Online Negotiations

-

An Analysis of the Importance and Impact of Different Negotiation Competencies

Bachelor's Thesis

Co-funded by the Erasmus+ Programme of the European Union

Hochschule Düsseldorf
University of Applied Sciences

Fachbereich Elektro- & Informationstechnik
Faculty of Electrical Engineering & Information Technology

Bachelor-Thesis

Optimizing Hybrid Lectures: An Effective Methodology for the Collection and Analysis of Eye-Tracking and Facial Expression Data – A Study on Attention and Engagement in Hybrid Learning Environments

Submitted by [redacted]
Matriculation number: [redacted]
From Düsseldorf, Germany

Prepared as part of the bachelor's examination for Industrial Engineering - Electrical Engineering at the Department of Electrical and Information Technology at the University of Applied Sciences Düsseldorf

Supervisor: Prof. Dr. Kati Lang
Second Examiner: Dr. Divya Seernani

List of Illustrations

List of Abbreviations

AOI – Area of Interest

AU – Action Unit

BEACON – Behavioral Analysis of Communication and Negotiation

BNCOW – Business Negotiation Competencies in Online World

ET – Eye-Tracking

FE – Facial Expression

FEA – Facial Expression Analysis

FACS – Facial Action Coding System

GDPR – General Data Protection Regulation

IMPULSE – Investigating Metrics: Participation Understanding Lecture Engagement

POV – Point of View

PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RDC – Remote Data Collection

TSL – Technical Sales Lab

UASD – University of Applied Sciences Düsseldorf

Abstract

This thesis proposes and validates a methodology for studying students' attention and emotional engagement in hybrid learning environments through Eye-Tracking and Facial Expression Analysis. Building on pilot tests and a main study at the Technical Sales Lab of the University of Applied Sciences Düsseldorf, the research integrates biometric data collection—Neon Invisible Glasses for in-person Eye-Tracking, standard webcams for remote participants, and dedicated webcams for in-person FEA—to capture real-time indicators of visual focus and emotional states. By integrating these data streams into iMotions for analysis and combining them with self-reported surveys, the study identifies differences in gaze behavior and emotional expression between in-person and remote students.

Results

The results of this study show that in-person participants exhibit faster orientation to the lecturer, more balanced attention between slides and lecturer and a wider range of emotional and facial expressions, suggesting greater social and cognitive engagement. Remote participants, by contrast, tend to fixate more on slides and display fewer expressive variations, highlighting the need for improved instructor visibility, clear slide transitions, and interactive cues. The results underscore the efficacy of ET and FEA as tools for data-driven course design. While the short 10-minute session and modest sample size limit generalizability, the workflow presented provides a scalable foundation for further research, potentially informing more inclusive and effective hybrid teaching strategies over longer periods of time and larger cohorts.

Conclusions

Despite the short session and modest sample size, ET and FEA provided actionable insights for optimizing hybrid lectures. By improving lecturer visibility, structuring slide transitions and adding interactive cues, the attention and engagement levels of remote students in hybrid settings can be better approximated to those of in-person students, providing a foundation for future studies.

Introduction

Background and context

The digital transformation of education has brought significant changes to teaching and learning formats in recent years. The shift to digital and hybrid formats has expanded access to educational content, offering students greater flexibility in how they engage with their studies. However, these advancements have also introduced new pedagogical challenges, as traditional in-person teaching relies on direct interaction and nonverbal communication, whereas digital learning environments require adapted instructional methods to sustain attention and engagement.

The COVID-19 pandemic acted as a catalyst for these changes, forcing universities worldwide to transition their lectures and assessments to digital formats. UNESCO reported that, at the peak of the pandemic in 2020, nearly 1.6 billion students across 190 countries were affected by the closure of educational institutions, prompting an unprecedented shift toward remote and hybrid teaching solutions.[1]

Even after lockdown restrictions were lifted, safety measures such as mandatory COVID-19 vaccinations and regular testing continued to impact in-person attendance, further solidifying hybrid learning as a permanent feature in higher education. These developments emphasize the growing importance of hybrid learning models in modern education.

Hybrid lectures combine in-person and remote participation. This allows students to attend lectures remotely while studying abroad or overcoming health-related barriers. Despite their benefits, hybrid learning environments pose significant challenges for educators. They face the task of ensuring equal engagement for both in-person and remote students while managing interactions across two distinct settings. Additionally, educators encounter the challenge of balancing the needs of both groups, so that remote students do not feel disadvantaged compared to those in the classroom. This involves overcoming technical limitations, sustaining student engagement in a dual-format environment, and fostering interactive discussions that seamlessly integrate remote and in-person participants.

[1] UNESCO, "COVID-19 Educational Disruption and Response," 2020.

Digital Education Monitor #2

A report from the *Digital Education Monitor #2* (Bertelsmann Stiftung, 2017) indicates that while many universities in Germany have robust technical infrastructures—such as stable campus-wide Wi-Fi and learning management systems—pedagogical innovation often remains limited. Instructors frequently employ digital tools only for basic tasks, such as uploading PDFs or slides, rather than more interactive methods like flipped classrooms or real-time discussions. Meanwhile, students generally welcome multimedia content and flexible learning opportunities but have varied motivations and digital proficiency levels. Nearly 80% of those surveyed felt that digital teaching methods were underutilized, calling for more diverse, interactive approaches and better integration of technology into instruction.[2]

Overall, the *Digital Education Monitor #2* findings indicate that hybrid lectures, backed by a robust digital strategy, can become a critical driver of modern, inclusive higher education—bridging technological possibilities with innovative teaching methods that genuinely enhance student learning.

This reveals a critical challenge: traditional evaluation methods, often reliant on retrospective self-assessments, fail to capture the real-time, dynamic interactions that occur in a hybrid lecture. Consequently, there is a pressing need for more quantitative measurement techniques—such as **Eye-Tracking (ET)** and **Facial Expression Analysis (FEA)**—to gather a comprehensive understanding of students' visual attention and emotional engagement.

Technical Sales Lab

Building on these challenges, the study is anchored in the work conducted at the **Technical Sales Lab (TSL)** of the **University of Applied Sciences Düsseldorf (UASD)** an interdisciplinary center where research, teaching, and practical applications in the field of technical and digital sales are combined. Under the leadership of Prof. Dr. Kati Lang, the TSL provides state-of-the-art, specially equipped facilities, enabling the implementation of innovative teaching formats,

[2] Bertelsmann Stiftung, *Monitor Digitale Bildung #2 – Die Hochschulen im digitalen Zeitalter*, 2017.

practice-oriented workshops, and close collaborations with renowned industry partners. This facility addresses the evolving market demands by not only teaching traditional sales and communication strategies but also integrating digital methods and technologies.[3]

As a result, the TSL offers an ideal platform for conducting practice-oriented research on current challenges in higher education and digital sales, fostering the development of forward-thinking solutions. The hardware and software licenses utilized for this study are provided by the TSL, ensuring a technically advanced research environment.

For further details on the TSL and its initiatives, please visit the official website of the University of Applied Sciences Düsseldorf:
https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/default.aspx

iMotions

iMotions A/S is a company, which specialized in human behavior research. It offers an integrated software platform designed for multimodal biometric research, enabling the simultaneous collection and synchronization of various data—such as eye trackers, facial expressions, and physiological devices.[4]

It streamlines the entire research process, from stimulus presentation to real-time data aggregation, allowing researchers to measure participants' cognitive and emotional responses with minimal manual intervention.

In this study, iMotions was utilized to import and synchronize data from remote and in-person participants. By integrating these data streams within a single platform, iMotions facilitated efficient annotation, confidence thresholding, and subsequent export of metrics for further analysis. This approach ensured consistency and comparability across the different study conditions, forming a robust foundation for investigating student engagement in hybrid lectures.

[3] Hochschule Düsseldorf, "Technical Sales Lab," Hochschule Düsseldorf. [Online]. Available: https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/default.aspx. Accessed: Mar. 1, 2025.

[4] iMotions, "iMotions – We Power Human Insight." [Online]. Available: https://imotions.com/. Accessed: Mar. 1, 2025.

Ring Project: „ET and FEA at the TSL”

The Ring Project is an interdisciplinary teaching and research initiative at the TSL of the UASD. It provides students with the opportunity to conduct practice-oriented research projects in collaboration with industry partners and academic institutions. Within this framework, students work in teams on real-world challenges in areas such as technical sales, digital business models, and innovative analytical techniques. The project bridges theoretical concepts with practical applications, fostering independent problem-solving under academic supervision.

As the first project at UASD to systematically explore the potential of ET and FEA, our team of five students focused on the technical implementation of hardware and software configuration for data collection at the TSL. Additionally, we examined potential applications for these technologies in higher education and corporate environments, with the objective of conducting one of the identified use cases at the TSL, generating meaningful data, and verifying the robustness of the setup.

During the course of the project, two key application areas for ET and FEA at the TSL were identified:

1. **Negotiation training for students:** Simulated negotiation scenarios can be analyzed using ET and FEA, allowing students to reflect on their negotiation strategies based on gaze behavior and emotional responses. This method provides valuable support in preparing for salary negotiations or contract discussions in professional settings.
2. **Assessment of didactic methods in teaching:** By analyzing ET and **Facial Expressions (FE)** data, it is possible to examine which instructional content and presentation styles capture students’ attention most effectively. This enables a targeted evaluation of teaching materials and lecture techniques, helping to minimize cognitive load and enhance student engagement.

After an intensive technical training phase in iMotions and a market research analysis of potential applications, the team decided to proceed with the evaluation of one of the identified use cases: **negotiations**.

At the TSL, the elective course *Digital Negotiations* (formerly *Market Psychology and Behavioral Management*) introduces students to digital negotiation processes and their unique challenges. It combines theoretical foundations with practical applications, equipping students with strategies for digital negotiation environments. As part of the course, an exploratory field test was conducted using eye-tracking and facial expression analysis to examine negotiation behavior. The scenarios were designed to elicit strong emotional responses, including negotiations over reserved parking spaces, job interviews, and IT service contracts.

Summary of Findings and Research Challenges

The Ring Project successfully identified two relevant use cases and confirmed that the TSL setup is robust and capable of producing reliable data when properly planned and executed.

However, several challenges and limitations emerged during the research process:

- **Data Interpretation**: A deep understanding of the theoretical foundations of emotion research and ET was necessary to derive valid conclusions from the data.
- **Defining Meaningful Metrics:** Clear metrics for measuring negotiation behavior and attention needed to be established to ensure reliable results.
- **Training in Analytical Tools:** The use of iMotions requires extensive training to conduct accurate analyses and ensure high data quality.

iMotions Workshop

Following the Ring Project, an iMotions Workshop was held at the TSL to address previously identified challenges and establish best practices for research at the UASD using ET and FEA.

The goal of this workshop was to provide in-depth training on study design, data quality assurance, and the use of iMotions for biometric analysis. This initiative was essential to further establish scientific research using ET and FEA at UASD. The workshop brought together experts from iMotions, scientific staff from the TSL, and students, allowing participants to gain hands-on experience with study design, data processing, and the application of iMotions analysis tools.

On the first day of the workshop, the specific challenges of data collection were intensively discussed. iMotions experts presented the capabilities and applications of the integrated analysis tools, focusing on the precise tracking of gaze patterns and emotional responses. Additionally, the **General Data Protection Regulation (GDPR)** officer of UASD was invited to discuss data protection regulations, including aspects such as cloud storage and ethical considerations. This ensured that participants understood the legal framework for conducting studies and the necessary precautions to comply with data privacy laws. The discussion provided a solid foundation for assessing the potential applications of ET and FEA within academic research at UASD.

On the second day of the workshop, a practical hands-on session took center stage. Through practical exercises, participants learned how to set up study configurations in iMotions and explored the key criteria for data collection and analysis. During this session, exclusion criteria for data analysis were defined, and participants engaged in extensive training with the integrated iMotions analysis tools to ensure precise interpretation of biometric data.

This workshop laid the foundation for three pioneering research projects, developed in close collaboration between students, faculty, and experts at the TSL:

Business Negotiation Competencies in Online World (BNCOW):

The **BNCOW project** was established to address the growing demands of digital business negotiations. With the increasing shift towards online interactions, accelerated by the COVID-19 pandemic, the project aims to equip students with the necessary skills to navigate virtual negotiation environments effectively.

In addition, research is being conducted on the use of FEA and ET in negotiations and customer behavior. This involves the application of advanced technologies such as the **iMotions software** and **Neon Glasses from Pupil Labs**.[5]

Progress reports on the BNCOW project, detailing research findings and developments, has been published on the Hochschule Düsseldorf website. For more information, visit:

https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/BNCOW.aspx

[5] Hochschule Düsseldorf, “BNCOW – Business Negotiation Competencies in Online World,” Hochschule Düsseldorf. [Online]. Available: https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/BNCOW.aspx. Accessed: Mar. 5, 2025.

Behavioral Analysis of Communication and Negotiation (BEACON):

The **BEACON project** explores the role of **nonverbal communication** in negotiations using ET and FEA. It aims to identify behavioral patterns that influence negotiation success and develop best practices for training. The findings will be applied in both academic research and practical negotiation strategies, following ethical guidelines.[6]

Investigating Metrics of Participation, Understanding, and Lecture Engagement (IMPULSE):

The **IMPULSE project** investigates the key factors influencing student comprehension, participation and engagement during lectures. By utilizing ET and FEA, the project aims to develop objective metrics for measuring learning involvement and optimizing teaching methodologies. Given the critical role of motivation and active engagement in academic success, IMPULSE analyzes cognitive attention patterns and emotional states to identify areas for instructional improvement. The study involves students from various disciplines in **real-world lecture settings**, measuring visual attention and emotional responses to compare the effectiveness of interactive teaching methods versus traditional lecture formats.[7]

After outlining the developments in education and the challenges faced by both educators and students, the focus shifted to UASD, which—through the TSL—serves as the perfect innovation hub to address the challenges of digital transformation in teaching. The historical progression of research at the TSL was presented, illustrating how knowledge in ET and FEA was developed and how pioneering research projects emerged. Building on the IMPULSE project, the objectives of this study were defined and are presented in the following section.

---

[6] Hochschule Düsseldorf, "BEACON: Behavioral Analysis of Communication and Negotiation," Hochschule Düsseldorf. [Online]. Available: https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/BEACON.aspx. Accessed: Mar. 5, 2025.

[7] Hochschule Düsseldorf, "IMPULSE: Investigating Metrics of Participation, Understanding, and Lecture Engagement," Hochschule Düsseldorf. [Online]. Available: https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/IMPULSE.aspx. Accessed: Mar. 5, 2025.

Objectives of the Study

The primary objective of this research, as part of the IMPULSE project, is to develop and validate an effective methodology for collecting and analyzing ET and **facial expression (FE)** data in hybrid learning environments. By examining both in-person and remote participants, the study aims to provide data-driven insights into visual attention and emotional engagement, supporting the optimization of hybrid lectures and the enhancement of pedagogical strategies.

To provide a clear understanding of how the research questions, expected outcomes, and measurement methods interrelate, Figure 1 below outlines the key elements of this study. This overview helps readers immediately see the logical connection between the theoretical assumptions, the hypotheses, and the measurement methods used to test them.

Research Question	Expected Outcome / Hypothesis	Method
RQ1: What technical and methodological challenges arise when collecting data in hybrid learning environments, and how can they be addressed through methodological adjustments?	Identification of key challenges that can be mitigated by a standardized workflow.	Field test results: process evaluations data quality checks
RQ2: What differences in gaze behavior and emotional responses can be observed between students attending in-person and those participating remotely?	**H1:** Visual attention and emotional engagement are more pronounced for in-person students compared to remote students.	Combined analysis of ET and FE data, supplemented by self-assessment surveys
RQ3: Which optimization strategies for hybrid learning environments can be derived from the analysis of collected data?	**H2:** The analysis of ET and FE data provides actionable insights for optimizing hybrid lectures.	ET metrics (gaze-, fixation-based metrics) FEA (engagement scores)

Figure 1: Overview of Research Questions, Expected Outcomes and Measurement Methods

As illustrated, the study addresses both practical challenges—such as methodological adjustments, data quality—and theoretical considerations, including potential differences in visual attention and emotional engagement between in-person and remote students.

This research aims to explore the current lecture dynamics in hybrid learning environments. By integrating ET and FEA in real classroom settings, the study seeks to provide insights that could help refine pedagogical strategies, improve hybrid lectures, and better address the challenges of engaging both in-person and remote students. While further research will be needed to fully validate these approaches, the findings may serve as a foundation for future innovations in digital education, contributing to the development of more interactive, inclusive, and effective hybrid learning environments.

Structure of the Thesis

The thesis begins with an **Introduction** that establishes the relevance of hybrid learning environments, outlines the research problem and goals, and situates the study within the TSL. The Introduction ends by explaining the main focus of the thesis—namely what will be examined—together with the research questions and hypotheses.

Next, the **Literature Review** defines key concepts of attention and engagement, presents essential ET metrics and introduces the foundations of FEA. The review also examines the hybrid learning context and identifies gaps in current research that this study aims to fill.

Following that, the **Methodology** details the research design, including common elements such as ethical approvals, stimuli, and the synchronization process. It summarizes the pilot study findings and shows how these informed methodological refinements for the three main studies: **Remote Data Collection (RDC)**, In-Person ET, and In-Person FEA.

In the **Results** section, the thesis presents ET and FEA findings. A comparative descriptive analysis of in-person and remote conditions follows, along with qualitative data from surveys.

The **Discussion** evaluates the findings in relation to the previously formulated hypotheses, addresses methodological considerations, and highlights the study's limitations.

Finally, the **Conclusion and Outlook** summarizes the study's key contributions, proposes recommendations for educators and outlines limitations that suggest avenues for future research.

A full **References** list and an **Appendix** of supporting materials, such as surveys conclude the thesis.

This structured approach ensures a clear and logical progression, guiding the reader from the theoretical foundation through the practical implementation to the critical discussion of the results. As a result, the study provides a comprehensive and systematic investigation into the integration of measurement methods in real-world hybrid learning environments.

Literature Review

This literature review aims to systematically explore the most widely recognized measurement techniques for ET and FEA, shedding light on their conceptual foundations and methodological nuances. By delineating the tools and frameworks necessary for robust data collection, the review elucidates how these methods enable the effective generation of empirically grounded insights that address the research questions at hand. Ultimately, serving as the underpinning for developing a refined methodological approach to investigate attention and engagement in real world hybrid learning environments.

Scope and Search Strategy

A systematic literature search was conducted to ensure a comprehensive and unbiased selection of relevant studies. The review followed a framework inspired by the **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)** 2020 guidelines (Page et al., 2021), ensuring transparency and reproducibility.[8]

[8] M. J. Page *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, p. n71, 2021.

