

# MAPPING THE CULTURAL MEANINGS OF PARENTAL IDENTITIES

Niyle Sutherland and Dr. Maria Cristina Ramos

## Background

- **Identity theory** conceptualizes identity as a set of **meanings** attached to **social positions** by which one can **define oneself** (Burke & Stets, 2023, 2025).
  - **Examples:** parent, teacher, athlete
- Meanings include **traits, expectations, and other content** that capture what it means to be a father, mother, and so on. (Stets & Burke, 2025).
  - **Examples:** caring, aggressive, responsible, stern
- Identity meanings exist both:
  - outside the individual (as **culturally or institutionally shared meanings**) and
  - inside the individual (as **internalized meanings**)
- Identity meanings guide behavior as individuals seek to **behave in ways that align with their identity meanings** (Burke & Stets, 2023).
  - **Examples:**
    - A father being stern
    - A medical student being studious

## Research Gaps & Aims

### Gaps:

- We lack systematic accounts that **distinguish types of identity meanings** and **their properties**.
- We lack a systematic and scalable **text-as-data approach** for **measuring identity meanings** in naturalistic discourse.
- We know little about **how identity meanings are expressed and maintained** in **online communities**.

### Aims:

- Document **different types of identity meanings**
- Identify meanings as expressed in **natural settings**
- Evaluate the use of scalable **text-as-data approaches** to measure identity meanings
- Understanding **cultural expectations** tied to parental roles
- Understanding how meanings are **socially enforced**
- Determining where meanings **converge or diverge** between mother and father

## Data

- The text that we are using as data for this project comes from **Parenting Stack Exchange**.
- This is a website where **parents can ask questions** about situations with their children.
  - People can **answer these questions** by commenting and giving their **opinions and advice**.
- The data includes text from both the questions and answers.
- The corpus spans from 2011 until January 2026 and includes:
  - **~7,000 questions**
  - **~23,000 answers**
  - **~52,000 users**

## Methods

- This project combines **qualitative content analysis** with **computational annotation**.

- We create **rules** for the computational annotation process and **establish conventions** for these rules.
  - This is called **creating a codebook**.

- We conduct the computational annotation process through **locally run large language models (LLMs)**.

- We use the annotated data to:
  - assess the **content and moral intensity** of expectations for mothers and fathers
  - assess **cultural consensus and contestation** surrounding parental norms

parental attributes and expectations

statistics modeling the prevalence, strength, and distribution of the attributes/expectations

Content of the analysis

norms across the data

reliability and validation checks

## Advantages and Limitations

### Advantages

- Developing and testing a structured codebook for computational annotation forces **conceptual clarity and theory alignment**.
- The data are **large**, which allows for **measuring meanings at scale**.
- The Q&A format explicitly elicits claims and standards of “good parenting,” making identity meanings **observable rather than inferred**.
- **Self-presentation bias is reduced** because posts are written for peers instead of for research.
- The data are being **measured in real time** because the forum is **constantly collecting data**.

### Limitations

- Since the data source was **not created for research purposes**, it may not have crucial information, such as demographics of posters.
- **Not all potentially useful data is accessible** - we originally intended to use parenting subreddits as data, but we were unable to access all of the data we would need due to Reddit’s policies.
- The results from this sample **cannot be generalized** to the population outside of the sample.
- The posts are **not completely naturally occurring** because of algorithms and moderators.
- Some information in posts might be **sensitive**; we minimize this risk via de-identification and obfuscation.

## References

Please **scan this QR code** to view our reference sheet.

