Anti-Vaccination Arguments: A Conceptual Taxonomy and a Machine-Learning Model

Stephan Lewandowsky

University of Bristol

@STWorg

The JITSUVAX Project

- Jiu Jitsu with misinformation in the age of COVID: Using refutation-based learning to enhance vaccine uptake and knowledge among healthcare professionals and the public

- Angelo Fasce (University of Coimbra)
- Dawn Holford (University of Bristol)
- Philipp Schmid (University of Erfurt)
- Luke Bates (Technical University of Darmstadt)
- Iryna Gurevych (Technical University of Darmstadt)
- Thomas Costello (MIT)

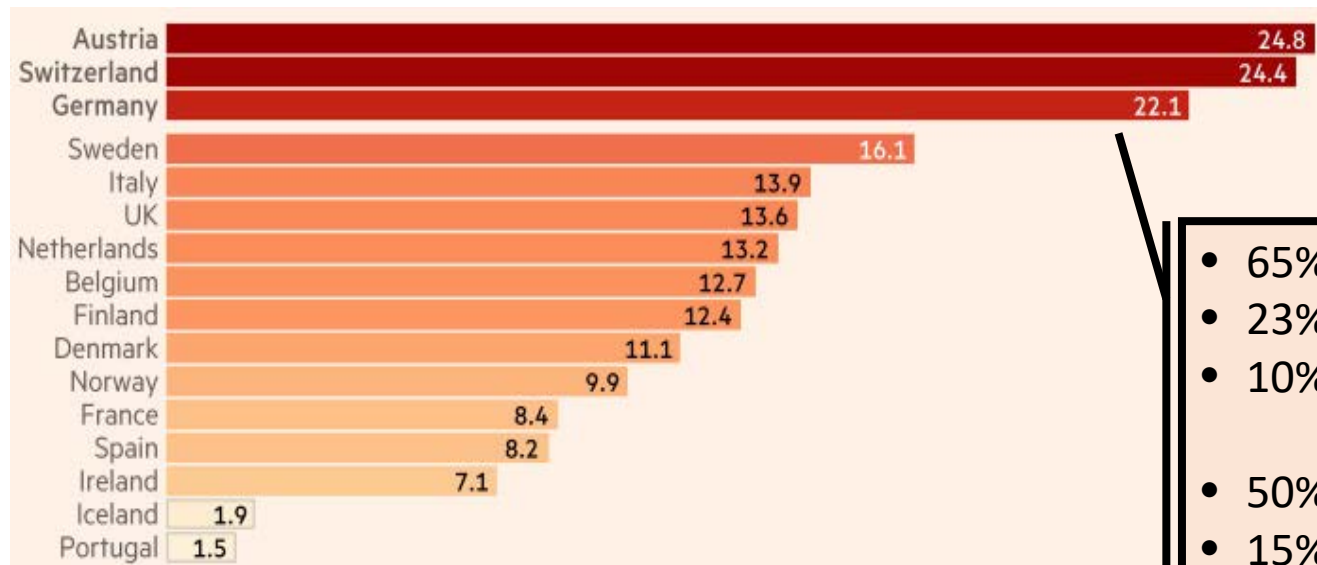# Europe Became More Pro-Vaccine After Rollout Commenced

% vaccinated or willing to get vaccinated against COVID-19

(YouGov, November 2021)

Co-funded by the Horizon 2020 program of the European Union
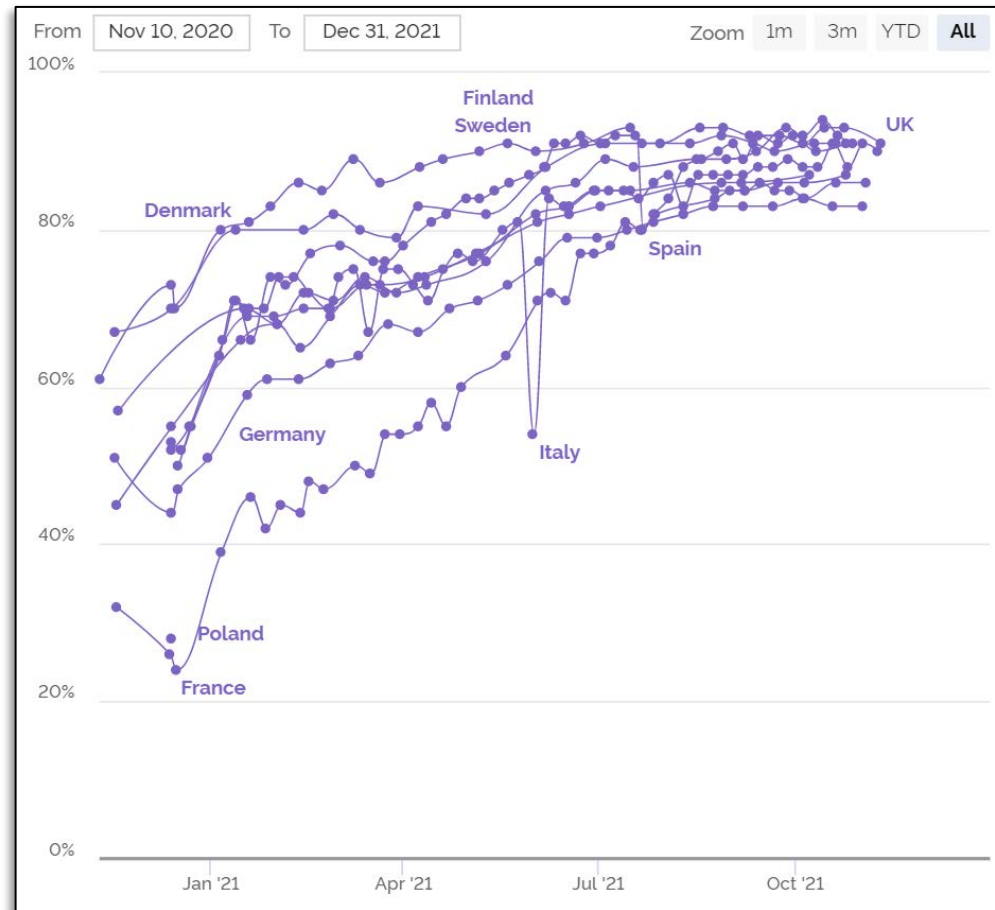
Residual Vaccine Hesitancy is Hardening

Share of unvaccinated people (October 2021)

- As of June 2022, actual uptake (first doses) mirrors survey data

Country	Value
Austria	24.8
Switzerland	24.4
Germany	22.1
Sweden	16.1
Italy	13.9
UK	13.6
Netherlands	13.2
Belgium	12.7
Finland	12.4
Denmark	11.1
Norway	9.9
France	8.4
Spain	8.2
Ireland	7.1
Iceland	1.9
Portugal	1.5

Source: FT analysis of figures from national sources and Our World in Data. Rates shown are as of November 9
© FT

- 65% definitely no
- 23% probably no
- 10% undecided or possibly yes

- 50% voted for extreme right
- 15% voted for other "anti-establishment" parties

We Need To Understand Where Hesitancy Is Coming From

Attitude roots:

"The beliefs, ideologies, fears, and identity issues that motivate people to want to reject scientific consensus"
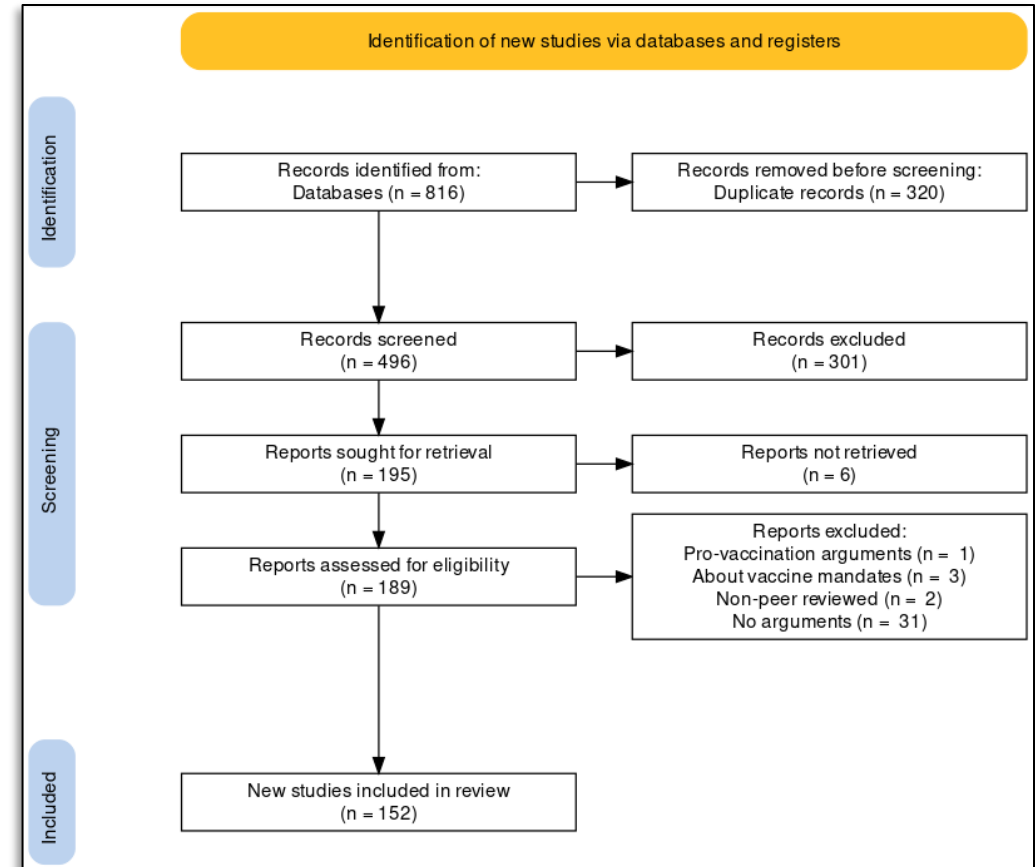
(Hornsey & Fielding, 2017)

Objectives

- Identify attitude roots by analyzing anti-vaccination arguments
- Create taxonomy of anti-vaccination arguments based on attitude roots and their manifestations
- Develop counter arguments or inoculation messages that are attitude-root-consonant.

Study 1: Systematic Literature Review (PRISMA)

Search string:

("vacc* refus* " OR "anti-vacc* " OR "anti-vax* " OR "vacc* oppon* " OR "vacc* oppos* " OR "vacc* skeptic* " OR "vacc* sceptic* ") AND (trope* OR meme* OR narrative* OR argumen* OR theme* OR rhetoric*)

Identification of new studies via databases and registers

Identification

Records identified from:
Databases (n = 816)

Records removed before screening:
Duplicate records (n = 320)

Screening

Records screened
(n = 496)

Records excluded
(n = 301)

Reports sought for retrieval
(n = 195)

Reports not retrieved
(n = 6)

Reports assessed for eligibility
(n = 189)

Reports excluded:
Pro-vaccination arguments (n = 1)
About vaccine mandates (n = 3)
Non-peer reviewed (n = 2)
No arguments (n = 31)

Included

New studies included in review
(n = 152)

Results of Systematic Review (*n*=152)

- 1967 – 2021

- Methods
  - qualitative (66.45%),
  - quantitative (3.95%)
  - mixed-methods (13.82%)
  - remainder conceptual

- Vaccines
  - multiple vaccines (61.18%)
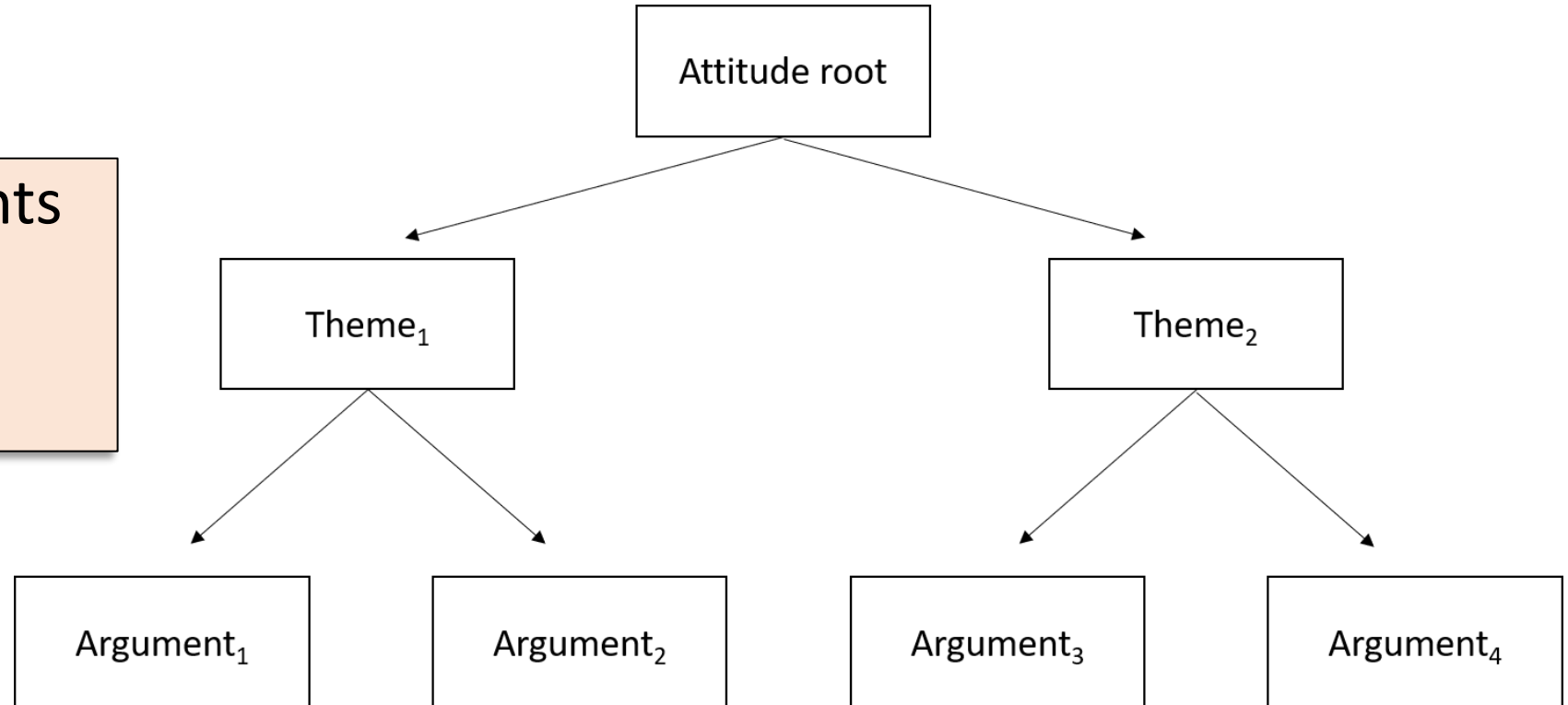 - HPV (9.87%)
  - COVID-19 (9.21%)
  - Influenza (7.89%)

47 countries

Argument Extraction

- 4 reviewers independently examined arguments
  - *specific level*: "HPV vaccine just encourages unsafe sex"
  - *abstract level*: "sexual promiscuity"

- Identify root
  - "sexual promiscuity" → moral concerns
  - some arguments assigned to >1 root

- Discrepancies between reviewers resolved by discussion

Argument Extraction

- 2,066 arguments
- 61 themes
- 11 roots

(Diagram: "Attitude root" at top, branching to $Theme_1$ and $Theme_2$. $Theme_1$ branches to $Argument_1$ and $Argument_2$. $Theme_2$ branches to $Argument_3$ and $Argument_4$.)

