modalities to collect timestamped data (see Section 2). For example, camera systems capture 5–30 FPS video, while audio systems sample microphones at 44 kHz (Ahuja et al., 2019). Systems measuring spatial information track instructor, student, and object locations instantly (Martinez-Maldonado, Echeverria, Schulte, et al., 2020). Systems determining posture measure changes in body part positions over time (Martinez-Maldonado, 2019). Attention-focused systems measure viewing angle changes (Ahuja et al., 2021). Audio systems capture amplitude data to infer interactions and activities (Donnelly et al., 2016; Schlotterbeck et al., 2021). Fundamentally, these systems output timestamped sensor values, enabling analysis of classroom environments and behaviours.

Based on these observations, we designed a unified format for taking input from a variety of sensing systems. This common representation comprises a list of (timestamp, *frame*) pairs, where the *frame* is defined as a set of key-value pairs containing information about people and objects in a classroom at a given time instant. *Location input* comprises location information (x,y,z) coordinates of instructors, students, and classroom objects (e.g., board, podium, projector wall). To incorporate people's posture information, we allow fine-grained location information to be specified for body key points (head, hand, knee, etc.) for each detected person in the frame. *Gaze input* comprises viewing angle information in terms of yaw (horizontal rotation), pitch (vertical tilt), and roll (sideways tilt) of the head for instructors and students. *Audio input* accepts raw audio signals across multiple audio sources (channels) with meta-information (e.g., sampling frequency, audio codec). To streamline the integration of various data inputs into a uniformly structured format, Edulyze provides a *DriverInterface* with implementation support. New sensing systems can seamlessly connect with Edulyze by constructing a single Python script, typically no more than a few lines of code, which converts their sensing system output to the required input for Edulyze. For illustration, past implementations, such as Moodoo, required approximately 250 additional lines, while Edusense needed about 350 lines for data preprocessing. Information in these formats from various data sources is compiled in a list of ordered, timestamped frames for instructor and students separately. As a pedagogy tool, Edulyze follows a configurable design to support various runtime optimizations. It is primarily designed for event-based execution, which can get triggered as soon as new classroom sessions are recorded and processed from underlying sensing systems. It can also be run manually on a single class session or multiple sessions, constituting an entire quarter- or semester-long course in bulk.

3.2 Edulyze Analytics Modules

Table 1. Set of core analytics modules developed for Edulyze.

Input Type	Analytics Module Description	Motivation and pedagogical relevance
Location	*Object Interaction*: Instructor interactions with objects in the classrooms (podiums, boards, etc.)	Cheong Yin Mei (2017) finds that instructor interaction with relevant objects in the classroom may influence teacher-student rapport.
	Movement Metrics: How much time does the instructor spend moving or standing still in a class?	Martinez-Maldonado, Mangaroska, and colleagues (2020) find that teacher stops and movement throughout the classroom can provide insights into how they interact with students during class.
	Principal Movement Direction: The directions and frequency of movements by the instructors (moving across board or podium, class centre or podium, etc.)	Instructors expressed an interest in understanding their principal movement over time (Xhakaj, 2021; White & Gardner, 2013).
	Top Locations: Top five locations in class for instructors	Where an instructor primarily moves about a classroom can reveal patterns in their spatial pedagogy (Martinez-Maldonado, Echeverria, Schulte, et al., 2020).
	Head/Body Movement: Instructor and student head/body movements and gestures	Student head movements and body positions can be an indication of engagement (Bosch et al., 2016; Andersen et al., 1979).
	Seating Arrangement: Student seating topography in class (facing front, in a circle, in small groups, etc.)	Where students sit in class may impact engagement and provide information to instructors about how to best engage their students (Xhakaj, 2021; Gao, Rahaman, et al., 2022).
Gaze	*Principal Gaze Direction*: Where instructors/students look most often	Gaze is an important metric of engagement and immediacy (Kuang et al., 2024). Gaze patterns can help instructors understand how they are using gaze within the classroom (Mangaroska et al., 2018).
	Prolonged Gaze: Where instructors direct their gaze for extended periods of time (at slides, students in front, etc.)	Instructors expressed interest in knowing how much they directed their gaze toward or away from students for periods of time. Instructors who look away from students for extended periods of time may be seen as more distant (Kuang et al., 2024; White & Gardner, 2013).
	Location-Specific Patterns: How instructor/students' gaze patterns vary based on where they are standing	Instructors expressed interest in knowing how their positioning affected student gaze (Xhakaj, 2021). Mutual gaze patterns are also indicators of student engagement and learning (White & Gardner, 2013, 2013).
	Aggregated Gaze Metrics: Overall, where the majority of students are focusing (at instructor location, looking down, toward board, etc.) in class	Because gaze might serve as a proxy for attention (Sümer et al., 2023), instructors wanted to know how their own gaze affected students' gaze as a proxy for attention.
Audio	*Silent vs. Non-silent Periods*: Detecting short and long periods of silence between lectures or discussions in class	These data can show how much verbal interaction and back-and-forth occurs between instructors and students (Gerritsen, 2018; Sedova et al., 2019).

Motivated by prior work in learning research, we selected a set of analytics modules to answer pedagogically relevant questions about *instructor behaviours*, *student behaviours*, and *interactions between student and instructor behaviours*. Previous explorations of classroom sensing systems for instructors show that instructors expressed interest in understanding their own teaching habits, such as egalitarian gaze, question asking, and engagement with students in different parts of the classroom (Xhakaj, 2021; Alzoubi et al., 2021). Instructors also expressed interest in understanding student behaviours in response to their teaching behaviours to see the effectiveness of different teaching strategies and practices (Xhakaj, 2021; Alzoubi et al., 2021). Student behaviours such as hand raise (Si et al., 2019), head movement and posture (Sümer et al., 2023), and student gaze (Ahuja et al., 2021) can serve as proxies for student engagement and learning. Allowing various kinds of classroom sensing systems introduces heterogeneity in data rates across various types of raw information. For example, Moodoo (Martinez-Maldonado, Echeverria, Schulte, et al., 2020) captures location information every second, whereas EduSense (Ahuja et al., 2019) extracts location information at 2–5 seconds of granularity based on the number of people in the classroom. One way to address these challenges is to aggregate information across time periods and use these results to study spatiotemporal trends. Inspired by prior work on manual classroom observation protocols (e.g., COPUS; Smith et al., 2013), we divide each session into a set of two-minute time periods, called blocks, with all data from start to end of the block processed at the block level. We use this derived information to get higher-order insights across blocks in a session or across different sessions for the same course, which in turn can be used to compare behaviours across different courses. Table 1 summarizes the set of analytics modules we developed for gaze, location, and audio modalities and also presents the motivation behind the development of these modules from prior literature.

