

Computational Affective Science:

Understanding How Emotion Shapes the Mind, Body, and World

Saif M. Mohammad

National Research Council Canada

✉ uvgotaif@gmail.com

🐦 [@SaifMMohammad](https://twitter.com/SaifMMohammad)



NLP for Affective Science:

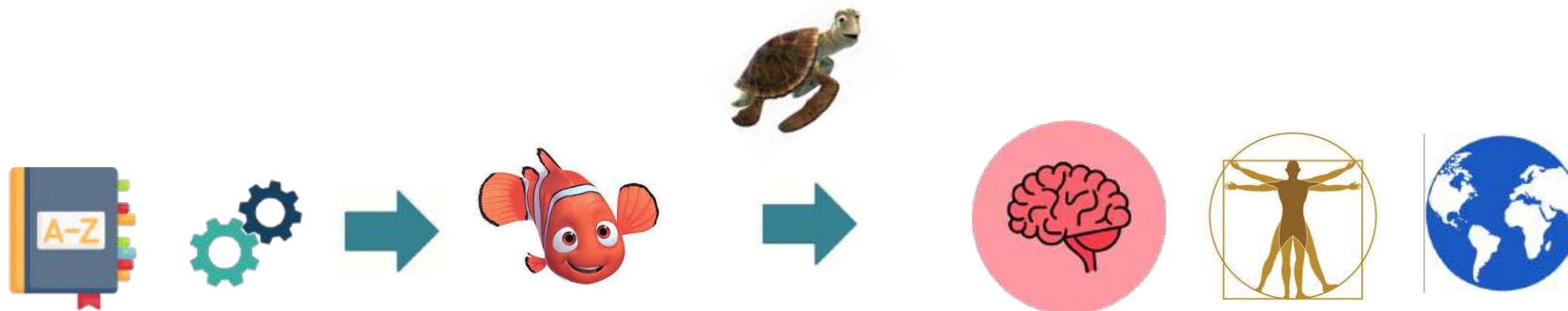
A window into emotions through language and computation



National Research
Council Canada

Conseil national de
recherches Canada

Canada



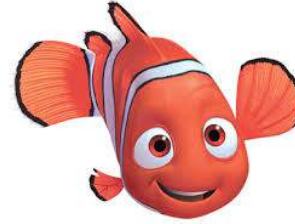
NLP for Affective Science:

A window into emotions (mind, body, health, and behavior) through language and computation

This Talk



- Core Theories of Affect and Emotion
 - What is Computational Affective Science?
 - Why it matters?
- Overview of the Areas of Work in CAS
 1. nature of affect; affect and the mind, body, world
 2. affective data
 3. affective tasks
 4. affective ethics
- Case Studies Exploring CAS Research



What are Affect and Emotions?

to start with...

Affect: the basic sense of feeling

Emotions: joy, sadness, fear, anger, etc.

Psychological Theories of Emotions

ON THE ORIGIN OF SPECIES

BY MEANS OF NATURAL SELECTION,

OR THE

PRESERVATION OF FAVOURED RACES IN THE STRUGGLE

FOR LIFE

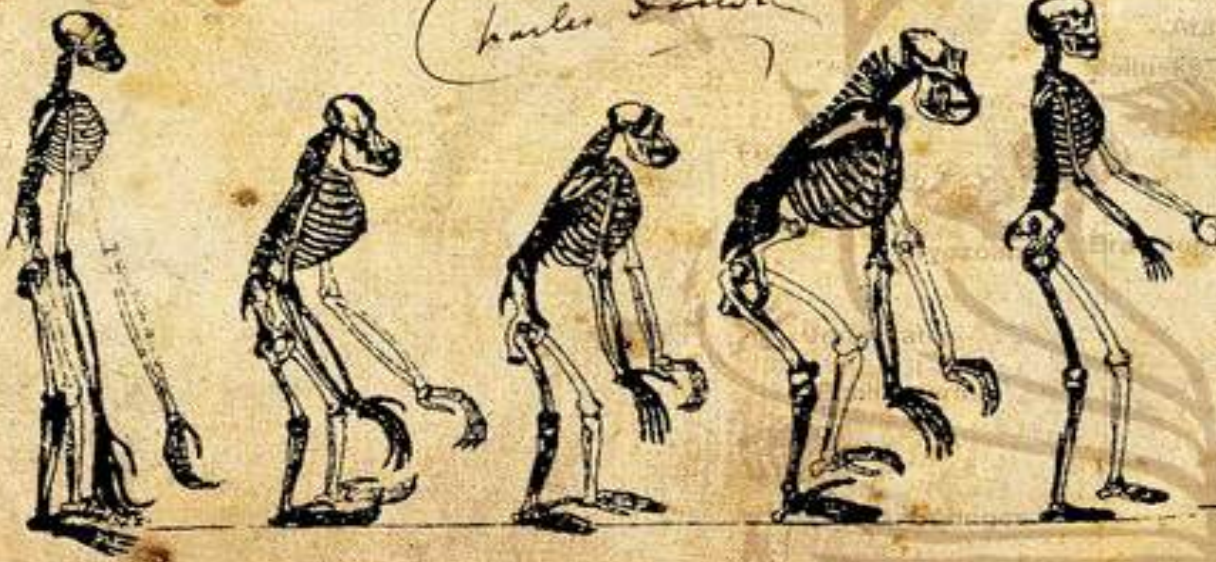


By CHARLES DARWIN, M.A.

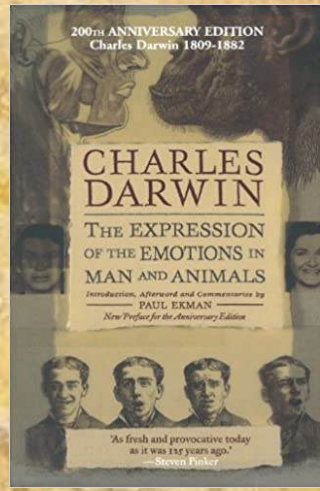
Charles Darwin

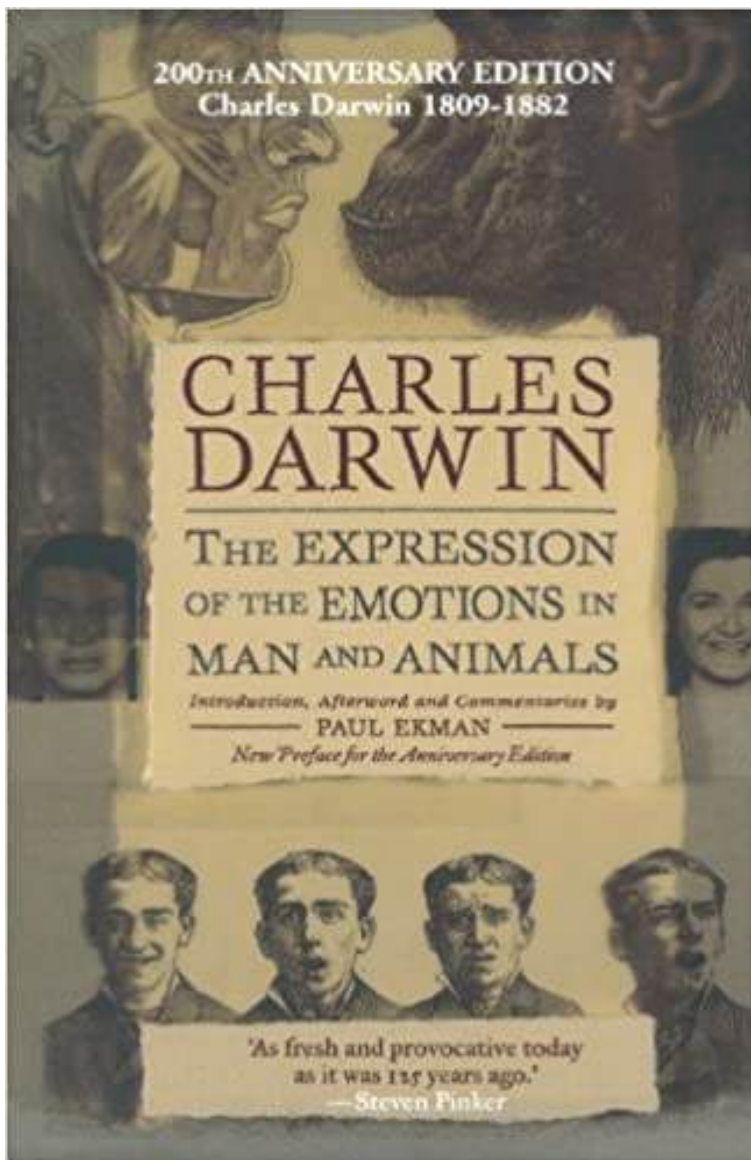


*I think
the letters A & B show
the position of the
front pedicle, B & D
with the position of the
the same could be
found. - heavy white*



Gibbon Orangutan Chimpanzee Gorilla Man





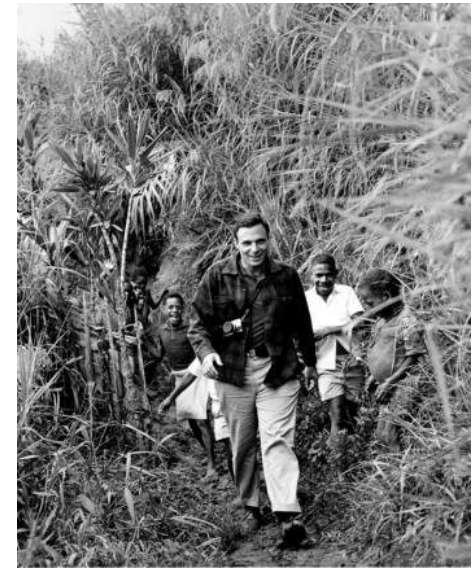
Theories of Emotion



Margaret Mead
Cultural anthropologist



Paul Ekman, Psychologist



- Mead, 1950s: culture determines emotion
- Paul Ekman, 1971: **Six** Universal Basic Emotions
 - Plutchik, 1980: **Eight** Basic Emotions
 - And many others



Plutchik's Emotion Wheel
Image credit: Julia Belyanevych

Core Dimensions of Connotative Meaning

Influential factor analysis studies (Osgood et al., 1957; Russell, 1980, 2003) have shown that the three most important, largely independent, dimensions of word meaning:

- **valence (V)**: positive/pleasure – negative/displeasure
- **arousal (A)**: active/stimulated – sluggish/bored
- **dominance (D)**: powerful/strong – powerless/weak

Thus, when comparing the meanings of two words, we can compare their V, A, D scores. For example:

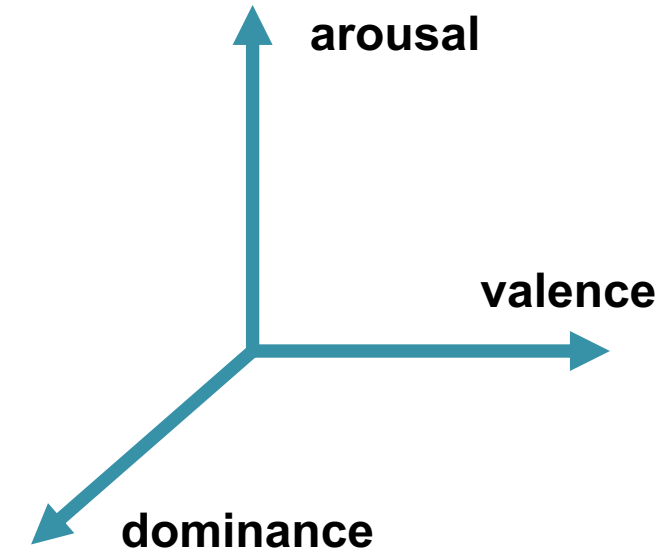
- *banquet* indicates more positiveness than *funeral*
- *nervous* indicates more arousal than *lazy*
- *queen* indicates more dominance than *delicate*



Osgood



Russell



Theories of Emotion



Margaret Mead
Cultural anthropologist



Paul Ekman
Psychologist and discoverer
of micro expressions.



