



TECHNISCHE
UNIVERSITÄT
DARMSTADT

GENERATION OF LOG FILES IN ONLINE STUDIES USING COMPUTER VISION

Brazilian-German Symposium on
Learning Analytics

Björn Mattes

OUTLINE

- 1** Why Online Studies?
- 2** Own Motivation for Log Files
- 3** Data Protection Issues
- 4** Idea: Screen Recordings
- 5** Use of Computer Vision

- 6** Creation of Log Data
- 7** Qualitative Analysis
- 8** Discussion

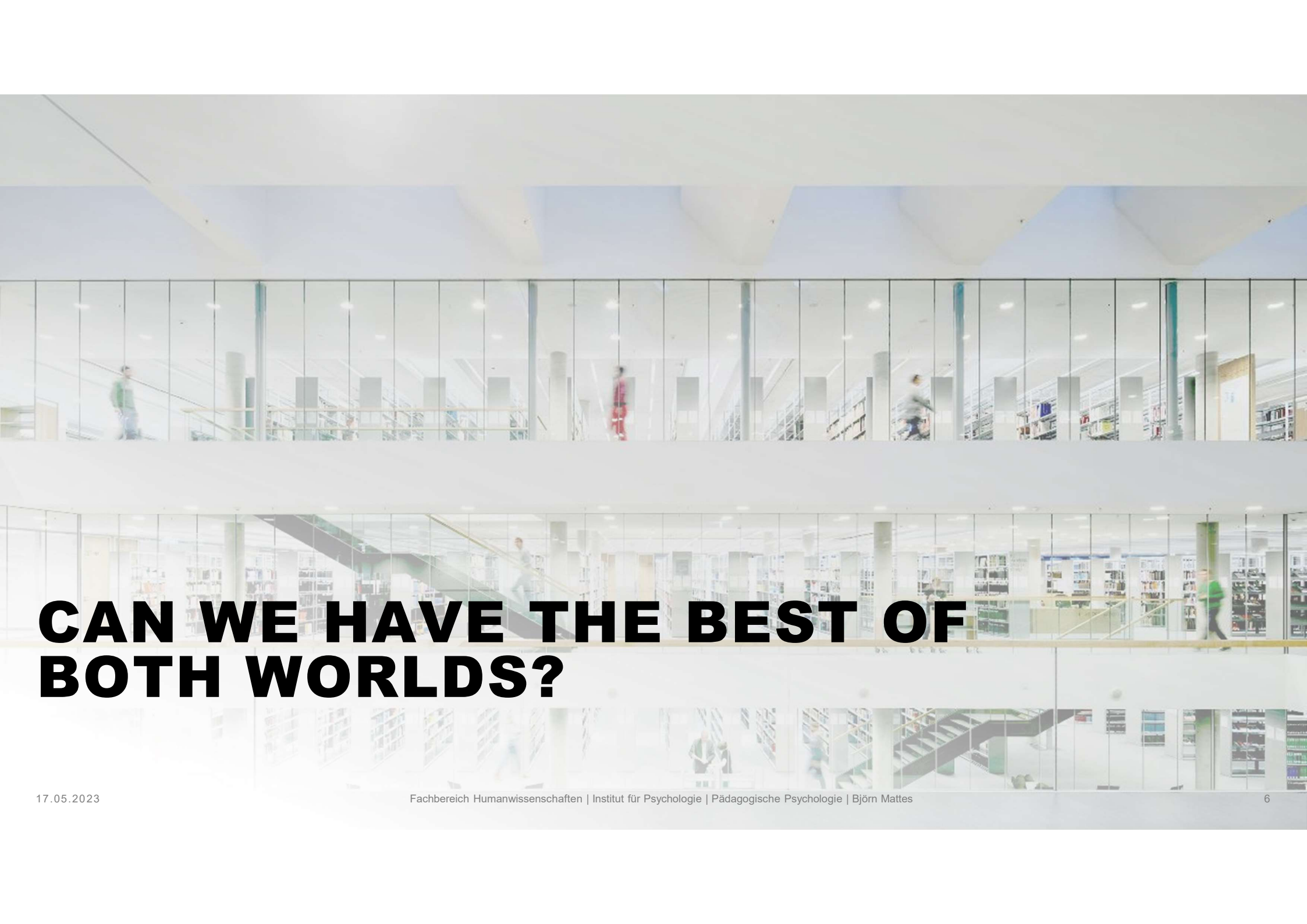
WHY ONLINE STUDIES?

REASONS TO DO ONLINE STUDIES

- **Temporal Flexibility:** Online studies offer participants the flexibility to engage at their own convenience
- **Cost and Resources:** Compared to lab-based studies, online research often requires less financial and physical resources
- **Reach:** Online research platforms provide access to diverse participant pool
- **Anonymity:** Online anonymity can reduce social desirability bias, leading to more honest and authentic responses