Location Modules We implemented six analytics modules using raw location data from instructors, students, and classroom objects (see Table 1). We classified instructors' head interactions with boards or podiums based on proximity within one and a half object lengths to an object's bounding box, tailored to accommodate varying camera distances in classrooms. For instructor movement metrics, we used the method proposed by Martinez-Maldonado, Echeverria, Schulte, and colleagues (2020), which is grounded in the notion of spatial pedagogy and defines instructor stop as a sequence of positioning data points that are a short distance apart in space and time (Lim et al., 2012). These interactions and movements are aggregated at block levels as proportions of time spent interacting versus moving. Principal movement directions are analyzed by collecting point clouds of instructor locations during a block, followed by principal component analysis to determine movement dynamics with explained variance. We also analyze instructors' and students' head and body movements by tracking absolute pixel changes and calculating minimum, maximum, and average movement metrics over a session block. This captures overall patterns and allows for analysis of student engagement through their movements and use of space (Bosch et al., 2016). Lastly, we assess classroom transitions by evaluating how quickly students settle and use silhouette analysis and k-means clustering to analyze seating patterns, determining student clusters and their classroom distribution per block.

Gaze Modules We implemented four gaze-specific analytics modules (see Table 1), using raw gaze data (roll, pitch, yaw) and head locations of instructors and students. The back camera captures the instructor's gaze, while the front camera records student gazes. Modules like principal gaze direction and prolonged gaze derive directly from these raw data, providing insights into attention patterns. For aggregate gaze metrics, we assess the focal points of the classroom collective per session. Due to differing camera perspectives, direct comparison of raw gaze angles between instructors and students is impractical. To address this, we compute categorical gaze directions (up/down, left/right, front/back) from raw values for a broader comparison of gaze interactions. These data are aggregated by block for instructors, individual students, and groups, enhancing our understanding of how gaze patterns correlate with the instructor's location throughout the session.

Audio Modules We developed an audio preprocessing pipeline that converts raw audio from multiple channels into Mel-frequency Cepstral coefficients (MFCC) and Mel spectrograms for chunks of 4,000 samples, i.e., 250 ms converted into the frequency domain across 128 bins, providing 4×128 values per second. Audio data can provide rich information on classroom behaviours. However, we encountered notable challenges in using detailed audio features, largely due to widespread environmental noise and ethical issues related to processing speech information. Our empirical experiments also confirmed that ambient audio lacks the robustness needed to provide a solid foundation for deriving dependable engagement metrics. As a result, we simplified our approach to primarily concentrate on detecting silence and speech while preserving a structure that supports the addition of more analytic modules for future research projects. For silence detection, we use featurized audio to aggregate spectrogram data at block levels and clustering to identify periods of low amplitude across all frequency bins. This block-level clustering adapts to different classroom noise backgrounds, and post-processing merges these to identify silence periods. We tested this with 15-minute segments from five diverse sessions, manually verifying silence at one-second intervals. Our module demonstrated 85% accuracy, 75% precision, and 70% recall in detecting silence, based on analysis of approximately 4,500 data points.


```
{
    "id": "4f793e67-7f98-4ba6-9a9c39e9a6ae",
    "keyword": "CS13_BG10_20201001_123000",
    "metaInfo": {
        "pipelineVersion": 2.1.0,
        "analysisStartTime": 1652198965,
        "totalRuntime": 127.6,
         ...
    },
    "debugInfo": {},
    "secondLevel": [{
        "secondInfo":{},
        "audio":{},
        "gaze":{},
        "location":{},
    },...],
    "blockLevel": [{
        "blockInfo":{},
        "audio":{},
        "gaze":{},
        "location":{},
    },...],
    "sessionLevel": {
        "audio":{},
        "gaze":{},
        "location":{},
    }}
```

Figure 2. Top level of Edulyze schema showing data aggregation for different sensing modalities (audio, gaze, and location) at different granularities for a single session.

```
{
  "blockInfo": {
  "unixStartSecond": 16522000165,
  "blockLength": 120,
  "id": 10},
 "audio": {
  "silenceFraction": 0.8,
  "teacherActivityType": "lecture",
  ...},
 "gaze": {
  "instructor": {
   "gazeDirection": 124.5,
   "principalGaze": {
    "direction": "left",
    ...},
  ...},
  "student": {}
 },
 "location": {
  "instructor": {},
  "student": {
   "ids": ["S5", "S8", ...],
   "isSettled": [true, true, ...],
   "maxHeadEntropy": [128.5,180.9,...],
   "seatingArrangement": "front_facing",
  ...}
 }}
```

Figure 3. A snippet of block-level information from the Edulyze schema showing detailed output(s) from audio, gaze, and location modules.

3.3 Edulyze Schema

One of the challenges with learning pedagogy research is the heterogeneity in the organization of output data from various sensing and processing modules. As a part of Edulyze, we created an extensible schema to represent session-level analytic data from classrooms. We characterized the output data generated from the various analysis modules in Edulyze to create a comprehensive and unified schema, written and validated using the Schema library in Python. Edulyze's schema organizes output from all analytics modules into a hierarchical tree format. Information at each level is stored so that partial queries can be made. For example, Edulyze allows access to different parts of sessions individually rather than requiring all of the session data to be accessed in one go, which is computationally expensive and considerably slower. Based on our current block-based design, we finalized three different categories of information aggregations: (i) the *second level*, which is a proxy for instantaneous features, incorporating raw information within a second; (ii) the *block level*, which is a way to broadly understand what is happening in different parts of a session; and (iii) the *session level*, which contains information compiled for complete sessions (see Figures 2 and 3). At each aggregation level, information is further separated based on the type of observation (or different behaviour modalities), such as gaze- or audio-related information (see Table 1). Once all of the analysis modules have been executed, Edulyze restructures the results from each of them into a predefined schema and runs a schema validation check to ensure that no unexpected data are posted to the final schema. Any schema updates due to changes in the output type or the addition of new modules are allowed with version updates in Edulyze, and our datastore server (see Section 3.4) can store different versions together and allow querying from runs from any of the pipeline versions.

3.4 Edulyze Datastore and APIs

The Edulyze datastore, implemented in GoLang, uses MongoDB, a NoSQL database, to maintain all processed data following the Edulyze schema, including analytic outputs. We provide a RESTful API for data insertion by processing and analysis modules and a GraphQL API for complex, efficient data queries. Each classroom session is uniquely identified by a course code, semester, date, time, and location and stored in individual MongoDB collections for streamlined search and organization.