Lisa Barrett
University Distinguished
Professor of Psychology,
Northeastern University

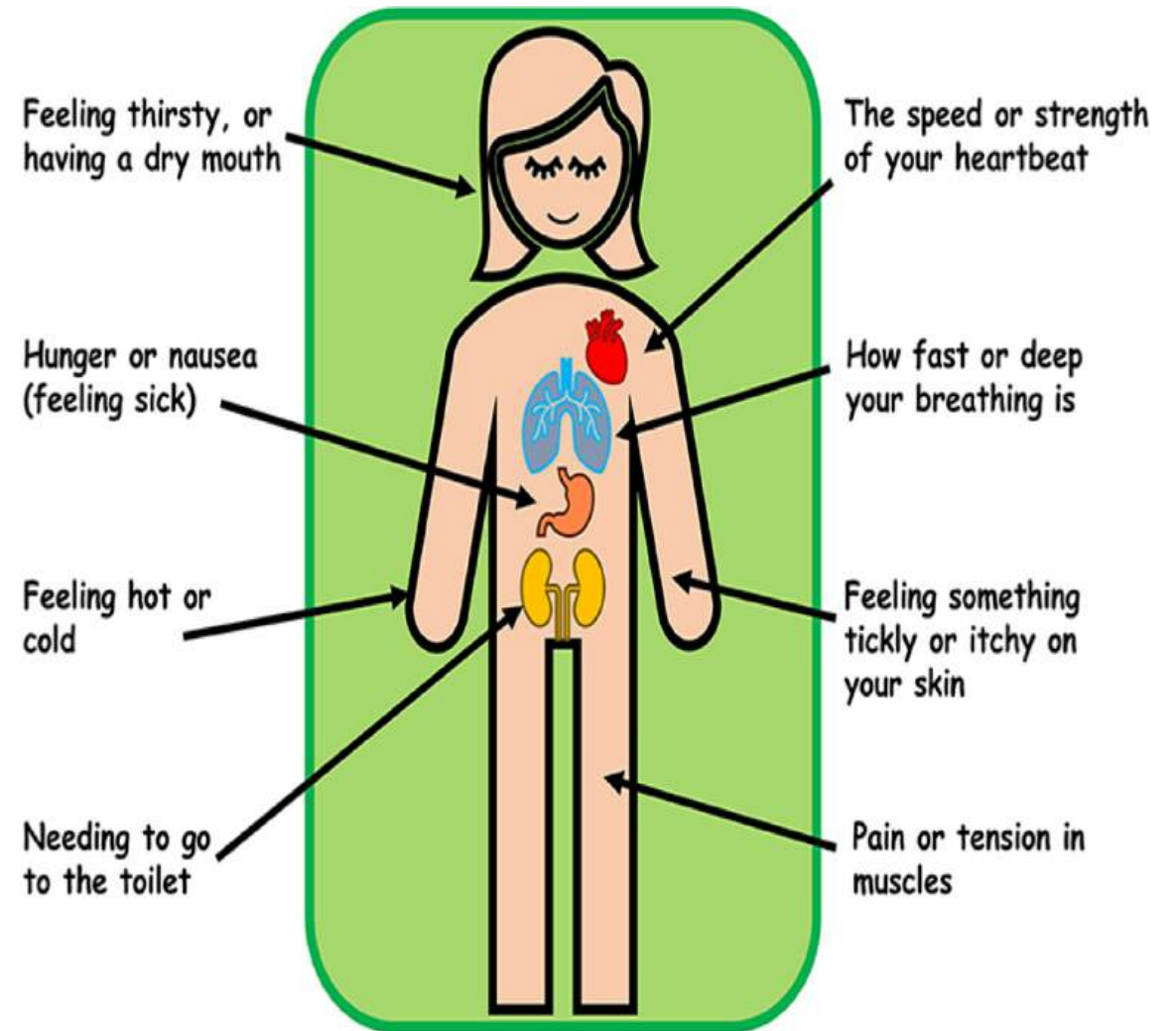
Theory of Constructed Emotion (Barrett, 2017)

- the brain **constructs** emotions
- important tenets of BET discredited (“basic” emotions)
- stress on variability




Interoception

the ability to sense signals about your bodily state

Interoceptive signals
sensory representations of the interior of the body (viscera)



Constructed Emotion

External Situation	Bodily Signal	Possible Emotions
 Public speaking in front of an audience	Heart races Palms sweat Stomach churns	Anxiety Excitement
 Walking alone at night and hearing footsteps behind you	Muscles tense Heart rate spikes Breathing shallow	Fear Apprehension Hypervigilance
 Sitting alone at café	Slow breathing Relaxed muscles	Loneliness Contentment

Affect

The basic sense of feeling:

- Transduced and summarized from **interoceptive signals**
 - sensory representations of the interior of the body (viscera)
- A feature of consciousness
 - occurs in every moment (whether you're aware of it or not)
- Key dimensions
 - **Valence**: displeasure to pleasure
 - **Arousal**: idle/sluggish to activated
 - **Dominance**: weak/loss of control to strong/having a sense of control

Emotions

Constructed by the brain using:

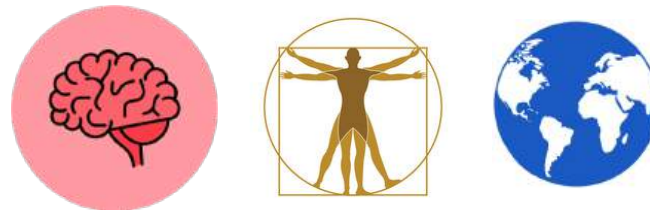
- affective and interoceptive signals
- "emotion concepts" from one's culture learned through socialization
- predictive coding

The brain categorizes the continuous affect into discrete categories (analogous to color perception)

- joy, sadness, fear, anger, etc.

Why Emotions and Affect Matter

- Determine human experience and behavior
- Condition our actions
- Central in organizing meaning
 - No cognition without emotion
- A window into
 - understanding our mind (cognition), body (health, well-being), how we make sense and interact with the world
 - the evolutionary forces that shaped us

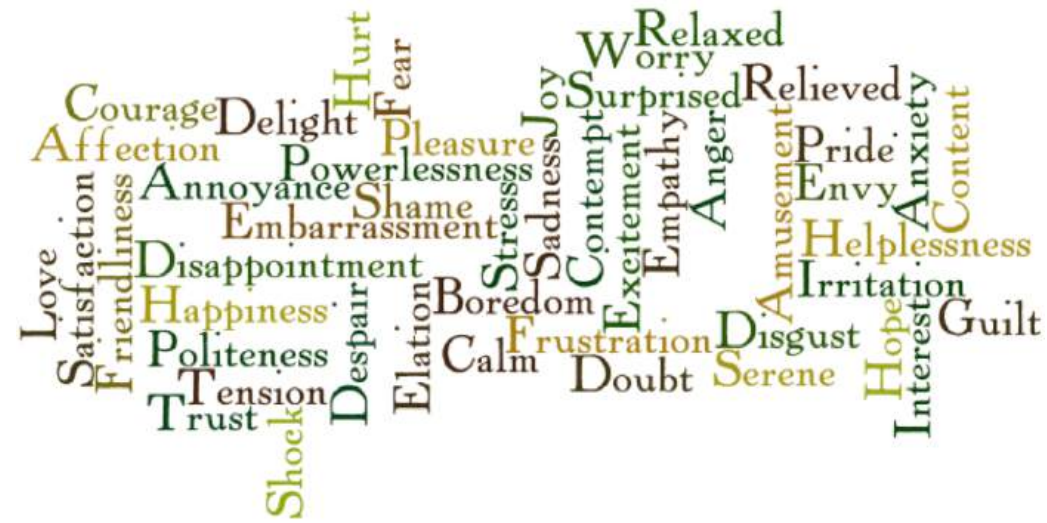


Affective Science

Interdisciplinary field focused on understanding emotions and affect

- How do affect and emotions work?
 - affective and emotional processes
 - affective neuroscience
 - emotion regulation
- How do they impact our mental health, physical health, and behaviour?
 - wellbeing, emotion-related disorders
- What agency do we have in managing our emotions?

The Language of Emotions



Language is a powerful way of expressing emotions

- can express numerous emotional shades
 - terms with fuzzy boundaries, overlapping meanings, socio-cultural influences, etc.
- usually conveyed by connotation (and not denotation)
 - can be subtle, direct, ambiguous, deceptive
 - can be creative
 - can be conscious expression or subconscious manifestations

Computational Affective Science (CAS) aka NLP for Affective Science

Computational Analysis of Emotions Through Language

- Challenging
 - see previous slide on language
- Powerful
 - makes use of large amounts of text
 - simple to complex NLP techniques
 - language impacts thought and how we construct emotions
linguistic relativity aka Sapir–Whorf hypothesis
- Complementary view to traditional affective science approaches in psychology
 - makes use of, complementary, ecologically valid data

This Talk



- Core Theories of Affect and Emotion
 - What is Computational Affective Science?
 - Why it matters?
- Overview of the Areas of Work in CAS
 1. nature of affect; affect and the mind, body, world
 2. affective data
 3. affective tasks
 4. affective ethics
- Case Studies Exploring CAS Research

CAS: Areas of Research

1. The Nature of Affect



The Relationship of Affect with:

a. the mind, cognition



b. the body



c. the world we interact with

2. Affective Data and Resources



3. Affective Tasks



4. Affective Ethics



This Talk



- Core Theories of Affect and Emotion
 - What is Computational Affective Science?
 - Why it matters?
- Overview of the Areas of Work in CAS
 1. nature of affect; affect and the mind, body, world
 2. affective data
 3. affective tasks
 4. affective ethics
- Case Studies Exploring CAS Research

This Talk



- Core Theories of Affect and Emotion
 - What is Computational Affective Science?
 - Why it matters?
- Overview of the Areas of Work in CAS
 1. nature of affect; affect and the mind, body, world
 2. affective data
 3. affective tasks
 4. affective ethics
- **Case Studies** exploring CAS Research

This Talk



- Core Theories of Affect and Emotion
 - What is Computational Affective Science?
 - Why it matters?
- Overview of the Areas of Work in CAS
 1. nature of affect; affect and the mind, body, world
 2. affective data
 - a. word-emotion datasets
 - b. sentence/post-emotion datasets
 3. affective tasks
 4. affective ethics
- **Case Studies** Exploring CAS Research



word --



2a: Word-Emotion Association Lexicons

Affect Datasets for Words

Lexicons for both **categorical emotions** as well as for **valence, arousal, and dominance**

- Lists of words associated with joy, sadness, fear, etc.
- Lists of words and their valence, arousal, and dominance scores

The NRC Valence, Arousal, and Dominance Lexicon (2018)

provides ratings of valence, arousal, and dominance for ~20,000 English words

<http://saifmohammad.com/WebPages/nrc-vad.html>

The NRC Word–Emotion Association Lexicon aka NRC Emotion Lexicon or EmoLex (2010)

provides associations for ~14,000 words with eight emotions

(anger, fear, joy, sadness,
anticipation, disgust, surprise, trust)

<http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

The NRC Emotion Intensity Lexicon aka Affect Intensity Lexicon (2018-19)

provides intensity scores for ~6000 words with four emotions

(anger, fear, joy, sadness)

<http://saifmohammad.com/WebPages/AffectIntensity.htm>

The NRC Word–Colour Association Lexicon (2010)

provides associations for ~14,000 words with 11 common colours

<http://saifmohammad.com/WebPages/lexicons.html>

Use of The Emotion Lexicons

- For research by the scientific community
 - Computational linguistics, psychology, digital humanities, robotics, public health research, social science, etc.
- To analyze text
 - Brexit tweets, Radiohead songs, Trump tweets, election debates,...
 - **Wishing Wall**, uses the NRC Emotion lexicon to visualize wishes.
Displayed in:
 - Barbican Centre, London, England, 2014
 - Tekniska Museet, Stockholm, Sweden, 2014
 - Onassis Cultural Centre, Athens, Greece, 2015
 - Zorlu Centre, Istanbul, Turkey, 2016
- In commercial applications





word --



Word-Emotion Association Lexicons

New Work...

New Dimensions

2024

WorryWords

- Anxiety-Calmness Lexicon

Theoretical Framework: Circumplex Model (Russell, 1980)

2025



The NRC Warmth and Competence Lexicon

- Warmth-Coldness
 - Sociableness-Unsociableness
 - Trustworthiness-Untrustworthiness
- Competence-Incompetence

Theoretical Framework: Stereotype Content Model (SCM), Social Cognition Theory (Fiske et al. 2002)

2026

The NRC MFT Lexicon

- Care-Harm
- Fairness-Cheating
- Loyalty-Betrayal
- Authority / Subversion
- Sanctity / Degradation

Theoretical Framework: Moral Foundation Theory (MFT) (Haidt, 2007)

Crowdsourcing Questionnaire

Consider trustworthiness to be a broad category that includes:
trustworthy, honesty, fairness, dependability, reliability, morality, virtuousness, sincerity, honorableness, etc.

Consider untrustworthiness to be a category that includes:
unfairness, dishonesty, untrustworthiness, dubiousness, immorality, sinfulness, insincerity, dishonorableness, etc.