REASONS TO DO “LAB” STUDIES

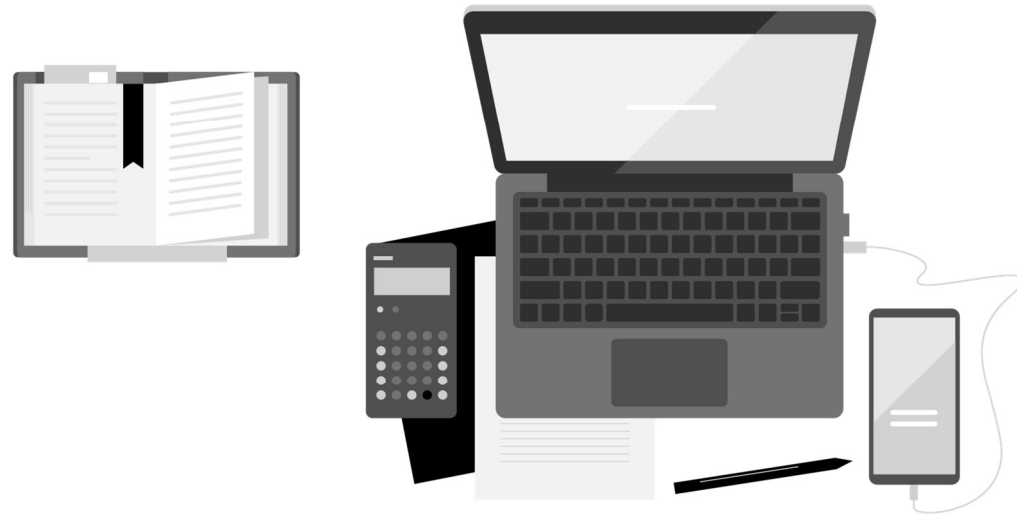
- **Control:** Lab-based research offers a controlled environment that minimizes external and confounding variables, increasing the internal validity of the study
- **Standardization:** Lab settings enable strict standardization of protocols and procedures across participants, promoting consistency
- **Participant Supervision:** Direct supervision of participants ensures compliance with the study protocol and reduces the likelihood of distractions



CAN WE HAVE THE BEST OF BOTH WORLDS?



MY OWN MOTIVATION FOR LOG FILES

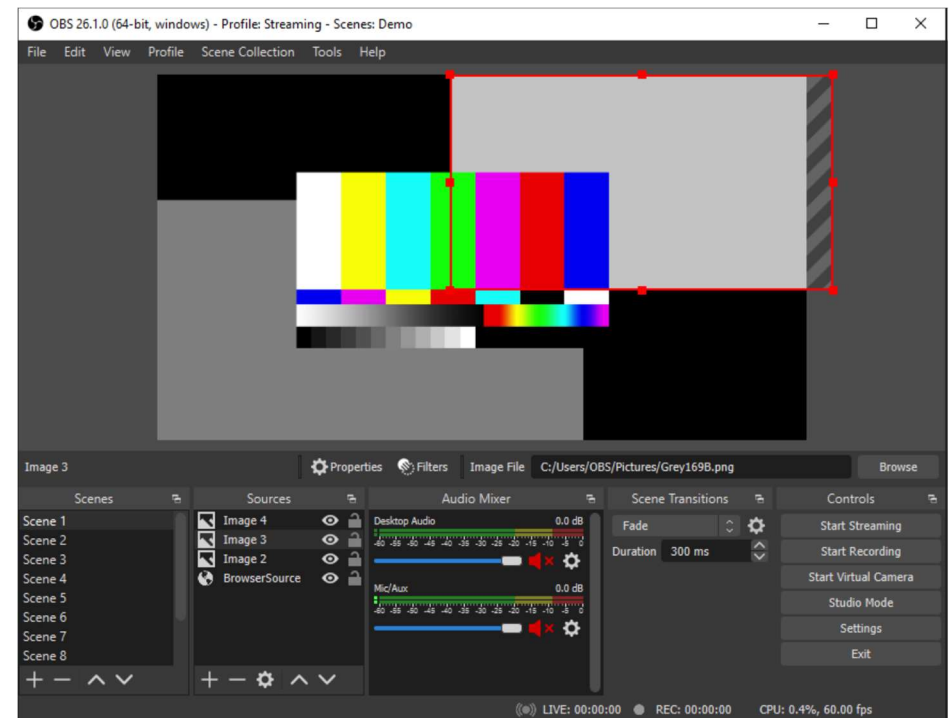


USE LOGGING-PROGRAMS?

- Significant privacy concerns are raised when software like **key-loggers** or **cloud-monitoring tools** are installed on study participants' devices
- The data collection process is often **non-transparent**, leaving participants unaware of the specific data being collected
- These tools can **covertly collect extensive data**, potentially violating data protection regulations
- Continued surveillance data collection can occur if the software is **not correctly uninstalled** post-study, compromising system security without the participants' awareness

IDEA: USE SCREEN RECORDINGS

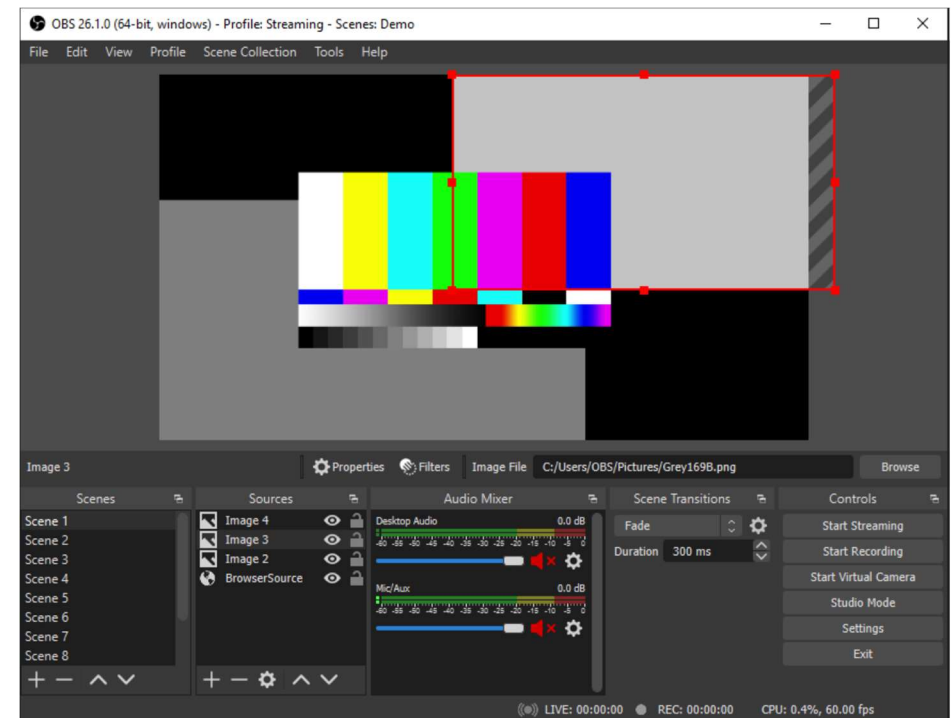
- We instruct participants to record their screen during the study
- Participants are asked to provide us with their screen recording after the study
- Participants maintain full control over their data, as they can choose whether or not to send us the video recordings