The search process involved the following four stages:

1. **Identification**: Databases such as Scopus, IEEE Xplore, ScienceDirect, ResearchGate, Litmaps, and Elicit were searched using keywords related to *"hybrid learning," "eye-tracking," "facial expression analysis," "attention," and "engagement."*
2. **Screening**: Duplicates were removed, and an initial evaluation was conducted based on titles and abstracts.
3. **Eligibility**: Full-text evaluation was performed to assess relevance to hybrid learning contexts.
4. **Inclusion**: Only peer-reviewed articles and information from official websites were considered.

To enhance efficiency, Litmaps was used to visualize citation networks, while Elicit facilitated abstract summarization to ensure a focused selection of literature. Through this search, a body of literature addressing attention, engagement, and their measurement methods in hybrid learning environments was identified.

By systematically reviewing existing research, this chapter identifies key theoretical and methodological gaps, which serve as the foundation for the methodological approach of this study.

Eye-Tracking in Educational Research

In contemporary academic research, ET has gained prominence as an essential methodology for examining student behavior and cognitive processes. Analyzing visual attention can provide valuable insights into the cognitive processes behind problem-solving, learning, and teaching. Research has shown that ET can help uncover how individuals engage with problem-solving tasks[9], enhance learning and memory processes[10] and improve multimedia learning strategies.[11] Additionally, studies have demonstrated the role of teachers' gaze in influencing student

[9] L. van Marlen, M. van Wermeskerken, H. Jarodzka, and T. van Gog, “Effectiveness of eye movement modeling examples in problem solving: The role of verbal ambiguity and prior knowledge,” *J. Comput. Assist. Learn.*, vol. 34, no. 6, pp. 717–727, 2018.
[10] M. Chan, C. R. Madan, and A. Singhal, “The influence of visual attention on learning and memory: Insights from eye-tracking studies,” *Psychon. Bull. Rev.*, vol. 29, pp. 1234–1250, 2022.
[11] T. van Gog and K. Scheiter, “Eye tracking as a tool to study and enhance multimedia learning,” *Learn. Instr.*, vol. 20, no. 2, pp. 95–99, 2010.

engagement and comprehension in both physical and online learning environments.[12]

By examining how students direct their gaze, educators and researchers can gain insights into how information is encoded and retrieved, facilitating the identification of learning barriers and the development of targeted teaching strategies.

Foundational Theories and Works

Before delving into the studies, it is important to understand how these innovative methodologies have evolved from earlier foundational theories and works. Building on the significance of ET and FEA in capturing cognitive and emotional dynamics, the following section examines the seminal works that have paved the way for current research.

Andrew T. Duchowski (2007) provided a foundational examination of how ET can serve as a direct measure of human attentional behavior across a range of real-world settings. By highlighting that eye movements offer insights into visual, cognitive, and attentional processes, his work underscores the importance of applying these methodologies in contexts where understanding user interaction is crucial. For the present study, Duchowski's perspective reinforces the value of integrating ET to capture nuanced behavioral data, which can inform more effective instructional design and enhance overall user engagement. Furthermore, *Eye Tracking Methodology* itself delivers a comprehensive overview of the technical, methodological, and theoretical aspects of ET—covering hardware requirements, data analysis methods, and real-world applications—making it a valuable reference for anyone seeking a deeper understanding of gaze-based research tools.[13]

Holmqvist et al. (2011) offer an extremely detailed presentation of all relevant aspects of gaze research in *Eye Tracking: A Comprehensive Guide to Methods and Measures*. The work not only explains the technical requirements of different Eye-Tracking systems but also provides concrete guidelines for planning and conducting studies, including study designs, data collection protocols and the targeted selection of participants. In addition, the authors address the methodological challenges of

---

[12] H. Jarodzka, N. Janssen, P. A. Kirschner, and G. Erkens, "Teacher's gaze and its relation to students' engagement and learning," *Comput. Educ.*, vol. 175, p. 104324, 2021.

[13] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 2nd ed. London, UK: Springer, 2007.

data analysis, offering recommendations on how best to segment, classify, and interpret gaze data. Practical considerations, such as avoiding measurement artifacts, ensuring high data quality, and reflecting on ethical aspects, are also discussed in depth. Due to this holistic approach, the book has become a standard reference in the field, widely cited by researchers seeking robust and transparent ET methods.[14]

Together, the works of Duchowski (2007) and Holmqvist et al. (2011) provide the methodological foundation for the ET measurement techniques applied in this study. While Duchowski's contributions emphasize the theoretical and practical significance of ET for understanding attentional and cognitive processes, Holmqvist et al. offer a comprehensive framework for the technical execution and methodological rigor of gaze-based research. By integrating these established principles, the present study ensures a structured and validated approach to capturing and analyzing visual attention dynamics.

Ekman's Facial Action Coding System

Paul Ekman's **Facial Action Coding System (FACS)** is a seminal work developed to systematically describe and categorize facial movements based on observable muscle actions. The system decomposes FE into individual components known as **Action Units (AUs)**, each corresponding to the contraction or relaxation of specific facial muscles.[15]

This approach allows researchers to capture subtle variations in expressions that may indicate a range of emotional states, such as joy, anger, or sadness.

In the context of measuring emotional engagement, FACS provides several key advantages. First, it offers an objective and standardized method to decode complex FE, enabling a more reliable assessment of emotional responses compared to self-reported measures. By linking specific AUs to corresponding emotional states, researchers can quantify and compare the intensity and frequency of these emotions across different contexts, such as hybrid learning environments. For instance, variations in AU activation may reflect students' levels of interest or

---

[14] K. Holmqvist, M. Nystrom, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford, U.K.: Oxford University Press, 2011.

[15] P. Ekman and W. V. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press, 1978.

frustration during a lecture, providing actionable insights into the effectiveness of instructional methods.
Furthermore, the integration of FACS into automated FEA tools—such as those implemented in platforms like iMotions with Affectiva—has significantly enhanced its applicability in real-time research settings. These tools leverage advanced algorithms and large datasets to interpret FACS-derived metrics, thus allowing for the dynamic monitoring of emotional engagement during learning sessions. By capturing the temporal patterns of FE, researchers can correlate these with other biometric data (e.g., ET) to obtain a comprehensive picture of how emotional and cognitive processes interact during learning.

Overall, Ekman's FACS remains a cornerstone in the field of FE research. Its methodological rigor and adaptability in digital analytics make it an invaluable tool for investigating the nuanced ways in which emotional engagement influences learning outcomes, particularly in innovative educational settings like hybrid lectures.

The foundational works outlined in this section have established the theoretical and methodological basis for utilizing ET and FEA in research.

Advancements in Engagement Measurement

Understanding student engagement is crucial for improving learning outcomes. Research on engagement measurement has evolved from theoretical models to advanced automated detection methods. The following studies provide an overview of key contributions in this field.

Fredricks, Blumenfeld, and Paris (2004) established an influential framework for school engagement, categorizing it into three dimensions: **behavioral engagement** (active participation and adherence to school norms), **emotional engagement** (positive or negative affective responses to school and teachers), and **cognitive engagement** (intensive mental effort to overcome academic challenges). This multidimensional perspective enables a nuanced understanding of the factors that

foster or hinder engagement and highlights engagement as a mechanism between the learning environment and academic achievement.[16]
Particularly relevant to the present study is the interactive nature of these engagement dimensions, as an emotional assessment of engagement provides deeper insights into learning motivation and interaction.

Whitehill et al. (2014) examined the automatic recognition of student engagement through FEA. Their study demonstrated that both human observers and machine learning algorithms could reliably distinguish between high and low engagement levels, with machine classifiers achieving comparable accuracy to human assessments. The results indicate that static facial features provide significant cues for engagement and that algorithmic models can efficiently process this information. Furthermore, their study found a correlation between automated engagement scoring and academic performance, underscoring the relevance of computer-based engagement analysis for educational applications.[17]

These findings provide a critical foundation for the present study, highlighting the role of FEA as a quantitative method to capture engagement in hybrid learning settings and refine instructional strategies accordingly.

Dewan, Murshed, and Lin (2019) investigated various methods for detecting engagement in online learning environments and identified computer-based approaches as particularly promising. Their analysis suggests that automatically detecting engagement through facial expressions is an effective and non-invasive alternative to traditional methods such as self-reports or observational checklists. The recognition of Facial Action Units plays a central role, as specific muscle movements correlate with affective states such as interest, frustration, or boredom. The authors emphasize that accurately capturing emotional responses can provide valuable insights into student engagement.[18]

[16] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris, *School Engagement: Potential of the Concept, State of the Evidence*, vol. 74, no. 1. London, UK: SAGE Publications, 2004.
[17] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. Movellan, *The Faces of Engagement: Automatic Recognition of Student Engagement from Facial Expressions*, vol. 5, no. 1. New York, NY, USA: IEEE Transactions on Affective Computing, 2014.
[18] M. A. Dewan, M. Murshed, and F. Lin, *Engagement Detection in Online Learning: A Review*, vol. 6, no. 1. Cham, Switzerland: Springer, 2019.

This perspective aligns with the present study, as it supports the methodological decision to rely on FEA for engagement assessment, ensuring a robust and scalable approach for evaluating learning interactions in hybrid environments.

De Carolis et al. (2019) explored the automatic detection and monitoring of student engagement by analyzing facial expressions, head movements, and gaze behavior. Their study is based on Flow Theory, which conceptualizes engagement as a combination of concentration, interest, and enjoyment, shaped by an optimal balance of challenge and skill. The authors developed a real-time engagement analysis system that utilizes a Long Short-Term Memory network to process facial data and classify engagement on a four-point scale. Their findings show that facial expressions, particularly AUs, provide a reliable basis for engagement recognition and correlate with subjective engagement assessments.[19]
These findings are particularly relevant to the present study, as they support the validity of emotion-based engagement detection and offer a strong methodological foundation for integrating FEA into hybrid learning environments.

The reviewed studies demonstrate that the automatic detection of engagement through FEA is a valid method for measuring learning interaction and motivation. While Fredricks et al. (2004) provide the theoretical foundation for the concept of engagement, Whitehill et al. (2014) and Dewan et al. (2019) show that automated Facial Action Unit based approaches can offer reliable insights into emotional states. The work of De Carolis et al. (2019) further confirms that integrating gaze behavior can contribute additional information to engagement detection.
For the present study, emotion-based engagement measurement through FEA is particularly significant, as it serves as the foundation for a data-driven optimization of hybrid learning formats.

[19] B. De Carolis *et al.*, “Engaged faces: Measuring and monitoring student engagement from face and gaze behavior,” in *Proc. IEEE/WIC/ACM Web Intell. Workshops*, 2019.

Understanding

To ensure a clear understanding of the ET methods used in this study, it is essential to define key gaze-related concepts as described by Holmqvist et al. (2011).

1. Gaze

Gaze refers to the direction in which a person is looking at a given moment. In eye-tracking analysis, gaze data helps determine where visual attention is directed and for how long. It consists of fixations and saccades.

2. Fixations

A fixation is a brief period during which the eye remains relatively stable on a specific location. Longer fixation durations may suggest deeper cognitive processing of the observed content, while shorter fixations may indicate rapid scanning or difficulty in processing information.

3. Saccades

Saccades are rapid eye movements (between fixations), enabling visual exploration. In ET, they help analyze scan patterns, cognitive load, and search efficiency. Longer saccades may indicate active exploration, while shorter ones may signal uncertainty or effort in processing information.

Analysis

Areas of Interest (AOIs) are clearly defined regions within a visual stimulus used as reference points for analyzing gaze behavior. Examining these regions enables precise insights into the distribution and intensity of visual attention.

Below are concise explanations of **Dwell Time**, **Revisit Count**, **Fixation Count**, **Fixation Duration** and **Time-to-First Fixation (TTFF)**, as described by Holmqvist et al. (2011).

Dwell Time

Refers to the total amount of time—encompassing both fixations and any associated saccades—that a viewer spends within a predefined AOI. A longer dwell time generally indicates sustained attention to that region.

Revisit Count

Indicates how many times a viewer’s gaze returns to the same AOI after leaving it. Multiple revisits may signal heightened interest or the need to re-check particular information.

Fixation Count

Represents the total number of fixations recorded within an AOI. A higher fixation count can suggest either increased interest or difficulties in processing the content located there.

Fixation Duration

Describes the average length of each fixation in milliseconds. Longer fixation durations often correlate with deeper cognitive processing or more detailed inspection of the stimulus.

Time-to-First Fixation

Measures the elapsed time from the onset of a stimulus (e.g., when an AOI first appears on screen) until the viewer’s initial fixation on that AOI. A shorter TTFF typically implies that the AOI quickly captures attention.

These AOI metrics provide valuable insights by:

- **Identifying key instructional content** (e.g., central information on slides or instructor interactions).
- **Evaluating** design elements that immediately capture students’ attention.
- **Detecting** potential attention gaps that may highlight areas for instructional improvement.

By integrating these metrics into the empirical analysis, targeted strategies can be developed to enhance both cognitive and emotional engagement, ultimately supporting improved learning outcomes.

Heatmaps: Interpretation and Relevance to Research Objectives

Heatmaps are powerful visualizations that overlay gaze data onto stimuli, providing insights into the distribution and intensity of visual attention.

Bojko (2009) examined heatmaps as a tool in ET research, focusing on their interpretation, methodological application, and limitations. By distinguishing between various types of heatmaps and discussing their strengths and potential misinterpretations, Bojko's research establishes best practices for their application in empirical studies, making it a fundamental reference for evaluating visual attention in diverse contexts, including educational research. Heatmaps serve as an intuitive tool to visualize attention distribution. These maps use a color-coded gradient, where red areas indicate the highest concentration of fixations (i.e., areas of prolonged visual focus), while cooler colors (yellow, green, blue) signify lower engagement levels or sporadic viewing. This makes heatmaps an effective means of interpreting attentional hotspots and identifying overlooked areas in instructional content.[20]

This study utilizes fixation-based heatmaps, which aggregate fixation locations and durations across multiple participants to visualize attention distribution in hybrid learning environments. These heatmaps help identify high-attention zones, ensuring that key instructional elements—such as lecture slides, instructor gestures—effectively capture student focus. Additionally, they reveal attention gaps, highlighting areas that receive little to no visual attention and may require instructional adjustments.

Metrics for measuring Engagement via FEA

The Engagement metrics were chosen to directly measure how emotionally involved students are.

Emotional states can be challenging to interpret; therefore, additional emotional, FE and behavioral metrics are utilized to provide a more comprehensive perspective.

Confusion is important because it can show whether students are struggling with the material, which might motivate them to think more deeply or signal the need for extra support.

Neutral provides a baseline to compare other emotional changes, such as moments of high engagement.

[20] A. Bojko, "Informative or misleading? Heatmaps deconstructed," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 53, no. 27, 2009.

To get more details about students' reactions, certain facial expressions were included. For example, Brow Furrow and Lip Press can indicate concentration or mild frustration; Brow Raise, Eye Widen, and Smile can show surprise, interest, or satisfaction. These additional cues help explain why a student might have a moderate engagement score but a high confusion level: they might be focusing hard on difficult content rather than becoming disengaged.

Finally, behavioral metrics like head movements (Pitch, Yaw, Roll) detect changes in posture or orientation. A learner might keep a neutral face but still turn away from the screen, which could suggest reduced attention. Adding these metrics helps confirm or clarify the emotional data.

Overall, these emotional and behavioral metrics create a clear picture of each student's experience. Engagement remains the main measure, while other indicators help resolve uncertainties and give a richer view of student behavior in a hybrid learning setting.

This combination of metrics allows for a comprehensive assessment of emotional engagement, integrating both the affective and behavioral dimensions to provide actionable insights for improving hybrid lectures.

Conclusion of the Literature Review

While previous studies confirm the potential of ET and FEA in education, their application in hybrid classrooms remains underexplored. Most research is conducted in controlled settings, limiting its relevance for real-world lectures. Hybrid learning introduces unique challenges, such as technological variability, engagement differences between in-person and remote students, and ethical concerns in data collection.

To address these gaps, this study develops a robust methodology ensuring valid and scalable data collection while accounting for technical, ethical, and practical constraints.

The following Methodology section details the implementation of ET and FEA, including the data collection process, experimental setup, and analysis procedures, to ensure a structured and reproducible approach to measuring student attention and engagement in hybrid lectures.

Methodology

This study adopts a mixed-method research design with a primarily quantitative focus to investigate how ET and FE data can be used to analyze student attention and engagement in hybrid lectures, as well as the methodological challenges associated with data collection. By integrating real-time biometric data with students' own perceptions, the methodology delivers a comprehensive view of attention and engagement in hybrid learning environments to address the research questions.

Leveraging the iMotions software platform, the research setup is designed for minimal disruption to live lectures. Hardware configurations can be quickly assembled and disassembled, allowing for flexible use in multiple real-world scenarios. Data is recorded separately and later imported into iMotions for analysis.

High-Level Overview:

Phase	Purpose	Studies/Activities
Field Test I	Pilot study to identify technical limitations and refine data collection processes.	Technical checks. Workflow optimization.
Field Test II	Main study, consisting of three targeted studies:	**Study 1 (Remote Data Collection):** ET and FEA via standard webcams for remote participants.
		Study 2 (In-Person ET): ET with Neon Invisible Glasses.
		Study 3 (In-Person FEA): FEA using C920 Full HD Webcams.

Figure 2: Overview of Field Tests and Targeted Studies

The data collection is organized into two main field tests, each containing of three targeted studies, as shown in Figure 2.

Common Aspects of All Studies

The common aspects shared by all studies—including ethical guidelines, participant selection, stimuli, synchronization procedures—are presented in the following section to ensure a consistent and reproducible research framework.

FEA Analysis

The facial coding algorithm in iMotions assigns a **confidence score** to each detected FE, indicating the certainty of classification. To ensure data reliability, a confidence threshold is applied, excluding low-certainty detections. R Notebook—a coding and documentation tool—is used to filter out segments below this threshold, enabling binary classification of expressions (present/absent).

After thresholding, data is aggregated across all participants within each study to analyze expression frequency and distribution. Users can set either a global or expression-specific confidence threshold, with 50% representing moderate reliability, 75% high reliability, and below 25% considered unreliable.

In this study, a **50% threshold** was selected to **balance inclusivity** and **accuracy** in engagement-related expressions.

