

HLV in Ideological Coherence in Political Speech

Karolina Zaczynska, Julian Schlenker, Ines Rehbein,
Simone Paolo Ponzetto
University of Mannheim

HLV-DP Workshop @ ESSLLI 2026
August 10-14, 2026

Project Context: UNCOVER – Uncovering Ideological Frames in Political Discourse



Previous work



- Moral frame extraction (Haidt et al., 2009; Graham et al., 2013)
 - Situation entity extraction (Smith, 2003)
 - Stance detection
 - Narrative roles (Shanahan et al., 2017)
 - etc.
- [i.a. Rehbein 2026a, Rehbein 2026b, Schlenker 2026]



This work

Political ideology annotation

Broader Motivation: Political Ideology annotation

Overall Goal:

- Investigate whether political actors & parties express coherent ideological positions across their discourse

Research interests:

- Are the positions expressed by a party **coherent or potentially contradictory**?
- Can ideological coherence reveal **within-party variation** across legislative periods?
- etc.

Broader Motivation: Political Ideology annotation

Overall Goal:

- Investigate whether political actors & parties express coherent ideological positions across their discourse

Research interests:

- Are the positions expressed by a party **coherent or potentially contradictory**?
- Can ideological coherence reveal **within-party variation** across legislative periods?

To study coherence, we first need to establish the **political position** expressed in individual excerpts.

Project Context: UNCOVER – Uncovering Ideological Frames in Political Discourse



Previous work



- Moral frame extraction (Haidt et al., 2009; Graham et al., 2013)
 - Situation entity extraction (Smith, 2003)
 - Stance detection
 - Narrative roles (Shanahan et al., 2017)
 - etc.
- [i.a. Rehbein 2026a, Rehbein 2026b, Schlenker 2026]



This work

Political ideology annotation

1. Political positioning of individual excerpts

2. Ideological coherence annotation between excerpts

Project Context: UNCOVER – Uncovering Ideological Frames in Political Discourse



Previous work

- Moral frame extraction (Haidt et al., 2009; Graham et al., 2013)
 - Situation entity extraction (Liu et al., 2003)
 - Stance detection
 - Narrative roles (Shafarman et al., 2017)
 - etc.
- [i.a. Rehbein 2026a, F. Schlenker 2026]

Focus of this talk

Work in progress

This work

Political ideology annotation

1. Political positioning of individual excerpts

2. Ideological coherence annotation between excerpts

Motivation: **HLV Case Study on Political Positioning**

- Ideological positioning in political speech is **inherently subjective**
- What can the annotation data tell us about Human Label Variation?

Research Questions:

- Where do annotators (dis)agree on political positioning?
- Does disagreement reflect **ambiguous political language** or **genuinely unclassifiable statements**?
- How consistent are annotation decisions **across annotators** and for **repeated annotations**?

Annotation Task Overview

Annotators

- **8 annotators** from political and social science
- Passed test phase

Input

- Excerpts from German parliamentary speeches
- 3 topics: Immigration, Pension, and Climate Change
- Each excerpt contains one **highlighted morally loaded statement**

Task 1: Political Positioning

- Given the excerpt, which political values does the speaker express?

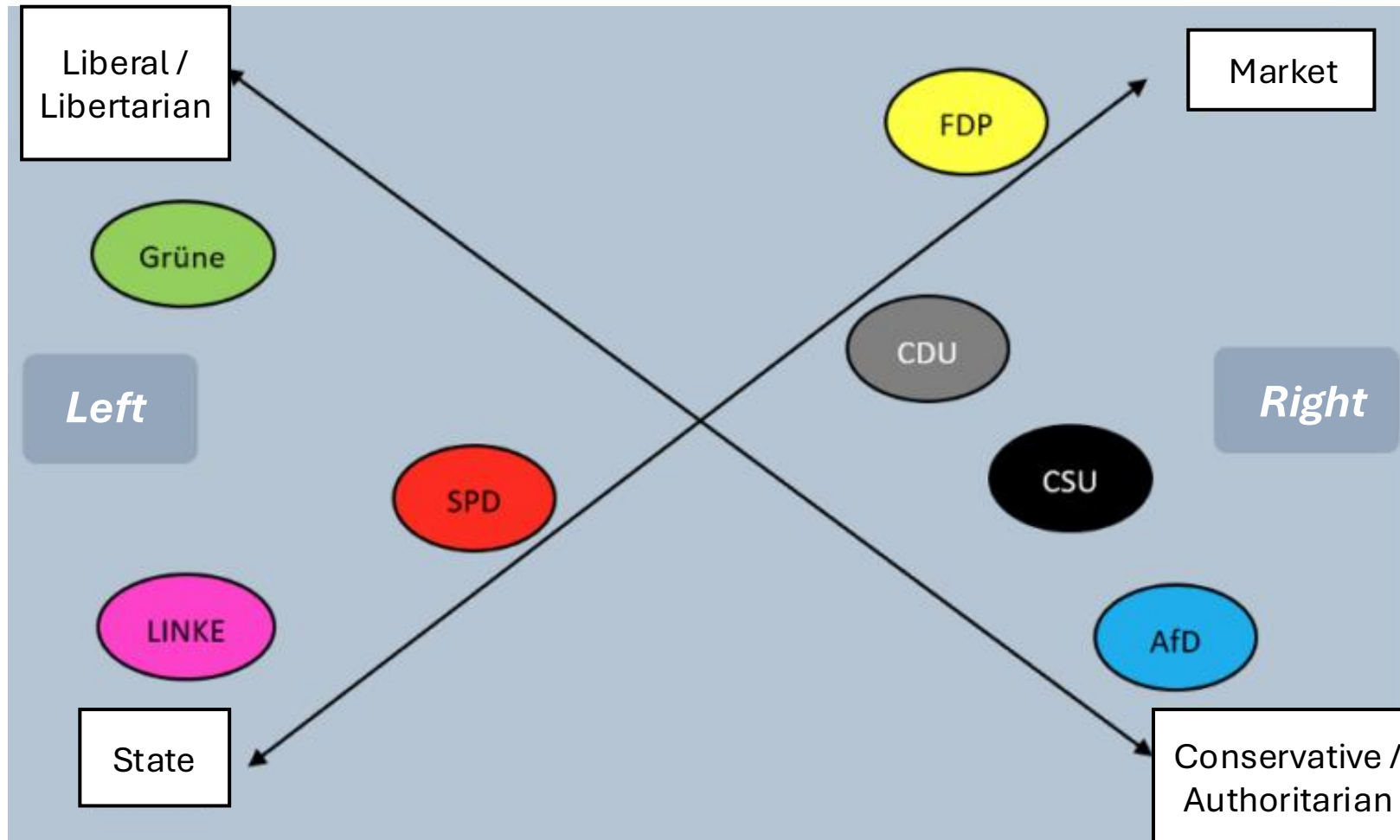
Task 2: Ideological Coherence

- Compare the anchor excerpt with **four candidate excerpts**
- Rate how similar their political positions are



Part II - Annotation Setup Political Positioning

Task 1: Political Positioning Annotation



Bergmann-Rusche-Image

Task 1: Political Positioning Annotation

Instance

“One last thing. You’re all **supporting the students in their Fridays for Future movement**. If you’re serious about that, then cut military spending and invest in climate protection and the welfare state instead of continuing to increase military spending.”

Socio-cultural scale
(libertarian ↔
authoritative)

Q1: Does the speaker represent more **liberal** or **authoritative** values?

liberal	rather liberal	rather authoritative	authoritative	underspecified
☐	☐	☐	☐	☐

Socio-economic scale
(state-oriented ↔ market-
freedom)

Q2: Does the speaker represent more **state-oriented** or **market-freedom** values?

Social justice	rather s-j	rather neoliberal	neoliberal	underspecified
☐	☐	☐	☐	☐

Task 1: Political Positioning Annotation

Instance

“One last thing. You’re all **supporting the students in their Fridays for Future movement**. If you’re serious about that, then cut military spending and invest in climate protection and the welfare state instead of continuing to increase military spending.”

Socio-cultural scale
(liberal ↔ authoritative)

Q1: Does the speaker represent more **liberal** or **authoritative** values?

liberal	rather liberal	rather authoritative	authoritative	underspecified
☒	☐	☐	☐	☐

Socio-economic scale
(state-oriented ↔ market-freedom)

Q2: Does the speaker represent more **state-oriented** or **market-freedom** values?

Social justice	rather s-j	rather neoliberal	neoliberal	underspecified
☒	☐	☐	☐	☐

Task 1: Political Positioning Annotation

Instance

“One last thing. You’re all **supporting the students in their Fridays for Future movement**. If you’re serious about that, then cut military spending and invest in climate protection and the welfare state instead of continuing to increase military spending.”

Socio-cultural scale
(liberal ↔ authoritative)

Q1: Does the speaker represent more **liberal** or **authoritative** values?

ORDINAL

liberal

☒

rather
liberal

☐

rather
authoritative

☐

authoritative

☐

underspecified

☐

Socio-economic scale
(state-oriented ↔ market-
freedom)

Q2: Does the speaker represent more **state-oriented** or **market-freedom** values?

ORDINAL

Social
justice

☒

rather
s-j

☐

rather
neoliberal

☐

neoliberal

☐

underspecified

☐

Task 1: Political Positioning Annotation

Instance

“One last thing. You’re all **supporting the students in their Fridays for Future movement**. If you’re serious about that, then cut military spending and invest in climate protection and the welfare state instead of continuing to increase military spending.”

Socio-cultural scale
(liberal ↔ authoritative)

Q1: Does the speaker represent more **liberal** or **authoritative** values?

ORDINAL

liberal

☒

rather
liberal

☐

rather
authoritative

☐

authoritative

☐

underspecified

☐

Socio-economic scale
(state-oriented ↔ market-
freedom)

Q2: Does the speaker represent more **state-oriented** or **market-freedom** values?

