

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Heilala, Ville; Lehesvuori, Sami; Hämäläinen, Raija; Kärkkäinen, Tommi

Title: Toward Scalable and Transparent Multimodal Analytics to Study Standard Medical Procedures : Linking Hand Movement, Proximity, and Gaze Data

Year: 2024

Version: Published version

Copyright: © 2024 Copyright held by the owner/author(s).

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Heilala, V., Lehesvuori, S., Hämäläinen, R., & Kärkkäinen, T. (2024). Toward Scalable and Transparent Multimodal Analytics to Study Standard Medical Procedures : Linking Hand Movement, Proximity, and Gaze Data. In SAC '24 : Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing (pp. 3-10). ACM. <https://doi.org/10.1145/3605098.3635929>



Toward Scalable and Transparent Multimodal Analytics to Study Standard Medical Procedures: Linking Hand Movement, Proximity, and Gaze Data

Ville Heilala*
University of Jyväskylä
Jyväskylä, Finland

Raija Hämäläinen
University of Jyväskylä
Jyväskylä, Finland

Sami Lehesvuori
University of Jyväskylä
Jyväskylä, Finland

Tommi Kärkkäinen
University of Jyväskylä
Jyväskylä, Finland

ABSTRACT

This study employed multimodal learning analytics (MMLA) to analyze behavioral dynamics during the ABCDE procedure in nursing education, focusing on gaze entropy, hand movement velocities, and proximity measures. Utilizing accelerometers and eye-tracking techniques, behaviorgrams were generated to depict various procedural phases. Results identified four primary phases characterized by distinct patterns of visual attention, hand movements, and proximity to the patient or instruments. The findings suggest that MMLA can offer valuable insights into procedural competence in medical education. This research underscores the potential of MMLA to provide detailed, objective evaluations of clinical procedures and their inherent complexities.

CCS CONCEPTS

• **Applied computing** → **Interactive learning environments**; • **Hardware** → **Signal processing systems**; • **Computing methodologies** → **Artificial intelligence**.

KEYWORDS

multimodal, eye-tracking, hand movement, proxemics, behaviorgram, medical education

ACM Reference Format:

Ville Heilala, Sami Lehesvuori, Raija Hämäläinen, and Tommi Kärkkäinen. 2024. Toward Scalable and Transparent Multimodal Analytics to Study Standard Medical Procedures: Linking Hand Movement, Proximity, and Gaze Data. In *The 39th ACM/SIGAPP Symposium on Applied Computing (SAC '24)*, April 8–12, 2024, Avila, Spain. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3605098.3635929>

1 INTRODUCTION

Measurement and data collection instruments structure how we gather research data, whereas models and theories structure how we

*Corresponding author



This work is licensed under a Creative Commons Attribution International 4.0 License. SAC '24, April 8–12, 2024, Avila, Spain
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0243-3/24/04.
<https://doi.org/10.1145/3605098.3635929>

define what qualifies as valuable information [48]. Once integrated into scientific practice, the instruments inspire new theoretical concepts and pave the way for their acceptance within the scientific community [19]. Learning analytics (LA) involves collecting and analyzing educational data to understand better and improve learning [10] and multimodal multichannel trace data is suggested to hold promising potential in providing richer insights into domains of learning across various educational settings [3]. However, learning and its processes are complex. Thus, the more comprehensive and transparent data on learners, environments, and interactions can be traced [49], the better possibilities of analyses and utilizations emerge. Utilizing diverse forms and sources of data, in other words, multiple data modalities [4] can enhance the precision and scope of understanding learner behaviors in their contexts [14].

Despite the potential for educational research, multimodal and multichannel data collection presents methodological challenges such as instrumentation errors, lack of accuracy and replicability, handling data with varying dimensions (e.g., sampling rates, temporal alignment), securing internal and external validity, and ensuring the reliability of measures [3, 52]. Also, when collecting such data, a major issue with many commercial and proprietary measurement systems is the lack of financial scalability, methodological transparency, and control over the underlying algorithms used for data collection and analysis, which can lead to questions about the reliability, validity, and ethics [12, 52]. Open-source technologies and accessible APIs of hardware instruments can provide promising approaches for constructing scalable and transparent measurement systems [e.g., 32]; however, the systems are often in prototyping stages and might sacrifice accuracy or portability for affordability [35].

This study represents a case of exploratory experimentation [6, 13, 45] aiming to construct instrumentation for multimodal measurement and analysis of behavior in the context of nurse education. Efficient teamwork is essential in health care, and multimodal approaches to analyze complex dynamical behavior could provide insight, for example, into collaborative practices between health care professionals in educational settings and the field [53]. Specifically, this study describes a minimum viable experiment (MVE) [13] to discover regularities concerning the complex dynamical behavior of a person conducting a medical ABCDE examination procedure (see, Section 2.3). The research aims to answer the following research question: What elements of the ABCDE procedure can

be reconstructed from the multimodal hand movement, proximity, and gaze data by mainly utilizing affordable technology?

2 BACKGROUND

2.1 Multimodal learning analytics

Multimodal learning analytics involves collecting, synchronizing, and analyzing various high-frequency data sources like video, logs, audio, and biosensors to study learning in various settings [4]. Different kinds of multimodal multichannel data streams are the key ingredients of MMLA, and Molenaar et al. [37] categorized them as physiological, behavioral, and contextual data. Physiological data, such as heart rate (HR) and electrodermal activity (EDA), have been associated with, for example, cognitive load management [e.g., 29] and emotional states [e.g., 51]. Behavioral data obtained, for example, using eye tracking and wearable motion detectors, can capture aspects of learners' activities like movement accuracy and situation awareness as they engage with learning content [7, 39, e.g.,]. Contextual data like video recordings, positioning data, and self-reported measures offer insights into learners' interactions and experiences within various environments and learning situations [e.g., 17, 21, 47].

MMLA is expected to produce relevant, accountable, and actionable representations and interpretations while respecting the privacy of the stakeholders [2, 14]. For this purpose, the use of interpretable and hyperparameter-free predictive models can produce minimal overhead, maximizing methodological transparency; an example of such a method is the *Extreme Minimal Learning Machine* (EMLM) with full data as reference points and *Mean-Absolute-Sensitivity* (MAS) to estimate the feature importances [33]. Ouhachi et al. [40] review concluded four trending themes of the MMLA research: 1) addressing different contexts of learning, 2) focusing on self-regulated and collaborative learning processes, 3) encapsulation of multisensory affections from heterogeneous data, and 4) use of modern tools and methods for data analysis. Specifically, MMLA research conducted in real educational contexts, in other words, in the wild, is suggested to hold the potential for providing personalized learning experiences [35]. Regarding MMLA, this research aims to integrate behavioral data from hand movement, proximity, and eye-tracking utilizing scalable and transparent approaches to facilitate research of collaborative learning processes in the wild.

