# Eye tracking capabilities in evaluating human factors during cognitive and motor tasks

**AUTHORS**
*Divya Seernani, Pernille Bülow, Jessica Wilson*
*iMotions A/S - Copenhagen, Denmark*

## INTRODUCTION

In order to test any process, product, or system, it is important to understand the underlying human factors at play. A reliable way to do this is by including biosensors in research. Simple sensors like eye tracking, facial expression analysis, galvanic skin response, and heart rate*[1-4]* can provide valuable information about how people work with information, for example by analyzing how they pay attention to instructions, analyze data, and identify errors. This study compares and contrasts one biosensor - eye tracking*[5]* - in the value it can provide for analysis of different systems.

## APPARATUS & ALGORITHMS

Smart Eye AI-X (60Hz) was used for the SBET condition, Pupil Invisible Glasses (200 Hz) for the Glasses condition, and iMotions Webcam based eye tracking algorithm*[6]* for the WebET condition.

iMotions 9.3.10 was used for fixation classification, gaze mapping and AOI analysis. Fixation classification was conducted using the iMotions I-VT filter for the SBET and Glasses condition and the iMotions I-HMM filter for the WebET Condition. Affectiva's AFFDEX algorithm*[7]* was used for facial emotion detection.

## METHODS

**A feasibility study was carried out in three eye tracking situations (N=12, 4 per condition).**

SBET Condition: screen based eye tracker, rendering high accuracy and precision of gaze data.

Glasses Condition: Eye tracking glasses which allow for flexible environments with high ecological validity.

WebET Condition: Webcam based eye tracking that allows a large-scale analysis across populations.

Area of interest (AOI) analysis was carried out to compare eye tracking metrics.

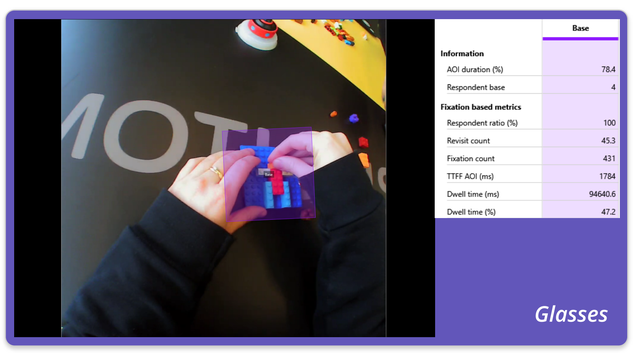
Facial expressions were explored as indicators of enjoyment and frustration during the task.

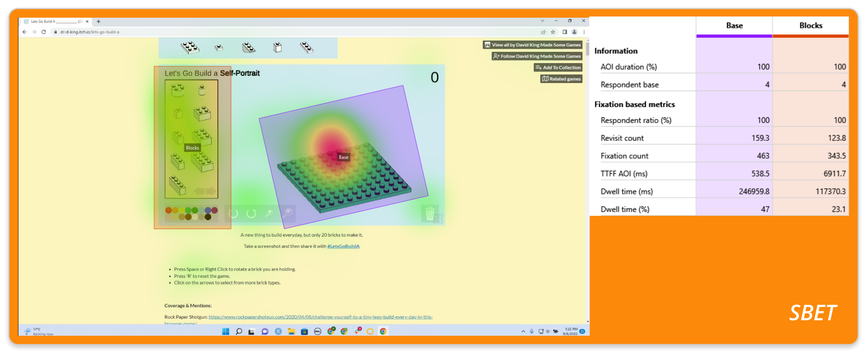
## EASE OF USE & DATA ACCURACY

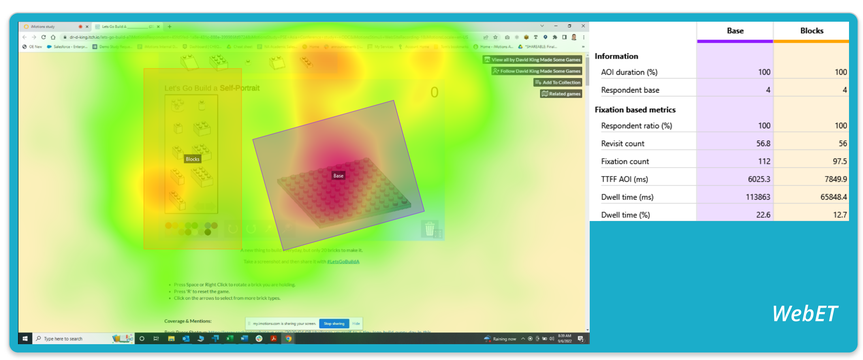
SBET vs WebET showed an **accuracy - scalability tradeoff**. While SBET required a resource intensive lab set-up, it also provided accurate data. The WebET provided ease of scalability, but much lowered accuracy.

SBET vs Glasses showed an **ecological validity to study resources tradeoff.** The Glasses allowed participants to use physical blocks lending higher ecological validity, but the analysis of the individualized, unique datasets proved labor intensive.