ORDINAL

Social
justice

☒

rather
s-j

☐

rather
neoliberal

☐

neoliberal

☐

underspecified

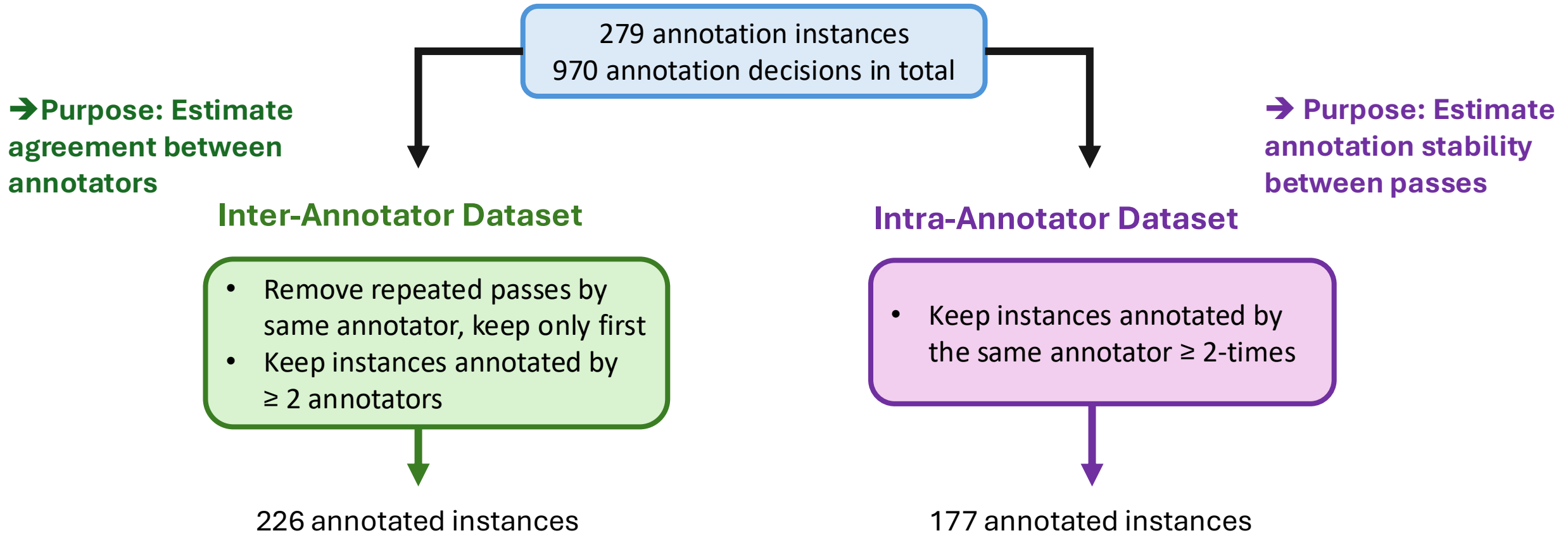
☐

Part III - Corpus

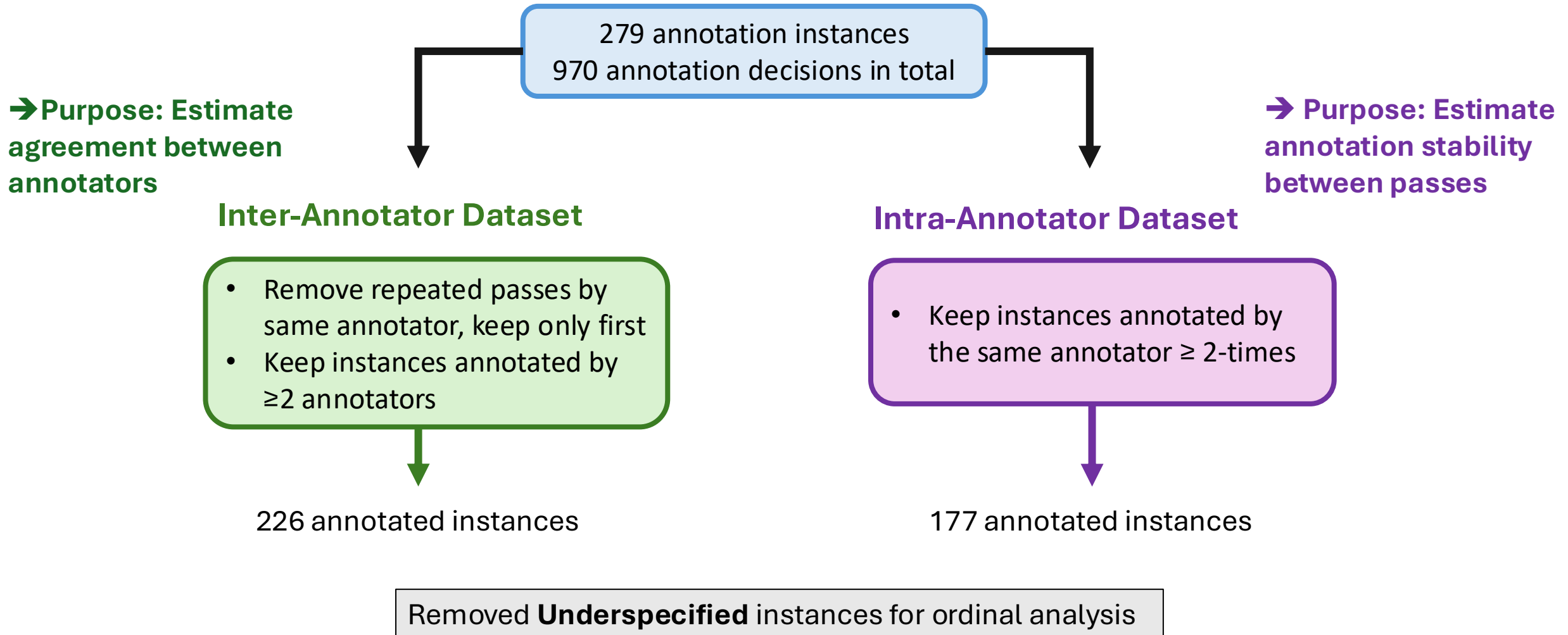
Analysis Datasets

- We have both annotation instances annotated by **more than one annotator...**
- and annotated by the **same annotator repeated times (more than one pass)**

Analysis Datasets

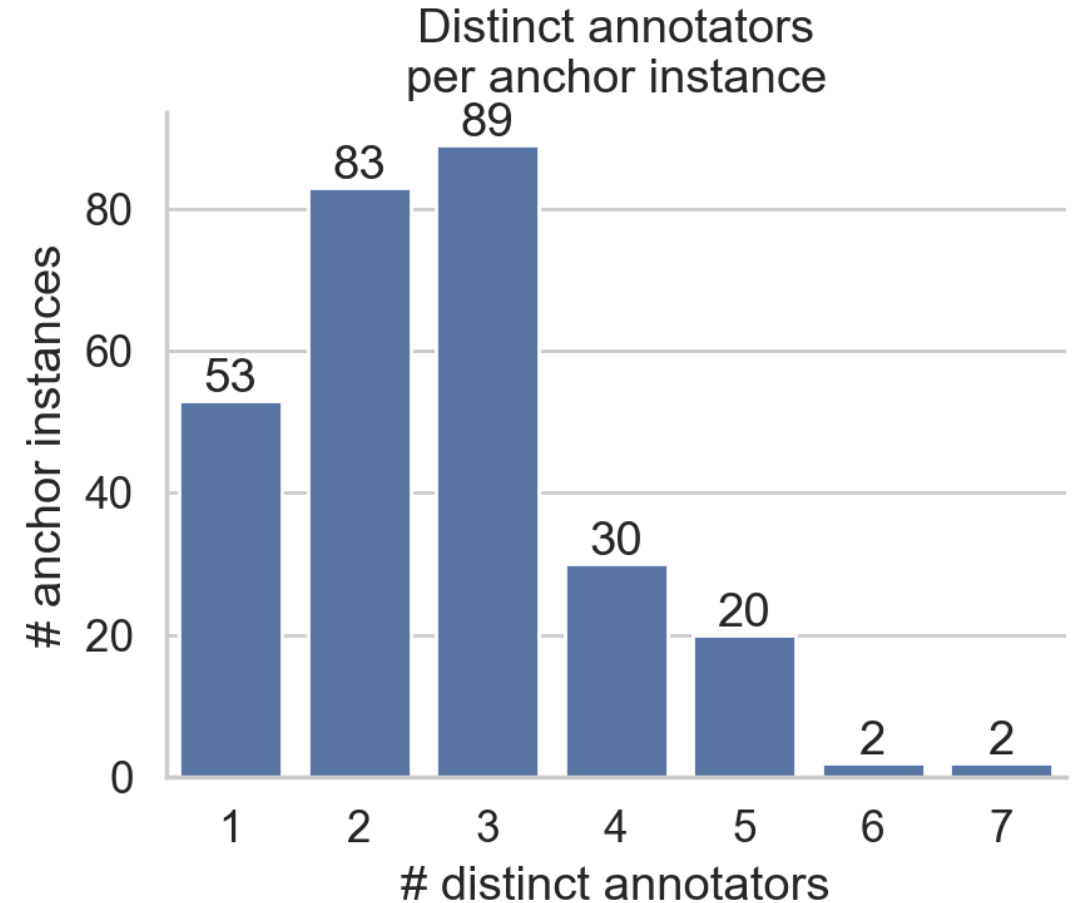


Analysis Datasets



Inter-Annotator Dataset

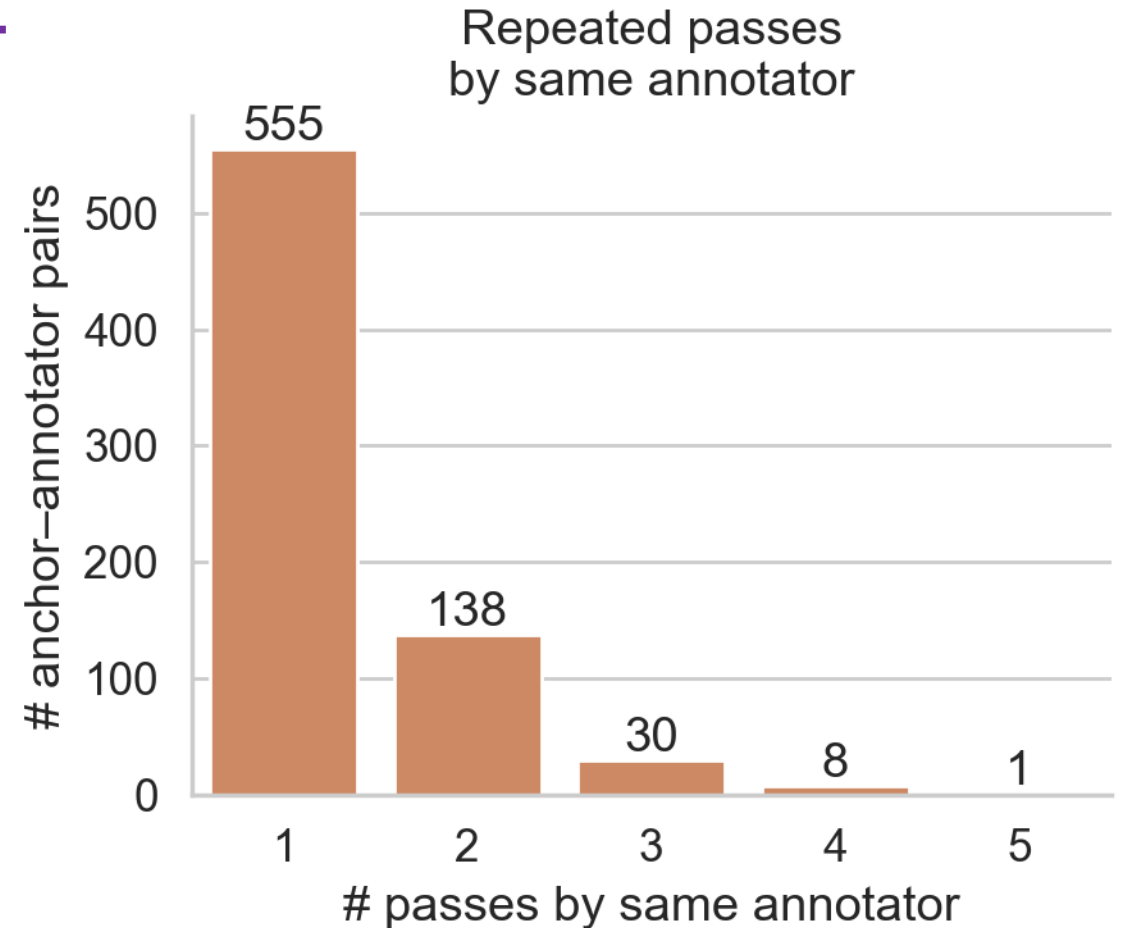
- 279 anchor instances
- 226 (81%) annotated by at least two annotators
- Median number of annotators per instance: 3.0
- Maximum: 7 annotators



Dataset: first annotation per annotator-instance

Intra-Annotator Dataset

- 177 annotated with repeated annotations
- Most repetitions occurred exactly twice (138)
- Maximum repetitions for one annotation instance: 5



Part IV - Analysis

“There is only one way to be accurate but there are many ways to be inaccurate.” (paraphrased from [Passonneau & Carpenter 2013](#))

Inter-Annotator-Agreement (IAA)

How much disagreement exists?