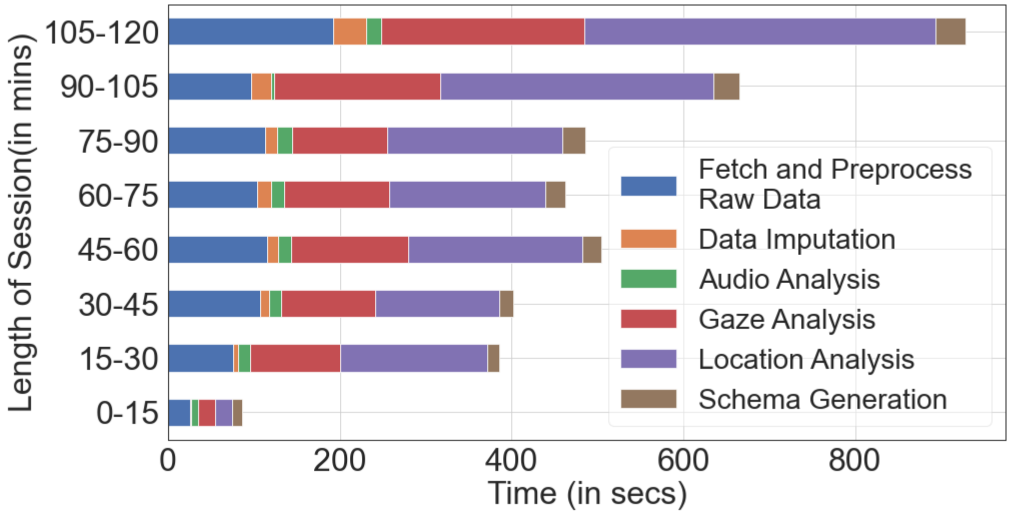

Figure 4. Time taken by various components in Edulyze over different session lengths.

Basic access control is implemented to restrict data access to authorized individuals only, enhancing security. We opted for GraphQL over the usual REST API due to its ease of use and expressive querying capabilities, allowing specific field requests in a single endpoint. This approach avoids the inefficiency of REST APIs, which either deliver complete data sets or need multiple endpoints for different data slices, thus optimizing computational efficiency in fetching analysis-specific data.

4. Implementation and Evaluation

We implemented Edulyze in Python (see footnote 1, page 2), integrating a wide variety of behavioural modules for classroom learning analytics. We focus on enabling learning pedagogy research on classroom data collected from a number of classroom sensing systems with minimal programming efforts using a general-purpose, high-level language. In order to support multiple Edulyze pipelines executing in parallel and provide flexibility of single or multiple (if needed) storage servers, our Edulyze datastore and its API are implemented separately in GoLang.

We created input driver interfaces to translate the output of different sensing systems to the Edulyze input format. Using our input driver interface, we created drivers for three different sensing systems, (a) EduSense (Ahuja et al., 2019), which uses OpenPose's vision library to extract instructor/student location information from a raw image frame; (b) ClassGaze (Ahuja et al., 2021), a gaze-tracking module developed based on face localization and 3D Dense Face alignment in images; and (c) Moodoo (Martinez-Maldonado, Echeverria, Schulte, et al., 2020), which uses sensors worn by instructors to track their location through classroom sessions.

We execute the Edusense and ClassGaze pipelines on data collected from our sensing system apparatus, which consists of a two-camera setup affixed to face the front and the back of the classroom that records independently to capture instructor and student data simultaneously. For Moodoo, we used the open-source data collected over 18 laboratory sessions across a single lab course, which consists of location information (x, y coordinates) for instructors throughout sessions. We hand-coded the location of the whiteboard and the podium relative to raw location data for all classrooms to enable modules that give semantically meaningful interpretation rather than just coordinates (e.g., instructor at board vs. the podium).

In Figure 4, we characterize the processing time (in seconds) taken by the various components of the Edulyze analysis pipeline. Note that our current Edulyze pipeline is designed to use a single CPU core without any multi-threading. To benchmark Edulyze modules, we execute it on a 64-core machine configured with 100 GB of RAM. The Edulyze datastore server is also a virtual machine connected to the same network switch to provide data access to fetch frame-level data. Since Edulyze primarily aggregates and processes data from frames, we measure the average time it takes to analyze sessions of different durations. Figure 4 shows that as session length increases, the overall time to run the entire Edulyze pipeline also increases proportionally. Focusing on the time breakdowns among different Edulyze modules, the three that take the most time are "fetch and formatting video data," "gaze analysis," and "location analysis."

Table 2. Summary of course data collected with example courses from dual camera setup.

Course Type	No. of Courses	Example Courses in the Category
Applied	9	Seminar/Project Courses, SE For Startups, Designing Educational Games
Theory	11	Intro. to Anthropology, Microeconomics, Calculus, Modern Chemistry
Language	2	Elementary French, Fables and Stories from Chinese Civilizations

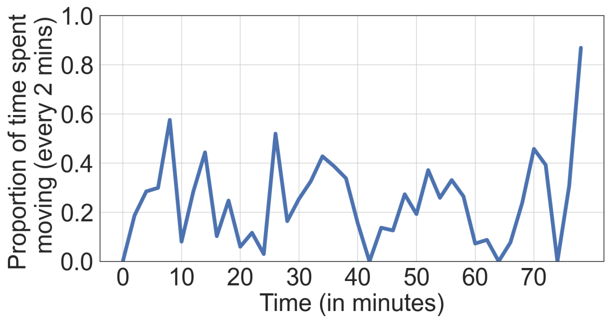

Figure 5. [Edusense] Changes in instructor movement over each block in a language class session.

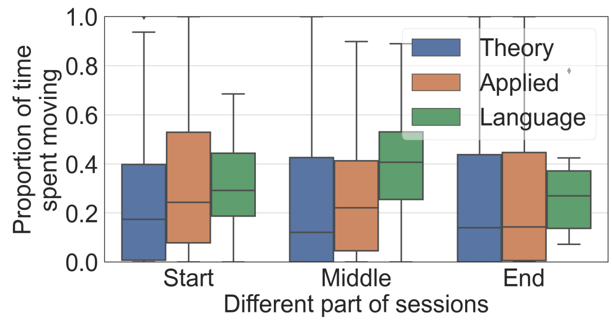

Figure 6. [Edusense] Comparison of teacher movement density over different types of courses, showing the distribution of movement intensity for different parts of classroom sessions across all sessions in the course.

5. Edulyze Use Cases for Classroom Analytics

We now present several exemplary research questions to demonstrate the breadth of data Edulyze can collect and process and the concrete insights that could be drawn from these data. These use cases are inspired by relevant research questions and are motivated by prior work, as stated in Section 3.3. After briefly reviewing our data collection process, we present how Edulyze enables analysis across various dimensions: short- and long-term temporal analyses (over single and multiple class sessions), spatial analyses (how teachers and students use classroom space), and analysis across courses (by course category or type).

For the camera-based sensing pipelines (Edusense and ClassGaze), we selected data from over 20 courses from three different semesters (spring, summer, and fall) at a mid-sized university in the United States. All data collected are from courses held prior to the COVID-19 pandemic. We carefully annotated these courses based on their descriptions and handpicked them from a pool of a total of 50+ courses from which we had video data to ensure diversity in course types, use of different physical classrooms, average session lengths, hours of student engagement, count of sessions, and count of students attending the course. To categorize course types, we used the course descriptions and course titles to label courses as “theory,” “applied,” or “language.” Table 2 shows a short summary of course data collected based on category and sample course titles. For Moodoo, we used their open-source dataset from 18 laboratory sessions from a single course (Martinez-Maldonado, 2020). Each lab exhibited one of three possible learning designs: (i) a prescribed lab, in which all students had to do the same experiment following a step-by-step guide; (ii) a project-based lab, in which students were asked to formulate a testing project; and (iii) a theory-testing lab, in which the teacher set up four or five experiments, and students had to move to one experiment at a time and predict the outcome of each without further guidance (Martinez-Maldonado, Echeverria, Schulte, et al., 2020). One benefit of analyzing data across different sensing systems is that we can ask the same research questions across different courses and contexts.

