Feeling key PRESSED:

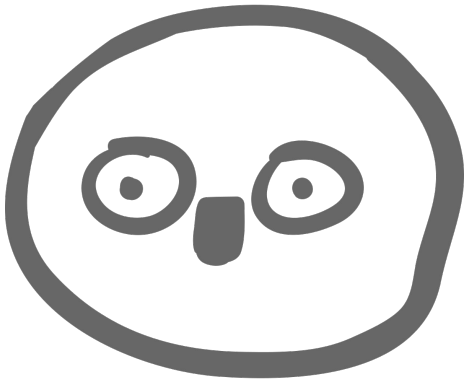
comparing the ways in which force and self-reports reveal emotion

Rubia Guerra
Oct 2022

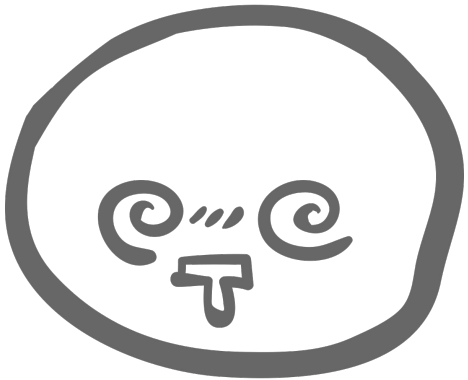
SPIN
SENSORY PERCEPTION & INTERACTION RESEARCH GROUP

HAVE YOU EVER FELT...

SO ECSTATIC YOU WERE OUT OF WORDS



...or so stressed that you could barely speak



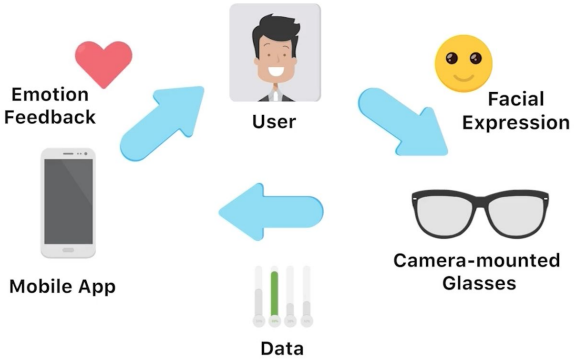
YOU ARE SO CLOSE TO FINISHING YOUR DAY

...but then you remember that you have a meeting. right. after. this. presentation.

Emotions are *dynamic* and *complex* experiences

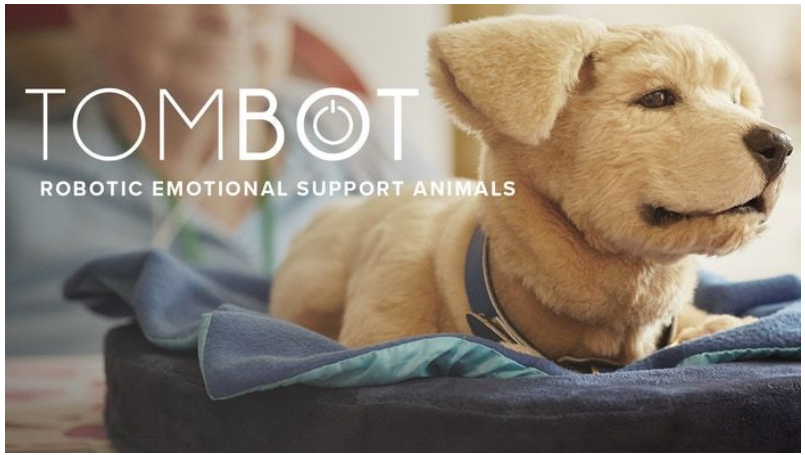
WHY DO WE CARE ABOUT EMOTIONS?

Estimating non-"verbalized" emotion



An end-to-end AI-enabled wearable platform for enhancing self-awareness of emotional health

EmoGlass, Yan et al. (CHI 2022)



RESEARCHER'S POSITION AND ETHICS CONSIDERATIONS

Overview

Motivation

Researcher Positionality

Computable Emotions

FEEL Dataset

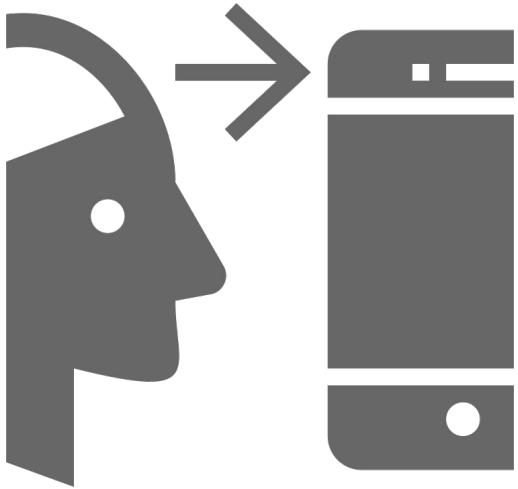
Research Questions + relevant Results (by RQ)

Limitations

Next steps

Summary

WHERE DO WE START?



WHAT ARE MODELS?

Models can be:

- representational tools

- adequate and limited to a purpose

Unfortunately, no model is morally neutral

Parker, W. S., & Winsberg, E. (2018). Values and evidence: How models make a difference. *European Journal for Philosophy of Science*, *8*(1), 125–142. https://doi.org/10.1007/s13194-017-0180-6

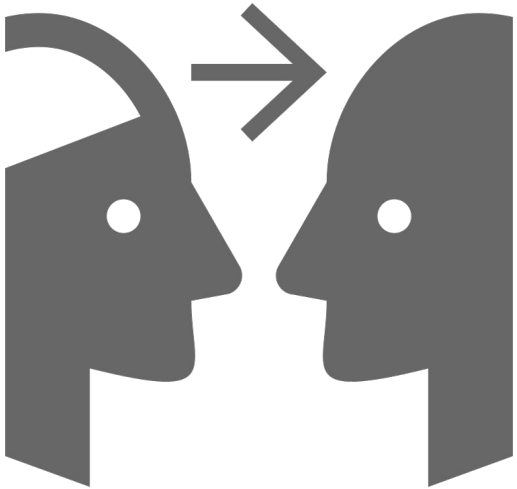
COMPUTATIONAL MODELS OF EMOTION

Estimate emotion given a set of inputs

Annotated datasets to train and evaluate models

- Emotion elicitation
- Inputs (or affective expression modalities)
- Emotion measurement instruments

Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion*, *59*, 103-126.

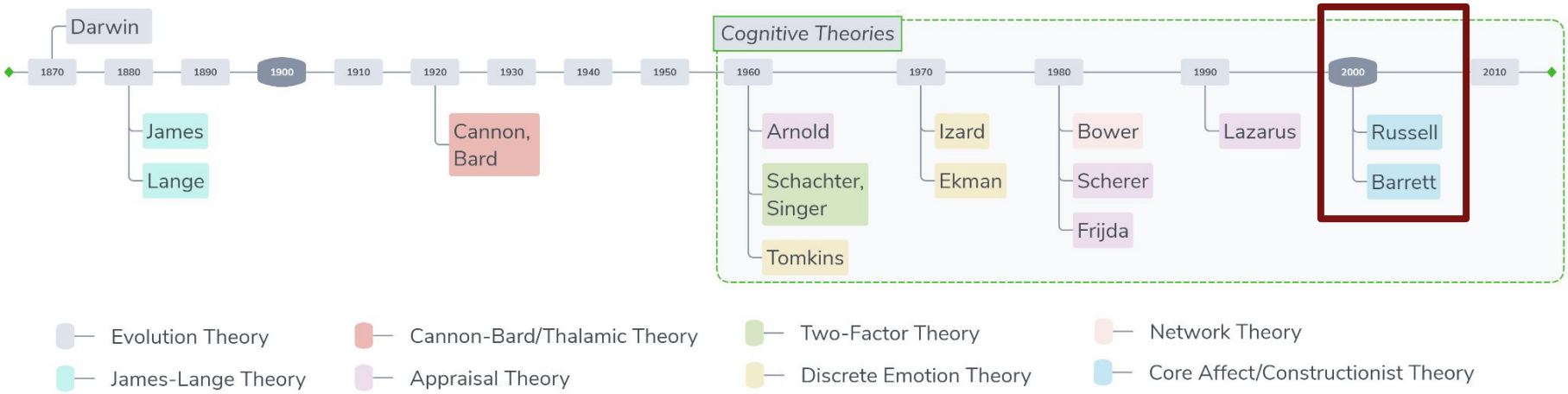


HOW DO WE START?

UNDERSTANDING EMOTION LANGUAGE

emotions are *dynamic* and *complex* experiences

...AND SO IS OUR UNDERSTANDING OF EMOTIONS



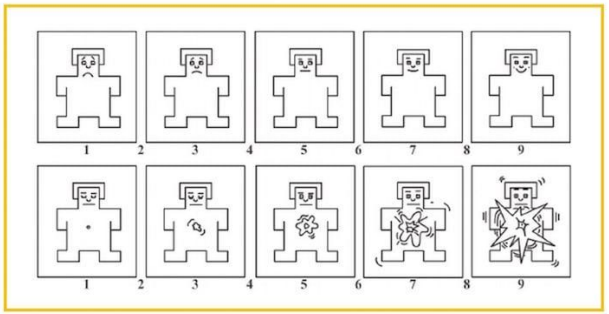
Thanapattheerakul, T., Mao, K., Amoranto, J., & Chan, J. H. (2018). Emotion in a Century: A Review of Emotion Recognition. *Proceedings of the 10th International Conference on Advances in Information Technology*, 1–8. https://doi.org/10.1145/3291280.3291788

SELF-REPORT FOR EMOTION MODELLING

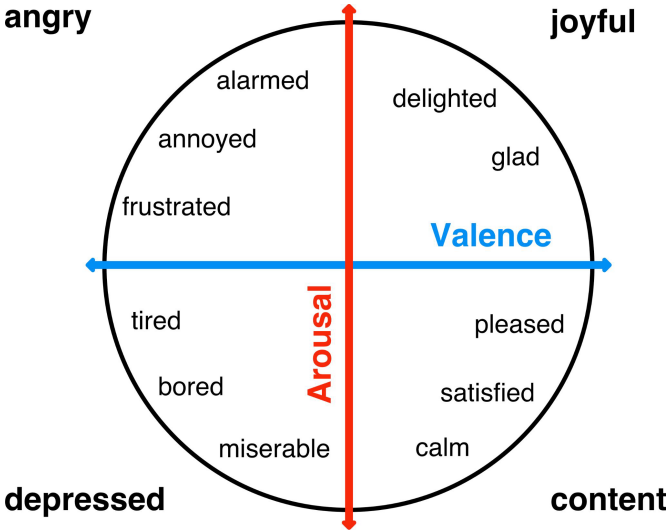
oversimplifies the experience

requires a lot of instruction

is non-exhaustive / low-density



SAM (Self-Assessment Manikin) scale

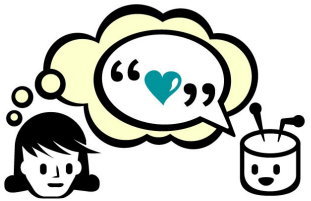


Circumplex model of affect

METAPHORS New representations that account for time, uncertainty, and multiplicity are needed.