Agreement is rather low. **Why?**

Scale	Krippendorff's alpha (ordinal)
Socio-economic	0.602
Socio-cultural	0.657

N annotations (first pass)	N excluded (Underspecified)	Underspecified rate
732	174	23.8%
732	103	14.1%

Dataset: first annotation per annotator-instance (**Underspecified excluded** for ordinal α)

Inter-Annotator-Agreement (IAA)

How much disagreement exists?

Agreement is rather low. **Why?**

Scale	Krippendorff's alpha (ordinal)
Socio-economic	0.602
Socio-cultural	0.657

N annotations (first pass)	N excluded (Underspecified)	Underspecified rate
732	174	23.8%
732	103	14.1%

Dataset: first annotation per annotator-instance (**Underspecified excluded** for ordinal α)

Annotator-Uncertainty

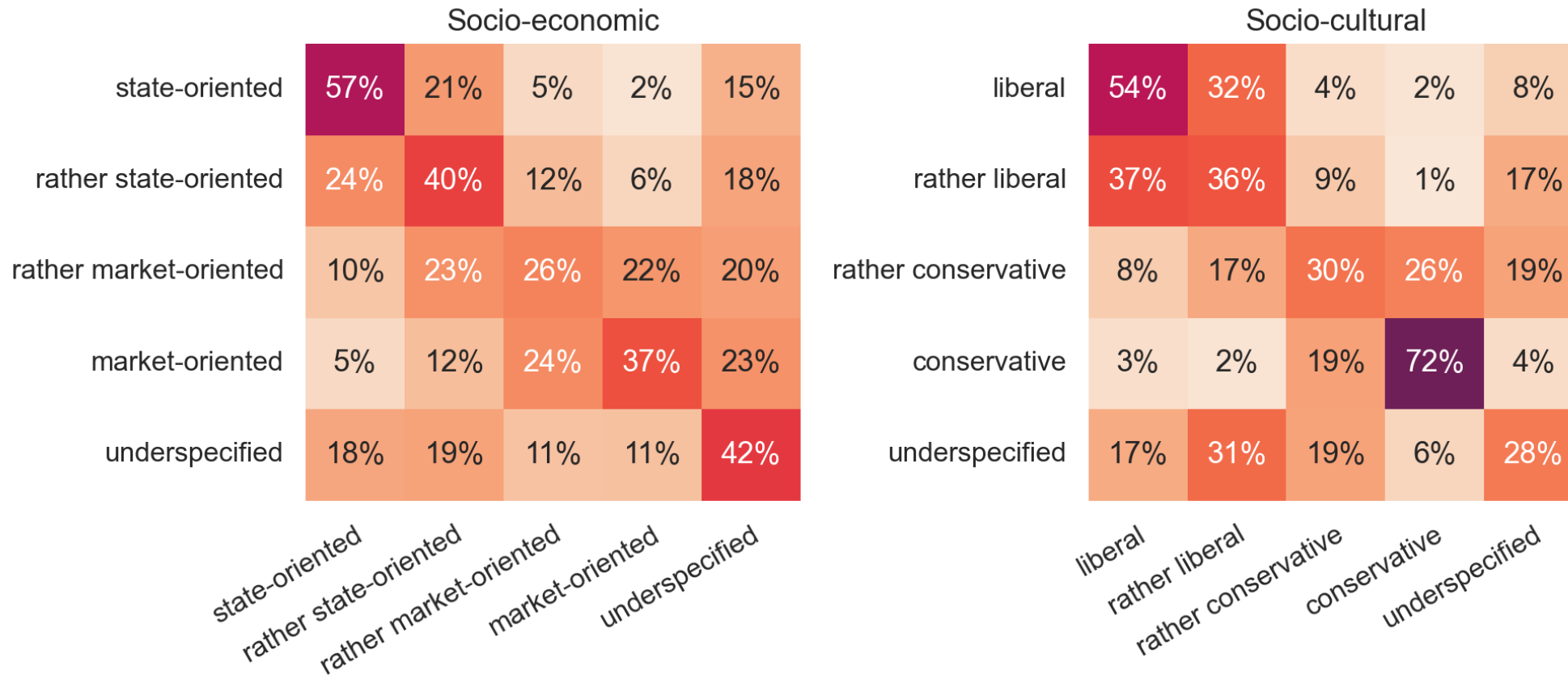
What makes an item difficult?

Or: Is disagreement caused by uncertain items or by competing interpretations?

Agreement	Uncertainty
Do annotators agree?	Can annotators decide at all?

Inter-Annotator-Agreement (IAA)

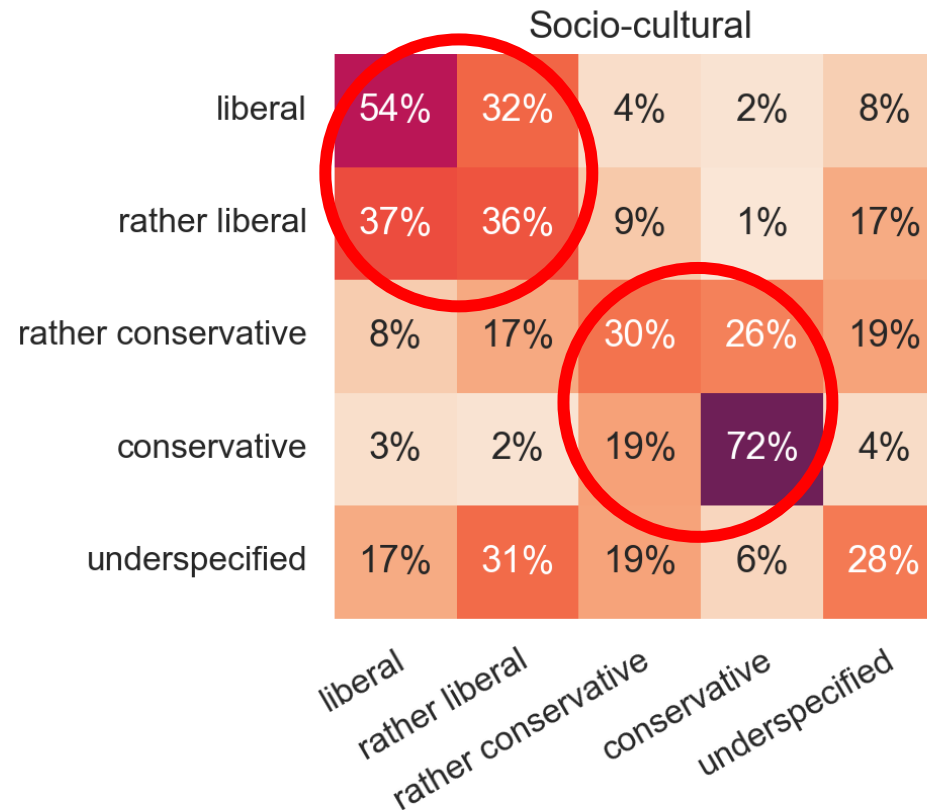
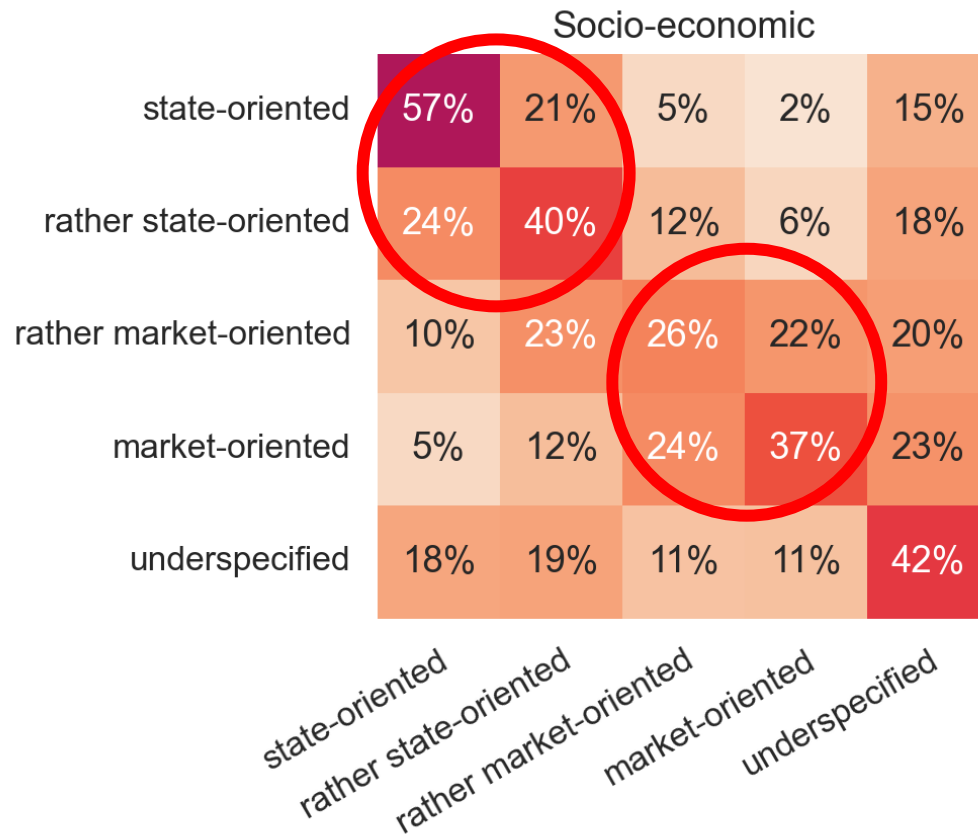
Where does disagreement occur?



Row-normalised confusion matrix, with all annotation comparisons aggregated.

Inter-Annotator-Agreement (IAA)

Where does disagreement occur?