Understanding Instructor Movement in Class/Lab Sessions Instructor movement around a classroom can reveal patterns about their spatial pedagogy strategies (Martinez-Maldonado, Echeverria, Schulte, et al., 2020). Figure 5 shows an example of a single course session in which this particular instructor shows periodic movement patterns, i.e., low movement followed by high movement, which suggests that the instructor is switching between going over the lecture content and then engaging with students regarding the content. Similarly, Figure 7 shows an example of a single lab session in which the instructor shows high movement during the beginning of the lab session in comparison to the end of the lab sessions, which suggests that the instructor is helping students set up experiments and then letting them work on their own. Researchers may be interested in looking at instructor movement across different types of courses or lab sessions as well. Figure 6 shows the variation in teacher movement during different parts of a class session for a theory, an applied, and a language course. Figure 8 shows the variation in teacher movement over different kinds of lab sessions for a single course. From these data, a researcher might be able to draw

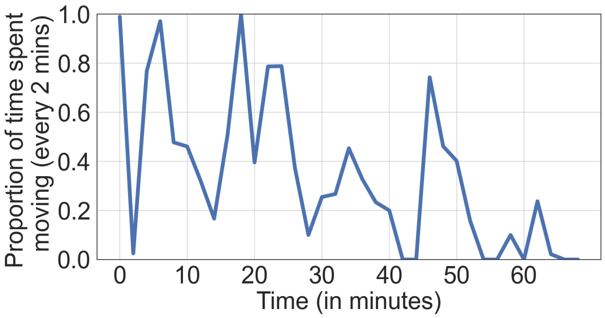

Figure 7. [Moodoo] Changes in instructor movement over each block in a theory lab.

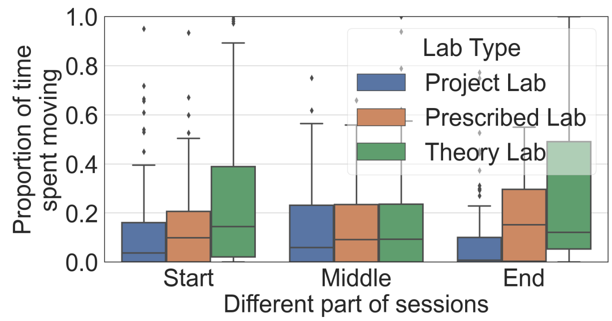

Figure 8. [Moodoo] Comparison of teacher movement density over different types of labs, showing the distribution of movement intensity for different parts of classroom sessions across different types of labs.

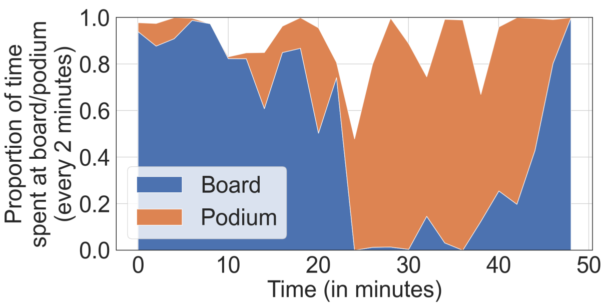

Figure 9. [Edusense] The proportion of time for each block in a class session that an instructor spends at the board/podium.

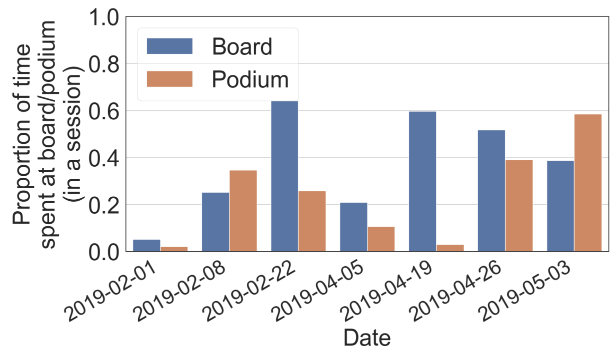

Figure 10. [Edusense] The proportion of time the instructor spends at the board/podium for multiple sessions of a course.

inferences about how course content might influence course activities and overall instructor movement in class. Instructors also mentioned interest in knowing how much time they spend in different parts of the classroom. Together, these data can help answer research questions about how instructors use the physical classroom space over time and across course activities in different class contexts and classroom configurations.

Interaction with Board/Podium in Class/Lab Sessions How instructors interact with relevant objects in the classroom (e.g., the board or podium) can influence instructor immediacy (Hesler, 1972). Instructors also mentioned interest in knowing how much time they spend at the board or podium in class. Figure 9 and Figure 11 show the time an instructor spends at the board or podium for a single class session and lab session, respectively. For this particular class session, the instructor spends most of the first half of class at the board and the second half at the podium, suggesting different types of teaching activities, such as writing on the board and then lecturing. For the lab session, the instructor spends very little time at the board or podium for most of the time, except at the start of lab sessions when students are settling down and during the middle of the session, which shows that most of the time is spent interacting with students. Figure 10 and Figure 12 show the proportion of time instructors spend at the board and podium across multiple sessions, which can give insights into the longitudinal distribution of class/lab activities. For a closer look at where instructors move throughout the classroom, Figure 13 shows a representation of principal movements in a classroom over multiple sessions to show an instructor's movement patterns. These data can help researchers and instructors draw inferences about an instructor's classroom orchestration and use of classroom objects.

Instructor Gaze Patterns Researchers may be interested in understanding student gaze patterns across different classroom contexts as well. Figure 14 shows the proportion of time an instructor directs their gaze at students during verbal interactions for a single class session. The figure suggests that the instructor predominantly looks at students while speaking. The time that the instructor looks away from students while speaking is brief, which suggests that eye contact with students is prevalent and may point toward the instructor's pedagogical strategy of maintaining direct gaze and eye contact with students. Variation across multiple class sessions may hint at different activities occurring during particular class sessions. For instructors' professional development, these data may provide an explanatory visualization for noticeable differences in student behaviour or engagement.

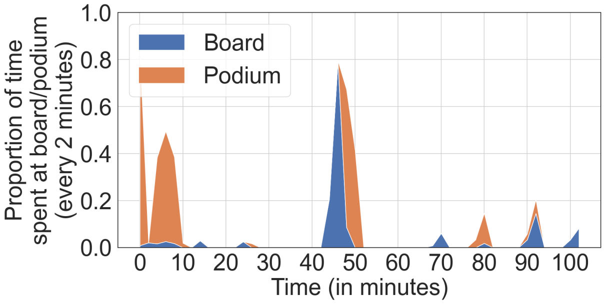

Figure 11. [Moodoo] The proportion of time for each block in a theory lab session that an instructor spends at the board/podium.