Jitsu VAX

11 Attitude Roots

Conspiracist ideation

Distrust

Unwarranted beliefs

Worldview and politics
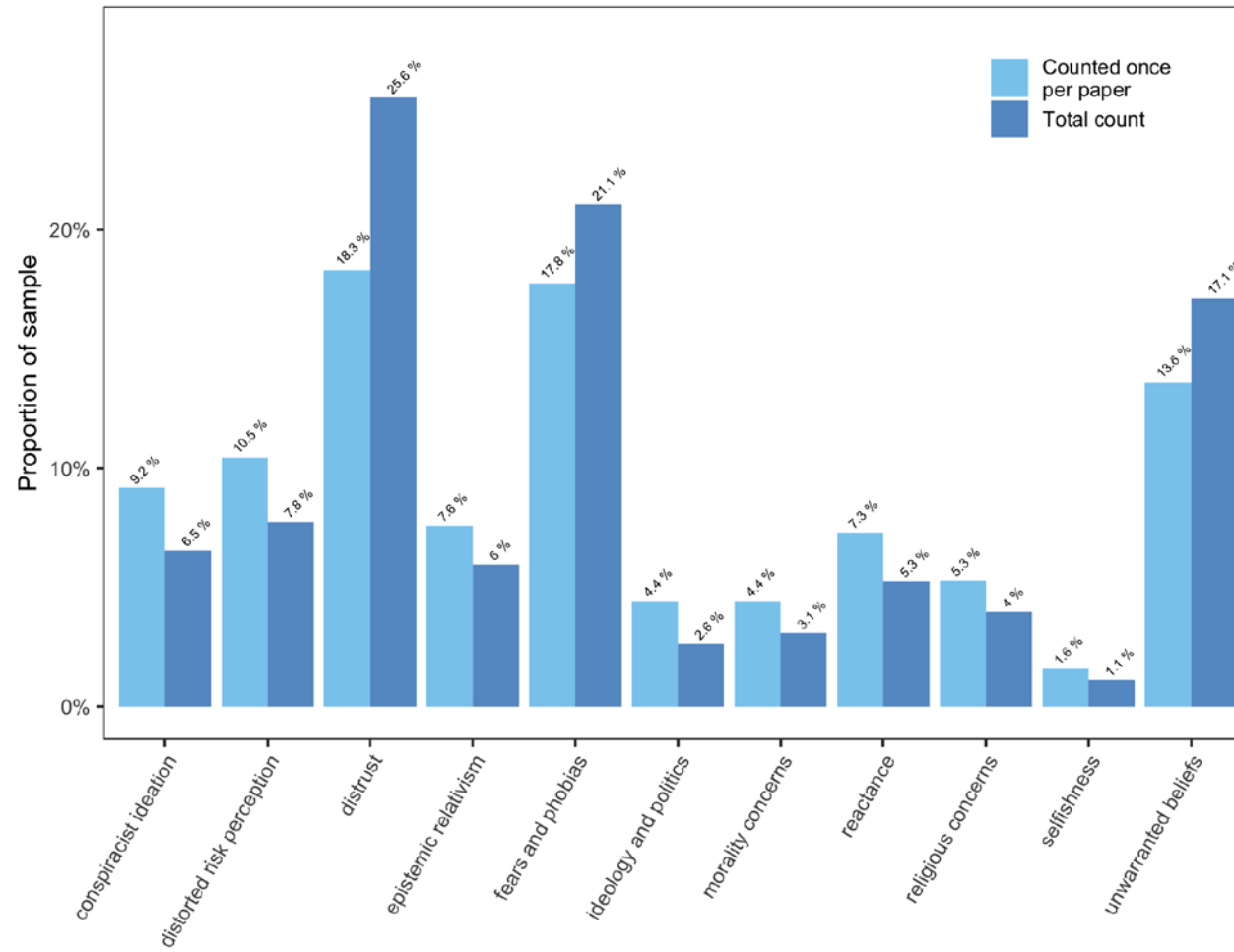
Religious concerns

Moral concerns

Fear and phobias

Distorted risk perception

Perceived self-interest

Epistemic relativism

Reactance

Distribution of Attitudinal Roots in Corpus

Distrust	18.3%
Fears and phobias	17.7%
Unwarranted beliefs	13.6%

…and 62 Themes

Epistemic relativism

Truth is relative — Rejection of scientific epistemology and objectivity, claiming that knowledge is relative.

Anecdotal evidence — Overreliance on testimonies and personal experiences that contradict scientific research.

Epistemic superiority — Alternative perspectives on health as more comprehensive and informed than science.

Individualistic epistemology — People are experts on their bodies and children (e.g., mother knows best, intuition and instinct).

All or nothing — Impossible expectations: We should not accept anything that is less than 100% effective.

Conspiracist ideation

Government cover-up — Health authorities and companies conspiring to withhold important information from the public.

Big Pharma — Pharmaceutical companies conspiring to profit over people or test new drugs.

Population control — Vaccines are a way to sterilize, depopulate or track the existing population.

Made up threat — Vaccine-preventable diseases are exaggerated or does not exist.

Targeting the disadvantaged — Vaccines are intended to harm disadvantaged groups or poor countries.
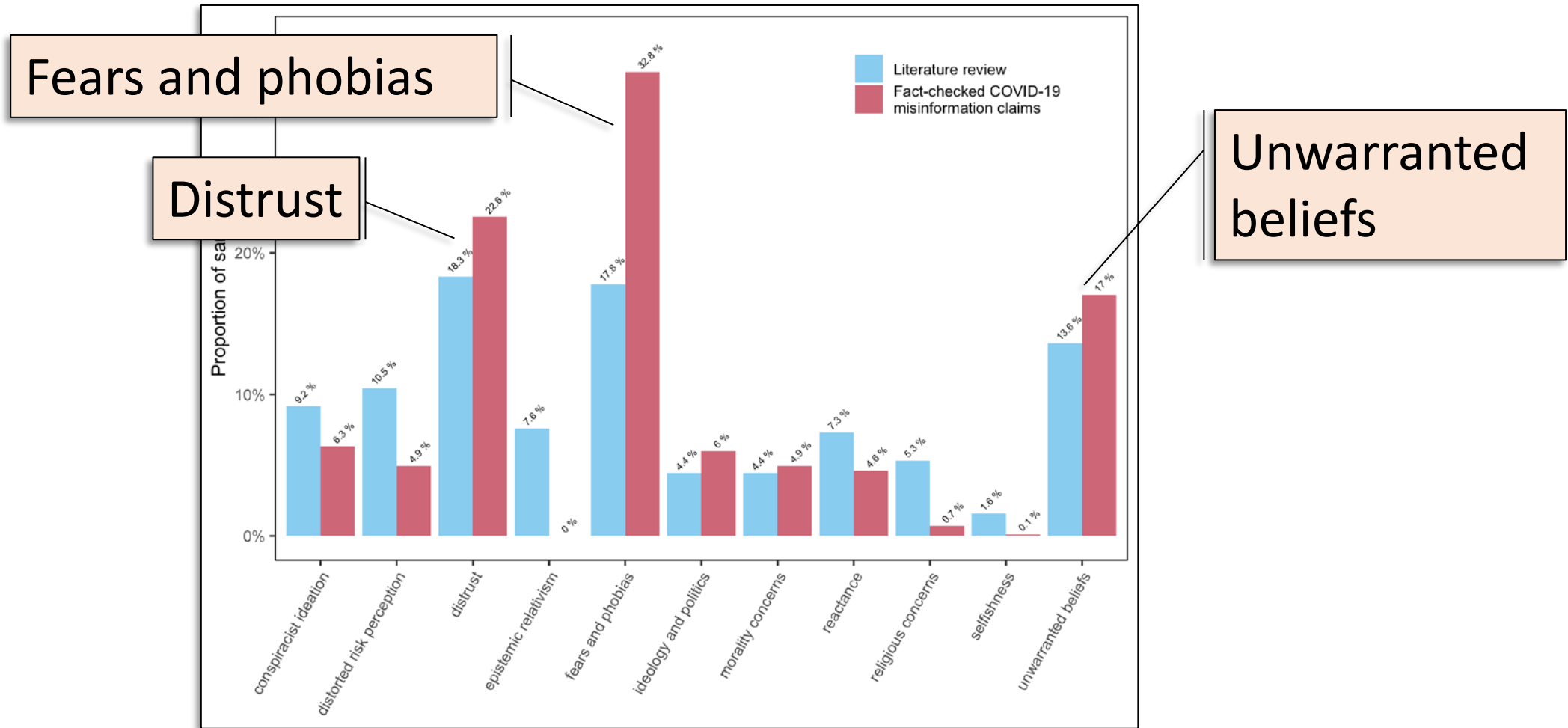
Unwarranted beliefs