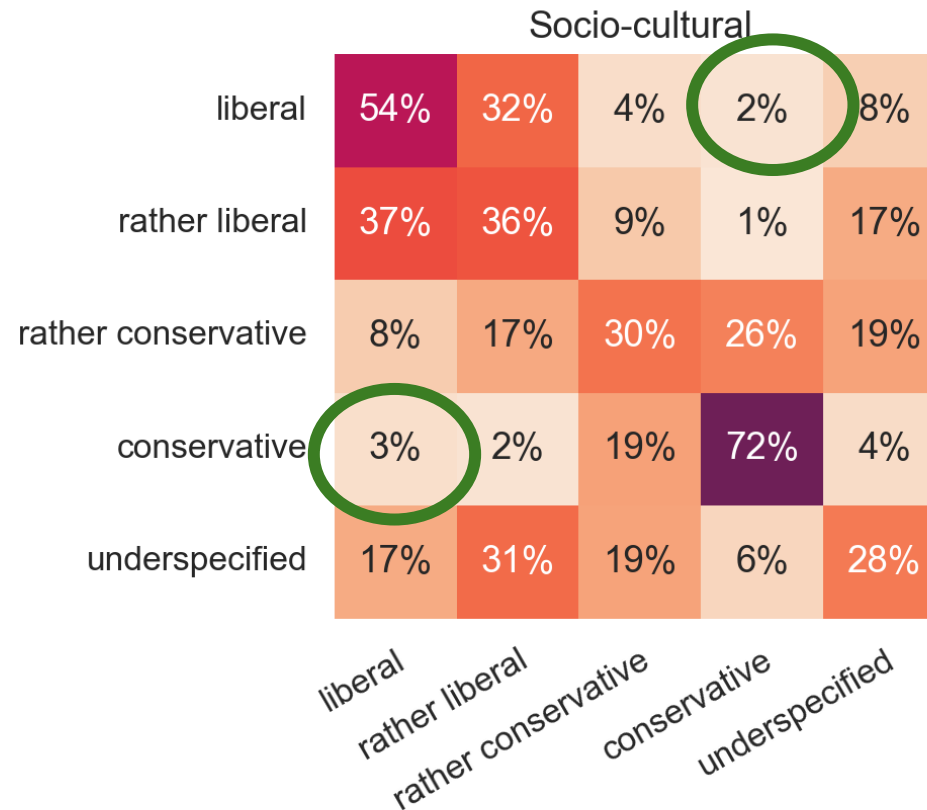
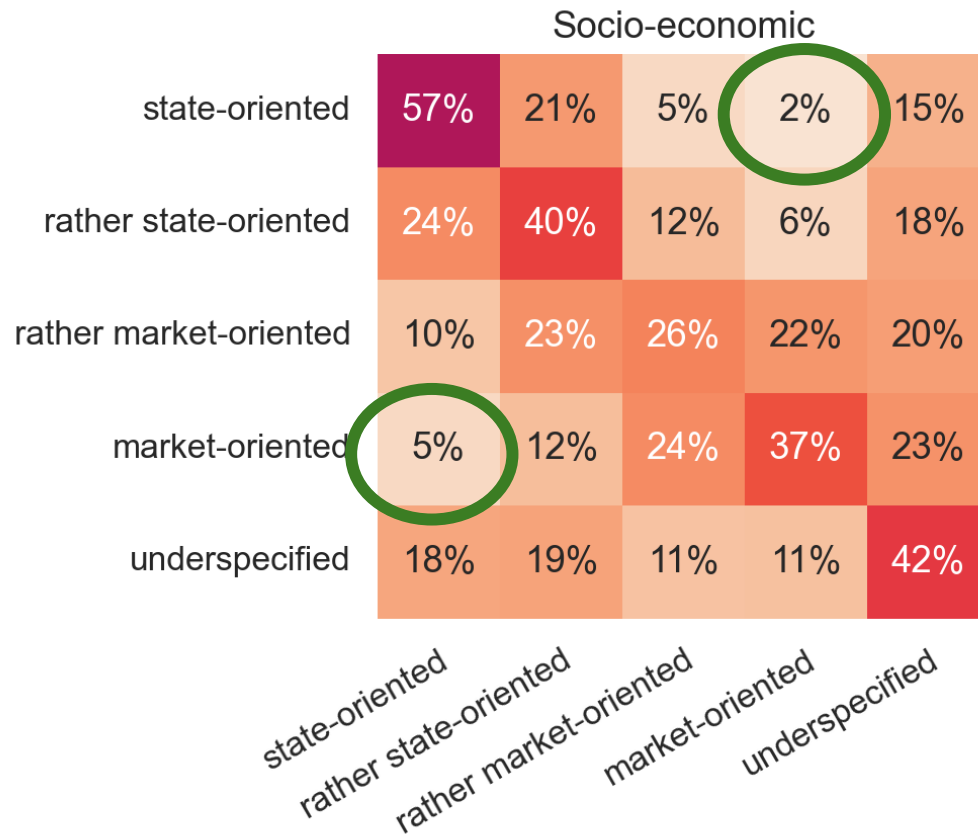


Disagreements
mostly occur
between
neighbouring
categories

Row-normalised confusion matrix, with all annotation comparisons aggregated.

Inter-Annotator-Agreement (IAA)

Where does disagreement occur?



Disagreements
mostly occur
between
neighbouring
categories

Political extreme
positions are rarely
confused

Row-normalised confusion matrix, with all annotation comparisons aggregated.

Uncertainty

Are some positions less clearly identifiable?

Underspecified rate by majority label

Annotation Instance

"We must protect..."

Annotation Decisions

Conservative
Conservative
Underspecified
Conservative



Majority label
=
conservative

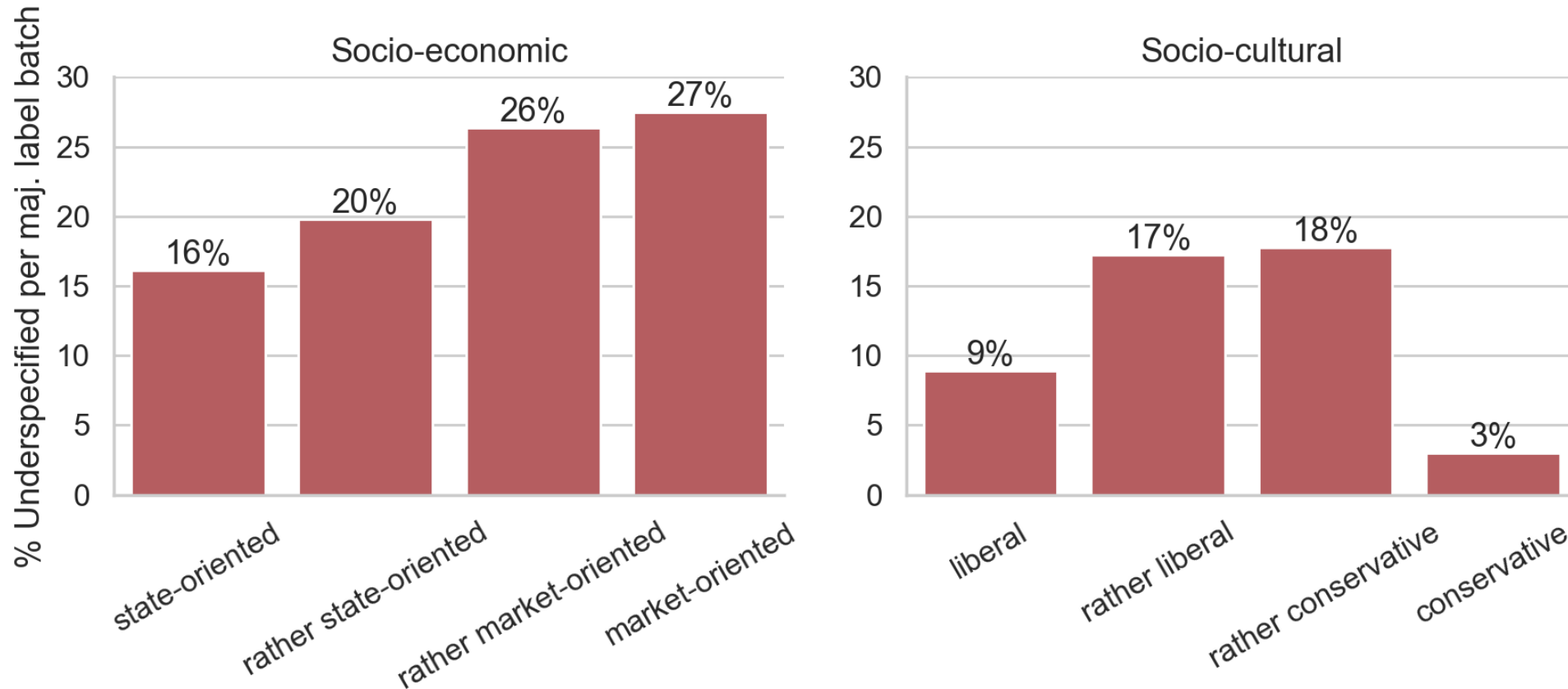


How high is Underspecified rate for **all instances with majority label X?**

Uncertainty

Are some positions less clearly identifiable?

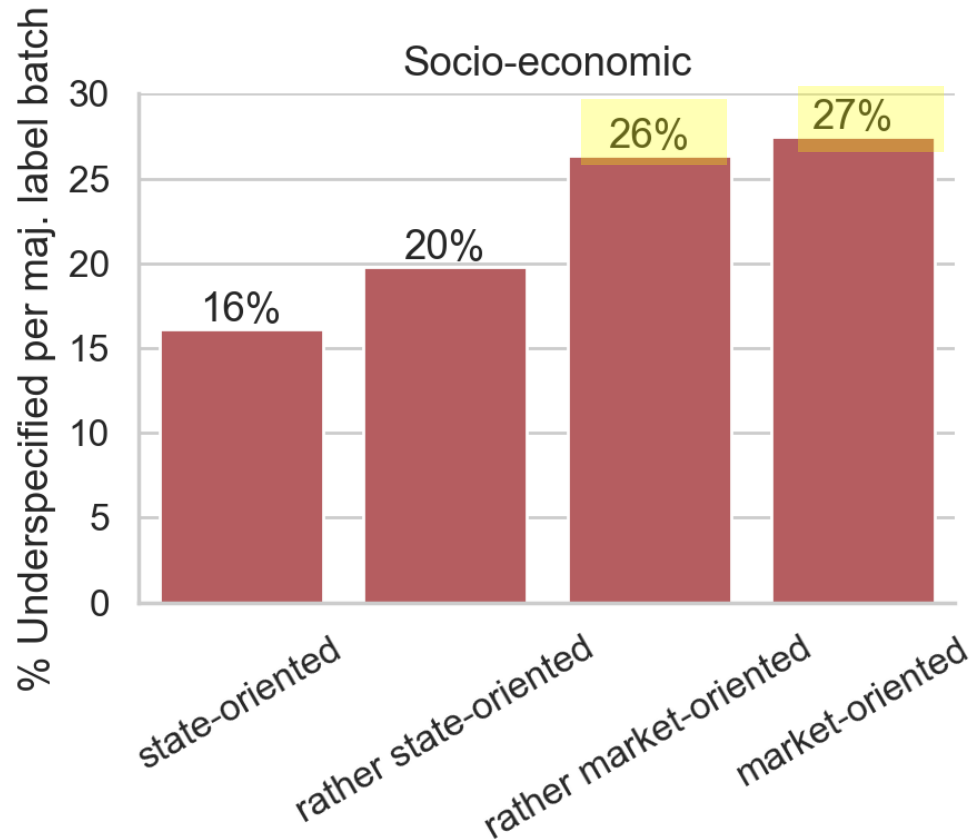
Underspecified rate by majority label



Uncertainty

Are some positions less clearly identifiable?