2.2 Behavioral dynamics in health care

Eye-tracking has become an instrumental tool in the medical field, for example, understanding cognitive load and assessing practitioner efficiency. Tokuno et al. [47] reviewed cognitive load assessment tools in surgical education, revealing a range of subjective and objective measures. Subjective tools included questionnaires like the NASA Task Load Index (NASA-TLX), while objective measures encompassed physiological parameters like heart rate variability, gaze entropy, gaze velocity, and pupil size. Also, gaze metrics have been utilized to assess non-technical skills like situation awareness in health care [e.g., 7]. Ahmadi et al. [1] evaluated the mental workloads of Intensive Care Unit (ICU) nurses during their 12-hour shifts, focusing on how stress impacts their eye movement metrics. Their results suggested that periods of high stress seem to be associated with increased eye fixations and gaze entropy and decreased

saccade duration and pupil diameter. Wright et al. [50] utilized mobile eye-tracking to analyze visual attention patterns during an ultrasound-guided anesthesia procedure to differentiate between proficiency levels of practitioners. Their results showed that experienced medical professionals had fewer visual fixations, spent less time on the procedure, and exhibited less visual entropy, suggesting that eye-tracking can offer objective measures for assessing procedural competence and distinguishing expertise levels.

Proxemics—the study of how humans perceive and use space—and examinations of body movements can provide useful information on behaviors in medical education. Already Momen and Fernie [38] utilized a wireless Sony game controller's hardware, including a 3-axis accelerometer, to identify six nursing activities around a patient to improve hand hygiene prompts. By attaching five sensors to a nurse's body and analyzing the movements, the research found that the 1-Nearest Neighbour classifier was the most effective in identifying the activities. Morita et al. [39] used Bluetooth accelerometers and optical position tracking to examine microsurgical technical skills. Fernandez-Nieto et al. [17] pointed out the importance of spatial abilities in nursing, especially in effective team interactions and clinical procedures. Using indoor positioning sensors, their research transformed raw positioning data from nursing education classes into meaningful proxemics constructs like co-presence in interactional spaces, socio-spatial formations, and presence in spaces of interest with the aim of facilitating nurses' reflection, learning, and professional development in simulation-based training. However, indoor positioning systems often require stationary installations bound to a specific space.

Overall, the multimodal analytical advancements in medical education emphasize the importance of real-time assessment and its challenges. Cloude et al. [9] considered metacognition and self-regulation in clinical reasoning and argued that medical education faces challenges in effectively analyzing learning during activities, as most educational settings utilize intermittent assessments that miss real-time information on knowledge, skills, and abilities, highlighting a need for approaches like MMLA. Furthermore, based on an MMLA implementation in nursing education, Martinez-Maldonado et al. [35] pointed out that MMLA systems need to be trustworthy and address data incompleteness while balancing high-quality data capture with portability and affordability of sensors and consider users' concerns about potential distractions and inconvenience due to being monitored.

2.3 The ABCDE approach

Healthcare professionals utilize various standard procedures when diagnosing patients. In this study, we focus on one such procedure that goes by the acronym ABCDE, which stands for Airway, Breathing, Circulation, Disability, and Exposure. It refers to a systematic protocol primarily used in emergency medicine but applies to other healthcare areas [46]. It serves as a universal approach for patient assessment and directs medical professionals, particularly nurses, in conducting an efficient and comprehensive assessment of a critically ill patient's condition [42]. The completion of the assessment involves five stages consisting of different simultaneous and continuous assessment and treatment steps [46]. The procedure starts with an airway assessment to ensure the patient has a

clear breathing passage. The patient’s respiratory rate and quality are then examined during the breathing analysis. The patient’s blood pressure and heart condition are evaluated during the circulation inspection. In the disability stage, neurological function is examined, typically through a quick assessment of the patient’s responses. The final step involves a prompt but thorough examination of the patient’s body to look for any additional symptoms of disease or trauma. The main aim of the ABCDE approach is that healthcare professionals can accurately prioritize treatments and interventions by consistently following a protocol that simplifies complex clinical situations, allowing them to establish common situational awareness among the medical team and save valuable time [46].

Despite the wide use of the ABCDE approach in various clinical settings, Schoeber et al. [42] found that healthcare professionals’ theoretical knowledge of the approach varies based on the professionals’ type of department, profession category, and age. The result suggests a need to more closely examine the underlying individual differences beyond theoretical knowledge. For example, eye-tracking has been successfully used in evaluating the medical professionals’ performance in the ABCDE approach. Fernández-Méndez et al. [16] utilized eye-tracking to study how lifeguards performed the ABCDE approach. They found that the lifeguards’ performance was misaligned with multimodal data: none of the lifeguards completed the approach correctly, but most of their visual fixations during the assessment procedure were shared between the essential areas for the approach, indicating that eye-tracking could be a valuable method for evaluating the performance of medical procedures. Lee et al. [31] utilized eye-tracking, log data, and self-reported cognitive load measurements to assess the performance of the ABCDE approach between experts and novices in a medical simulation game. Their results indicated that experts outperformed novices regarding speed, accuracy, and cognitive load, associated with higher prior knowledge.

3 MATERIALS AND METHODS

3.1 Experimental setting

Two nurse educators specialized in critical care conducted the ABCDE procedure on an actor patient. The experiment was conducted in a classroom, simulating a real medical examination room with real medical equipment and a hospital bed (Figure 1). An actor patient played the role of a patient who had arrived from an appendectomy, a common surgery operation. The task of the participating nurses was to conduct the ABCDE procedure for evaluating the patient’s condition. The participant was required to work close to the patient’s bed and utilize the instrument table positioned 6 meters away from the center of the hospital bed. Multimodal measurement was used to record the participants’ hand movements, gaze dynamics, and proximity data. Both participants performed the procedure twice, and the data were collected for the initial and repeated experiments.