FRAMING Framework needs to account for, negotiated, and reported on between people.

METHODOLOGY Theory and methods need to be explained for different meanings of activities, instruments.



Bucci, P. H., Cang, X. L., Mah, H., Rodgers, L., & MacLean, K. E. (2019). Real Emotions Don't Stand Still: Toward Ecologically Viable Representation of Affective Interaction.

2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), 1–7. https://doi.org/10.1109/ACII.2019.8925534

FEEL DATASET

Collected in 2018

Over 400 researcher hours and ~50 participant hours

EEG and FSR data from 16 participants

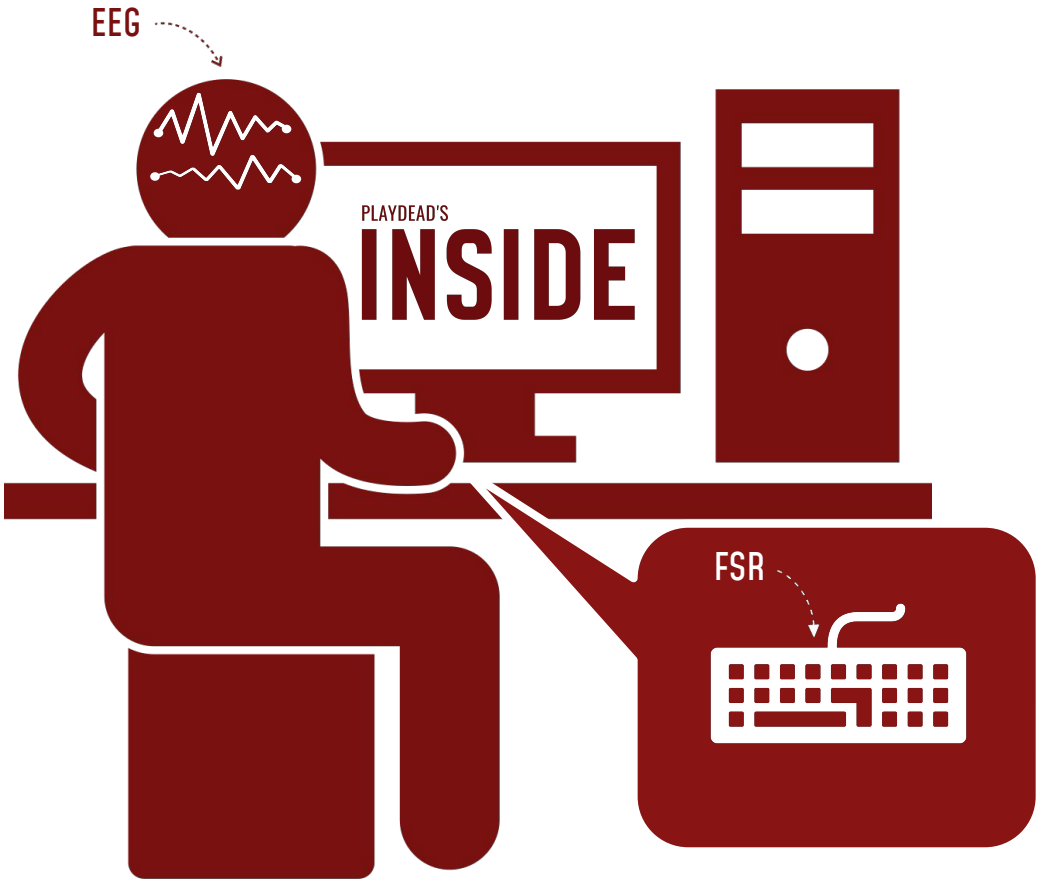
Co-designed multi-pass emotion self-reports