Worldview/politics

Study 2: Debunks of COVID-19 Vaccine Misinformation

- Sample from https://covid19misinfo.org

- Collects debunks COVID-19 related misinformation, based on large ensemble of fact checkers

- We selected COVID-19 vaccination misinformation that has been debunked (N=623)

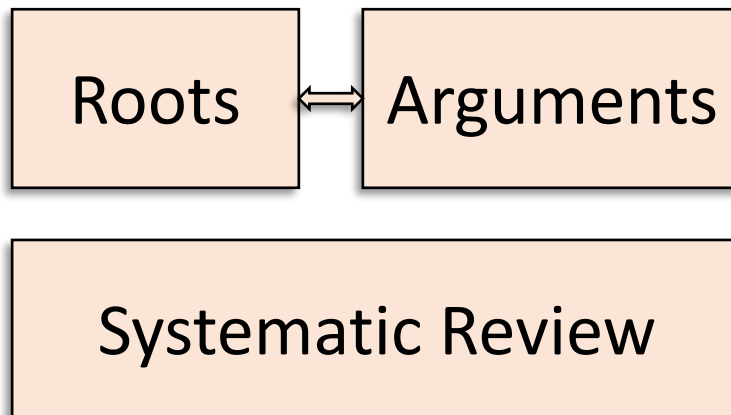- 2 reviewers independently assigned each headline claim (4-88 words, M=15.2) to attitude roots
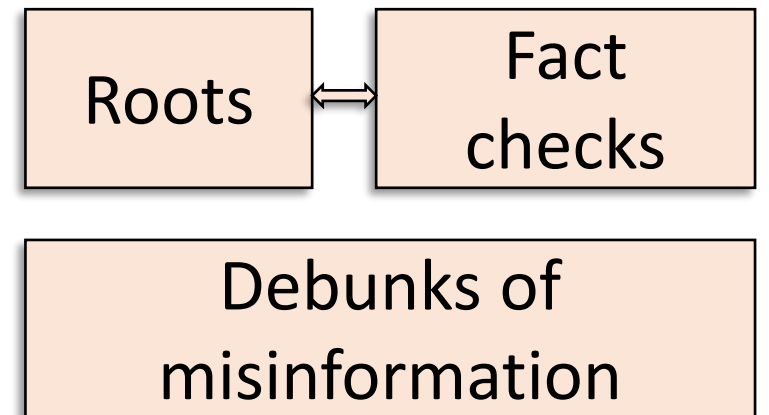
Comparing Study 1 and Study 2

Fears and phobias

Distrust

Unwarranted beliefs

Co-funded by the Horizon 2020 program
of the European Union

Confirmatory Topic Modeling

Study 1

Study 2

Roots ↔ Arguments

Systematic Review

Roots ↔ Fact checks

Debunks of misinformation
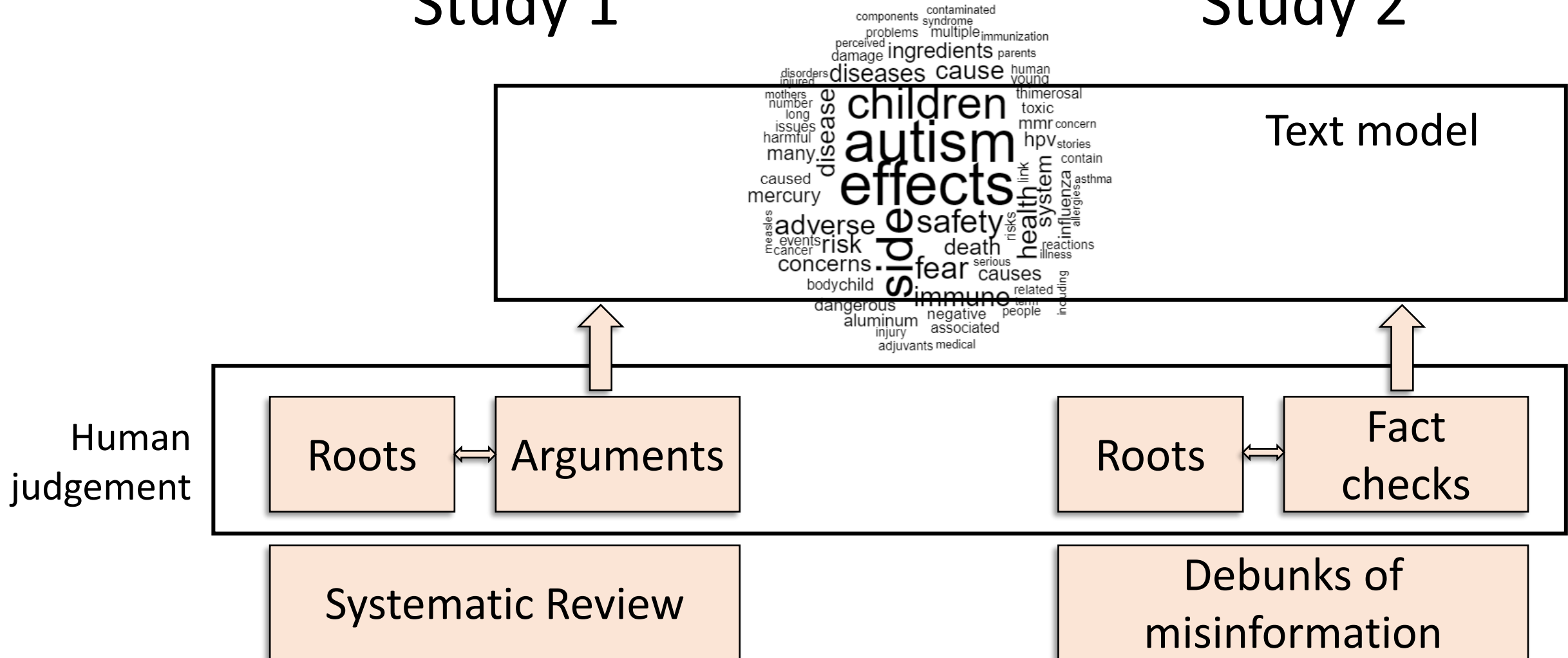
I. Latent Dirichlet Allocation (LDA; Blei, 2012)

- Unsupervised "Bag of Words" algorithm

- Uses a corpus of documents to
  - assign words to topics (β = probability of word in topic)
  - topics to documents (γ = probability of topic)

- Ideally, once trained on a corpus, topic model can generalize to new documents

"Arts"	"Budgets"	"Children"	"Education"
NEW	MILLION	CHILDREN	SCHOOL
FILM	TAX	WOMEN	STUDENTS
SHOW	PROGRAM	PEOPLE	SCHOOLS
MUSIC	BUDGET	CHILD	EDUCATION
MOVIE	BILLION	YEARS	TEACHERS
PLAY	FEDERAL	FAMILIES	HIGH
MUSICAL	YEAR	WORK	PUBLIC
BEST	SPENDING	PARENTS	TEACHER
ACTOR	NEW	SAYS	BENNETT
FIRST	STATE	FAMILY	MANIGAT
YORK	PLAN	WELFARE	NAMPHY
OPERA	MONEY	MEN	STATE
THEATER	PROGRAMS	PERCENT	PRESIDENT
ACTRESS	GOVERNMENT	CARE	ELEMENTARY
LOVE	CONGRESS	LIFE	HAITI