IDEA: USE SCREEN RECORDINGS

Computer Vision is employed to extract pertinent information and generate log files via

- Object detection
- Scene segmentation
- Text analysis



COMPUTER VISION

- Computer vision is an AI field that trains computers to **interpret and understand** the visual world, allowing them to extract and analyze information from images or videos.
- The objective of computer vision is to imitate human vision capabilities by **mimicking processes** like perception and understanding of visual material.
- Machine learning algorithms are used to **identify patterns in visual data** for various applications, including autonomous vehicles, medical image analysis, surveillance, security, or industrial quality inspection.

COMPUTER VISION

- **Object Detection** and Recognition identifies objects within an image or video feed and their locations
- **Image Classification** categorizes an entire image into a class or label
- **Semantic Segmentation** partitions an image into segments, each representing an object or a part of an object
- **Motion Analysis** or Activity Recognition understands the movements of objects between consecutive video frames
- **Facial Recognition** identifies or verifies a person from a digital image or a video frame, widely used in security systems and other areas like unlocking smartphones

IS THE FIELD MATURE ENOUGH TO BE USED BY NON EXPERTS?

- Bachelor's practicum in Informatics: A team of 4 - 5 students from the **third-semester** focuses on learning and experiencing the systematic execution of a software project.
- Members of the university can apply to participate in the **role of a client** who wants software to be developed in an agile way.
- The project was further pursued by a **student research assistant** from the Cognitive Science department.

THE SOLUTION: HOW IT WORKS

Optical Character Recognition with **Tesseract-OCR**:

- Tesseract-OCR was developed by HP in 1985
- It has been open-source since 2005 and is now developed by Google
- The program attempts to extract characters from an image

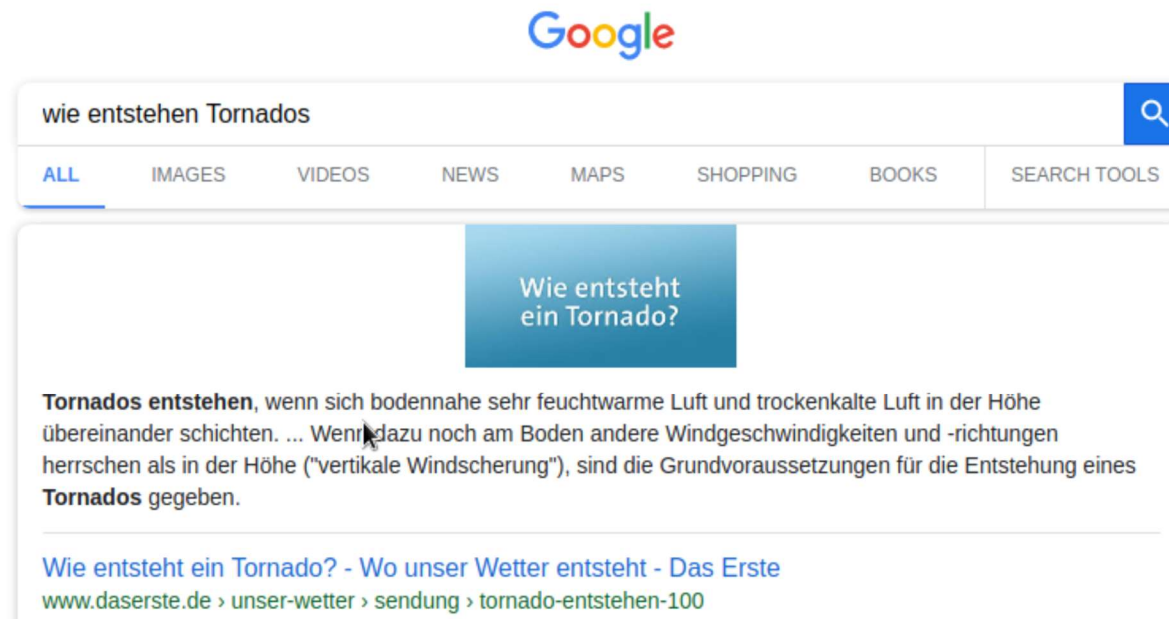


OpenCV and **EAST Network**:

- OpenCV is a free Computer Vision Library with bindings for languages such as Python and R
- It can process a video frame by frame and apply Computer Vision methods
- The EAST Network is a deep neural network specialized in text detection
- It identifies bounding boxes of text areas in an image

TEXT RECOGNITION IN VIDEOS

google.com/search?hl=en-DE&source=hp&biw=&bih=&q=wie+entstehen+Tornados&btnG=Google+Search&ifsig=ALs-wAMAAAAAYZy4trSUiLi52LsxivWms8ROGK...



OpenCV

EAST-Detector

Tesseract-OCR

TEXT RECOGNITION IN VIDEOS



TECHNISCHE
UNIVERSITÄT
DARMSTADT

https://www.google.com/search?hl=en-DE&source=hp&biw=&bih=&q=wie+entstehen+tornados%3F&btnG=Google+Search&iflsig=ALs-wAMAAAAAYZy_Fj8S-... K...



OpenCV

EAST-Detector

Tesseract-OCR

TEXT RECOGNITION IN VIDEOS

https://www.google.com/search?hl=en-DE&source=hp&biw=&bih=&q=wie+entstehen+tornados%3F&btnG=Google+Search&iflsig=ALs-wAMAAAAAYZy_Fj8S-...

[...] &q=wie+entstehen+tornados [...]