PRIMARY EMOTION ACTIVITY → EEG, FSR

TASK 2

TASK 3

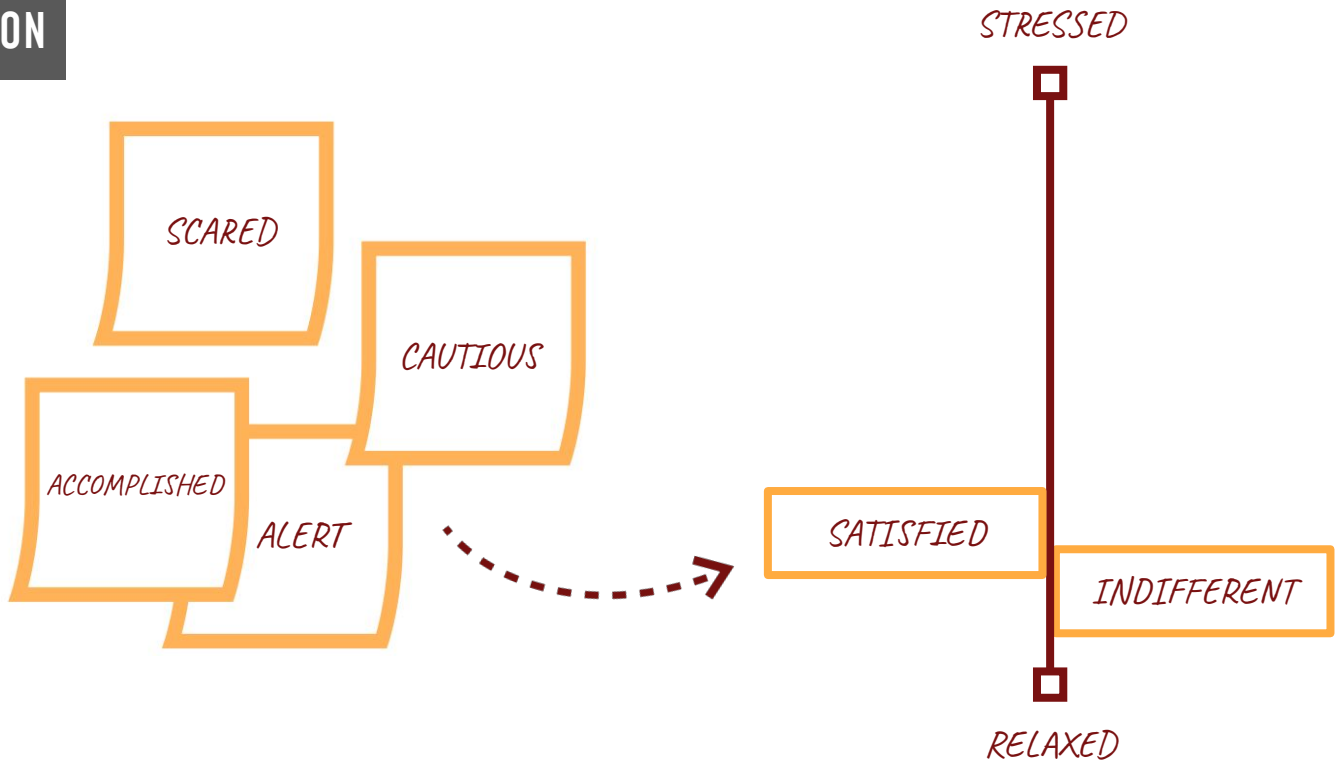
TASK 4

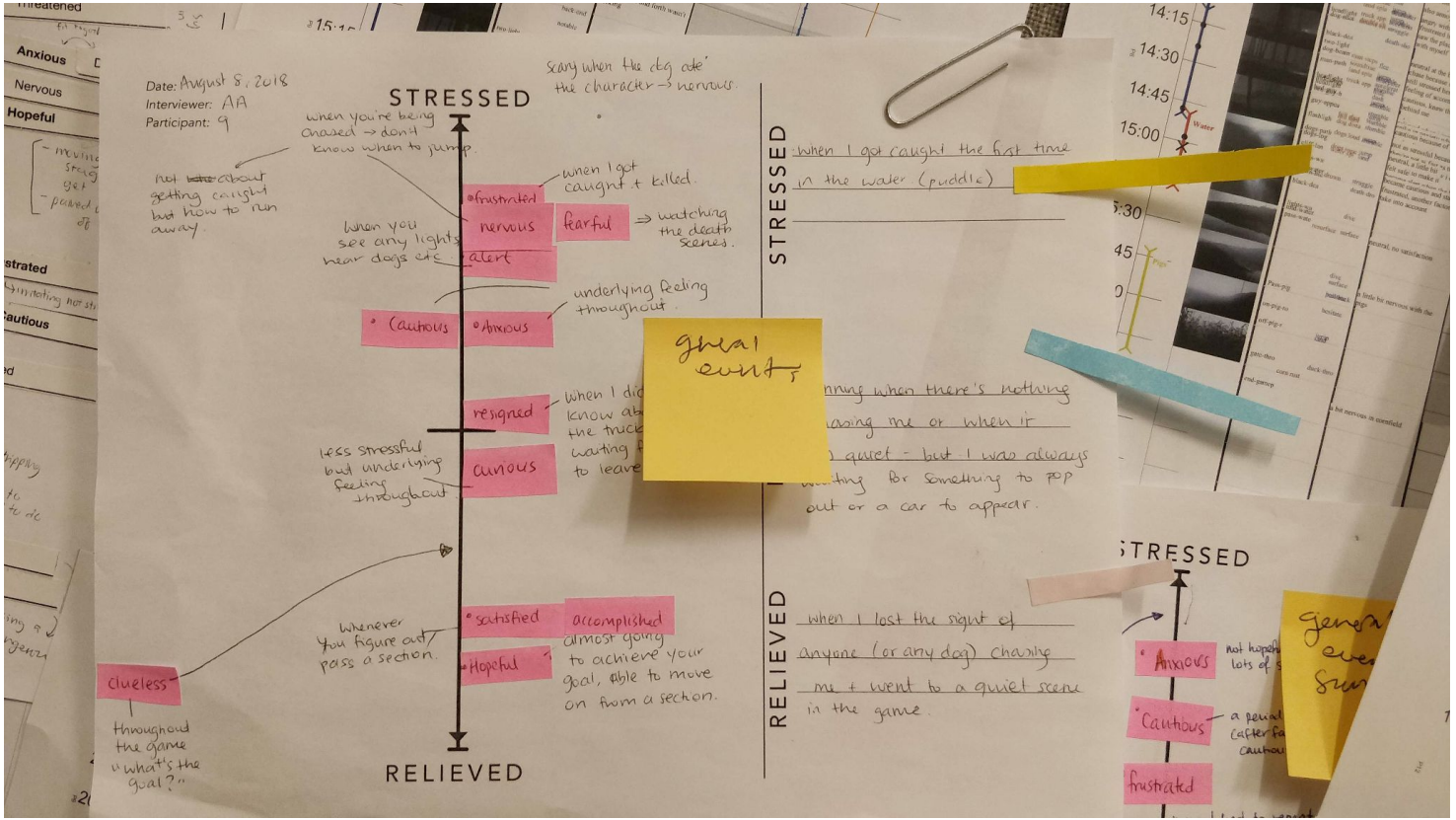


EMOTION WORD CALIBRATION

TASK 3

TASK 4

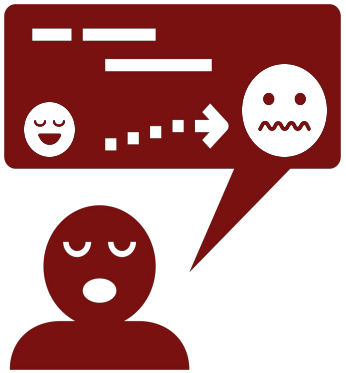


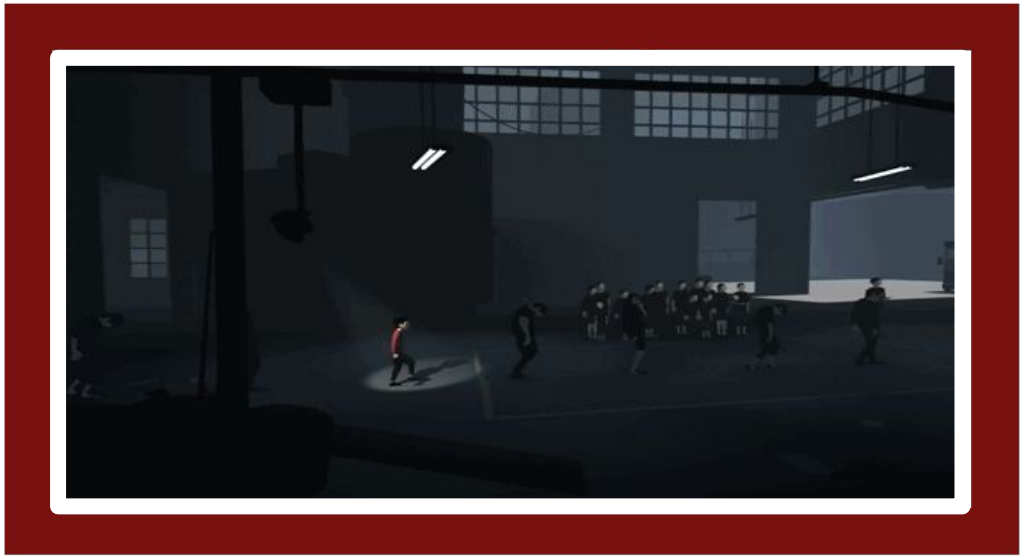

Date: August 8, 2018
Interviewer: AA
Participant: 9
STRESSED
RELIEVED
scary when the dog ate' the character → nervous.
when you're being chased → don't know when to jump.
not about getting caught but how to run away.
frustrated
nervous
fearful
⇒ watching the death scenes.
when you see any lights near dogs etc.
alert
underlying feeling throughout.
Cautious
Anxious
resigned
curious
less stressful but underlying feeling throughout
satisfied
accomplished
almost going to achieve your goal, able to move on from a section.
Hopeful
whenever you figure out/ pass a section.
clueless
throughout the game "what's the goal?"
when I got caught the first time in the water (puddle)
when I lost the sight of anyone (or any dog) chasing me + went to a quiet scene in the game.
Anxious
Cautious
frustrated
14:15
14:30
14:45
15:00
Water

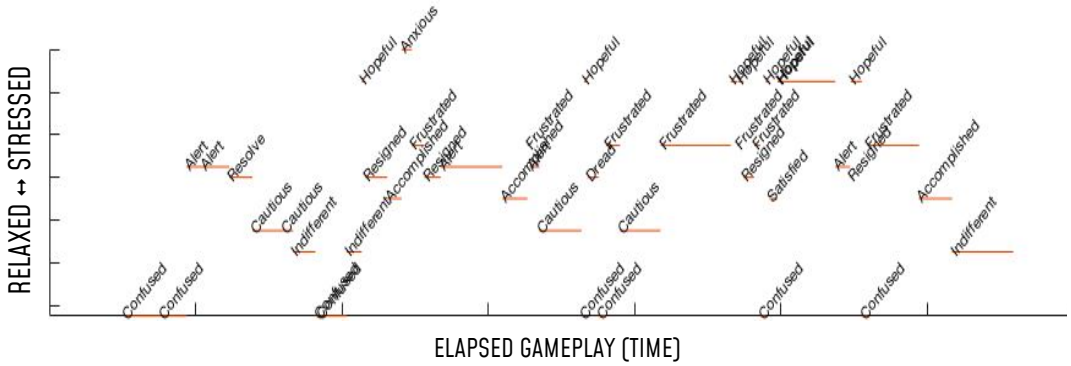
EMOTION WORD CALIBRATION

INTERVIEW → TIMELINE WITH CALIBRATED WORDS

TASK 4







PRIMARY EMOTION ACTIVITY → EEG, FSR

EMOTION WORD CALIBRATION

INTERVIEW → TwCW

CONTINUOUS ANNOTATION → CA

MULTI-PASS CO-CREATED SELF-REPORTS

Contextualized emotion calibration scales

Continuously annotated emotion (on emotion scales) High frequency

Emotion experience interview Low frequency

Cang, X. L., Guerra, R. R., Bucci, P., Guta, B., Rodgers, L., Mah, H., … MacLean, K. E. (2022 (to appear)). Choose or Fuse: Enriching Data Views with Multi-Label Emotion Dynamics. IEEE 10th Int'l Conf on Affective Computing & Intelligent Interaction (ACII).

HOW DO WE ANALYZE EMOTION DATA?

Personalized models emotion expression at an individual level

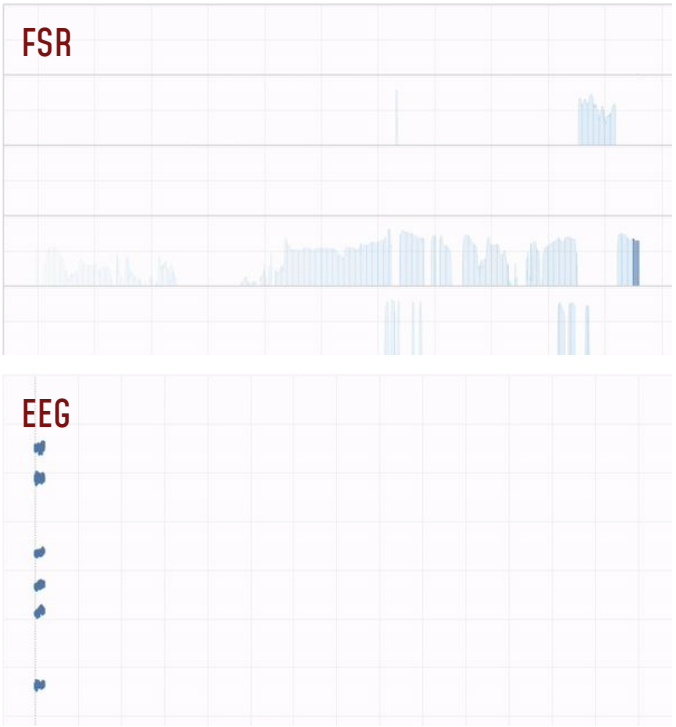
More than one computational method to analyze data

