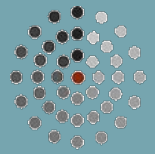
Measuring Group Flow of Jazz Musicians through Face Emotion Recognition and Body Signals

Leeban Morgan and Peter Gloor

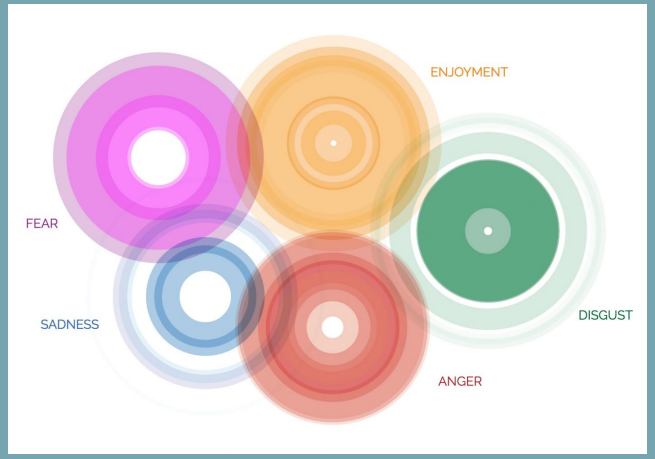


MIT CENTER FOR COLLECTIVE INTELLIGENCE

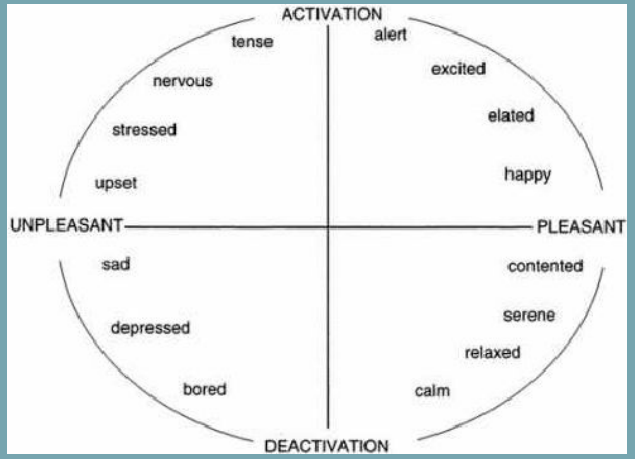
Motivation & Background

Emotions

In order to predict one's emotions, we first need to have our emotions defined for us.



Discrete Model
Paul Ekman's Theory (Ekman & Friesen, 1971).

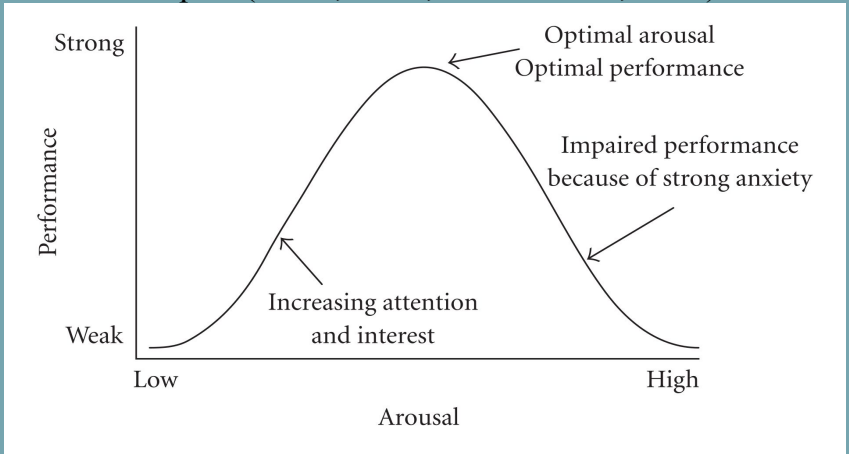


Dimensional Model:
Circumplex Model of Effect (Posner, Russell, & Peterson, 2005).

