Norwegian University of Science and Technology

IT3915 - Preparatory Project Masters in Informatics

Preparatory Project Masters in Informatics

Author
Jesper Barfod

December 11, 2024 (Autumn 2024)

NTNU
Norwegian University of
Science and Technology

Developing an Open-Source MMLA Tool for BPMN Comprehension Studies

ABSTRACT

This paper introduces an open-source tool designed to advance BPMN comprehension research by integrating multimodal biometric data collection and analysis. Utilizing Cognitive Load Theory (CLT) and the Affective Learning Framework (ALF), the tool combines EyeMind for eye-tracking with OpenFace for facial expression analysis. Addressing limitations in existing methods, the tool streamlines the experimental workflow and supports dynamic BPMN models. Initial implementation demonstrates its potential to enhance research accuracy, with future iterations planned to incorporate EEG and physiological data for deeper insights.

1 INTRODUCTION

Research into the comprehension of *Business Process Model Notation* (BPMN) has traditionally relied on self-report methods, which struggle to capture the complexity of cognitive and emotional processes involved in understanding visual conceptual models. Despite the potential of neuroscience and biometric techniques—widely applied in linguistics and software engineering—their use in BPMN studies remains limited.

To address this gap, this paper presents an open-source tool that combines biometric data collection with advanced analysis methods, enhancing research into BPMN comprehension. Building on Krogstie and Sharma's 2024 findings [1], the tool integrates EyeMind for eye-tracking and OpenFace for emotion analysis. By offering an all-in-one solution, it supports the entire experimental workflow—from setup and data collection to initial analysis and data export—streamlining processes and reducing potential errors. This integration sets it apart from existing tools by addressing challenges such as ecological validity, scalability, and compatibility with dynamic BPMN models, enabling more effective and comprehensive studies.

The remainder of the paper is structured as follows: Section 2 introduces the 2024 paper by Krogstie and Sharma, alongside relevant theory and an explanation of key concepts in this research field. Section 3 details the features of the integrated tool I have created, including descriptions of the two integrated applications, EyeMind [2] and OpenFace [3]. Section 4 provides documentation on how to download, set up, and use the tool. Finally, Section 5 offers a summary of the paper, draws conclusions, and outlines potential directions for future work.

2 BACKGROUND

In their 2024 paper [1], Krogstie and Sharma collected experimental data from a variety of biometric sources — including a wristband with physiological monitoring, an EEG cap, an eye-tracking device, and a webcam — to examine participants' model comprehension of BPMN models while solving related tasks. By applying *Multimodal Learning Analytics* (MMLA) principles, the authors investigate whether integrating diverse biometric data sources can offer a more comprehensive and accurate understanding of how learners process information. Traditionally, self-report methods have been employed to assess participants' comprehension of models. However, biometric sensors present a significant advantage by providing objective and replicable measurements. In contrast, self-reports are typically collected post hoc (i.e., after the task is completed), which can compromise their validity due to the inherent inaccuracies of recall.

The goal of studying model comprehension is to deepen our understanding of how humans process and interpret conceptual models. This insight can guide the development of more effective models, improving usability and ensuring their appropriate application. The study adopted an exploratory methodology, as integrating neuroscience techniques and biometric data into the study of conceptual models is still in its infancy. Notably, no standardized terminology exists for this area of research. In a literature review on the topic, Krogstie proposes the term *neuro-conceptualization* could be used [4]. Similar fields, like *neuroIS* (Information Systems) and *neuroSE* (Software Engineering), use comparable biometric methods but they too remain underexplored.

While neuroimaging devices such as fMRI provide highly accurate data, their invasive nature compromises ecological validity, requiring participants to remain stationary in an MR machine. To address this limitation, Krogstie and Sharma opted for less invasive methods, including EEG caps and wristbands with physiological sensors. Although EEG detects only surface-level brain activity, it provides excellent temporal resolution — capturing changes in brain activity within milliseconds [5] — and is noninvasive, allowing participants greater mobility. These simpler biometric techniques balance practicality with the ability to yield meaningful results in realistic settings.

MMLA combines two major theoretical frameworks, the *Cognitive Load Theory (CLT)* and the *Affective Learning Framework (ALF)*. Krogstie and Sharma utilize both theories, recognizing their value in improving our understanding of participant model comprehension.

CLT can shortly be described as a theory of what factors influence *working memory load*. Working memory is the immediate memory, responsible for thinking "here-and-now". It has a very limited capacity of about four to nine objects [6, 7]. However, this number can be increased or decreased based on three factors according to CLT.