AOI Metrics

AOI Metrics	Metric Name	General Definition	Purpose	Link to RQs
Gaze-Based Metrics	**Respondent Count**	Number of participants who watched an AOI.	Ensures enough valid data are collected to compare visual attention in Remote vs. In-Person groups.	**RQ1:** Methodological Rigor
	Revisit Count	Tracks how many times the gaze returns to an AOI after leaving it. Multiple revisits can imply re-inspection or ongoing interest in content.	Detects repeated engagement with specific elements, indicating deeper involvement or rechecking of information (e.g., key slide sections).	**RQ2:** Key Lecture Elements
Fixation-Based Metrics	**Fixation Count**	Number of fixations in an AOI. Higher counts may reflect interest, complexity, or difficulty in processing.	Highlights which parts of the lecture (slides or lecturer) spark more cognitive effort or curiosity, guiding instructional design.	**RQ2:** Group Differences, **RQ3:** Key Lecture Elements
	Fixation Duration (ms)	Average length of each fixation. Longer durations often indicate deeper cognitive processing.	Suggests how intensively certain information is examined, shedding light on which content requires more focus.	**RQ2:** Group Differences, **RQ3:** Key Lecture Elements
	Dwell Time (%)	Percentage of total viewing time spent in an AOI. Prolonged dwell may signal higher informativeness or stronger interest in that region.	Reveals which aspects of the lecture (slides, lecturer) receive sustained gaze, helping identify elements that strongly capture attention.	**RQ3:** Key Lecture Elements

Figure 3: Overview of AOI Metrics for ET Studies

The table above presents the AOI metrics explained earlier in the literature review, used for the ET studies, and links them to the research questions.

Ethics and Approval

All procedures described in this research were **reviewed and approved** by the **Ethics Committee** of Hochschule Düsseldorf, ensuring compliance with institutional guidelines and data protection regulations. Prior to participation, each student received a **written explanation** of the study's purpose, procedures, and data handling measures. By signing an **informed consent** form, participants explicitly agreed to the collection and processing of their data (including video recordings, ET, FEA), with the understanding that:

1. **Participation was voluntary**
   Students could withdraw at any time without academic or personal penalties.
2. **Anonymity and confidentiality**
   All data was pseudonymized and securely stored on password-protected PCs in the TSL, which are accessible only to authorized personnel via keycard access. The data was used solely for academic and analytical purposes and was deleted after the analysis was completed.
3. **Right to information**
   Participants could request access to their data at any point or inquire about the study's general findings.
4. **Data protection**
   The methodology and analysis conformed to relevant data privacy laws (e.g., GDPR), as well as institutional guidelines set by the **Data Protection Officer** at UASD.

Adhering to these ethical standards not only **safeguards participant welfare** but also strengthens the **credibility** of the collected data and any conclusions drawn from it.

Inclusion Criteria

The following criteria applied to all participants:

- **Academic Background:** Participants had to be officially enrolled in the course where the studies were conducted to ensure familiarity with the learning context and content.
- **Language Proficiency:** Participants needed to have sufficient proficiency in the **language used during the study** to fully understand the learning content and tasks.

- **Attendance Requirement:** Participants were required to be present (either in-person or remotely) for the full duration of the session.
- **High Quality Data:** Based on iMotions indicators (ET Accuracy and Face Detection percentage).

Stimuli

The session lasted approximately **10 minutes** and included the following components:

- **Presentation Slides:** Structured slides covering **core Customer Relationship Management (CRM) concepts**.
- **Real-Time Lecture Input:** The lecturer delivered verbal explanations and live commentary for the slides, providing context and elaboration on key topics.
- **Interactive Questions:** The lecturer posed questions to the students, encouraging active participation and engagement.
- **Self-Assessment Survey:** At the end of the session, participants completed a self-assessment survey, allowing them to reflect on their lecture experience.

Lecture Environment

All participants attended a lecture on CRM, delivered in a regularly scheduled hybrid lecture in the assigned classroom and simultaneously broadcast via Microsoft Teams for remote participants. The experimental setup was carefully planned and implemented within 30 minutes before the start of the lecture.

Seating & Camera Setup

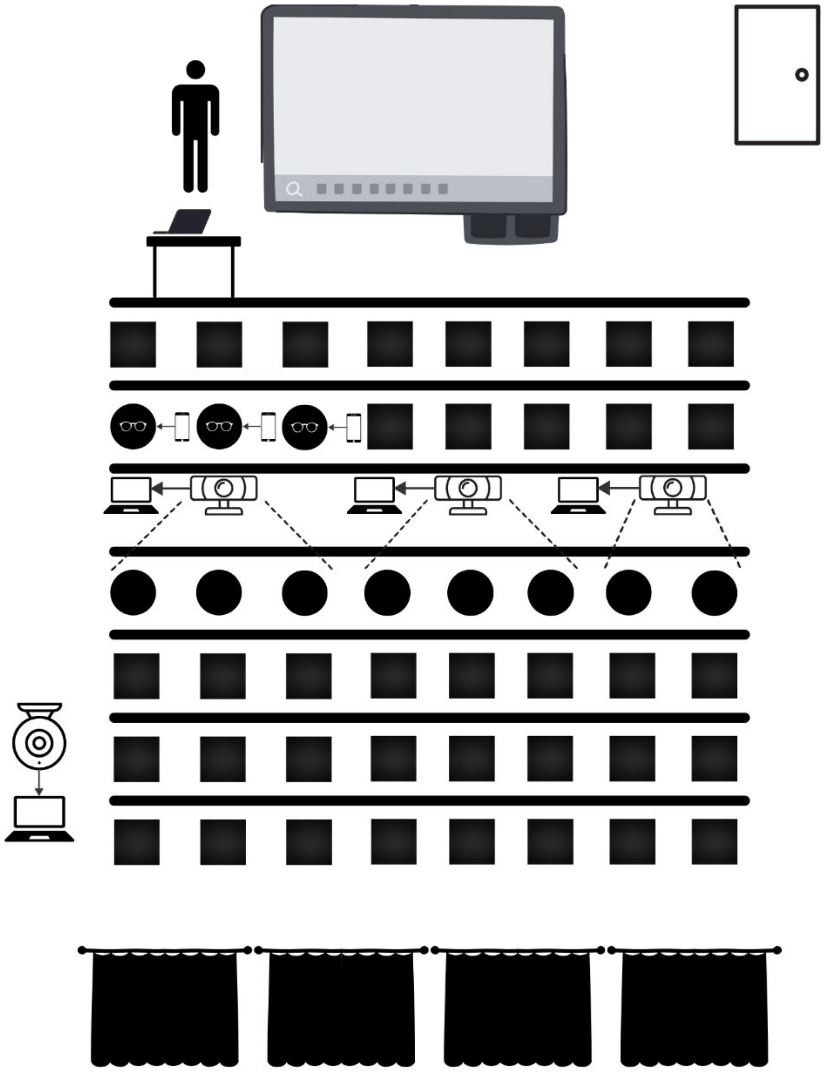


Figure 4: Classroom Setup

Above is the sketch of the classroom setup, including all hardware used for the in-person studies.

Round chairs represent the seating positions of the **in-person FEA** participants. **Round chairs with glasses** represent the seating positions of the **in-person ET** participants.

Specific areas were designated to ensure participant comfort, ethical considerations and optimal data collection:

- **Reserved Area for Non-Participants:** A section to the right of the in-person ET participants and the row in front of them was designated for students who did not wish to be recorded, ensuring voluntary participation.
- **Study 1: Remote Data Collection (RDC):** A backup room was booked to accommodate additional remote participants in case an insufficient number of students joined the lecture from home.

- **Study 2: In-Person ET:** A dedicated section was arranged for three ET participants to ensure optimal conditions.
- **Study 3: In-Person FEA Camera Placement:**
  - Cameras for the FEA study were positioned in the row behind the ET participants.
  - The eight participants in the in-person FEA group were seated one row behind the cameras. To prevent background interference, no one was allowed to sit behind them.
- **General Lecture Recording:** An **environmental camera** was placed in the room to capture an overview of the lecture setting, including:
  - The lecturer
  - The presentation slides

Lighting Conditions

To optimize the visual conditions, initial lighting adjustments were made before data collection:

- **Curtains were closed** to minimize backlight interference behind the FEA participant group, ensuring clearer facial recognition.
- **Room lighting** settings were adjusted to achieve consistent and even illumination.

Synchronization

At the beginning of the lecture, the instructor performed a **clap**, which was clearly captured in all recordings and served as the universal synchronization point. Event markers were then defined based on the environmental camera recording to establish consistent annotation durations. These standardized annotations were applied across all three studies, ensuring uniform segmentation of lecture events. The **time durations of the annotations** were determined based on the environmental camera recording, ensuring precise segmentation of lecture events.

The following figure presents the predefined event markers used for synchronization:

Event Marker	Duration (ms)
Clap	1
1 Introduction	23571
2 Slide Transition	5562
3 Outline of Lecture	53503
4 Slide Transition 2	5744
5 Definition CRM	160382
6 Slide Transition 3	4289
7 Relationship Management	410452

Figure 5: Event Markers for Synchronization and Data Analysis

This **standardized synchronization approach** served three key purposes:

1. **Within-Study Comparability** → Ensured that participants within the same study could be reliably compared.
2. **Cross-Study Comparability** → Prepared the data in a way that allows for descriptive comparisons between different study conditions.
3. **Analysis** → All **exported metrics and heatmaps** were based on these annotations, linking metrics to the segments with fixed durations.

By applying identical annotations across all datasets, this approach minimized timing inconsistencies, ensuring that any observed differences in attention and engagement were driven by participant behavior rather than variations in data collection timing.

General Workflow

A generalized four-step workflow was developed to ensure a structured and efficient procedure across all studies, comprising of:

1. **Preparation:**
    - Setup of Hardware & Software
    - Participant briefing & informed consent
2. **Data Collection:**
    - Instructor Clap for synchronization
    - Lecture recording (RDC, ET, FEA)

3. **Integration:**
    - Import into iMotions
    - Pseudonymization to protect participant confidentiality
4. **Analysis:**
    - Synchronization and application of annotation durations
    - Quantitative: Export key metrics
    - Qualitative: Heatmaps, Surveys

Field Test I - Pilot study

The pilot study was designed with several key objectives in mind. It aimed to assess the feasibility of the initial procedure and methods under actual classroom conditions in a hybrid lecture environment, evaluating both hardware and software performance—including the calibration protocols and synchronization using a pre-established clap marker. Additionally, the study sought to identify potential technical limitations and methodological challenges that might arise during data collection.

Participants

Study	Notes
RDC	17 participants, 9 completed
In-Person ET	3 participants, 3 completed
In-Person FEA	6 participants (3 participants from study 2 were included), 6 completed

Figure 6: Participants Pilot Study

Findings and Adjustments

The findings from the pilot study are presented in the table below, organized by study, with the most significant modifications needed for adjusting the procedure and methods for each study highlighted in bold:

Study	Category	Finding	Adjustment
Study 1: RDC	**Synchronization**	**Clap was not visible.**	**Check instructor position.**
	Data Quality	Respondent 0002 showed only 57% face detection.	Exclude from analysis.
		Respondent 0015 experienced signal loss during the session.	Exclude from analysis.
	Stimuli	AOIs were insufficiently defined.	Define AOIs to differentiate between slide text and lecturer.
Study 2: In-Person ET	**Data Quality**	**Brightness and contrast issues made slides unclear in ET recordings.**	**Adjust contrast settings for better clarity of slides in recordings. Backup: additional environment cam.**
	Distractions/ GDPR	One participant sent WhatsApp messages.	Provide clear pre-session instructions forbidding unrelated activities during data collection.
	Stimuli	AOIs did not account for detailed slide elements.	Define AOIs to differentiate between slide text and lecturer.
Study 3: In-Person FEA	**Data Quality**	**ET glasses obstructed eyebrows, affecting FE metrics.**	**Exclude ET-Group from this study.**
	Documentation	Lack of clear documentation for slide content and AOIs during analysis.	Implement additional environment cam.
	Capacity	Only 6 participants could be recorded due to hardware limitations.	Purchased additional webcam to increase recording capacity to 8 participants.
	Participant Setup	**Background faces in recordings interfered with FE data quality.**	**Prevent students from sitting behind the FE group.**

Figure 7: Findings and Adjustments for Methodology

The key findings from Field Test I – Pilot Study led to several **refinements**.

For **Study 1**, synchronization issues, such as the visibility of the clap marker, were addressed by adjusting the instructor's position.

For **Study 2**, visibility issues in ET recordings were resolved by adjusting contrast settings and introducing an environment camera as a backup. To minimize distractions and ensure GDPR compliance, clear pre-session instructions were introduced to prevent unrelated activities, such as mobile phone use, during data collection.

For **Study 3 (In-Person FEA)**, methodological adjustments were achieved by **excluding the ET group** due to the obstruction of facial expression metrics by ET glasses. The classroom setup was reorganized to prevent **background faces** from interfering with FE data quality. Additionally, an **extra webcam** was added to increase recording capacity from six to eight participants, and an **environment camera** was implemented to improve documentation of slide content and AOIs.

Conclusion

These refinements were integrated into both the **overall methodology** and the **specific procedures and methods** detailed in the next section for each study to enhance data quality and ensure reliable engagement measurement.

Procedures and Methods used for each Study

The specific procedures and methods presented are refined based on the findings from the pilot study. The following section details their implementation for each study, emphasizing tailored adaptations for remote and in-person data collection.

Study 1: Remote Data Collection

Objective:

Study 1 collects ET and FE data remotely using standard webcams installed in the participants' PCs. The goal is to measure predefined metrics related to visual attention and emotional engagement among remote participants.

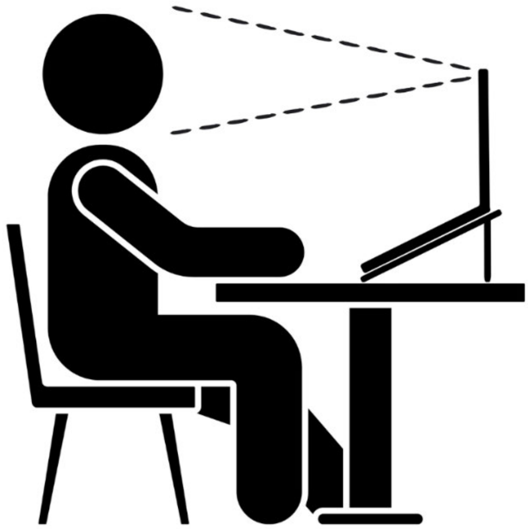


Figure 8: RDC Sketch

Materials Used:

- Standard webcams installed in participants' PCs
- Microsoft Teams for lecture participation and screen recording

Procedure Overview:

All steps follow the **common four-step workflow** introduced earlier.

1. Preparation

First, the **RDC** study is configured in iMotions Desktop and then uploaded to iMotions Cloud. Participants are informed in advance about the procedure and receive the following documents:

- An instruction document for the RDC study
- A consent form
- The link to the iMotions study, shared via Microsoft Teams

2. Data Collection

At the beginning of the session, participants receive verbal instructions regarding the study procedure, including the type of data collected, optimal lighting conditions, and correct head positioning. They then sign the consent form.

Once data collection is activated in iMotions Cloud, the researchers wait for all participants to complete the head and webcam check, pre-calibration, and join the Microsoft Teams call.

To synchronize biometric data with the recorded videos, the researcher performs a synchronization clap. In addition to biometric data collection, participants complete a standardized Likert scale survey immediately after the lecture as a second task in the iMotions study to capture subjective self-assessments. Finally, data collection in iMotions Cloud is stopped, and participants send their signed consent forms to the researcher.

3. Integration, Preprocessing, and Synchronization

After data collection is completed, the RDC study is exported from iMotions Cloud and downloaded. The data is then imported into iMotions Desktop for further processing, including:

- Checking data quality to identify issues and exclude unsuitable participants
- Pseudonymizing participant names and entering biometric data (age and gender)
- Post-processing data with Affectiva

4. Analysis

Synchronization

Event markers are then applied, starting with the clap cue for each video recording. The environment camera recording serves as a reference to consistently apply fixed-time event annotations (e.g., Clap, Introduction, Slide 1, Slide Transition 1, Slide 2) across all recordings.

ET Analysis

The analysis is conducted using predefined AOIs, with 'Lecturer' and 'Slide' designated as AOIs for the entire lecture duration. The following figure illustrates the student’s **point of view (POV)** along with the defined AOIs used in this study.

[redacted]

Figure 9: RDC student's POV with AOIs

Finally, the ET, FE metrics, heatmaps, and survey results described earlier are exported for further analysis.

Study 2: In-Person ET

Objective:

Study 2 collects ET data from in-person participants using Neon Invisible Eye-Tracking glasses to analyze visual attention during the lecture.

Materials Used:

- Three **Neon Invisible ET glasses**, connected via USB-C to a **OnePlus 10 Pro 5G**, with data stored locally using the **Neon Companion App**.
- **Survey** for additional qualitative data.

Procedure Overview:

All steps adhere to the **common four-step workflow** described earlier.

1. Preparation

Prior to the study, all necessary materials are prepared. Surveys and informed consent forms are printed and labeled with identifiers (ET_1, ET_2, ET_3) corresponding to the assigned ET glasses. The classroom is arranged accordingly. Mobile devices are updated to the latest OS version, and the Neon Companion App is checked for updates. Sufficient storage space is ensured for recordings, and batteries are fully charged. The ET glasses are cleaned and positioned in the classroom for efficient deployment. Additionally, contrast settings are optimized to ensure the slide content is clearly visible in the recordings.

2. Data Collection

At the beginning of the session, participants receive verbal instructions covering calibration procedures, the importance of minimizing head movement during the study, and the fact that all recordings will be collected and analyzed.

Participants then sign the informed consent form. Once consent is obtained, recordings are started on each mobile device. Then, the participants of Study 1 are awaited to complete the calibration process and join the Microsoft Teams call. After this the instructor performs a synchronization clap to establish a reference point for later alignment of the recordings. The lecture is then conducted while the ET data is continuously recorded. After the lecture, recordings are stopped on all mobile devices, and participants complete the survey. Finally, the ET glasses are packed up, and all informed consent forms and survey responses are collected and archived for further processing.

3. Integration and Preprocessing

Following data collection, the recorded ET data is exported from the Neon Companion App. Mobile devices are connected via USB-C to the computer for file transfer. A dedicated study folder (e.g., ET) is created, with separate subfolders for each participant labeled with pseudonymized IDs (ET_1, ET_2, ET_3).

The study is then set up in iMotions Desktop with the study type set to "Import Glasses Data." The corresponding study folder (ET) is selected for import, and fish-eye correction is applied to the data. A data quality check is performed to identify potential issues and exclude unsuitable participants. Biometric details, such as age and gender, are entered for each respondent, followed by post-processing steps.

4. Analysis

For each respondent, the introduced annotation procedure is applied, beginning with the clap marker for synchronization.

Eye-Tracking Analysis

The analysis is conducted using predefined AOIs, with ”Lecturer” and ”Slide” for the entire lecture duration.