Motivation & Background

Flow

- Flow is the mental state of elatedness resulting from being energized and immersed into an activity while fully enjoying it (Czikzentimihay 1990).
    - Recently, the flow concept has been extended from a state defined for individuals to groups.
- Studies have gauged the “group flow,” which is when a group of people is in an extremely focused and productive state, of jazz musicians in the past (Gloor, Oster, & Fischbach, 2014).



Yerkes–Dodson curve for a difficult task (Diamond DM, et al. 2007).

Motivation & Background

Emotion Detection

- Gauging one's emotions is a heavily researched area, and has quite a few applications, like promoting safe driving, detecting lies, and detecting mental health patient's emotions in unobtrusive ways
- Emotion detection also varies based on the method:

Speech

Facial

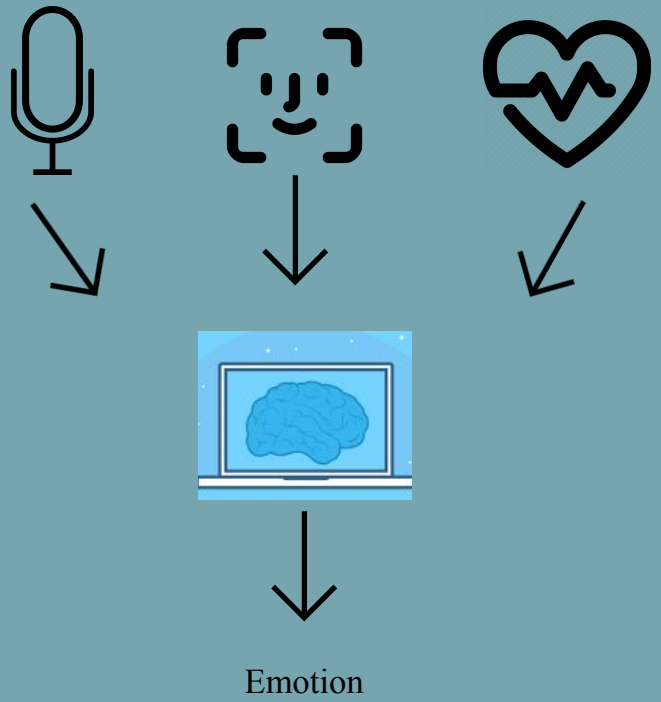
Physiological

- There are some limits with these methods, though
  - Speech and facial data can be skewed (Shu et. al, 2017)
  - Emotions are extremely complicated and using one method to gauge them would not be very accurate (Qiu, et. al, 2017)

Motivation & Background

Multimodal Methods

- To get around these limits, people have used multiple modalities
  - For example, using both facial and speech data to predict emotions
  - This produces more accurate results (Ullah et. al, 2017)
- There is still limited research in the dynamics between each modality, though.
  - For example, how do the predictions of speech, facial, and physiological emotions vary?
- This research area is called inter-modal dynamics.
- Research on this has been conducted on individuals between two groups (Gloor, Araño, & Guerrazzi, 2019), but there is limited research on the inter-modal dynamics of one group.



Motivation & Background

Research Question

How would the physiological signals and emotions of the performers relate to their facial emotions, and what does this tell us about their group flow?

Data Collection

Experimental Setup

- Collected facial and physiological data at four hour jazz rehearsal in the Berklee School of Music
- 30 performers (string players, percussionists, singers, and dancers) attended this rehearsal
- 10 of them had their data collected
    - Provided with smartwatches with the Happimeter app
    - Recorded with webcam

Data Collection

Data and methods

Facial Emotions

- The video of participants was fed through the Python face_recognition package
- Then, the faces were put into a CNN trained on four different datasets and pre-trained on ImageNet
- Emotions were Happy, Sad, Angry, Surprise, Neutral, and Fear
- ~1400 emotions recorded

Body Signals

- The Happimeter recorded accelerometer values, heart rate, and sound values of participants throughout the experiment
- Took absolute value of accelerometer values and added them together to get movement
- Took standard deviation of each participant's movement as well
- ~200 of each metric recorded

