Embracing Ambiguity and Subjectivity in Emotion Research

EASE WORKSHOP @ ACII 2024

<ease.ewi.tudelft.nl>

Workshop Overview

- Key Themes and Presenters
  - Theoretical Foundations
    - Girard (T), Dudzik (P)
  - Modeling and Evaluation Approaches
    - Woodland (T), Wu (P)
  - Database and Annotation Design
    - Yannakakis (T), Kunc (P), Viswanath (P)
  - Future Challenges and Opportunities
    - Plenary Discussion, Future Planning
- Workshop Organizers
  - Jeffrey Girard (KU)
  - Vidhyasaharan Sethu (UNSW)
  - Bernd Dudzik (TU Delft)
  - Carlos Busso (UT Dallas)
  - Emily Mower Provost (UMich)
  - Shrikanth Narayanan (USC)
- Workshop Position Papers
  - 7 submitted, 4 accepted

Theoretical Foundations of Ambiguity and Subjectivity

JEFFREY GIRARD, PHD

UNIVERSITY OF KANSAS

<affcom.ku.edu>

Presentation Roadmap

1. Introduction and Motivation
2. Statistical Foundations
3. Philosophical Foundations
4. Psychological Foundations

Introduction & Motivation

WORKSHOP RATIONALE

PROPOSED TERMINOLOGY

QUESTIONS/APPLICATIONS

Workshop Rationale

- In ACII research, we often study affective perceptions and experiences
  - **Perceptions:** Third-party inferences about what another person is feeling
  - **Experiences:** First-person reports about what you are feeling/experiencing
- We may explain, predict, or generate perceptions and experience
- We often assume there is **a single "correct" perception or experience**
  - e.g., we take the majority vote or average rating from several participants
- This approach simplifies things, but **what are we losing** by doing so?

Workshop Rationale

- In practice, perceptions and reactions usually vary across individuals
- There are thus **distributions** of perceptual and experiential ratings
- We argue that ACII would benefit from **embracing this variability**
  - Instead of treating variability as a nuisance, we should model and study it
- People are already beginning to explore aspects of this in the field!
  - This workshop brings them together and raises awareness of these issues
  - Our goals are to establish **shared goals, directions, terminology, and methods**

Proposed Terminology

- **Perceptual Variability:** Observers may perceive the same stimuli differently
- **Experiential Variability:** Individuals may experience the same events differently
- **Ambiguity:** variability related to *uncertainty* (e.g., missing or conflicting info)
  - Influenced by features of the stimulus being perceived or experienced
- **Subjectivity:** variability related to *personal perspectives* (e.g., feelings or beliefs)
  - Influenced by attributes of the individual doing the perceiving or experiencing
- **Error:** variability related to mistakes, accidents, and misunderstandings

