

## Workshop

# Digital measurement of mental health using audio and video data



European  
Commission

Horizon 2020  
European Union funding  
for Research & Innovation



Fachgruppe Gesundheitspsychologie  
*Peer Mentoring Award 2024*

# Overview

## Workshop timeline (approximate)

**10:30** Introduction & theoretical part [L Puhlmann; L Meine]

**12:00** Discussion

**~12:30** Preparation for hands-on part (as needed)

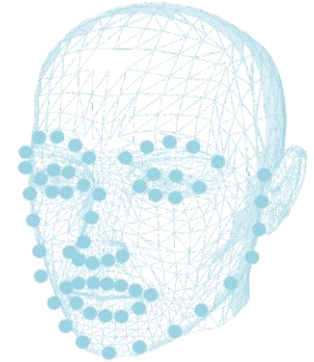
**[12:30/45** Lunch]

**13:30** hands-on part [M Worthington; L Meine; R Pardhi]

**[16:15** Coffee break]

**~16:30** Wrap-up & further discussion

**17:30** Workshop ends.



# Digital measurement of mental health using audio and video data - what, why and how?

**Lara Puhlmann**

Leibniz Institute for Resilience Research, Mainz, Germany



European  
Commission

Horizon 2020  
European Union funding  
for Research & Innovation



Fachgruppe Gesundheitspsychologie  
Peer Mentoring Award 2024

# Digital Biomarkers of mental health workshop

## Introduction

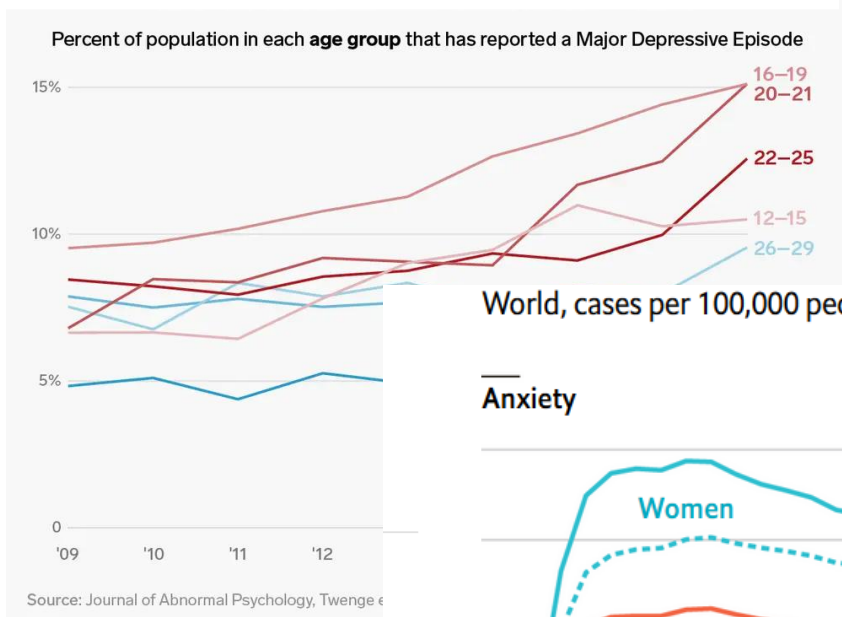
- What: Background & Definitions
- Why digital measurement of audio and video data?
- How: Introduction to Study design: interview-derived Biomarkers

## State of the evidence

## Two study examples

# Need for innovative healthcare solutions

## Depression rates by age, 2009-2017

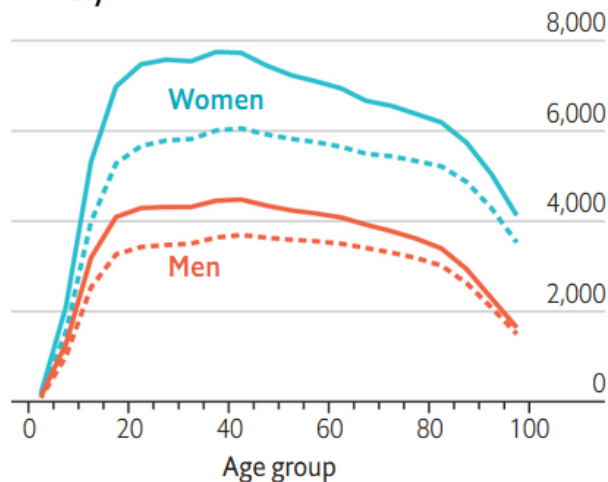


<https://www.businessinsider.com/depression>

## Depression and Anxiety Cases pre- and during pandemic

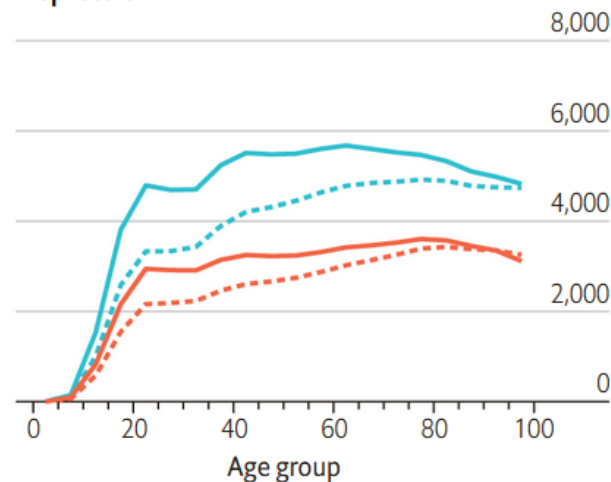
World, cases per 100,000 people

### Anxiety



-- Before covid-19 pandemic — During pandemic

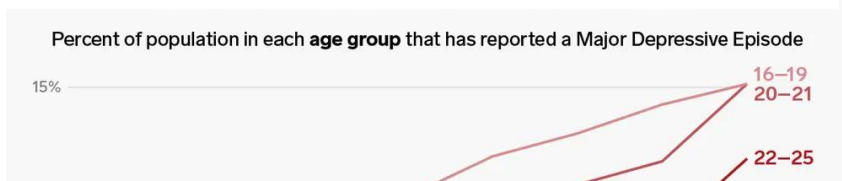
### Depression



Source: "Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the covid-19 pandemic", by D.F. Santomauro et al., *The Lancet*, 2021

# Need for innovative healthcare solutions

## Depression rates by age, 2009-2017

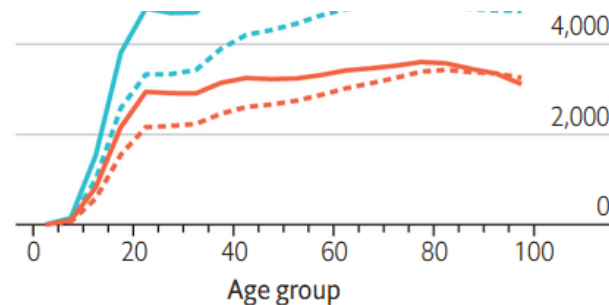
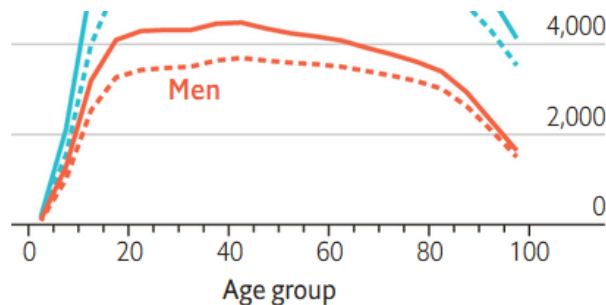


## Need for:

- Refined diagnosis
- Early detection
- Dense monitoring
- Individualised treatment

Source: Journal of Abnormal Psychology, Twenge et al.