3.2 Apparatus

The measurement system (Figure 2) consisted of wireless and wired sensors and recording devices connected to a Raspberry Pi 4 (8 GB)

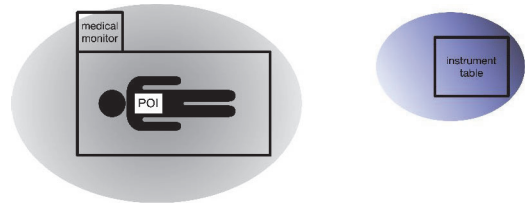


Figure 1: Experimental setting where the point of interest (POI) indicates the reference point for proximity estimation. The gray area indicates an area close to the patient, whereas the blue area is a specific area far from the patient.

microcomputer that served as a hub for collecting and synchronizing the data streams and forwarding them to a recording laptop through the Lab Streaming Layer (LSL). The system’s architectural design is aimed at being extendable for adding additional measurement instruments and scalable for measuring multiple subjects. Apart from the eye-tracking device, the other devices were relatively affordable and accessible and utilized open-source technologies. In this study, the system was capable of real-time measurement and synchronization of five data modalities: hand movements using wireless accelerometers, proximity estimation based on Bluetooth Low Energy (BLE) signal, eye tracking using Tobii Pro Glasses 3, video recording and discrete markers used for real-time annotation. Markers were used as a reference point for evaluating the latency of each individual measurement device. In general, the highest latency of the system was assessed to be approximately 50 ms. Wearable accelerometers and the eye tracker were wireless, allowing free and safe movement of the participants.

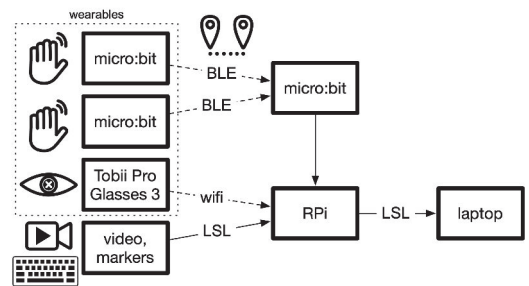


Figure 2: An overview of the apparatus used to measure and record all five data streams.

Hand movement. The micro:bit is a small, versatile, affordable, and programmable open-source ARM-based microcontroller intended for educational and learning purposes, focusing on teaching children the fundamentals of programming and electronics. It includes a variety of sensors, input and output options, and an environment for block-based programming. For example, the micro:bit contains a built-in 3-axis accelerometer that can detect motion, orientation, and tilt. Promoting the constructionist approach by encouraging building interactive projects and engaging with technology, the micro:bit facilitates hands-on learning and fosters creativity and problem-solving skills (Austin et al., 2020). Micro:bit

has been used in several studies relating to computing education (e.g., Andersen, 2022). However, to our knowledge, it has not been utilized as a measurement device for scientific work. This study aimed to pilot and evaluate micro:bit as a scientific instrument. Thus, two micro:bit devices were connected to an add-on shield to enable battery power and wireless wristband use. The devices were attached to the wrists of the participants. The built-in 3-axis accelerometer measured hand movements with a sampling rate of 40 Hz. Devices sent the raw accelerometer signal values using BLE connection to a third micro:bit connected to the RPi receiver.

Proximity. Spatial behavior in terms of proximity was measured using the Received Signal Strength Indicator (RSSI) of the accelerometers, which were connected using BLE to the third micro:bit serving as a receiver. RSSI in Bluetooth technology is a metric that quantifies the power level of a received radio signal. It is commonly used to estimate the distance between devices, as signal strength typically decreases with increasing distance. By employing wireless Bluetooth-based instruments and RSSI, researchers can utilize proximity-based methods [e.g., 36]. RSSI values can be influenced by various factors, such as environmental conditions and obstacles that interfere with radio waves [e.g., 27]. In this study, no significant structures were interfering with the Bluetooth signal. Thus, the raw RSSI signal values between hand movement sensors and the receiver were used, and the signal was calibrated based on the closest and farthest distance to the point of interest (POI). The third receiver micro:bit was placed on the chest of the actor patient, serving as the POI (Figure 1). The farthest point was chosen to be the table containing some of the medical instruments the participants had to use in the procedure. The experiment was designed spatially so that the participant moved mainly around the patient’s bed and between the bed and the medical instrument table.

Gaze. Tobii Pro Glasses 3 eye tracking device (sampling rate 50 Hz) was used to record participants’ eye movements. The raw signal was the (x, y) coordinate of the participants’ gaze in the visual measurement plane of the device. The coordinate values were continuous and in [0, 1]. Blinks were coded as missing values because they caused interruptions in the measurement signal. Tobii Pro Glasses 3 API was used to communicate with the eye-tracking device. The synchronization of the accelerometer and eye-tracking signals was verified by asking the subject to fixate gaze on a stationary point and perform a slow vertical head movement while the micro:bit was attached to the forehead of the subject wearing the eye-tracking glasses. For example, a similar approach has been used to synchronize eye-tracking and motion-capture systems [5]. Figure 3 shows the synchronized vertical head movement (up and down) measured using the micro:bit and the slowly changing vertical eye movement when fixating on a stationary point. The use of raw gaze signals in this study allowed context-free analysis without the need to define areas of interest (AOI).

Video and markers. The video recorder setup consisted of a laptop and a webcam. The webcam stored the raw video file on the laptop’s local hard drive and sent the video stream’s frame numbers over LSL to the recording laptop. Synchronized frame numbers enabled the synchronization of the video with other multimodal data. Also, the laptop was used to send and synchronize keyboard markers

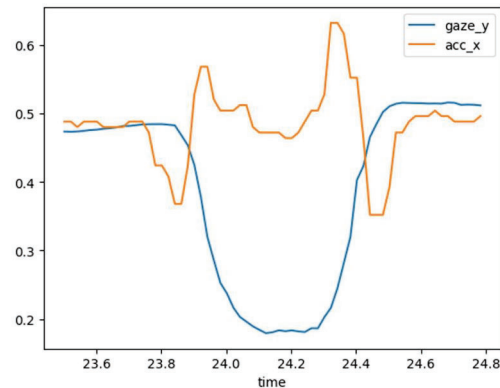


Figure 3: Synchronized head movement (acc_x) and eye movement (gaze_y)

over LSL for live annotation of the experiment. Markers were used to sequence the multimodal data according to the steps and phases of the ABCDE procedure. Synchronized video and markers served as ground truth for validating the analysis results.

3.3 Analysis procedure

To reduce the incoherence in the RSSI signal, proximity was operationalized based on the strongest signal of the accelerometers in the right hand (rh) and left hand (lh) for each time point t , $RSSI_{max}^t = \max(RSSI_{rh}^t, RSSI_{lh}^t)$. Based on the calibration measurements in the experiment, the signal was discretized as a binary variable to indicate the time points when the participant was working beside the patient and beside the medical instrument table. A missing value suggested that the participant was located somewhere in the intermediate space. Hand movements were operationalized using the velocity of the movement. Before calculating velocity, the signal was preprocessed by applying a Savitzky–Golay filter for denoising [24, 41].