Q1. <term> is often associated with feeling:

- | | |
|---|------------------------------|
| 3: very trustworthy | -1: slightly untrustworthy |
| 2: moderately trustworthy | -2: moderately untrustworthy |
| 1: slightly trustworthy | -3: very untrustworthy |
| 0: not associated with being trustworthy or untrustworthy | |
-

Crowdsourcing Questionnaire

Consider social warmth to be a broad category that includes:
*warmness, sociableness, generosity, helpfulness,
tolerance, understanding, thoughtfulness, etc.*

Consider social coldness to be a broad category that includes:
*coldness, antisocialness, ungenerosity, unhelpfulness,
intolerance, indifferent, thoughtlessness, etc.*

Q1. <term> is often associated with feeling:

- | | |
|---|---------------------------|
| 3: very sociable | -1: slightly unsociable |
| 2: moderately sociable | -2: moderately unsociable |
| 1: slightly sociable | -3: very unsociable |
| 0: not associated with being sociable or unsociable | |
-

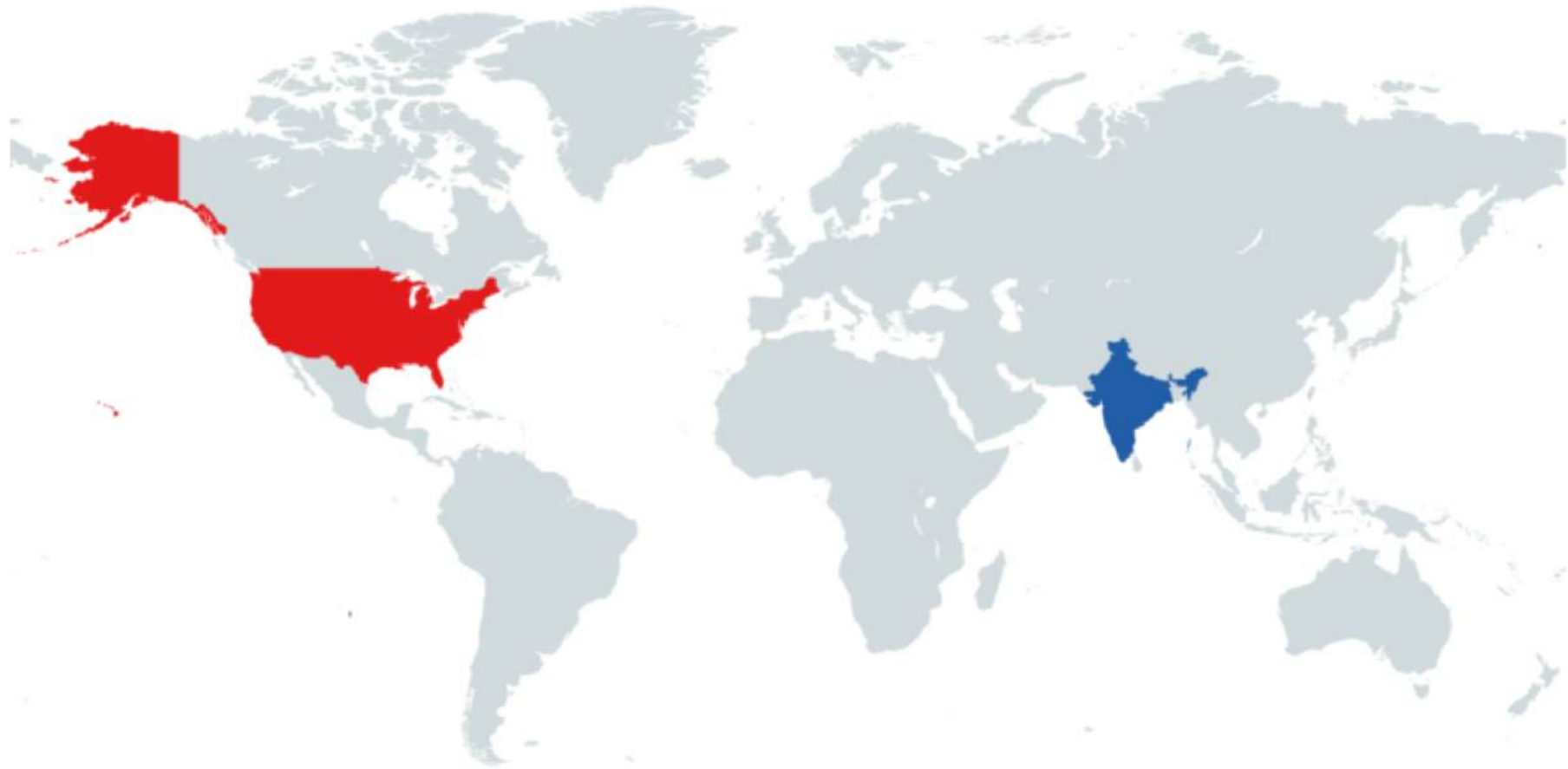
More Coverage

- 44k Unigrams
- 20k Multiword Expressions (MWEs)



Locations

USA, India



Created with mapchart.net

Some Things We Learned About Emotions

- Variability
 - cultural
 - context
 - person-to-person
 - what/how you ask about emotions
 - more agreement for valence; less for fear, sadness, etc.
 - associated vs. evoked
 - greater agreement with comparative annotations
- Substantial common ground (SHR over 0.8)

Draw inferences at aggregate level



- determine broad trends



Warmth And Competence

Warmth And Competence

Primary dimensions we assess people and groups on:

- Warmth (W) 
 - Sociability (S): friendliness, gregariousness, and conviviality 
 - Trust (T): morality, goodness, sincerity, and integrity 
- Competence (C) 
 - ability, power, dominance, and assertiveness

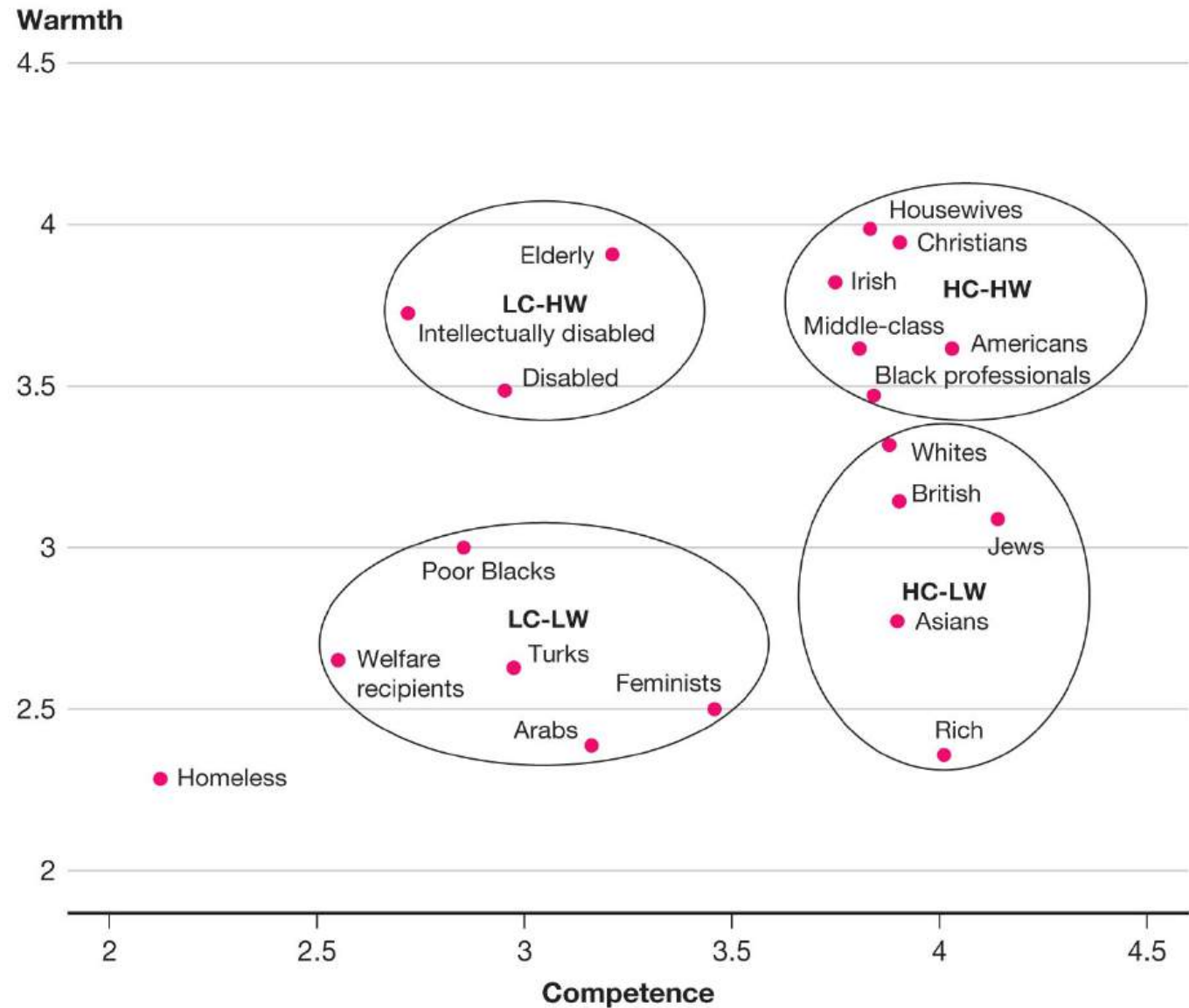
[Decades of social cognition and stereotype research. Notably by Susan Fiske and colleagues. Evolutionary benefits.]

American Stereotypes

Cuddy, Fiske, and Glick 2008

Tied to who is considered to be in their:

- ingroup
- outgroup



Why create language resources for WCTS?

In Psychology and Social Cognition

- What kind of WCTS assessments do children develop first?

In Computational Social Science, NLP

- Levels of WCTS in public discourse (climate change, vaccines, etc.)

In HCI and NLP

- Perceptions of WCST of people towards artificial agents

In Digital Humanities and NLP

- Role do WCTS in developing compelling characters and story arcs

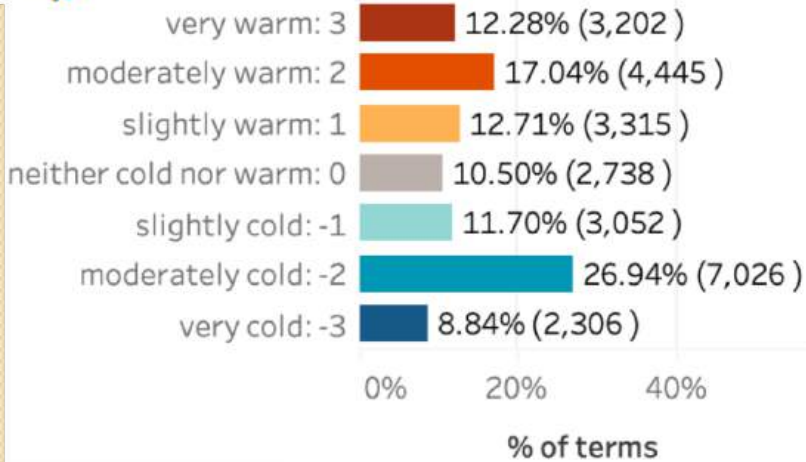
In Commerce

- Warmth, competence, trust, and sociability towards one's product on social media

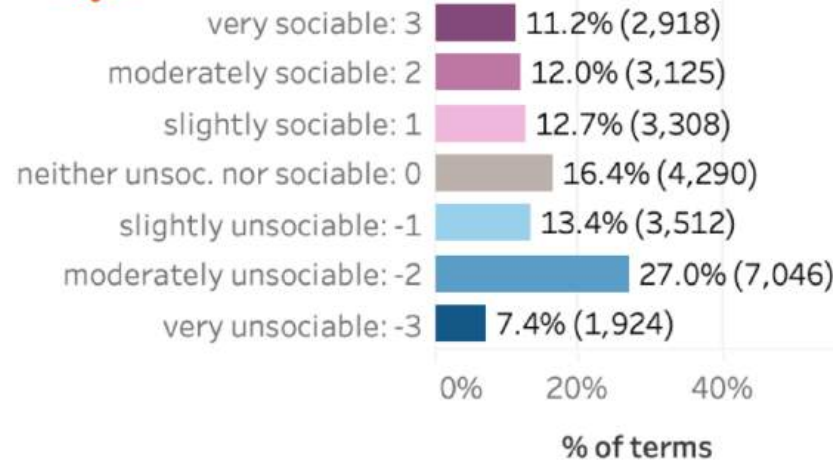
Class Distributions



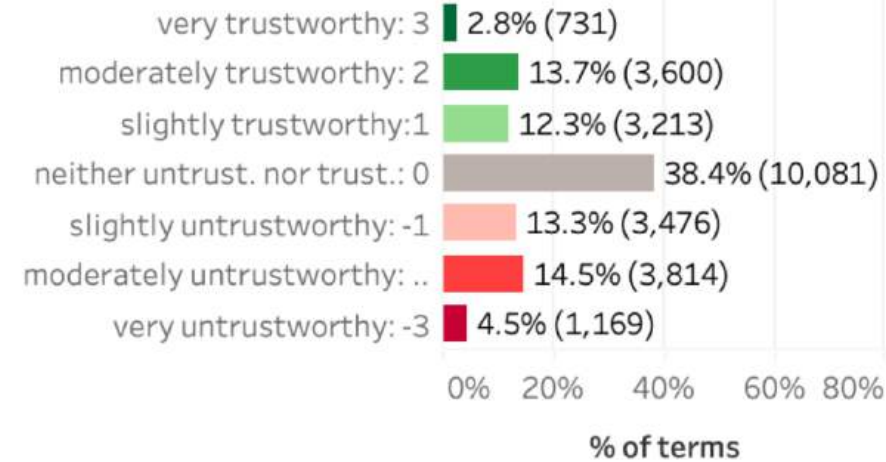
warmth (group)



sociability (group)



trust (group)





Experiments



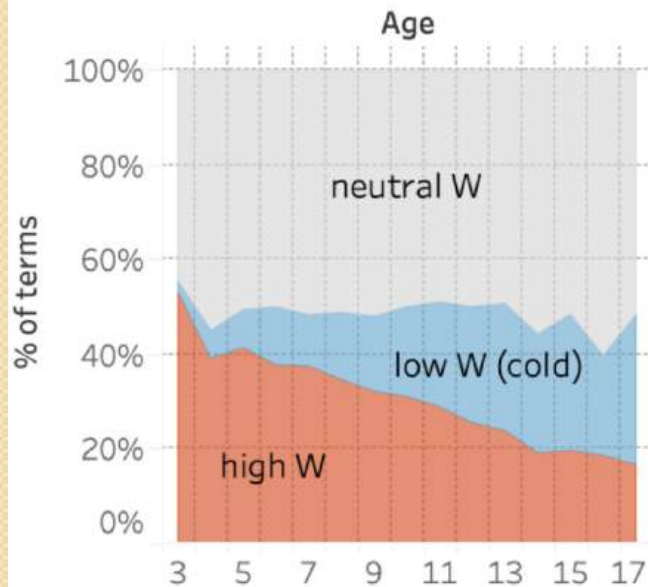
1. At what rate do children acquire WCST words?