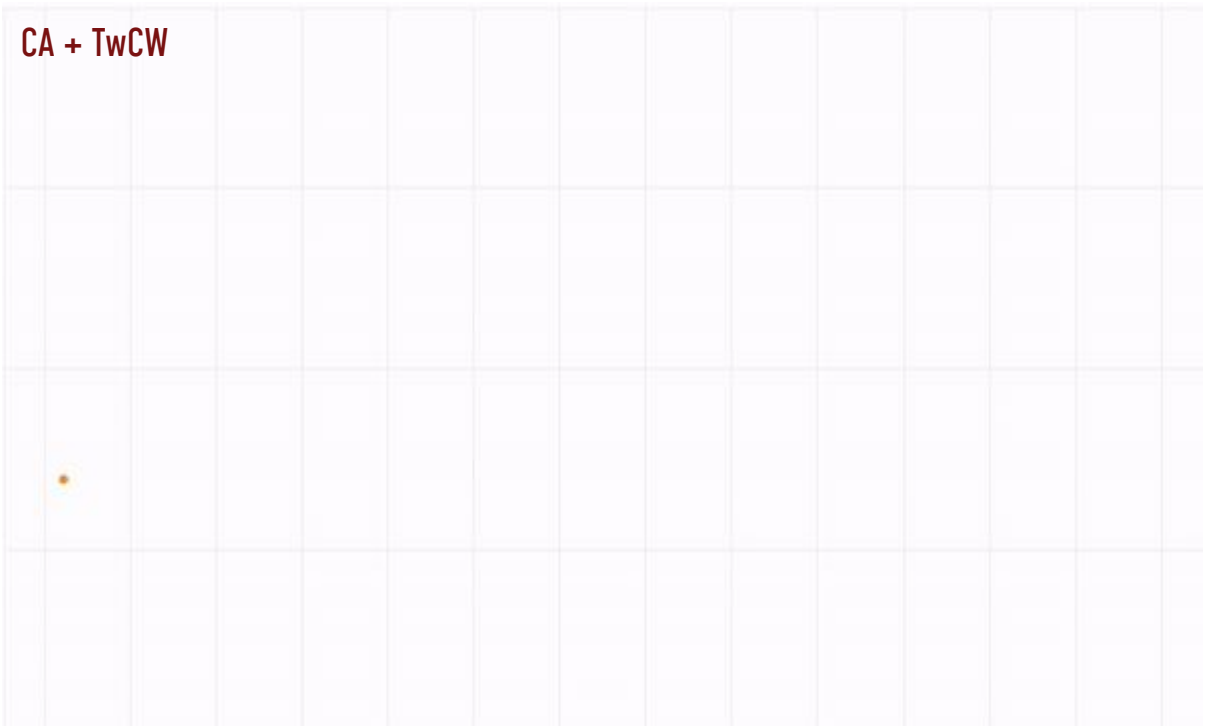
Multiple algorithms and feature selection **methods** if using Machine Learning

Azari, B., Westlin, C., Satpute, A. B., Hutchinson, J. B., Kragel, P. A., Hoemann, K., ... & Barrett, L. F. (2020). Comparing supervised and unsupervised approaches to emotion categorization in the human brain, body, and subjective experience. *Scientific reports*, *10*(1), 1-17.

[RQ1] what does keypress force tell us about emotions during tense gameplay?

First, what are the labels?



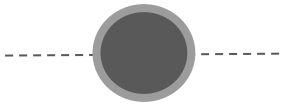


FORCE EEG AND EMOTION-LABELLED DATASET

CA + TwCW

FORCE EEG AND EMOTION-LABELLED DATASET

EMOTION METAPHORS



emotion-as-position

emotion-as-direction

Cang, X. L., Guerra, R. R., Bucci, P., Guta, B., Rodgers, L., Mah, H., ... MacLean, K. E. (2022 (to appear)). Choose or Fuse: Enriching Data Views with Multi-Label Emotion Dynamics. IEEE 10th Int'l Conf on Affective Computing & Intelligent Interaction (ACII).

RQ1: WHAT DOES KEYPRESS FORCE TELL US ABOUT EMOTIONS IN A TENSE GAMEPLAY?

FORCE EEG AND EMOTION-LABELLED DATASET

[RQ1] APPROACH

Keypress force encodes emotion most model *typing*

EEG is well studied → compare FSR models to EEG

FEEL DATASET

keypress from 5 keys (←↑→↓ and **alt**)

EEG (out of scope)

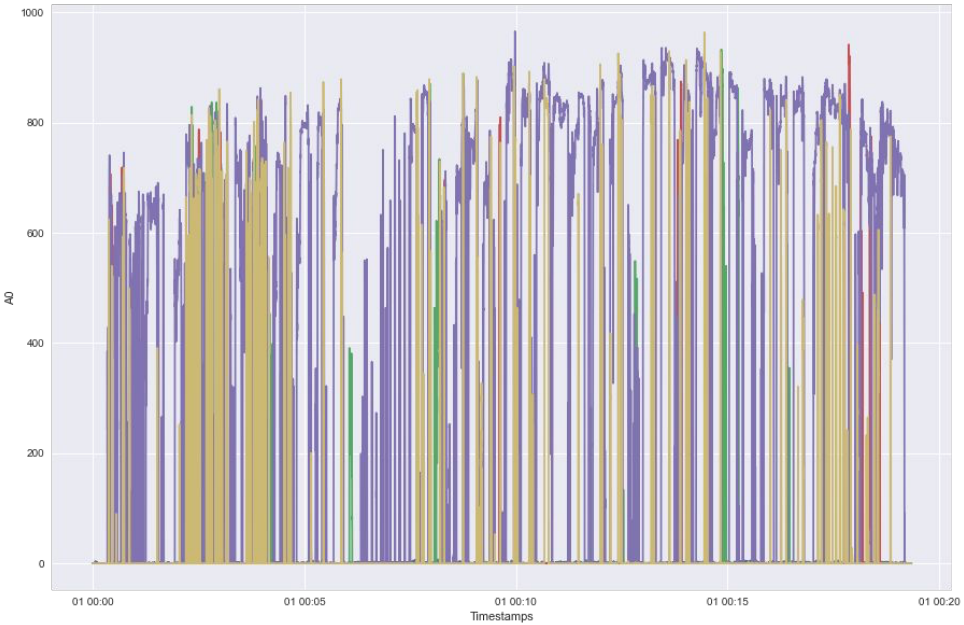
FEATURES

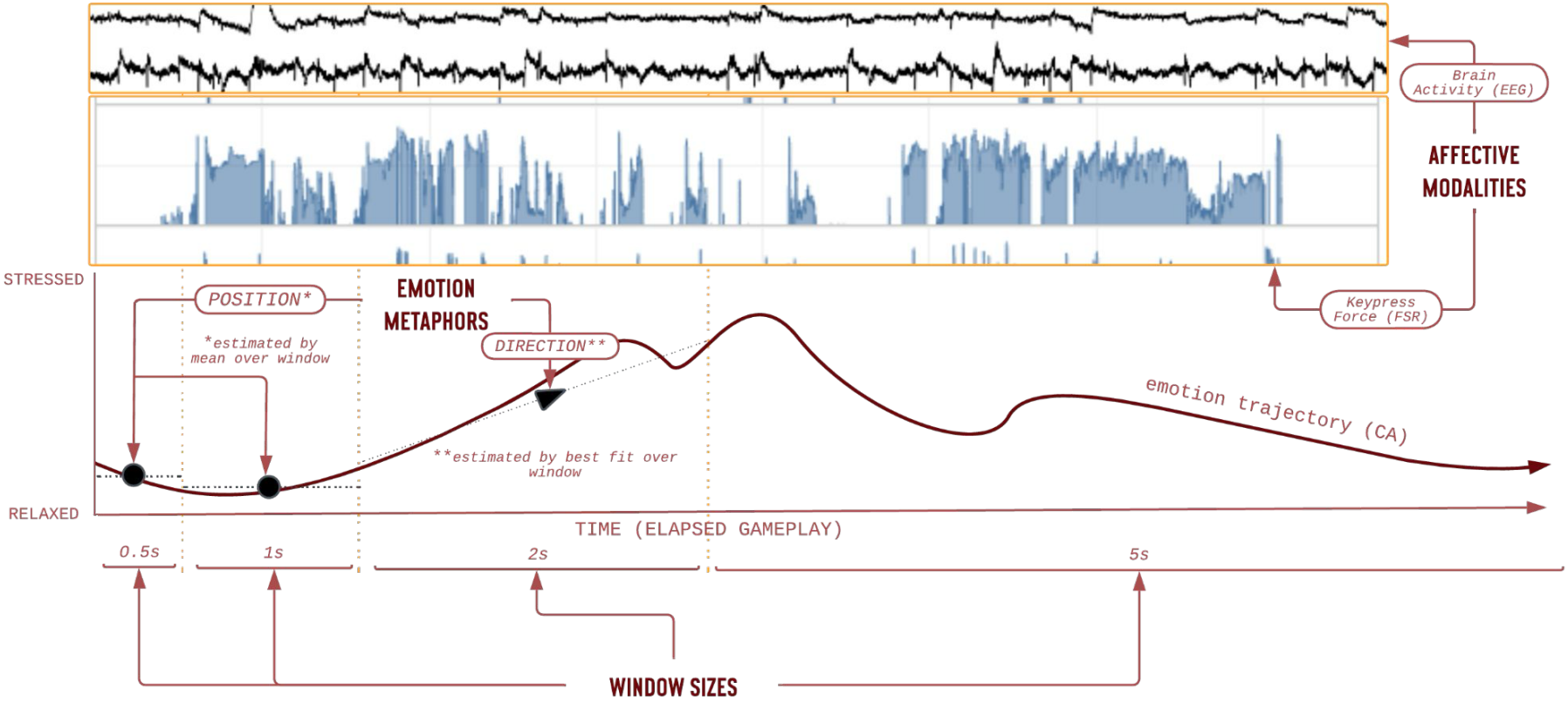
statistical analysis max, min, var, area under the curve (CANG et al., TAC'21)

