

Learning with AI based Videos

Patrick Struger, BSc
Institute of Human-Centred Computing (HCC)

29.01.2025

Outline

- ▶ Introduction
- ▶ Method
- ▶ Results
- ▶ Discussion

Introduction

Learning with AI based Videos

Objective

Analyzing the use of learning videos, created with AI human avatars compared to real video presenters

Aim to identify differences in terms of

- Cognitive Engagement
- Emotional Engagement
- Personal Perception

Empirical study based on qualitative and quantitative analysis of 55 participants



Motivation

Progressive development of AI in the learning sector

- Content Generation
- Intelligent Tutoring
- Learning Analytics

Lack of research using AI tools considering

- Cognitive Engagement
- Emotional Engagement



Learning Content Delivery

...can be communicated via audio, video, animations, slides, text-based etc.

Educational Videos

highly effective educational tool when considering

- Cognitive Load
- Student Engagement
- Active Learning

Instructor Presence

concerned as one of the fundamental design aspects in learning videos

- Social Presence
- Gesture Effects
- Para-social Interaction
- Split Attention Effect

with AI

new approaches and tools based on *Generative Artificial Intelligence (GAI)*

- Content Generation
- Virtual Instructors
- Quality Assurance

Learning Engagement

Behavioral LE

observable actions and participation of students during learning activities

- watching behavior
- taking notes
- answering embedded questions

Emotional LE

refers to feelings and attitudes while interacting with the content

- Emotional Impressions
- Social Engagement
- Personal Perception

Cognitive LE

mental effort invested in processing and understanding the content

- focusing attention
- analyzing information
- applying critical thinking
- connecting new knowledge with prior understanding

Focus of this study

Focus of this study

Measurement of Learning Engagement

based on published peer-reviewed systematic review by *Struger, Brünner and Ebner (2024)*

- Metadata
- Objective Measurements Used for this study
- Gamification Indicators
- Affective Computing Used for this study
- Personal Feedback Used for this study

Research Questions

RQ1: Does the use of AI avatars in learning videos affect their quality from the learners' perspective?

RQ2: Is the recognition software *FaceReader Online* able to track reliable emotional states while watching learning videos?

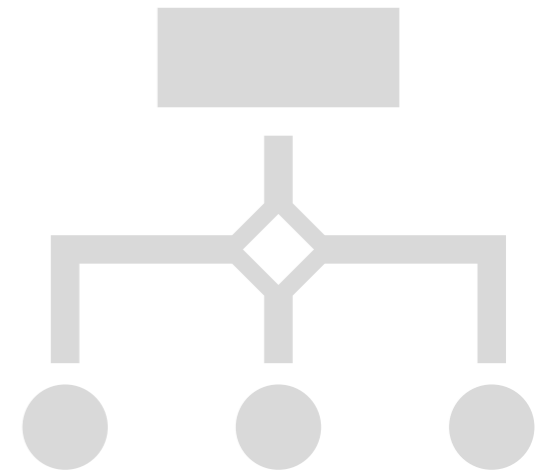
RQ3: Is there a significant difference in emotion and therefore in Emotional Learning when using AI generated presenters?

Method

Learning with AI based Videos

Implementation

- Preprocessing
- Experiment Setup
- Evaluation



Used Tools

- **HeyGen** (*Text-to-Video TTV*)
Preprocessing + Experiment Setup
- **FaceReader Online** (*Emotion Recognition System ERS*)
Preprocessing + Experiment Setup + Evaluation
- **Whisper** (*Automatic Speech Recognition ASR*)
Evaluation