OpenCV

EAST-Detector

Tesseract-OCR

SAVED AS LOGFILES

[...] &q=wie+entstehen+tornados [...]

```
start_time, end_time, url, search_query
0.0, 21.0, "- 0 — p&biv=&bih=&q-wie+entstehen tornados?3F&btnG=Googles Search&iflsig=
ALs- WAMAAAAAYZy_Fiss- ... H)", "wie entstehen tornados"
21.0, 128.0, "de.wikipedia.org/wiki/Tornado",,
128.0, 143.0, "example.com/Umfrage",,
143.0, 201.0, "de.wikipedia.org/wiki/Tornado",,
201.0, 587.0, "example.com/Umfrage",,
```

OpenCV

EAST-Detector

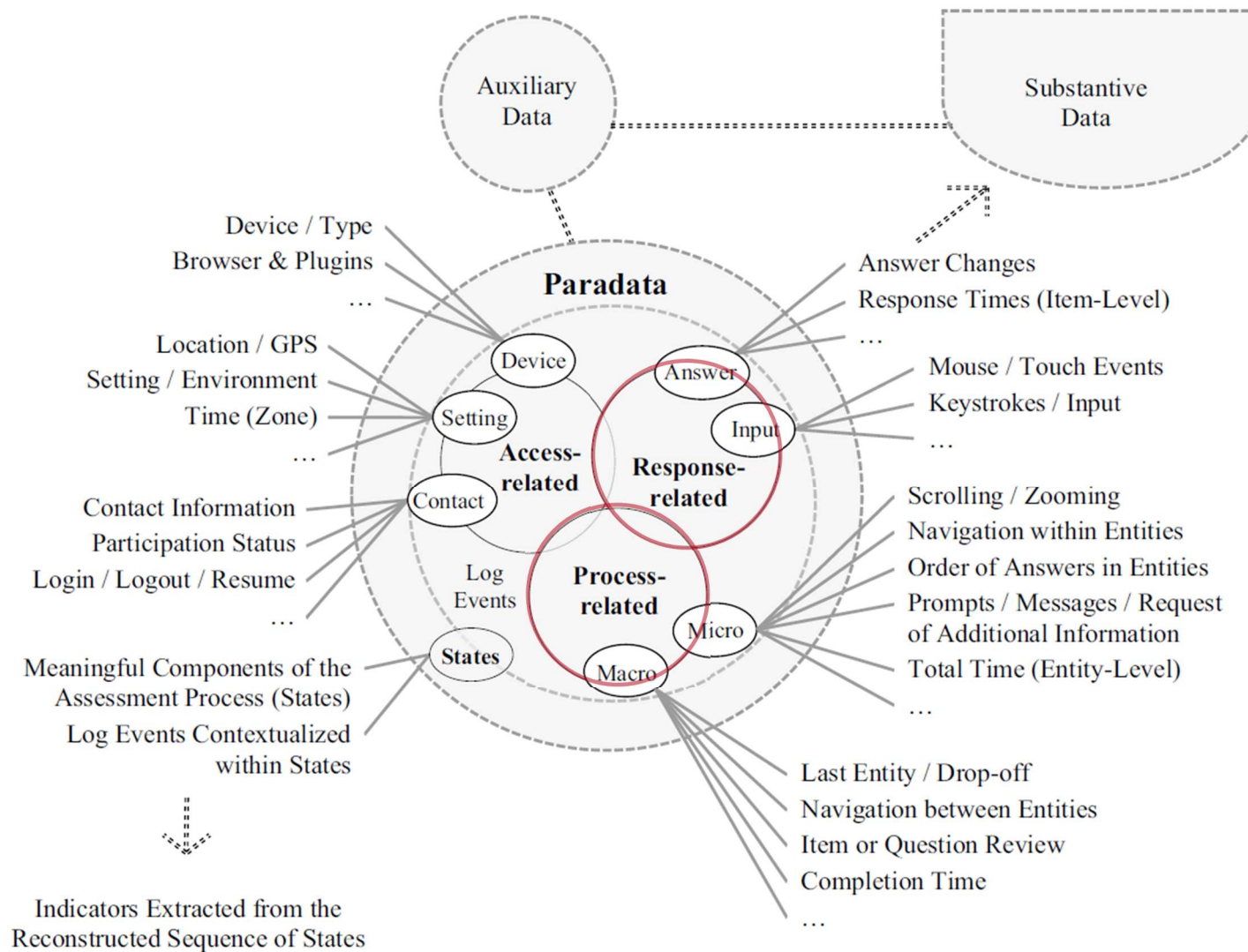
Tesseract-OCR

WHAT WERE WE ABLE TO EXTRACT?

- Search queries
- URLs
- Time on tabs & tab switches

```
start_time, end_time, url, search_query
0.0, 21.0, "- 0 — p&biv=&bih=&q-wie+entstehen tornados?3F&btnG=Googles Search&iflsig=
ALs- WAMAAAAAYZy_Fiss- ... H)", "wie entstehen tornados"
21.0, 128.0, "de.wikipedia.org/wiki/Tornado",,
128.0, 143.0, "example.com/Umfrage",,
143.0, 201.0, "de.wikipedia.org/wiki/Tornado",,
201.0, 587.0, "example.com/Umfrage",,
```

WHAT DO WE THINK IS POSSIBLE?



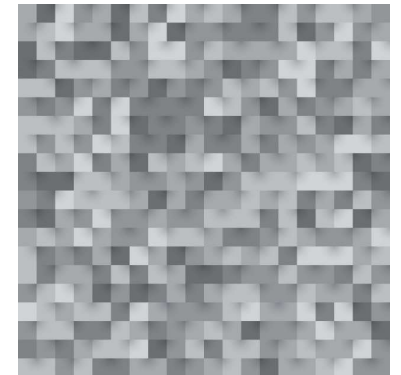
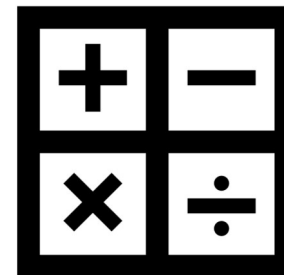
Kroehne, U., & Goldhammer, F. (2018). How to conceptualize, represent, and analyze log data from technology-based assessments? A generic framework and an application to questionnaire items. *Behaviormetrika*, 45(2), 527-563.