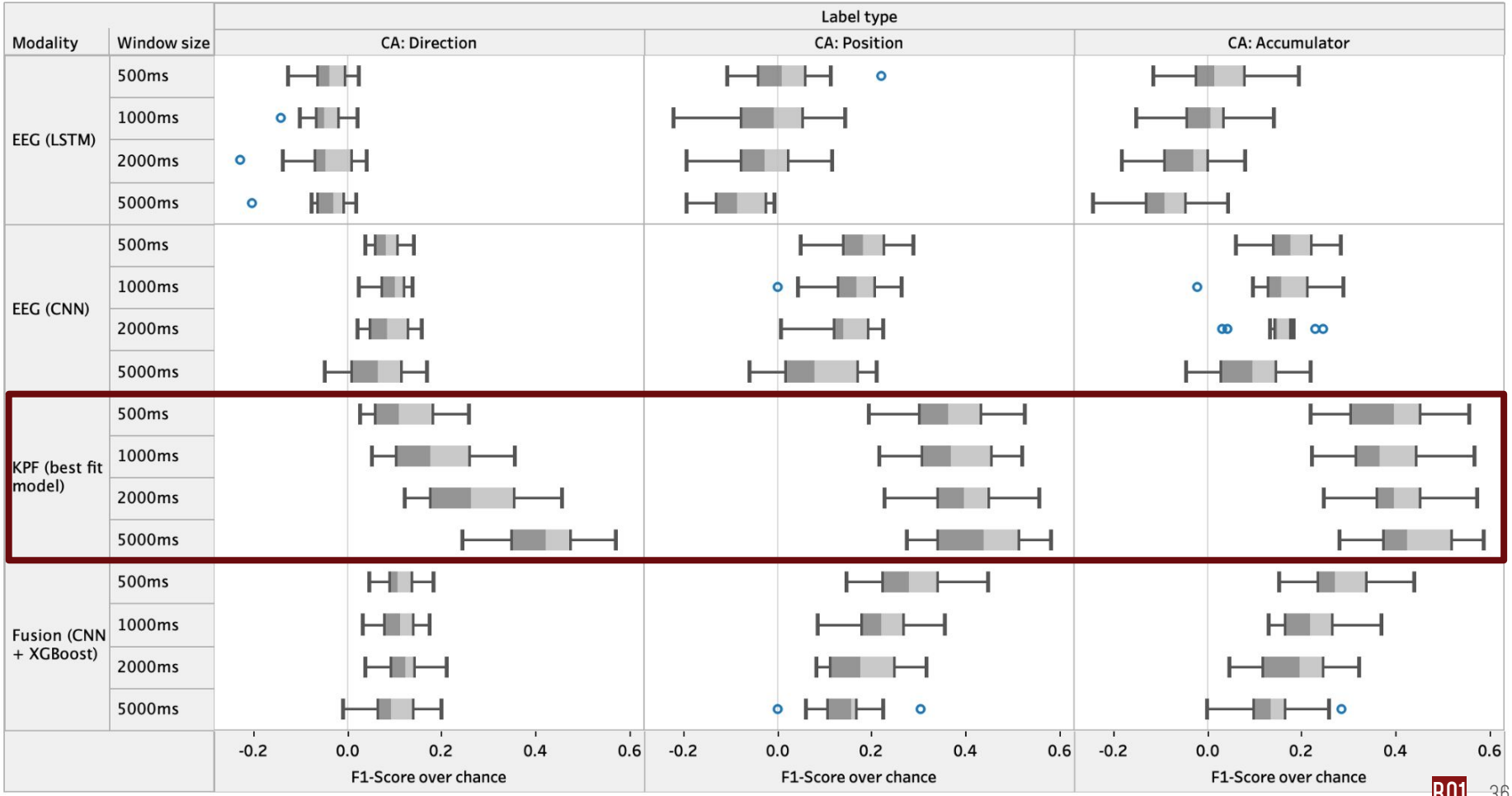
frequency analysis peak count, fundamental frequency, max amplitude, cepstral analysis (CANG et al., TAC'21)

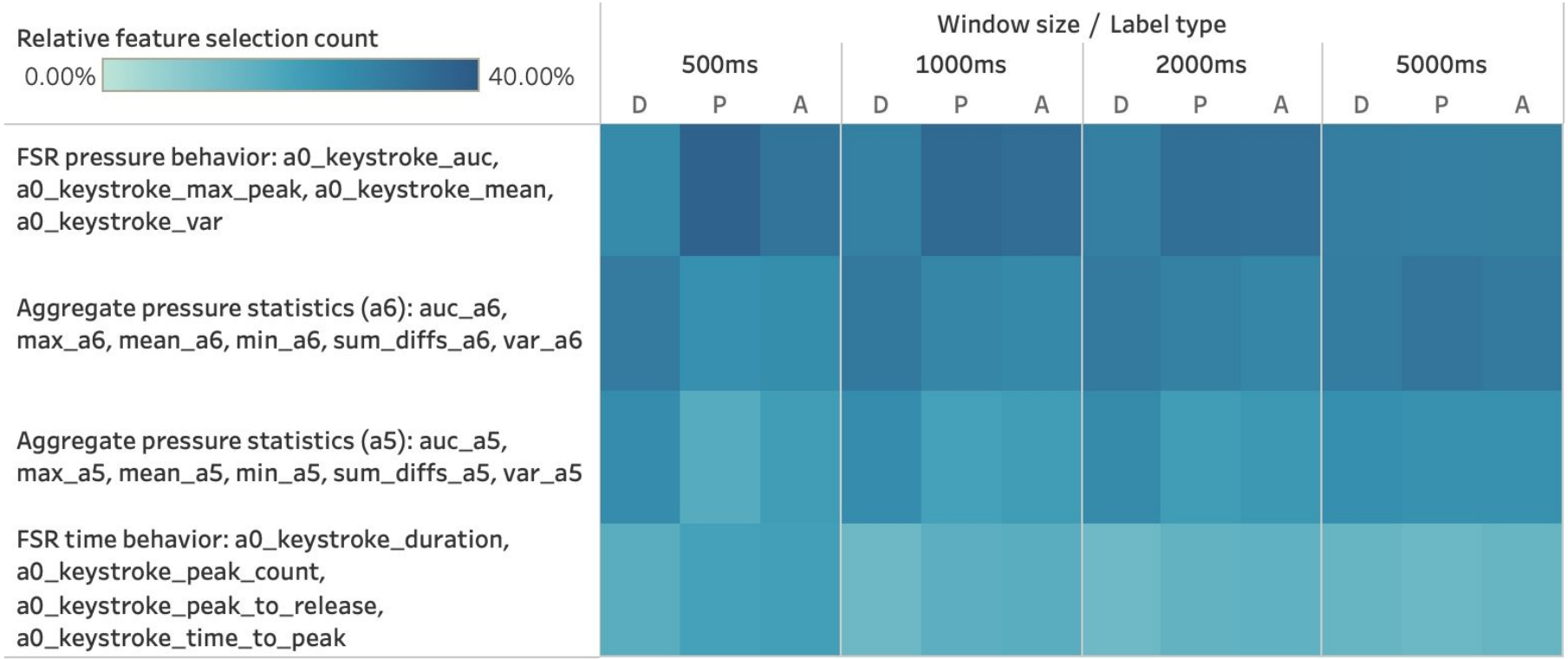
keystroke dynamics key up, key down, count of active keys (Epp et al., CHI 2011)

ADSR envelope time to peak, peak to release, keystroke duration, max peak

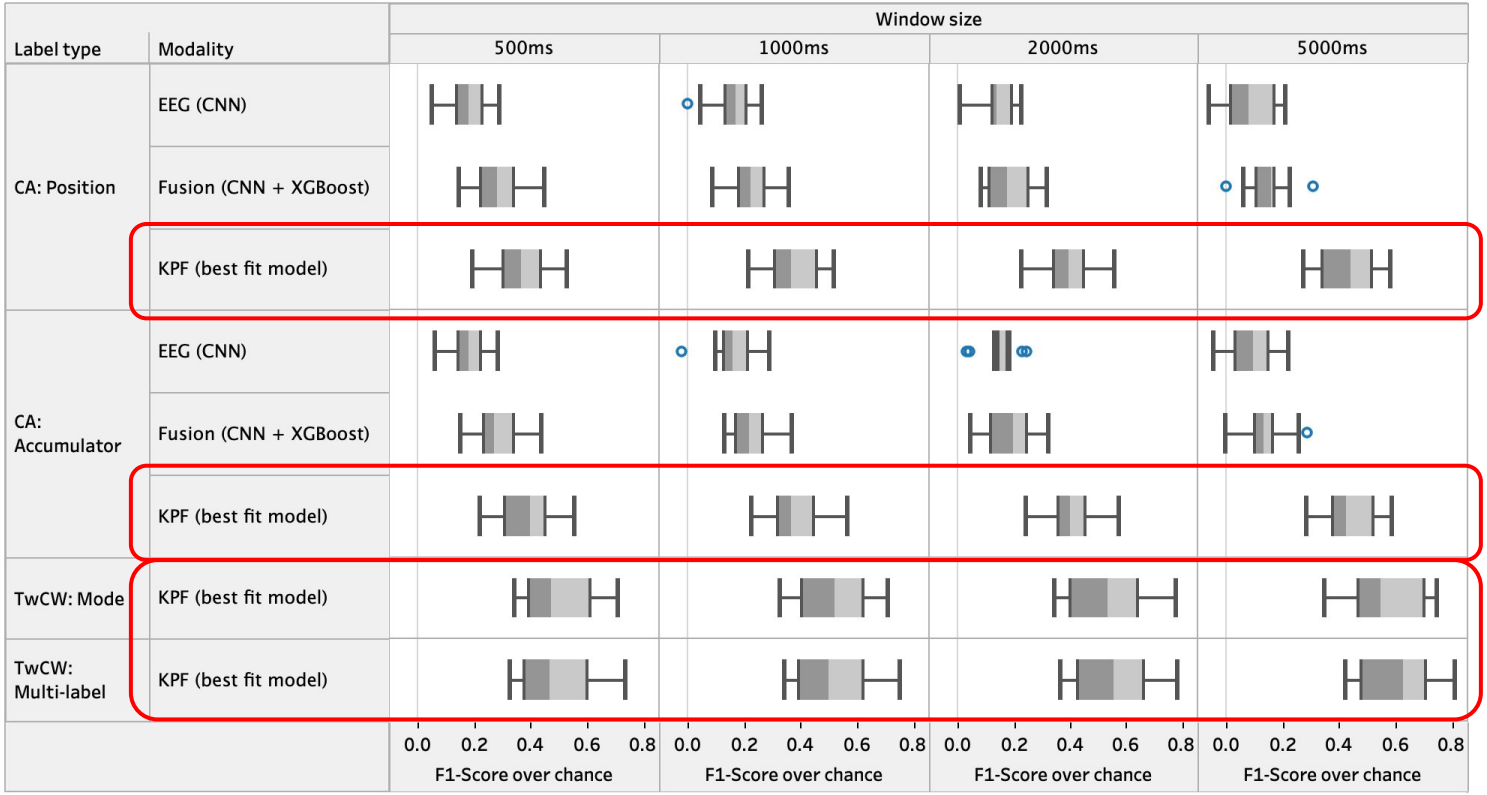



Brain Activity (EEG)
AFFECTIVE MODALITIES
Keypress Force (FSR)
STRESSED
POSITION*
*estimated by mean over window
EMOTION METAPHORS
DIRECTION**
**estimated by best fit over window
emotion trajectory (CA)
RELAXED
TIME (ELAPSED GAMEPLAY)
0.5s
1s
2s
5s
WINDOW SIZES