Rate at which children acquire WCST words

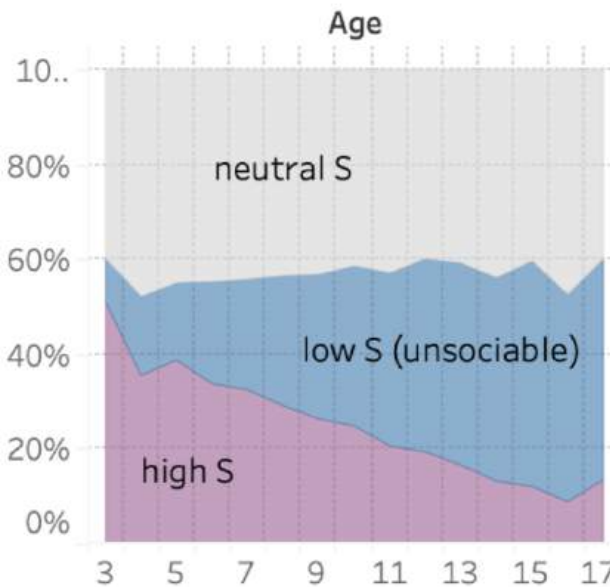


Graphs created using: Words Of Warmth, NRC VAD, Age of Acquisition lexicons.

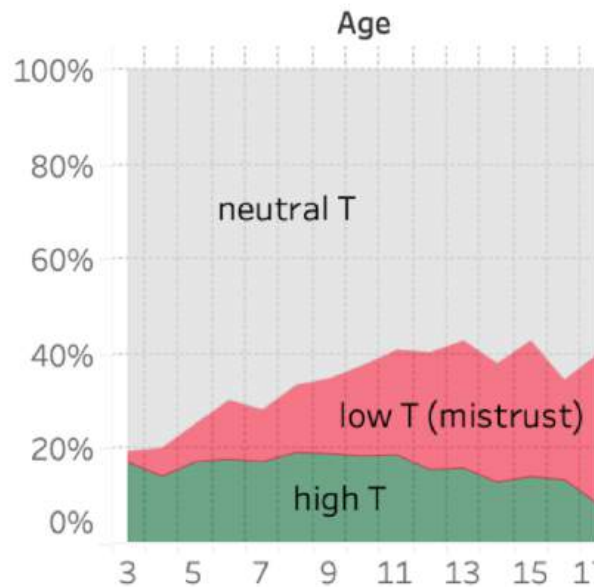
(a) warmth (W)



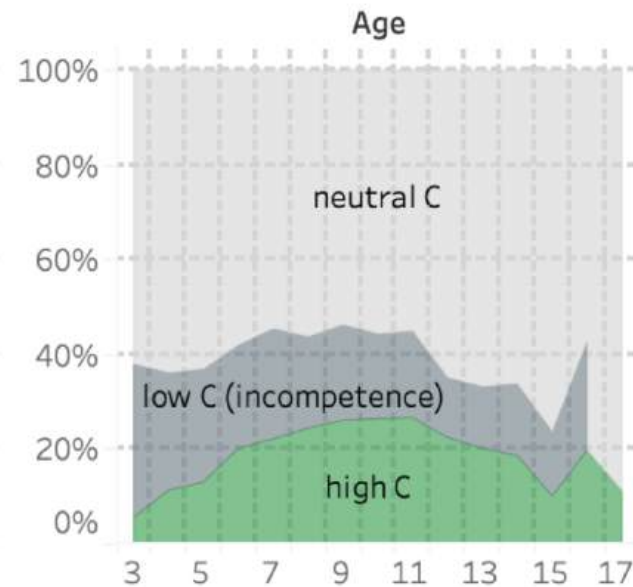
(b) sociability (S)



(c) trust (T)



(d) competence (C)



W/S/T scores -1.5 to 1.5 : neutral; ≤ -1.5 : low; ≥ 1.5 : high. C scores -0.33 to 0.33 : neutral; -1 to -0.33 : low C; 0.33 to 1 : high C.

- Higher % for polar W words vs. polar C words: consistent with the primacy of valence hypothesis (not primacy of C).
- Higher % for polar S words as opposed to polar T words in early years: S is more important than T (and morality).
- Among the polar words, the early years are marked with a greater % of high-WST words, as well as low-C words.



Experiments

2. Case Studies of W and C Stereotypes

Measuring Stereotypes

Windows into stereotype towards various targets: Two methods

- **Direct:** Direct Target Lookup in the WCTS lexicon
- **Co-terms:** Examining WCTS of terms co-occurring with the target terms in the TUSC dataset

American and Canadian geo-located posts on X from 2015 to 2021 (**Vishnubhotla and Mohammad, 2022**).

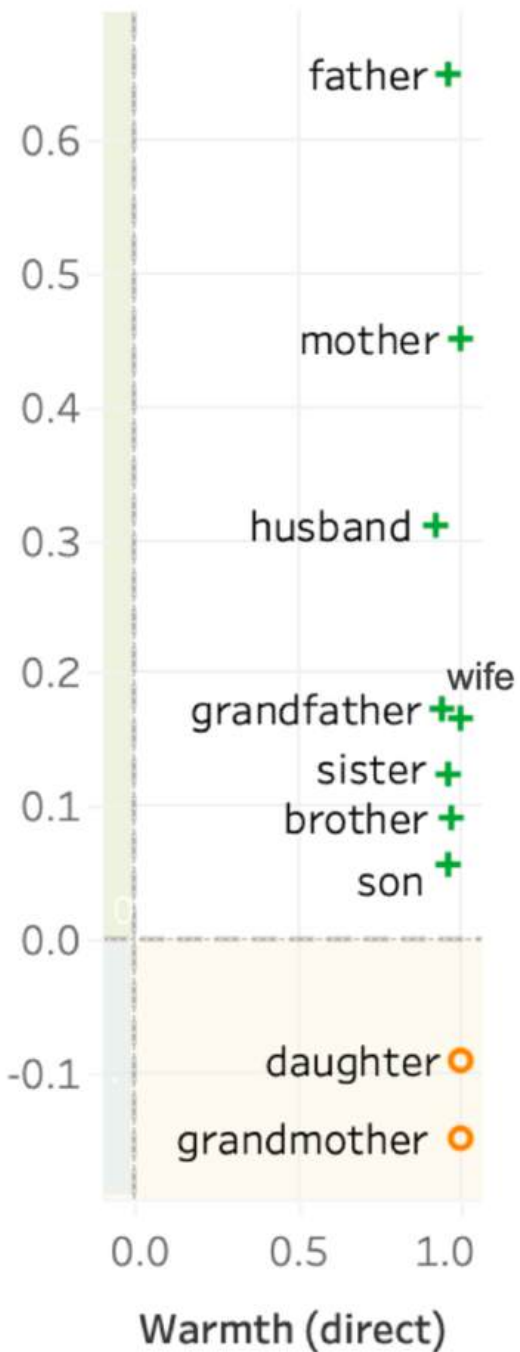
Average W and C scores of all 3.1 million posts in the corpus: 0.5001, 0.1370 (in a range from -1 to 1).

Ethical Considerations

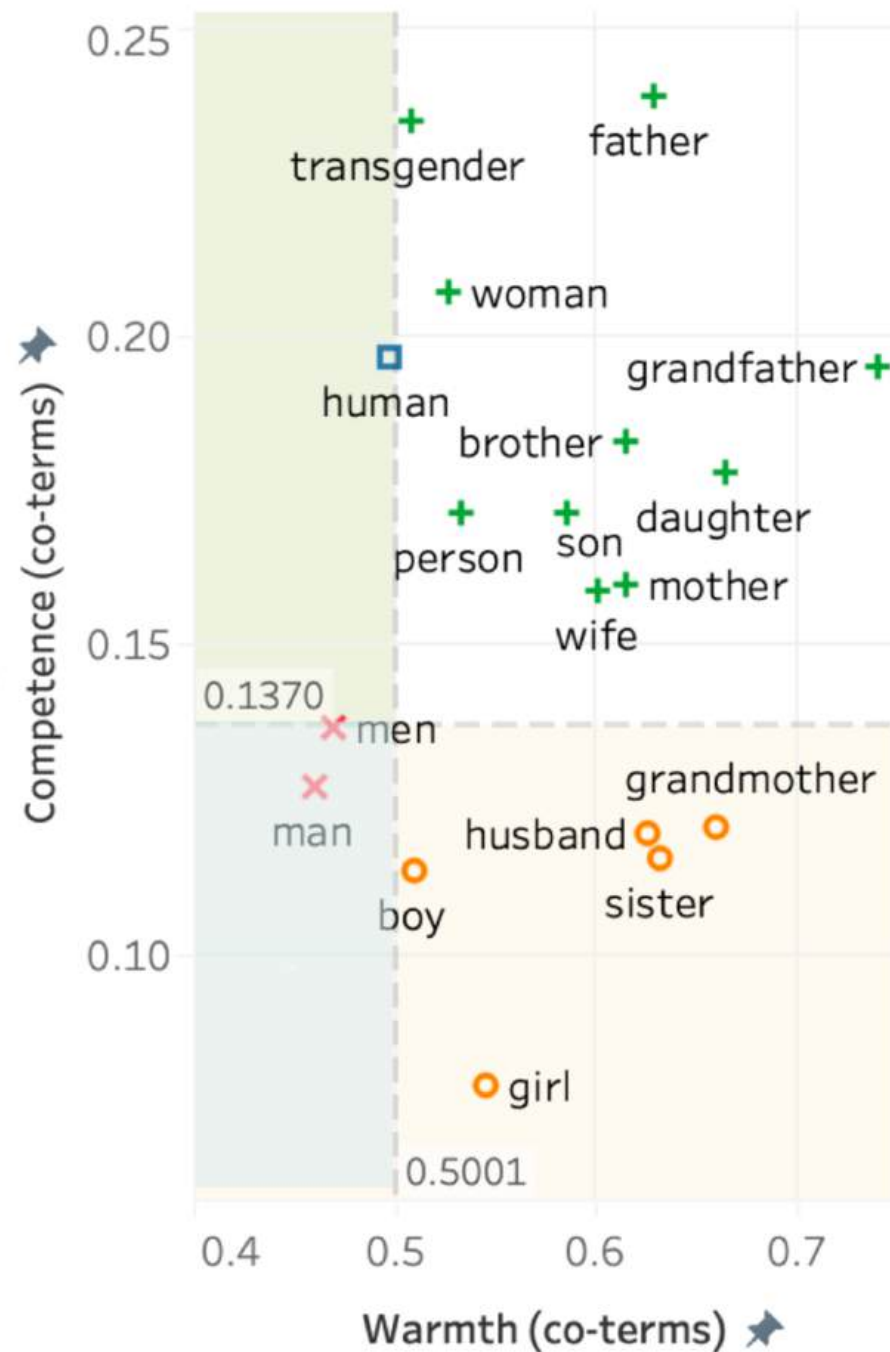


- Associations and stereotypes; not inherent properties
- Consider coverage, domain, ambiguity, socio-cultural effects, etc.
- Ethics Sheet for Emotion Recognition (Mohammad, 2022) [CL Journal]
- Best Practices in the Use of Emotion Lexicons (Mohammad, 2020)

a. gender (direct)



b. gender (cotermis)