<https://www.businessinsider.com/depression>



Source: "Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the covid-19 pandemic", by D.F. Santomauro et al., *The Lancet*, 2021

## New markers for objective symptom assessment

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

**Clinical Presentation**

**Neurotransmitter occupancy**  
**Circuit activity**  
**Protein expression**  
**Clinical phenotype**

### **Self-report (subjective, biased)**

[retrospective recall biases (eg, to downplay or overestimate symptoms; state-dependency), cognitive limitations (eg, memory of episodes and environment, causal inference), social stigma]

# Biomarkers of disorders

Article | Published: 16 March 2022

## Reproducible brain-wide association studies require thousands of individuals

Scott Marek , Bre

*Nature* 603, 654–6

21k Accesses | 1

### Performan phenotype

Marc-Andre S  
John-Dylan H

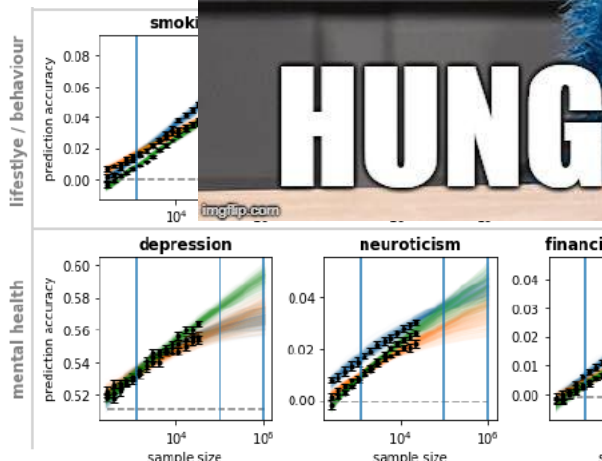
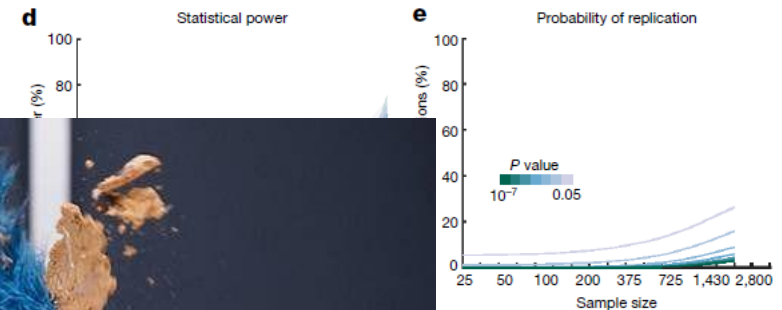





Figure 2: Linear models are operating far below ceiling accuracy for most target phenotype predictions.



the HPA axis  
een  
response

Boushra Dalile <sup>a, b</sup> , Danique La Torre <sup>a, b</sup>, Kristin Verbeke <sup>a, 1</sup>, Lukas Van Oudenhove <sup>a, b, c, 1</sup>, Bram Vervliet <sup>d, 1</sup>

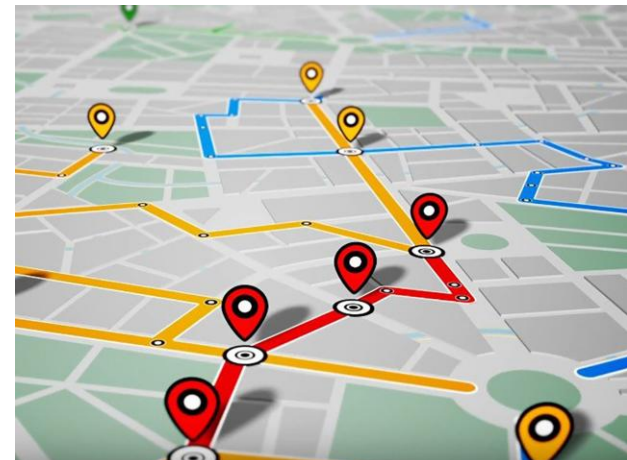
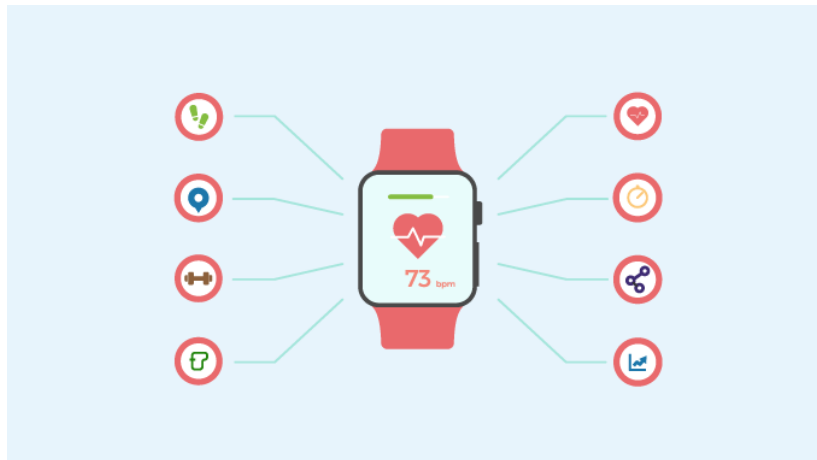
Show more 

+ Add to Mendeley  Share  Cite



# Digital Phenotyping for mental health measurement

Digital phenotyping: *The “frequent and constant measurement of human phenotypes in situ based on data from smartphones and other personal digital devices”*



Onella & Rauch, 2016, *Neuropsychopharmacology*

# Digital biomarkers (DBMs)

## Biomarker definition:

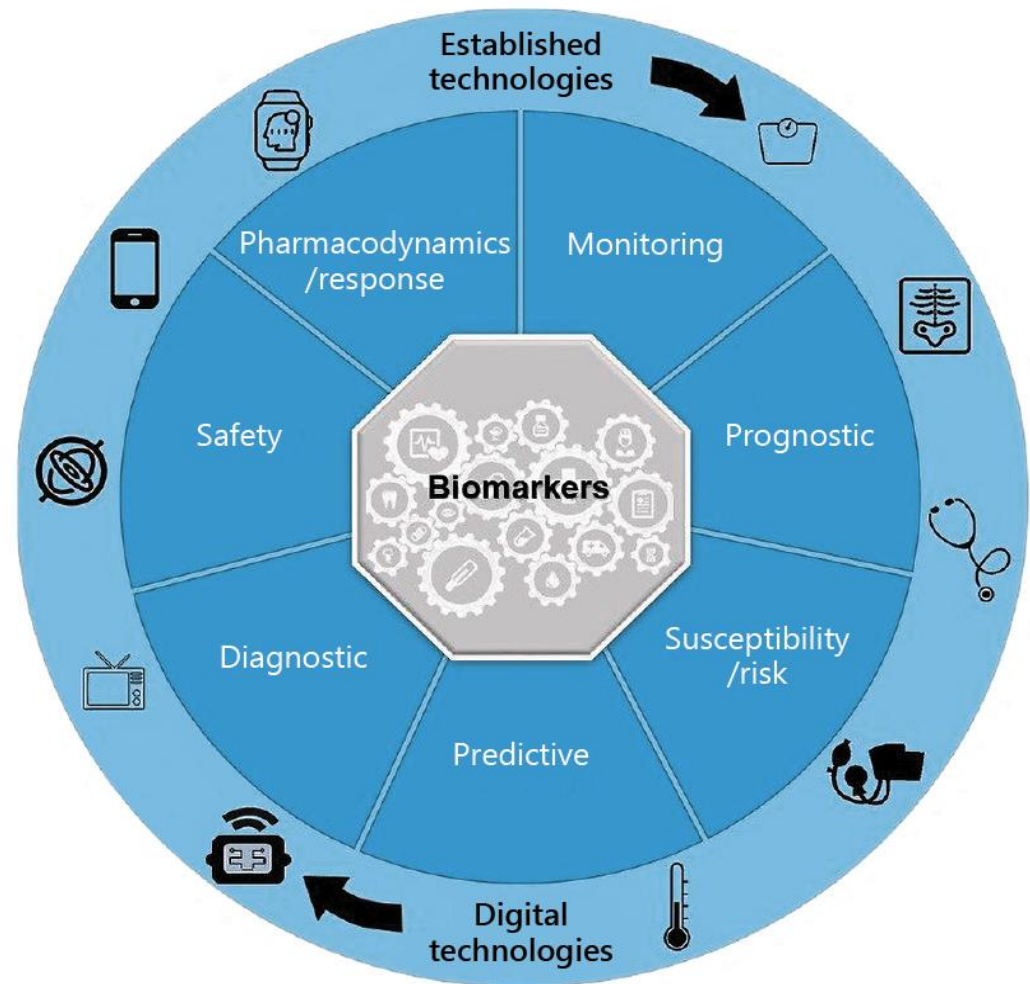
An objectively measurable characteristics that serve as indicators for biological or pathogenic processes, or response to therapeutic intervention