The *Intrinsic Load* is the inherent difficulty of the subject being learned. Some subjects are inherently harder to focus on and learn - however this inherent difficulty will also be influenced by the learner's prior knowledge of the subject. The *extraneous load* is the difficulty imposed by the way the subject is being taught. Lastly, *germane load* is the attention the learner allocates to the task. The goal of CLT is thus to maximize germane load (attention), while minimizing extraneous load (difficulty arising from the teaching approach).

Krogstie and Sharma explore cognitive processes related to CLT by utilizing EEG and eye-tracking technologies to measure factors such as fatigue, attention, cognitive load, and working memory load. They also account for intrinsic load by assessing participants' prior knowledge of the material through some initial questions conducted before the tasks are presented - ensuring participant expertise is factored into the analysis.

The other framework, Affective Learning Framework, highlights the influence emotions have on learning. Factors such as personal interest, positive emotions, external rewards, social learning environments, and self-regulated learning can enhance engagement, retention, and the practical application of knowledge. Shortly put, to increase learner comprehension, the material being learnt should be engaging and rewarding. Korgstie and Sharma operationalized ALF by using a wristband for physiological data, tracking variables such as heart rate variability (HRV), an indicator of stress. Facial data scanning to assess emotional states and eye-tracking technology to measure attention levels was also used.

The result from this study showed some similarities with another study Sharma et al. had conducted regarding study on code comprehension [8] (and not model comprehension), however with a few marked differences. For instance, frustration negatively correlate with model comprehension, whereas it had a positive association in code comprehension tasks. Krogstie and Sharma attributed this to differences in task nature: the code comprehension tasks involved active problem solving and debugging, where frustration can drive progress. In contrast, the model comprehension tasks were designed to focus primarily on understanding, which should not evoke frustration among participants who are familiar with BPMN models and confident in their abilities.

The experiment demonstrated that using multiple biometric sensors to assess model comprehension is a viable and potentially more accurate alternative to self-reports. Krogstie and Sharma concluded that this methodology is a promising but underexplored avenue for research. They outlined areas for future investigation, emphasizing the potential of multimodal data integration.

However, the study faced limitations. Of the 68 participants, data from only 57 were usable due to errors during data collection. Improving the environment for data collection could lead to fewer lost entries and enhanced ecological validity. Additionally, it may inspire other researchers to conduct similar experiments, further exploring the topic, validating existing findings, or challenging prior conclusions.

3 DEVELOPMENT & FEATURES

The development of this tool builds on the existing open-source applications *EyeMind* and *OpenFace* to create a comprehensive system for BPMN comprehension experiments. By integrating eye-tracking and facial behavioral analysis, the tool supports multimodal learning analytics with enhanced ecological validity. Future plans include the integration of EEG and physiological wristband data, which are further detailed in Section 5.1.

3.1 EyeMind Eye Tracking

The vast majority of neuro-conceptualization studies rely on eye-tracking (ET) as a primary source of biometric data. In fact, all but one of the 43 papers Krogstie examined in his literature review on neuro-conceptualization employed ET [4]. Given its prevalence, ET was selected as the foundation for developing the integrated tool.

The open-source application EyeMind served as an excellent foundation for creating this integrated tool. It handles all ET-specific modelling tasks, from setting up process models and associated questions to conducting experiments and analyzing the collected data.

A common and major limitation of experiments in this domain has been their reliance on static stimuli, often consisting of small, non-interactive process models [2]. Researchers have mostly used simple models in experiments due to the significant challenges posed by dynamic stimuli. When scrolling or zooming during experiments, the on-screen layout changes, causing the ET-recorded coordinates to misalign with parts of the model. By using static models that fit entirely within a single screen, researchers avoided these complications, ensuring that model elements remained in consistent positions. However, static stimuli greatly reduce the ecological validity of experiments, failing to represent the true size and complexity of BPMN models used in real-world scenarios. EyeMind overcomes these limitations by automatically accounting for scrolling and zooming, ensuring accurate ET data even with dynamic stimuli.

EyeMind uses the Tobii SDK to access eye-tracking data, and thus only supports eye-tracking devices which provide Tobii SDK access. For process modelling, it utilizes BPMN 2.0, and includes support of sub-processes to model complex workflows. However, BPMN models are not created directly within the application; users can design models using their preferred tools, such as *Camunda*, and import them into EyeMind.

A limitation with the Tobii eye-trackers is that glasses with strong prescriptions or thick lenses can interfere with the equipment's performance. However, thin glasses or contact lenses typically work without issue.