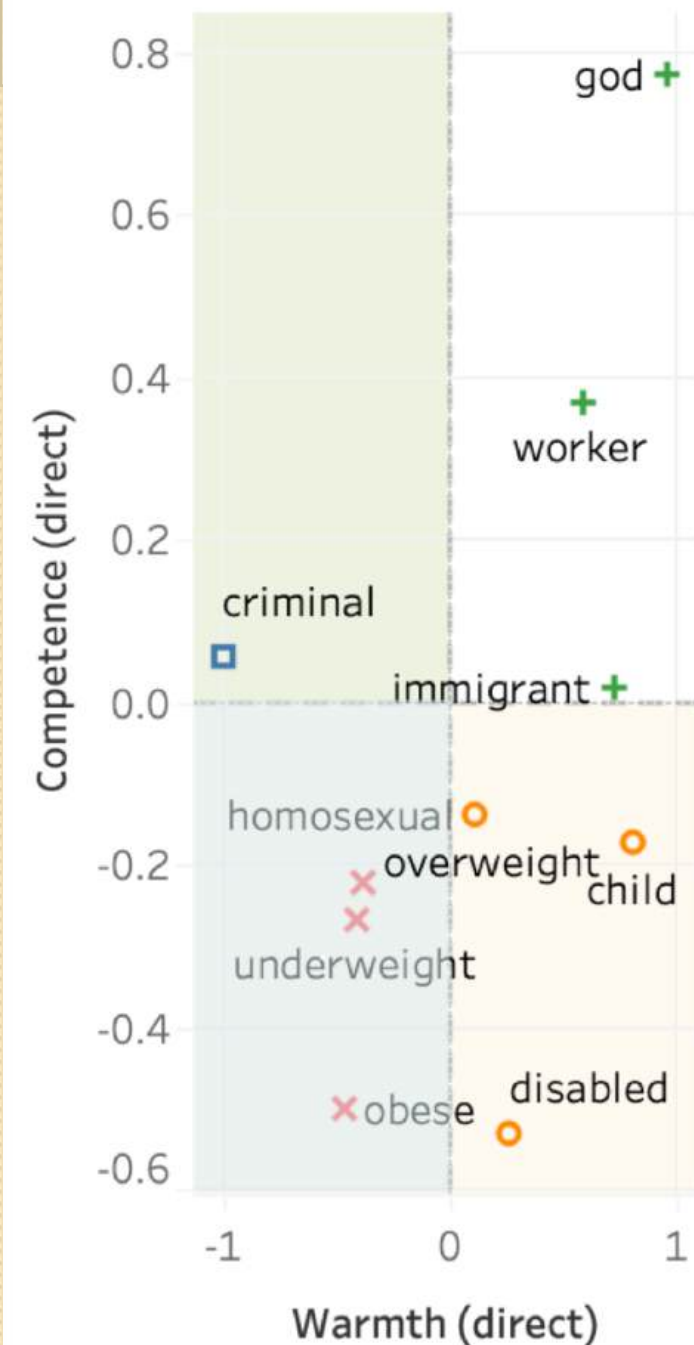


Relations

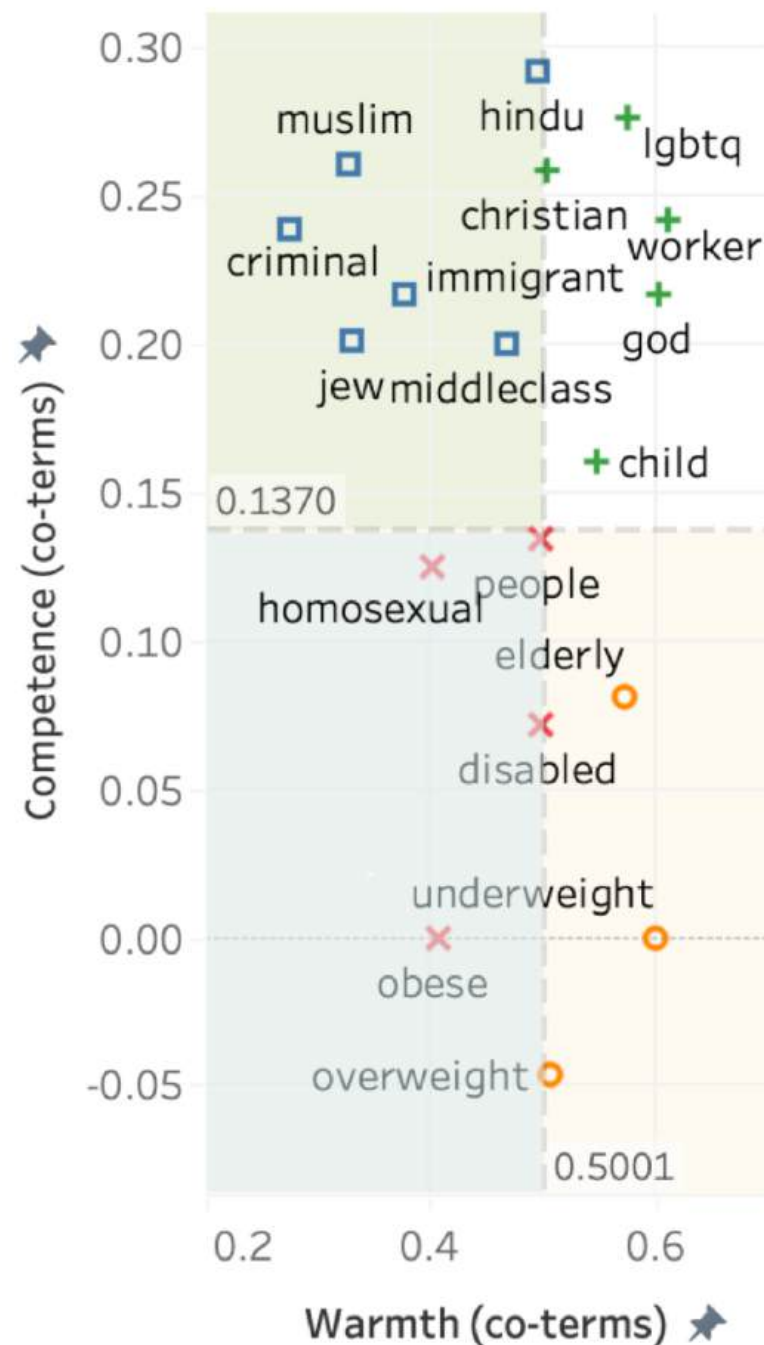


- Direct WC: people consider all these terms as high W; substantial variations in their C.
- In contrast, the co-term plots show that our language has marked differences for these terms for C and W.
- Clear gender stereotypes reflected in these scores

a. social groups (direct)



b. social groups (cotermis)

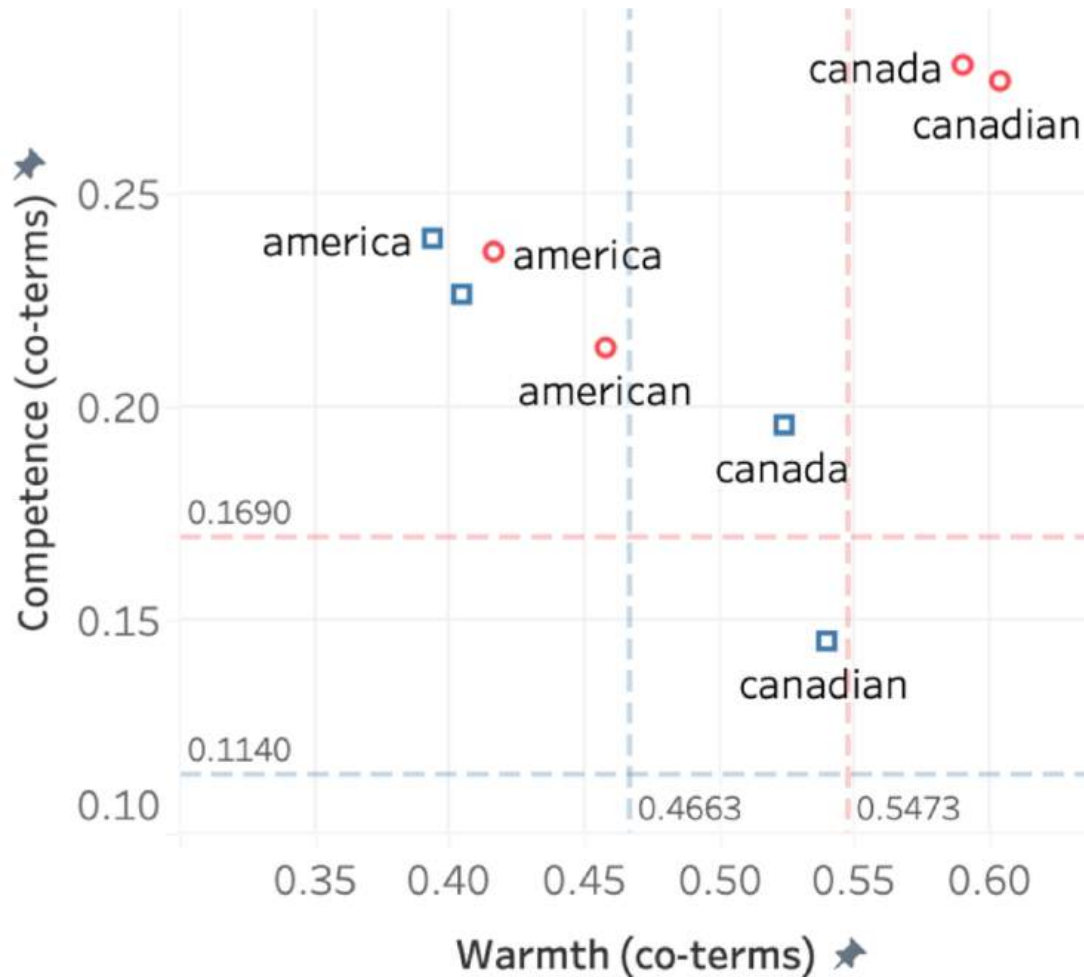


Social Groups



- *muslim, jew, immigrant*: low-W scores (consistent with known negative stereotypes in US, CA)
- *elderly, underweight*: low-C, high-W; *overweight*: even lower C score
- *obese*: low-W, low-C scores
- *god*: high direct W and C scores; but the discourse around *god* on X is such that the term gets lower co-terms-based C score than many other terms

In-group – Out-group Dynamics



The blue dashed lines indicate the average W and C scores of all posts by Americans and red dashed lines indicate the averages for Canadians.

- posts by Canadians in general have higher W and C scores than posts by Americans
- Canadians view themselves as more competent and much warmer than their neighbour (consistent with ingroup and out-group stereotypes)
- while Americans view themselves as more competent than Canadians, they too perceive Canadians as warmer (the *Canadians are nicer* stereotype overrides the outgroup stereotype)

This Talk



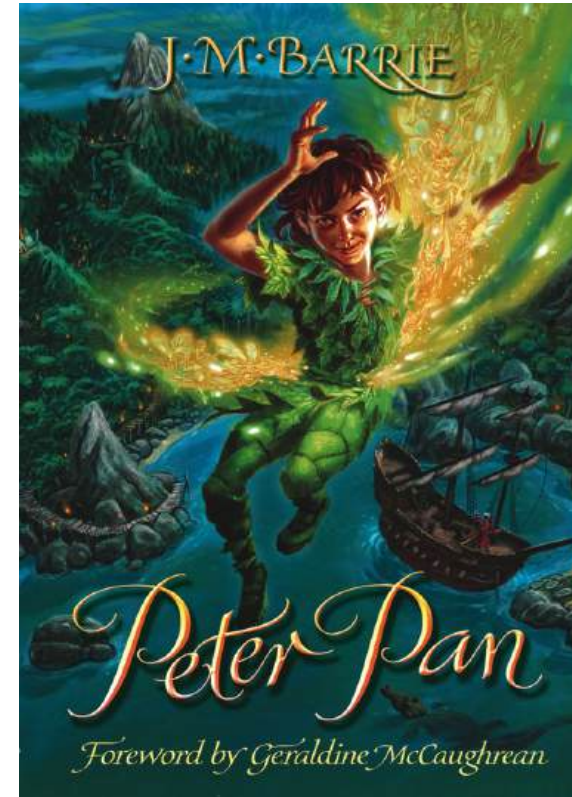
- Core Theories of Affect and Emotion
 - What is Computational Affective Science?
 - Why it matters?
- Overview of the Areas of Work in CAS
 1. **nature of affect**; affect and the mind, body, world
 2. affective data
 3. **affective tasks**
 - a. sentence/post-level classification/regression/generation
 - b. aggregate-level (emotion arcs, group comparisons, etc.)
 4. affective ethics
- **Case Studies Exploring CAS Research**



3b: Affective Tasks

aggregate-level (emotion arcs, group comparisons, etc.)

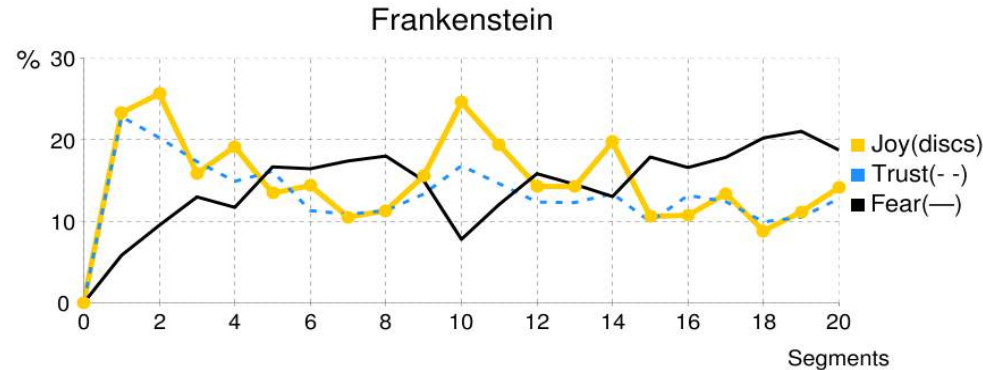
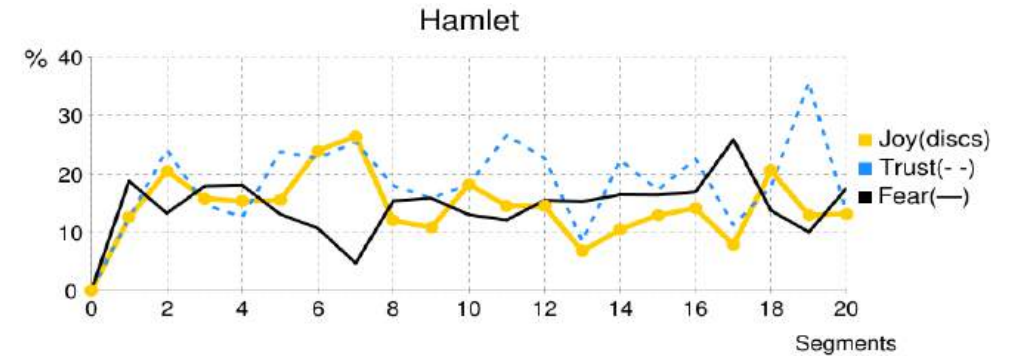
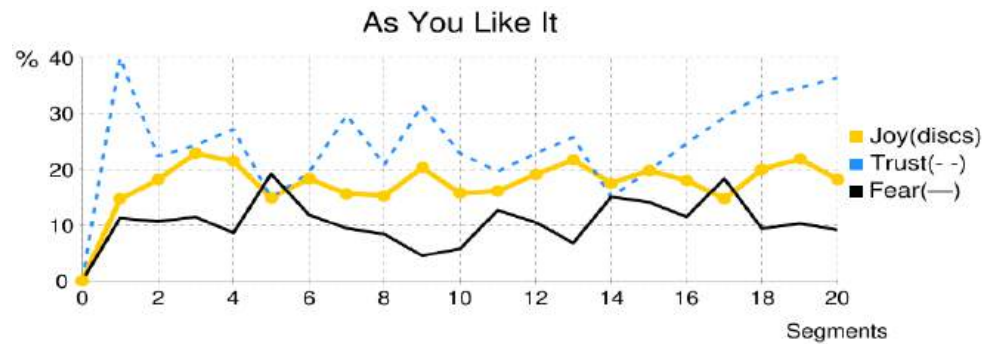




Emotions Arcs in CAS (mind, body, world), Commerce, Psychology, Health, and Humanities

Back in 2011:

Tracking Emotions in Stories



From *Once Upon a Time to Happily Ever After: Tracking Emotions in Novels and Fairy Tales*, Saif Mohammad, In Proceedings of the ACL 2011 Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH), June 2011, Portland, OR.