## Digital biomarker:

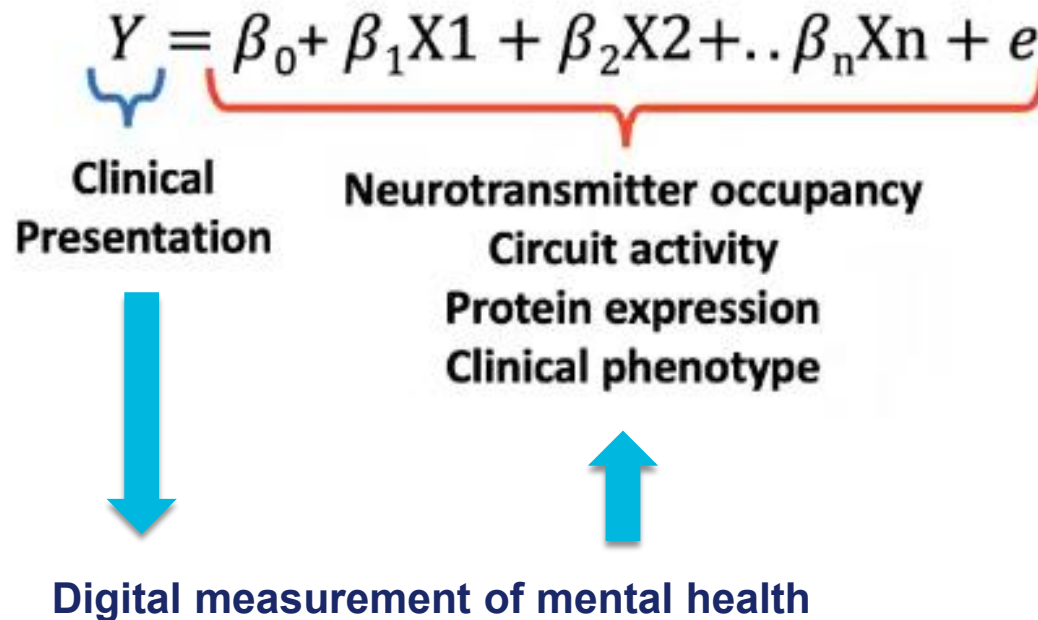
Same principle, but based on observable data (e.g. behavior, language) that is *digitally* measured

## Advantages:

- Objective (automatic)
- Sensitive (rich data)
- Scalable



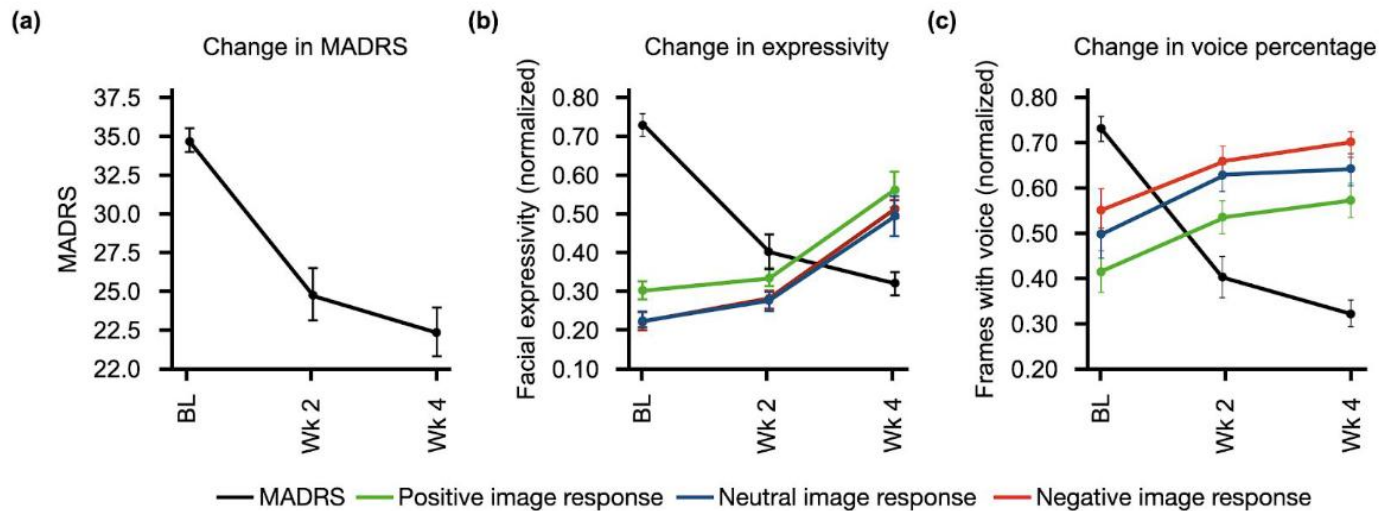
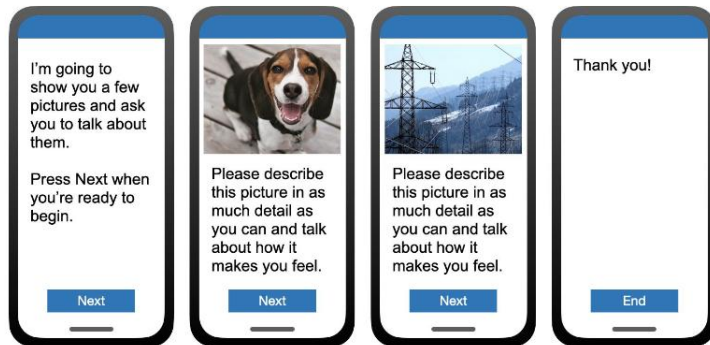
# New Markers for objective symptom assessment



# Study design with interview-based DBMs

- Currently predominantly used for clinical populations / clinical interviews; conceptually much wider applicability
- Facial and vocal expressivity broadly fall into three categories:
  - **Spontaneous behaviour**
  - **Evoked behaviour** (e.g., emotional response to stimuli)
  - **Behavior on cue** (e.g. ‘make a happy face’; ‘make a sustained vowel sound’)
- Relevance may differ by patient populations, but also algorithms (e.g., Low et al., 2020)

# Example studies: evoked responses



Abbas et al., 2021, *Frontiers in digital health*

# Interview-based digital biomarkers

## Interview settings...

- are a longstanding, validated approach
- build on clinical concepts of nonverbal behaviour in psychopathology (e.g., Kring & Stuart, 2008)
- can be used for screening and diagnosis
- provide rich subjective information
- ensure that data is provided consensually; accessibility can be standardized

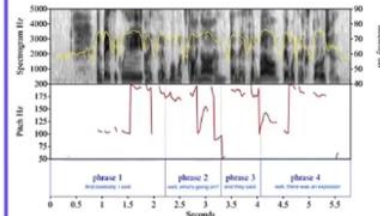


# Interview-based digital markers: Key modalities

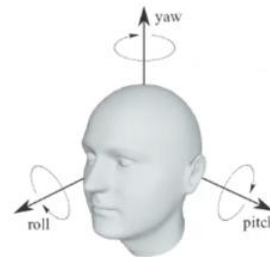
## Facial Expressivity Using Facial Action Coding System



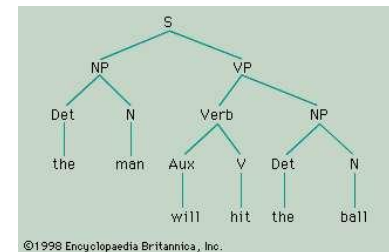
## Voice & Speech



## Movement



## Content analysis

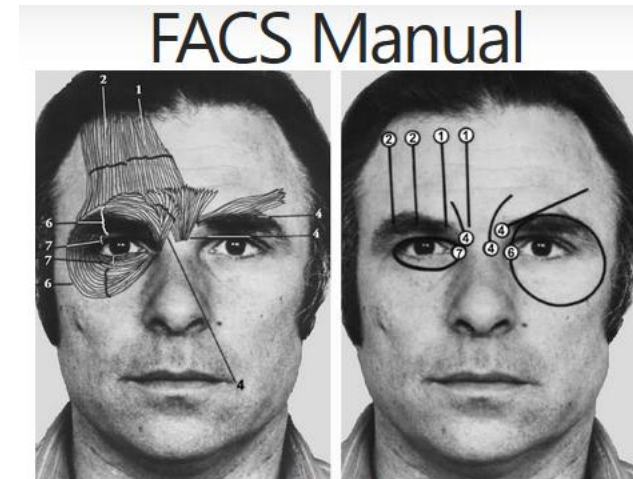




# Facial Action Unit Coding System (FACS)

Paul Ekman, Wallace V. Friesen, 1978

Upper Face Action Units					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
					
Inner Brow Raiser	Outer Brow Raiser	Brow Lowerer	Upper Lid Raiser	Cheek Raiser	Lid Tightener
*AU 41	*AU 42	*AU 43	AU 44	AU 45	AU 46
					
Lid Droop	Slit	Eyes Closed	Squint	Blink	Wink
Lower Face Action Units					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
					
Nose Wrinkler	Upper Lip Raiser	Nasolabial Deepener	Lip Corner Puller	Cheek Puffer	Dimpler
AU 15	AU 16	AU 17	AU 18	AU 20	AU 22
					