Underspecified rate by majority label

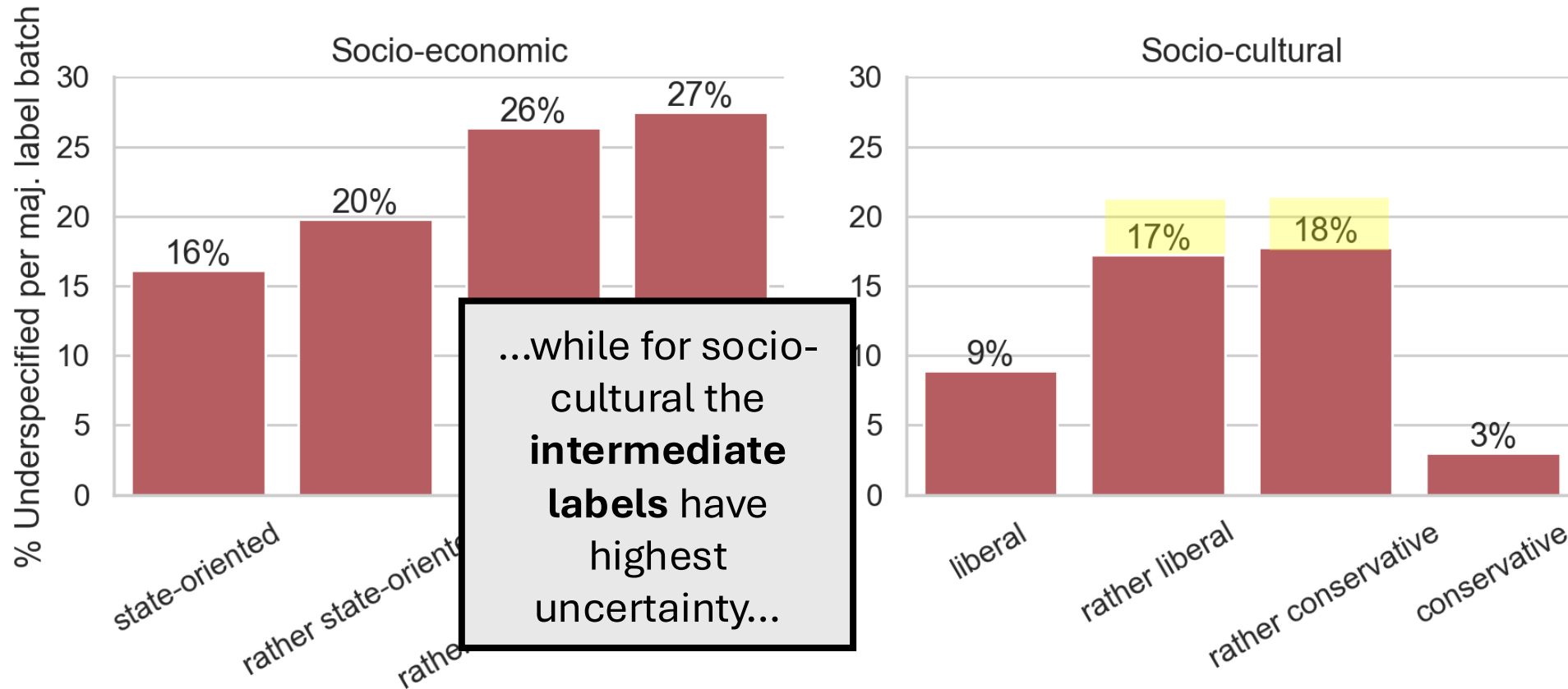


Market-oriented
positions are more
frequently judged as
underspecified...

Uncertainty

Are some positions less clearly identifiable?

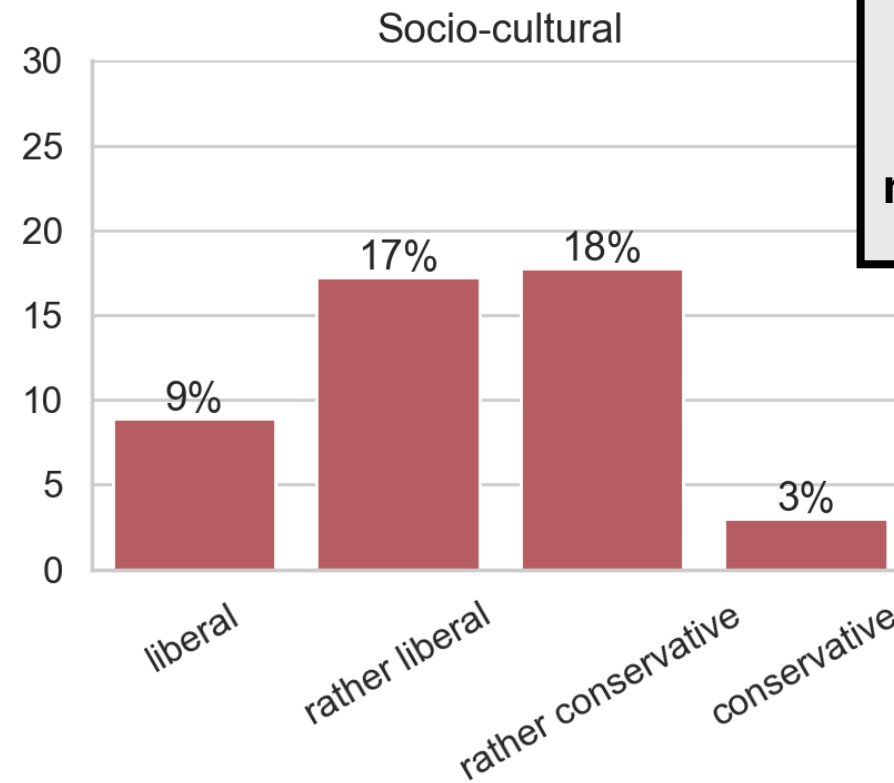
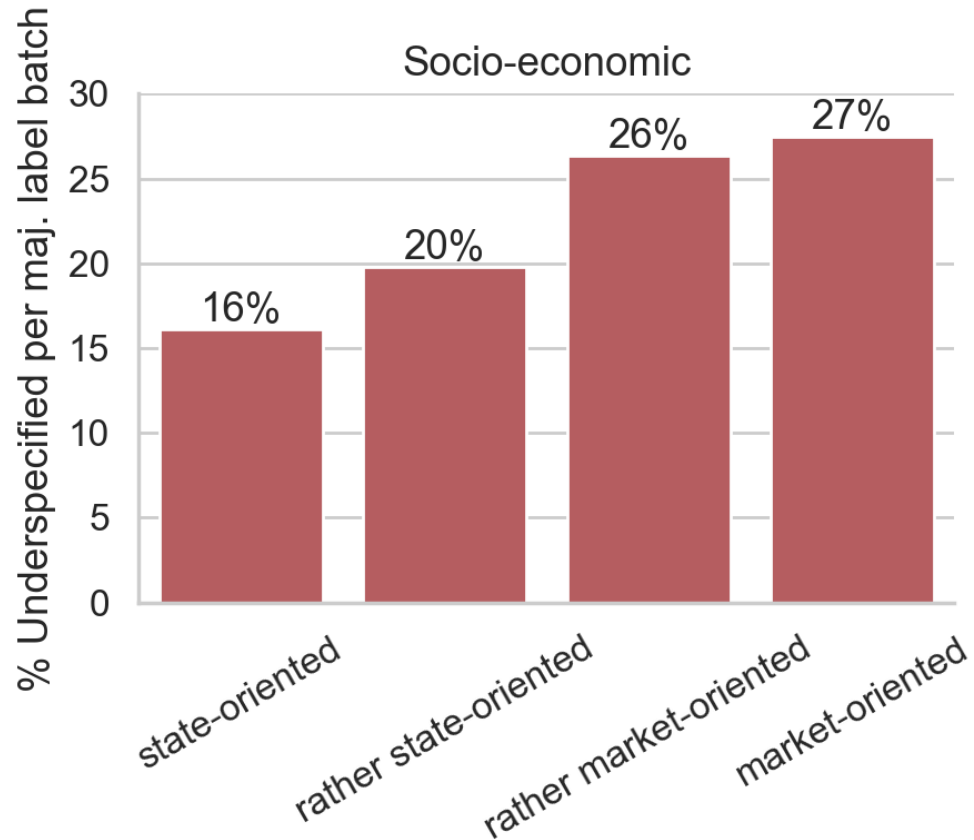
Underspecified rate by majority label



Uncertainty

Are some positions less clearly identifiable?

Underspecified rate by majority label



...but also
**generally lower
rate** per maj. label.

Corpus Examples: Underspecified or Market-Oriented?

*“Unlike asylum seekers, pensioners have spent their lives in Germany, worked here for decades and paid taxes. After a long working life, they have to find a way to make ends meet on a small pension, whilst asylum seekers are already receiving many times the amount of pocket money they would receive in other European countries. Whilst foreigners **transfer several billion every year to their home countries**, many pensioners cannot even afford to buy presents for their grandchildren on their birthdays or at Christmas.”*

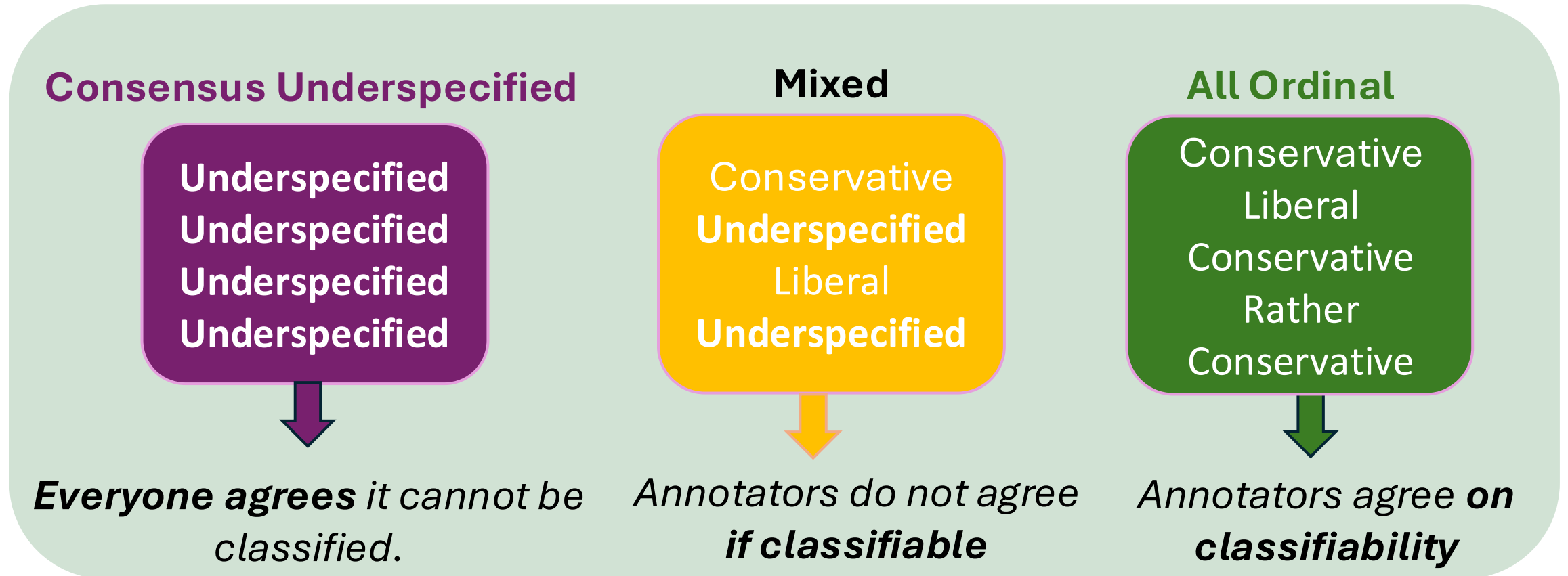
[Die Rentner haben im Gegensatz zu den Asylbewerbern ihr Leben in Deutschland verbracht, hierzulande Jahrzehnte gearbeitet und Steuern gezahlt. Nach einem langen Erwerbsleben müssen sie schauen, wie sie mit einer mickrigen Rente über die Runden kommen, während Asylbewerber bereits jetzt ein Vielfaches des Taschengelds bekommen, was sie in anderen europäischen Ländern bekommen würden. Während Ausländer **jedes Jahr etliche Milliarden in ihre Heimatländer transferieren**, können sich viele Rentner nicht einmal die Geschenke für ihre Enkel zum Geburtstag oder zu Weihnachten leisten.]