Entropy provides a useful metric for understanding the degree of variability, disorder, or unpredictability in the studied data or system. For example, entropy has been used to examine the development of attention to faces [18] and webpage aesthetics [20]. Stationary gaze entropy reflecting the overall spatial dispersion of gaze was used to operationalize gaze dynamics between explorative (i.e., wider gaze dispersion) and exploitative (i.e., limited gaze dispersion) phases where lower entropy is interpreted to indicate more exploitative, spatially focused, and coherent visual focus [18, 22, 30, 44].

In the context of information theory, entropy quantifies the uncertainty or randomness of a set of outcomes or events. Entropy can be quantified using the Shannon entropy [43], defined as the average Shannon information content of an outcome [34]. In other words, it quantifies the average amount of information needed to describe an outcome from a random variable following a given probability distribution. Measured gaze data concerns two coordinate variables. Using the logarithm base 2, Shannon entropy is measured in bits [43] and the joint entropy of two variables is [34]: $H(x, y) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P(x, y) \log_2 \frac{1}{P(x, y)}$.

Raw gaze signal was preprocessed using cubic spline imputation [e.g., 18] to deal with the missing coordinate values caused by blinks. A probability distribution of the continuous gaze measurement signal was needed for calculating the joint entropy. To create a probability distribution of the continuous gaze measurement signal, the data was discretized into equally sized bins representing the state space of gaze behavior. In other words, the discretization divided the measurement plane of the eye tracker as a 100 x 100 matrix, each cell depicting the probability of gazing at that section of the visual plane during a specific time period. Entropy was calculated for a sliding window of 5 seconds. To evaluate the robustness of the approach, different discretization group sizes (i.e., 10, 25, 50, 75) and sliding windows (i.e., 2, 3, 4, 6) were tested. However, the results were qualitatively the same.

The measurements were visualized using a behaviorgram, a graphical representation that visually depicts patterns of behavior, interactions, or activities over time. Like exploratory data analysis, visual analytics aims to uncover knowledge and acquire insight from complex data sets [11]. Behaviorgrams can be used in visual analytics to understand the behavior of individuals or groups in contexts such as psychology and human-computer interaction [e.g., 8]. The custom extended behaviorgram (Figure 4) presented in this study exploits dimensional stacking and the dense pixel technique [25, 26] to visualize temporal relationships of all the measured dimensions.

The central axis of the behaviorgram represents temporal hand movement velocities as an accelerograph. The accelerograph is asymmetrical concerning the central line, the lower part representing the right hand and the upper part representing the left hand. Color coding of the accelerograph exploits the dense pixel technique, a sort of heatmap that depicts higher RSSI signal strength in a brighter color, indicating higher proximity to the POI. The accelerograph's upper temporal segment illustrates the participant's binary position (i.e., beside the patient, beside the instrument table). The lower segment depicts the gaze entropy, where the mean entropy was set as the threshold for marking a segment as denoting low entropy (i.e., more coherent and spatially focused visual perception). Furthermore, the extended behaviorgram was reduced to a more simplified behaviorgram (Figures 5 and 6). The simplified behaviorgram captures proximity and combines the dimension of hand movement and gaze entropy, specifically illustrating the participant's behavioral dynamics concerning the patient. Behaviorgrams were discretized into broad behavioral phases based on video observation and marker annotations.

4 RESULTS

Based on visual analytics, the expected steps in the ABCDE procedure, and validation according to video recordings and annotations, behaviorgrams were found to consist of four phases (Figures 5 and 6). In Phase I, the nurse retrieved medical instruments and attached them to the patient, which involved hand movements close to the patient while maintaining high visual focus when handling the instruments. Phase II consisted of monitoring respiration frequency by observing the patient's chest, confirming visually other vital signs from the medical monitor (IIa), and using a stethoscope to auscultate the patient's chest (IIb). In phase II, the nurses were

positioned either close to the patient or in the intermediate space between the patient and the instrument table. Phase II corresponds to the Breathing assessment in the ABCDE approach, and it is characterized by high visual attention as measured using gaze entropy because the tasks require focusing on the patient and the monitor. Phase IIa involved only visual observations without hand movements, and in Phase IIb, some hand movements can be seen in the behaviorgram because of the use of a stethoscope.

The latter part of the ABCDE procedure (Phase III) involved fetching instruments from the instrument table and performing small examination operations close to the patient (e.g., measuring body temperature and giving medication). Thus, Phase III corresponds to the Circulation, Disability, and Exposure assessments in the ABCDE approach, and it is characterized by changes in proximity and alternating hand activity combined with low gaze entropy (i.e., coherent visual perception). Phase IV consisted of retrieving a checklist and reviewing the patient's condition according to the list. The phase was characterized by changes in proximity, a few short periods of hand movements, and visual focus.

Accelerographs representing hand movement velocity showed specific dynamical patterns based on the phases of the ABCDE procedure. The preparation (Phase I) involved attaching medical instruments to the patient, which is seen as a continuous period of high hand movements in all behaviorgrams. Specifically, hand disinfection was clearly shown as having high peaks in velocity and a higher distance from the patient because it was performed at the instrument table. For example, Nurse 1 performed three hand disinfections in Phase III in the repeated experiment (Figure 5b). On the other hand, the phases where the nurse mainly observed the patient's condition visually were characterized by low hand activity (Phase IIa, IIb, and IV).

5 DISCUSSION AND CONCLUSION

The results of this study underscore the potential of using multimodal learning analytics in understanding behavioral dynamics in the medical field. Utilizing relatively affordable technology and visual analytics, the research was able to trace the different phases of the ABCDE procedure and discern the behavioral patterns associated with each phase. The clear co-occurrence of hand movement activity, gaze entropy, and spatial location across various stages suggest that these metrics provide insights into the dynamics of the procedure. Notably, the low gaze entropy indicated periods of consistent visual perception throughout the procedure, suggesting that medical professionals frequently alternate between explorative and exploitative gazes, especially during intricate procedures. This is particularly significant when considering the importance of visual attention in medical tasks and how it can influence the outcome of procedures.

Integrating multimodal data into a behaviorgram revealed distinct visual patterns based on the different phases of the ABCDE procedure. The results showed that preparation for the procedure, breathing assessment, Circulation/Disability/Exposure phase, and review phase could be identified. Different periods of eye-hand coordination can be distinguished when combining gaze entropy with information from the accelerograph (i.e., high hand activity and high visual focus, low hand activity and high visual focus).