Preprocessing

Participant Acquisition

through course lecture and personal acquisition



Theoretical Preprocessing

video topics, screenplays, scripts, post assessment



Technical Preprocessing

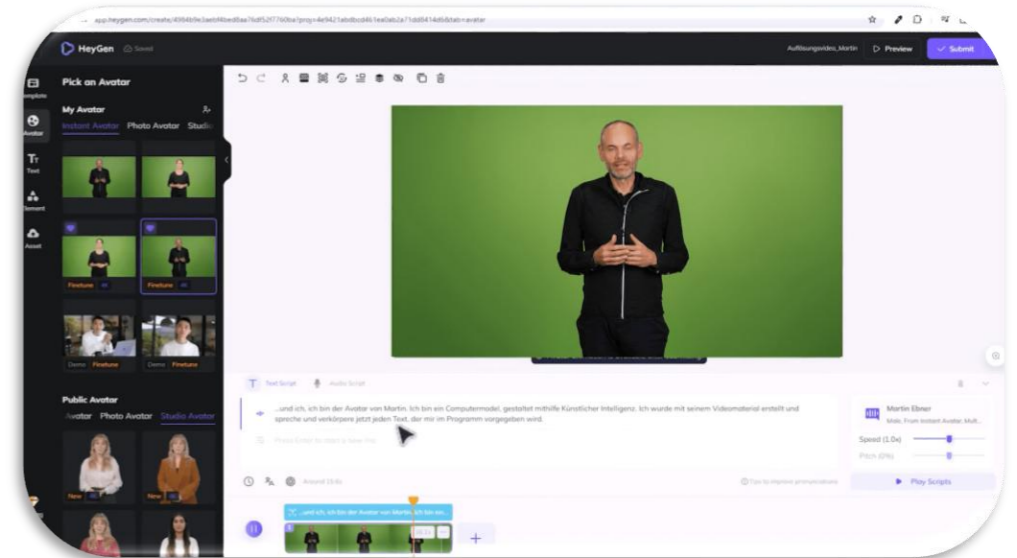
video creation, instructor generation (TTV), initializing project for ERS