Annotation decisions social-economic

- 3 market-oriented
- 1 rather-market oriented
- 2 underspecified

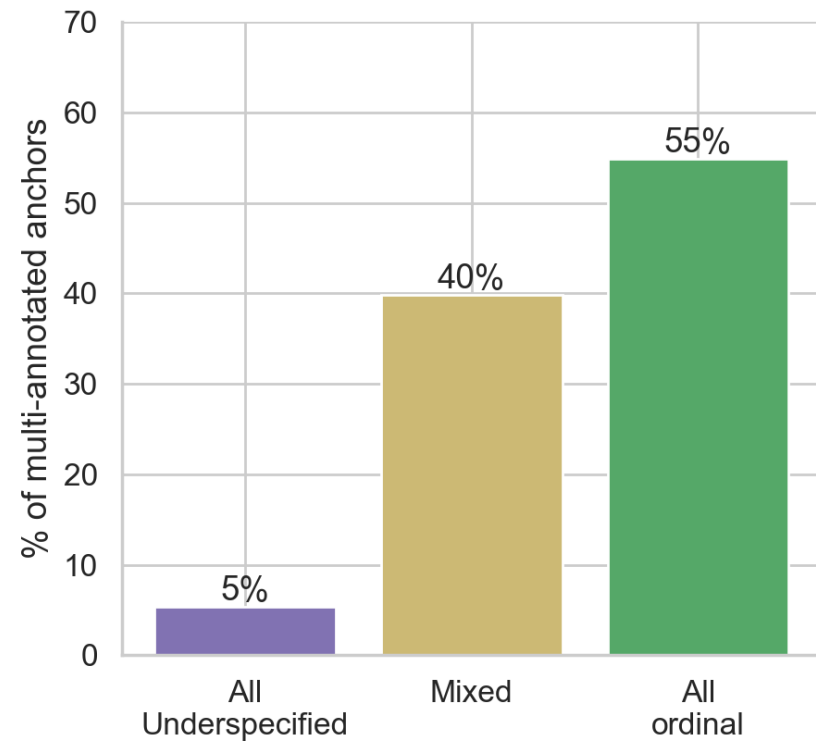
Uncertainty & Item Ambiguity

Are some statements not classifiable and others ambiguous?