CAVEATS

- Frame by frame video analysis with text recognition is time consuming. Optimisation measures include
 - Optimisation of the code and libraries
 - Analysis of specific time windows
 - Parallelisation of analyses
 - Use of high performance clusters
- Text recognition is prone to errors
 - multiple frames can be taken into consideration (autoregression of frames)
- Large heterogeneity observed in screen recordings; approximately 9% of videos were unusable due to video codec errors or very poor quality

bih=&q=wie+entstehen+tornados%3F&

&q=wie+entstehen+tornados
&q-wie+entstehen lornados



OUTLOOK: GROUND TRUTH VALIDATION

- We begun to record participants' screens **in the lab** in conjunction with 'real' logging software to estimate the degree of accordance (agreement)
- A cooperative study is being planned with a task from a **large educational research institution**, where log files are genuinely collected on a dedicated server, serving as the ground truth

INTERIM DISCUSSION

- Screen recordings could served as a new method for capturing objective data in online studies
- A Computer Vision pipeline was developed to automatically extract relevant information from screen recordings
- Outlook:
 - General advancements in Computer Vision for speed and accuracy
 - Potential applications to more complex video information, such as scrolling behavior

BUT WAIT, WE NOW HAVE VIDEOS!

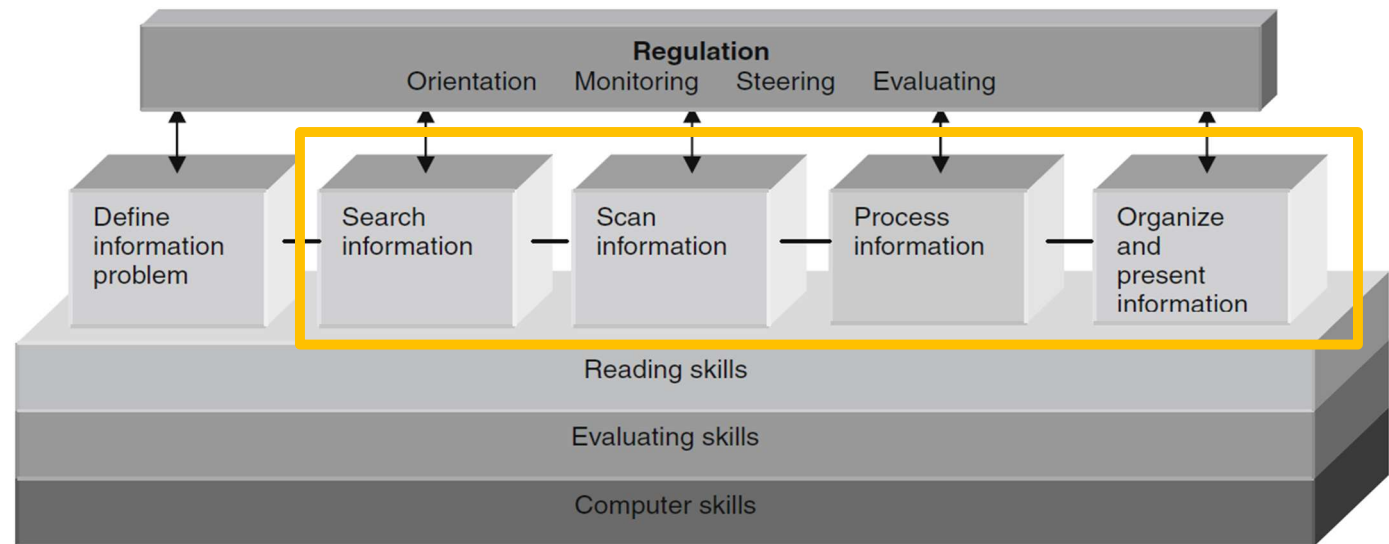
QUALITATIVE ANALYSIS

- In comparison the automatic video analysis, only **6% of videos** were non-codable for human evaluators due to defects, extremely poor quality, and artefacts of the recordings.
- **Non-adherence to experimental conditions** could be identified based on screen recordings:
 - Approximately 12% of participants were flagged for exclusion due to breaches of study instructions
 - Participants were found browsing the Internet during the No-Internet condition
 - Some participants initiated internet searches before submitting their predictive ratings

BUT WAIT, WE NOW HAVE VIDEOS!

QUALITATIVE ANALYSIS INTERNET SEARCH PROCESS

- Many models of Self-Regulated Learning (SRL) express the importance of exploring time and order effects in SRL processes (Panadero, 2017)
- The IPS-I Model (Brand-Gruwel et. al., 2009) defines distinct phases of the internet search process

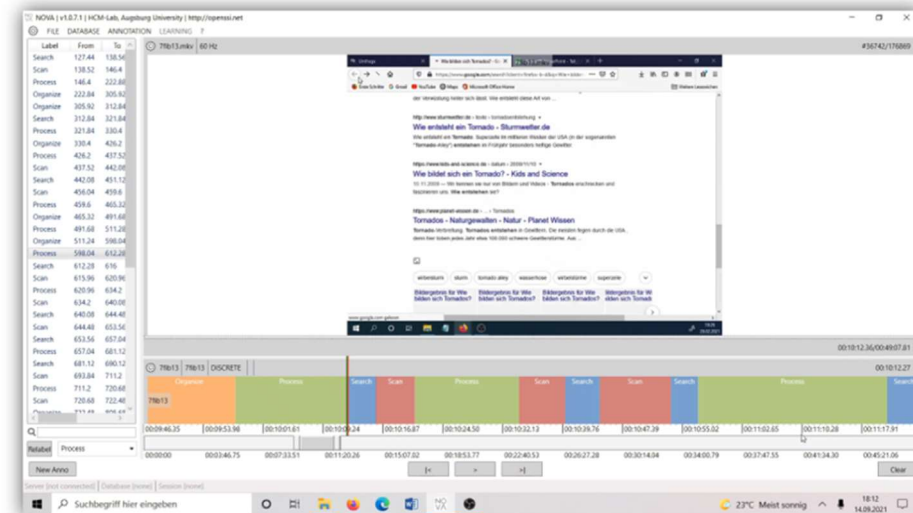


IPS-I-Model by Brand-Gruwel et. al. (2009)

BUT WAIT, WE NOW HAVE VIDEOS!