Physiological Emotions

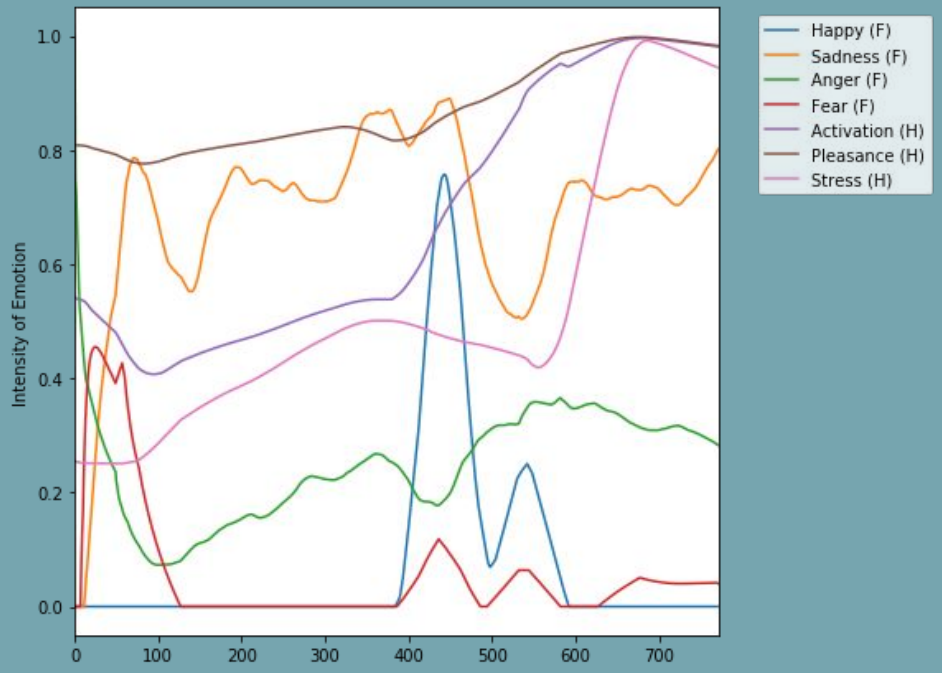
- We used the Happimeter's machine learning model to predict the intensity of the performers' levels of activation, pleasance, and stress (Budner et al., 2017).
- They were labeled with zero, one, or two
- ~300 of each emotion recorded per person

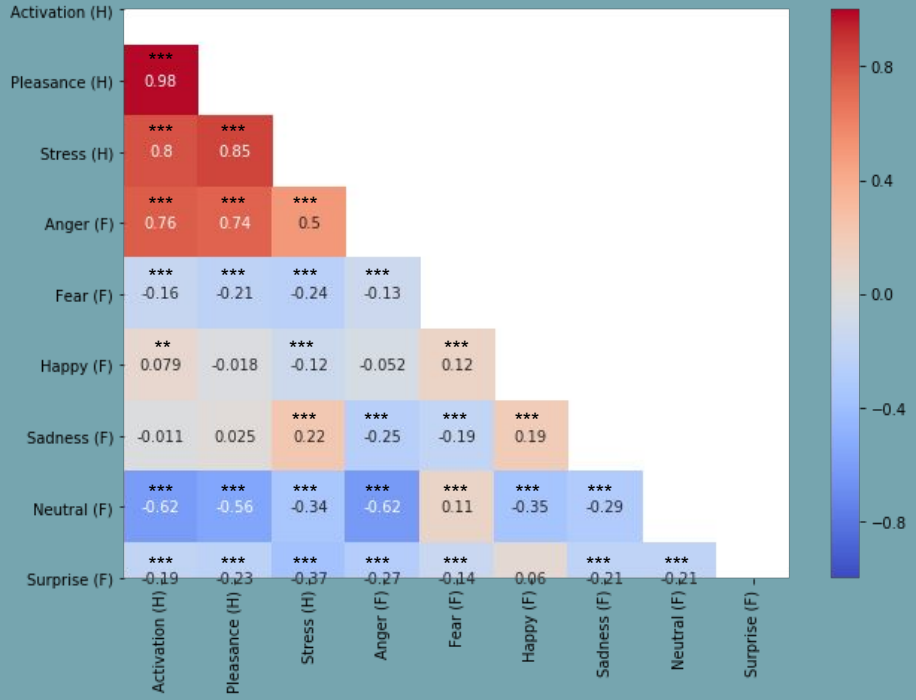
Results

Correlations: PER vs FER

F: FER, H: Happimetter emotion

(* <0.05, ** <0.01, *** <0.001).