Lip Corner Depressor	Lower Lip Depressor	Chin Raiser	Lip Pucker	Lip Stretcher	Lip Funneler
AU 23	AU 24	*AU 25	*AU 26	*AU 27	AU 28
					
Lip Tightener	Lip Pressor	Lips Part	Jaw Drop	Mouth Stretch	Lip Suck



\$350.00

## Description

### PRODUCT INCLUDES

FACS Manual - 527-page PDF

Investigator's Guide - 197-page PDF

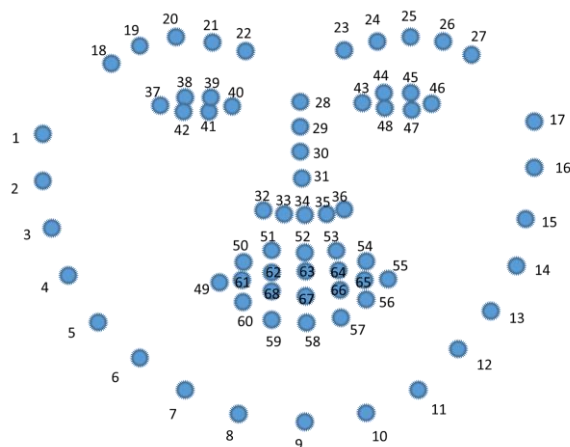
Score-Checker App

Example photos and videos



# Interview-based digital markers: Computation

## 1. Capture facial landmarks + displacement



## 2. Extrapolate underlying 'Action Units'

Action unit number	Description
AU1	Inner brow raiser
AU2	Outer brow raiser
AU4	Brow lowerer
AU5	Upper lid raiser
AU6	Cheek raiser
AU7	Lid tightener
AU9	Nose wrinkler
AU12	Lip corner puller
AU15	Lip corner depressor
AU16	Lower lip depressor
AU20	Lip stretcher
AU23	Lip tightener
AU26	Jaw drop

## 3. Derive emotion

Emotion	EMO	Action Units
Happiness	hap	6 + 12
Sadness	sad	1 + 4 + 15
Surprise	sur	1 + 2 + 5 + 26
Fear	fea	1 + 2 + 4 + 5 + 7 + 20 + 26
Anger	ang	4 + 5 + 7 + 23
Disgust	dig	9 + 15 + 16
Contempt	con	12 + 14




# State of the evidence

# State of the evidence

Mainly cross-sectional studies comparing patient groups to healthy controls



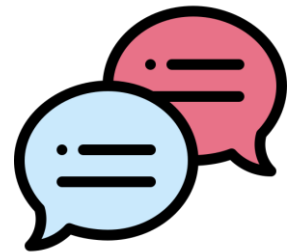
Findings show marked differences between groups in ...

- speech characteristics 
- vocal features 
- facial and emotional expressivity 

icons from flaticon.com

# Studies on clinical populations vs. controls report ...

- Patients with schizophrenia, depression, and healthy controls can be classified based on **articulation, pause rate, and variability in speech** (Berardi et al., 2023)
- Greater use of **first-person pronouns** is associated with depression, anxiety, and PTSD (Edwards & Holzman, 2017; Zimmermann et al., 2017; Teferra & Rose, 2023; Kleim et al., 2018)
- Low **specificity** of autobiographical memories is a core feature of multiple disorders (Mistica et al., 2024)



icon from flaticon.com

# Studies on clinical populations vs. controls report ...

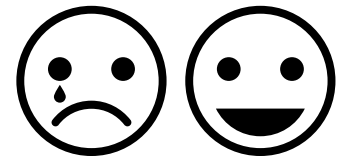
- Monotonous **pitch**, lower **speech prevalence**, lower **loudness**, and lower **harmonics-to-noise ratio** is associated with depression (Albuquerque et al., 2021; Koops et al., 2023; Alpert et al., 2001; Min et al., 2023; Wang et al., 2019)
- Later **speech onset** is related to both depression and PTSD (DeVault et al., 2013)
- Individuals with symptoms of PTSD tend to have a more **tense voice** (Scherer et al., 2014)
- In anxiety, **jitter** appears to be increased (Fuller et al., 1992)



icon from flaticon.com

# Studies on clinical populations vs. controls report ...

- Reduced overall **expressivity**, less expression of **happiness** and more expression of **sadness**, and reduced **eye contact** in patients with depression (Gaebel & Wölwer, 2004; Renneberg et al., 2005; Trémeau et al., 2005; Min et al., 2023; Rottenberg et al., 2005; Philippot et al., 2003)
- Attenuated expressions of **fear**, **surprise**, and happiness in patients with anxiety (Benecke & Krause, 2005; Buchheim & Benecke, 2007)
- More **startle responses**, a greater tendency to display **anger**, shorter and less intense **smiles**, and lower **gaze** in posttraumatic stress disorder (Kirsch & Brunnhuber, 2007; Scherer et al., 2014)



# State of the evidence

Studies comparing acute stress to a control condition in healthy participants



Findings show significant differences between conditions in ...

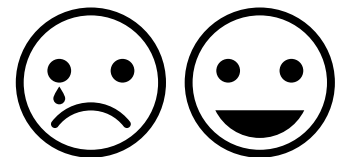
- vocal features (⌈{
- facial and emotional expressivity 😞 😊

However, there are not very many studies yet

# Studies on acute stress vs. control report ...



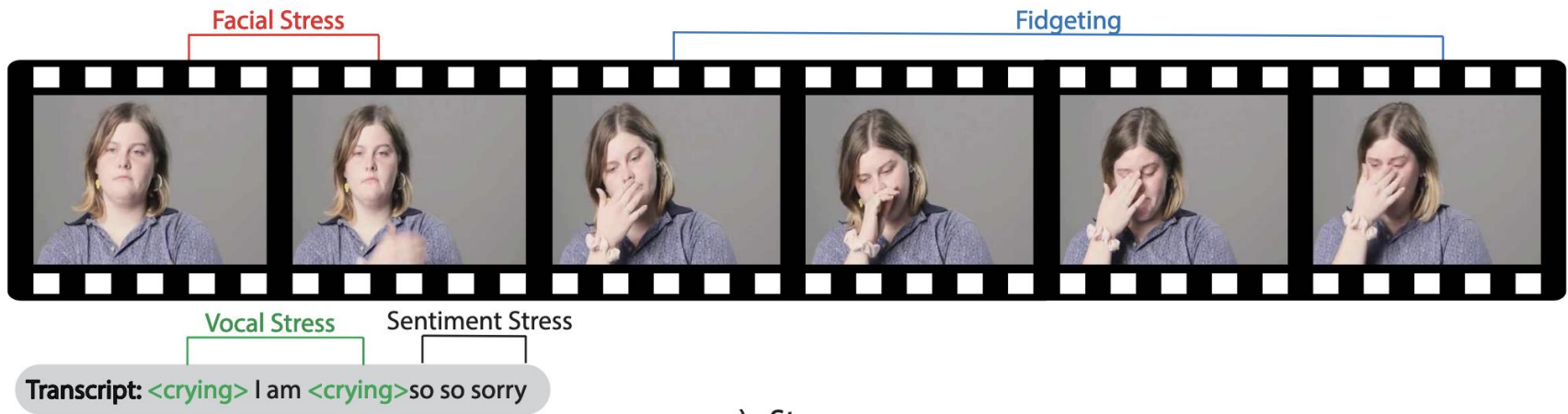
- Increased **pitch** in stress compared to control (Giannakakis et al., 2022)
- Differences in **shimmer, harmonicity** (Tavi, 2017)
- Louder mean **speech volume** and higher pitch is associated with higher cortisol levels (Baird et al., 2021; Pisanski et al., 2016)
- **Smiling** acts as a buffer of stress linked to lower cortisol levels, whereas expressions of **anger, fear**, and **disgust** are linked to higher cortisol levels (Blasberg et al., 2023)
- Less positive facial expressions, reduced **head movements**, and a more static **gaze** were observed during stress compared to a non-stressful control condition (Müller, 2024)





# Acute stress vs. control

Ghose et al., 2025



a) Stress



b) No Stress

# State of the evidence

**Summary:** Specific features have robustly been associated with stress and stress-related disorders

However, more research is needed on ...