Psychological Questions

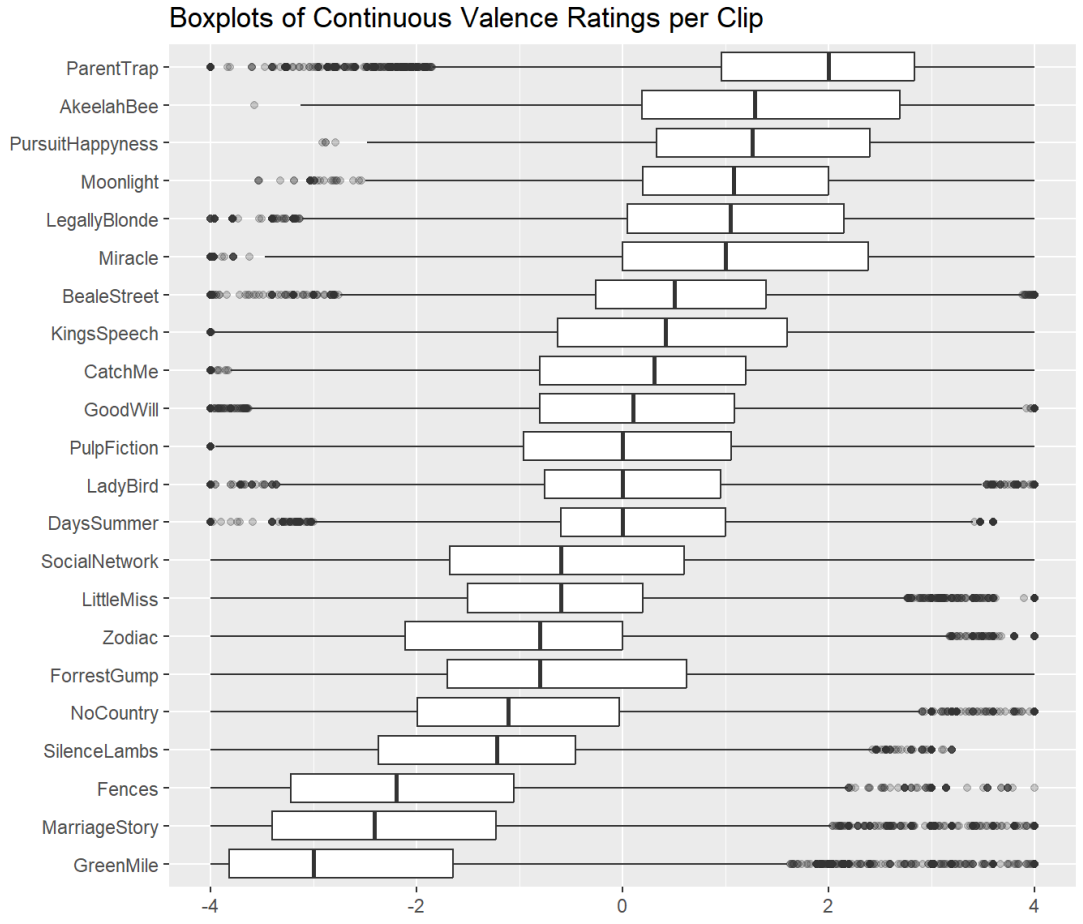
- What **features of the stimulus** explain how ambiguous it is?
- What **attributes of the participant** explain their subjective ratings?
- What **aspects of the surrounding context** influence these ratings?
- What **consequences** (e.g., socially and clinically) are there for producing ambiguous signals or being "unusual" in your perceptions and reactions?
- Can we **help people** to send less ambiguous affective signals or to "normalize" their affective perceptions and experiences? Should we?

Computational Questions

- Can we **predict** how ambiguous a stimulus is?
- Can we **control** how ambiguous a generated stimulus is?
- Can we predict the ratings' **distributional shape** from stimulus features?
- Can we predict a **rater's place** in the distribution from their attributes?
- Can we **personalize** insights, recommendations, ads, etc. for individuals based on their patterns of affective perception and experience?