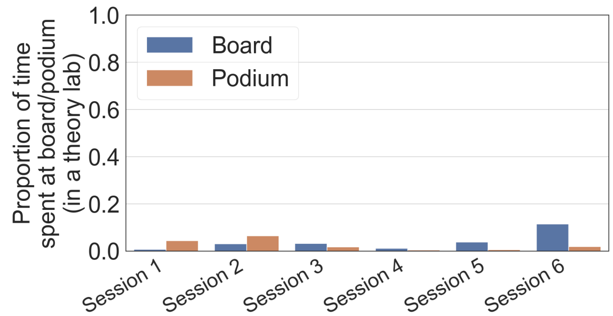

Figure 12. [Moodoo] The proportion of time the instructor spends at the board/podium for multiple theory labs.

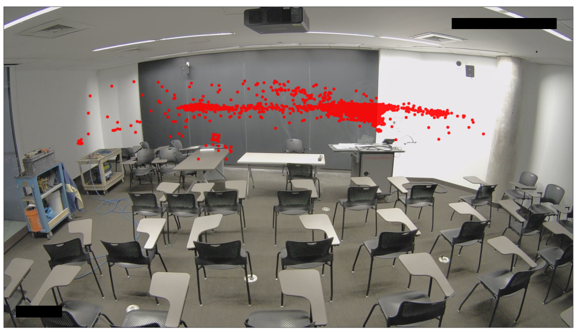

Figure 13. [Edusense] The principal movement patterns for the instructors across multiple class sessions in a course.

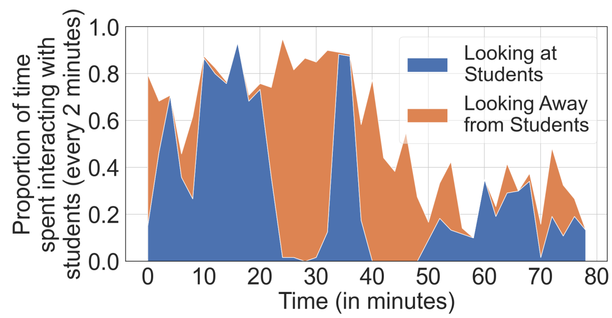

Figure 14. [ClassGaze] Instructor gaze patterns during student interactions in a single class session.

Student Gaze Patterns and Head Movement Student head movement and gaze can also be proxies for engagement and attention (Bosch et al., 2016; Sümer et al., 2023). As an example, Figure 15 shows the fraction of students who look down throughout a single class period. These data can reveal when most students are attentive during lecturing or when students might be synchronously taking notes. Instructors expressed interest in whether student gaze follows their gaze in the classroom. Mutual gaze patterns between teachers and students can show greater student engagement (White & Gardner, 2013). Figure 16 shows times when a majority of students are sharing this specific instructor's gaze during a single class session. This may help instructors understand where students are directing their attention and whether they are sharing joint attention (e.g., all students looking at slides).

Comparisons across Course Types Researchers may be interested in how instructor behaviours differ depending on the context of the class. Meta-data, such as the class type, can be included as part of Edulyze's input to analyze metrics according to these categories. For example, Figures 17 and 18 show the amount of time instructors in a theory, an applied, and a language course spend at the board and podium during different sections of the class. Researchers might draw conclusions about the type of class activities that instructors in each of these categories of classes might conduct based on these data. Instructors or professional development consultants might also use these data to understand how different instructors use different teaching strategies depending on the content of the course.

6. Discussion & Future Work

In addition to describing the implementation of our Edulyze system, we have also presented analyses for a set of example case study questions across three different sensing systems and different classroom contexts that may be of interest to pedagogical researchers. Here we discuss implications for future sensing and analytics systems.

6.1 Relevant Classroom Analytics for Educational Stakeholders

Edulyze bridges the gap in understanding classroom behaviours and their impact on learning and engagement through multi-modal data analysis (Worsley, 2018; Oviatt, 2018; Cukurova et al., 2020). This paper showcases a subset of research questions that can be explored using Edulyze, offering a starting point for researchers to correlate different data sources to learning outcomes and engagement. The platform's scope for this study enables multi-modal classroom data analyses to address pertinent research queries. The flexibility of Edulyze extends to different stakeholders who can harness the system's adaptability

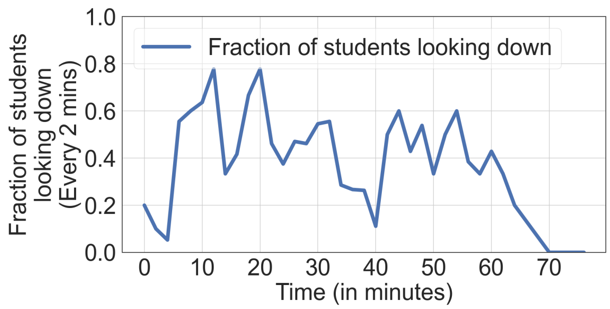

Figure 15. [ClassGaze] Fraction of students looking down in a single class session over time.

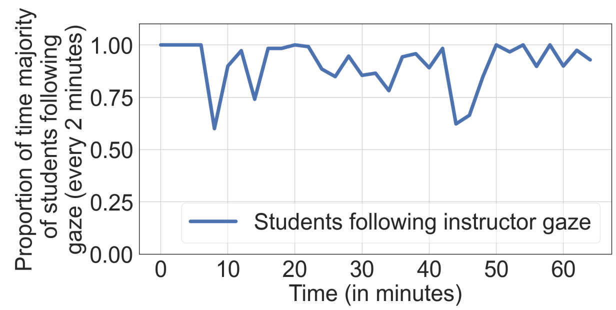

Figure 16. [ClassGaze] Fraction of time that a student's gaze is following the instructor's gaze.

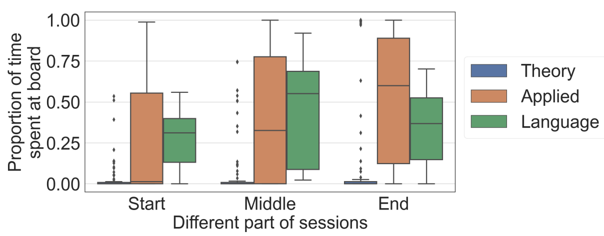

Figure 17. [EduSense] Comparison of instructors' board use over different types of courses, showing the distribution of instructors' use of the board at the start, middle, and end of a class across all sessions in the course.