Source: <https://www.heygen.com/>



Source: <https://facereader-online.com/>



Source: HeyGen User Interface

Experiment Setup

Micro Learning Courses



12 min

Facial Analysis
Post Assessment

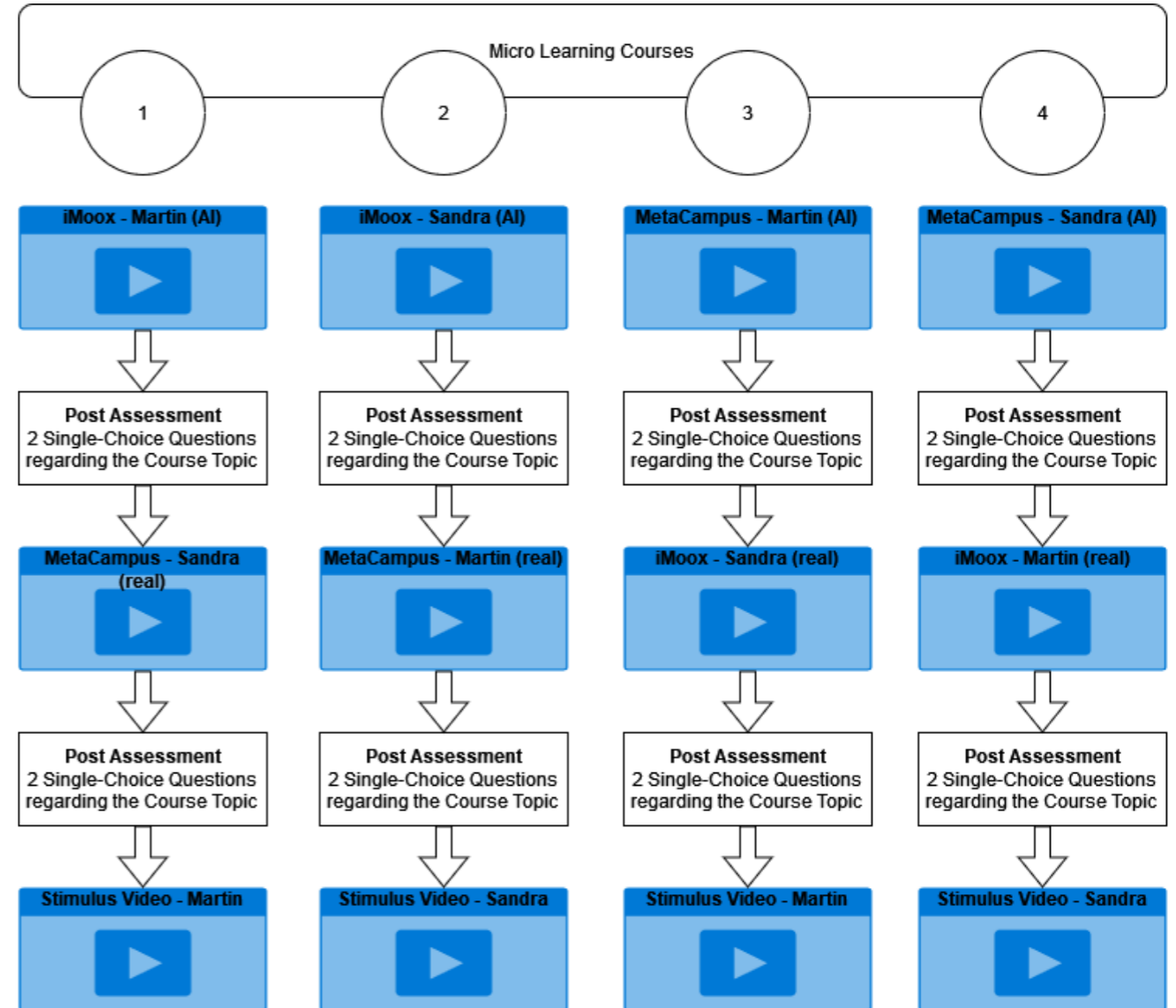


Structured Interviews



10 min

Recorded face-to-face
interviews with guided
questions



Evaluation

Quantitative Analytics

Emotional Classification
Post Assessment

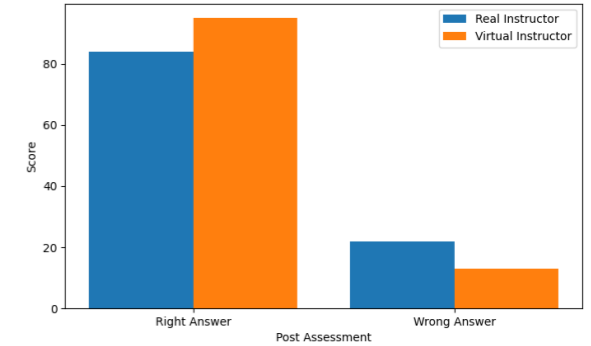
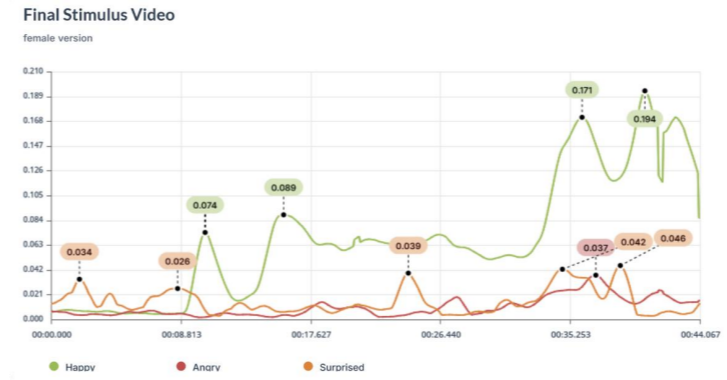


Qualitative Analytics

Interview Transcripts



Source:
<https://facereader-online.com/>



Whisper OpenAI

Qualitative Content Analysis
by Philipp Mayring (2019)

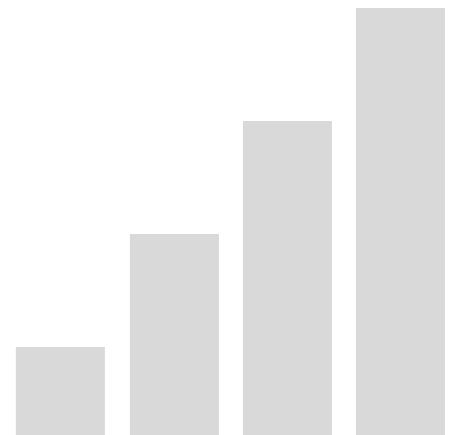
Disadvantages of AI in Learning Environments			
Category	Description	Example	Frequency
Potential for Cheating or Plagiarism	This category portrays the potential of generated content in terms of irresponsible use.	"One can certainly expect an increased potential for cheating or plagiarism, for example in text generation for academic papers."	5
Losing the Learning Process	This category describes that some participants found that using AI can lead to loss in the learning process itself.	"When we rely on AI, the learning process is lost or laziness is encouraged."	4

Results

Learning with AI based Videos

Outline

- Post Assessment ($RQ1$)
- Emotional Classification ($RQ3$)
- Structured Interviews ($RQ1 + RQ3$)
- Stimulus Video ($RQ2$)