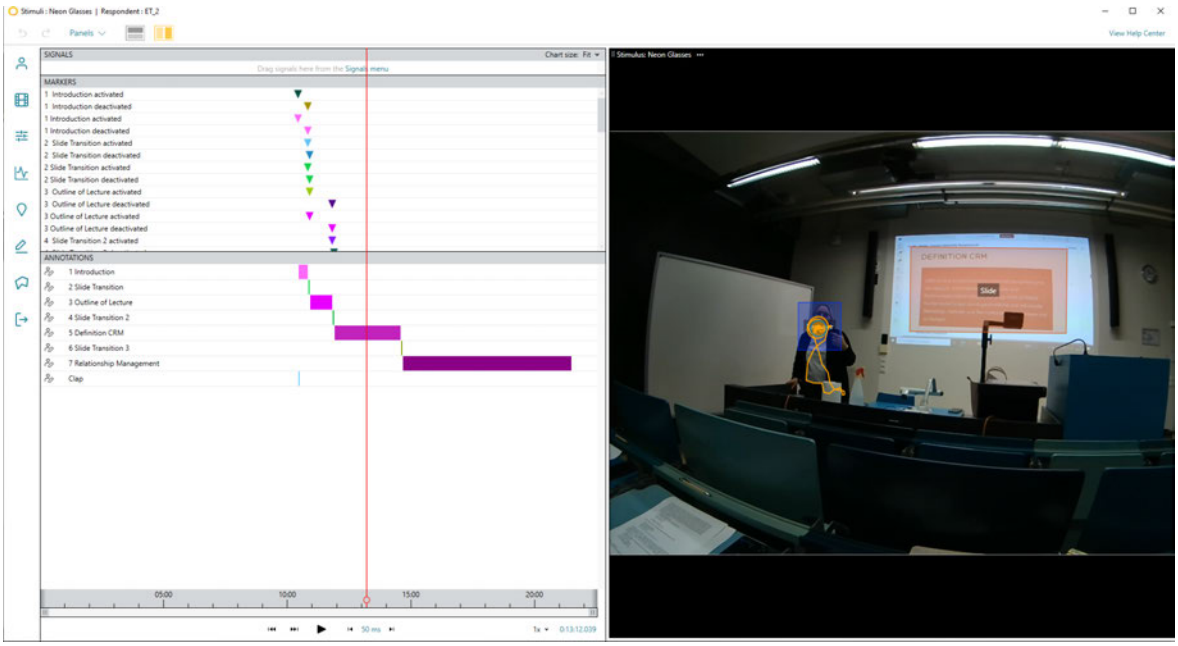


Figure 10: ET POV with AOIs Definition Slide

The figure above illustrates the POV of an in-person ET student, along with the defined AOIs used in this study. The **orange circle** represents the **fixation point**, while the **orange line** visualizes the **gaze trajectory**.

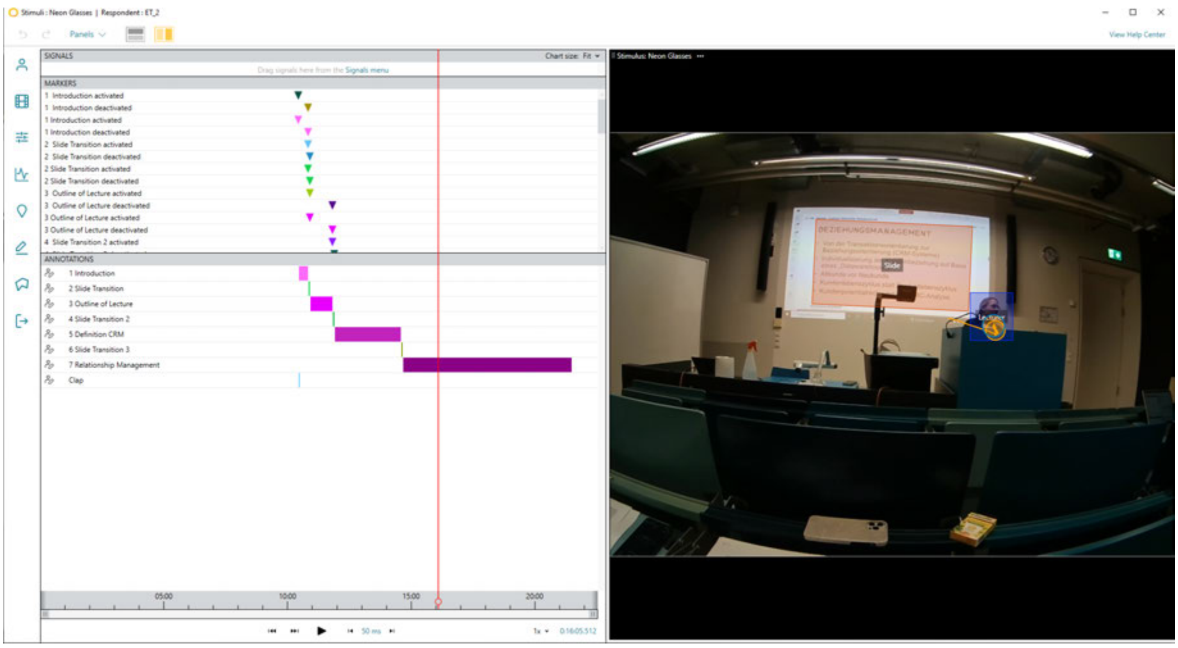


Figure 11: ET POV with AOIs Relationship Slide

The figure above illustrates the in-person ET student's POV, including the defined AOIs used in this study. Unlike the previous figure, this image showcases how the professor utilizes the space and involves students by posing questions.

Finally, the ET data and heatmaps described earlier are exported for further analysis.

By following this structured approach, high quality ET data is systematically collected, processed, analyzed, and visualized to address the key objectives of this study.

Study 3: In-Person FEA

Objective:

Study 3 collects FE data using **Logitech C920 HD Pro Webcams** mounted on JOBY JB01549-BWW TelePods 325 in the classroom. Each webcam is connected to a dedicated PC to ensure that if any PC or camera encounters difficulties, data remains available. At the specified distance, each recording captures three faces. In Study 3, **eight participants** were **divided into three groups**, with Group 1 and Group 2 consisting of three participants and Group 3 of two participants.

Materials:

- Three Logitech C920 HD Pro Webcams – Full HD 1080p

- Three JOBY JB01549-BWW Telepods 325
- Survey for additional qualitative data

Procedure Overview:

All steps adhere to the **common four-step workflow** described in the overall methodology.

1. Preparation

Before conducting the study, all necessary equipment is prepared. Webcams are cleaned to ensure optimal recording quality, and sufficient storage space is available on the PCs used for recording.

Surveys and informed consent forms are printed and labeled accordingly: FEA_1-3, FEA_2_1-3, FEA_3_1-2. These are placed in the classroom according to the seating plan outlined earlier in the setup sketch.

Webcams are mounted on TelePods at eye level and connected to the designated PCs. The setup is thoroughly checked to confirm that all faces are visible in the video, that the cameras are at the correct eye level and that the lighting conditions in the classroom are adequate.

2. Data Collection

At the start of the session, participants receive verbal instructions on the study procedure, including details on the type of data collected and the required head positioning. Each participant then signs the informed consent form.

Recording is started on all PCs, followed by a synchronization clap performed by the instructor to provide a reference marker for later data alignment. The lecture is conducted while the facial expression data is continuously recorded. At the end of the session, recordings on all PCs are stopped.

Participants then complete the survey, after which the webcams and PCs are packed up. All informed consent forms and survey responses are collected and archived for further processing.

3. Integration, Preprocessing, and Synchronization

Following data collection, the recorded videos are exported to the PC where iMotions is installed. A main folder (e.g., FEA) is created to organize the data.

Each video is duplicated based on the number of participants present in the recording. For example, if a video contains three participants, it is cloned three times. Each participant's data is then stored in a dedicated subfolder labeled with a **pseudonymized ID** (e.g., FEA_1_1, FEA_1_2, FEA_1_3, FEA_2_1, FEA_2_2 etc.) A new study is created in iMotions Desktop using the study type "Import Face Recordings." The FEA folder is selected for import. A data quality check is conducted to identify any issues, and unsuitable respondents are excluded. A static face box is defined for each participant, and videos are replayed to ensure that the participant's head remains within the designated face box for the entire study duration. Biometric details such as age and gender are entered for each participant, followed by post-processing using Affectiva.

4. Analysis

Event markers are then applied for each respondent, starting with the clap marker for synchronization purposes.

Facial Expression Analysis

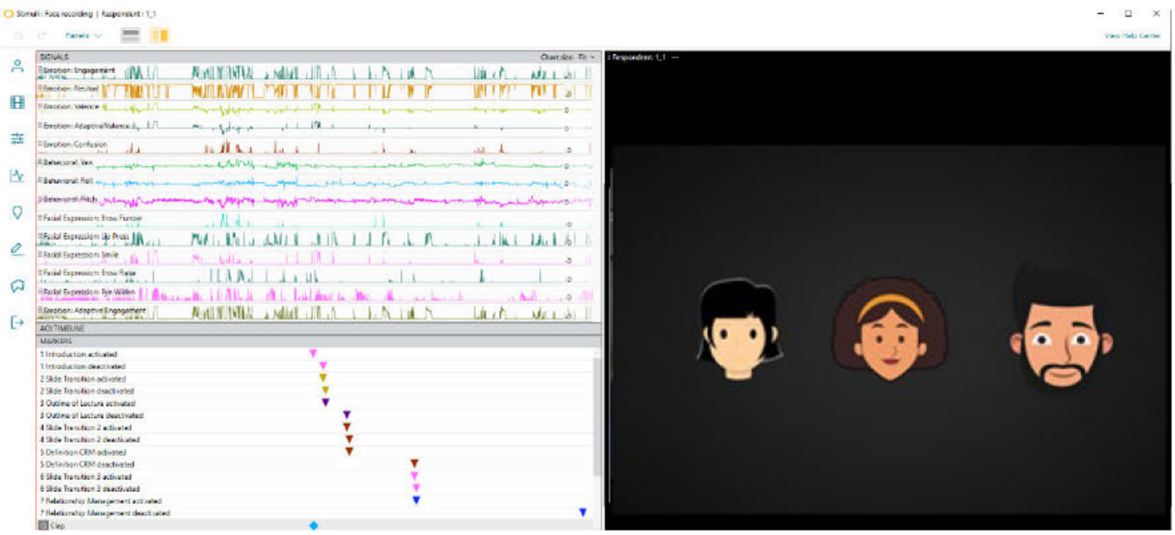


Figure 12: FEA in iMotions Desktop

The image above is a screenshot from iMotions Desktop. On the left side, the top section displays the signals of the selected emotions, facial expressions, and behavioral data. Below that, the event markers are visible. On the right side, the recording of three students is shown. To ensure data privacy, the original footage has been blacked out and replaced with sketched avatar heads for illustration purposes.

Finally, the FE metrics described earlier are exported for further analysis. By following this structured approach, high-quality FE data is systematically collected, processed and analyzed to ensure reliable insights into emotional engagement.

Field Tests II – Main Study

The **Main Study** represents the final implementation of the **refined methodologies** for each study, incorporating the **refinements and adjustments** made based on insights from the **Pilot Study**. This chapter documents the day of the main data collection and provides an overview of its execution.

Study Day: Setting Up the Experiment

The day began with a rapid setup at the designated location. Within just 30 minutes, the entire setup was assembled according to the pre-planned setup (*see setup sketch).* Equipment was positioned, software was initialized, and final checks were conducted to ensure a smooth data collection process.

Objectives

The **Main Study** comprised **three parallel and synchronized data collection processes**:

1. **RDC** via **iMotions Cloud**,
2. **In-Person ET** using **Neon Invisible Glasses**,
3. **In-Person FEA** using **stationary webcams**.

All data collection procedures were conducted simultaneously during a 10-minute hybrid lecture on CRM. This synchronized approach enabled an integrated analysis within iMotions, ensuring seamless data correlation across the different modalities.

The regularly scheduled hybrid lecture took place in Lecture Hall 05.03.001 in early 2025 and was simultaneously streamed via Microsoft Teams, allowing participation from both remote and in-person students. In total, 30 students attended the lecture and data from 20 students were collected.

Category	Study 1: RDC	Study 2: In-Person ET	Study 3: In-Person FEA
Participants	9	3	8
Excluded	3	0	1
Final	6	3	7

Figure 13: Recording Overview

Exclusions shown in the figure above were based on quality criteria, ensuring that only high-quality data was retained for analysis.

Conclusion of the Methodology

The methodology presented in this chapter proved to be successfully implementable, ensuring a structured, robust and reliable approach to data collection. Standardized data processing workflows enabled comparability across datasets, ensuring methodological validity. As a result, this study has established a foundation for future research at the UASD, demonstrating that even complex studies can be effectively conducted in real-world classroom environments.

Results

The results presented in this section were generated from the previously described Field Test II – Main Study.

Note: Described values in this section focus on measured occurrences without interpretation of *why* these patterns arose. Further discussion and analysis of potential causes or implications appear in the subsequent Discussion section.

Aggregated AOI Metrics

	Study 1: RDC (6 recordings)							Study 2: In-Person ET (3 recordings)						
AOI's/ segment	Gaze based metrics			Fixation based metrics				Gaze based metrics			Fixation based metrics			
	Respondent count	Dwell time (%)	Revisit count	Fixation count	Fixation duration (ms)	TTFF AOI (ms)	Dwell time (%)	Respondent count	Dwell time (%)	Revisit count	Fixation count	Fixation duration (ms)	TTFF AOI (ms)	Dwell time (%)
Introduction	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min	0.39 min
Lecturer	6	13.59	4.83	5	997.62	3587.08	21.7	3	47.47	8	32.67	146.54	2971.83	19.89
Slide	6	21.64	5.83	12.5	342.55	5349.83	19.34	3	29.43	6	23.67	127.2	4319.5	12.72
Slide Transition	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min	0.09 min
Lecturer	4	6.5	0.75	1.5	570.46	1368.88	13.7	1	24.59	1	3	187	294.5	10.09
Slide	5	58.11	1.4	7.6	423.5	699.5	50.56	3	76.43	1	16	128.93	875.5	36.25
Outline of Lecture	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min	0.89 min
Lecturer	5	11.47	7.2	6.6	2236.64	1540.7	16.93	3	31.45	8	67.33	145.39	4170.17	19.5
Slide	6	55.28	13	47.17	558.34	362.17	48.87	3	59.86	16	129	150.36	323.5	36.17
Slide Transition 2	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min
Lecturer	0	NA	NA	NA	NA	NA	NA	3	11.13	0.67	2	159	2371	5.54
Slide	6	67.92	1.67	7.17	619.09	656.25	57.98	3	82.05	1.33	16.67	136.24	389.83	39.17
Definition CRM	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min	2.67 min
Lecturer	6	6.76	18.33	14.5	1541.05	35555.5	12.86	3	57.22	44.33	324.33	174.15	11242.83	35.1
Slide	6	45.63	30.33	124.83	489.77	695.42	39.79	3	38.03	42.67	231.67	132.21	114.5	19.18
Slide Transition 3	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min	0.10 min
Lecturer	0	NA	NA	NA	NA	NA	NA	3	14.33	1	1.5	185.12	620.75	7.47
Slide	6	63.97	1	6.5	707.52	588.75	49.9	3	75.75	0.67	10	128.43	1316.83	31.19
Relationship Management	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min	6.84 min
Lecturer	6	4.57	34.33	37	522.65	97939.92	6.5	3	60.37	143.67	862.67	195.49	3465.83	40.6
Slide	6	50.46	88.83	320.67	592.51	646	45.96	3	22.25	73	329.67	125.66	276.17	10.15

Figure 14: Aggregated AOI Metrics

The gaze-based and fixation-based metrics in the figure above reveal clear differences between the **RDC group** on the **left** and the **In-Person ET group** on the **right**.

1.1 Gaze-Based Metrics

1.1.1 Respondent Count

Indicates the number of participants who viewed a specific AOI (Lecturer or Slides).

- **RDC Group**
  - 100% of participants fixated on both AOIs (Lecturer and Slides) in segments where the content was explicitly explained by the lecturer (e.g., “Introduction,” “Definition CRM,” “Relationship Management”).
- **In-Person ET Group**
  - 100% of participants fixated on both AOIs in most segments.
  - One exception was the “Slide Transition” segment, where only 1 out of 3 participants fixated on the Lecturer.

1.1.2 Dwell Time (%)

Percentage of total gaze time spent on an AOI (fixations + saccades).

- **RDC Group**
  - Dwell Time was generally higher on Slides (e.g., “Slide Transition 2”: 57.98%) compared to In-Person ET (39.17%).
- **In-Person ET Group**
  - More balanced distribution between Lecturer and Slides in some segments (e.g., “Definition CRM”: 35.1% Lecturer vs. 19.18% Slides; “Relationship Management”: 40.6% vs. 10.15%).

1.1.3 Revisit Count

Number of times participants returned to an AOI after looking elsewhere.

- **RDC Group**
  - Lecturer: Consistently lower revisit counts (e.g., 34.33 in “Relationship Management” vs. 143.67 for In-Person ET).
  - Slides: Sometimes higher than In-Person ET (e.g., 88.83 in “Relationship Management” vs. 73).
- **In-Person ET Group**
  - Lecturer: Higher revisit rates across segments.
  - Slides: Revisit counts were slightly lower or comparable.

1.2 Fixation-Based Metrics

1.2.1 Fixation Count

Number of individual fixations per AOI.

- **RDC Group**
  - **Lecturer**: Fewer fixations overall (e.g., “Definition CRM”: 14.5 vs. 324.33 in In-Person ET).
  - **Slides**: Also lower, though closer in some segments (e.g., 124.83 in “Definition CRM” vs. 231.67 in In-Person ET).
- **In-Person ET Group**
  - **Lecturer**: Substantially higher fixation counts in content-rich segments (e.g., “Definition CRM”: 324.33).
  - **Slides**: Consistently higher (e.g., “Outline of Lecture”: 129 vs. 47.17 in RDC).

1.2.2 Fixation Duration (ms)

Average duration of each fixation on an AOI.

- **RDC Group**
  - **Lecturer**: Significantly longer fixations in some segments (e.g., “Definition CRM”: 1541.05 ms vs. 174.15 ms in In-Person ET), but shorter in others (e.g., “Relationship Management”: 522.65 ms vs. 862.67 ms).
  - **Slides**: Generally, longer than in-person (e.g., “Outline of Lecture”: 558.34 ms vs. 150.36 ms).
- **In-Person ET Group**
  - **Lecturer**: Shorter average fixation durations (e.g., 174.15 ms in “Definition CRM”), except in “Relationship Management,” where durations were longer.
  - **Slides**: Consistently shorter than in RDC.

1.2.3 Time to First Fixation (TTFF)

Time it took to first fixate on an AOI.

- **RDC Group**
  - **Lecturer**: Considerably longer TTFF (e.g., “Definition CRM”: 35,555.5 ms; “Relationship Management”: 97,939.92 ms).
  - **Slides**: Also longer, but differences were less extreme.
- **In-Person ET Group**

- **Lecturer**: Much shorter TTFF values (e.g., 11,242.83 ms in “Definition CRM”).
- **Slides**: Lower TTFF overall compared to RDC.

1.2.4 Dwell Time (%) – Fixation-Based

Percentage of total fixation time spent on each AOI.