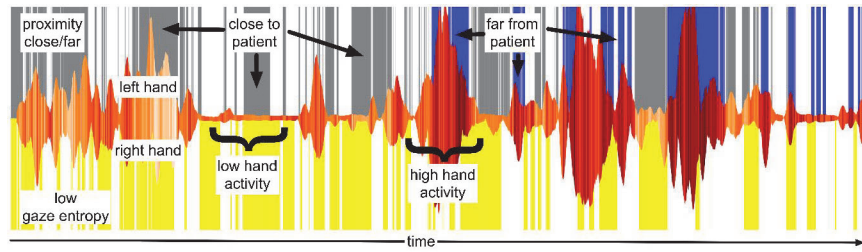


Figure 4: An example of a behaviorgram fusing the dynamics of the multimodal data

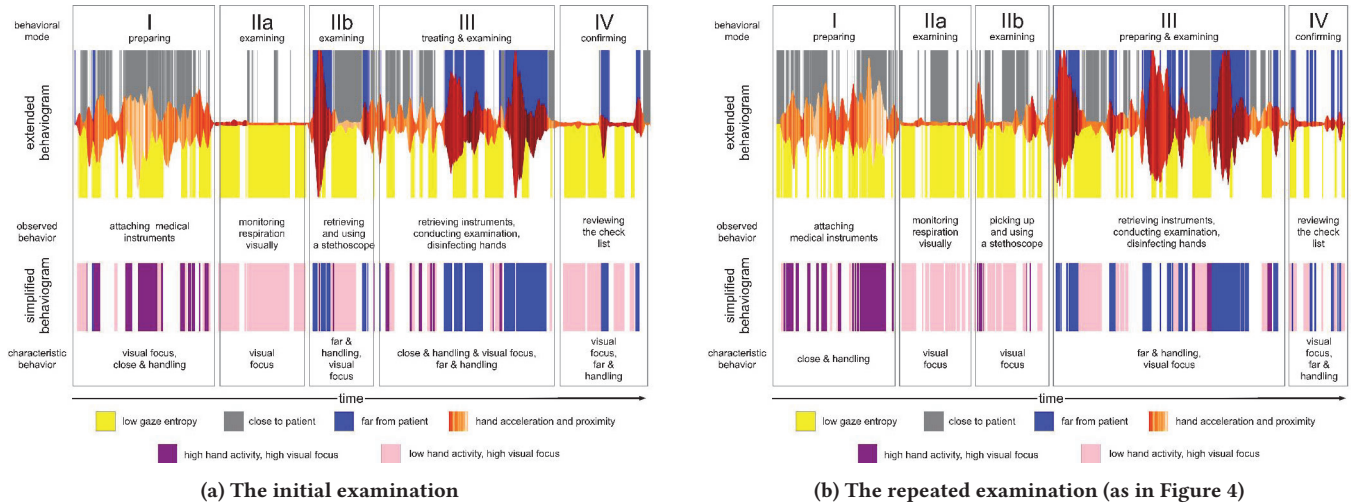


Figure 5: Behaviorgrams of Nurse 1

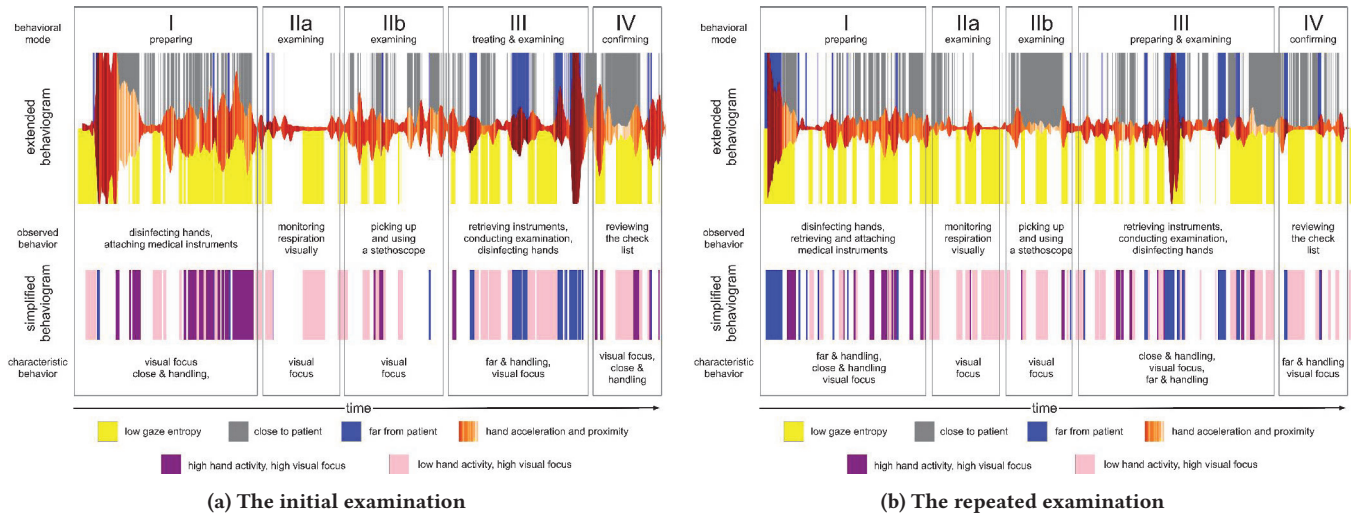


Figure 6: Behaviorgrams of Nurse 2

Specifically, phases characterized mainly by visual observation displayed visual focus and reduced hand activity, thus allowing for the differentiation between manual and observational phases of the procedure. The proximity measure captured the movement of the

nurse between the patient and the instrument table. In general, multimodal behaviorgrams and results based on visual analytics could be linked with actual behavioral dynamics during the procedure.

The results highlight that multimodal multichannel data collection approaches could and should be examined for validity before feeding the data to complex machine learning and artificial intelligence algorithms. Before engaging with more complex analysis techniques, it can be helpful to utilize visual analytics to examine the potential patterns in the data. Such an approach could assist in validating the measurement procedures and facilitating transparency of the more complex analyses. The results provided initial evidence of validity and reliability: results aligned with visual analytics and observed behavior in marker-annotated video recordings for both the initial and repeated examinations for both subjects. In other words, the results showed evidence of within-subject and between-subjects similarity.