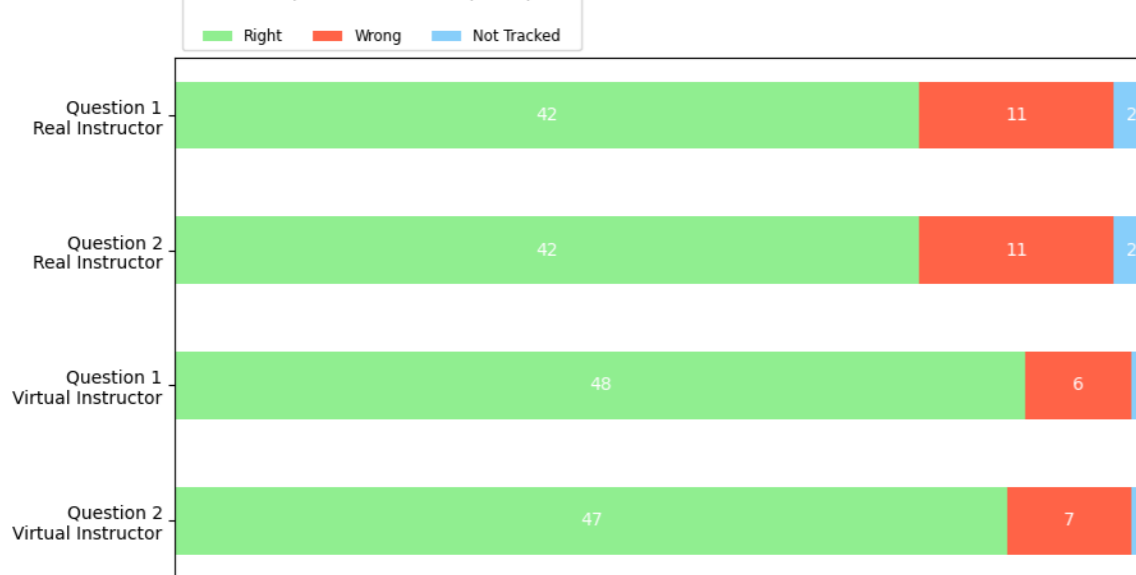
Post Assessment

55 Participants – each 2 questions per course topic / instructor

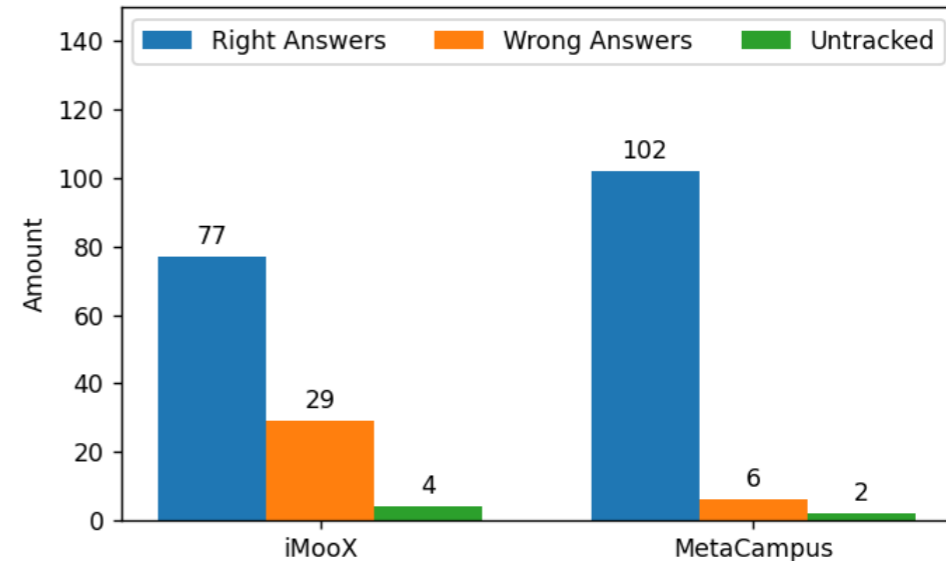
Answers — Virtual Instructor

		Right		Wrong		Not Finished		Total	
		n	%	n	%	n	%	n	%
Answers — Real Instructor	Right	72	65.45%	10	9.09%	2	1.82%	84	76.36%
	Wrong	19	17.27%	3	2.73%	0	0%	22	20%
	Not Finished	4	3.64%	0	0%	0	0%	4	3.64%
Total		95	86.36%	13	11.82%	2	1.82%	110	100%