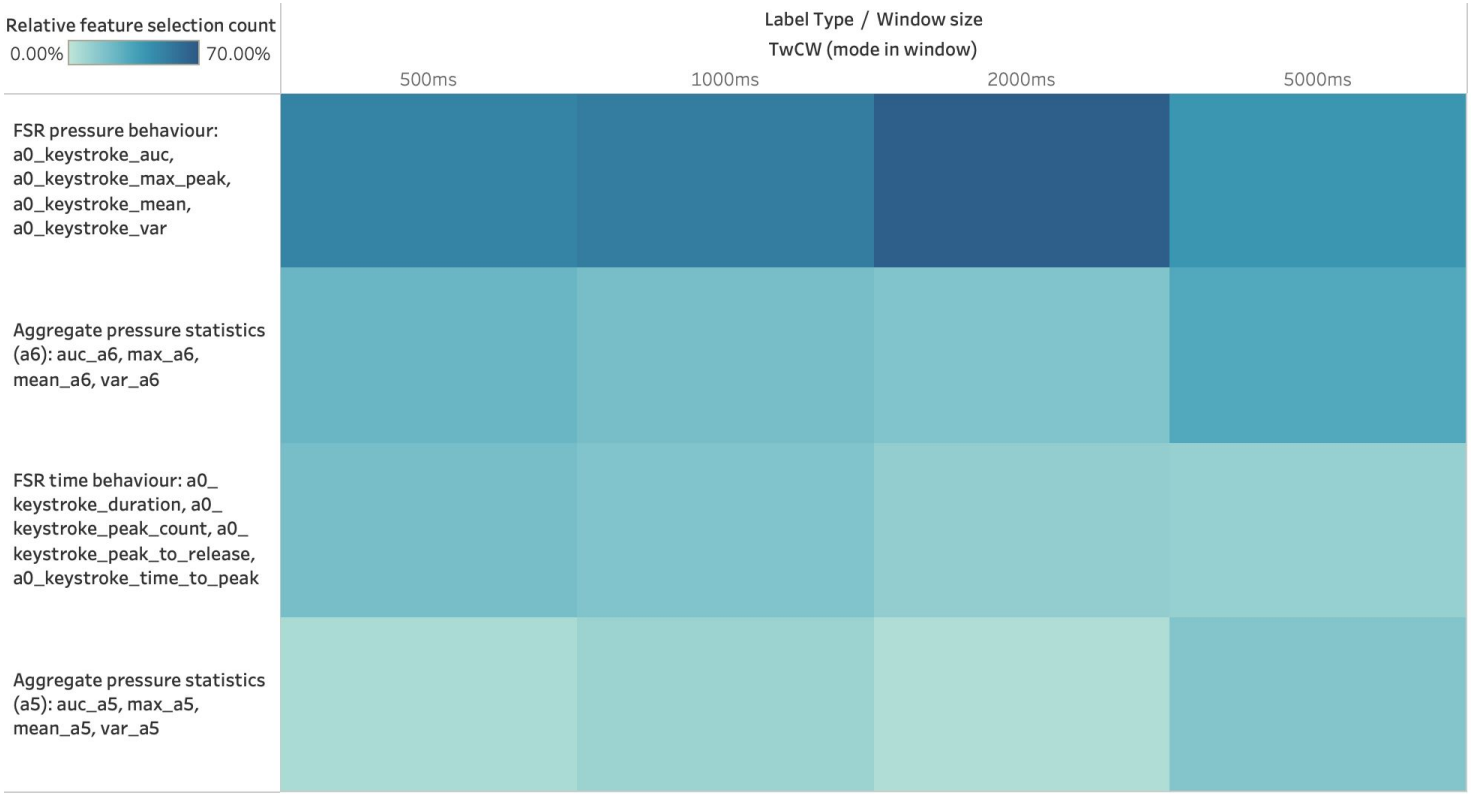

Label type
Modality
Window size
CA: Direction
CA: Position
CA: Accumulator
EEG (LSTM)
EEG (CNN)
KPF (best fit model)
Fusion (CNN + XGBoost)
500ms
1000ms
2000ms
5000ms
-0.2
0.0
0.2
0.4
0.6
F1-Score over chance



RELATIVE FEATURE SELECTION COUNT BY WINDOW AND LABEL TYPE



Comparing TwCW and CA on emotion-as-position



RELATIVE FEATURE SELECTION COUNT BY WINDOW AND LABEL TYPE

WHAT DOES KEYPRESS FORCE TELL US ABOUT EMOTIONS IN A TENSE GAMEPLAY?

FSR VS. EEG VS. FUSION

Keypress force is a better predictor of dynamic emotion than brain **activity estimated** by EEG from comparing modalities on CA

Manual **touch pressure** encodes valuable emotion content Feature importance analysis

RQ2: HOW DOES ACCOUNTING FOR SUBJECTIVE AND DYNAMIC EMOTION IMPROVE COMPUTATIONAL EMOTION MODELLING?

[RQ2] APPROACH

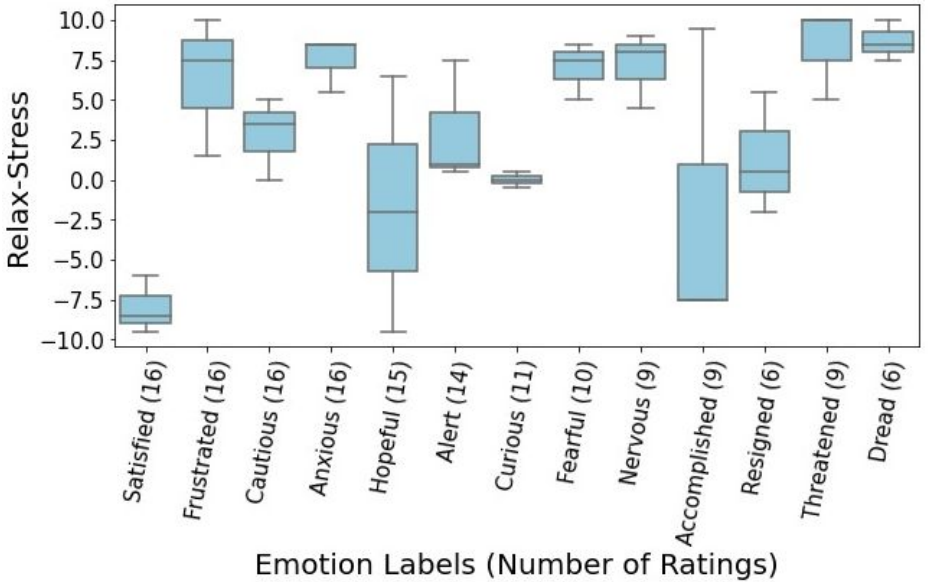
Do user-centred emotion reports add new information?

Do people rank common emotion words similarly? In what ways does labelling data differ by pass?

What are effective ways to incorporate multi-pass labels to model dynamic emotion?

Do people rank common emotion words similarly?

Emotion word calibration inter-rater reliability test



*Work by Chuxuan Zhang

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, *15*(2), 155-163.

Do people rank common emotion words similarly?