- **RDC Group**
  - **Lecturer**: Substantially lower fixation share (e.g., “Definition CRM”: 12.86%; “Relationship Management”: 6.5%).
  - **Slides**: Higher than in In-Person ET (e.g., “Slide Transition”: 50.56%).
- **In-Person ET Group**
  - **Lecturer**: Higher fixation time shares (e.g., “Definition CRM”: 35.1%; “Relationship Management”: 40.6%).
  - **Slides**: Lower compared to RDC.

Visualization of Heatmaps

In this section, the **fixation-based** heatmaps for all seven segments are presented and linked to previously described fixation-based metrics. On the **left**, each figure displays the heatmap for the **RDC group**, and on the **right**, the heatmap for the **In-Person ET group**.

Notes on Heatmap Visualization

- **Color Gradients**: Warmer shades **(red/orange/yellow)** correspond to **higher density** of fixations, whereas cooler shades **(green)** indicate moderate to low fixation density.
- **Variability Among Participants**: Each heatmap aggregates data from all participants in that group, providing a holistic picture of where attention clusteres within each segment.

Segment 1: Introduction

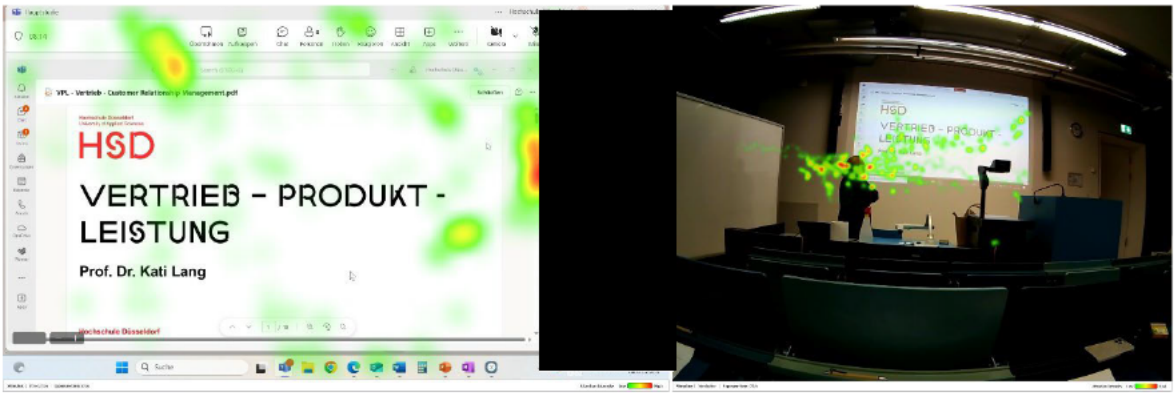


Figure 15: Heatmap Comparison (Segment 1: Introduction)

In the **Introduction** segment, the **RDC heatmap** shows a **strong focus on the lecturer (red)**, aligning with aggregated fixation-based AOI metrics indicating **higher average dwell time** and **longer fixation duration** among remote participants. The **In-Person ET heatmap** also centers attention on the lecturer but exhibits more frequent fixations overall, consistent with fixation-based metrics reporting **higher fixation counts** in the classroom environment.

Segment 2: Slide Transition

The observations of these segments provide a snapshot of how attention momentarily shifts during slide changes, although the segment's short duration limits further detailed analysis.

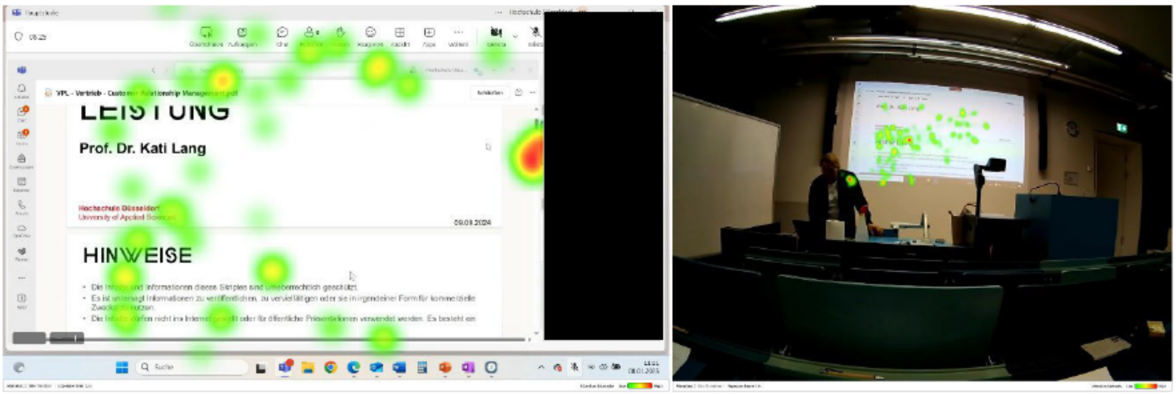


Figure 16: Heatmap Comparison (Segment 2: Slide Transition)

The **RDC heatmap** visualizes relatively balanced fixations between the lecturer and slides, with a slight emphasis on the lecturer.

In contrast, the **In-Person ET heatmap** displays a red area on the slides. And **more overall fixations** distributed across slide elements, as visualized by the more pronounced green areas. These observations mirror the measured values for the

ET group, which recorded **higher fixation counts** and **greater dwell time** on slides compared to the lecturer.

Segment 3: Outline of Lecture

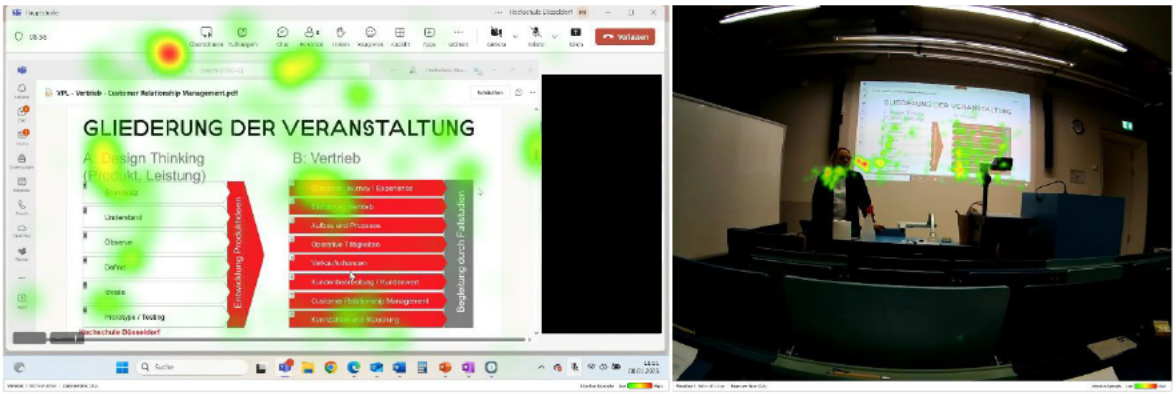


Figure 17: Heatmap Comparison (Segment 3: Outline of Lecture)

During the **Outline of Lecture**, the lecturer verbally explained the session's plan. The **RDC heatmap** displays more intense fixations on the slide content, consistent with the measured fixation counts (e.g., **6.6 on the lecturer** vs. **47.17 on the slide**). In contrast, the **In-Person ET heatmap** shows a red cluster around the lecturer along with numerous green points spread across the slides, showing a **higher overall fixation volume.** The measured counts likewise reflect **67 fixations** on the lecturer and **129** on the slides.

Segment 4: Slide Transition 2

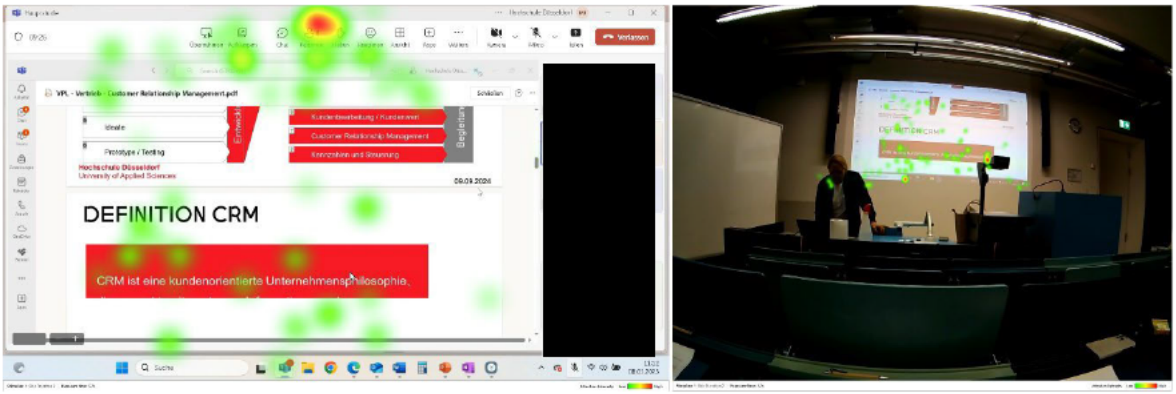


Figure 18: Heatmap Comparison (Segment 4: Slide Transition 2)

In the **RDC heatmap**, fixations are relatively scattered (as shown by the green zones), yet a prominent **red cluster** also appears **outside the primary lecture materials** (i.e., neither on the slide nor the lecturer).

By contrast, the **In-Person ET** visualization shows widespread fixations across both the lecturer and slides, with a small red concentration on key slide elements.

This limited time window offers only a snapshot of how attention shifts during slide changes, aligning with the numerical fixation data but making in-depth analysis challenging.

Segment 5: Definition CRM

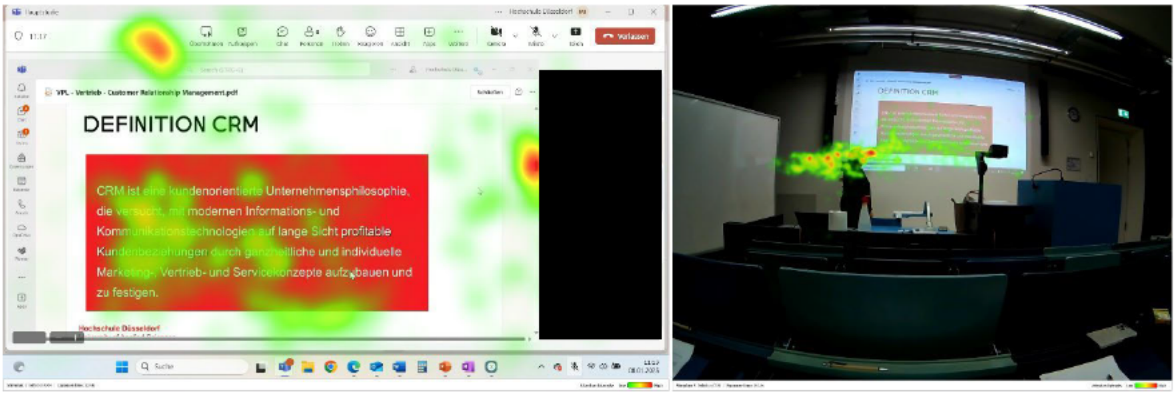


Figure 19: Heatmap Comparison (Segment 5: Definition CRM)

Definition CRM is a **content-heavy** segment pivotal to the lecture's objectives. Given its significance, the analysis below provides a more detailed account of how each group directed their attention.

1. **RDC Group**

   The **RDC heatmap** shows a modest cluster of fixations on the lecturer, with most of the attention focused on the slide content. This pattern aligns with the fixation-based data, which record **14.5 fixations** on the lecturer—a comparatively small figure alongside the **124.83 fixations** on the slides. However, the average fixation duration on the lecturer was **1541.05 ms**, indicating fewer but more sustained glances. Meanwhile, the **TTFF** for the lecturer reached **35,555.5 ms**, revealing a delayed onset of lecturer-focused attention in the remote environment.

2. **In-Person ET Group**

   By contrast, the **In-Person ET heatmap** depicts a dense concentration of fixations around the lecturer as well as a substantial spread across the slide. The metrics confirm **324.33 fixations** on the lecturer and **231.67** on the slides—both significantly higher than in the RDC group. Although the average fixation duration on the lecturer was **174.15 ms**, implying shorter but more frequent glances, the **TTFF** on the lecturer stood at **11,242.83 ms**, indicating that in-person participants oriented to the lecturer much sooner than their remote counterparts.

Segment 6: Slide Transition 3

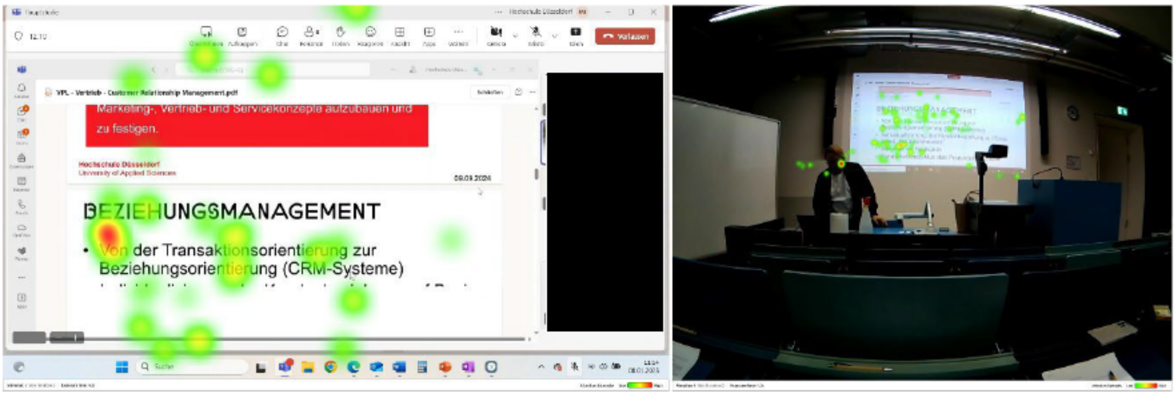


Figure 20: Heatmap Comparison (Segment 6: Slide Transition 3)

RDC Group

The fixation-based metrics show **no recorded fixations** on the lecturer **(respondent count = 0)**. Remote participants focused entirely on the new slide during this brief transition. In contrast, the slides received multiple fixations **(fixation count = 6.5)** at an average duration of **707.52 ms**, which underscores a relatively sustained attention on emerging content.

In-Person ET Group

The in-person heatmap and metrics confirm that some attention was still directed toward the lecturer (e.g., **fixation count = 1.5**), although most fixations concentrated on the slides (fixation count = 10). Despite having more fixations overall, the average **fixation duration** on the slides (**128.43 ms**) was **shorter** compared to the RDC group, suggesting more frequent yet briefer glances. This pattern aligns with in-person participants' tendency to quickly shift between the lecturer and the transitioning slide.

Segment 7: Relationship Management

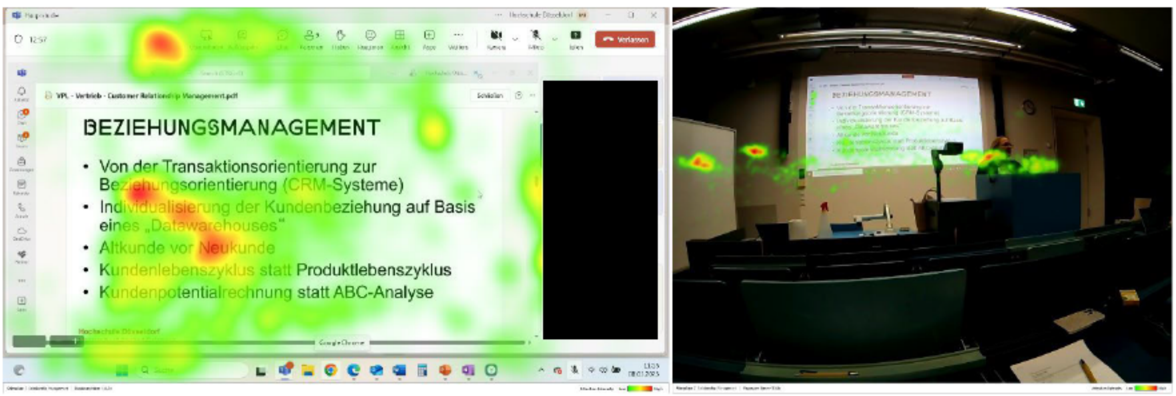


Figure 21: Heatmap Comparison (Segment 7: Relationship Management)

Segment 7, served as a thorough recap of the session's essential topics. It was the longest segment and saw the lecturer make full use of the space, moving around to maintain engagement. Additionally, the lecturer posed questions throughout, testing the participants' understanding and reinforcing the key takeaways.

RDC Group

The **RDC heatmap** suggests a limited focus on the lecturer, reflected in the fixation-based metrics: **37 fixations** on the lecturer and a notably long **TTFF** of **97,939.92 ms**. Meanwhile, participants recorded **320.67 fixations** on the slides. Dwell time percentages also indicate a smaller portion of attention allocated to the lecturer compared to the slides.

In-Person ET Group

In contrast, the **In-Person ET heatmap** reveals a substantial cluster around the lecturer, supported by **862.67 fixations**—significantly more than RDC—and a considerably shorter **TTFF** of **3,465.83 ms**. Although the slides also received a high number of fixations (329.67), the lecturer commanded a higher dwell time percentage overall. This distribution illustrates that in-person participants devoted more frequent and earlier attention to the speaker, even while examining the slide content.

2. Facial Expression Analysis

The figures below (Figure 22: Aggregated FEA for RDC and Figure 23: Aggregated FEA for In-Person FEA) present the exported **emotional metrics** and **facial expression metrics** for both groups across the segments. Despite notable differences between the groups, some common patterns emerge: **Confusion** was the least observed emotional state, while **Neutral** was the most prevalent. **Engagement** consistently ranked between these two.