Example 1

Distribution of Dynamic Valence Ratings of 22 movie clips by 104 participants

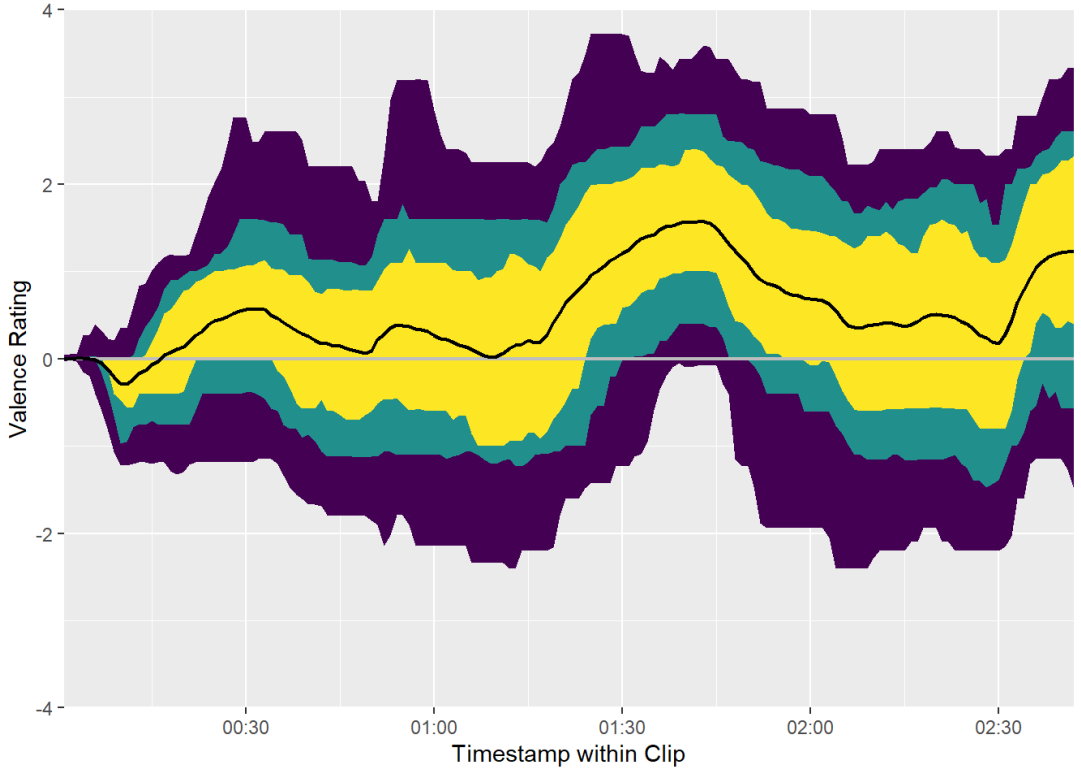


<dynamos.mgb.org>

Example 2

Time Series Distribution of Dynamic Valence Ratings per Second of the "BealeStreet" movie clip by 101 subjects

Black=Median
Yellow=Inner 50%
Green=Inner 70%
Purple=Inner 90%



<dynamos.mgb.org>

Statistical Foundations

CLASSICAL TEST THEORY

GENERALIZABILITY THEORY

CROSS-CLASSIFIED MODELS

Classical Test Theory

- CTT assumes there is a single true score T_i for stimulus i
- Our observed score X_i includes some random error E_i

$$X_i = T_i + E_i$$

$$E_i \sim Normal(0, \sigma_E)$$

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2$$

- We can't access T_i directly, so we measure X_i many times and combine
- If the errors are random as assumed, they should cancel each other out
- We don't care who the raters are – we just need several for each stimulus

Classical Test Theory

- Reliability in CTT is the proportion of X variance that is T variance

$$R = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2}$$

- With 2 raters, we estimate it using a correlation between their scores
- With 3+ raters, we estimate it using a one-way intraclass correlation (ICC)
- Reliability in CTT can give us a sense of the amount of variability present
- But in CTT, error is "monolithic" and uninteresting (i.e., merely a nuisance)

Generalizability Theory

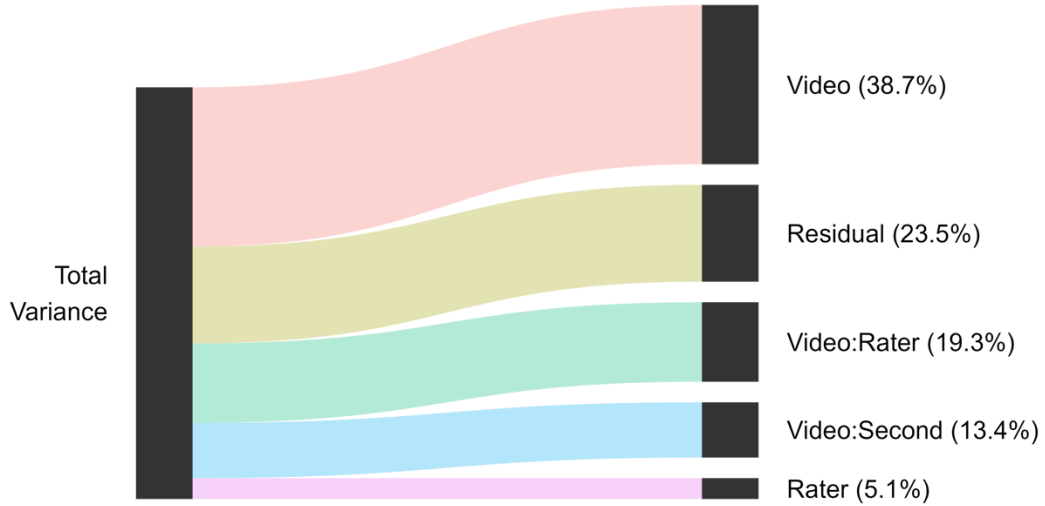
- GT is an extension of CTT that allows for more nuanced partitioning of X_i
- If we track raters, then we can include rater and rater-by-stimulus effects