Emotion word calibration

*Work by Chuxuan Zhang

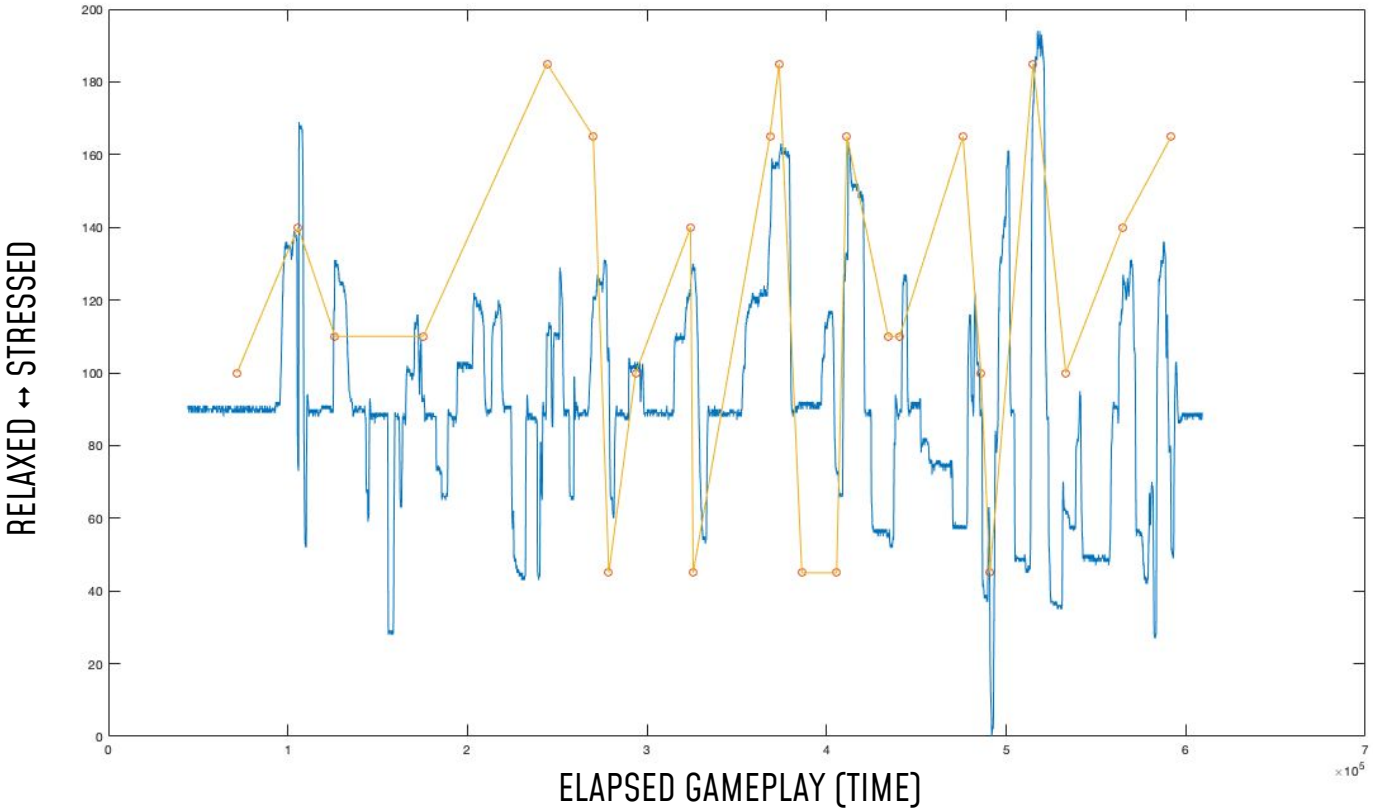
Differences in how individuals **interpret** emotion words inter-rater reliability test

highlighting the importance of personalized models

Do user-centred emotion reports add new information?

Continuous annotation Downsampled at 30Hz

Timeline with Calibrated Words Less than 0.1Hz


RELAXED ↔ STRESSED
ELAPSED GAMEPLAY (TIME)
×10^5
CONTINUOUS ANNOTATION (CA)
CALIBRATED WORDS (TwCW)

Do user-centred emotion reports add new information?

Continuous annotation Low-pass and downsampled at 0.05Hz

Timeline with Calibrated Words Resampled at 0.05Hz

Are CA and TwCW interchangeable? Pearson's correlation for signal similarity

Can we use CA forecast TwCW? Granger causality test

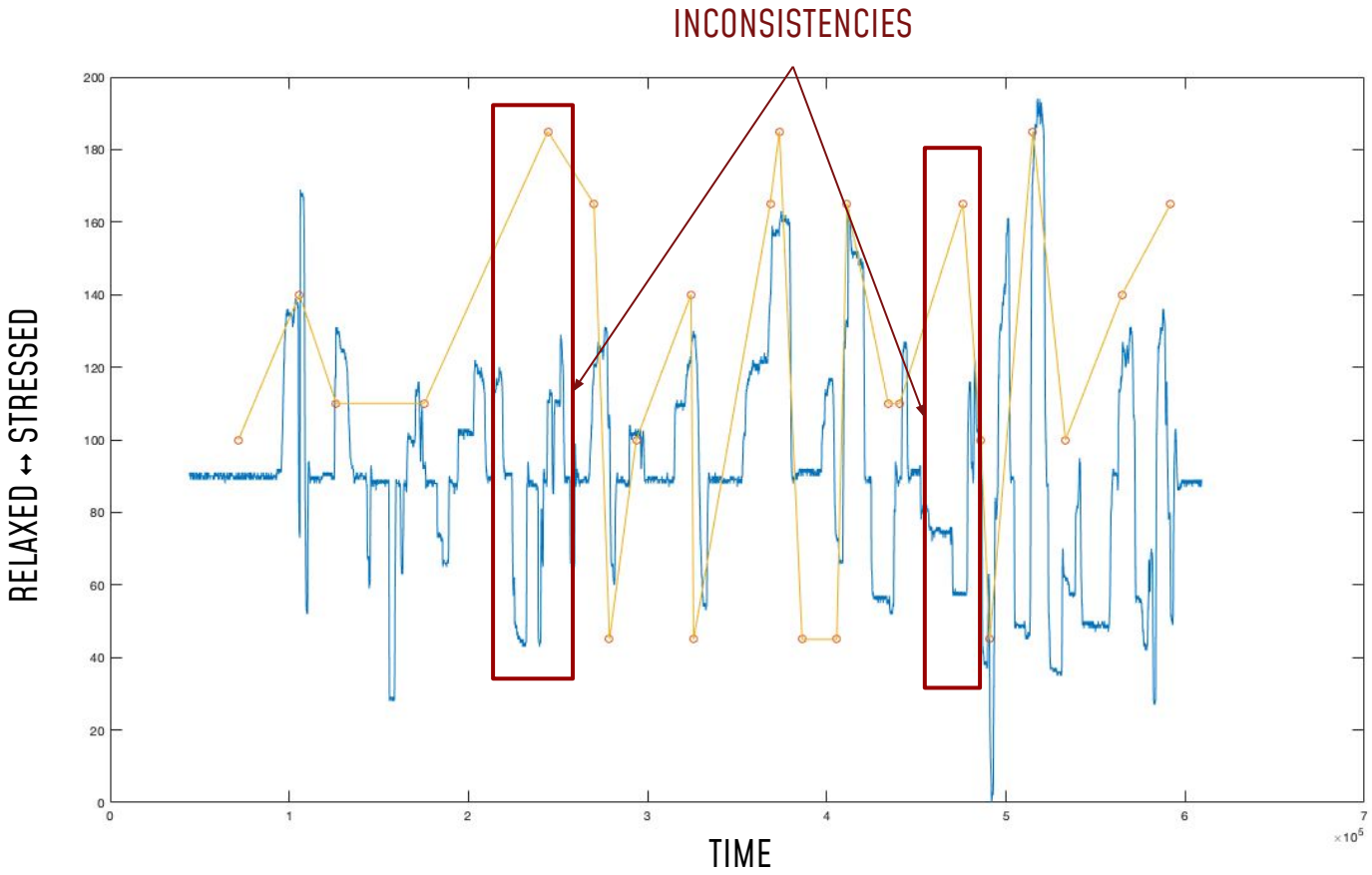
Lütkepohl, H. (2013). *Introduction to multiple time series analysis*. Springer Science & Business Media.

Do user-centred emotion reports add new information?

Continuous annotation
Timeline with calibrated words

Although individuals' self-reports **are not interchangeable** Pearson's correlation for signal similarity

both are **appropriate as responses** to the **same stimulus** Granger causality test


INCONSISTENCIES
RELAXED ↔ STRESSED
TIME
CONTINUOUS ANNOTATION (CA)
CALIBRATED WORDS (TwCW)

What are effective ways to incorporate multi-pass labels to model dynamic emotion?

Continuous annotation position or direction

Timeline with Calibrated Words position

What are effective ways to incorporate triangulated labels to model dynamic emotion?

INCORPORATE

Continuous annotation position or direction

Timeline with Calibrated Words position

BY

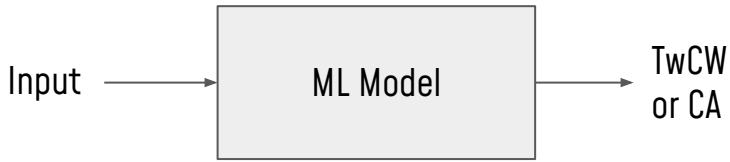
1. Combining labels
2. Extracting new information: instance probabilities
3. Resolving conflicts: label probabilities

Liu, W., Wang, H., Shen, X., & Tsang, I. (2021). The emerging trends of multi-label learning. IEEE transactions on pattern analysis and machine intelligence.

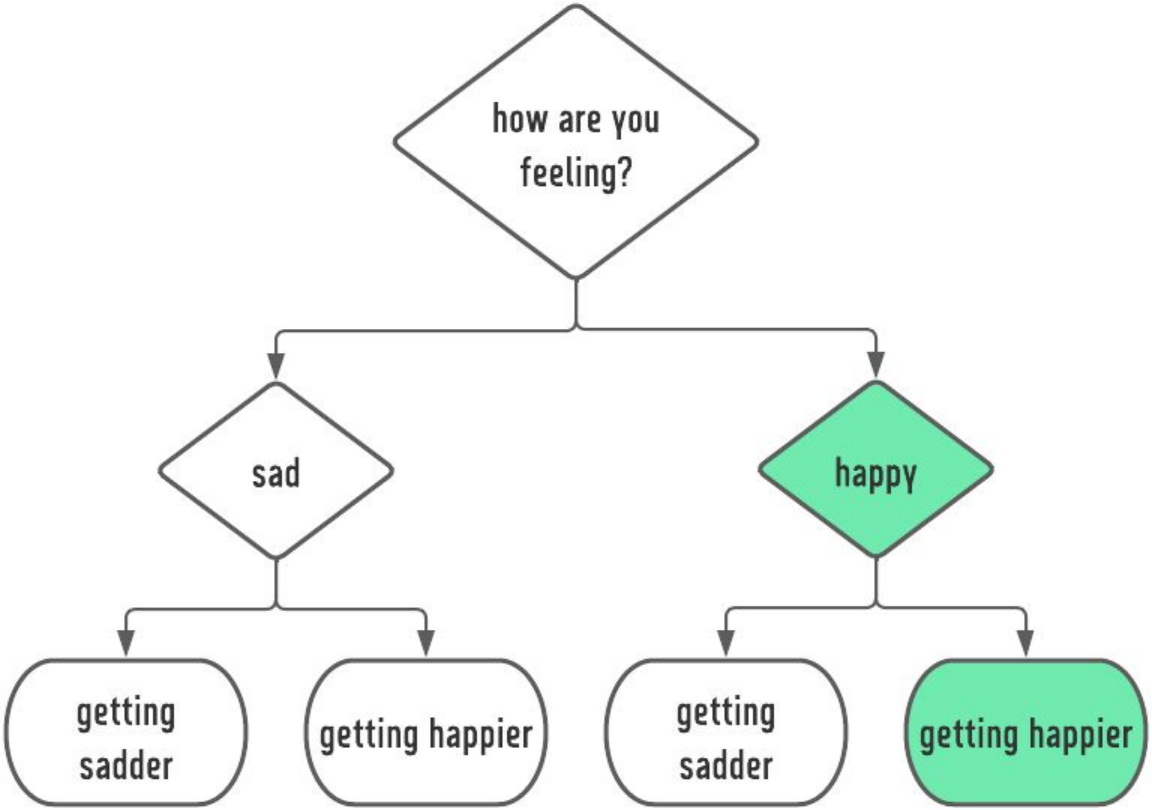
SINGLE PASS LABELS
(example: Keypress Force Classification Results shown earlier)