Discussion

Correlations: PER vs FER

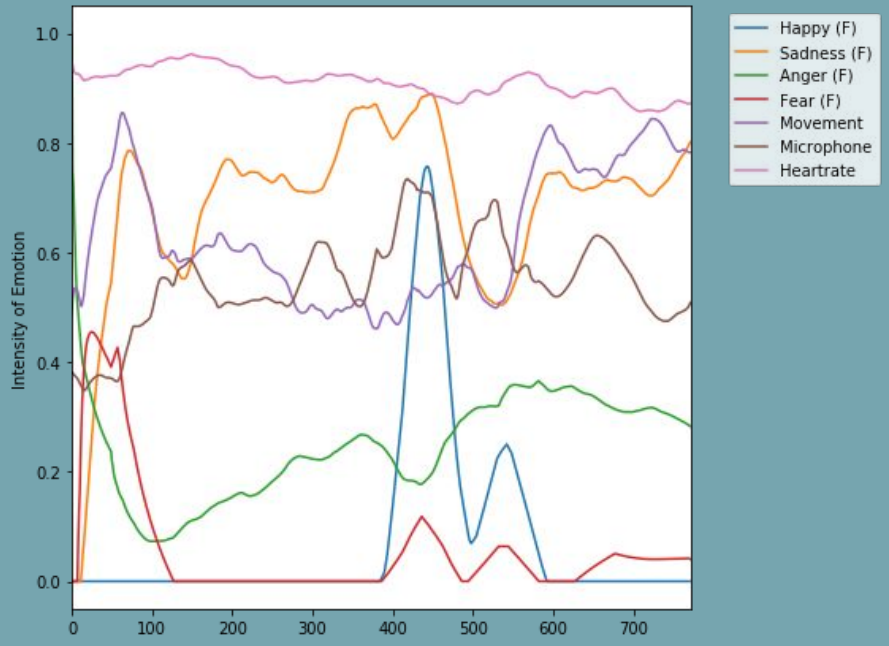
- Activation, pleasance, and stress all have strong, significant correlations ($r > 0.75^{***}$), reflecting the trends seen in the graph, as all rise throughout the performance
    - This can be an indicator of group flow, which has also been called “eustress” or positive stress
- An interesting result is the very weak and insignificant correlation between pleasance and happiness, and the strong, positive correlation between anger, and activation, pleasance, and stress, ($r > 0.50^{***}$)
    - This could be because the FER captured angry, instead of happy, faces while the musicians were experiencing eustress.
- Other interesting results are:
    - The negative correlation neutrality had with activation, pleasance, stress, anger, and happiness ($-.34^{***} < r < -.62^{***}$)
    - Fear had a weak negative correlation with activation, pleasance, stress, and anger (r was between -0.13^{***} and -0.24^{***}). This could suggest that as fear decreased, the group flow increased, hinting at a negative relationship between the two.