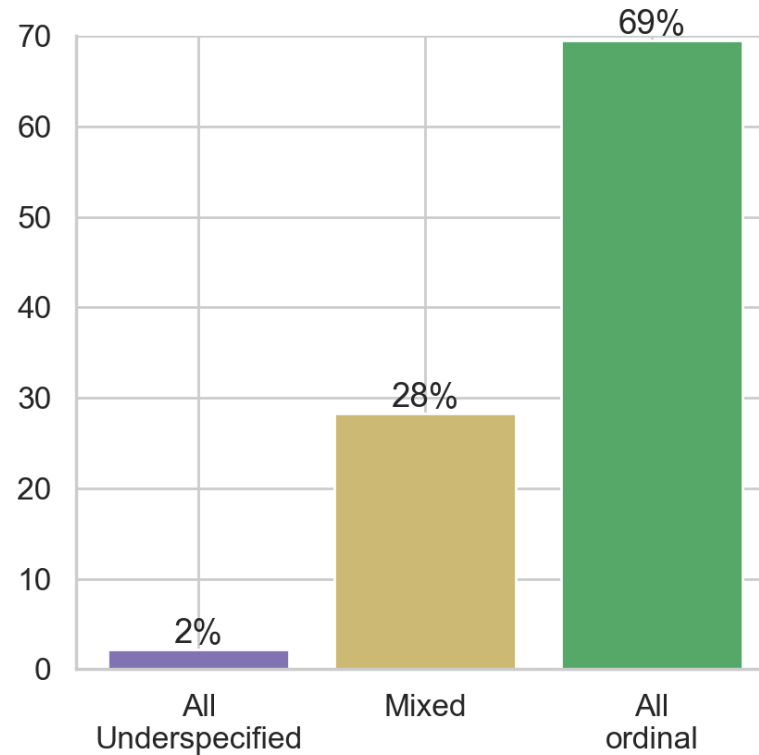


Item Ambiguity

Socio-economic

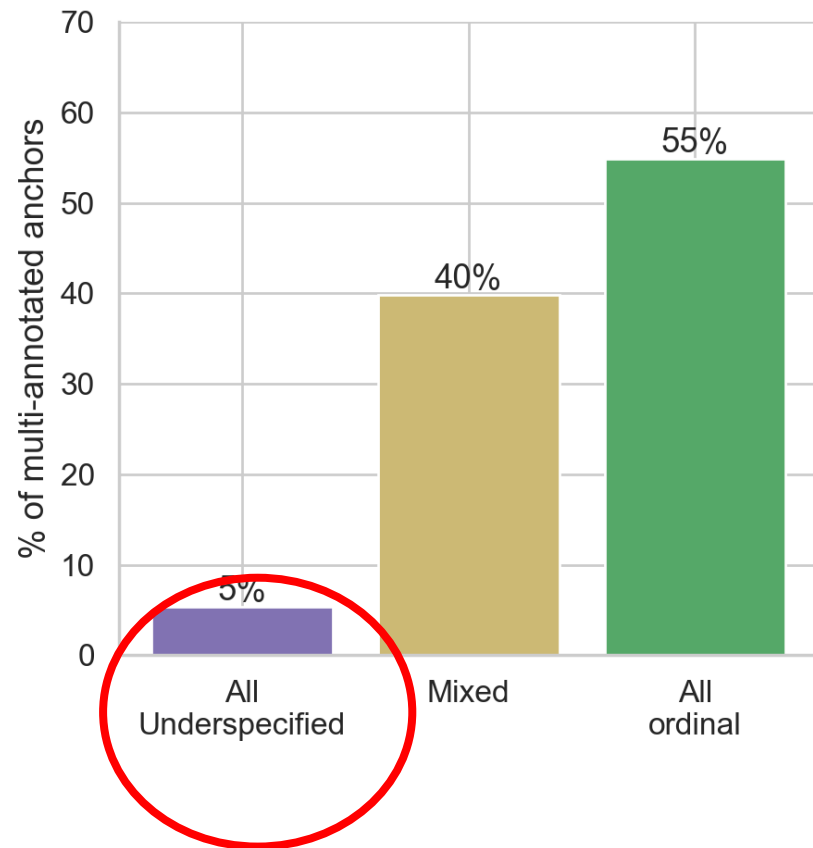


Socio-cultural

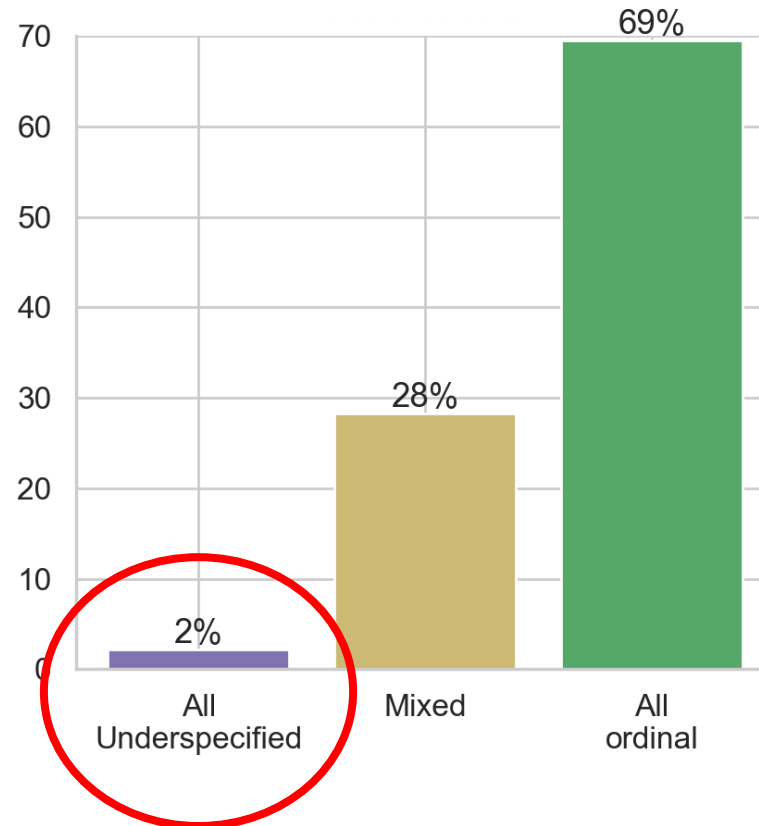


Item Ambiguity

Socio-economic

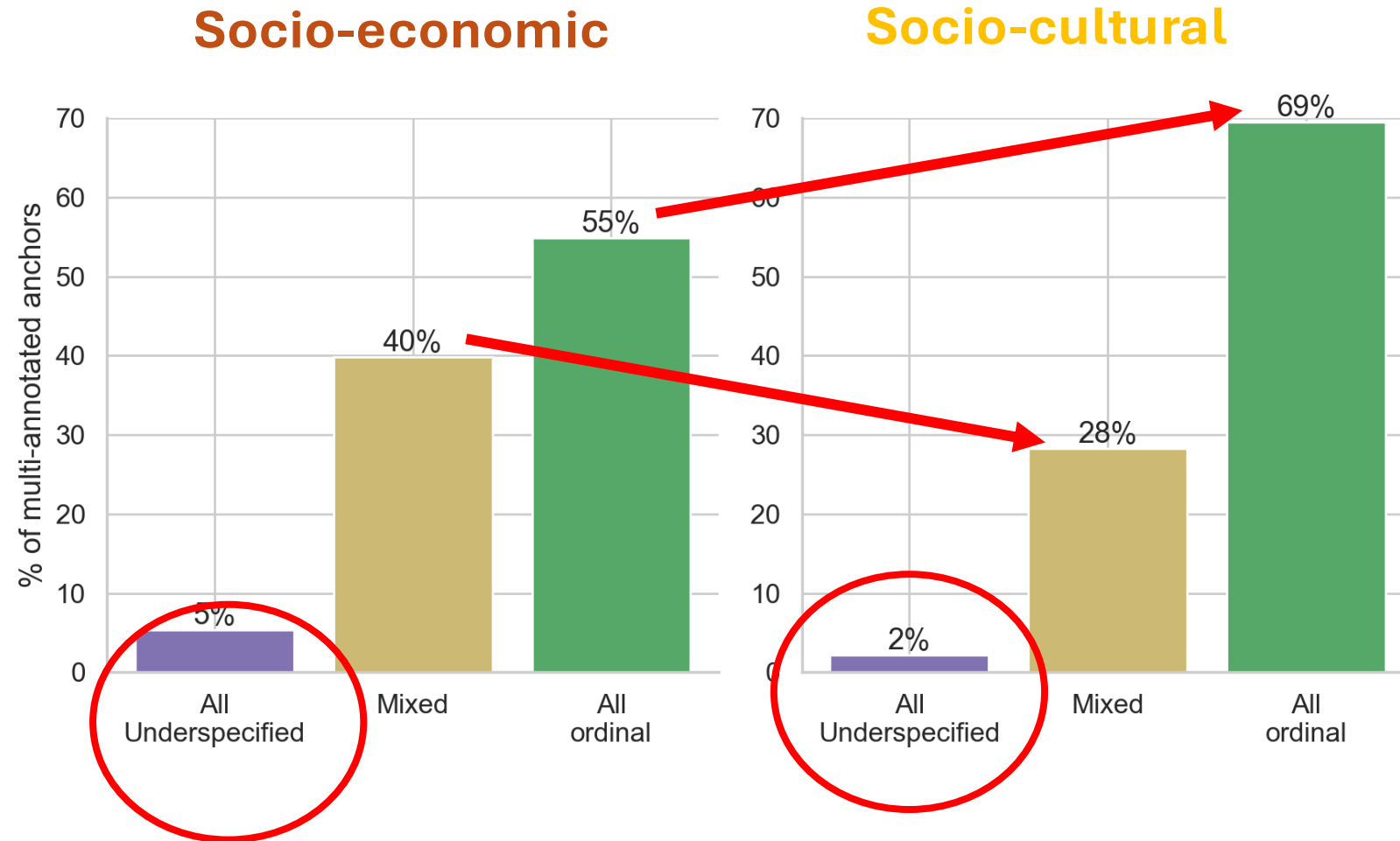


Socio-cultural



- **Only 5% / 2% of items are unanimously considered underspecified.**

Item Ambiguity



- **Only 5% / 2% of items are unanimously considered underspecified.**
- Socio-cultural **less Mixed** and **more All Ordinal** compared to socio-economic

Results **Inter Annotator Agreement**

- Overall agreement is **rather low** (Krippendorff's α).
- Most disagreement is **not random**:
 - disagreements are concentrated around neighbouring ordinal labels
- Unanimous Underspecified judgments are relatively uncommon
- Annotators often agree that a label can be assigned, but disagree about which label best fits.
 - Human label variation therefore mostly reflects **interpretative differences**, not inherently unclassifiable moral expressions.

Intra-Annotator Dataset

Inter-annotator disagreement tells us how annotators differ from one another.

- How consistent are **individual annotators when they revisit an item?**

Intra-Annotator Consistency (IAC)

How consistently do the same annotators label the same instance?

Alpha higher than for IAA (0.6-0.65)

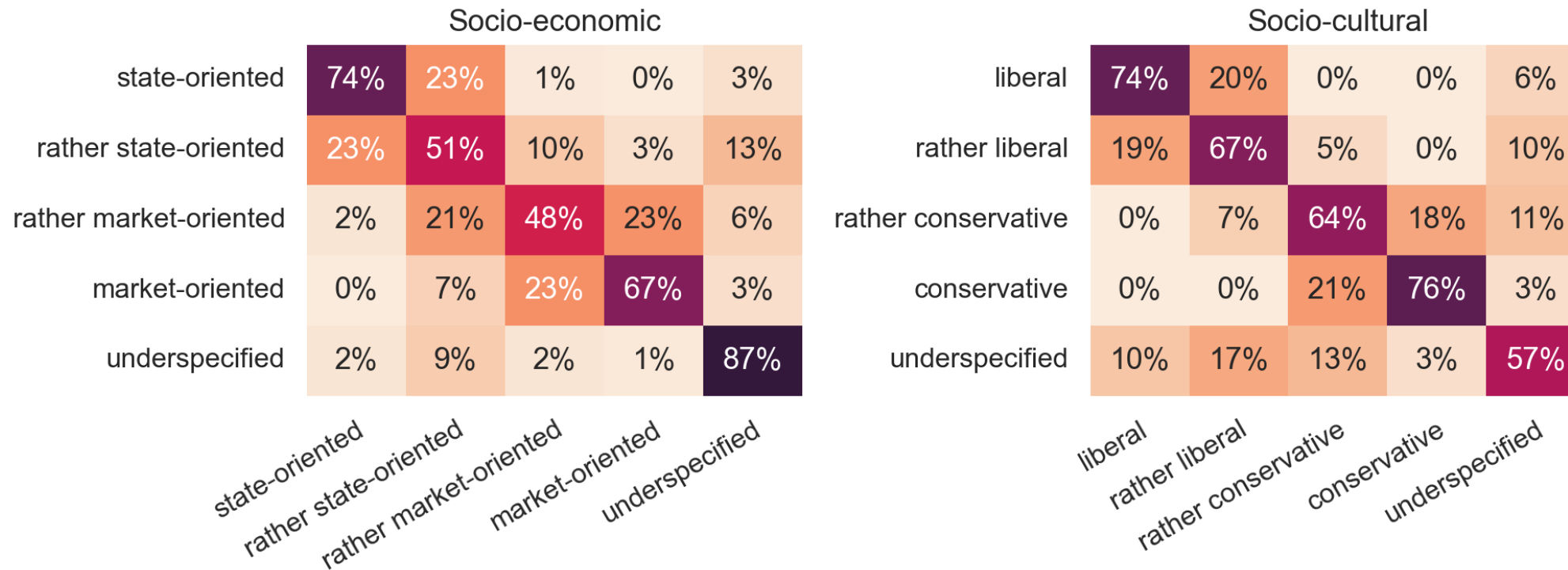
Scale	Krippendorff's alpha (ordinal)
Socio-economic	0.82
Socio-cultural	0.90

N annotation instances	N total repeat annotations	Underspecified rate
177	423	31.0%
177	423	16.3%

Dataset: repeated same-annotator instances (**Underspecified excluded** for ordinal α)

IAC Label Confusions

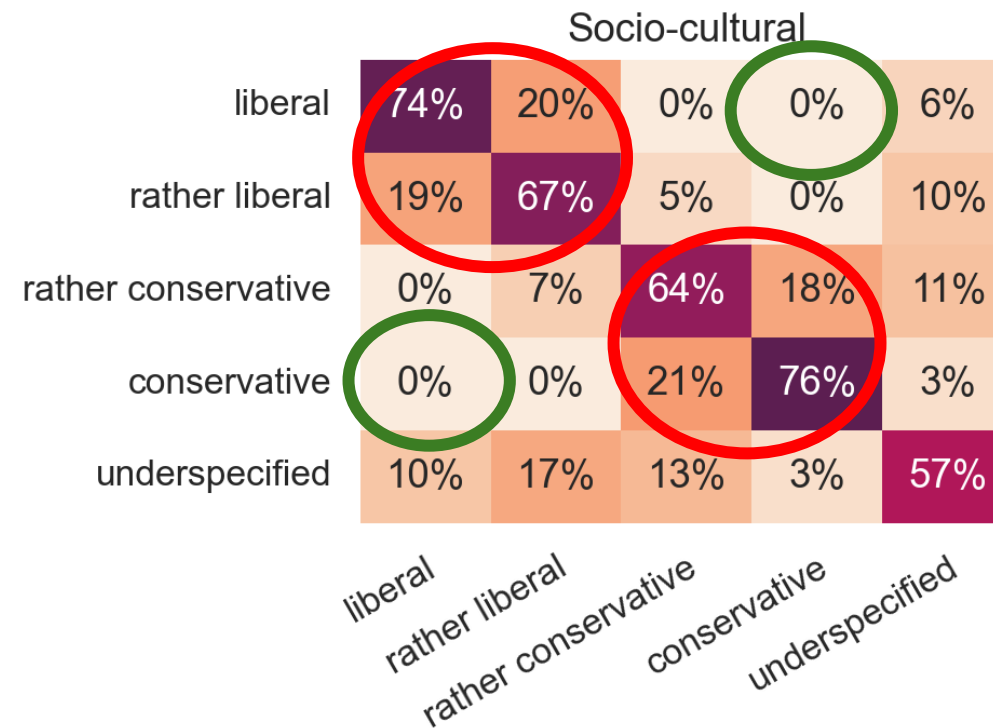
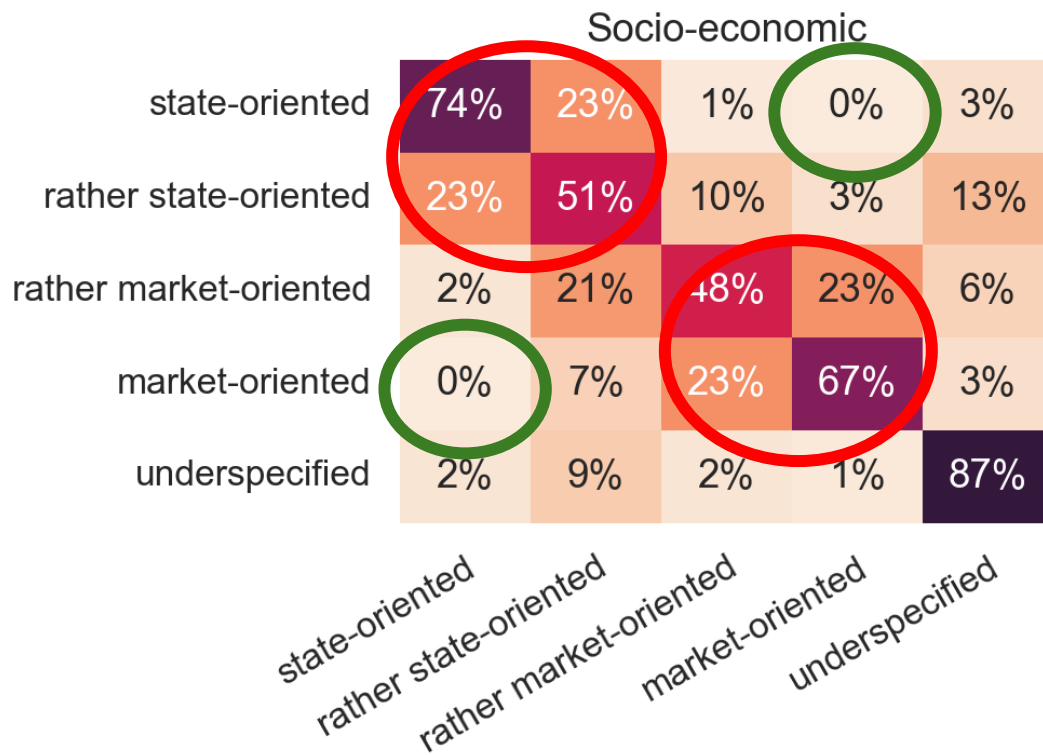
What kind of decision changes do annotators make?



Comparing all annotation decisions made by the same annotator on the same instance.