QUALITATIVE ANALYSIS INTERNET SEARCH PROCESS

- Two independent raters coded the start end times of the IPS-I phases
- 'Scan' and 'Process' are sometimes hard to distinguish (7% of all rated seconds are classified as a mismatch)
- We used the agreed upon starting and ending times of the phases only for further analysis

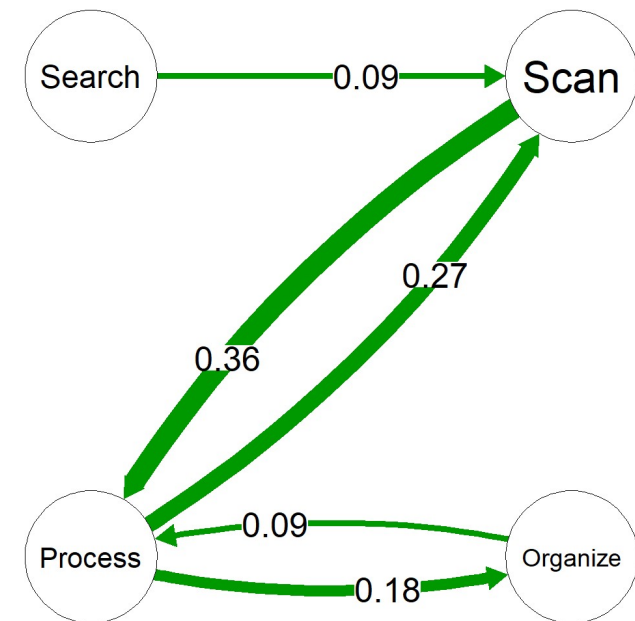


BUT WAIT, WE NOW HAVE VIDEOS!

QUALITATIVE ANALYSIS INTERNET SEARCH PROCESS

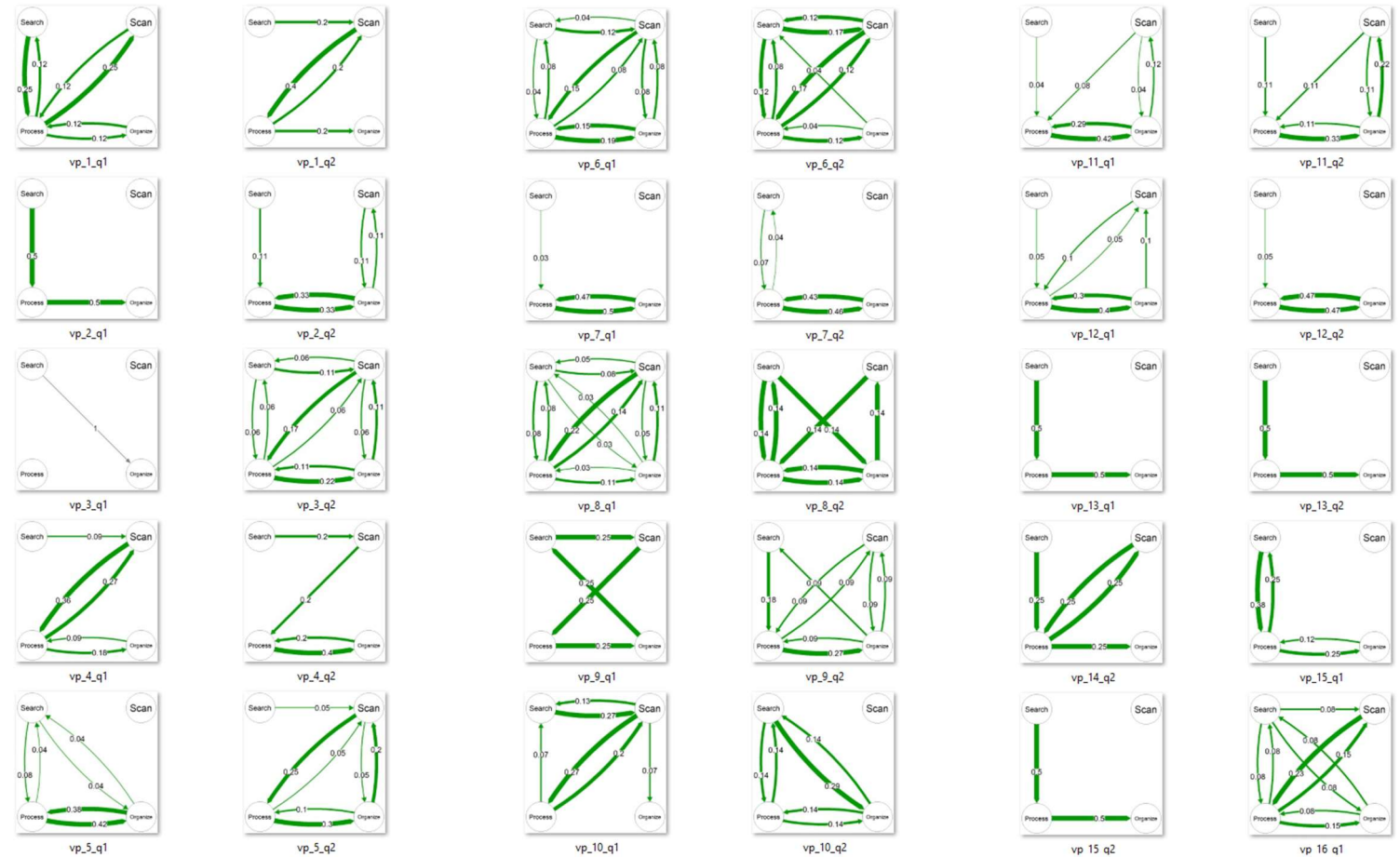
From the screen recordings we derived the following process parameters:

- Absolute and relative times in the phases of the IPS-I task
- Number of state transitions
- State transition probabilities



BUT WAIT, WE NOW HAVE VIDEOS!