The multimodal multichannel measurement in this study utilized relatively affordable technology (i.e., Raspberry Pi, micro:bit), enabling the techniques to be scaled for multiple subjects. The results showed that micro:bit has the potential to produce accurate multimodal measurement data while Raspberry Pi functions as a recording device. The expensive part of the instrumentation was the eye-tracking device; however, more affordable devices are potentially being introduced to the market as technology advances. Scalable and transparent measurement and analysis of behavioral dynamics can enable research in the wild, which refers to approaches to studying and understanding human behavior and technology interactions in real-world, everyday settings, as opposed to controlled lab environments [23, 35]. For example, such approaches can enable research in medical situations where an observer can not enter the room of a patient [e.g., 15]. Furthermore, Kolbe and Boos [28] pointed out the limitations of traditional team research methods in healthcare, which often focus on static descriptions rather than dynamic team processes over time. They suggested that more profound insights into the intricacies of teamwork can be achieved by adopting methodological approaches that consider dynamics, such as event- and time-based observations, social sensor-based measurement, and micro-level coding. Thus, potential applications of the approach presented in this research could include the analysis of situation awareness, professional noticing, joint visual attention in collaborative tasks, understanding the dynamics of patient care, and exploring how medical instruments are handled in real-world scenarios. These insights could inform training programs, process improvements, and even technology design for healthcare contexts. However, it is worth being aware of and clearly defining the limits and scope of the multimodal approaches; in other words, “noting one’s paradigm’s relatively well-marked perimeter is a hallmark of sound and responsible science” [48, p. 288]. In conclusion, this research adds value to medical education by emphasizing the importance of integrating multimodal measures to understand medical professionals’ behavior during standard procedures comprehensively.

5.1 Limitations and future research

Like other similar studies implementing an exploratory approach, this study can be criticized due to its lack of a control group and the generalisability of the results. Furthermore, while behaviorgram provides a comprehensive overview of behavior over time, it does not capture nuanced processes underlying specific actions. Finally,

the study utilized a very small sample size, and the generalizability of the results to other medical procedures beyond the ABCDE approach remains to be explored. Future studies need to utilize more comprehensive analysis techniques and delve deeper into the individual differences among professionals and how these might influence the observed behaviors. Furthermore, the critical issue for future work is to examine how to bridge observed behavioral dynamics with cognitive functions and outcomes.

ACKNOWLEDGMENTS

The work was supported by the Research Council of Finland under Grant number 353325.

REFERENCES

- [1] Nimma Ahmadi, Farzan Sasangohar, Jing Yang, Denny Yu, Valerie Danesh, Steven Klahn, and Faisal Masud. 2022. Quantifying Workload and Stress in Intensive Care Unit Nurses: Preliminary Evaluation Using Continuous Eye-Tracking. *Human factors* 0, 0 (May 2022), 187208221085335. <https://doi.org/10.1177/00187208221085335>
- [2] Haifa Alwahaby, Mutlu Cukurova, Zacharoula Papamitsiou, and Michail Giannakos. 2022. The Evidence of Impact and Ethical Considerations of Multimodal Learning Analytics: A Systematic Literature Review. In *The Multimodal Learning Analytics Handbook*, Michail Giannakos, Daniel Spikol, Daniele Di Mitri, Kshitij Sharma, Xavier Ochoa, and Rawad Hammad (Eds.). Springer International Publishing, Cham, 289–325. https://doi.org/10.1007/978-3-031-08076-0_12
- [3] Roger Azevedo and Dragan Gašević. 2019. Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Computers in human behavior* 96 (July 2019), 207–210. <https://doi.org/10.1016/j.chb.2019.03.025>
- [4] Paulo Blikstein. 2013. Multimodal learning analytics. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK '13)*. Association for Computing Machinery, New York, NY, USA, 102–106. <https://doi.org/10.1145/2460296.2460316>
- [5] Birgitta Burger, Anna Puupponen, and Tommi Jantunen. 2018. Synchronizing eye tracking and optical motion capture: How to bring them together. *Journal of eye movement research* 11, 2 (May 2018), 1–16. <https://doi.org/10.16910/jemr.11.2.5>
- [6] R M Burian. 1997. Exploratory experimentation and the role of histochemical techniques in the work of Jean Brachet, 1938-1952. *History and philosophy of the life sciences* 19, 1 (1997), 27–45.
- [7] Jackie S Cha and Denny Yu. 2022. Objective Measures of Surgeon Non-Technical Skills in Surgery: A Scoping Review. *Human factors* 64, 1 (Feb. 2022), 42–73. <https://doi.org/10.1177/0018720821995319>
- [8] Richard Chen, Filip Jankovic, Nikki Marinsek, Luca Foschini, Lampros Kourtis, Alessio Signorini, Melissa Pugh, Jie Shen, Roy Yaari, Vera Maljkovic, Marc Sunga, Han Hee Song, Hyun Joon Jung, Belle Tseng, and Andrew Trister. 2019. Developing Measures of Cognitive Impairment in the Real World from Consumer-Grade Multimodal Sensor Streams. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (Anchorage, AK, USA) (KDD '19)*. Association for Computing Machinery, New York, NY, USA, 2145–2155. <https://doi.org/10.1145/3292500.3330690>
- [9] Elizabeth B Cloude, Megan D Wiedbusch, Daryn A Dever, Dario Torre, and Roger Azevedo. 2022. The Role of Metacognition and Self-regulation on Clinical Reasoning: Leveraging Multimodal Learning Analytics to Transform Medical Education. In *The Multimodal Learning Analytics Handbook*, Michail Giannakos, Daniel Spikol, Daniele Di Mitri, Kshitij Sharma, Xavier Ochoa, and Rawad Hammad (Eds.). Springer International Publishing, Cham, 105–129. https://doi.org/10.1007/978-3-031-08076-0_5
- [10] G Conole, D Gašević, P Long, and G Siemens. 2011. Message from the LAK 2011 General & Program Chairs. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*. ACM, New York, NY, 1–2.
- [11] Wenqiang Cui. 2019. Visual Analytics: A Comprehensive Overview. *IEEE Access* 7 (2019), 81555–81573. <https://doi.org/10.1109/ACCESS.2019.2923736>
- [12] Lida Z David, Maaike D Endendijk, and Piet Van den Bossche. 2022. Investigating Interaction Dynamics: A Temporal Approach to Team Learning. In *Methods for Researching Professional Learning and Development: Challenges, Applications and Empirical Illustrations*, Michael Goller, Eva Kyndt, Susanna Paloniemi, and Crina Damşa (Eds.). Springer International Publishing, Cham, 187–209. https://doi.org/10.1007/978-3-031-08518-5_9
- [13] Berna Devezer and Erkan Buzbas. 2021. Minimum Viable Experiment to Replicate. <http://philsci-archive.pitt.edu/21475/>
- [14] Daniele Di Mitri, Jan Schneider, Marcus Specht, and Hendrik Drachslers. 2018. From signals to knowledge: A conceptual model for multimodal learning analytics.