Answers to post assessment of participants



Answers to Post Assessment grouped by Course Topic



Emotional Classification

through FaceReader face-recognition

Source:
<https://www.youtube.com/watch?v=0vIJ-8gXMII>

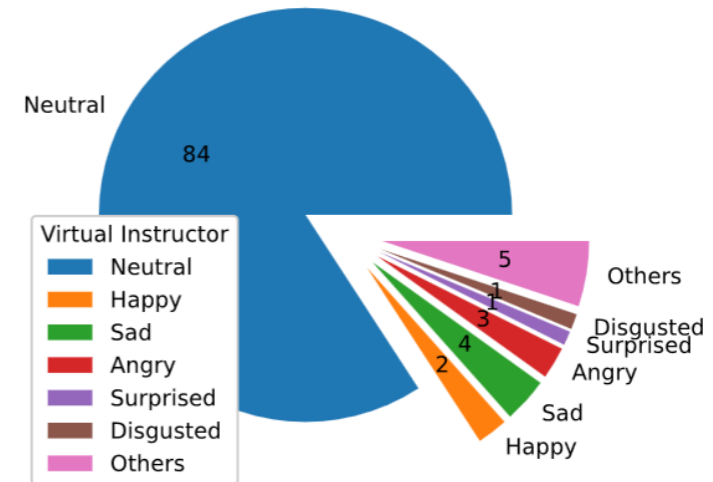
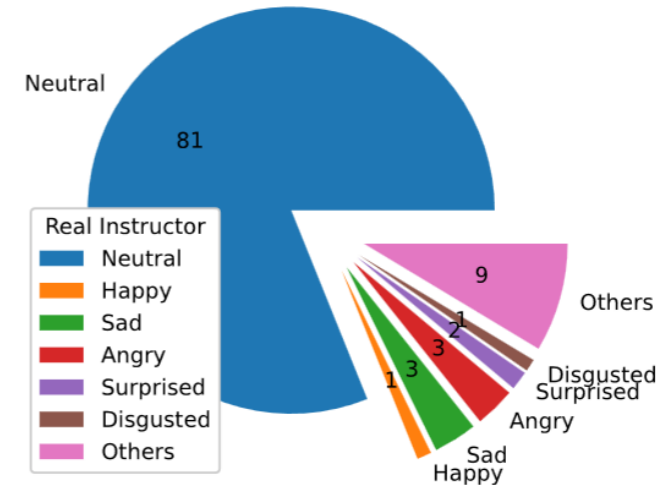


Emotional Classification

55 Participants – 6 standard emotions

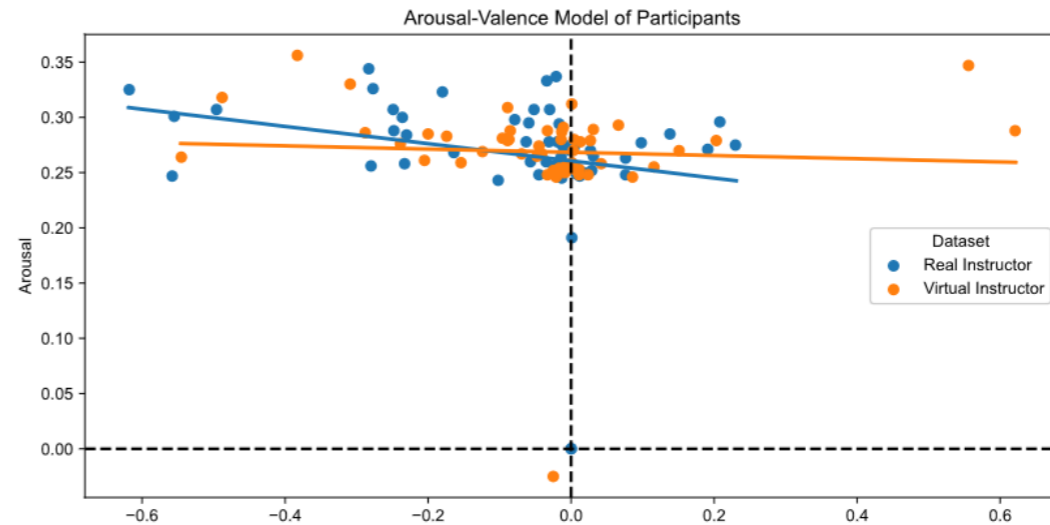
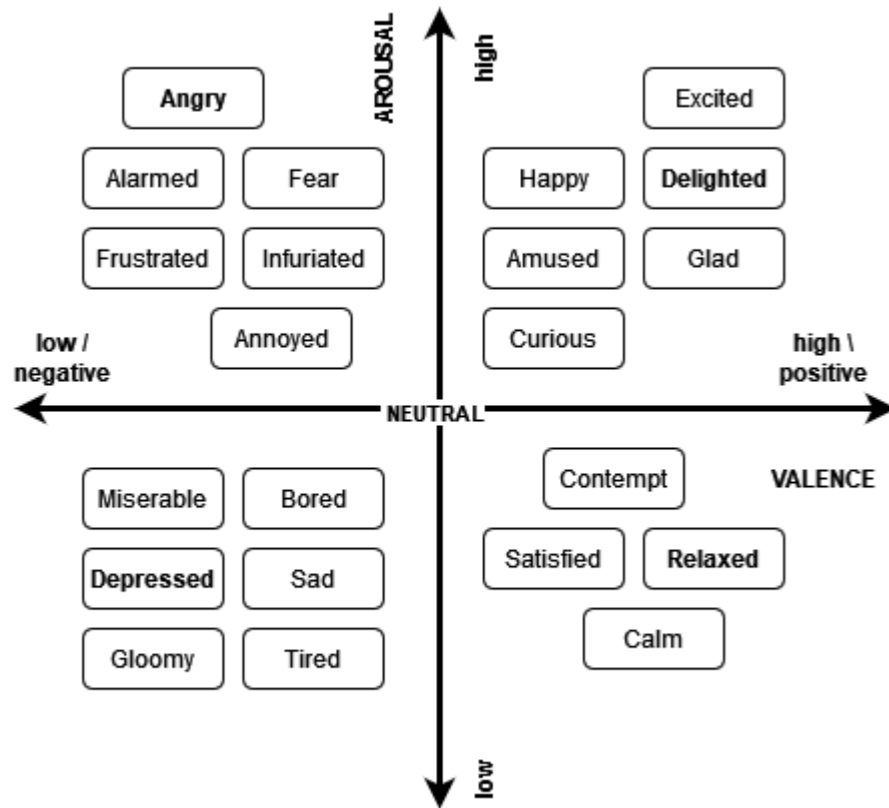
Real Instructor	Percentage of Emotional States					
	Neutral	Happy	Sad	Angry	Surprised	Disgusted
MetaCampus Sandra	0.728	0.020	0.073	0.054	0.006	0.005
MetaCampus Martin	0.780	0.019	0.018	0.020	0.010	0.032
iMooX Sandra	0.890	0.003	0.024	0.033	0.021	0.001
iMooX Martin	0.846	0.004	0.023	0.021	0.023	0.001
	0.811	0.0115	0.0345	0.032	0.015	0.00975