- Subclinical populations who are stressed (e.g., first-responders)
- Prospective associations of features with later mental health
- Comparison of different disorders (not just patients vs. controls)

→ Next steps for prevention and precision medicine

# DMMH for resilience research

# A Preventive Approach to Mental Health Disorders

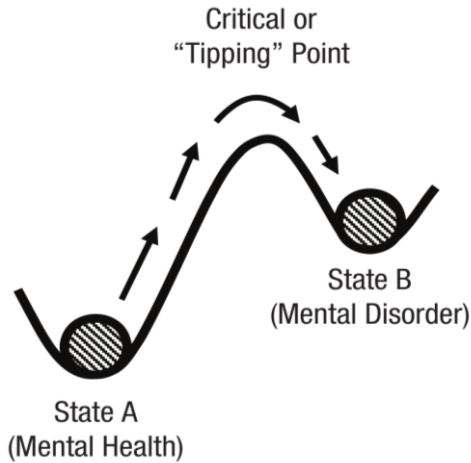
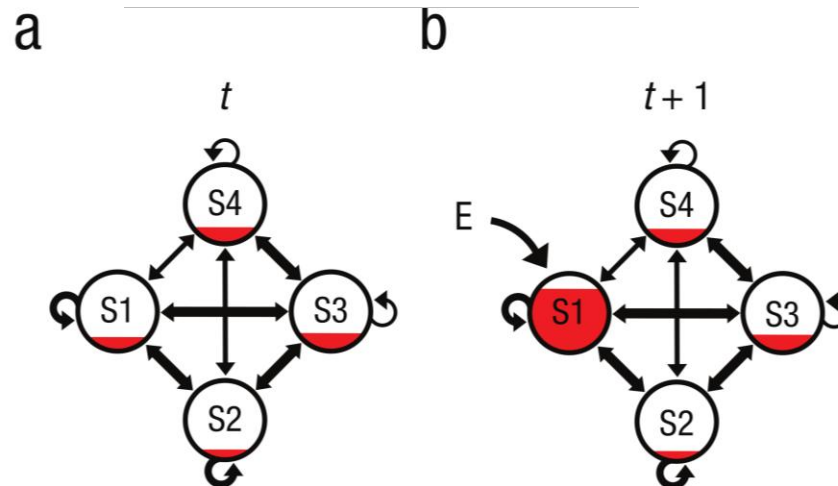


Fig 1. A diathesis-stress model of mental health network-state transitions

Fig 2. Resilience Factors (RFs) as nodes in dynamic symptom networks.



Stressor exposure (E)  
activates symptom  
nodes (S)

Borsboom, 2017, *World Psychiatry*  
Kalisch et al., 2019, *Persp. Psych. Science*

# A Preventive Approach to Mental Health Disorders

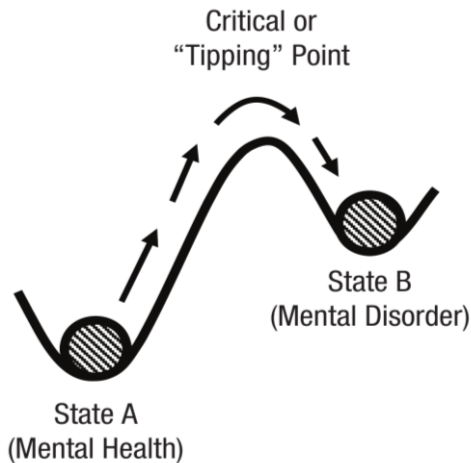
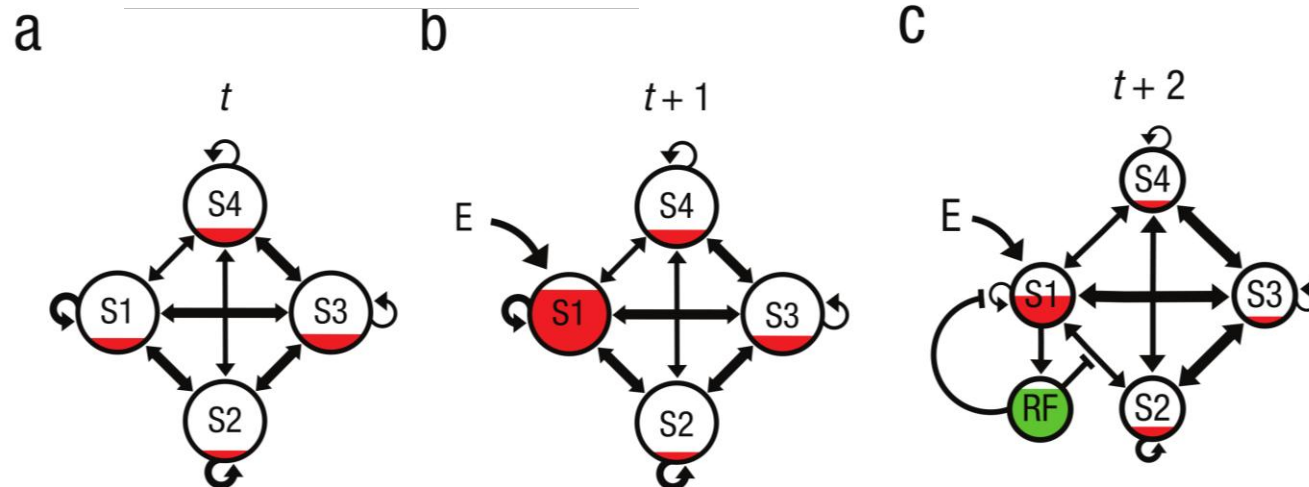


Fig 1. A diathesis-stress model of mental health network-state transitions

Fig 2. Resilience Factors (RFs) as nodes in dynamic symptom networks.



Stressor exposure (E) activates symptom nodes (S), which are downregulated by a Resilience factor (RF)

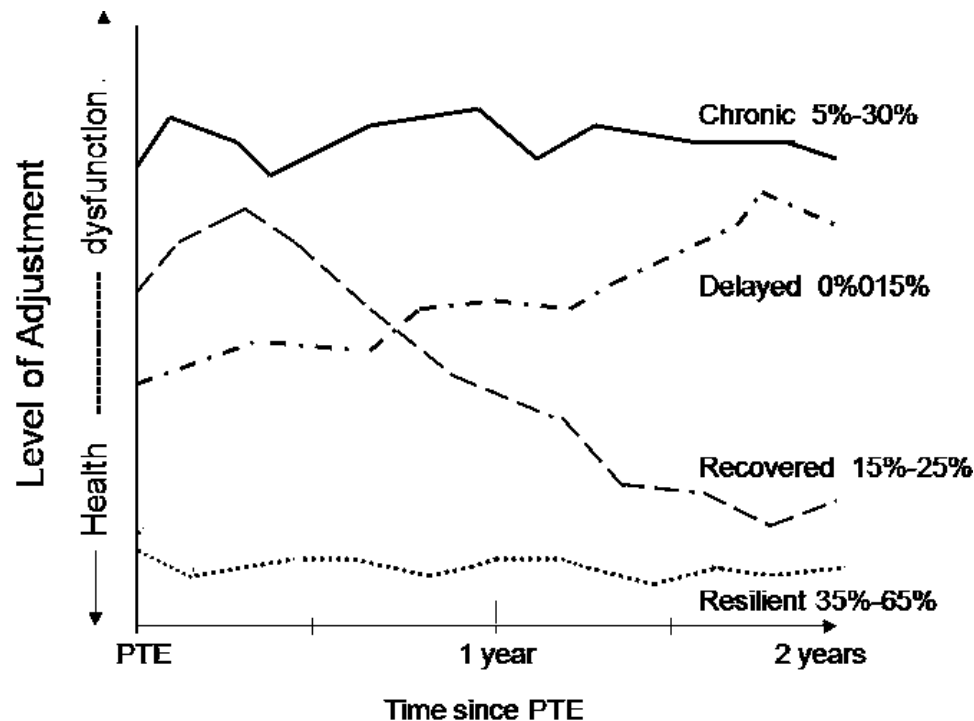
Borsboom, 2017, *World Psychiatry*  
Kalisch et al., 2019, *Persp. Psych. Science*

# Resilience Research

*Resilience is “the maintenance or quick recovery of mental health during and after periods of adversity”*