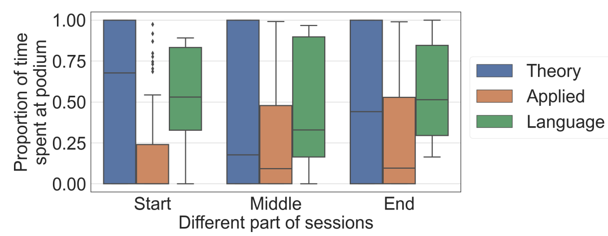

Figure 18. [EduSense] Comparison of instructors' time at the podium over different types of courses, showing the distribution of instructors' time at the podium at the start, middle, and end of a class across all sessions in the course.

to investigate a broader spectrum of inquiries. For instance, course type as meta-data could reveal how content influences classroom interactions. Stakeholders might integrate additional meta-data like student achievement or teacher evaluations to probe various research dimensions. School administrators could leverage these data for institutional queries such as optimizing space and classroom configuration (Vujovic et al., 2022) and understanding environments conducive to learning by merging student performance, physiological measures, and multi-modal data. While traditional classroom sensing systems provide pre-set metrics (Alzoubi et al., 2021; Martinez-Maldonado, Echeverria, Schulte, et al., 2020; Xhakaj, 2021), Edulyze's configurability allows instructors to adopt the role of researchers, defining their own variables to explore teaching strategies and their impacts on student engagement. This could include analyzing performance across multiple courses or testing different instructional approaches. Future work involves translating data into actionable insights, potentially developing "sandbox" interfaces for instructors to experiment with tailored analytics (Martinez-Maldonado, Echeverria, Nieto, & Buckingham Shum, 2020; Alzoubi et al., 2021; Xhakaj, 2021).

Through our case studies, we also demonstrate Edulyze's feasibility for collecting and processing multi-modal data from different classroom environments. One limitation of existing systems is that they are tied to the physical layout or context of the setting in which they are deployed (e.g., laboratory-based classrooms; Martinez-Maldonado, Mangaroska, et al., 2020; Martinez-Maldonado, Echeverria, Schulte, et al., 2020), limiting the generalizability of the analysis. Edulyze's analytics engine can process similar data outputs for a variety of courses in a variety of classroom layouts, as demonstrated by the several classes in which Edulyze was deployed. By collecting a variety of relevant data under a unified processing schema, researchers can draw associations and connections between several variables. Interested stakeholders could use Edulyze on top of other existing sensing systems to process and interpret whatever low-level data are available without the need to build an entirely custom analytics engine for that particular system.

6.2 Analytics for Understanding Conversational Dynamics

Our setting was pedagogical research within classrooms, but the core of Edulyze's data is conversational dynamics and interaction, both verbal and non-verbal. Researchers interested in understanding interactional behaviours in any conversational setting could benefit from Edulyze's analytics engine. Within the classroom, how teachers use silence as a strategy could give insight into in-class participation and discussion (Alerby, 2020). Teachers could use these analytics for equity goals related to distributed participation and attention for all students (Reinholz & Shah, 2018). Outside of the traditional classroom, these multi-modal analytics could be useful in professional learning, especially in professions where interaction with others is a vital aspect of success. For example, recent work has explored the use of multi-modal analytics in helping medical students improve their non-verbal behaviours (Liu et al., 2016) and psychotherapists improve dialogue with clients (Hirsch et al., 2018).

Edulyze's flexible and extendable analytics could allow interested stakeholders to correlate different types of data to answer interesting questions, promote reflection, and gain insights into professional performance. This is a ripe area for future work since the issues of scale present in professional development for education exist across domains as well.

6.3 Limitations and Future Work

We focus on Edulyze's analytics processing and the data it outputs for researchers interested in questions regarding classroom behaviour, student learning, and engagement. Edulyze provides human-interpretable insights across multiple dimensions, but users might still need some technical competence to process output from Edulyze based on their own needs. Future work could examine how to design visualizations for instructors or professional development professionals to make sense of multi-modal classroom analytics. Also, our current analytical focus within Edulyze has predominantly catered to traditional classroom setups where an instructor leads from the front and students are organized in rows or clusters. This configuration is prevalent and provides a straightforward framework for data collection and analysis by establishing a clear primary focal point. However, this configuration might not capture the full scope of interaction dynamics present in more informal or variably structured learning environments. Future research can focus on adapting the system to function effectively in diverse physical settings, thereby widening our understanding and enhancing the utility of learning dynamics analytics in various educational contexts. Our scope in this paper is audio, gaze, and location inputs, because these are the inputs most stated within our motivating prior work and based on the availability of open-source data from classroom sensing systems. Other inputs, such as physiological measures, could also provide rich data about engagement and learning. An area for future work could be incorporating these different data inputs, which would enable a broader set of research questions for analysis. Further, setting up the underlying sensing system can be resource-intensive (Ahuja et al., 2019), which might be a bottleneck in constrained settings. This paper presented a demonstration of Edulyze's capabilities and potential value for researchers and other stakeholders to analyze multi-modal classroom analytics. Future work could investigate what value these stakeholders find in Edulyze's analytics through empirical evaluations or further user interviews. Similarly, another area for future work is analytics for students and what data would be meaningful for them.

7. Conclusion

In this paper, we describe Edulyze, a novel MMLA system that provides meaningful insights to a variety of educational stakeholders. Edulyze is motivated by prior work in learning research and MMLA. It translates data from multiple sensing systems with a unified schema, allowing for customizable and configurable analytics without the need to create extensive analysis scripts and custom programs for each sensing system. We demonstrated Edulyze's utility through six real-world example case studies using data collected from 250 sessions across 25 courses. Through our analyses, we highlight the breadth and flexibility of Edulyze's data processing schema for enabling researchers and other educational stakeholders to make sense of raw classroom sensing data in real-world classrooms at scale.

Acknowledgements

The authors would like to thank other original contributors of Edusense (Ahuja et al., 2019) for guidance during the ideation process and development support for initial versions of the Edulyze system. We are grateful to Franceska Xhakaj and all of the students and instructors who participated in the data collection. We also thank Carnegie Mellon University's CyLab Security and Privacy Institute and JP Morgan Chase for their support on smart buildings research. Finally, we thank the anonymous reviewers for their constructive feedback, which helped improve the clarity of this paper.

Declaration of Conflict of Interest

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

Funding

The publication of this article received financial support from the James S. McDonnell Foundation and the National Science Foundation (NSF awards SaTC-1801472, IIS-1822813, IIS-1747997, CSR-1526237, and STEM Education Postdoctoral Fellowship #2222530).

References

Ahuja, K., Kim, D., Xhakaj, F., Varga, V., Xie, A., Zhang, S., Townsend, J. E., Harrison, C., Ogan, A., & Agarwal, Y. (2019). Edusense: Practical classroom sensing at scale. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *3*(3), 71. https://doi.org/10.1145/3351229

Ahuja, K., Shah, D., Pareddy, S., Xhakaj, F., Ogan, A., Agarwal, Y., & Harrison, C. (2021). Classroom digital twins with instrumentation-free gaze tracking. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (CHI 2021), 8–13 May 2021, Yokohama, Japan (p. 484). ACM. https://doi.org/10.1145/3411764.3445711

Alerby, E. (2020). The significance of silence in pedagogical settings. In *Silence within and beyond pedagogical settings* (pp. 43–50). Springer International Publishing. https://doi.org/10.1007/978-3-030-51060-2_4

Alzoubi, D., Kelley, J., Baran, E., B. Gilbert, S., Karabulut Ilgu, A., & Jiang, S. (2021). Teachactive feedback dashboard: Using automated classroom analytics to visualize pedagogical strategies at a glance. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (CHI 2021), 8–13 May 2021, Yokohama, Japan (p. 312). ACM. https://doi.org/10.1145/3411763.3451709