Virtual Instructor	Percentage of Emotional States					
	Neutral	Happy	Sad	Angry	Surprised	Disgusted
MetaCampus Sandra	0.875	0.011	0.021	0.014	0.012	0.001
MetaCampus Martin	0.921	0.005	0.014	0.026	0.020	0.002
iMooX Sandra	0.775	0.075	0.024	0.018	0.007	0.037
iMooX Martin	0.797	0.005	0.079	0.041	0.005	0.007
	0.842	0.024	0.035	0.025	0.011	0.012



Emotional Classification

55 Participants – arousal valence model



Arousal - the intensity or energy level of an emotion

Valence - how positive or negative the emotion feels

Structured Interviews

41 Participants – inductively formed categories during qualitative content analysis

Perception

- Conspicuousness in pauses, spelling and voice modality
- Artificially movement

Disdvantages of AI

- Potential for cheating
- Losing the learning process

Personal Perspective

- Translation using AI
- Preference for real Instructor

Learning Experience

- Equality of understanding
- Equality of learning
- Social distance
- Distraction by the avatar

Threats about AI

- Credibility, Trust and Ethics
- Privacy concerns
- Fear of job losses

Advantages of AI

- Overcoming language barriers
- Efficient and time saving
- Fast content generation

Future of AI

- Desired future use
- Quality assurance
- Labeling
- Individual Learning

AI Workshop-Day

11 Participants – graduating class of The Federal Upper Secondary School Monsberger Graz (aged 17-21)

Perception

- Slow perception of language and movement
- emphasis of words inappropriate
- gestures seemed blunt and inhuman