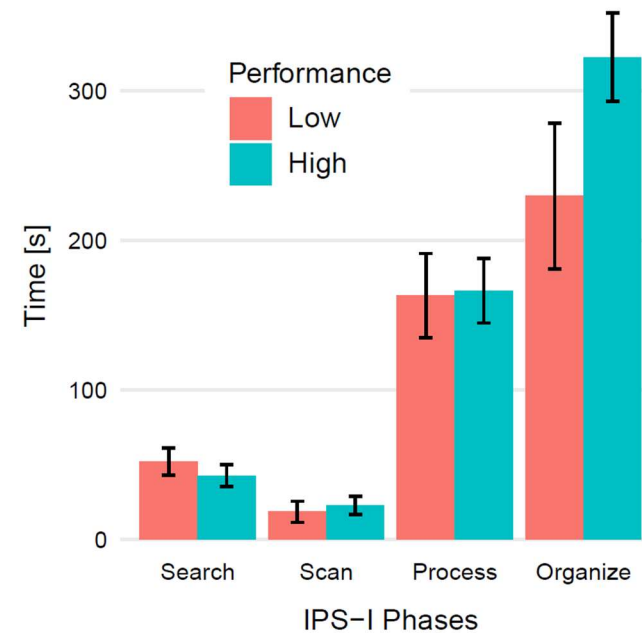
- High inter- and intraindividual differences in the IPS-I task
- Linear as well as non-linear processes



BUT WAIT, WE NOW HAVE VIDEOS!

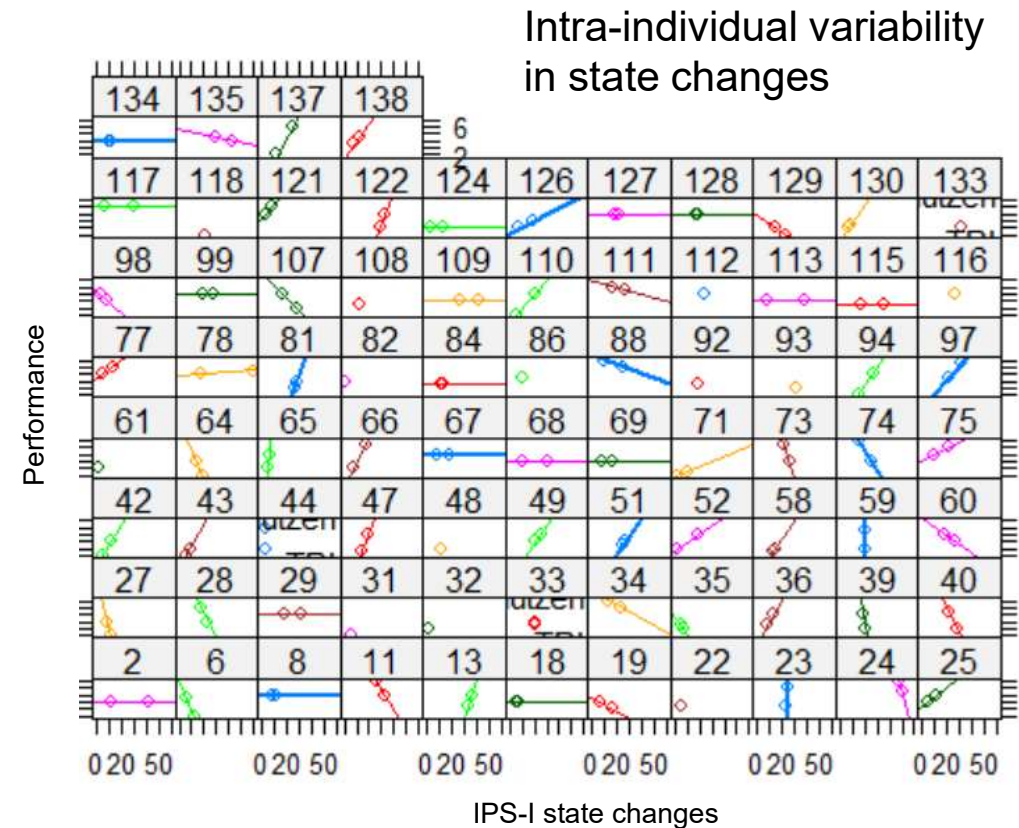
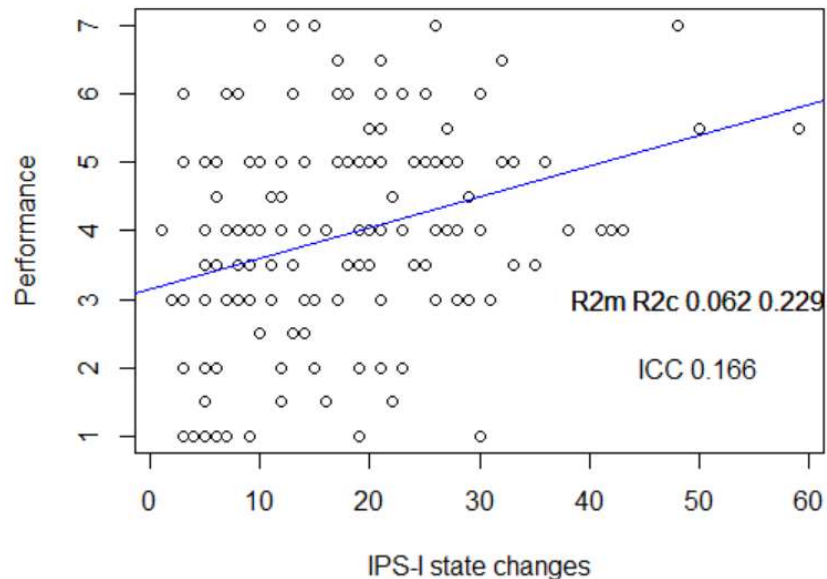
ABSOLUTE TIME IN PHASES PREDICTING PERFORMANCE