Kalisch et al., 2017, *Nat Hum Behav*

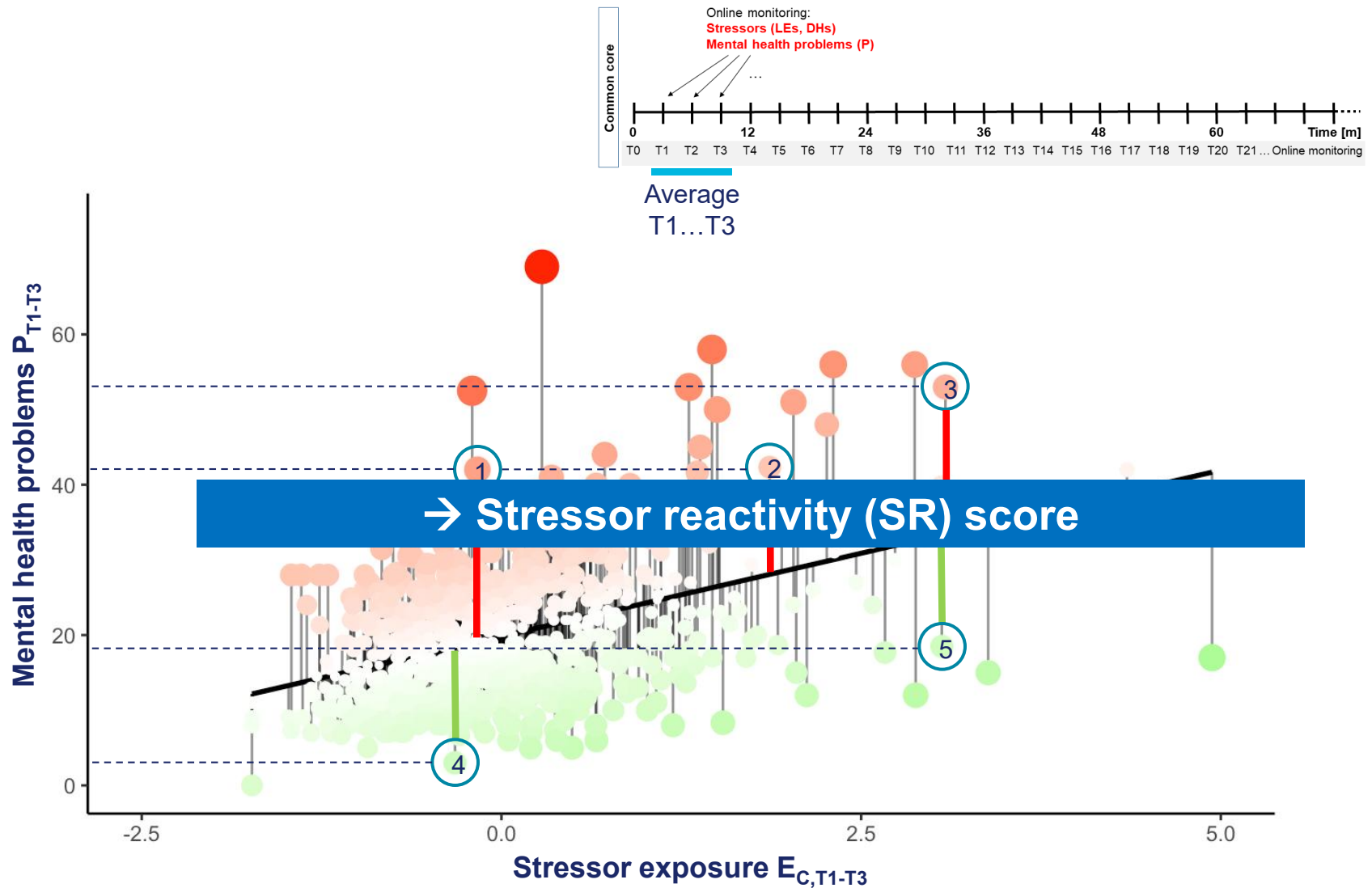


Prototypical trajectories of adjustment following a potentially traumatic event

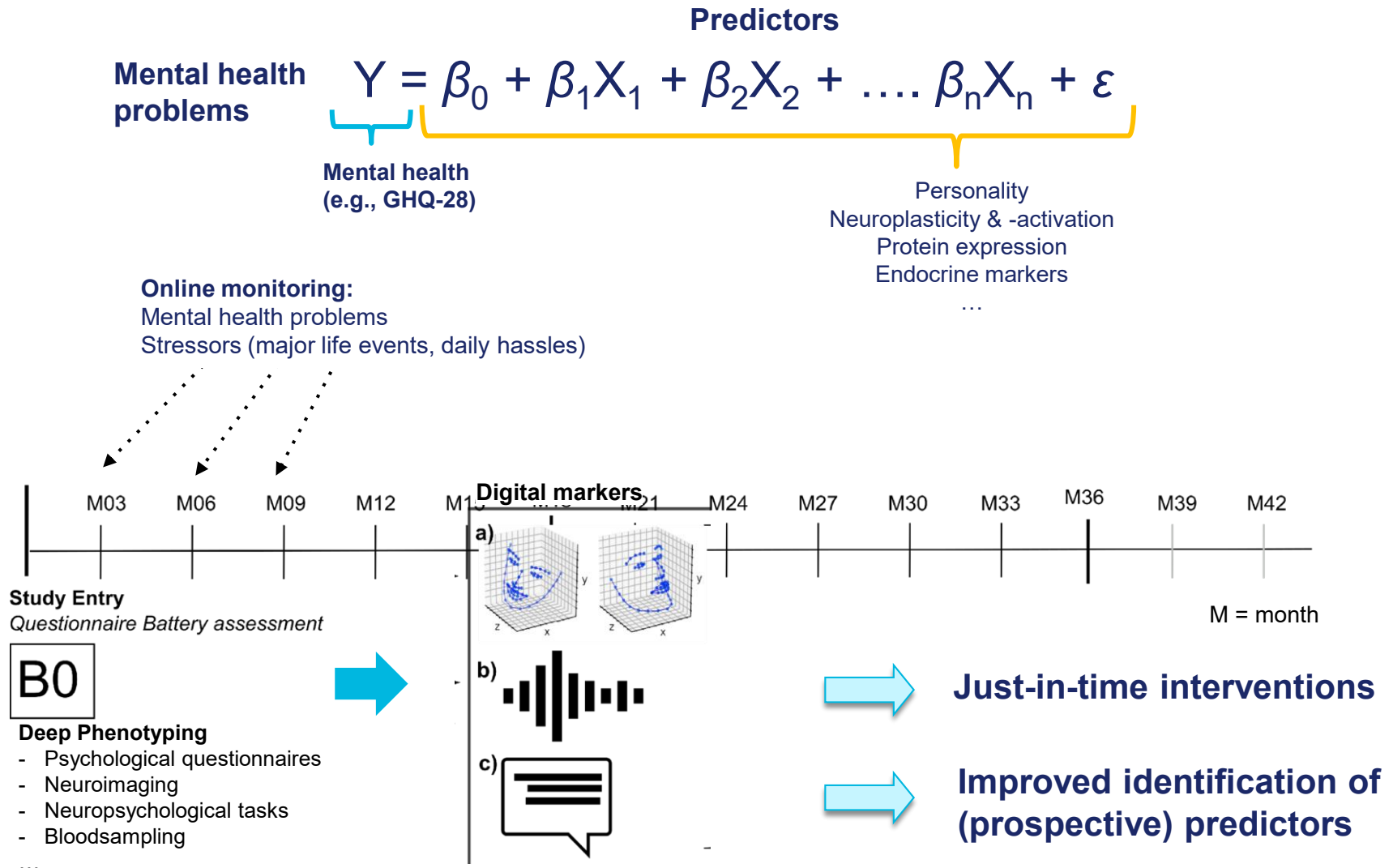
Bonanno, Westphal, Mancini, 2010  
*Annual Review of Clinical Psychology*