	Segment	Introduction	Slide Transition	Outline of Lecture	Slide Transition 2	Definition CRM	Slide Transition 3	Relationship Management
	Duration	0.39 min	0.09 min	0.89 min	0.10 min	2.67 min	0.10 min	6.84 min
Emotional metrics (%)	Engagement	27.88	22.19	19.19	20.31	19.49	10.81	18.41
	Confusion	1.02	1.77	1.79	0.41	0.97	0.08	0.97
	Neutral	92.94	99.22	91.41	97.25	92.81	98.94	91.99
Facial expression metrics (%)	Brow furrow	26.01	35.58	24.81	22.39	17.34	17.01	27.22
	Brow raise	0.65	0.04	0.33	0.09	0.69	0	0.57
	Eye widen	7.47	6.27	4.66	2.93	3.69	9.68	4.14
	Lip press	2.89	2.25	1.99	5.64	2.51	2.13	3.23
	Smile	0.48	0	0.31	0	0.9	0	1.51

Figure 22: FEA for RDC Group

	Segment	Introduction	Slide Transition	Outline of Lecture	Slide Transition 2	Definition CRM	Slide Transition 3	Relationship Management
	Duration	0.39 min	0.09 min	0.89 min	0.10 min	2.67 min	0.10 min	6.84 min
Emotional metrics (%)	Engagement	27.88	22.19	19.19	20.31	19.49	10.81	18.41
	Confusion	1.02	1.77	1.79	0.41	0.97	0.08	0.97
	Neutral	92.94	99.22	91.41	97.25	92.81	98.94	91.99
Facial expression metrics (%)	Brow furrow	26.01	35.58	24.81	22.39	17.34	17.01	27.22
	Brow raise	0.65	0.04	0.33	0.09	0.69	0	0.57
	Eye widen	7.47	6.27	4.66	2.93	3.69	9.68	4.14
	Lip press	2.89	2.25	1.99	5.64	2.51	2.13	3.23
	Smile	0.48	0	0.31	0	0.9	0	1.51

Figure 23: FEA for In-Person Group

2.1 Emotional Metrics

2.1.1 Engagement

- **RDC**
  - **Min**: 10.81% in **Slide Transition 3**
  - **Max**: 27.88% in **Introduction**
  - Other segments ranged between 18.41% and 22.19%.
  - **Difference**: 27.88% − 10.81% = **17.07%**
- **In-Person FEA**
  - **Min**: 9.75% in **Slide Transition 3**
  - **Max**: 37.95% in **Slide Transition**
  - Other segments ranged between 13.54% and 29.18%.
  - **Difference**: 37.95% − 9.75% = **28.20%**

2.1.2 Confusion

- **RDC**
  - **Min**: 0.08% in **Slide Transition 3**
  - **Max**: 1.79% in **Outline of Lecture**
  - Other segments ranged from 0.41% to 1.77%.
  - **Difference**: 1.79% − 0.08% = **1.71%**
- **In-Person FEA**
  - **Min**: 0% in **Slide Transition 3**
  - **Max**: 6.81% in **Slide Transition**
  - Remaining segments typically stayed at or below 1.41%.
  - **Difference**: 6.81% − 0% = **6.81%**

2.1.3 Neutral

- **RDC**
  - **Min**: 2.93% in **Slide Transition 2**
  - **Max**: 9.68% in **Slide Transition 3**

- Remaining values ranged between 3.69% and 7.47%.
- **Difference**: 9.68% − 2.93% = **6.75%**

- **In-Person FEA**
  - **Min**: 75.51% in **Slide Transition**
  - **Max**: 90.16% in **Relationship Management**
  - Most other segments varied between 75.94% and 89.00%.
  - **Difference**: 90.16% − 75.51% = **14.65%**

2.2 FE Metrics

2.2.1 Brow Furrow

- **RDC**
  - **Min**: 17.01% in **Slide Transition 3**
  - **Max**: 35.58% in **Slide Transition** (the first one)
  - Other segments ranged from 17.34% to 27.22%.
  - **Difference**: 35.58% − 17.01% = **18.57%**
- **In-Person FEA**
  - **Min**: 0% in **Slide Transition 3**
  - **Max**: 6.41% in **Slide Transition** (the first one)
  - Remaining segments ranged between 0.30% and 1.71%.
  - **Difference**: 6.41% − 0% = **6.41%**

2.2.2 Brow Raise

- **RDC**
  - **Min**: 0% in **Slide Transition 3**
  - **Max**: 0.69% in **Definition CRM**
  - Other measurements fell between 0.04% and 0.65%.
  - **Difference**: 0.69% − 0% = **0.69%**
- **In-Person FEA**
  - **Min**: 0% in **Slide Transition 2**
  - **Max**: 9.89% in **Outline of Lecture**
  - Other segments were around 0.67% to 7.93%.
  - **Difference**: 9.89% − 0% = **9.89%**

2.2.3 Eye Widen

- **RDC**
  - **Min**: 2.93% in **Slide Transition 2**
  - **Max**: 9.68% in **Slide Transition 3**

    - Remaining values ranged between 3.69% and 7.47%.
    - **Difference**: 9.68% − 2.93% = **6.75%**
- **In-Person FEA**
    - **Min**: 0% in **Slide Transition** (the first one)
    - **Max**: 10.71% in **Introduction**
    - Other segments were between 2.13% and 8.27%.
    - **Difference**: 10.71% − 0% = **10.71%**

2.2.4 Lip Press

- **RDC**
    - **Min**: 1.99% in **Outline of Lecture**
    - **Min**: 5.64% in **Slide Transition 2**
    - Other segments ranged from 2.13% to 3.23%.
    - **Difference**: 5.64% − 1.99% = **3.65%**
- **In-Person FEA**
    - **Min**: 1.34% in **Slide Transition 3**
    - **Max**: 7.89% in **Outline of Lecture**
    - The rest of the segments reported values from 2.81% to 7.73%.
    - **Difference**: 7.89% − 1.34% = **6.55%**

2.2.5 Smile

- **RDC**
    - **Min**: 0% (Slide Transition, Slide Transition 2, Slide Transition 3)
    - **Max**: 1.51% in **Relationship Management**
    - Introduction 0.48%, Outline of Lecture 0.31%, Definition CRM 0.9%.
    - **Difference**: 1.51% − 0% = **1.51%**
- **In-Person FEA**
    - **Min**: 0% (Slide Transition, Slide Transition 3)
    - **Max**: 19.69% in **Introduction**
    - Other segments ranged from 0.58% to 2.81%.
    - **Difference**: 19.69% − 0% = **19.69%**

Summary of emotional Engagement

The following figure visualizes the emotional engagement identified in both studies:

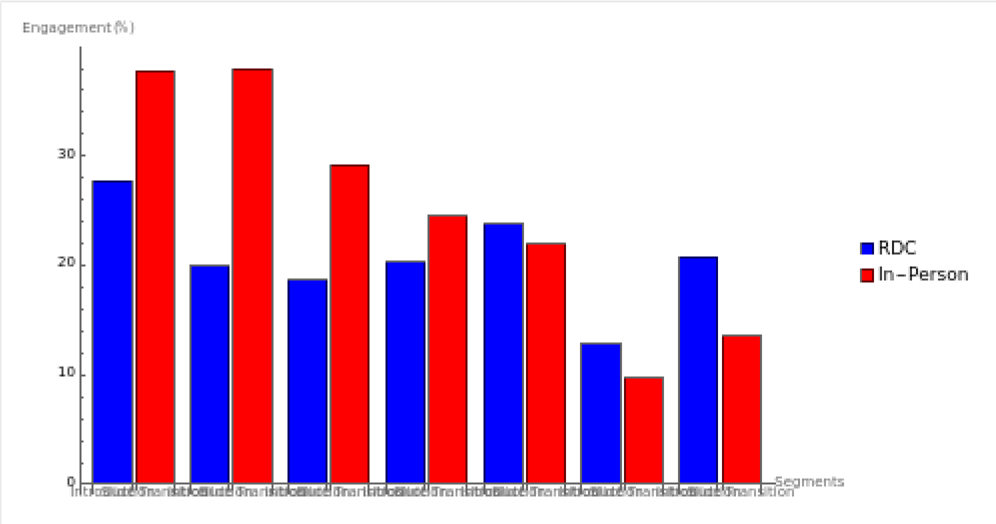


Figure 24: Emotional Engagement (%) for each Segment

The figure above shows emotional engagement (%) for each Segment. **Left** begins with **segment 1** Introduction and **ends** with **segment 7** Relationship Management.

3. Survey Results

The tables below present the **survey responses** collected from participants in each study group. Participants rated their experiences based on key **engagement-related factors**, such as **attention, motivation, and external distractions**. Responses were measured on a **Likert scale ranging from 1 to 5**, where:

Legend for the Scale (1-5):

5 = Fully agree (e.g., "I was highly motivated"),

4 = Rather agree

3 = Partially agree

2 = Rather disagree

1 = Fully disagree (e.g., "I was not motivated at all").

Respondent	Gender	Age	Technical Issues	Study Process Clear	Consistent Attention	Following Topics	Content Understandable	Focused on Content	Motivated	Content Interesting	Distracted Thoughts	Environment Influence
15	Female	20	No	Yes	4	5	5	3	5	5	3	2
17	Male	27	No	Yes	2	5	5	5	3	5	3	3
14	Male	20	No	Yes	3	2	4	3	4	4	4	5
12	Male	22	No	Yes	4	4	(No answer)	4	4	4	2	4
6	Male	29	No	Yes	4	5	5	4	5	5	3	4
3	Female	23	No	Yes	3	4	4	3	3	3	4	4
Average	-	**23.5**	-	-	**3.33**	**4.17**	**4.6**	**3.67**	**4**	**4.33**	**3.17**	**3.67**

Figure 25: RDC Survey Responses

Respondent	Gender	Age	Technical Issues	Study Process Clear	Consistent Attention	Following Topics	Content Understandable	Focused on Content	Motivated	Content Interesting	Distracted Thoughts	Environment Influence
ET_1	Female	30	No	Yes	5	5	5	5	5	5	2	4
ET_2	Male	24	No	Yes	4	5	5	4	5	5	2	2
ET_3	Female	21	No	Yes	5	5	5	5	5	4	2	1
Average	-	**25.00**	-	-	**4.67**	**5.00**	**5.00**	**4.67**	**5.00**	**4.67**	**2**	**2.33**

Figure 26: In-Person ET Survey Responses

Respondent	Gender	Age	Technical Issues	Study Process Clear	Consistent Attention	Following Topics	Content Understandable	Focused on Content	Motivated	Content Interesting	Distracted Thoughts	Environment Influence
FEA_1_1	Female	27	No	Yes	4	4	5	5	5	5	3	3
FEA_1_2	Male	23	No	Yes	4	4	4	4	5	5	5	4
FEA_1_3	Male	24	No	Yes	4	5	4	4	3	3	2	4
FEA_2_1	Male	22	No	Yes	3	4	4	3	4	5	5	3
FEA_2_2	Male	22	No	Yes	5	5	5	3	5	4	3	2
FEA_2_3	Male	22	No	Yes	1	5	5	3	3	4	5	3
FEA_3_1	Male	21	No	Yes	3	5	5	3	3	3	5	3
FEA_3_2	Male	22	No	Yes	4	5	4	3	4	5	1	4
Average	-	**22.88**	-	-	**3.5**	**4.63**	**4.5**	**3.5**	**4**	**4.25**	**3.63**	**3.25**

Figure 27: In-Person FEA Survey Responses

3.1 Technical Issues & Study Process Clarity

- None of the participants in any study reported technical issues.
- The study process was clear to all participants across groups.

3.2 Attention and Understanding of Content

"Consistent Attention" (Maintaining focus throughout the session)

- RDC Group reported: **3.33** (lowest)
- ET Group reported: **4.67** (highest)
- FEA Group reported: **3.5** (slightly higher than RDC)

→ The ET group rated themselves as having the best ability to stay focused.

"Following Topics" (Keeping up with the material)

- RDC Group reported: **4.17**
- ET Group reported: **5.00**
- FEA Group reported: **4.63**

→ Again, the ET group reported the highest ease in following topics, while the FEA group rated themselves slightly better than RDC.

"Content Understandable" (Perceived clarity of the material)

- RDC Group reported: **4.6**
- ET Group reported: **5.00**
- FEA Group reported: **4.5**

→ All groups rated the content as fairly understandable, but ET gave themselves the highest score.

3.3 Motivation and Interest in the Content

"Motivated" (How motivated they felt)

- RDC Group reported: **4.00**
- ET Group reported: **5.00**
- FEA Group reported: **4.00**

→ The ET group self-reported the highest motivation, while RDC and FEA had identical self-assessments.

"Content Interesting" (How engaging the content was perceived to be)

- RDC Group reported: **4.33**
- ET Group reported: **4.67**
- FEA Group reported: **4.25**

→ The ET group rated their experience as the most interested in the content, while the other two groups had similar ratings.

3.4 Distractions and Environmental Influence

"Distracted Thoughts" (How often they felt distracted)

- RDC Group reported: **3.17**
- ET Group reported: **2.00** (lowest distraction)
- FEA Group reported: **3.63** (highest distraction)

→ The ET group reported being the least distracted, while FEA self-reported the highest level of distraction.

"Environment Influence" (How much their surroundings affected their learning)

- RDC Group reported: **3.67**
- ET Group reported: **2.33**
- FEA Group reported: **3.25**

→ The ET group felt the least affected by their environment, while RDC reported the highest impact.

Overall Experience (Self-Reported)

1. **The ET group consistently rated themselves the highest** in motivation, attention, clarity, and focus.
2. **The FEA group positioned themselves between ET and RDC**, but they also reported the **highest distraction levels**.
3. **The RDC group reported the most challenges**, particularly in terms of attention.

Discussion

The following section summarizes the analyses of the ET, FE, and survey results, along with the interpretation and practical implications best recommendations for optimizing hybrid lecture design.

Summary and Interpretation

1. Eye-Tracking and Heatmaps

1.1 Summary

RDC participants exhibited fewer, yet longer fixations, primarily directed at the slides. They also had a delayed TTFF for the lecturer. Heatmaps confirm this distinction: for instance, in the introduction segment, remote participants spent considerable amount of dwell time focused on the lecturer, whereas in-person heatmaps revealed a dynamic, alternating gaze between lecturer and slides.

In-Person participants showed frequent, shorter fixations on both the lecturer and the slides, as well as lower TTFF. These patterns indicate swift orientation to AOIs.

1.2 Interpretation

Physical presence appears to provide immediate nonverbal and social cues, resulting in a more balanced visual distribution between slides and lecturer. The lower TTFF and higher fixation counts seen in in-person data underscore the advantages of face-to-face settings for continuous visual engagement. Conversely, remote learners often dedicate longer fixations to slides, which implies reduced real-time interaction with the lecturer's visual signals.

2. Facial Expression Analysis

2.1 Summary

RDC participants tended to stay in a neutral expression more frequently, with fewer observable emotional changes. Although confusion levels remained low across both settings, in-person participants recorded more engagement peaks (e.g., brief moments of increased facial movement).

In-Person participants displayed a wider range of facial expressions, including lower levels of neutrality and periodic peaks in positive affect.

2.2 Interpretation

A broader emotional variability in the in-person group suggests stronger social and cognitive engagement, likely driven by direct face-to-face interaction and peer influence. The elevated neutrality among remote participants may reflect limited social feedback and fewer prompts to exhibit emotion during online participation.

3. Survey Analysis

3.1 Summary

(a) Attention & Understanding

RDC participants reported lower attention scores (3.33) and moderate ease in following topics (4.17), aligning with the ET data that showed slower fixation on the lecturer and a stronger focus on slides.

In-Person ET participants reported the highest attention (4.67) and had no difficulty following the lecture topics (5.00).

(b) Motivation & Interest

RDC and **FEA** participants ranked their motivation slightly lower (4.00). These self-reports parallel the FEA patterns, where in-person participants exhibited more dynamic emotional responses.

In-Person ET participants indicated the greatest motivation (5.00), finding the content highly interesting.

(c) Distractions & Environmental Influence

RDC participants reported moderate-to-high distraction (3.17) and the strongest environmental impact (3.67).

FEA participants noted the highest distractions overall (3.63), possibly due to being aware of cameras.

In-Person ET consistently indicated least distraction and minimal environmental influence, matching sustained gaze metrics from the Eye-Tracking data.

3.2 Interpretation:

RDC participants experiences were consistent with their gaze data: heavy reliance on slides, lower attention scores, and higher reported external distractions.

In-person participants felt more attentive, motivated, and less distracted, aligning with the supportive nature of a shared physical environment.

4. Combined Interpretation

Taken together, these findings form a comprehensive picture of attention and emotional engagement in hybrid lectures, and it is particularly noteworthy that the self-reported survey results align closely with the quantitative data from ET and FEA.

Remote Challenges: Delayed lecturer fixation, prolonged focus on slides, and higher neutrality, possibly due to limited instructor visibility and environmental distractions (noise, background stimuli).

In-Person Advantages: Rapid orientation to the lecturer, frequent shifts between slides and instructor, and greater emotional range—all supported by direct social presence and nonverbal cues.

Optimizing Hybrid Lectures

The following Practical Implications build on these findings, offering concrete steps to address each challenge and enhance hybrid lecture engagement for remote learners.

Practical Implications

The combined findings from Eye-Tracking, Facial Expression Analysis, and Survey data underscore the importance of a well-prepared lecture design, structured lecture flow, and mindful remote participation protocols to foster high engagement in hybrid learning environments.

Implications from ET

Enhanced Lecturer Visibility

Longer TTFF in remote conditions suggests the need for a larger, more centrally placed lecturer video window. By making the lecturer's presence a primary visual anchor, students can orient themselves more quickly to core content.

Encourage Gaze Shifts

Interactive elements—such as live polls or quick Q&A sessions—guide learners to switch from slides to the instructor, thereby mimicking the more dynamic, in-person gaze pattern.

Implications from FEA

Promoting Emotional Engagement

The higher neutrality among remote students indicates limited emotional expression online. Fostering face-to-face-like behaviors, such as video-on policies or instructor-led discussions, can prompt students to share more visible reactions.

Strengthening Social Presence

Direct interaction with the lecturer and peers, even in a virtual space (e.g., breakout rooms, small-group chats), can replicate the collaborative spirit of a physical classroom. Instructors who greet students by name, respond to nonverbal cues, and encourage spontaneous feedback often see stronger emotional engagement.

Survey Insights

Managing Distractions

Higher distraction and environmental impact for remote participants point towards the need for setting guidelines—like turning off mobile notifications or finding a quiet space—to maintain focus.

Sustaining Motivation & Interest

In-person groups reported the highest motivation, partly due to direct instructor support. Simple gestures like personalized feedback or short energizer activities can help online students remain motivated.

These implications highlight the challenges lecturers face in today's hybrid learning environments, emphasizing their pivotal role in simultaneously managing both in-person and remote participants—balancing visual attention, emotional engagement, and the complexities of coordinating two classrooms in real time. The next section, **Recommended best practices**, offers a guide to effectively address these challenges through practical teaching strategies.

Recommended Best Practices

Below is a concise set of best practices aimed at boosting remote participants' attention and engagement in hybrid lectures. Because of limitations, which will be discussed later, the study session lasted only 10 minutes. The recommendations primarily address immediate, short-term attention patterns. However, they can also serve as a starting framework for longer lectures: instructors may need to reinforce or adapt these measures to maintain engagement over extended teaching periods.

1. Improve Lecturer Visibility and Social Presence

Enhanced Video Framing: Provide a clear, consistently visible video feed of the lecturer. Use larger or side-by-side views rather than relegating the instructor's image to a small corner. This can reduce the **time to first fixation** on the lecturer and encourage remote participants to reference the lecturer more frequently.

Frequent Verbal Cues: RDC participants primarily fixated on both AOIs only when the lecturer explicitly referenced the slides or concepts. Thus, short verbal signposts (e.g., "Now, please look at this slide," or "Focus on me for a moment") can help guide remote attention more effectively.

2. Structure Slide Transitions to Minimize Confusion

Predictable Slide Changes: Since confusion and neutral expressions spiked in RDC during slide transitions, incorporate brief on-screen indicators (e.g., short summary statements) to smooth the handover between topics.