Labels inferred by researcher

Words with top values of β

Jitsu VAX

"Arts"	"Budgets"	"Children"	"Education"
NEW	MILLION	CHILDREN	SCHOOL
FILM	TAX	WOMEN	STUDENTS
SHOW	PROGRAM	PEOPLE	SCHOOLS
MUSIC	BUDGET	CHILD	EDUCATION
MOVIE	BILLION	YEARS	TEACHERS
PLAY	FEDERAL	FAMILIES	HIGH
MUSICAL	YEAR	WORK	PUBLIC
BEST	SPENDING	PARENTS	TEACHER
ACTOR	NEW	SAYS	BENNETT
FIRST	STATE	FAMILY	MANIGAT
YORK	PLAN	WELFARE	NAMPHY
OPERA	MONEY	MEN	STATE
THEATER	PROGRAMS	PERCENT	PRESIDENT
ACTRESS	GOVERNMENT	CARE	ELEMENTARY
LOVE	CONGRESS	LIFE	HAITI

The William Randolph Hearst Foundation will give $1.25 million to Lincoln Center, Metropolitan Opera Co., New York Philharmonic and Juilliard School. "Our board felt that we had a real opportunity to make a mark on the future of the performing arts with these grants an act every bit as important as our traditional areas of support in health, medical research, education and the social services," Hearst Foundation President Randolph A. Hearst said Monday in announcing the grants. Lincoln Center's share will be $200,000 for its new building, which will house young artists and provide new public facilities. The Metropolitan Opera Co. and New York Philharmonic will receive $400,000 each. The Juilliard School, where music and the performing arts are taught, will get $250,000. The Hearst Foundation, a leading supporter of the Lincoln Center Consolidated Corporate Fund, will make its usual annual $100,000 donation, too.

Seeded LDA
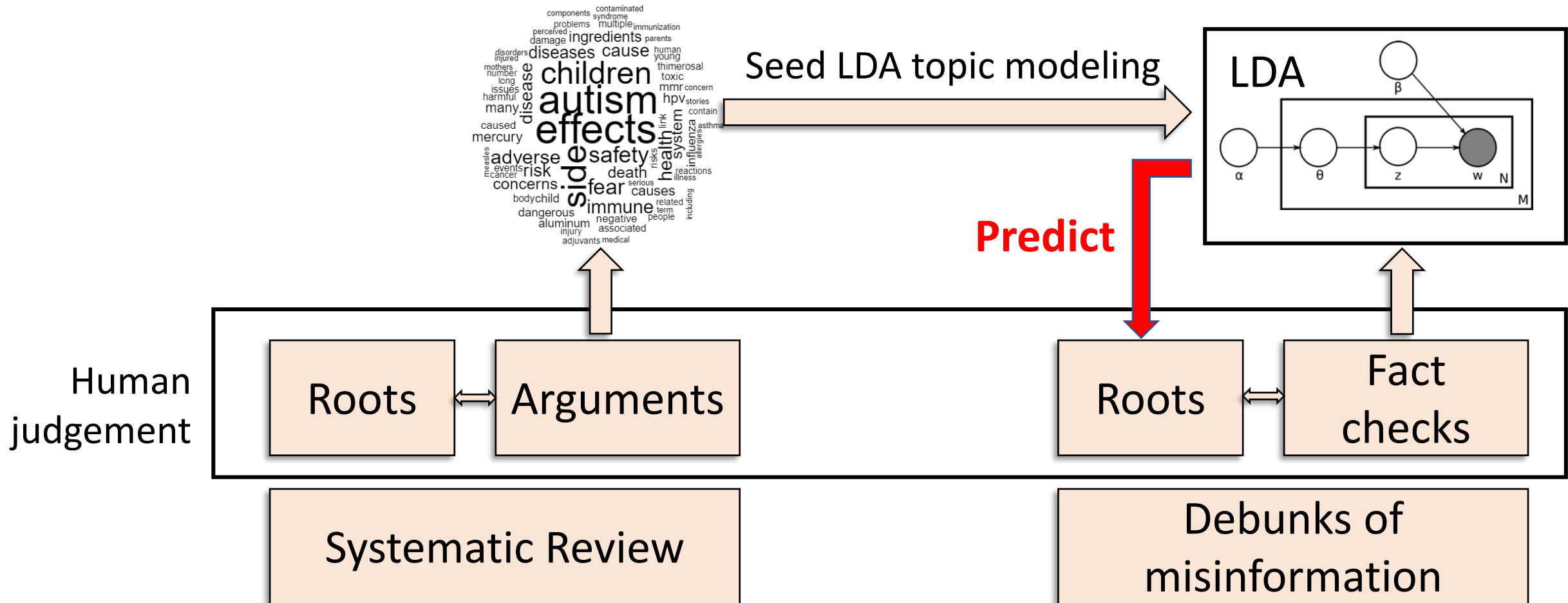
- Top 8 words of each attitude root in Study 1 used as seeds
  - weighted (β) to ensure blending of seed words with others identified by LDA
  - seeding predefined 11 topics corresponding to roots
- LDA done on full text of debunks in Study 2
  - corpus of 134K words
- Predominant topic (γ) for each document constitutes prediction of root
- Root predicted correctly 14.9% (chance = 9.09%)
- Top-2 accuracy 31.6% (chance = 18%)

Seeded LDA

- Top 8 words of each attitude root
  – weighted (β) to ensure blending of s
  – seeding predefined 11 topics cores
- LDA done on full text of debunks i
  – corpus of 134K words
- Predominant topic (γ) for each do
- Root predicted correctly 14.9% (ch
- Top-2 accuracy 31.6% (chance = 1

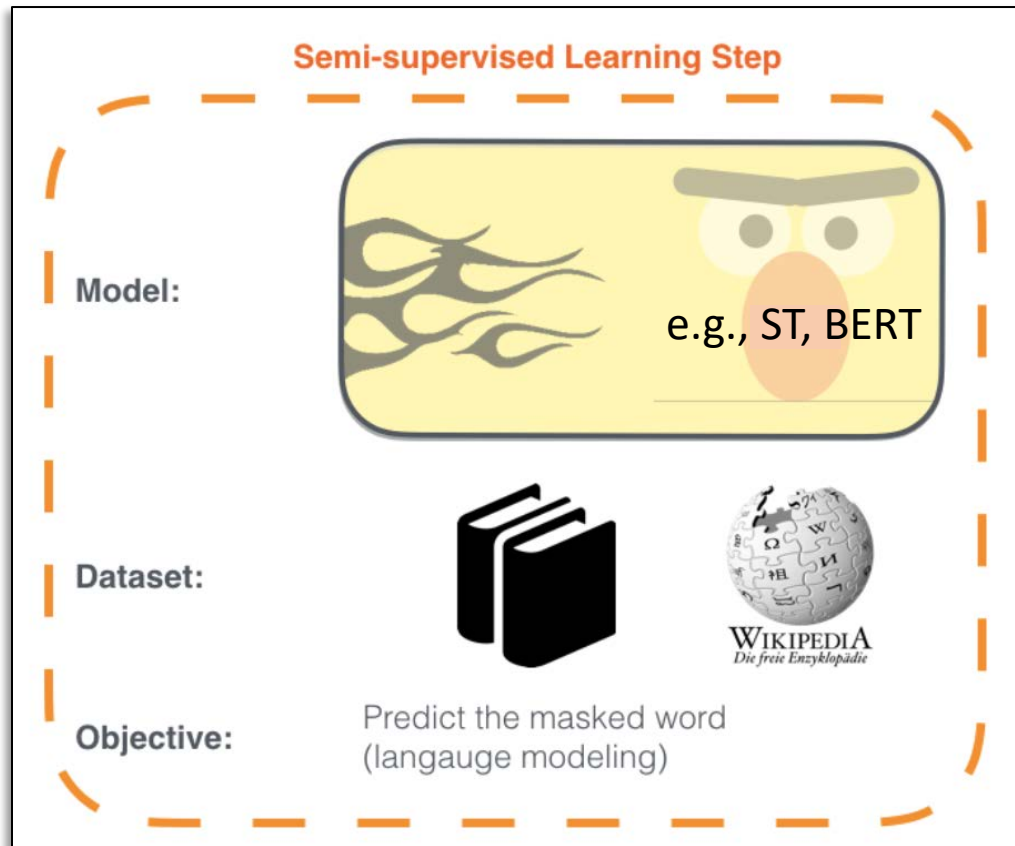Signal clearly detected albeit with low precision. But:

Study 1:

- 54 years of academic research

- 90% vaccines other than COVID

Study 2:

- ≅ 1 year COVID

- Debunks, not original arguments
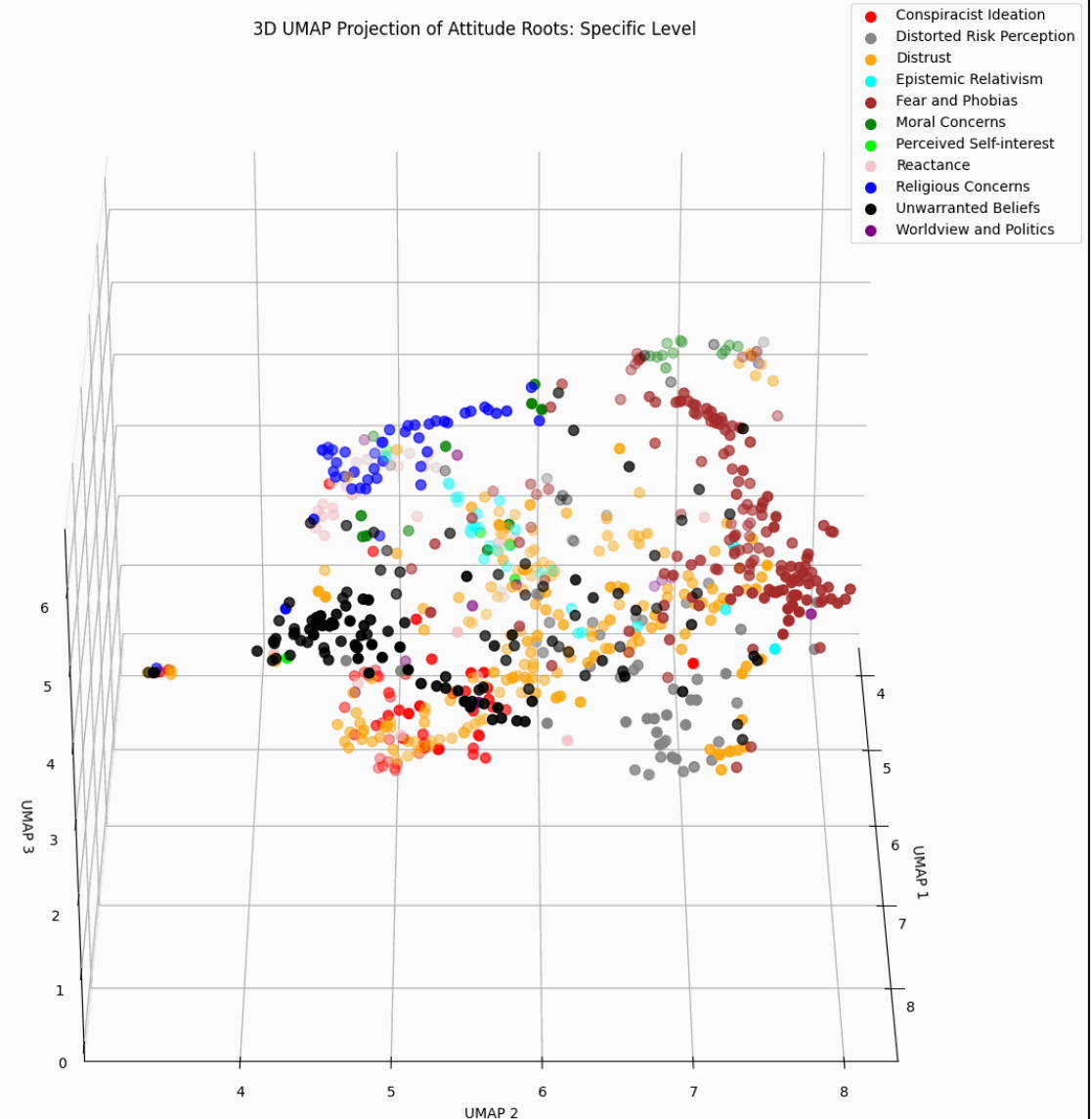
- Humans classified headlines, LDA used full text

II. Transformer Based Modeling

Semi-supervised Learning Step

Model: e.g., ST, BERT

Dataset: WIKIPEDIA Die freie Enzyklopädie

Objective: Predict the masked word (langauge modeling)

- Pretrained transformer
- Converts text into context-sensitive vector representation
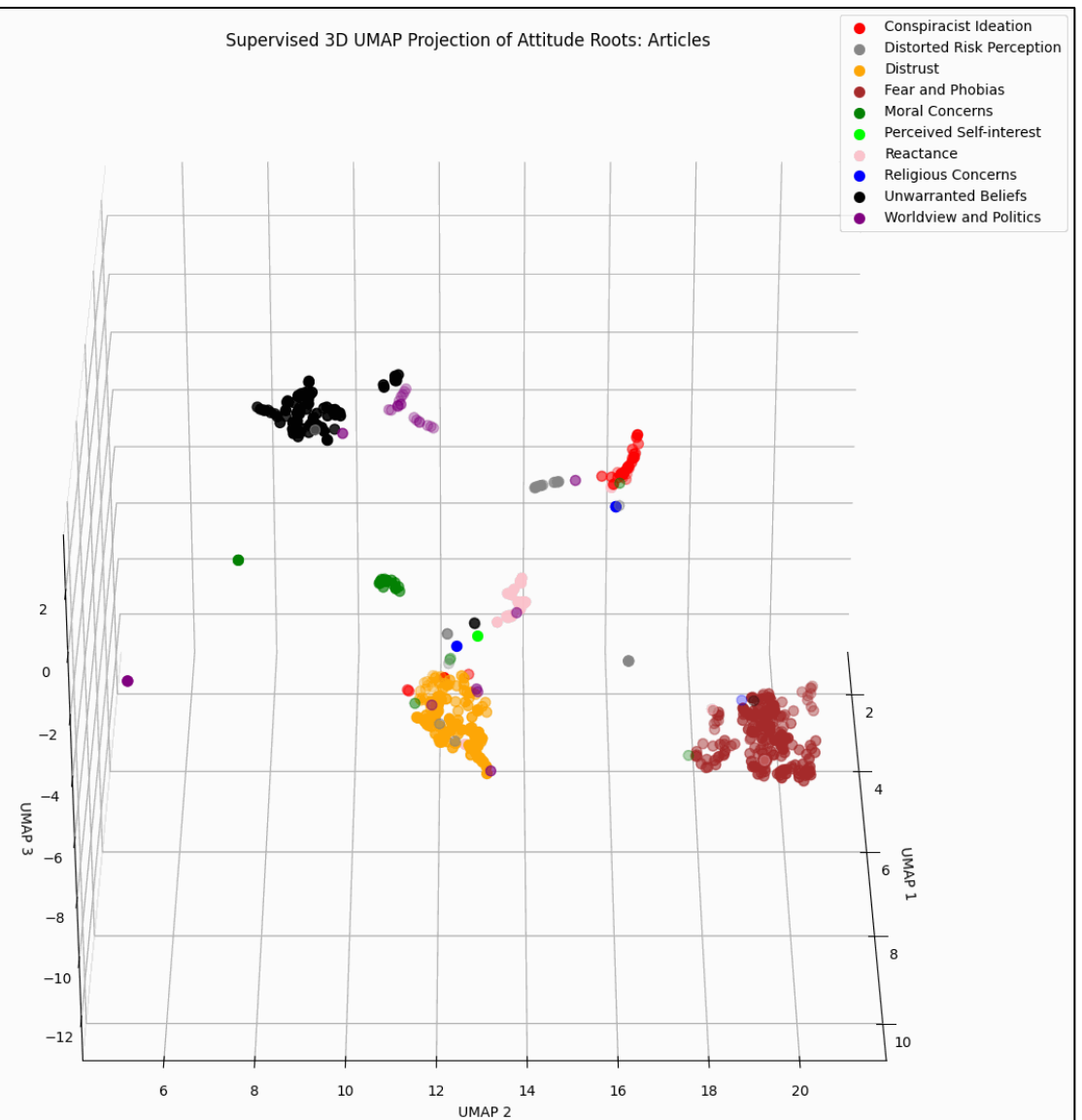- Optionally, can be fine tuned on new corpus