Learning Experience

- can support work
- makes no difference in learning compared to a real person

Future of AI

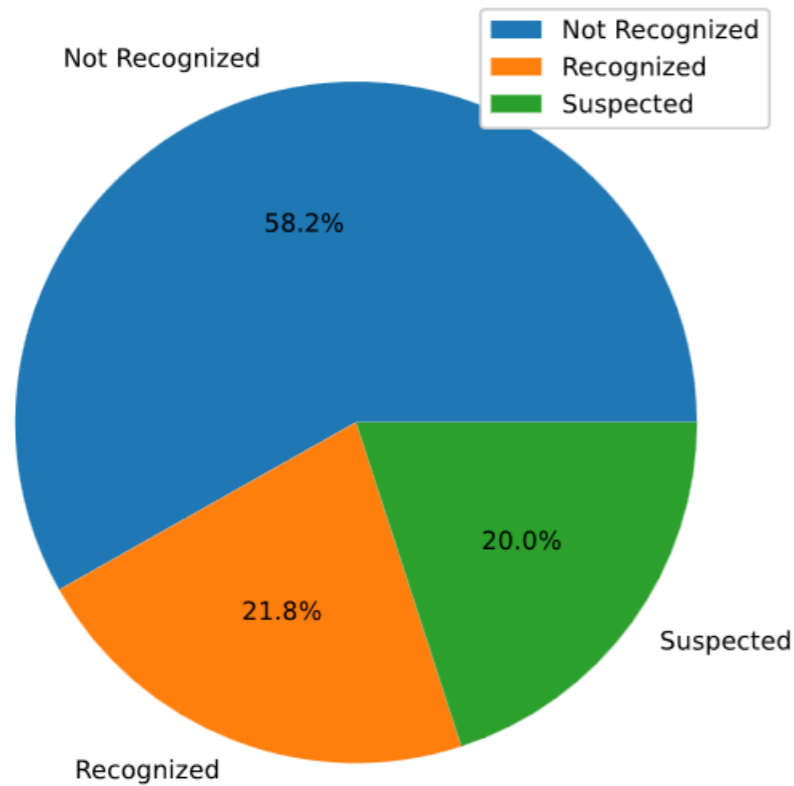
- information to be processed and made available faster
- Negatively influence the learning process
- Relying too much on AI
- constantly available for feedback



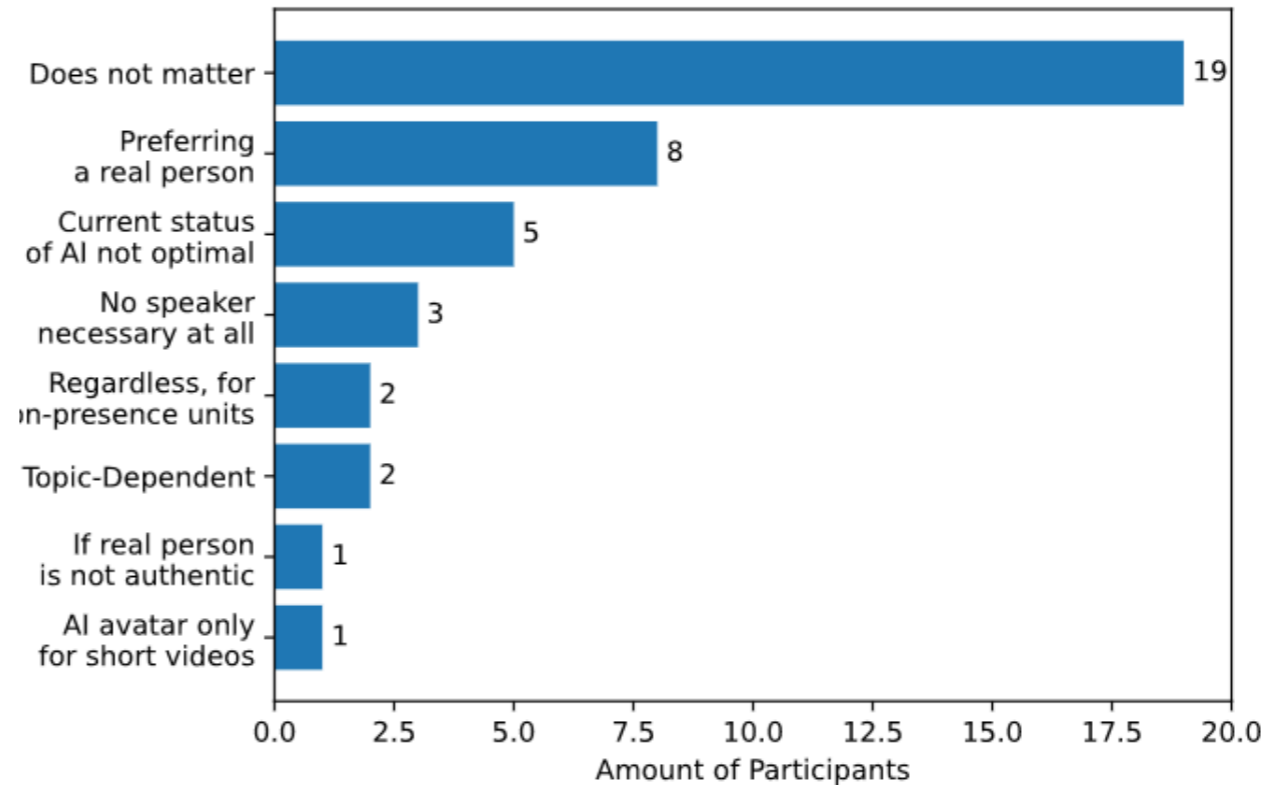
Additional Metrics

55 Participants – recognition and preferences

Number of participants recognizing AI instructor



Statements regarding the preference of AI generated instructors in learning videos