Interactive Summaries: At each transition, pose a quick question or poll. This can re-engage remote participants and reduce confusion spikes by creating an active rather than passive shift in content.

3. Integrate Interactive Elements to Boost Engagement

Real-Time Q&A or Chat Features: Encourage remote learners to submit questions or respond to polls regularly. Survey data showed RDC participants experienced more distractions; scheduled interactive prompts every few minutes can **anchor attention** and keep them cognitively more involved.

Breakout Sessions: For longer lectures, small-group discussions (virtual or in-person) allow participants to **process material actively**, which can raise emotional engagement levels.

4. Encourage Balanced Attention Between Slides and Lecturer

Slide Design: Keep slides concise, with key points visually emphasized, but not overly dense. In-person participants naturally balanced lecturer and slide attention, whereas RDC participants tended to fixate heavily on slides, especially when the layout was text-heavy.

5. Address Environmental Distractions for Remote Learners

Clear Guidelines on Study Environment: The survey showed higher self-reported distractions and environmental impact in RDC participants. Providing best-practice guides (e.g., "find a quiet space," "reduce background apps," or "mute notifications") can limit external influences.

Camera Policy: Encouraging remote participants to keep their cameras on (where feasible) may foster reciprocal social presence, promoting more balanced visual attention rather than uninterrupted slide fixation.

6. Manage Cognitive Load and Confusion

Segmented Content Delivery: Break material into shorter segments or mini lectures, each followed by a quick check-in or summary. The FEA data revealed confusion notably increased in remote conditions during transitions, so bite-sized pacing could reduce cognitive overload.

Immediate Clarification: For complex segments, plan short recap periods. Ask for feedback, encourage questions, or provide short quizzes to ensure remote learners do not remain confused for extended durations.

7. Leverage FEA Insights to vary Emotional Tone

Introduce Moments of Positivity: Smiling and emotional engagement higher for in person participants. Using brief humor, personal anecdotes, or warm greetings can spark positive emotional responses—even in remote conditions.

Monitor Facial Feedback (If Available): In advanced hybrid setups, real-time FEA could help the instructor gauge remote participants' engagement or confusion. Adjusting the lecture pace on-the-fly if neutral/confused expressions surge can keep participants more emotionally invested.

8. Provide Ongoing Feedback Channels

Post-Lecture Surveys/Reflections: The data confirmed that participants willingly offer insights on motivation and distraction levels. Continuing short post-lecture feedback forms can refine teaching strategies iteratively.

Lecturer–Student Check-Ins: Directly asking remote learners about any issues or sources of confusion mid-lecture helps them reorient attention and fosters a sense of lecturer accessibility, mirroring in-person spontaneity.

Because the session was only 10 minutes, these tips primarily address quick orientation and short bursts of engagement. In a longer format, these strategies might require periodic reinforcement or extended interactive elements (e.g., breakout rooms every 30 minutes).

Conclusion

Implementing 360° cameras, using presentation modes in Microsoft Teams, and regular interactive cues help to address the findings from the Analysis. These recommended best practices were derived from the lecturer's successful application of various didactic methods—integrating targeted questioning, purposeful room-space utilization, expressive gestures, and emotional cues (e.g., tone of voice) for the in-person environment. Such strategies proved instrumental in fostering strong attention and engagement among in-person participants.

By focusing on structured balanced lecturer-slide visibility, and environmental distraction reduction, remote participants might more closely emulate the attention and emotional engagement observed in in-person classes.

Limitations

In this section, Limitations concerning the study design and analytical methods will be addressed.

Practical Constraints and Resource Demands

The implementation of in-person ET required specialized infrastructure and hardware (e.g., eye-tracking glasses), as well as considerable time to accurately manually define AOIs within the iMotions software. Training staff to use these systems and scheduling participants proved to be time-intensive and resource-heavy.

Different Measurement Methods in ET

The study employed two different ET modalities: webcam-based eye tracking (RDC) and mobile ET glasses (In-Person). Discrepancies in hardware precision, calibration procedures, and environmental conditions may have introduced measurement bias and reduced the direct comparability of results.

Single, Short-Duration Data Collection

Data was gathered only once per participant over a relatively brief session. This limited time frame may not fully capture natural fluctuations in attention or emotional states, thus restricting the study's ability to observe long-term attention and engagement patterns. Furthermore, most participants cannot maintain a relatively still head position for much longer than ten minutes, making extended data-collection sessions less practical. For a full hybrid lecture, intermediate recalibrations and regular head-position checks would likely be required to preserve data reliability over a longer duration.

Small Sample Size

Resource limitations and the ethical requirement of voluntary participation led to a modest participant pool. Such a restricted sample makes it challenging to generalize findings across broader populations or diverse educational contexts.

Potential Distortion of Authentic Behaviors

RDC Calibration and Posture Constraints: Remote participants had to follow a calibration process and remain still during the session, which does not accurately reflect typical online learning behaviors.

Distraction by FEA Cameras in In-Person Settings: Participants in the FEA group were visibly aware of cameras placed directly in front of them, possibly altering their emotional or attentional responses. This aligns with self-reported questionnaire data, wherein these participants reported higher "Distracted Thoughts" compared to other groups.

Conclusion of Limitations

Taken together, these limitations highlight that the study's logistical complexities, diverging measurement methods, and limited sample might have affected both the scope and ecological validity of the findings. Although the results offer valuable initial insights, future research should account for these limitations by adopting standardized measurement techniques, considering longer or repeated data collection to better reflect genuine classroom dynamics. Potential solutions to these issues are discussed in the **Outlook** and **Recommendations for Future Research** sections.

Research Questions and Hypotheses

This section addresses the research questions (RQ1–RQ3) and the hypotheses (H1–H2).

1. Research Question 1 (RQ1)

RQ1: What technical and methodological challenges arise when collecting data in hybrid learning environments, and how can they be addressed through methodological adjustments?

Technical Challenges:

None of the participants reported serious technical difficulties (e.g., unclear about the study procedure or software setup). This **smooth process** highlights the importance of **well-prepared infrastructure**—including stable internet connections, properly positioned cameras, and **clear participant instructions**—to ensure minimal disruption and high data quality.

Methodological Challenges:

External Distractions: Ambient noise, multitasking, or pop-up notifications can interfere with remote measurements, necessitating explicit guidelines (e.g., quiet space, no extra browser tabs).

In-Person Groups (ET/FEA)

Camera Awareness: Some participants felt distracted by the study equipment, especially in the FEA group, which was corroborated by self-reported higher distraction levels in surveys.

Rapid Setup Constraints: Coordinating ET glasses and stationary webcams for FEA within a short pre-lecture window requires streamlined protocols as shown in the specific procedures part in the methodology to maintain consistent data quality.

Conclusion for RQ1:

A robust technical setup, and clearly defined participation requirements are essential for reliable data collection in hybrid scenarios. Tailored protocols are needed to address distinct environmental factors in remote versus in-person sessions (e.g., standardized instructions, checks for proper camera alignment).

2. Research Question 2 (RQ2)

RQ2: What differences in gaze behavior and emotional responses can be observed between students attending in person and those participating remotely?

Gaze Behavior:

In-Person ET Group: Viewed both instructor and slides frequently, switching more often (shorter fixations).

RDC: Focused predominantly on slides; fewer total fixations on the instructor, and longer Time-to-First-Fixation (TTFF).

Emotional Responses

In-Person FEA: Higher engagement levels, more emotional variability.

Remote FEA (RDC): Consistently high neutrality (up to 90%), plus higher confusion during slide transitions.

Confusion

RDC participants showed higher confusion, especially during slide transitions, whereas in-person confusion was lower but accompanied by more signs of cognitive effort (brow furrow).

Conclusion for RQ2:

In-person participants divided their gaze more evenly between the instructor and the slides and showed a higher range of emotions and facial expressions. In contrast, remote participants were more slide-focused, took longer to look at the instructor, and displayed fewer emotional indicators, with a significant proportion of neutral expressions.

3. Research Question 3 (RQ3)

RQ3: What optimization strategies for hybrid lectures can be derived from the analysis of the collected data?

The data suggests multiple strategies:

Instructor Presence

A larger or more engaging instructor video feed can counteract remote participants' delayed fixation.

Interactive Elements and Structured Transitions

Clear verbal cues and mini summaries reduce confusion spikes, particularly during slide changes.

Distractions and Social Presence

Muting notifications and encouraging interactive elements (e.g., Q&A, polls) address higher distractions; remote learners benefit from explicit social cues to stimulate emotional engagement.

Conclusion for RQ3

Addressing **instructor visibility**, **structured slide transitions**, and **interactive prompts** can help remote participants **emulate** in-person attention and emotional engagement levels.

4. Evaluation of the Hypotheses

4.1 Hypothesis H1

H1: "Visual attention and emotional engagement are more pronounced for in-person students than for remote students."

Evidence:

ET: In-person students viewed the instructor more often, switching between slides and lecturer.

FEA: In-person participants displayed more varied emotional reactions and FE.

Self-Reports: In-person groups rated higher attention and motivation.

Outcome: H1 is supported. In-person participants demonstrate higher and more varied attention/engagement compared to remote participants.

4.2 Hypothesis H2

H2: "The analysis of ET and FE data yields actionable insights for optimizing hybrid lectures."

Evidence:

Identified Weaknesses: Remote conditions showed delayed lecturer fixation and increased confusion at slide transitions.

Potential Solutions: Strengthening instructor presence, adding interactive elements, and reducing distractions.

Survey Alignment: Remote participants reported lower attention and higher distraction, corroborating ET/FEA findings.

Outcome: H2 is confirmed. The data clearly reveals potential points for improvement, validating that ET and FEA can inform tailored strategies to enhance hybrid learning experiences.

Overall Summary

Overall, this thesis achieved its aims by integrating Eye-Tracking, Facial Expression Analysis, and survey data to illuminate key strategies for improving hybrid lectures. The Main Study findings demonstrate that in-person scenarios (In-Person ET/FEA) offer advantages in terms of balanced gaze distribution, faster fixation on the instructor, and more varied emotional engagement. In contrast, remote participants focus more on slides, take longer to notice the instructor, and exhibit a higher proportion of neutral expressions. Coupling ET, FEA, and survey data confirms both hypotheses, offering robust insights into where and how to refine hybrid lecture design. Furthermore, this study serves as a foundational investigation under the IMPULSE project, setting the stage for more extensive data-driven research on optimizing lecture content.

Outlook

Automated AOIs in iMotions

Recent updates in the iMotions software offer Automated AOIs, a feature that automatically detects and labels areas of interest within a video or screen recording. This functionality reduces the substantial manual effort formerly required to specify AOIs for ET analyses, minimizing time-intensive labor and the risk of human error. By streamlining data processing, Automated AOIs can significantly accelerate the research workflow and allow larger, more complex datasets to be handled more efficiently.

Audio Analysis with Audeering

Future work may also consider integrating Audeering, a platform that employs audio-based analytics to assess vocal qualities such as pitch, tonal variation and emotional nuances. By incorporating tonal analysis alongside eye-tracking and facial expression data, researchers could gain a more holistic understanding of lecture delivery. Tracking subtle shifts in vocal intensity or emotional undertones can help clarify how lecturer tonality influences student engagement and attention, thus enriching multi-modal datasets.

Continued Workshops and Knowledge Transfer

Project representatives from **BNCOW, BEACON** and **IMPULSE** will remain available to conduct further workshops at the UASD. This ongoing training ensures that expertise in iMotions, data analysis and educational research remains institutionalized. Regular sessions for faculty and research staff can mitigate the risk of knowledge loss, foster collective competencies in handling larger datasets and expand on future experiments in learning environments.

Time and Resource Considerations

Although the new software solutions (e.g., Automated AOIs, audio analytics) **may** ease the analytical burden, researchers must **allocate** sufficient **time** and **resources** for:

1. Processing large volumes of multimodal data (ET, FE, audio streams).
2. Training academic and technical staff to use and interpret these tools effectively.
3. Managing the complex logistics required for **longer or more comprehensive** data-collection sessions.

Recommendations for Future Research

1. **Dedicated Studies on Slide Design**

An in-depth exploration focusing **solely** on the interplay between **slide layout, text density, and student gaze** could clarify which design elements most **effectively** capture and sustain attention.

2. **Analysis of Lecturer Tonality**

A separate study examining **vocal modulation, pitch range, and emotional undertones** may shed light on how **tone** impacts engagement and confusion, especially in remote contexts. Integrating **Audeering** would facilitate these analyses.

3. **Longitudinal Research**

Conducting extended or repeated-measures experiments over an entire semester **could** evaluate whether the **optimization strategies** proposed here—such as improved camera setups—consistently **improve** visual attention and emotional engagement over time.

Recommendations for Implementation

1. **Larger Sample Sizes**

Future studies **might** aim for **50 or more** participants, potentially recorded with a 4K camera in a **large lecture hall**. This expanded participant pool may yield **statistically robust** findings and **finer-grained** analyses of engagement patterns.

2. **Long-Term and Multimodal Data**

Integrating **audio analytics** (e.g., with Audeering) and **repeated** measurement points (e.g., a semester-long hybrid course) **could** confirm whether short-term trends persist. It would also offer a **richer dataset** for exploring the relationship between **visual attention, facial expression, and auditory cues**.

3. **Streamlined Presentation Modes**

Employing **Presentation Mode** in Microsoft Teams or similar platforms **could** reduce extraneous on-screen elements, helping remote participants stay focused on essential lecture content. In large-scale hybrid sessions, this approach **may** help **synchronize** remote and in-person experiences.

By expanding study durations, integrating advanced tools like **Automated AOIs** and **Audeering**, and consolidating **institutional training** on these methodologies, future research can **further refine** hybrid lecture practices. These measures **may** enhance the **effectiveness and inclusivity** of hybrid learning environments for diverse student populations.

Concept: Semester-Long Multi-Modal Data Collection in a Large Lecture Hall

1. **Lecture Hall Setup**
    - **Over 100 Students**: The hall is divided so that one side seats participants **who consent** to record data, while the other side remains off-camera for those who opt out.
    - **4K Camera**: A high-resolution camera is permanently mounted at the front, capturing the consenting section in real time (potentially 50+ students at once). This setup ensures high-quality facial and body cues for FEA, even from a distance.

- **Ethical Compliance**: All consenting participants sign detailed forms explaining the data usage, privacy measures, and their right to withdraw, thereby maintaining robust ethical standards.

2. **Live iMotions Integration**
   - **Desktop PC Installation**: A dedicated computer in the lecture hall runs iMotions, initiating a **multimodal study** at the start of each class.
   - **Automatic Event Markers**: Whenever the lecturer advances a slide (e.g., clicks “Next”), a signal could automatically generate onset events in iMotions, removing the need for extensive manual timestamping post-lecture.
   - **Audio Analytics (Audeering)**: Alongside video data, the system records the lecturer’s vocal tonality, pitch fluctuations, and emotional undertones. This additional layer enriches the dataset by correlating vocal cues with student reactions.
3. **Ongoing Feedback via QR-Code Surveys**
   - **In-Class Surveys**: At the lecture’s conclusion, the final slide displays a **QR code** linked to a short questionnaire. Students scan it on their devices, providing immediate self-reflections on attention, motivation, and perceived distractions.
   - **Qualitative Insights**: The resulting feedback complements the quantitative metrics (Eye-Tracking, FEA, audio data) with subjective impressions—helping correlate measured engagement with students’ own experiences.
4. **Minimal Additional Effort**
   - **Streamlined Workflow**: Because all hardware and software (camera, microphones, iMotions) are permanently installed, initiating each study session merely requires the lecturer to **start the recording**. This “one-click” approach dramatically reduces setup time across the semester.
   - **Automatic Data Synchronization**: With event triggers tied to slide transitions, the data is neatly segmented for subsequent analysis, freeing researchers from extensive post-processing tasks.

5. **Comprehensive Analytics**
    - **Facial Expression Analysis (FEA)**: Large-group FEA captures collective trends (e.g., spikes in confusion or neutrality) and identifies consistent emotional responses to content or delivery style.
    - **Audio/Tonality via Audeering**: Subtle shifts in the lecturer's tone can be correlated with real-time changes in student engagement or confusion, linking **vocal cues** to **nonverbal feedback**.
    - **Eye-Tracking (if feasible)**: While large-group ET is still evolving, even partial gaze capture—like transitions between the screen and lecturer—could validate more granular attention patterns.

Overall Summary

The results of the main study highlight clear differences between the **in-person (ET/FEA**) and **RDC** groups. In-person scenarios result in balanced gaze distribution, faster fixation on the instructor and richer emotional expression. In contrast, remote participants spend more time fixating on slides, take longer to shift attention to the lecturer and show a predominance of neutral facial expressions. Combining ET, FEA and survey results not only confirms both hypotheses, but also reveals concrete areas for improvement:

- **Technical & Methodological**: Stable setups, camera positioning, standardized protocols and minimal equipment intrusiveness are essential.
- **Attention & Emotional Responses**: Remote learners benefit from enhanced lecturer framing, structured transitions and interactive cues that replicate some of the in-person advantages.
- **Actionable Insights**: Clear environmental guidelines, personal framing, and frequent lecturer prompts can help reduce distractions and improve remote emotional engagement.