Creating Emotion Arcs

- Lexicon-only approach
- ML approaches (sometimes making use of lexicons)

Lexicon-only approaches

- Pros
 - simple, accessible
 - interpretable
 - low-carbon
 - domain-free
- Cons
 - not highly accurate at instance level (context, long-distance dependencies)

Evaluating Emotion Arcs

Very little work!! No dataset of gold arcs.

Evaluating Emotion Arcs

Consider tracking anger in tweets associated with vaccines (week by week)

- Manually annotate 300,000 individual tweets from 2018 to 2024
- Take the percentage of tweets marked as joy in every week to create the emotion arc

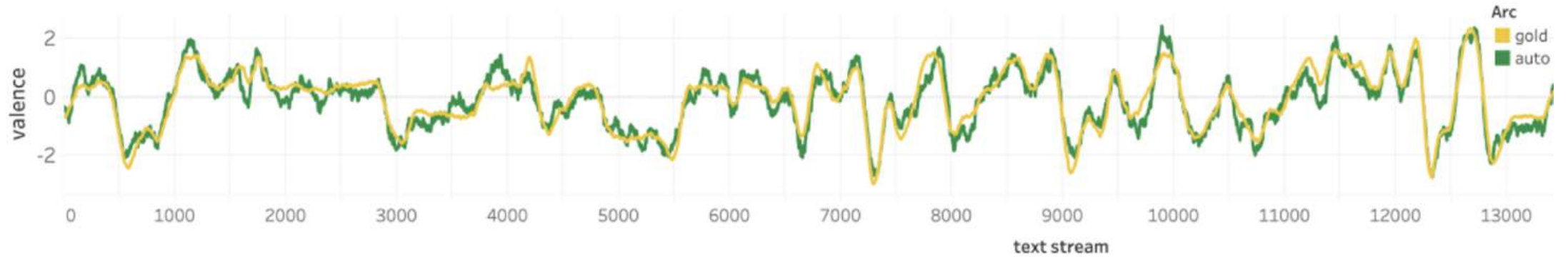
Annotating data is a bottleneck

2023 EMNLP: Evaluating Emotion Arcs Across Languages

- make use of existing emotion datasets (usually 2 to 5K instances)
- sample instances with replacement to generate random but non-trivial arcs
- create gold emotion arcs as usual



Daniela Teoderescu



2023 EMNLP: Generating High-Quality Emotion Arcs Using Emotion Lexicons

- Used 36 datasets that had emotion-labeled sentences/tweets to create gold arcs
- For various affect categories, multiple languages, and other characteristics

Key Conclusions:

- lexicon-only based methods are extremely accurate
- aggregating information from hundreds of tweets/instances to create points of the emotion arcs very powerful

Used WorryWords to

Track the change of anxiety in streams of text

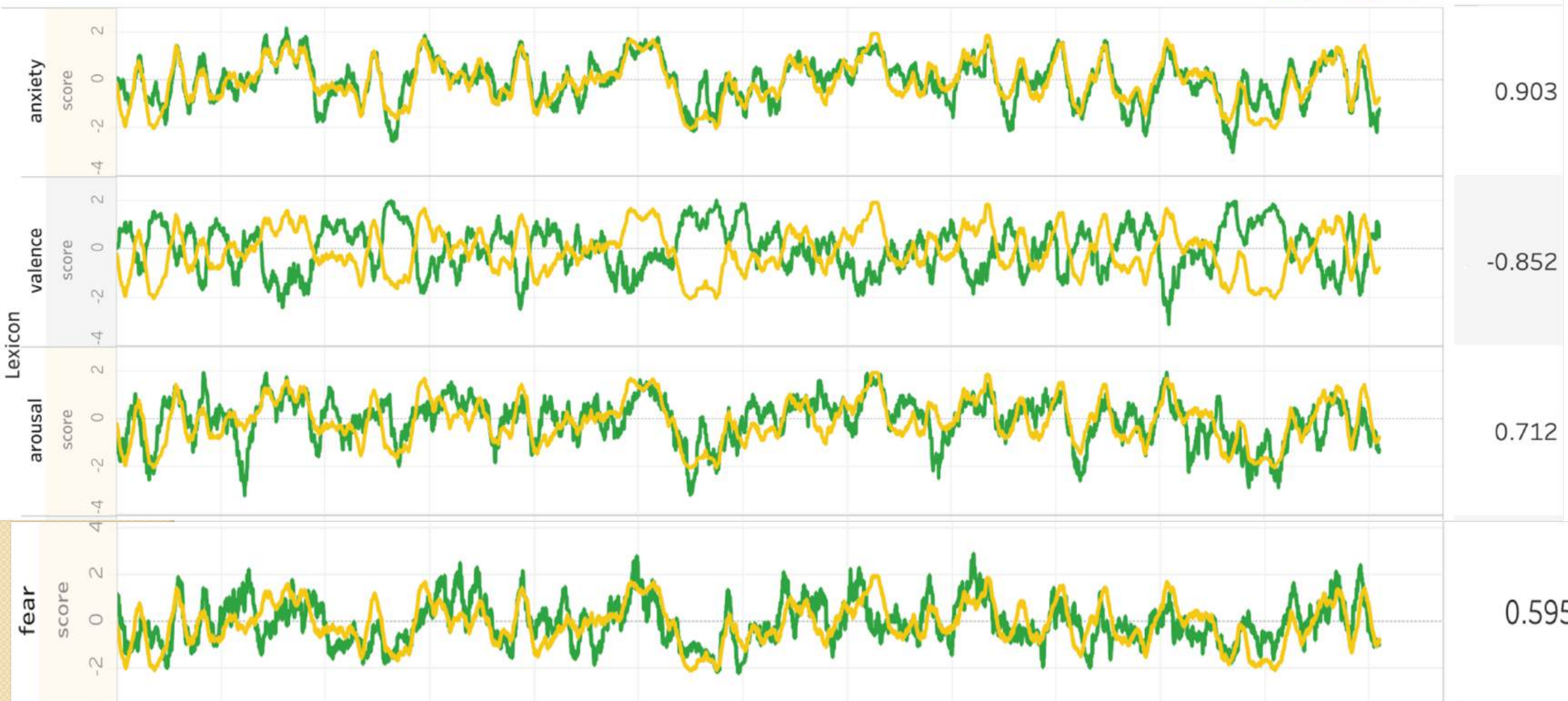




Emotion Arcs

Type
gold pred

Spearman
Correlation



This Talk



- Core Theories of Affect and Emotion
 - What is Computational Affective Science?
 - Why it matters?
- Overview of the Areas of Work in CAS
 1. nature of affect; affect and the mind, body, world
 2. affective data
 3. affective tasks
 - a. sentence/post-level classification/regression/generation
 - b. aggregate-level (group comparisons, emotion arcs, etc.)
 4. affective ethics
- Case Studies Exploring CAS Research



1: Affect and the Mind

aggregate-level (emotion arcs)

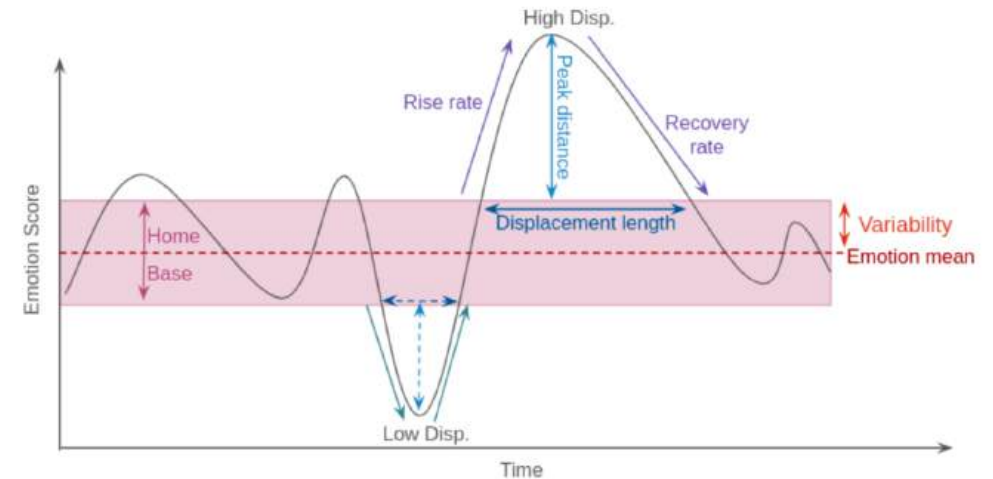
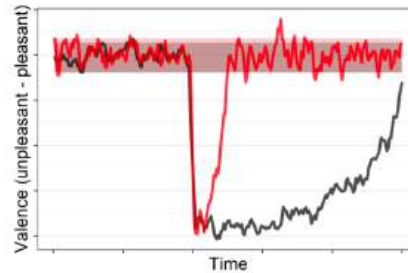
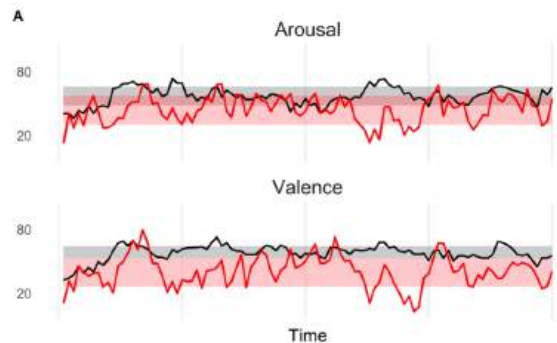


Emotion Dynamics (from Psychology)



Study of change in emotional state with time

- intensive longitudinal data (repeated self-reports of emotional state)
- quite difficult to obtain such data



Another window into emotions is through our words:

- E.g., if happier, we are likely to utter more happiness-associated words

Utterance Emotion Dynamics: study of change in emotion words over time
(Hipson and Mohammad, 2021)



PLOS One



Will Hipson

2021: Emotion Dynamics of Fictional Characters

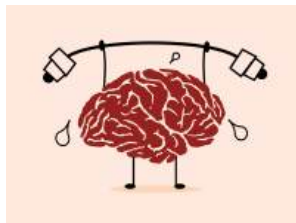


LREC

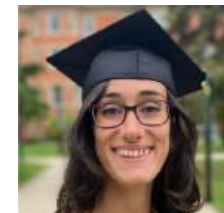


Krishnapriya (KP) Vishnubhotla

2022: Tweet Emotion Dynamics Emotion Word Usage in Tweets from US and Canada



EMNLP



Daniela Teodorescu



Tiffany Cheng



Alona Fyshe

2023: Language and Mental Health: Measures of Emotion Dynamics from Text as Linguistic Biosocial Markers.





We saw...

Emotion Dynamics: Individual Emotion Arcs

next...

Emotion Granularity: Pairs of Emotion Arcs

Emotion Granularity aka Emotion Differentiation (from Psychology)

Some people:

- recognize, identify, describe their feelings using **precise** terms
 - like guilt, anger, frustration, or helplessness
- can **reliably** describe these concepts using language
 - distinguishing between angry and sad, elated and content, etc.