- Journal of Computer Assisted Learning* 34, 4 (2018), 338–349.
- [15] Farina Faiz, Yoshinori Ideno, Hiromichi Iwasaki, Yoko Muroi, and Sozo Inoue. 2021. Multilabel Classification of Nursing Activities in a Realistic Scenario. In *Activity and Behavior Computing*, Md Atiqur Rahman Ahad, Sozo Inoue, Daniel Roggen, and Kaori Fujinami (Eds.). Springer Singapore, Singapore, 269–288. https://doi.org/10.1007/978-981-15-8944-7_17
- [16] Felipe Fernández-Méndez, Martín Otero-Agra, Cristian Abelairas-Gómez, Nieves María Sáez-Gallego, Antonio Rodríguez-Núñez, and Roberto Barcala-Furelos. 2019. ABCDE approach to victims by lifeguards: How do they manage a critical patient? A cross sectional simulation study. *PloS one* 14, 4 (April 2019), e0212080. <https://doi.org/10.1371/journal.pone.0212080>
- [17] Gloria Fernandez-Nieto, Roberto Martinez-Maldonado, Vanessa Echeverria, Kirsty Kitto, Pengcheng An, and Simon Buckingham Shum. 2021. What Can Analytics for Teamwork Proxemics Reveal About Positioning Dynamics In Clinical Simulations? *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1 (April 2021), 1–24. <https://doi.org/10.1145/3449284>
- [18] Michael C Frank, Edward Vul, and Scott P Johnson. 2009. Development of infants' attention to faces during the first year. *Cognition* 110, 2 (Feb. 2009), 160–170. <https://doi.org/10.1016/j.cognition.2008.11.010>
- [19] Gerd Gigerenzer. 1991. From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological review* 98, 2 (April 1991), 254–267. <https://doi.org/10.1037/0033-295X.98.2.254>
- [20] Zhenyu Gu, Chenhao Jin, Danny Chang, and Liqun Zhang. 2021. Predicting webpage aesthetics with heatmap entropy. *Behaviour & information technology* 40, 7 (May 2021), 676–690. <https://doi.org/10.1080/0144929X.2020.1717626>
- [21] Ville Heilala, Päivikki Jääskelä, Mirka Saarela, Anna-Stina Kuula, Anne Eskola, and Tommi Kärkkäinen. 2022. "Sitting at the Stern and Holding the Rudder": Teachers' Reflections on Action in Higher Education Based on Student Agency Analytics. In *Digital Teaching and Learning in Higher Education: Developing and Disseminating Skills for Blended Learning*, Leonid Chechurin (Ed.). Palgrave Macmillan, Cham, 71–91. https://doi.org/10.1007/978-3-031-00801-6_4
- [22] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost van de Weijer. 2011. *Eye Tracking: A comprehensive guide to methods and measures*. OUP Oxford, Oxford.
- [23] Edwin Hutchins. 1995. *Cognition in the Wild*. MIT Press, Cambridge, MA.
- [24] Malek Karaim, Aboelmagd Noureldin, and Tashfeen B Karamat. 2019. Low-cost IMU Data Denoising using Savitzky-Golay Filters. In *2019 International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, IEEE, Piscataway, 1–5. <https://doi.org/10.1109/ICCSPA.2019.8713728>
- [25] Daniel A Keim. 2001. Visual exploration of large data sets. *Commun. ACM* 44, 8 (Aug. 2001), 38–44. <https://doi.org/10.1145/381641.381656>
- [26] Daniel A Keim. 2002. Information visualization and visual data mining. *IEEE transactions on visualization and computer graphics* 8, 1 (Jan. 2002), 1–8. <https://doi.org/10.1109/2945.981847>
- [27] Dae-Yeob Kim, Soo-Hyung Kim, Daeseon Choi, and Seung-Hun Jin. 2015. Accurate Indoor Proximity Zone Detection Based on Time Window and Frequency with Bluetooth Low Energy. *Procedia computer science* 56 (Jan. 2015), 88–95. <https://doi.org/10.1016/j.procs.2015.07.199>
- [28] Michaela Kolbe and Margarete Boos. 2019. Laborious but Elaborate: The Benefits of Really Studying Team Dynamics. *Frontiers in psychology* 10 (June 2019), 1478. <https://doi.org/10.3389/fpsyg.2019.01478>
- [29] Joni Lämsä, Joonas Mannonen, Ari Tuhkala, Ville Heilala, Arto Helovu, Ilkka Tynkynen, Emilia Lampi, Katriina Sipiläinen, Tommi Kärkkäinen, and Raija Hämäläinen. 2023. Capturing cognitive load management during authentic virtual reality flight training with behavioural and physiological indicators. *Journal of computer assisted learning* 39, 5 (April 2023), 1553–1563. <https://doi.org/10.1111/jcal.12817>
- [30] Sara Lanini-Maggi, Ian T Ruginski, Thomas F Shipley, Christophe Hurter, Andrew T Duchowski, Benny B Briesemeister, Jihyun Lee, and Sara I Fabrikant. 2021. Assessing how visual search entropy and engagement predict performance in a multiple-objects tracking air traffic control task. *Computers in Human Behavior Reports* 4 (Aug. 2021), 100127. <https://doi.org/10.1016/j.chbr.2021.100127>
- [31] Joy Yeonjoo Lee, Jeroen Donkers, Halszka Jarodzka, and Jeroen J van Merriënboer. 2019. How prior knowledge affects problem-solving performance in a medical simulation game: Using game-logs and eye-tracking. *Computers in human behavior* 99 (Oct. 2019), 268–277. <https://doi.org/10.1016/j.chb.2019.05.035>
- [32] Maximilian Xiling Li, Mario Nadj, Alexander Maedche, Dirk Ifenthaler, and Johannes Wöhler. 2021. Towards a physiological computing infrastructure for researching students' flow in remote learning. *Technology Knowledge and Learning* 27, 1 (Sept. 2021), 365–384. <https://doi.org/10.1007/s10758-021-09569-4>
- [33] Joakim Linja, Joonas Hämäläinen, Paavo Nieminen, and Tommi Kärkkäinen. 2023. Feature selection for distance-based regression: An umbrella review and a one-shot wrapper. *Neurocomputing* 518 (2023), 344–359.
- [34] David J C MacKay. 2003. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, Cambridge.