Despite this limitation, the analysis capabilities of EyeMind are extensive and designed to ensure high data accuracy. This includes, gaze correction tools for addressing potential miscalculation issues by allowing researchers to review samples of recorded gaze points and apply offset corrections. Morever, Eyemind includes fixation filtering, highlighting areas of focus - and also heatmap generation providing a visual representation of gaze patterns, and highlighting areas of interest (AOIs) on the model. The tool also enables the export of detailed experimental data in CSV and JSON formats, including gaze coordinates, fixations, submitted answers, and clicks.

The EyeMind application consists of a graphical user interface and a backend server. The GUI served as a foundation for integrating additional open-source tools, while the backend was modified to incorporate code from other open-source tools.

3.2 OpenFace Emotion Analysis

To facilitate automatic facial expression analysis, the open-source tool OpenFace was integrated into the EyeMind application. OpenFace has been celebrated as a state-of-the-art technology in its field [3]. It provides researchers with a powerful, freely accessible, and customizable tool for analyzing facial expressions. OpenFace requires non-specialized hardware, functioning effectively with only a standard webcam.

One of OpenFace's key features is its ability to detect *Facial Action Units* (AUs), which are fundamental components of facial expressions [3]. AUs represent subtle facial movements, such as the wrinkling of the nose (AU9). Specific combinations of these units correspond to universal expressions of emotions, such as happiness, anger, or fear. For instance, the expression of happiness is characterized by the activation of AU6 (cheek raiser) and AU12 (lip corner puller). This system of categorizing facial movements, known as the *Facial Action Coding System* (FACS), was developed in the late 1970s based on the anatomy of facial muscles and their movements [9]. Over the years, FACS has been extensively validated through psychological studies and has become a cornerstone in research on emotional states.

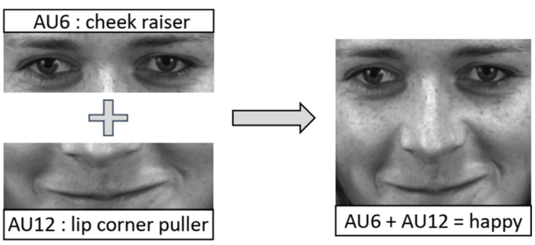

Figure 1: Example of two combined AUs forming an expression [10]

OpenFace employs a streamlined version of FACS, focusing on 18 of the 30 recognized AUs, while omitting less significant ones. Additionally, OpenFace not only detects the presence of these action units but also measures their intensity, providing nuanced insights into facial expressions. Beyond AU detection, OpenFace supports other features such as facial landmark location, head pose estimation, and eye gaze detection. However, these features were not utilized in this implementation. Eye gaze detection and analysis are already comprehensively handled by EyeMind, and the other features were deemed unnecessary for the specific requirements of this project.

The integration of OpenFace's AU detection into EyeMind involved modifications to the existing server code. During experiments, the EyeMind server now initiates asynchronous video recording, followed by automated analysis and the generation of facial action unit data using OpenFace. The data from OpenFace and EyeMind is synchronized with matching timestamps. To prioritize privacy, raw video files are deleted after processing; however, this setting can be easily adjusted to meet alternative requirements if necessary.

Table 1: Key Features of the Integrated Applications

Application	Features
EyeMind	Supports full workflow: setup, execution, analysis, and export
	Graphical User Interface (GUI)
	Scrolling, zooming, dynamic AOIs
	BPMN 2.0 (incl. sub-processes)
	Gaze correction, fixation filtering, and heatmaps
	Data export (JSON & CSV)
OpenFace	Action Unit Detection

4 DOCUMENTATION

This section outlines the setup process for the application, detailing the hardware and software requirements. For complete installation instructions and an in-depth manual covering all features, refer to the GitHub repository.

4.1 System Requirements

Hardware Requirements:

- An eye tracker compatible with the Tobii SDK[1]
- A webcam

Ensure these devices are connected to the computer running the application. Note that some Tobii eye trackers require a USB 3 port.

Software Requirements:

- 64-bit Windows 10 or 11[2]
- Python 3.8.6 (link, choose AMD64 on win32)
- Pip (link)
- NPM and Node.js (link)
- Tobii Eye Tracker Manager (link)

[1]This application has been tested with the *Tobii Pro X3-120* but should work with other Tobii eye trackers supporting the SDK.
[2]Windows 11 is currently supported but will not be in future versions with wristband data.

- For discontinued eye trackers, download Tobii Pro Eye Tracker Manager 2.6.1 instead (link).

- Visual Studio 2017 (link) or the 64-bit Visual C++ redistributable package (link)

4.2 Installation & Usage

For detailed explanations of the installation process and the application's features, please refer to the GitHub repository.