Andersen, J. F., Andersen, P. A., & Jensen, A. D. (1979). The measurement of nonverbal immediacy. *Journal of Applied Communication Research*, *7*(2), 153–180. https://doi.org/10.1080/00909887909365204

Bosch, N., D'mello, S. K., Ocumpaugh, J., Baker, R. S., & Shute, V. (2016). Using video to automatically detect learner affect in computer-enabled classrooms. *ACM Transactions on Interactive Intelligent Systems*, *6*(2), 17. https://doi.org/10.1145/2946837

Cukurova, M., Giannakos, M., & Martinez-Maldonado, R. (2020). The promise and challenges of multimodal learning analytics. *British Journal of Educational Technology*, *51*(5), 1441–1449. https://doi.org/10.1111/bjet.13015

Donnelly, P. J., Blanchard, N., Samei, B., Olney, A. M., Sun, X., Ward, B., Kelly, S., Nystran, M., & D'Mello, S. K. (2016). Automatic teacher modeling from live classroom audio. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization* (UMAP 2016), 13–17 July 2016, Halifax, Nova Scotia, Canada (pp. 45–53). ACM. https://doi.org/10.1145/2930238.2930250

Fong, C. J., Dillard, J. B., & Hatcher, M. (2019). Teaching self-efficacy of graduate student instructors: Exploring faculty motivation, perceptions of autonomy support, and undergraduate student engagement. *International Journal of Educational Research*, *98*, 91–105. https://doi.org/10.1016/j.ijer.2019.08.018

Foster, T. J., Justice, L., Villasanti, H. G., Irvin, D., & Messinger, D. (2024). Classroom sensing tools: Revolutionizing classroom-based research in the 21st century. *Topics in Early Childhood Special Education*, *0*(0), 02711214231220800. https://doi.org/10.1177/02711214231220800

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Gao, N., Marschall, M., Burry, J., Watkins, S., & Salim, F. (2022). Understanding occupants' behaviour, engagement, emotion, and comfort indoors with heterogeneous sensors and wearables. *Scientific Data*, *9*, 261. https://doi.org/10.1038/s41597-022-01347-w

Gao, N., Rahaman, M. S., Shao, W., Ji, K., & Salim, F. D. (2022). Individual and group-wise classroom seating experience: Effects on student engagement in different courses. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *6*(3). https://doi.org/10.1145/3550335

Gao, N., Shao, W., Rahaman, M. S., & Salim, F. D. (2020). n-Gage: Predicting in-class emotional, behavioural, and cognitive engagement in the wild. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(3), 79. https://doi.org/10.1145/3411813

Gerritsen, D. (2018, October). *A socio-technical approach to feedback and instructional development for teaching assistants* [Doctoral dissertation, Carnegie Mellon University]. https://doi.org/10.1184/R1/7195181.v1

Heng, B., Cheong, C., & Taib, F. (2017). Instructional proxemics and its impact on classroom teaching and learning. *International Journal of Modern Languages and Applied Linguistics*, *1*(1), 69–85. https://doi.org/10.24191/ijmal.v1i1.7637

Hesler, M. W. (1972). *An investigation of instructor use of space* [Doctoral dissertation, Purdue University]. https://www.proquest.com/openview/d0b8ef498d382cf5d2ba338d443730c6/1

Hirsch, T., Soma, C., Merced, K., Kuo, P., Dembe, A., Caperton, D. D., Atkins, D. C., & Imel, Z. E. (2018). "It's hard to argue with a computer": Investigating psychotherapists' attitudes towards automated evaluation. In *Proceedings of the 2018 Designing Interactive Systems Conference* (DIS 2018), 9–13 June 2018, Hong Kong, China (pp. 559–571). ACM. https://doi.org/10.1145/3196709.3196776

Karen Chen, L., & Gerritsen, D. (2021). Building interpretable descriptors for student posture analysis in a physical classroom. In I.-H. S. Hsiao, S. S. Sahebi, F. Bouchet, & J.-J. Vie (Eds.), *Proceedings of the 14th International Conference on Educational Data Mining* (EDM 2021), 29 June–2 July 2021, online (pp. 713–717). International Educational Data Mining Society. https://educationaldatamining.org/EDM2021/virtual/static/pdf/EDM21_paper_26.pdf

Kuang, Z., Wang, F., Andrasik, F., & Hu, X. (2024). Instructor's direct gaze not body orientation affects learning. *Journal of Computer Assisted Learning*, *40*(2), 731–741. https://doi.org/10.1111/jcal.12917

Li, I., Dey, A., & Forlizzi, J. (2010). A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI 2010), 10–15 April 2010, Atlanta, Georgia, USA (pp. 557–566). ACM. https://doi.org/10.1145/1753326.1753409

Lim, F. V., O'Halloran, K. L., & Podlasov, A. (2012). Spatial pedagogy: Mapping meanings in the use of classroom space. *Cambridge Journal of Education*, *42*(2), 235–251. https://doi.org/10.1080/0305764X.2012.676629

Liu, C., Calvo, R. A., & Lim, R. (2016). Improving medical students' awareness of their non-verbal communication through automated non-verbal behavior feedback. *Frontiers in ICT*, *3*. https://doi.org/10.3389/fict.2016.00011

Mangaroska, K., Sharma, K., Giannakos, M., Trætteberg, H., & Dillenbourg, P. (2018). Gaze insights into debugging behavior using learner-centred analysis. In *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (LAK 2018), 7–9 March 2018, Sydney, Australia (pp. 350–359). ACM. https://doi.org/10.1145/3170358.3170386

Martinez-Maldonado, R. (2019). "I spent more time with that team": Making spatial pedagogy visible using positioning sensors. In *Proceedings of the Ninth International Conference on Learning Analytics and Knowledge* (LAK 2019), 4–8 March 2019, Tempe, Arizona, USA (pp. 21–25). ACM. https://doi.org/10.1145/3303772.3303818

Martinez-Maldonado, R. (2020). Moodoo: Indoor positioning analytics for characterising classroom teaching. https://gitlab.erc.monash.edu.au/rmat0024/moodoo/-/tree/master/test/Merged%5C%20dataset%5C202018-2019

Martinez-Maldonado, R., Echeverria, V., Fernandez-Nieto, G., Yan, L., Zhao, L., Alfredo, R., Li, X., Dix, S., Jaggard, H., Wotherspoon, R., Osborne, A., Buckingham Shum, S., & Gašević, D. (2023). Lessons learnt from a multimodal learning analytics deployment in-the-wild. *ACM Transactions on Computer-Human Interaction*, *31*(1), 8. https://doi.org/10.1145/3622784

Martinez-Maldonado, R., Echeverria, V., Nieto, G. F., & Buckingham Shum, S. (2020). From data to insights: A layered storytelling approach for multimodal learning analytics. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI 2020), 25–30 April 2020, Honolulu, Hawaii, USA, 1–15. https://doi.org/10.1145/3313831.3376148