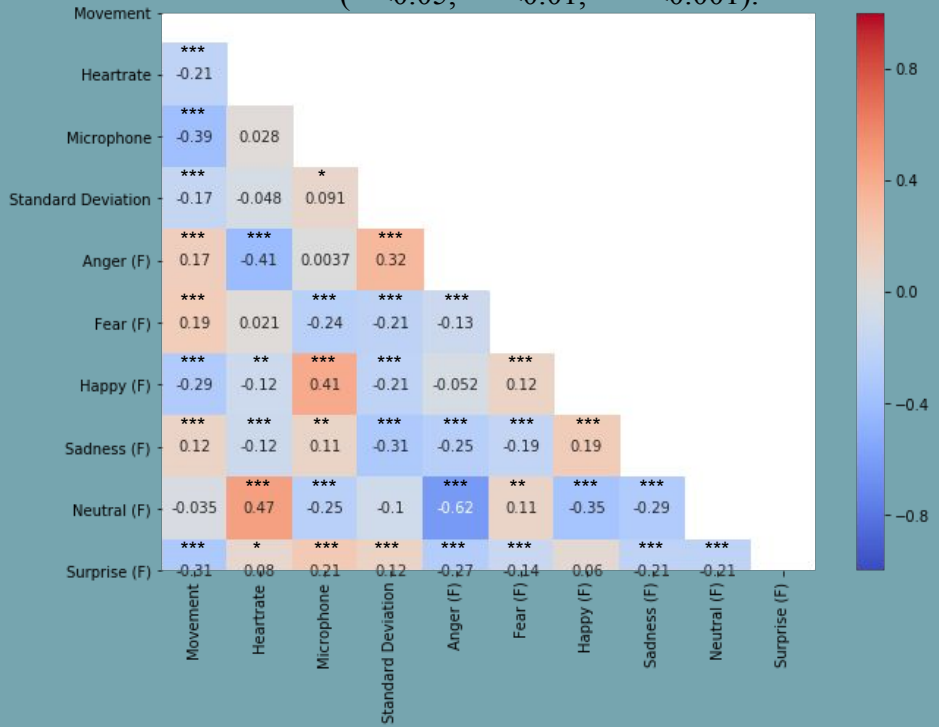
Results

Correlations: FER vs Physiological Signals

F: FER, H: Happimetter emotion

(* <0.05, ** <0.01, *** <0.001).





Discussion

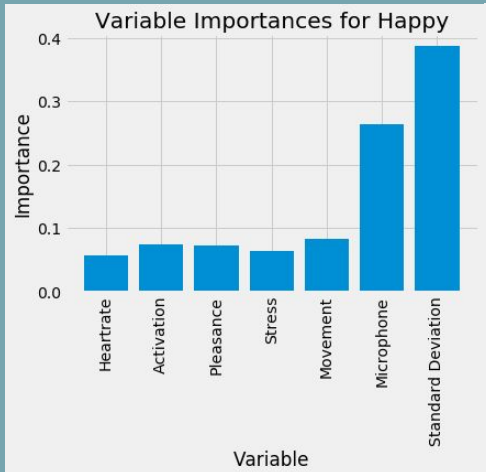
Correlations:FER vs Physiological Signals

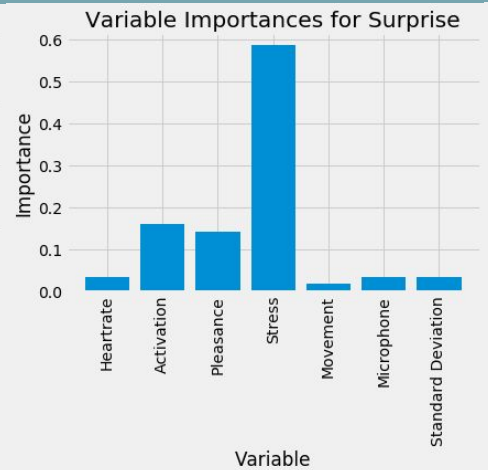
- Heart Rate tended to increase as neutrality did ($r = 0.47^{***}$) and decrease as anger and happiness increased
  - $r = -0.41^{***}$ and -0.12^{**}, respectively
- The volume of the performance increased as the surprise and happiness levels increased as well
  - $r = 0.21^{***}$ and 0.41^{***}
- On the other hand, fear and neutrality decreased as the volume increased
  - $r \approx -0.25^{***}$ for both
- The standard deviation of the performers' movement had negative correlations with their fear, happiness, and sadness levels