$$\sigma_X^2 = \sigma_S^2 + \sigma_R^2 + \sigma_{S:R}^2 + \sigma_E^2$$

- This allows us to model mean rater differences across all stimuli (σ_R^2) as well as rater differences in response to specific stimuli ($\sigma_{S:R}^2$)
- The proportion of total variance attributable to each source is revealing
- GT can also accommodate more complex designs with additional facets

Example 3

Statistical Analysis Quantifying the Different Sources of Variance in the Dynamic Valence Ratings



<dynamos.mgb.org>

Cross-Classified Models

- Now that we have estimated how much variance occurs across raters, videos, and other facets, we can try to **explain that variance**!
- Mixed-effects and cross-classified models can accomplish this
  - *(However, they usually assume linear relationships and normal distributions)*
- For instance, does knowing a rater's sex help us predict their ratings?

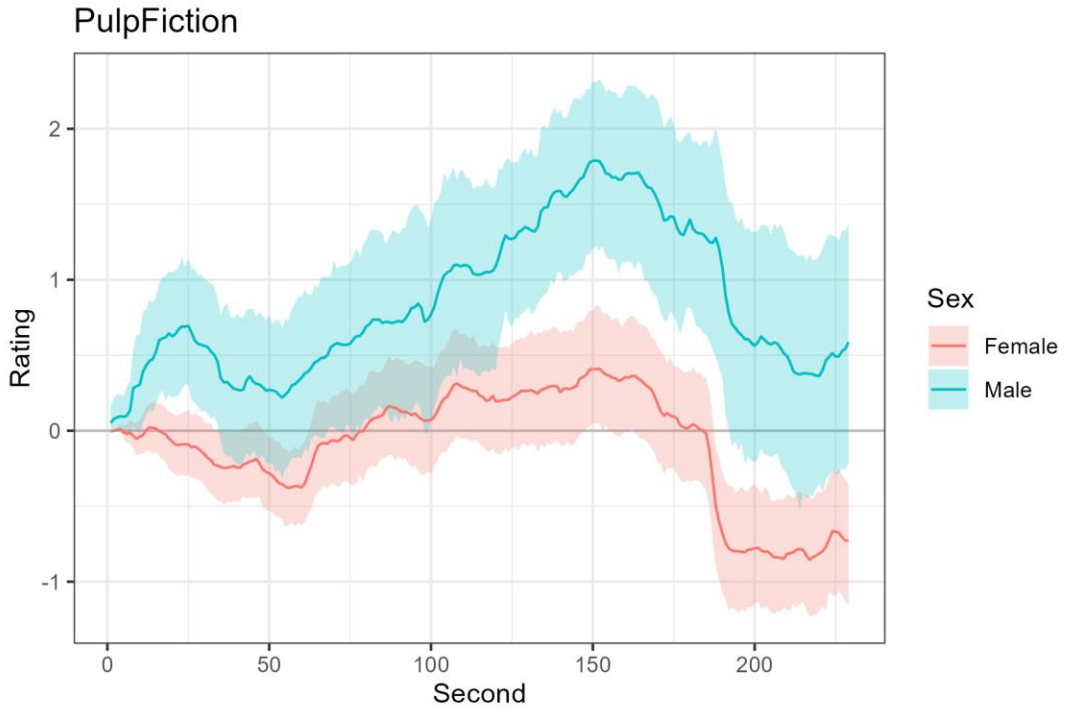
CTT: `rating ~ 1 + (1 | video)`

GT: `rating ~ 1 + (1 | video) + (1 | rater) + (1 | video:rater)`

CCM: `rating ~ 1 + sex + (1 + sex | video) + (1 | rater) + (1 | video:rater)`

Example 4

Time Series Distributions of Dynamic Valence Ratings per Second of the "PulpFiction" movie clip by Rater Sex