<i>Predictors</i>	Performance				<i>df</i>
	<i>Estimates</i>	<i>F-value</i>	<i>p</i>		
Search	-0.0010	-0.27	0.790		97
Scan	-0.0010	-0.17	0.867		97
Process	-0.0006	-0.43	0.667		97
Organize	0.0030	3.63	<0.001***		97
Random Effects					
σ^2	1.20				
$\tau_{00 \text{ id}}$	0.70				
ICC	0.37				
N_{id}	67				
Observations	104				
Marginal R^2 / Conditional R^2	0.123 / 0.447				



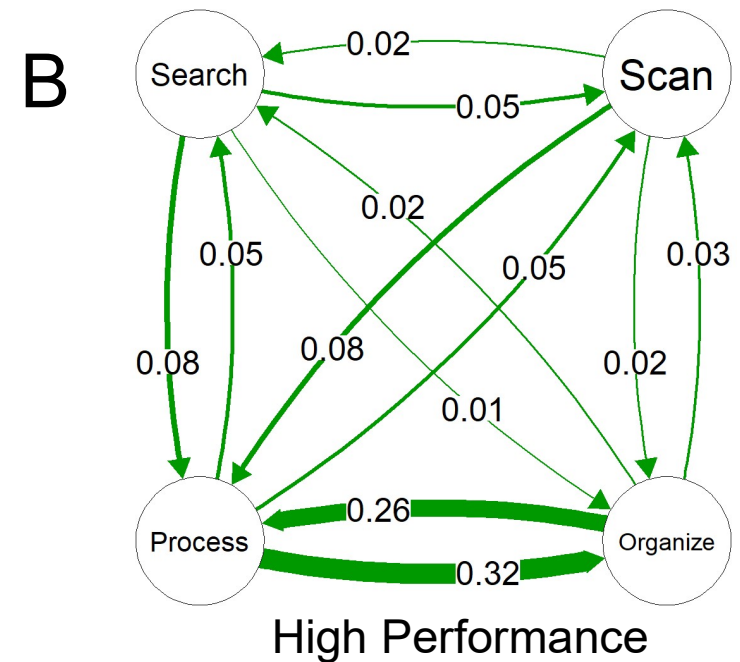
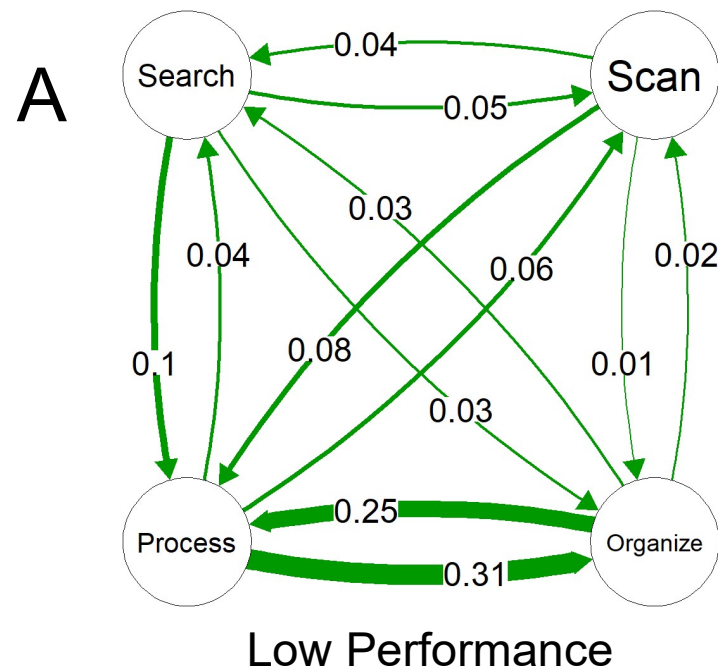
BUT WAIT, WE NOW HAVE VIDEOS!

STATE CHANGES PREDICTING PERFORMANCE



BUT WAIT, WE NOW HAVE VIDEOS!

TRANSITION PROBABILITIES ON PERFORMANCE



BUT WAIT, WE NOW HAVE VIDEOS!

FURTHER QUALITATIVE INSIGHTS

A note from a video rater:

Participant describes in the first five minutes that the question is not clearly defined (universe question). Towards the end, the participant copies text from Wikipedia into the answer field and starts organising. As she is running out of time, she deletes all of the text again and notes that the question cannot be answered clearly.

SUMMARY MANUAL VIDEO CODING

- Manual coding of screen recordings was **viable**
- Interesting **process parameters** could be derived
- **New ideas** for automatic video analysis could be generated:
 - Condition detection
 - Instruction violation



GENERAL SUMMARY

- This work focused on the possibility of log-file creation through screen recordings, utilising **Tesseract-OCR** and **OpenCV** for search process video analysis
- Screen recordings are a **data protection-compliant tool** for generating objective online study data and offer **new opportunities in empirical education**
- Our log-file generation approach will be **publicly available soon** under Apache License 2.0



DANKE

Björn Mattes

Pädagogische Psychologie

Institut für Psychologie

Fachbereich Humanwissenschaften

Technische Universität Darmstadt



DANKE

Björn Mattes

Pädagogische Psychologie

Institut für Psychologie

Fachbereich Humanwissenschaften

Technische Universität Darmstadt