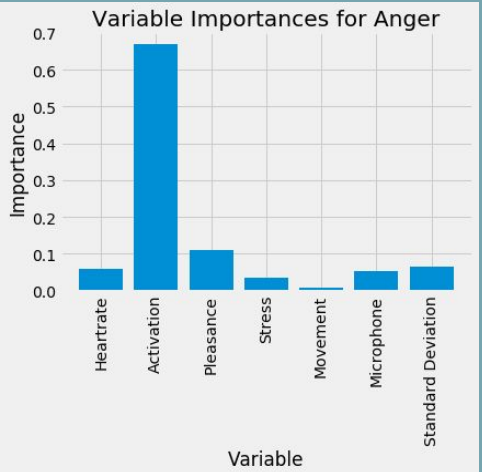
Results

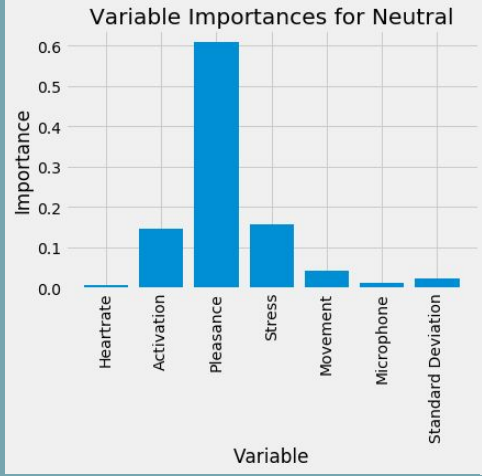
Random Forest Regression

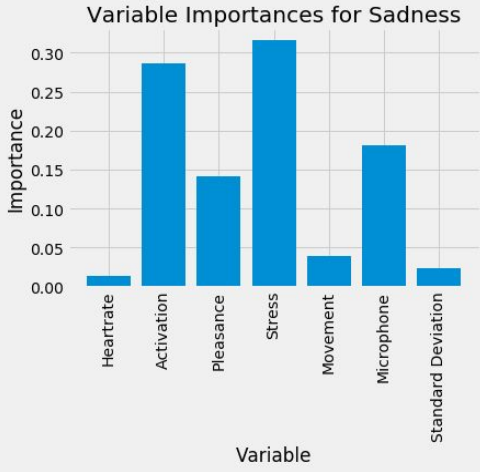
Emotion Predicted	R-Square	Mean Absolute Error
Anger	0.9671	0.00596
Fear	0.9381	0.00078
Happiness	0.9866	0.00069
Sadness	0.9849	0.00545
Neutrality	0.9741	0.00548
Surprise	0.9688	0.00665

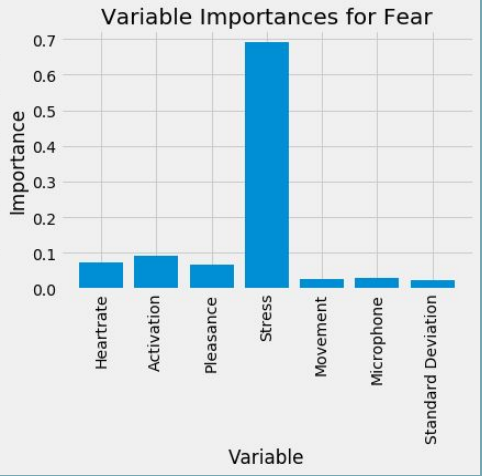












Conclusion

- There are many significant relationships between PER, FER, and physiological signals
  - “Anger” might not really be anger, but a so far not classified new emotion tentatively called “flow” driving eustress and group flow

- The random forest regression lets us see some non-linear relationships between emotions and physiological signals, as well, giving us insight into important interactions

- For further research, we should look into the anger facial emotion and what it actually represents
  - Potentially adding in more emotions to the FER model, like disgust, could deal with this

- Overall, this paper still addresses the relationships between FER, PER, and physiological signals alongside looking into flow, which could be helpful when it comes to virtual mirroring for the performers
  - With insight into these things, they could learn to increase their flow and have more productive rehearsals