# Resilience – outcome-based operationalization

*FRESHMO paradigm: frequent stressor and mental health monitoring*



# Identifying predictors of mental health





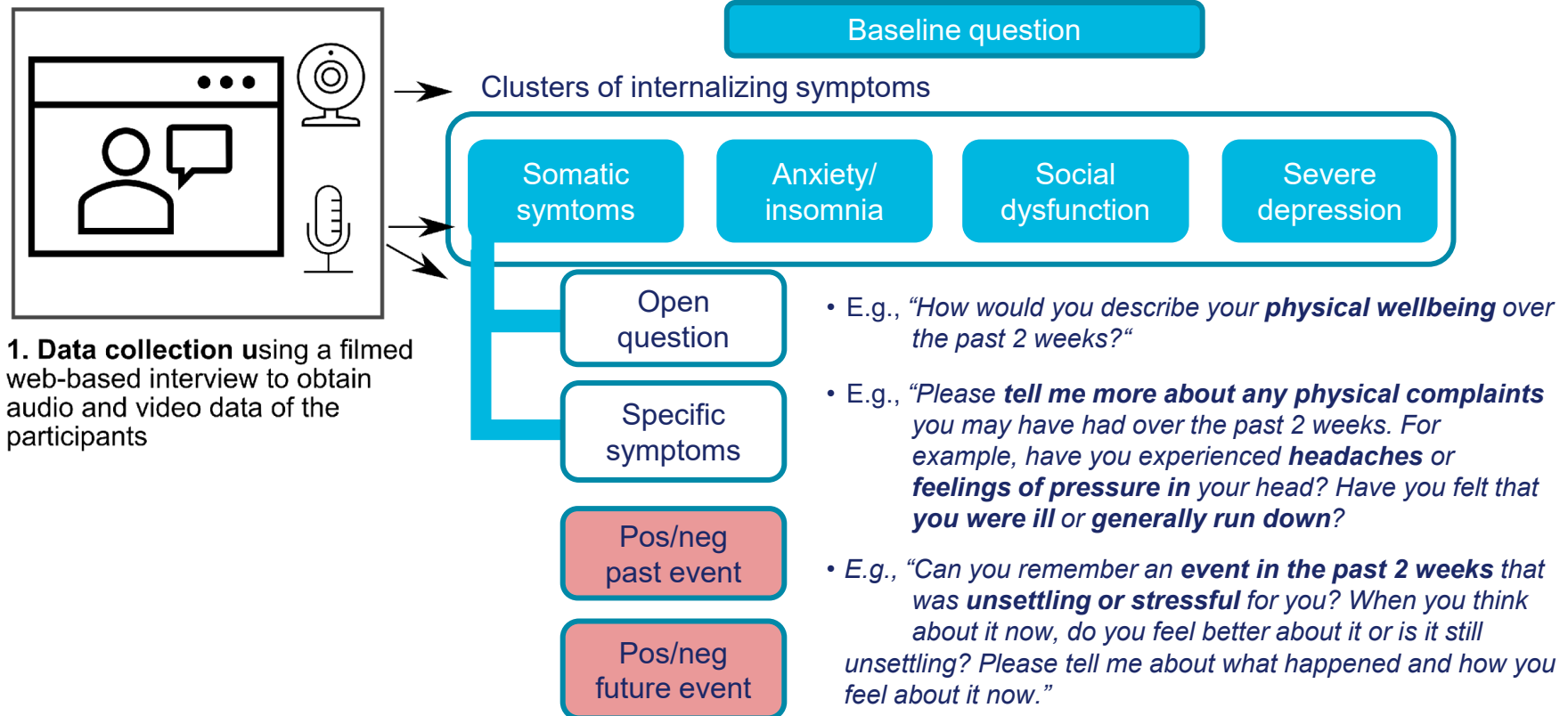
# Questions



# Example Studies

# Study Design (DynaM-INT)

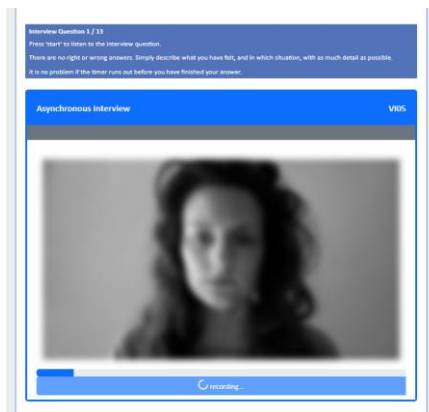
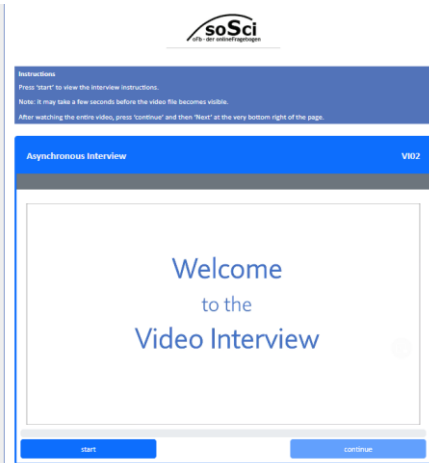
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \beta_n X_n + \varepsilon$$



Bögemann\*, Riepenhausen\*, Puhlmann\*  
et al., 2023, *BMC Psychol*

# Data collection in DynaM-INT

Hanno Burger



Baseline characterization phase					Ecological momentary intervention phase																				Follow-up phase											
Month 1 Week					Month 2 Week				Month 3 Week				Month 4 Week				Month 5 Week				Month 6 Week				Month 7 Week				Month 8 Week							
1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4					
✓																																				
👤																																				
MRI																																				
🩸																																				
💩																								🩸												
📄																								📄												
📱			📱		📱																			📱												
🏠											🏠													🏠												
📱																																				
📱																																				
📱																																				
📱																																				
€																																				

Legend		
✓	M.I.N.I. interview and urine drug test	📱 Video interview
📱	Neuropsychological tests	📱 Calibration week: EMA (10/day) + EPA (16h/day)
MRI	Neuroimaging battery (incl. training)	📱 Training week: EMA (3/day) + EMI (3/day)
🔴	Blood sample	📱 Booster week: EMA (10/day) + EPA (16h/day) + JITAI EMI (max. 4/day)
📄	Stool sample	📱 Practice week: EMA + EMI (self-triggered)
📄	Online extended questionnaire battery	€ Reimbursement
📱	Online monitoring questionnaire	📱 Lottery (up to 5 vouchers of €100 per site)

Bögemann\*, Riepenhausen\*, Puhmann\*  
et al., 2023, *BMC Psychol*

Lara Puhmann  
Leibniz Institute for Resilience Research (LIR)

# HMZ STRESS dataset



$N = 105$  Swiss medical students undergoing stressful internships

Prior to internship start, we recorded videos of the students talking about ...

- Neutral, negative, and positive experiences
- Negative and positive future events

(adapted from Galatzer-Levy et al., 2021)

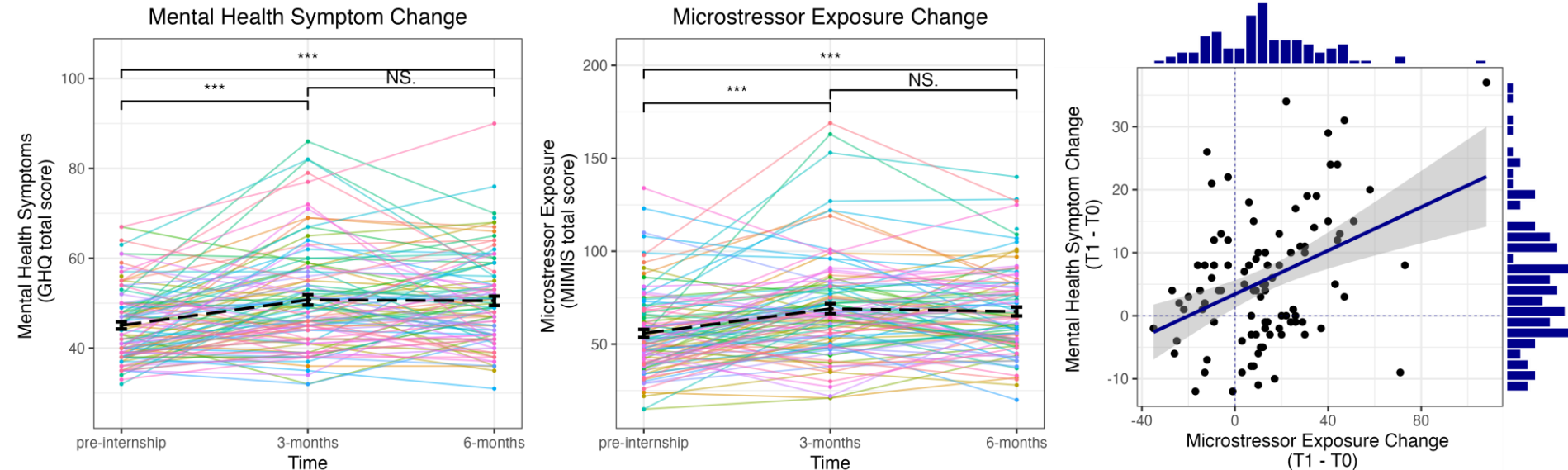
We are currently extracting speech characteristics, vocal features, and facial and emotional expressivity as predictors of stress resilience at 3 and 6 months following internship start.



# HMZ STRESS dataset

Resilience is operationalised as the residual resulting from regressing change in mental health on stress exposure to obtain one continuous outcome.

Considerable differences in change in mental health and stress may be (partly) explained by video- and audio-derived features assessed at baseline.



For example, we expect that students who are more emotionally expressive, prove more resilient.