Martinez-Maldonado, R., Echeverria, V., Schulte, J., Shibani, A., Mangaroska, K., & Buckingham Shum, S. (2020). Moodoo: Indoor positioning analytics for characterising classroom teaching. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Proceedings of the 21st International Conference on Artificial Intelligence in Education* (AIED 2020), 6–10 July 2020, Ifrane, Morocco (pp. 360–373, Vol. 12163 of Lecture Notes in Artificial Intelligence). Springer. https://doi.org/10.1007/978-3-030-52237-7_29

Martinez-Maldonado, R., Mangaroska, K., Schulte, J., Elliott, D., Axisa, C., & Buckingham Shum, S. (2020). Teacher tracking with integrity: What indoor positioning can reveal about instructional proxemics. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(1), 22. https://doi.org/10.1145/3381017

Martinez-Maldonado, R., Schulte, J., Echeverria, V., Gopalan, Y., & Buckingham Shum, S. (2020). Where is the teacher? Digital analytics for classroom proxemics. *Journal of Computer Assisted Learning*, *36*(5), 741–762. https://doi.org/10.1111/JCAL.12444

Ochoa, X., Domínguez, F., Guamán, B., Maya, R., Falcones, G., & Castells, J. (2018). The RAP system: automatic feedback of oral presentation skills using multimodal analysis and low-cost sensors. In *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (LAK 2018), 7–9 March 2018, Sydney, Australia (pp. 360–364). ACM. https://doi.org/10.1145/3170358.3170406

Ochoa, X., Worsley, M., Weibel, N., & Oviatt, S. (2016). Multimodal learning analytics data challenges. In *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge* (LAK 2016), 25–29 April 2016, Edinburgh, UK (pp. 498–499). ACM. https://doi.org/10.1145/2883851.2883913

Oviatt, S. (2018). Ten opportunities and challenges for advancing student-centered multimodal learning analytics. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (ICMI 2018), 16–20 October 2018, Boulder, Colorado, USA (pp. 87–94). ACM. https://doi.org/10.1145/3242969.3243010

Prieto, L. P., Sharma, K., Dillenbourg, P., & Jesús, M. (2016, April). Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors. In *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge* (LAK 2016), 25–29 April 2016, Edinburgh, UK (pp. 148–157). ACM. https://doi.org/10.1145/2883851.2883927

Pryor, J. B., Gibbons, F. X., Wicklund, R. A., Fazio, R. H., & Hood, R. (1977). Self-focused attention and self-report validity. *Journal of Personality*, *45*(4), 513–527. https://doi.org/10.1111/j.1467-6494.1977.tb00168.x

Reinholz, D., & Shah, N. (2018). Equity analytics: A methodological approach for quantifying participation patterns in mathematics classroom discourse. *Journal for Research in Mathematics Education*, *49*(2), 140–177. https://doi.org/10.5951/jresematheduc.49.2.0140

Rodríguez-Triana, M., Prieto, L., Vozniuk, A., Shirvani Boroujeni, M., Schwendimann, B., Holzer, A., & Gillet, D. (2016). Monitoring, awareness and reflection in blended technology enhanced learning: A systematic review. *International Journal of Technology Enhanced Learning*, *9*(2/3), 126–150. https://doi.org/10.1504/IJTEL.2017.10005147

Saquib, N., Bose, A., George, D., & Kamvar, S. (2018). Sensei: Sensing educational interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(4), 161. https://doi.org/10.1145/3161172

Schlotterbeck, D., Uribe, P., Araya, R., Jimenez, A., & Caballero, D. (2021). What classroom audio tells about teaching: A cost-effective approach for detection of teaching practices using spectral audio features. In *Proceedings of the 11th International Conference on Learning Analytics and Knowledge* (LAK 2021), 12–16 April 2021, Irvine, California, USA (pp. 132–140). ACM. https://doi.org/10.1145/3448139.3448152

Sedova, K., Sedláček, M., Švařı́ček, R., Majcík, M., Navrátilová, J., Drexlerova, A., Kychler, J., & Šalamounová, Z. (2019). Do those who talk more learn more? The relationship between student classroom talk and student achievement. *Learning and Instruction*, *63*, 101217. https://doi.org/10.1016/j.learninstruc.2019.101217

Si, J., Lin, J., Jiang, F., & Shen, R. (2019). Hand-raising gesture detection in real classrooms using improved R-FCN. *Neurocomputing*, *359*(100), 69–76. https://doi.org/10.1016/j.neucom.2019.05.031

Smith, M. K., Jones, F. H. M., Gilbert, S. L., & Wieman, C. E. (2013). The classroom observation protocol for undergraduate STEM (COPUS): A new instrument to characterize university stem classroom practices. *CBE—Life Sciences Education*, *12*(4), 618–627. https://doi.org/10.1187/cbe.13-08-0154

Sümer, Ö., Goldberg, P., D'Mello, S., Gerjets, P., Trautwein, U., & Kasneci, E. (2023). Multimodal engagement analysis from facial videos in the classroom. *IEEE Transactions on Affective Computing*, *14*(2), 1012–1027. https://doi.org/10.1109/TAFFC.2021.3127692

Vujovic, M., Hernández-Leo, D., Martinez-Maldonado, R., Cukurova, M., & Spikol, D. (2022). Multimodal learning analytics and the design of learning spaces. In M. Giannakos, D. Spikol, D. Di Mitri, K. Sharma, X. Ochoa, & R. Hammad (Eds.), *The multimodal learning analytics handbook* (pp. 31–49). Springer International Publishing. https://doi.org/10.1007/978-3-031-08076-0_2

White, J., & Gardner, J. (2013). *The classroom x-factor: The power of body language and non-verbal communication in teaching*. Routledge. https://doi.org/10.4324/9780203818701

Woolfolk, A. E., & Brooks, D. M. (1985). The influence of teachers' nonverbal behaviors on students' perceptions and performance. *The Elementary School Journal*, *85*(4), 513–528. https://doi.org/10.1086/461418

Worsley, M. (2018). Multimodal learning analytics' past, present, and, potential futures. In R. Martinez-Maldonado, V. Echeverria, L. P. Prieto, M. J. Rodriguez-Triana, D. Spikol, M. Curukova, M. Mavrikis, X. Ochoa, & M. Worsley (Eds.), *Companion Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (LAK 2018), 7–9 March 2018, Sydney, Australia (paper 5, Vol. 2163). CEUR-WS.

Xhakaj, F. (2021). *Investigating how to support teachers in their teaching and help them improve their practices through data and technology* [Doctoral dissertation, Carnegie Mellon University]. https://www.proquest.com/openview/aac5ecc54e2f2d54df0f1202066ee87f/1

Yan, L., Zhao, L., Gasevic, D., & Martinez-Maldonado, R. (2022). Scalability, sustainability, and ethicality of multimodal learning analytics. In *Proceedings of the 12th International Conference on Learning Analytics and Knowledge* (LAK 2022), 21–25 March 2022, online (pp. 13–23). ACM. https://doi.org/10.1145/3506860.3506862