either CA **or** TwCW

"standard" Machine Learning pipeline



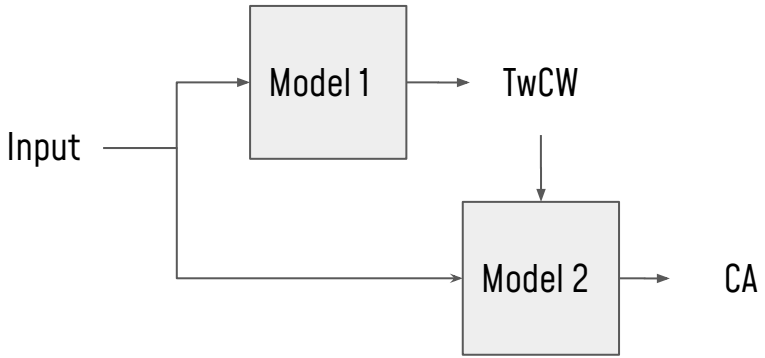
We took a hierarchical approach to three different multi-pass labeling models



MULTI-PASS LABEL COMBINATION

What if we

output **a combination of** TwCW and CA...



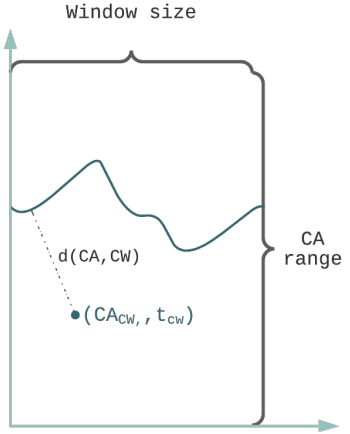
HIERARCHICAL MULTILABEL MODEL

MULTI-PASS LABEL RESOLVING

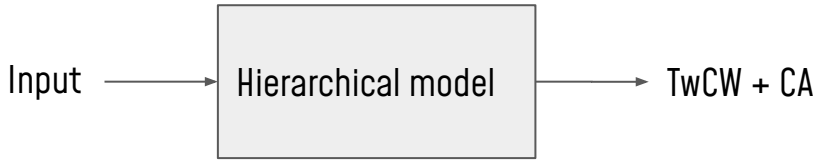
MULTIPASS LABELS

...or a **extract information** from both CA and TwCW

p(TwCW|CA) = 1 - d(CA, CW)



Resolving labels by using p(TwCW|CA)



LABEL THRESHOLDING MODEL

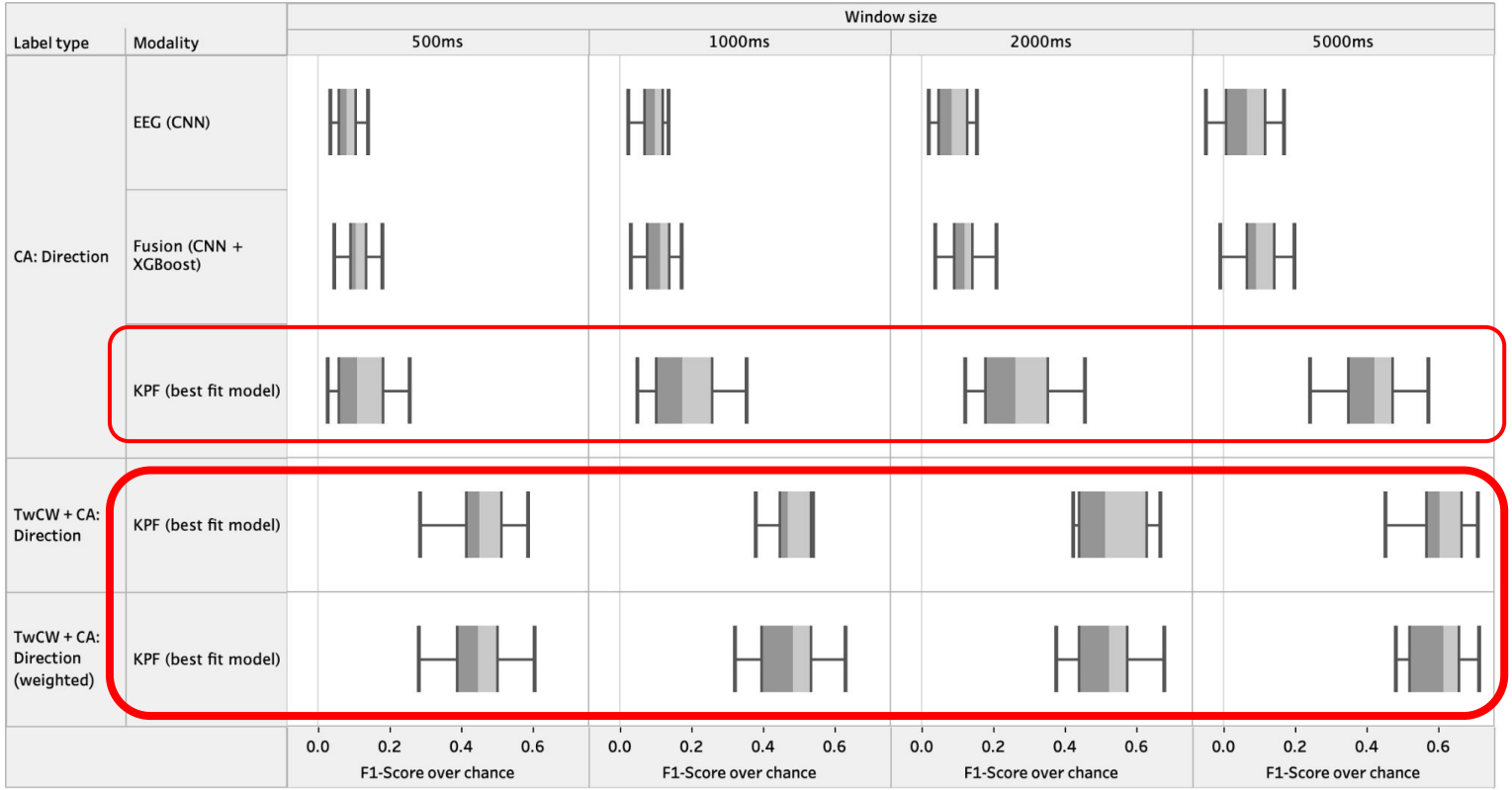
Weight inputs by using p(TwCW|CA)

INSTANCE THRESHOLDING MODEL

WHAT ARE EFFECTIVE WAYS TO INCORPORATE TRIANGULATED LABELS TO MODEL DYNAMIC EMOTION?

1. Combining self-reports
2. Extracting new labels
3. Resolving conflicts

Evaluate our proposed modelling schemes using FSR data



emotion-as-direction

WHAT ARE EFFECTIVE WAYS TO INCORPORATE TRIANGULATED LABELS TO MODEL DYNAMIC EMOTION?

1. Combining self-reports
2. Extracting new labels
3. Resolving conflicts

Emotion words are sparse, but **meaningful** Better emotion-as-position with TwCW

Combining self-reports **boosts performance** and enables better **contextualized outputs** Hierarchical classification with CA + TwCW

LIMITATIONS

1D touch

Scope

Single emotion scale

Calibration change?

NEXT STEPS: TOUCHTALES

Pre- and post-calibration

Multiple self-reports

Different scales

Repeated sessions

Participant emotion baseline + personal context

Spontaneous 2D touch

with pressure and *potentially* sheer NOVEL!

EVERYTHING IN A SLIDE

- Analyzed data from FEEL
  - ↳ Manual touch pressure encodes valuable emotion content
  - ↳ Multiple self-report passes are hard work, but it pays off
  - ↳ New study coming soon, be our participant

Karon MacLean
Laura Cang
Bereket Guta
Shay Zhang
Shinmin Hsu
Nao Rojas
Qianqian Feng
Paul Bucci
Devyani MacLaren
Elizabeth Reid
Kattie Sepehri
Preeti Vyas
Kemba

Hannah Elbaggari
Soheil Kianzad
Unma Desai
Yuna Watanabe
Dongwook Yoon
Heather O'Brien
Joanna McGrenere
Laura Ballay
Margo Seltzer
Robert Xiao
Tamara Munzner
William Bowman
Yan Liu

Adriana, Augusto, Saul, Samuel, Huong, Sebastian, Churro

Diane, Cristiano, Gabriel, Artur, Isabela, Camila, Izabella, Vinicius, Henrico, Raquel, Amir, Paulette, Joe, Nichole, Shaurya, Binx, Reggie, Blue, Sumi

Do you ever look at someone and wonder **what's going on inside** their head?