# HMZ STRESS dataset

## Our **research questions**:

- Which features are associated with stress resilience?
- Which modality shows the best predictive performance?
- Which interview topic is most relevant? (e.g., negative vs. positive)



## **Dataset specifics:**

- Video was recorded in the lab, under carefully controlled conditions
- But we used a standard camera
- Participants spoke in Swiss German to ensure authenticity
- Video includes audio from both participant and interviewer

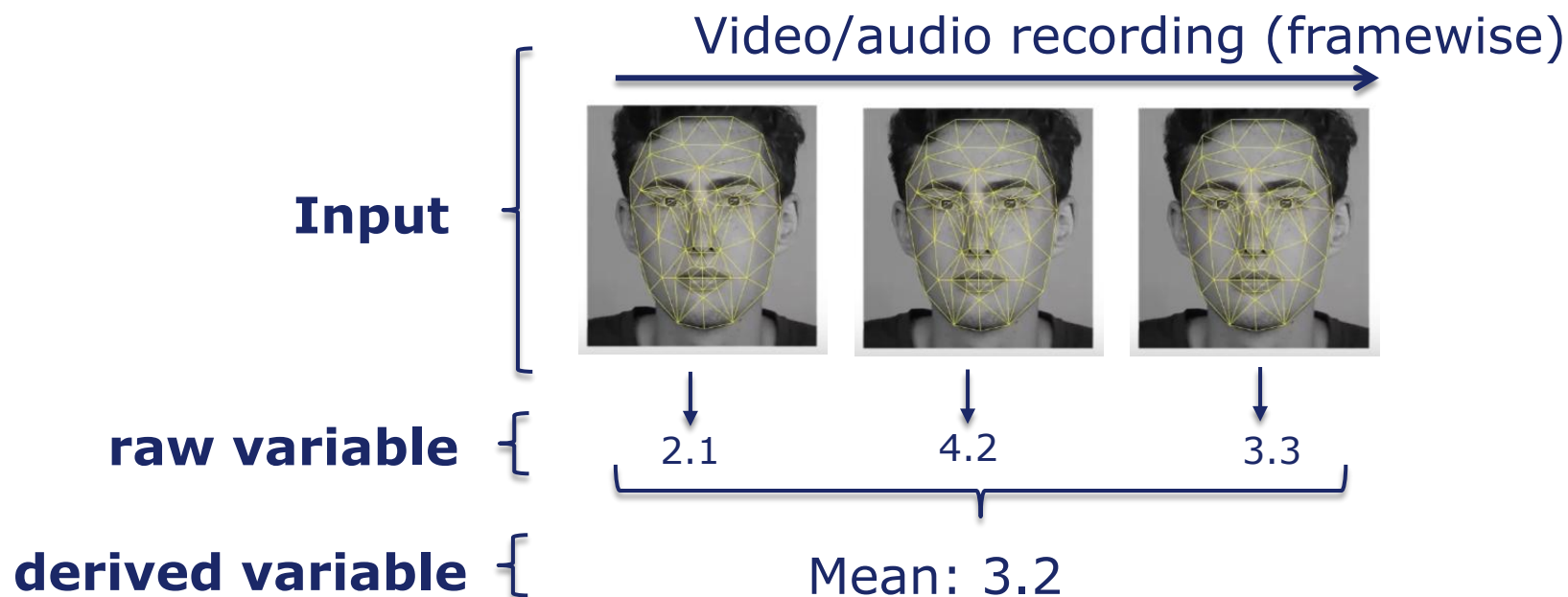
# Questions





# Intro data processing & analysis

# Interview-based digital markers: **Computation**



$$\text{Head movement in frame } n = \sqrt{(x_n - x_{n-1})^2 + (y_n - y_{n-1})^2 + (z_n - z_{n-1})^2}, \text{ where}$$

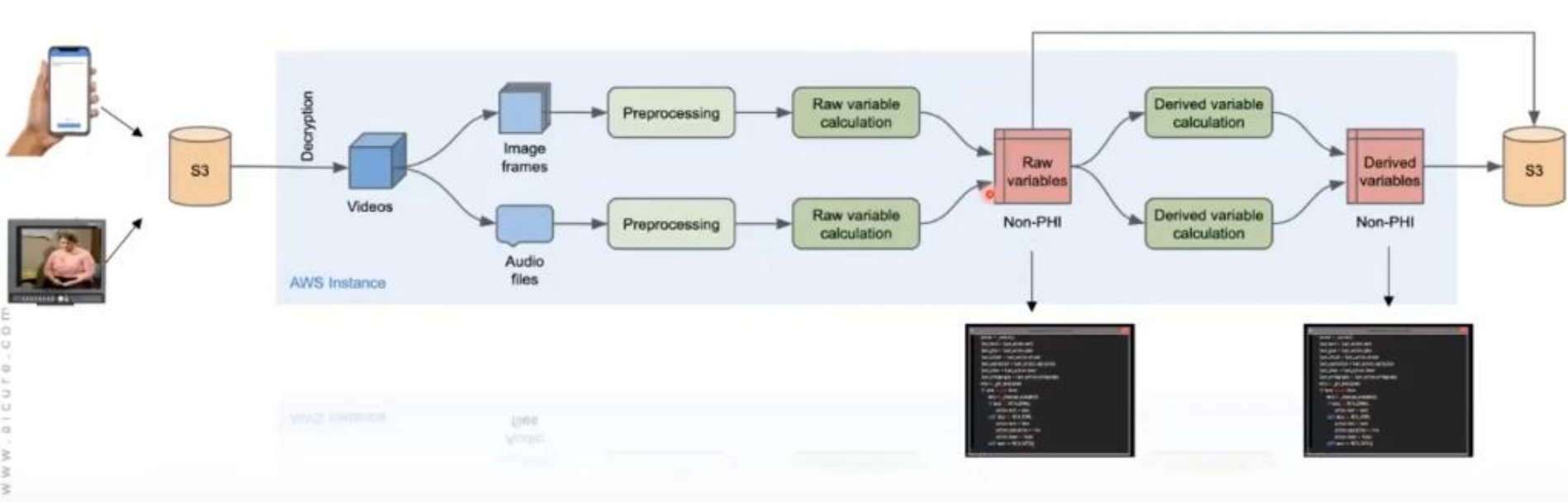
$x_n$  = head position in the x plane during the current frame,

$y_n$  = head position in the y plane during the current frame, and

$z_n$  = head position in the z plane during the current frame

Average per interval: defined time-window, question, entire interview

# Interview-based digital markers: **Computation**



# Data file structure

AutoSave OFF								
Home Insert Draw Page Layout Formulas Data Review View Tell me								
A1	fx   Filename							
	A	B	C	D	E	F	G	H
1	Filename	fac_asymma	fac_asymma	fac_asymma	fac_asymma	fac_asymma	fac_asymma	fac_asymma
2	/app/raw_data/subj02_timepoint01.mp4	1.5469	0.2258	0.7189	0.2018	1.7094	0.2808	2.0664
3	/app/raw_data/subj02_timepoint02.mp4	2.0387	1.3988	1.4166	1.3279	1.961	0.9172	2.5537
4	/app/raw_data/subj01_timepoint01.mp4	2.4274	0.5606	2.4448	0.4966	2.5277	0.5253	2.3518
5	/app/raw_data/subj01_timepoint02.mp4	3.4568	0.6572	2.3452	0.4872	2.1673	0.5161	5.0964
6								
7								
8								
9								
10								

# Considerations & Discussion

## 5 Tips for study design for conceptually driven research with audio and video markers

- Work with **your** research question – what is relevant?
- Look for ‘real-world’ differences.
- Reduce measurement noise (movements; face visibility)
- Ensure consistency of measurement settings.
- Preselect features.

## 5 Tips for quality checks of audio and video markers

- Test set up with your own example videos.
- Examine distributions.
- Examine feature time series and compare with the raw video data.
- Look for typical effects (e.g., gender differences) & correlations.
- Compare different processing Pipelines.

# Important Considerations

- Should length of video/audio be standardised?
- When should you baseline-correct for extraction of expressivity features?
- With multiple speakers, the audio needs to be partitioned (speaker diarisation)
  - Difference between interview and conversation
  - Which part of the video should be processed (speaking and/or listening)?
- Interpretability of DBMs (emotion expression != emotion)
- Ethical aspects (e.g., EU AI act regulation on emotion)



# Thank you!

## Workshop Team

Laura Meine

Rituja Pardhi

Tim Stäheli

Michelle Worthington

Lara Puhlmann

Anzar Abbas & OpenWillis Team



**DGP**s

Fachgruppe Gesundheitspsychologie  
Peer Mentoring Award 2024



Leibniz Institute for  
Resilience Research



**University of  
Zurich** <sup>UZH</sup>



**UNIVERSITÄT  
BIELEFELD**



**TECHNISCHE  
UNIVERSITÄT  
DRESDEN**



European  
Commission

**Dyna**<sup>•••</sup>**MORE**

Horizon 2020  
European Union funding  
for Research & Innovation

**DFG**

CRC 1193  
RESILIENCE



CRC 1193 "Neurobiology of  
Resilience"  
Mainz/Frankfurt



UNIVERSITÄTS**medizin.**  
MAINZ

**NIC**

NeuroImaging  
Center