References

- Bertelsmann Stiftung, *Monitor Digitale Bildung #2 – Die Hochschulen im digitalen Zeitalter*. Gütersloh, Germany: Bertelsmann Stiftung, 2017. [Online]. Available: https://www.bertelsmann-stiftung.de/fileadmin/files/Projekte/28_Bildung_Monitor_Digitale_Bildung/Studie_Hochschulen_Vollversion_final.pdf
  Accessed: [Date].
- A. Bojko, "Informative or misleading? Heatmaps deconstructed," in *Proc. Human Factors and Ergonomics Society Annu. Meeting*, vol. 53, no. 27, 2009, pp. 210–214.
- M. Chan, C. R. Madan, and A. Singhal, "The influence of visual attention on learning and memory: Insights from eye-tracking studies," *Psychon. Bull. Rev.*, vol. 29, pp. 1234–1250, 2022.
- B. De Carolis, A. M. Novellis, N. Plantamura, and V. Striani, "Engaged faces: Measuring and monitoring student engagement from face and gaze behavior," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. Workshops*, 2019, pp. 197–200.
- M. A. Dewan, M. Murshed, and F. Lin, *Engagement Detection in Online Learning: A Review*, vol. 6, no. 1. Cham, Switzerland: Springer, 2019.
- A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 2nd ed. London, U.K.: Springer, 2007.
- P. Ekman and W. V. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA, USA: Consulting Psychologists Press, 1978.
- J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris, "School Engagement: Potential of the Concept, State of the Evidence," *Rev. Educ. Res.*, vol. 74, no. 1, pp. 59–109, 2004.
- K. Holmqvist, M. Nystrom, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford, U.K.: Oxford Univ. Press, 2011.
- Hochschule Düsseldorf, "BNCOW – Business Negotiation Competencies in Online World," [Online]. Available:
  https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/BNCOW.aspx
  Accessed: Mar. 5, 2025.
- Hochschule Düsseldorf, "BEACON: Behavioral Analysis of Communication and Negotiation," [Online]. Available:
  https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/BEACON.aspx
  Accessed: Mar. 5, 2025.
- Hochschule Düsseldorf, "IMPULSE: Investigating Metrics of Participation, Understanding, and Lecture Engagement," [Online]. Available:
  https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/IMPULSE.aspx
  Accessed: Mar. 5, 2025.
- Hochschule Düsseldorf, "Technical Sales Lab," [Online]. Available:
  https://ei.hs-duesseldorf.de/forschung/technical_sales_lab/Seiten/default.aspx
  Accessed: Mar. 1, 2025.
- iMotions, "iMotions – We Power Human Insight," [Online]. Available:
  https://imotions.com/
  Accessed: Mar. 1, 2025.
- H. Jarodzka, N. Janssen, P. A. Kirschner, and G. Erkens, "Teacher's gaze and its relation to students' engagement and learning," *Comput. Educ.*, vol. 175, p. 104324, 2021.
- L. van Marlen, M. van Wermeskerken, H. Jarodzka, and T. van Gog, "Effectiveness of eye movement modeling examples in problem solving: The role of verbal ambiguity and prior knowledge," *J. Comput. Assist. Learn.*, vol. 34, no. 6, pp. 717–727, 2018.
- M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. Tetzlaff, A. Akl, and G. E. Moher, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, 2021.
- UNESCO, "COVID-19 Educational Disruption and Response," 2020. [Online]. Available:
  https://en.unesco.org/themes/education-emergencies/coronavirus-school-closures
- T. van Gog and K. Scheiter, "Eye tracking as a tool to study and enhance multimedia learning," *Learn. Instr.*, vol. 20, no. 2, pp. 95–99, 2010.

Appendix

Eidesstattliche Erklärung

Hochschule Düsseldorf
University of Applied Sciences

HSD

Eidesstattliche Versicherung

[redacted]

Name, Vorname — Matrikelnummer

Hiermit versichere ich an Eides Statt, dass ich die in elektronischer Form abgenommene schriftliche Prüfung

Bezeichnung der Prüfung: Abschlussarbeit (Bachelor Thesis), PNr.: [redacted]
(Modul, Prüfungsnummer)

Prüfer/in: Lang, Kati

Prüfungs-/Abgabedatum: 26.03.2025

eigenhändig erbracht habe. Bei der Bearbeitung habe ich keine unzulässigen Hilfsmittel benutzt und mich nicht der unerlaubten Hilfe Dritter bedient. Ich habe keine anderen als die angegebenen Quellen benutzt und, soweit von der Aufgabenstellung vorgesehen, die aus fremden Quellen direkt oder indirekt übernommenen Inhalte als solche kenntlich gemacht.

Ich bin darüber belehrt, dass die vorsätzlich oder auch nur fahrlässig falsche Abgabe einer eidesstattlichen Versicherung nach §§ 156, 161 StGB* strafbar ist.

[redacted]

Ort, Datum — Unterschrift

* **§ 156 StGB - Falsche Versicherung an Eides Statt**

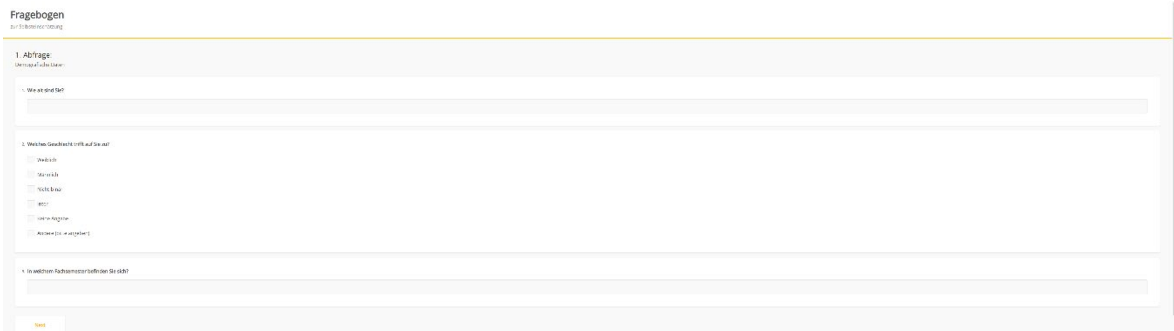
Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

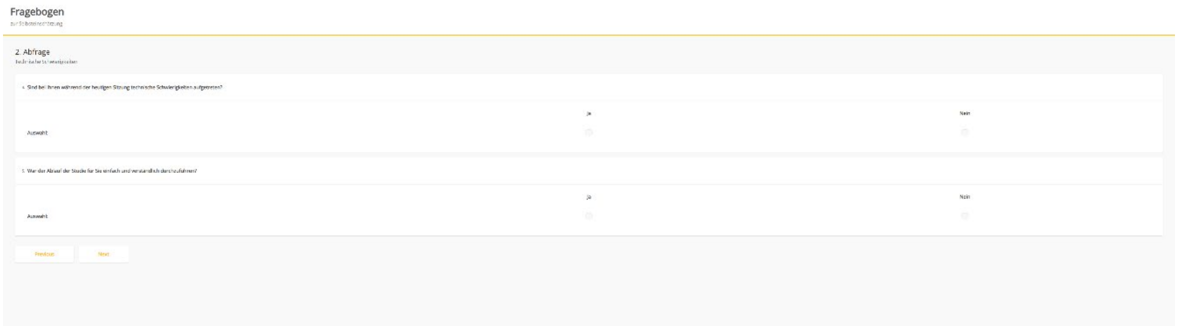
§ 161 StGB - Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt

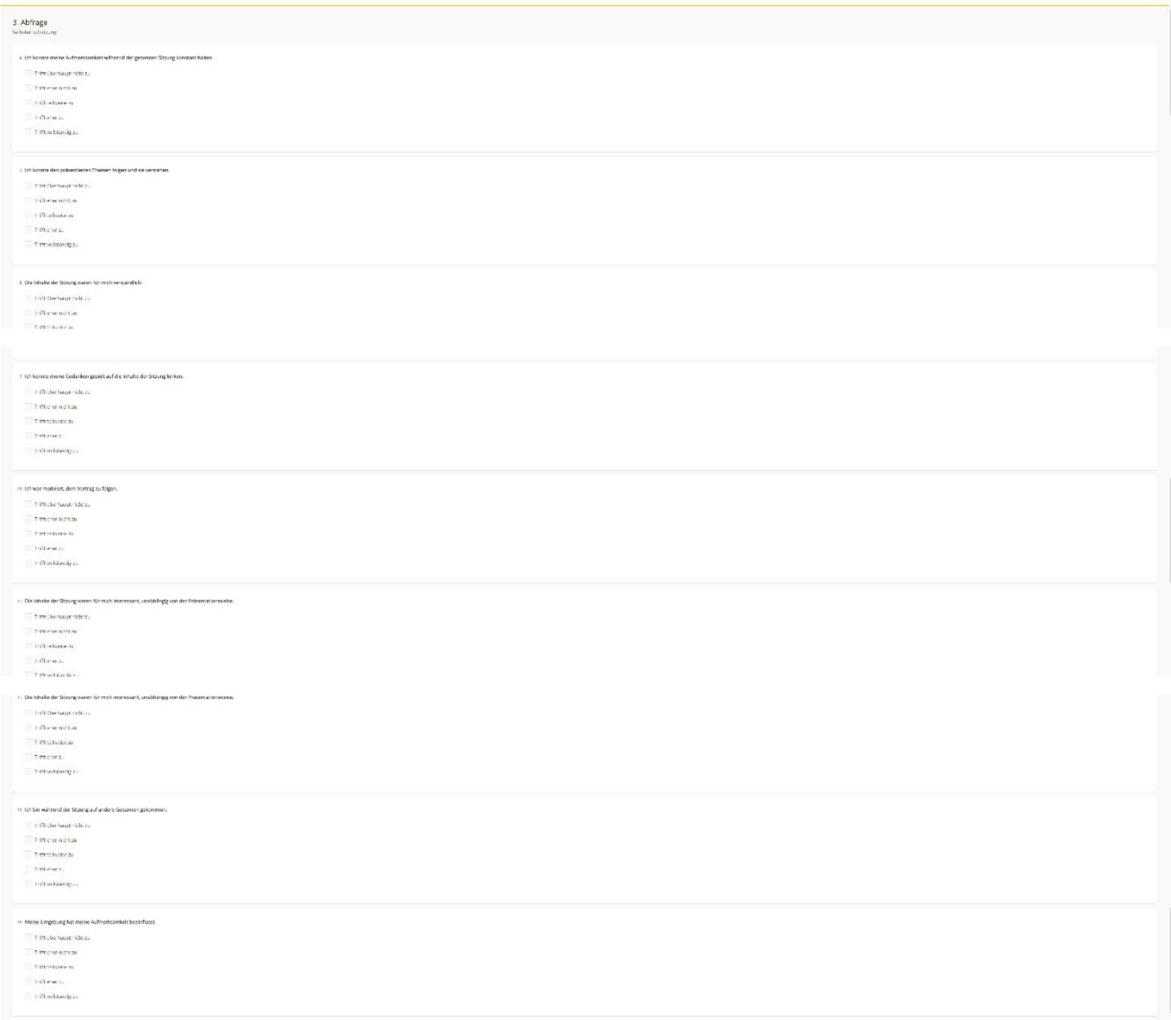
(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.

(2) Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.

Survey Remote







Survey In-Person

1. Abfrage

Demografische Daten:

Wie alt sind Sie? (Freitext)

Welches Geschlecht trifft auf Sie zu?

☐ Weiblich ☐ Männlich ☐ Nicht-binär ☐ Inter ☐ Keine Angabe

In welchem Fachsemester befinden Sie sich? (Freitext)

---

2. Abfrage

Technische Schwierigkeiten:

Sind bei Ihnen während der heutigen Sitzung technische Schwierigkeiten aufgetreten?

Auswahl: ☐ Ja ☐ Nein

War der Ablauf der Studie für Sie einfach und verständlich durchzuführen?

Auswahl: ☐ Ja ☐ Nein

3. Abfrage

Ich konnte meine Aufmerksamkeit während der gesamten Sitzung konstant halten.

☐ Trifft überhaupt nicht zu ☐ Trifft eher nicht zu ☐ Trifft teilweise zu ☐ Trifft eher zu ☐ Trifft vollständig zu

---

Ich konnte den präsentierten Themen folgen und sie verstehen.

☐ Trifft überhaupt nicht zu ☐ Trifft eher nicht zu ☐ Trifft teilweise zu ☐ Trifft eher zu ☐ Trifft vollständig zu

---

Die Inhalte der Sitzung waren für mich verständlich.

☐ Trifft überhaupt nicht zu ☐ Trifft eher nicht zu ☐ Trifft teilweise zu ☐ Trifft eher zu ☐ Trifft vollständig zu

---

Ich konnte meine Gedanken gezielt auf die Inhalte der Sitzung lenken.

☐ Trifft überhaupt nicht zu ☐ Trifft eher nicht zu ☐ Trifft teilweise zu ☐ Trifft eher zu ☐ Trifft vollständig zu

---

Ich war motiviert, dem Vortrag zu folgen.

☐ Trifft überhaupt nicht zu ☐ Trifft eher nicht zu ☐ Trifft teilweise zu ☐ Trifft eher zu ☐ Trifft vollständig zu

---

Die Inhalte der Sitzung waren für mich interessant, unabhängig von der Präsentationsweise.

☐ Trifft überhaupt nicht zu ☐ Trifft eher nicht zu ☐ Trifft teilweise zu ☐ Trifft eher zu ☐ Trifft vollständig zu

---

Ich bin während der Sitzung auf andere Gedanken gekommen.

☐ Trifft überhaupt nicht zu ☐ Trifft eher nicht zu ☐ Trifft teilweise zu ☐ Trifft eher zu ☐ Trifft vollständig zu

---

Meine Umgebung hat meine Aufmerksamkeit beeinflusst.

☐ Trifft überhaupt nicht zu ☐ Trifft eher nicht zu ☐ Trifft teilweise zu ☐ Trifft eher zu ☐ Trifft vollständig zu

Synchronization and Annotation Procedure

To ensure **precise alignment of all data**, a synchronization procedure based on the **Clap Marker** was developed and implemented. This approach provided **a synchronization** across all studies.

The **Clap Time** for each participant was documented in the **iMotions standard format** (**hh:mm:ss:ms**, 00:00:00:000) and entered into a **reference table**. The iMotions Clap Marker timestamp was initially converted into milliseconds, ensuring high-precision calculations without rounding errors in the reference table.

Following the initial **Clap Marker**, the **standardized annotation durations** for key lecture events were extracted from the **Environmental Cam recording**.

After defining the **annotation durations**, the **start times** of each event marker were systematically calculated for all participants. The **start time of each subsequent event** was calculated as follows:

$$Start\_Time_{(n)} = Start\_Time_{(n-1)} + Duration_{(n-1)}$$

The **annotation timestamps** were then reformatted into the standard iMotions time notation (hh:mm:ss:msmsms)manually entered into **iMotions** for each dataset.

RDC Calibration

Pre-Study

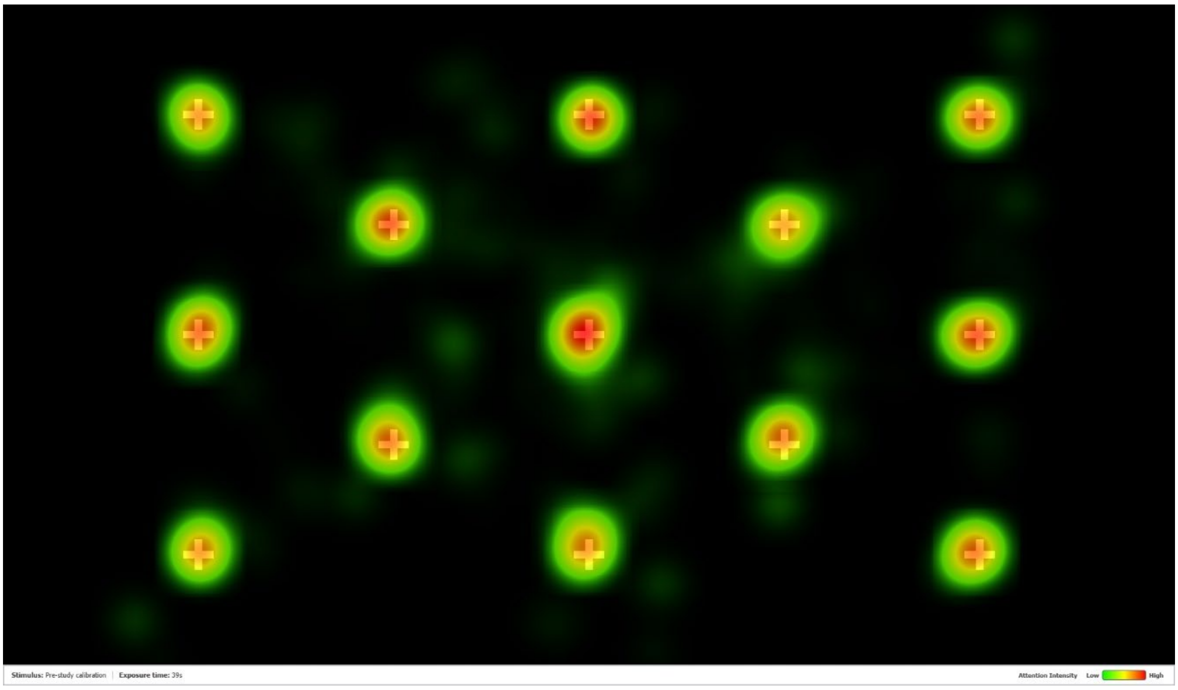


Figure 28: Aggregated Heatmap Pre-Study Calibration

Post-Study

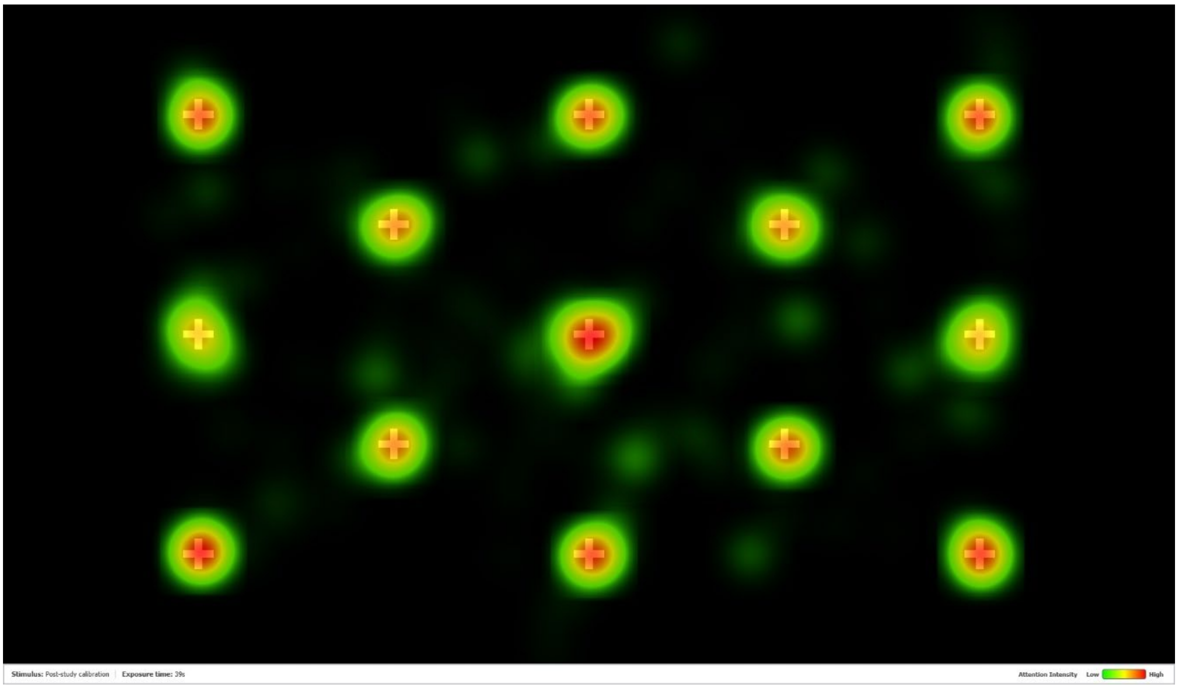


Figure 29: Figure 28: Aggregated Heatmap Post-Study Calibration

The following AI models and tools were used to enhance the fluency of the text and to create images: ChatGPT, GPT-4, Litmaps AI, Deepl, Canva and other advanced language models and tools.