Study 1: ST

- Unsupervised with dimensionality reduction by UMAP

- No fine tuning
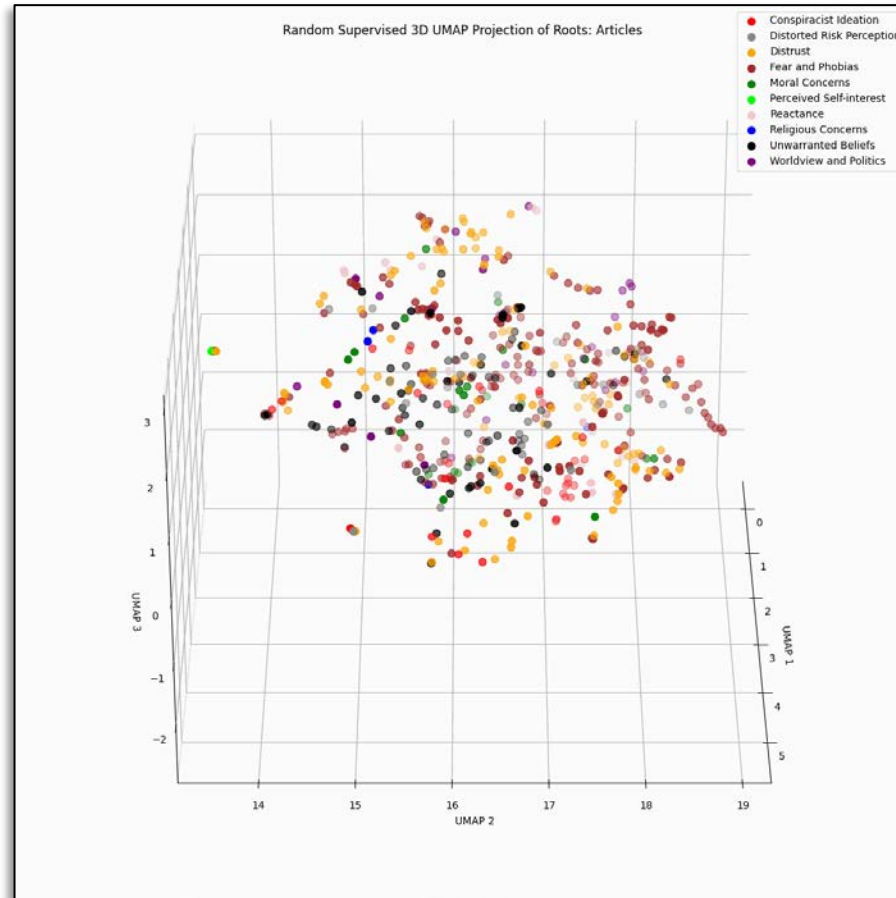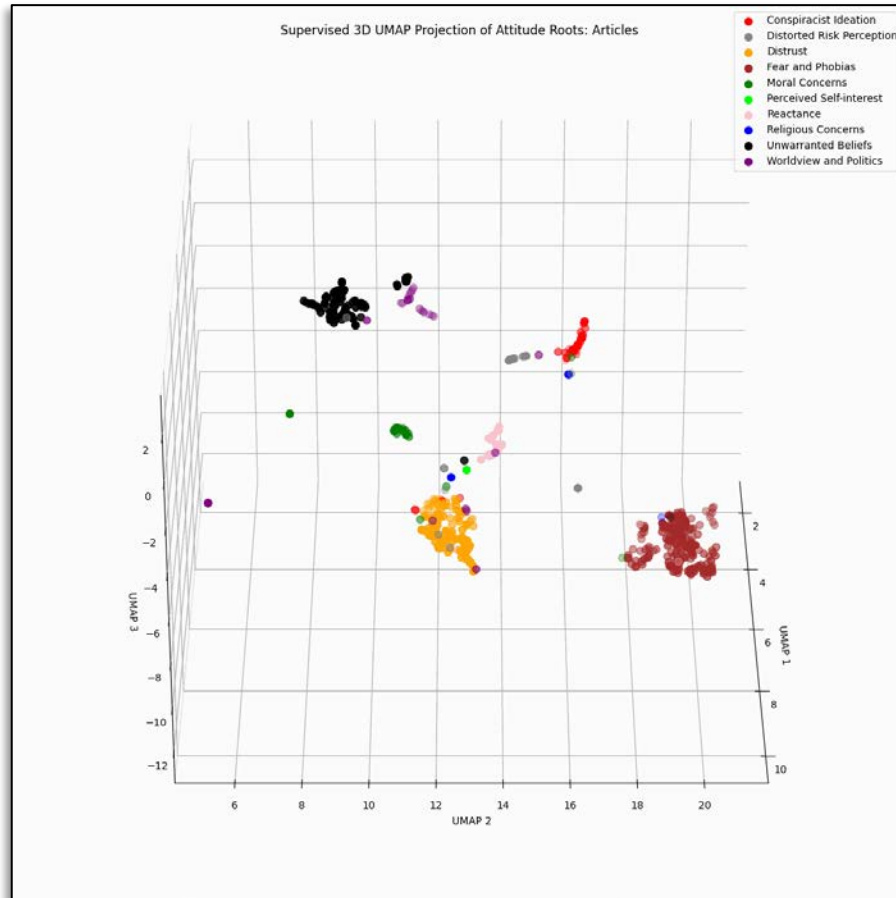
- No supervision

- No parameter estimation

3D UMAP Projection of Attitude Roots: Specific Level

Legend:
- Conspiracist Ideation
- Distorted Risk Perception
- Distrust
- Epistemic Relativism
- Fear and Phobias
- Moral Concerns
- Perceived Self-interest
- Reactance
- Religious Concerns
- Unwarranted Beliefs
- Worldview and Politics

Study 2: ST

- Supervised UMAP
- Used roots to guide UMAP

Supervised 3D UMAP Projection of Attitude Roots: Articles

Legend:
- Conspiracist Ideation
- Distorted Risk Perception
- Distrust
- Fear and Phobias
- Moral Concerns
- Perceived Self-interest
- Reactance
- Religious Concerns
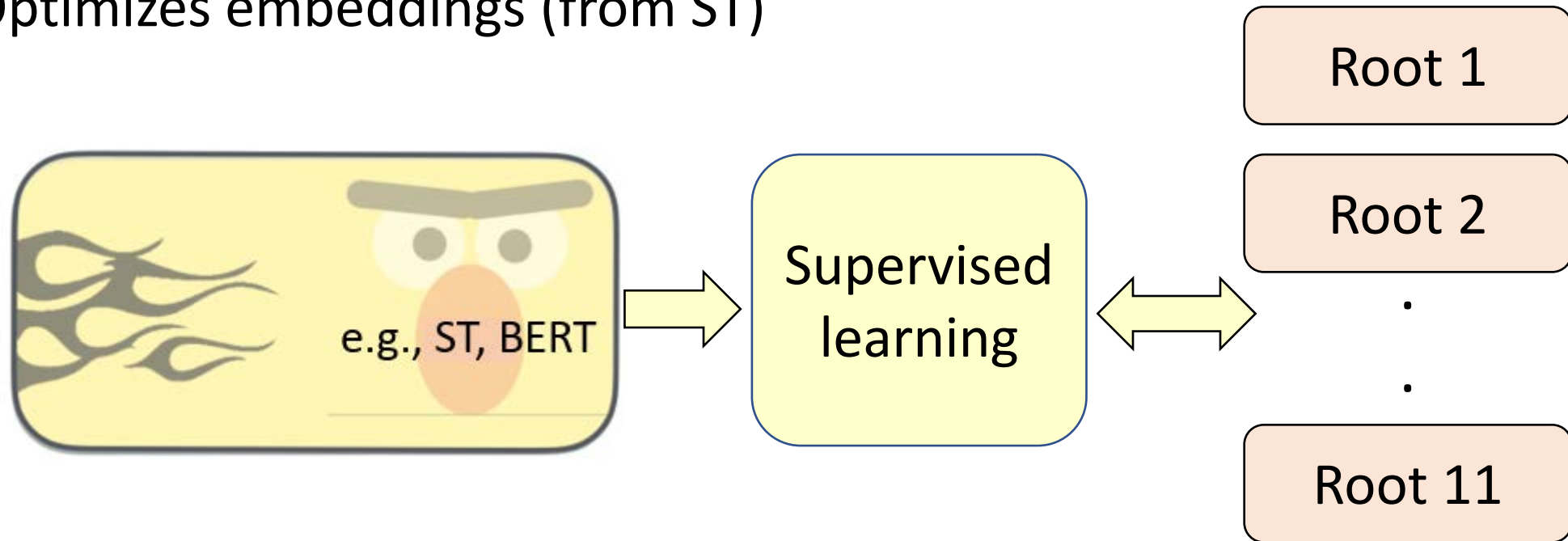- Unwarranted Beliefs
- Worldview and Politics