Lisa Barrett

Others:

- tend to use more broad terms to convey emotions
 - a general sense of feeling bad or feeling low
- co-endorsing multiple emotions

Emotion granularity (**Barrett et al., 2001**)

- this ability to experience and categorize emotions in very specific terms
- the degree of not co-endorsing multiple emotions

Link between Emotion Granularity and Health

- Mental health (Erbas et al., 2014, 2018)
 - depression (Starr et al., 2017)
 - anxiety (Seah et al., 2020)
 - borderline personality disorder (Dixon-Gordon et al., 2014, Suvak et al., 2011)
 - show less neural reactivity to rejection (Kashdan et al., 2014)
- Physical health (Hoemann et al., 2021)
 - cardiovascular physiological activity and stress (Bonar et al., 2023)
- Behavior
 - maladaptive behaviours such as binge drinking, aggression, and self-injurious behavior (Dixon-Gordon et al., 2014, Kashdan et al., 2015)
 - school behaviour (Brackett, Rivers, Reyes, & Salovey, 2012)
 - eating disorders (Selby et al., 2013)
 - less likely to retaliate aggressively (i.e., verbally or physically assault) against someone who has hurt them (Pond et al., 2012)

Emotion Granularity from Text (our work)

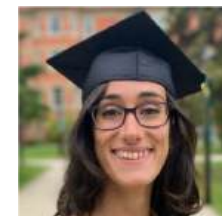
- To what extent are we co-endorsing multiple emotions ****in text****?
 - through connotations and not necessarily denotations
- Compute emotion arcs for various emotions
- Compute **emotion granularity (EG)**: correlation between pairs of emotion arcs
- Show that text by those who have self-disclosed to have certain mental health conditions (depression, PTSD, ADHD, etc.) have significantly lower EG than text by control group

EMNLP 2024:

Emotion Granularity from Text: An Aggregate-Level Indicator of Mental Health



Krishnapriya (KP)
Vishnubhotla



Daniela Teodorescu



Mallory Fedman



Kristen Lindquist



Sophie Wu



Jan Philip Wahle

The Language of Interoception:

Examining Embodiment and Emotion Through a Corpus of Body Part Mentions

Impact of the body on

- cognition, thoughts, how we conceptualize and perceive things
- emotions

Work from

- Medicine
- Psychology
- Affective Science
- Cognitive Science
- Philosophy

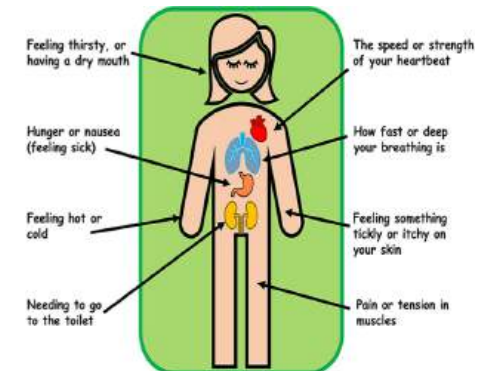
Key Theories

- Embodiment Theory
 - our physical body and its interactions with the world are fundamental to our cognitive processes, emotions, and experiences (not separate)
- The Theory of Conceptual Metaphor (Lakoff & Johnson, 1980)
 - abstract concepts (like time, emotion, power, and relationships) map to concrete source domains (like body, movement, or space)
e.g., *grasp* (a physical act) to mean *understand*
 - abstract domains are understood via source domains like the body, movement, or space.

Body Part Mentions (BPMs) offer a rich entry point into conceptual metaphor. Studying them can help reveal how language encodes embodied thought.

Key Theories (continued)

- **Mindfulness Theory**
 - developing an awareness of the present-moment
> positive outcomes (reduced stress, improved emotional regulation, well-being)
- **Theory of Constructed Emotion**
 - Emotions are actively constructed by the brain in each moment based on past experiences, context, and sensory information
- **Theories on Interoception**
 - Better interoception positively correlated with better physical and mental health
 - Emotional regulation (Zamariola et al., 2019)
 - Emotional decision-making (Dunn et al., 2010)
 - Emotional granularity (ability to distinguish emotions) (Zamariola et al., 2024)





The Language of Interoception:

Examining Embodiment and Emotion Through a Corpus of Body Part Mentions

Can we use language to shed light on this connection of the **body** with the mind, emotions, health, and behavior?



Body Part Mentions

Instances of language where words referring to parts of the body are used.

Examples:

My head hurt after the smog from last week.

Her chest tightened after hearing the news.

His back ached after the long flight.

I could feel my heart racing.

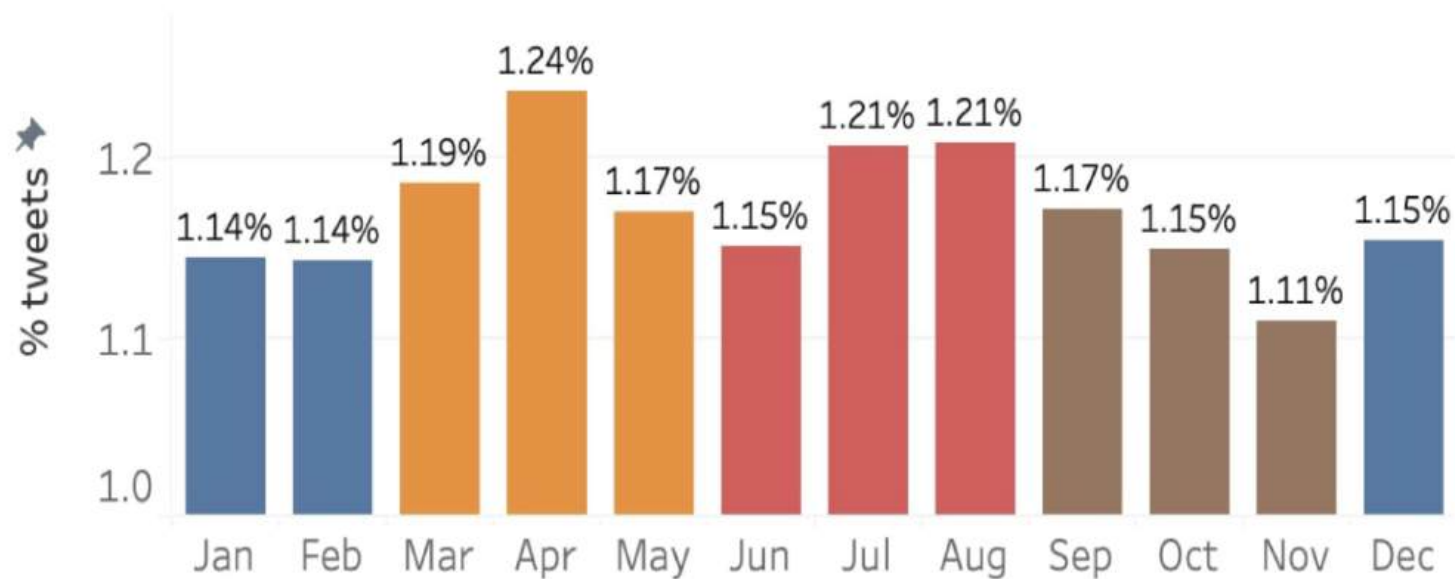
She rolled her eyes at the comment.

The hair on my arms stood up.

Body Research Questions

1. To what extent do we use body-related words? 5 to 10%
2. To what extent do we talk about our own body parts versus others' body parts?
3. Which of our body parts do we refer to most often? Do we refer to our body differently in different online contexts?
4. Does the time of day/week/year impact whether we refer to our body?
5. Do individuals in different regions refer to their bodies at different frequencies?

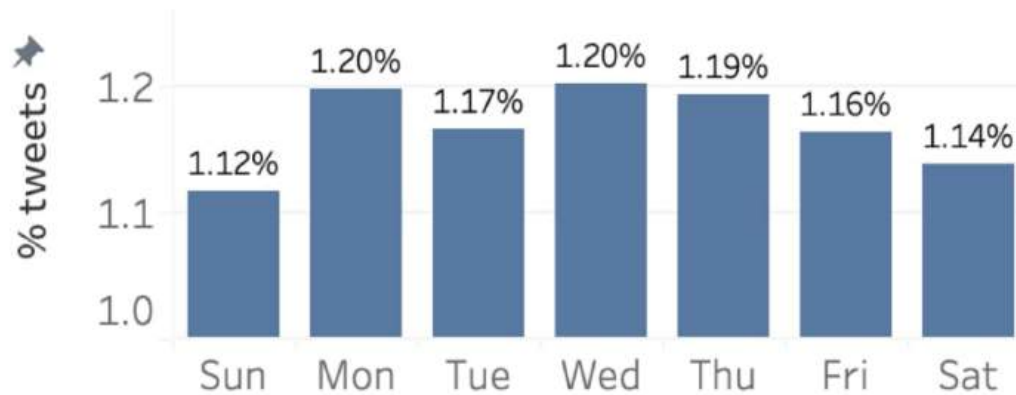
B4.a. Mentions of “my <BPM>” by Month



Mentions of “my <BPM>” peak in spring and summer, decline in fall, and stay lowest in winter.

- May reflect increased bodily awareness in warmer, more active months (sunlight, time outside).

B4.b. Mentions of “my <BPM>” by Day of Week

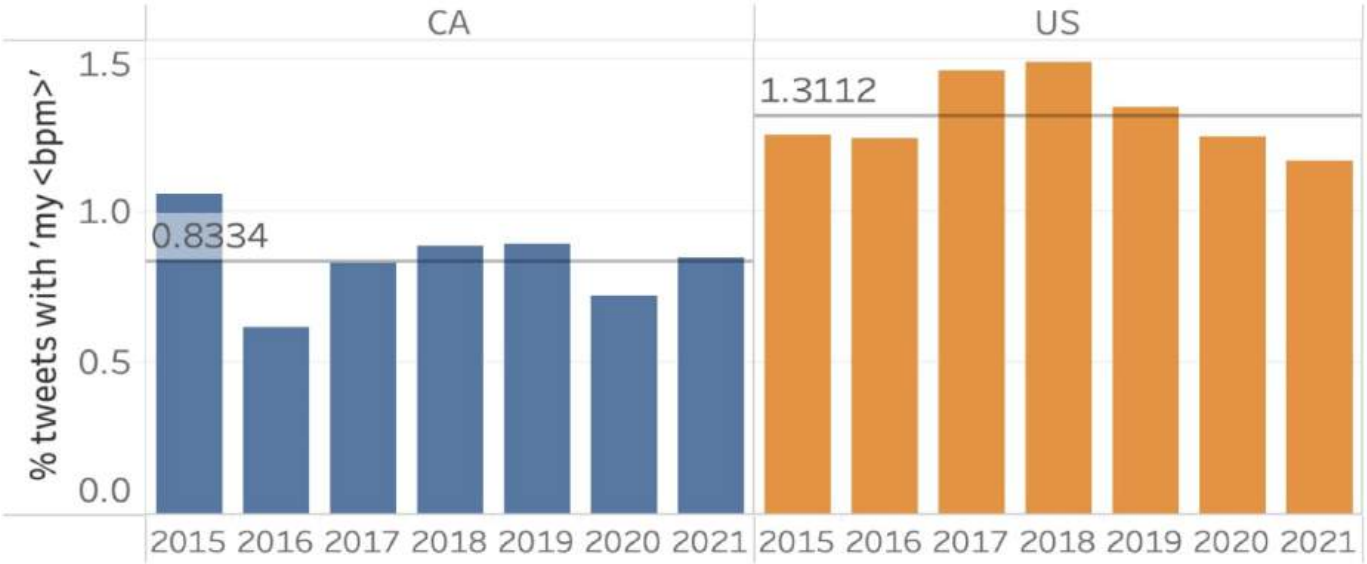


- Usage rises from Weekends to mid-week, then declines
- Suggests a connection to the structure and fatigue of the work week

References to our bodies are not static—they reflect seasonal, environmental, and social rhythms in our lives.

B5.a. Do individuals in different regions refer to their bodies at different frequencies?

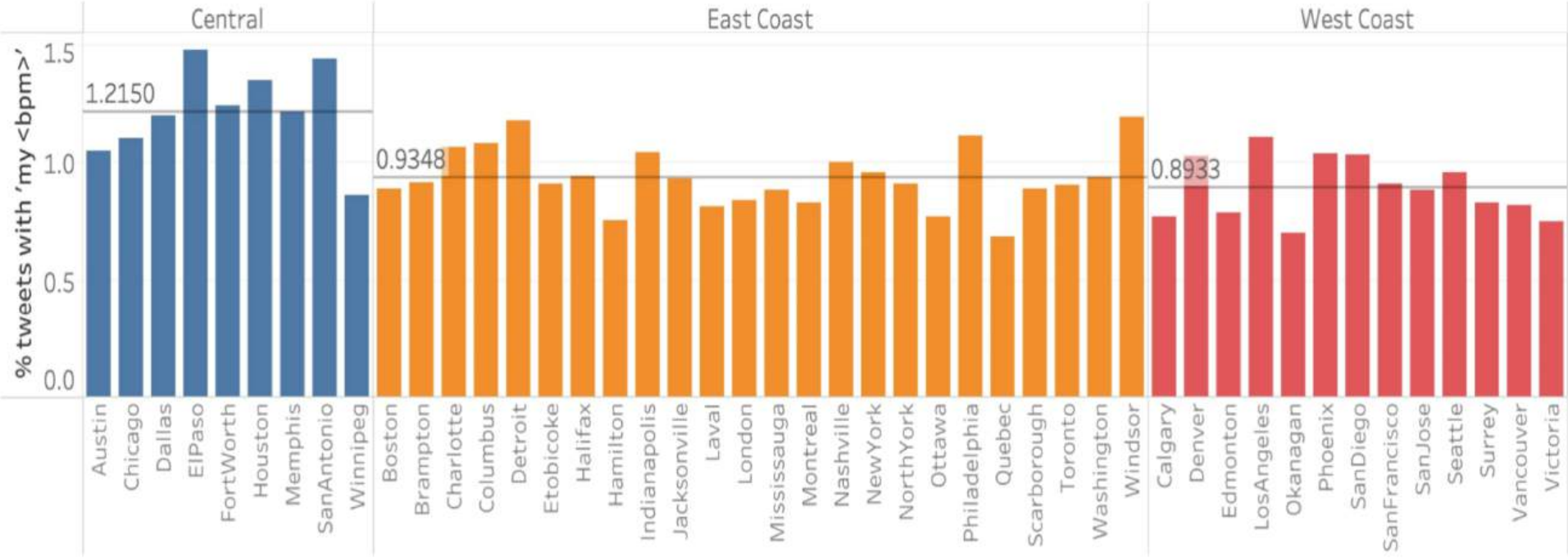
Canada vs. US across Years



Americans refer to their bodies more than Canadians in tweets.