- [35] Roberto Martinez-Maldonado, Vanessa Echeverria, Gloria Fernandez-Nieto, Lixiang Yan, Linxuan Zhao, Riordan Alfredo, Xinyu Li, Samantha Dix, Hollie Jaggard, Rosie Wotherspoon, Abra Osborne, Simon Buckingham Shum, and Dragan Gašević. 2023. Lessons Learnt from a Multimodal Learning Analytics Deployment In-the-wild. *ACM Trans. Comput.-Hum. Interact.* 31, 1 (Sept. 2023), 1–41. <https://doi.org/10.1145/3622784>
- [36] Roberto Martinez-Maldonado, Jurgen Schulte, Vanessa Echeverria, Yuveena Gopalan, and Simon Buckingham Shum. 2020. Where is the teacher? Digital analytics for classroom proxemics. *Journal of Computer Assisted Learning* 16 (May 2020), 1. <https://doi.org/10.1111/jcal.12444>
- [37] Inge Molenaar, Susanne de Mooij, Roger Azevedo, Maria Barnert, Sanna Järvälä, and Dragan Gašević. 2023. Measuring self-regulated learning and the role of AI: Five years of research using multimodal multichannel data. *Computers in human behavior* 139 (Feb. 2023), 107540. <https://doi.org/10.1016/j.chb.2022.107540>
- [38] Kaveh Momen and Geoff R Fernie. 2010. Nursing activity recognition using an inexpensive game controller: An application to infection control. *Technology and health care: official journal of the European Society for Engineering and Medicine* 18, 6 (2010), 393–408. <https://doi.org/10.3233/THC-2010-0600>
- [39] Akio Morita, Yasuo Murai, and Mamoru Mitsuishi. 2022. Current and Future Microsurgical Skills Assessment. In *Learning and Career Development in Neurosurgery: Values-Based Medical Education*, Ahmed Ammar (Ed.). Springer International Publishing, Cham, 349–356. https://doi.org/10.1007/978-3-031-02078-0_30
- [40] Hamza Ouhachi, Daniel Spikol, and Bahtijar Vogel. 2023. Research trends in multimodal learning analytics: A systematic mapping study. *Computers and Education: Artificial Intelligence* 4 (Jan. 2023), 100136. <https://doi.org/10.1016/j.caeai.2023.100136>
- [41] Ronald W Schafer. 2011. What Is a Savitzky-Golay Filter? [Lecture Notes]. *IEEE Signal Processing Magazine* 28, 4 (July 2011), 111–117. <https://doi.org/10.1109/MSP.2011.941097>
- [42] Nino H C Schoeber, Marjolein Linders, Mathijs Binkhorst, Willem-Pieter De Boode, Jos M T Draaisma, Marlies Morsink, Anneliese Nusmeier, Martijn Pas, Christine van Riessen, Nigel M Turner, Rutger Verhage, Cornelia R M G Fluit, and Marije Hogeveen. 2022. Healthcare professionals' knowledge of the systematic ABCDE approach: a cross-sectional study. *BMC emergency medicine* 22, 1 (Dec. 2022), 202. <https://doi.org/10.1186/s12873-022-00753-y>
- [43] C E Shannon. 1948. A mathematical theory of communication. *The Bell system technical journal* 27, 3 (1948), 379–423. <https://doi.org/10.1145/584091.584093>
- [44] Brook Shiferaw, Luke Downey, and David Crewther. 2019. A review of gaze entropy as a measure of visual scanning efficiency. *Neuroscience and biobehavioral reviews* 96 (Jan. 2019), 353–366. <https://doi.org/10.1016/j.neubiorev.2018.12.007>
- [45] Friedrich Steinle. 1997. Entering New Fields: Exploratory Uses of Experimentation. *Philosophy of science* 64, S4 (1997), S65–S74. <https://doi.org/10.1086/392587>
- [46] Troels Thim, Niels Henrik Vinther Krarup, Erik Lerkevang Grove, Claus Valter Rohde, and Bo Løfgren. 2012. Initial assessment and treatment with the Airway, Breathing, Circulation, Disability, Exposure (ABCDE) approach. *International journal of general medicine* 5 (Jan. 2012), 117–121. <https://doi.org/10.2147/IJGM.S28478>
- [47] Junko Tokuno, Tamara E Carver, and Gerald M Fried. 2023. Measurement and Management of Cognitive Load in Surgical Education: A Narrative Review. *Journal of surgical education* 80, 2 (Feb. 2023), 208–215. <https://doi.org/10.1016/j.jsurg.2022.10.001>
- [48] Philip H Winne. 2019. Paradigmatic Dimensions of Instrumentation and Analytic Methods in Research on Self-Regulated Learning. *Computers in human behavior* 96 (July 2019), 285–289. <https://doi.org/10.1016/j.chb.2019.03.026>
- [49] Marcelo Worsley, Roberto Martinez-Maldonado, and Cynthia D'Angelo. 2021. A New Era in Multimodal Learning Analytics: Twelve Core Commitments to Ground and Grow MMLA. *Journal of Learning Analytics* 8, 3 (2021), 10–27.
- [50] Andrew G Wright, Rahool Patel, Koralay Perez-Edgar, Xiaoxue Fu, Kayla Brown, Sanjib Adhikary, and Adrian Zurca. 2022. Eye-Tracking Technology to Determine Procedural Proficiency in Ultrasound-Guided Regional Anesthesia. *The journal of education in perioperative medicine : JEPM* 24, 1 (Jan. 2022), E684. https://doi.org/10.46374/volxxiv_issue1_zurca
- [51] Elaheh Yadegaridehkordi, Nurul Fazmidar Binti Mohd Noor, Mohamad Nizam Bin Ayub, Hannyyzura Binti Affal, and Nornazlita Binti Hussin. 2019. Affective computing in education: A systematic review and future research. *Computers & education* 142 (Dec. 2019), 103649. <https://doi.org/10.1016/j.compedu.2019.103649>
- [52] Lixiang Yan, Linxuan Zhao, Dragan Gasevic, and Roberto Martinez-Maldonado. 2022. Scalability, Sustainability, and Ethicality of Multimodal Learning Analytics. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (Online, USA) (LAK22). Association for Computing Machinery, New York, NY, USA, 13–23. <https://doi.org/10.1145/3506860.3506862>
- [53] Linxuan Zhao, Zachari Swiecki, Dragan Gasevic, Lixiang Yan, Samantha Dix, Hollie Jaggard, Rosie Wotherspoon, Abra Osborne, Xinyu Li, Riordan Alfredo, and Roberto Martinez-Maldonado. 2023. METS: Multimodal Learning Analytics of Embodied Teamwork Learning. In *LAK23: 13th International Learning Analytics and Knowledge Conference* (Arlington, TX, USA) (LAK2023). Association for Computing Machinery, New York, NY, USA, 186–196. <https://doi.org/10.1145/3576050.3576076>