Study 2: Correct Supervision vs. Random

Study 1 and Study 2: Predict Roots

- Introduce fine tuning (SetFit)
- Optimizes embeddings (from ST)

e.g., ST, BERT

Supervised learning

Root 1

Root 2

.

.

Root 11

Macro F1 Results

Study 1		Study 2
	Two-step fine-tuning: SetFit	**54.0**
	Zero-shot: SetFit	**23.3**
	Standard fine-tuning: SetFit	**50.9**
	Baseline: ST + logistic regression	**43.7**
	Majority-class: "fear and phobias"	**14.8**

Conclusions from Transformer Modeling

- Strong signal can be extracted from both studies

- Fine-tuning yields additional boost

- There is transfer from Study 1 to Study 2 – confirming LDA modeling

Applying The Taxonomy

- Identify attitude roots by analyzing anti-vaccination arguments
- Create taxonomy of anti-vaccination arguments based on attitude roots and their manifestations
- Develop counter arguments or inoculation messages that are attitude-root-consonant.

An Empathetic Approach To Rebuttals

- JITSUVAX is developing an empathetic approach
  - patients' psychological attributes are considered an asset
  - people do not like being challenged
  - Jiu-jitsu: don't fight opponent head-on, leverage their force and use it against them

- Empathetic affirmation of patients' concerns.
- Provide counter-arguments that align with motivations (i.e., fit with the attitude root)

Example – Alternative Medicine

Argument: Traditional and natural remedies have a similar record of healing and no side effects.

Affirmation: Many people use alternative and traditional medicine without negative consequences. It is understandable to want to complement our medical treatments with remedies that fit our lifestyles and outlook on health. Many of these alternative remedies are even presented as safer and more natural than medical treatments, so they are very appealing.

Example – Alternative Medicine

Rebuttal: Vaccines have a longstanding record of effectiveness in boosting the body's natural immune response, especially when it comes to dangerous diseases that our immune systems struggle with. For example, COVID-19 vaccinations have saved nearly half a million lives in less than a year through generating natural immune protection without suffering the symptoms of the disease. Vaccinations are thus our first line of defense for these diseases. <u>Or to put it another way: vaccinations are nature at its best.</u>

# Example – Alternative Medicine

Rebuttal: Vaccines have a longstanding record of effectiveness in boosting the body's natural immune response, especially when it comes to dangerous diseases that our immune systems struggle with. For example, COVID-19 vaccinations have saved nearly half a million lives in less than a year through generating natural immune protection without suffering the symptoms of the disease. Vaccinations are thus our first line of defense for these diseases. Or to put it another way: vaccinations are nature at its best.

- Series of experiments shows that this approach works

Rebuttal Website For Practitioners

- Searchable data base indexed by roots and themes
- Provides empathic rebuttal
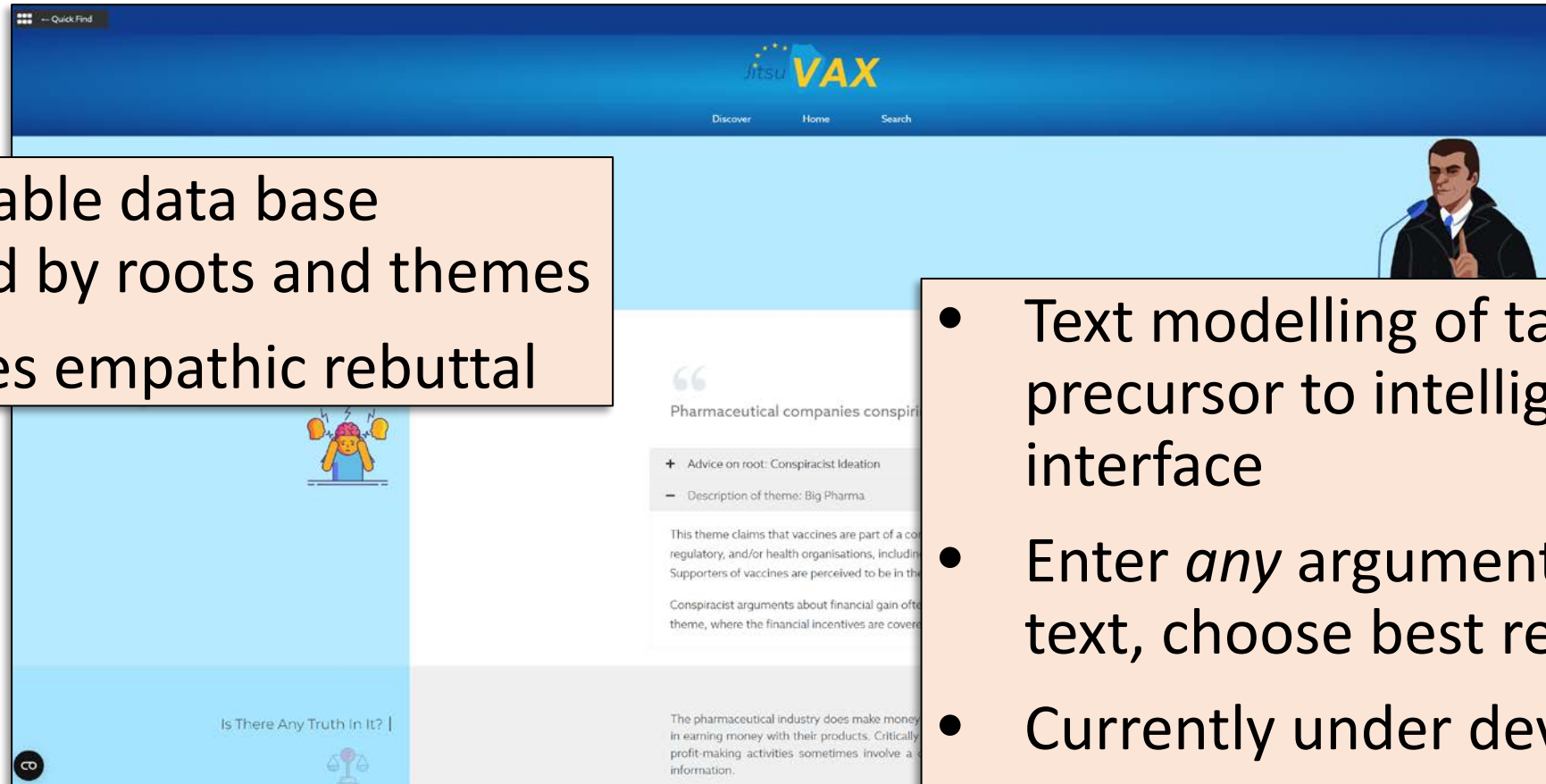
- Text modelling of taxonomy precursor to intelligent interface
- Enter *any* argument in free text, choose best rebuttal
- Currently under development

Website

https://jitsuvax.info/