<dynamos.mgb.org>

Philosophical Foundations

PHENOMENOLOGY

EXISTENTIALISM

DECONSTRUCTIONISM

Phenomenology

A philosophical tradition focused on the *direct examination and description of conscious experience*, aims to uncover fundamental structures of perception, thought, and meaning as they present themselves to consciousness (Husserl, Merleau-Ponty)

- Emotions **vary from person to person**, even in response to similar stimuli.
- Emotions are **complex**, multi-layered, and sometimes contradictory.
- Emotions are influenced by **past experiences** and **future anticipations**.
- Emotions are influenced by **cultural, social, and historical** contexts.
- Parts of others' emotional experiences will always be **inaccessible** to us.

Existentialism

A philosophical tradition focused on *individual freedom, choice, and responsibility*, asserting that individuals *create their own meaning and essence* through their actions in an inherently meaningless or absurd world. (Sartre, Camus, de Beauvoir)

- Emotions **depend on our choices and interpretations** of our circumstances.
- **Ambiguity is inevitable** and a fundamental part of the human condition.
- Ambiguity often produces fear, anxiety, dread, and a desire for certainty.
- We must **embrace our freedom and the inherent ambiguity of existence**.
- Emotions are shaped by our awareness of **how other people perceive us**.

Deconstructionism

A philosophical tradition focused on uncovering and challenging the inherent contradictions and assumptions within concepts, revealing how meaning is not fixed but rather *constructed through language and cultural context*. (Derrida, Kristeva)

- Meaning **cannot be pinned down** to a single, definitive interpretation.
- Emotions are always **subject to interpretation** and contextual factors.
- Emotions are **not neatly separable**; they bleed into one another and are shaped by overlapping cultural, linguistic, and personal factors.
- Emotions are **performative acts** and the process of expressing an emotion can actively shape and alter the emotional experience itself.

Psychological Foundations

DECISION SCIENCE

ATTRIBUTION & APPRAISAL

CONSTRUCTIVISM

Decision Science

An interdisciplinary field that explores *how individuals make choices*, often highlighting the ways in which human behavior deviates from traditional models of rationality due to *cognitive biases, emotions, and social influences*. (Kahneman, Loewenstein)

- People often use simplified mental/emotional **shortcuts to make decisions**.
- We are more likely to perceive emotions that **align with our expectations**.
- **Our emotional states** can influence how we perceive others' emotions.
- Thus, human behavior and choices are **not always rational or consistent**.
- People often struggle to anticipate how they will feel in the future.

Attribution & Appraisal

Psychological traditions and theories focused on how *individuals' interpretations of events* shape their emotional and behavioral responses. (Weiner, Lazarus, Scherer)

- Emotions are shaped by **how we interpret causes and significance.**
- People tend to attribute their own emotions and behaviors to **situational factors** but others' emotions and behaviors to more **dispositional factors**.
- **Individual differences** influence how we interpret and feel about events.
- Emotions can **shift over time** as information is gathered and processed.
- Complex situations can lead to ambiguity and **mixed emotions**.

Constructivism

A psychological theory that *emotions are interpretations of bodily sensations* constructed within a given situational framework through the dynamic interplay of *individual experiences, cultural context, and cognitive processes*. (Barrett, Russell)

- Individuals **dynamically interpret** and give meaning to their experiences.
- Emotions are influenced by the **words available** to describe them, which can vary widely across cultures and even across individuals.
- Perceiving others' emotions involves interpreting complex social cues through one's own **subjective lens**. Observers must draw on their own knowledge and expectations to make sense of others' expressions.

Conclusions

Conclusions

- There are reasons to doubt the existence of a single "true" score
  - We should strive to collect and study *distributions* of scores
- Emotions are complex, dynamic, and difficult to pin down
  - We should strive to use more comprehensive assessments of emotion
- Many factors combine to influence our perceptions and reactions
  - We should strive to measure relevant thoughts, states, traits, and context
- We often fill in the gaps in our knowledge with subjective heuristics
  - We should strive to model these biases in our data to control for them

Thank you! Questions / Discussion

Workshop website: https://ease.ewi.tudelft.nl

My lab website: https://affcom.ku.edu