5 CONCLUSION

This paper has introduced a novel open-source tool designed to enhance BPMN comprehension studies by handling multimodal biometric data collection and analysis. Building on the frameworks of Cognitive Load Theory (CLT) and Affective Learning Framework (ALF), the tool uses EyeMind for eye-tracking to analyze cognitive states, and OpenFace for analysis of affective states during modeling tasks.

The need for this tool arose from limitations in existing methods for data collection. These methods which lacked scalability, ecological validity, and streamlined workflows. In contrast, this application offers an all-in-one solution that supports the entire experimental workflow — from setup and data collection to initial analysis and data export. Its ability to handle large and dynamic BPMN models further enhances its applicability.

Drawing from the methodologies introduced by Krogstie and Sharma in 2024, this tool opens new avenues for investigating multimodal data's role in BPMN comprehension. Its potential will be further explored in upcoming studies.

5.1 Future Work

Some areas have been identified for future improvement and development of the application, both in terms of technical enhancements and experimental validation:

1. **Exploring Alternatives to OpenFace:** While once a cutting-edge tool for facial expression analysis, OpenFace was last updated in 2019 and no longer represents the state-of-the-art. A strong modern alternative is *LibreFace*, which offer significant advancements in accuracy, speed, robustness, and feature set. For example, LibreFace outperforms OpenFace by 7% in Pearson Correlation Coefficient (PCC) on the *DISFA dataset*, a widely recognized benchmark featuring spontaneous facial expressions in naturalistic settings [11]. Additionally, LibreFace operates at nearly twice the speed of OpenFace and benefits from regular updates, with the most recent release in September 2024.

2. **Integration of Physiological Data:** Future iterations of EyeMind should include the integration of wristband physiological data, such as those provided by the Empatica E4, as well as EEG data. This would enable greater multimodal data analysis, providing richer insights into cognitive and emotional states during experiments.

3. **Bug Fixing:** Several bugs have been identified in the current version of EyeMind, all of which existed prior to integration. While these issues are not critical, resolving them would enhance the user experience and improve overall usability. The identified bugs include:

   - BPMN models cannot be removed from the experiment setup as the REMOVE button is non-functional.
   - Dragging and dropping multiple BPMN files fails to maintain the correct order.
   - Reordering BPMN files does not work.
   - The heatmap feature does not update properly when switching between time units (microseconds, milliseconds, seconds).

4. **Automated Score Analysis:** The current version of EyeMind logs participant answers but does not verify or record their correctness in the log files. Incorporating this functionality would streamline workflows; however, external verification is not particularly burdensome since the output CSV and JSON files require external analysis regardless.

5. **Comprehensive Testing:** Additional testing of the integrated tool is necessary to assess its reliability, particularly in handling edge cases. Running simulated mock experiments could provide a clearer practical assessment of how well the tool functions during real-world applications.

6. **Conducting Experiments:** The next phase involves conducting full-scale experiments, which are scheduled to commence in the upcoming semester of spring 2025.

References

[1] John Krogstie and Kshitij Sharma. Enhancing our understanding of business process model comprehension using biometric data. 2024.

[2] Amine Abbad-Andaloussi, Daniel Lübke, and Barbara Weber. Conducting eye-tracking studies on large process models using eyemind. *SoftwareX*, 101564, 2023.

[3] Tadas Baltrušaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. Openface 2.0: Facial behavior analysis toolkit. In *2018 IEEE International Conference on Automatic Face and Gesture Recognition (FG 2018)*, pages 59–66. IEEE, 2018.

[4] John Krogstie. Using techniques from neuroscience in information systems modelling research: A systematic mapping study. In *Proceedings of the Norwegian University of Science and Technology (NTNU)*, Norway, 2024. Norwegian University of Science and Technology (NTNU).

[5] Barbara Weber, Thomas Fischer, and René Riedl. Brain and autonomic nervous system activity measurement in software engineering: A systematic literature review. *Journal of Systems and Software*, 180:110835, 2021.

[6] George A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63(1):81–97, March 1956.

[7] Nelson Cowan. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1):87–185, 2000.

[8] Katerina Mangaroska, Kshitij Sharma, Dragan Gašević, and Michail Giannakos. Exploring students' cognitive and affective states during problem solving through multimodal data: Lessons learned from a programming activity. *Journal of Computer Assisted Learning*, 38(1):40–59, 2022.

[9] Paul Ekman and Wallace V. Friesen. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, 1978.

[10] Ankita Jain and Dhananjoy Bhakta. Micro-expressions: a survey. *Multimedia Tools and Applications*, 83, 11 2023.

[11] Di Chang, Yufeng Yin, Zongjian Li, Minh Tran, and Mohammad Soleymani. Libreface: An open-source toolkit for deep facial expression analysis. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, January 2024. To appear.