B5.b. Do individuals in different regions refer to their bodies at different frequencies?

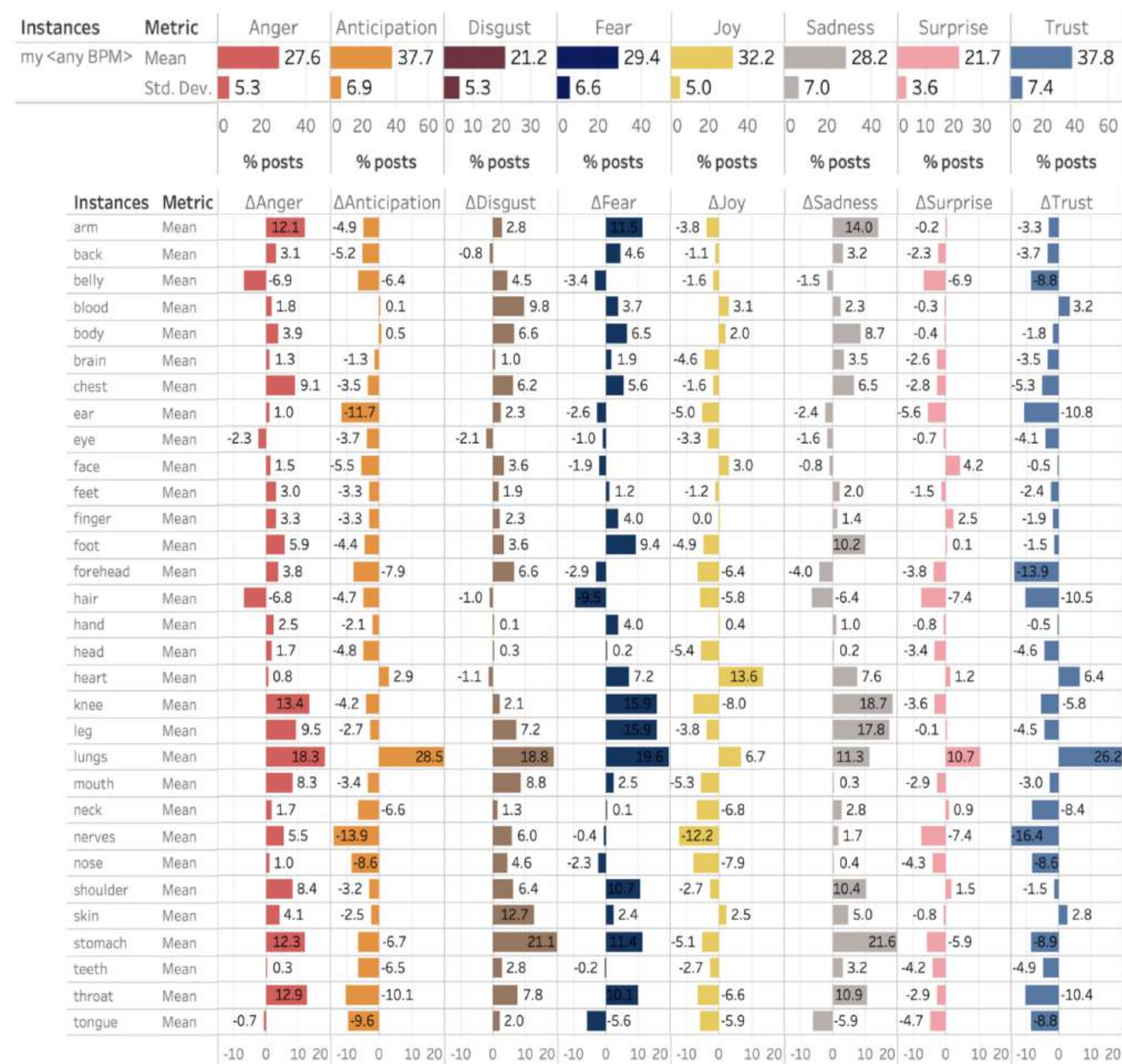
US cities: Central, East Coast, West Coast



Central cities show higher BPM usage than coastal cities.

Body and Affect Research Questions

1. Do posts with body part mentions have markedly different emotional associations?
2. What is the impact of explicitly embodied emotion on the emotions expressed through body part mentions?
3. Do individual body part mentions co-occur with markedly different emotion distributions?



BPMs and Emotions

- Different body parts are associated different emotions
 - my stomach → sadness
 - my chest → anger
- Negative emotions (anger, fear, sadness) dominate many high-frequency BPMs.

Are BPMs correlated with health outcomes?

	Freq. Mental Distress		Freq. Phys. Distress		Life Expectancy		Physical Inactivity	
	Spearman's r	p -value	Spearman's r	p -value	Spearman's r	p -value	Spearman's r	p -value
a. Number of tweets	-0.170	0.418	-0.167	0.425	0.290	0.160	-0.243	0.242
b. Prop. of <Fear word> tweets	-0.230	0.231	-0.370	0.054	0.160	0.403	-0.460	0.014
c. Prop. of "my <BPM>" tweets	0.497	0.012	0.721	0.000	-0.409	0.043	0.704	0.000
d. Prop. of "<BPM>" tweets	0.527	0.007	0.553	0.004	-0.613	0.001	0.539	0.006

- Proportion of BPM tweets is a strong predictor of negative health outcomes
- Emotion word use weakly correlated (at best)

This Talk



- Core Theories of Affect and Emotion
 - What is Computational Affective Science?
 - Why it matters?
- Overview of the Areas of Work in CAS
 1. nature of affect; affect and the mind, body, world
 2. affective data
 3. affective tasks
 - a. sentence/post-level classification/regression/generation
 - b. aggregate-level (group comparisons, emotion arcs, etc.)
 4. affective ethics
- Case Studies Exploring CAS Research



4: Affective Ethics

...



Emotion Recognition: Task

1. Inferring emotions felt by the speaker



Emotion Recognition: Task

1. Inferring emotions felt by the speaker
2. Inferring emotions of the speaker as perceived by the reader/listener
3. Inferring emotions that the speaker is attempting to convey
4. Inferring emotions evoked in the reader/listener
5. Inferring emotions of people mentioned in the text
6. Inferring whether what is described is good for pre-determined target of interest

Emotion Recognition: Task

1. Inferring emotions felt by the speaker
2. Inferring emotions of the speaker as perceived by the reader/listener
3. Inferring emotions that the speaker is attempting to convey
4. Inferring emotions evoked in the reader/listener
5. Inferring emotions of people mentioned in the text
6. Inferring whether what is described is good for pre-determined target of interest
7. Inferring the intensity of the emotions discussed above
8. Inferring patterns of speaker's emotions over long periods of time, across many utterances; including the inference of moods, emotion dynamics, and emotional arcs
9. Inferring speaker's emotions/attitudes/sentiment towards a target product, movie, person, idea, policy, entity, etc.
10. Inferring emotionality of language used in text (regardless of whose emotions)
11. Inferring how language is used to convey emotions such as joy, sadness, loneliness, hate, etc.
12. ...

Emotion Recognition: Task

1. Inferring emotions felt by the speaker
2. Inferring emotions of the speaker as perceived by the reader/listener
3. Inferring emotions that the speaker is attempting to convey
4. Inferring emotions evoked in the reader/listener
5. Inferring emotions of people mentioned in the text
6. Inferring whether what is described is good for pre-determined target of interest
7. Inferring the intensity of the emotions discussed above
8. Inferring patterns of speaker's emotions over long periods of time, across many utterances; including the inference of moods, emotion dynamics, and emotional arcs
9. Inferring speaker's emotions/attitudes/sentiment towards a target product, movie, person, idea, policy, entity, etc.
10. Inferring emotionality of language used in text (regardless of whose emotions)
11. Inferring how language is used to convey emotions such as joy, sadness, loneliness, hate, etc.
12. ...

All of these come with...

Benefits, Potential Harms, Ethical Considerations

Theories of Emotion



Margaret Mead
Cultural anthropologist



Paul Ekman
Psychologist and discoverer
of micro expressions.



Lisa Barrett
University Distinguished
Professor of Psychology,
Northeastern University

Theory of Constructed Emotion (Barrett, 2017)

- the brain **constructs** emotions
- important tenets of BET discredited (“basic” emotions)
- stress on variability

Computational Analysis of Emotions and Automatic Emotion Recognition (AER)

A force that helps unlock:

- how emotions work
- how they relate to our health, language, behavior, social interactions,...
- numerous commercial applications that benefit society

A tool for substantial harm, e.g.:

- mass application on vulnerable populations
- unreliable approaches
- privacy concerns
- perpetuation of physiognomy



[Strategies](#) [Topics](#) [Regions](#) [Up Close](#) [Tools](#) [Multimedia](#)

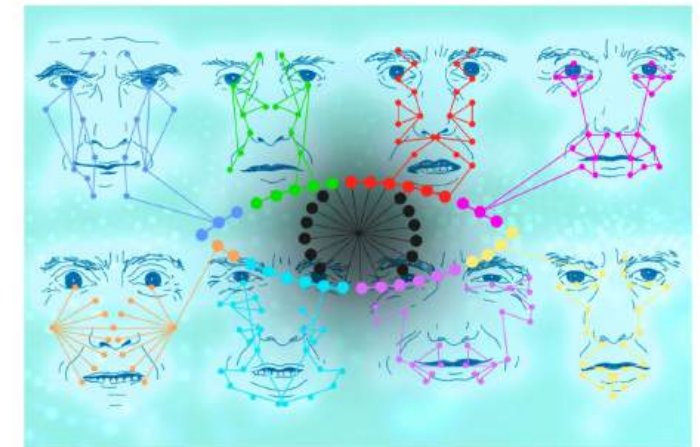
[Partnerships](#)

How emotion recognition software strengthens dictatorships and threatens democracies

Given that the idea of using emotion recognition technology as a tool of governance is an entirely flawed premise, a ban makes the most sense.

By: James Jennion

[Español](#)



Ethics Sheet for

Automatic Emotion Recognition and Sentiment Analysis



Medium Blog Post in summer of 2021:

<https://medium.com/@nlpscholar/ethics-sheet-aer-b8d671286682>



CL Journal June 2022

Ethics Sheet for Automatic Emotion Recognition and Sentiment Analysis

Saif M. Mohammad*

The importance and pervasiveness of emotions in our lives makes affective computing a tremendously important and vibrant line of work. Systems for automatic emotion recognition (AER) and sentiment analysis can be facilitators of enormous progress (e.g., in improving public health and commerce) but also enablers of great harm (e.g., for suppressing dissidents and manipulating voters). Thus, it is imperative that the affective computing community actively engage with the ethical ramifications of their creations. In this paper, I have synthesized and organized information from AI Ethics and Emotion Recognition literature to present fifty ethical considerations relevant to AER. Notably, the sheet fleshes out assumptions hidden in how AER is commonly framed, and in the choices often made regarding the data, method, and evaluation. Special attention is paid to the implications of AER on privacy and social groups. Along the way, key recommendations are made for responsible AER. The objective of the sheet is to facilitate and encourage more thoughtfulness on why to automate, how to automate, and how to judge success well before the building of AER systems. Additionally, the sheet acts as a useful introductory document on emotion recognition (complementing survey articles).

Template



50 considerations grouped under:

- *Task Design*
 - *Data*
 - *Method*
 - *Impact and Evaluation*
 - *Implications for Privacy and Social Groups*
- } common phases in system development

TASK DESIGN

A. Theoretical Foundations

1. Task Design and Framing
2. Theoretical Models and their Implications
3. Meaning and Extra-Linguistic Information
4. Wellness and Health Implications
5. Aggregate Level vs. Individual Level Prediction

B. Implications of Automation

6. Why Automate
7. Embracing Diversity
8. Participatory/Emancipatory Design
9. Applications, Dual Use, Misuse
10. Disclosure of Automation

DATA

C. Why This Data

11. Types of data
12. Dimensions of data
- D. Human Variability–Machine Normativeness
13. Variability of Expression, Conceptualization
14. Norms of Emotions Expression
15. Norms of Attitudes

... 50!

Shared Task on Emotions

SemEval-2025 Task 11: Bridging the Gap in Text-Based Emotion Detection

Labeled emotion datasets for **35** languages

-- most from Africa and Asia

Most popular task in CodaLab for the year 2024

SemEval 2025 Best Shared Task Award

ACL 2025 Best Resource Paper Award





Affect in GenAI Development and its Ethics

- Ethics of AI Assistants
- Anthropomorphization of LLMs
- Human-Machine Relationships
- Socio-affective Alignment for AI Systems



Computational Affective Science

A Window into emotions, mind, body, health, and behavior through language and computation

Slides, Papers, Datasets, Lexicons, Code

Available at: www.saifmohammad.com/WebPages/CAS-tutorial.html

✉ vk22priya@gmail.com, saif.mohammad@nrc-cnrc.gc.ca, uvgotsaif@gmail.com

🐦 [@krishnapriyaVi5](https://twitter.com/krishnapriyaVi5), [@SaifMMohammad](https://twitter.com/SaifMMohammad)