Stimulus Video

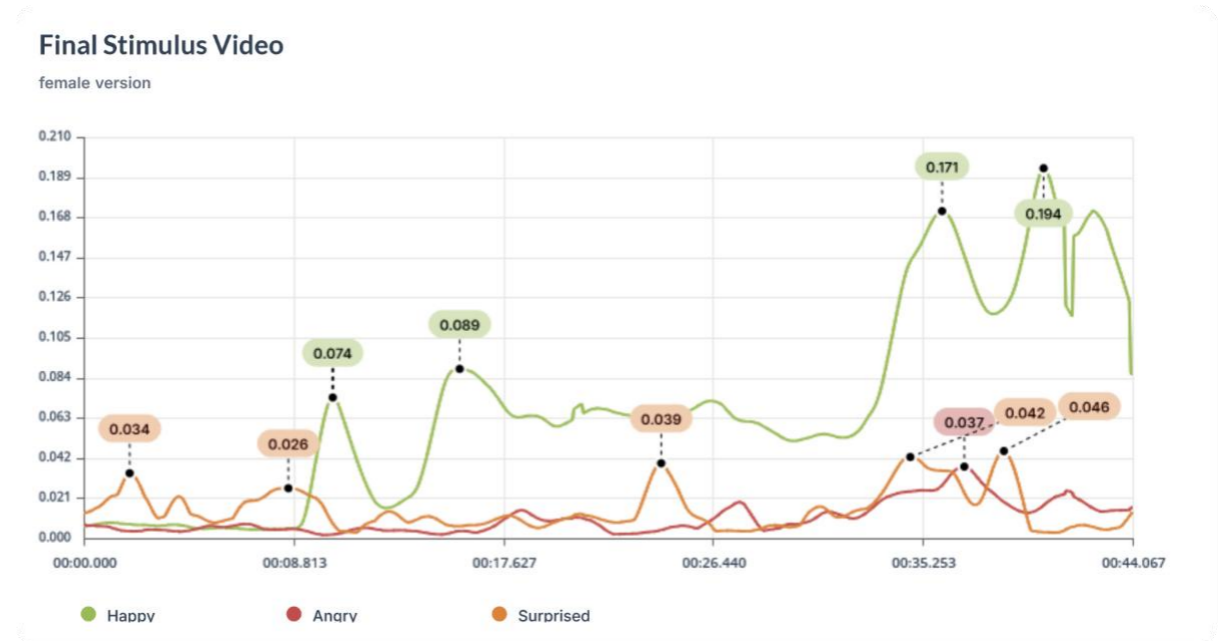
55 Participants

- validation of emotional classification
- peaks of emotional indicators
 - after 8 seconds - "Surprised"
 - after 35 seconds – "Happy"



Source:

<https://doi.org/10.3217/ggn-cr-sg773> CC BY Lehr- und Lerntechnologien, TU Graz



Discussion

Learning with AI based Videos

Answering Research Questions

RQ1: Does the use of AI avatars in learning videos affect their quality from the learners' perspective?

- Post-assessments revealed slightly better performance
- Perceived effectiveness of AI-generated instructors
- Predominantly neutral emotional state
- Arousal-valence model confirmed minimal emotional distinctions

However the lack of natural gestures and emotional authenticity was reported

Answering Research Questions

RQ2: Is the recognition software *FaceReader Online* able to track reliable emotional states while watching learning videos?

- proved effectiveness in classifying predefined emotional states during the stimulus video

Experienced downsides:

- long processing times
- occasional errors
- reduced accuracy for participants with glasses or squinting eyes

Answering Research Questions

RQ3: Is there a significant difference in emotion and therefore in Emotional Learning when using AI generated presenters?

- Quantitative results indicate no significant differences in emotional responses
- Challenge of maintaining attention due to AI avatars' repetitive gestures
- Distraction due to occasional inappropriate facial expressions
- Lack of tone modulation can limit the avatars' ability to communicate

Learning with AI based Videos

Patrick Struger, BSc
Institute of Human-Centred Computing (HCC)

29.01.2025