WebET vs Glasses showed both to have **higher ecological validity**. While the Glasses allow participants to use actual blocks, WebET allows participants to conduct screen based studies from the comfort of their homes. The Glasses condition is **resource intensive**, but the WebET data is less accurate.



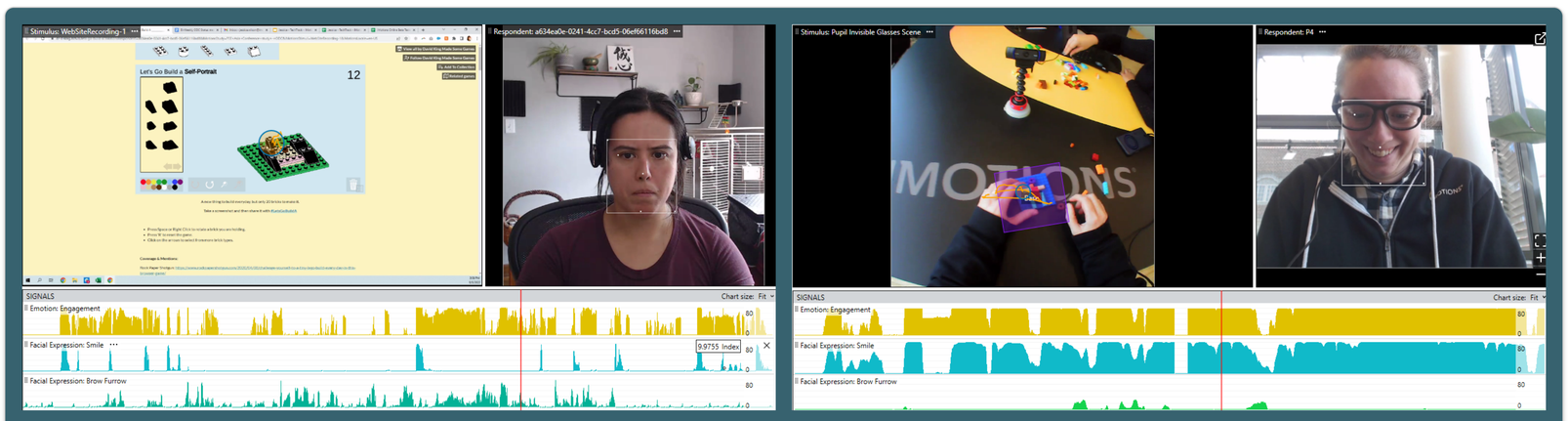




## COMPARABILITY OF DATA

Area of Interest (AOI) analysis showed **comparable aggregate Dwell time (~47% of time)** on the building board for the Glasses and SBET condition. The WebET condition had a much smaller Dwell time (~22%). The heatmap also reflects the **variability of the data inherent to** WebET, suggesting that many of the true fixations on the building board may have been too noisy and outside the scope of the AOI.

## FACIAL EXPRESSION ANALYSIS




*Probability graphs on a scale of 0 to 100 for Engagement (yellow), Smile (teal) and Brow Furrow (green) from a participant doing the task at home (left) or with the experimenter guiding the study (right).*

Metrics, such as smile and brow furrow, measured well in all three setups. However, people had the highest percentage of facial engagement in the Glasses condition, where the experimenter was **interacting with the participants.**

People were also more likely to smile when the experimenter was present, and more likely to furrow their brow (likely a **sign of concentration**) when doing the task online, in a lab, or at home in the SBET and WebET conditions.

## SUMMARY OF CONDITIONS COMPARING SETUPS

| | SBET | Glasses | WebET |
|---|---|---|---|
| **Accuracy** | High accuracy | Lower accuracy, larger field of view | Lowest accuracy |
| **Granularity of analysis** | Smaller AOIs (e.g. on blocks) are possible | Increased analysis time to track all AOIs dynamically | Recommended with larger AOIs to account for accuracy errors |
| **Ecological validity** | Controlled environment | High ecological validity | Higher ecological validity, restricted to a screen |
| **Scalability** | Time consuming to collect data | Time consuming to collect and analyse | Easily scalable |
| **Facial Expression** | Integrates easily | Affected by movements | Integrates easily |
| **Multi-sensor integration** | Possible to integrate with GSR, EMG, ECG, and EEG | Easy to integrate with GSR, ECG, EMG Movement artifacts for EEG | Not possible with the Online platform |
| **Cost effectiveness** | Requires dedicated hardware and study specific resources | Requires dedicated hardware and study specific resources | Webcam based. Minimum human resources. |

**READ THE FULL ARTICLE HERE**

REFERENCES

1. Mele & Federici, *Cogn Process* (2012)
2. Hazlett, *CHI '06* (2006)
3. Hardy et al., *Proc Hum Factors Ergo Soc*, (2013)
4. Hankins & Wilson, *Aviat Space Enviro Med* (1998)
5. Holmqvist et al., *Behavior research method* (2022)
6. iMotions, *"Webcam ET Accuracy Whitepaper v2"* (2022)
7. Bishay et al., *arXiv.org* (2022)