IAC Label Confusions

What kind of decision changes do annotators make?



Disagreements
mostly occur
between
neighbouring
categories

Political
extreme
positions not
confused

Comparing all annotation decisions made by the same annotator on the same instance.

Results Intra-Annotator Consistency

- Most repeated annotations remain unchanged ($\alpha > 0.8$).

Similar observation to IAA:

- Revisions mostly involve **neighbouring ordinal labels**
- Extreme ideological reversals do not occur
- Individual annotations are **largely stable**, suggesting that inter-annotator disagreement primarily reflects **different interpretations** rather than inconsistent annotation behaviour.

Part V - Takeaway & Limitations

Conclusions

- Political position annotation is not unreliable because annotators behave inconsistently
- Rather, different annotators **systematically interpret the same discourse differently**
- *socio-cultural* seems to be slightly easier than *socio-economic*

Future Work:

- Correlation analysis with parties, Moral Frame values, topics
- Analysis of Coherence annotations of political positions

Limitations

- **Intra-Annotator Dataset: Uncontrolled temporal interval** between repeated annotations
 - memory effects cannot be ruled out
- **Unequal distribution of repetitions** (up to 5 for same-annotator, 7 for iaa)
 - not every item has the same number of annotations
 - not every annotator repeated the same number of items

Thank you for listening!

Email: zaczynska@uni-potsdam.de

GePaDeU Corpus in ANNIS:

https://linguistischeempirieannis.uni-muenster.de/annis/#_q=&ql=aql&_c=&cl=5&cr=5&s=0&l=10



GePaDeU Paper:

<https://aclanthology.org/2026.lrec-1.141/>

Literature

- Rebecca J. Passonneau and Bob Carpenter. 2013. [The Benefits of a Model of Annotation](#). In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, pages 187–195, Sofia, Bulgaria. Association for Computational Linguistics.
- Ines Rehbein, Julian Schlenker, Lars Ostertag, and Simone Paolo Ponzetto. 2026. [GePaDeU - a Multi-layer Corpus of German Parliamentary Debates with Rich Semantic and Pragmatic Annotations](#). In *Proceedings of the Fifteenth Language Resources and Evaluation Conference*, pages 1794–1810, Palma de Mallorca, Spain. ELRA Language Resource Association.
- Ines Rehbein, Maris Leander Buttmann, Julian Schlenker, and Simone Paolo Ponzetto. 2026. [Appeal, Align, Divide? Stance Detection for Group-Directed Messages in German Parliamentary Debates](#). In *Proceedings of the Fifteenth Language Resources and Evaluation Conference*, pages 299–318, Palma de Mallorca, Spain. ELRA Language Resource Association.
- Ines Rehbein, Ines Reinig, and Simone Paolo Ponzetto. 2025. [Moral Framing in Politics \(MFIP\): A new resource and models for moral framing](#). In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 34643–34663, Suzhou, China. Association for Computational Linguistics.
- Julian Schlenker, Ines Rehbein, Lilly Brauner, Florian Ertz, Ines Reinig, and Simone Paolo Ponzetto. 2026. [GePaDeSE: A New Resource for Clause-Level Aspect in German Parliamentary Debates](#). In *Proceedings of the Fifteenth Language Resources and Evaluation Conference*, pages 3617–3632, Palma de Mallorca, Spain. ELRA Language Resource Association.

Image

- Bergmann-Rusche-Skala taken from:
Bergmann, Knut / Rusche, Christian, 2022,
Verhandlungspositionen im Bundesrat. Eine politökonomische
Analyse, https://regierungsforschung.de/wp-content/uploads/2022/03/22032022_regierungsforschung.de_Bergmann-Rusche_Bundesrat.pdf, Duisburg

Appendix

Annotation Task 1: Ideological Positioning

Instruction. Below, in the gray box, you see a short excerpt from a speech in the German Bundestag. The bold text presents a statement from the speech, shown in the context in which it was made. Please assess the speaker's ideological position on the *socio-economic scale* and the *socio-cultural scale*. Select the corresponding value on each Likert scale.

Reference

“One last thing. You’re all **supporting the students in their Fridays for Future movement**. If you’re serious about that, then cut military spending and invest in climate protection and the welfare state instead of continuing to increase military spending.”

(Original German: “Ja. Einen letzten Satz. Sie alle **unterstützen die Schüler und Schülerinnen im Hinblick auf Fridays for Future**. Wenn Sie das ernst meinen, dann rüsten Sie ab, und investieren Sie in Klimaschutz und in den Sozialstaat, statt weiter aufzurüsten.”)

Q1. Socio-economic scale: How would you place the speaker’s socio-economic position along the poles *social justice* ↔ *neoliberal market freedom*?

Social

Rather social

Rather neoliberal

Neoliberal

Underspecified

Q2. Socio-cultural scale: How would you place the speaker’s socio-cultural position along the poles *liberal* ↔ *conservative*?

Liberal

Rather liberal

Rather conservative

Conservative

Underspecified

Figure 1: Annotation Task 1: Rating task where annotators position the reference on a socio-economic and socio-cultural scale.

Annotation Task 2: Ideological Coherence

Instruction. Read the four comparison statements below and assess whether they express an ideological position similar to the statement from Task 1 shown on the left. Rate each candidate separately.

Reference
“One last thing. You’re all **supporting the students in their Fridays for Future movement**. If you’re serious about that, then cut military spending and invest in climate protection and the welfare state instead of continuing to increase military spending.”

Q3. How similar are the ideological positions expressed in the reference and candidate frame?

Candidate 1
“You can certainly claim that, but at least it is not described in your motion. Sustainability therefore always rests on three pillars: social, environmental, and economic considerations. This motion completely lacks that balance.”
(Original German: “Das können Sie ja reklamieren, aber das ist zumindest nicht in Ihrem Antrag beschrieben. Deshalb ist Nachhaltigkeit immer eine Säule von sozialen, ökologischen und wirtschaftlichen Aspekten. Dies fehlt in diesem Antrag völlig.”)

Similar

Rather similar

Rather dissimilar

Dissimilar

Underspecified

Candidate 2
“But none of this has to happen; with the right measures, some of which I would like to mention, we can tackle climate change in a sensible way. At the international level, above all, we must advance emissions trading.”
(Original German: “Doch das alles muss nicht sein; denn mit den richtigen Maßnahmen, von denen ich Ihnen gerne ein paar nenne, können wir den Klimawandel vernünftig bekämpfen. Auf internationaler Ebene müssen wir dazu vor allem den Emissionshandel voranbringen.”)

Similar

Rather similar

Rather dissimilar

Dissimilar

Underspecified

Candidate 3
“It must now act quickly and provide the regions with prospects for the future. The experts must then carefully assess by how many gigawatts coal-fired power generation can be reduced without jeopardizing security of supply; after all, none of us wants that.”
(Original German: “Sie muss jetzt schnell arbeiten und den Regionen Perspektiven aufzeigen. Die Fachleute müssen dann eingehend prüfen, um wie viel Gigawatt die Kohleverstromung reduziert werden kann, ohne die Versorgungssicherheit zu gefährden; denn das wollen wir alle nicht.”)

Similar

Rather similar

Rather dissimilar

Dissimilar

Underspecified

Candidate 4
“Take a look at the forecasts: The UN High Commissioner for Refugees predicts that climate change alone will result in more than 6 million additional displaced people every year.”
(Original German: “Werfen Sie einen Blick auf die Prognosen: Der UN-Flüchtlingskommissar prognostiziert, dass wir allein durch den Klimawandel mehr als 6 Millionen weitere Vertriebene Jahr für Jahr erleben werden.”)

Similar

Rather similar

Rather dissimilar

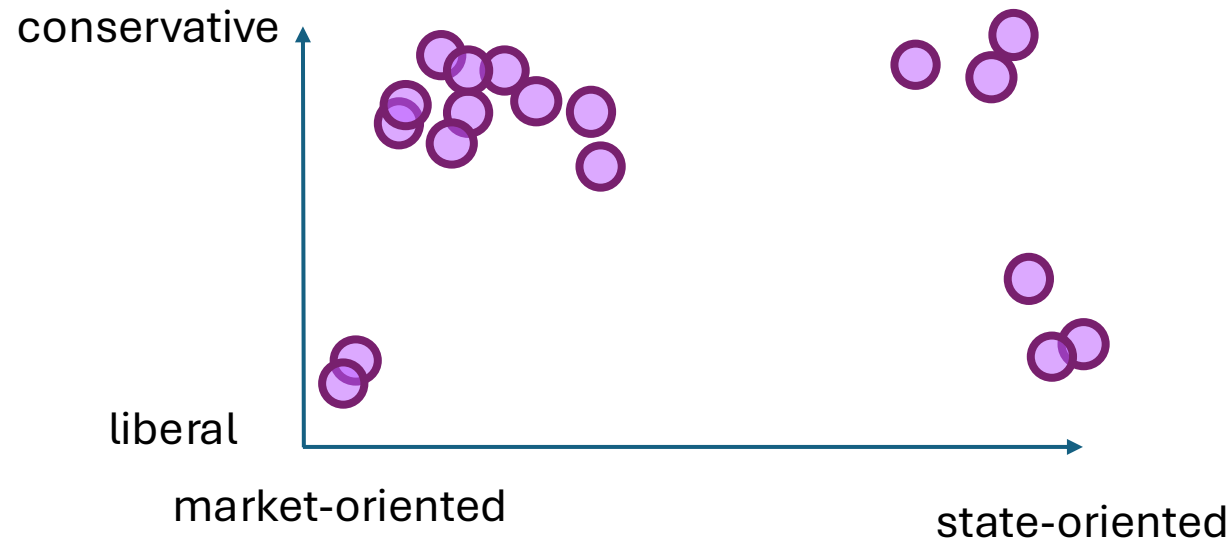
Dissimilar

Underspecified

Figure 2: Annotation Task 2: Comparison task where annotators judge the ideological coherence between the reference and four candidates.

Ideological Coherence Ideas

- Goal: create a distribution of political scales
- Based on coherence of moral expressions create gold standard for classification



IAC Label Confusions

Repeat-pass label co-occurrence per annotator — Socio-economic
(trend only; cannot rule out memorisation effects)

A1

state-oriented	100%	0%	0%	0%	0%
rather state-oriented	0%	0%	0%	0%	0%
rather market-oriented	0%	0%	50%	0%	50%
market-oriented	0%	0%	0%	0%	0%
underspecified	0%	0%	17%	0%	83%

state-oriented
rather state-oriented
rather market-oriented
market-oriented
underspecified

A2

state-oriented	0%	100%	0%	0%	0%
rather state-oriented	33%	67%	0%	0%	0%
rather market-oriented	0%	0%	100%	0%	0%
market-oriented	0%	0%	0%	100%	0%
underspecified	0%	0%	0%	0%	100%

state-oriented
rather state-oriented
rather market-oriented
market-oriented
underspecified

A3

state-oriented	92%	8%	0%	0%	0%
rather state-oriented	25%	50%	25%	0%	0%
rather market-oriented	0%	25%	0%	75%	0%
market-oriented	0%	0%	13%	87%	0%
underspecified	0%	0%	0%	0%	0%

state-oriented
rather state-oriented
rather market-oriented
market-oriented
underspecified

A4

state-oriented	86%	0%	0%	0%	14%
rather state-oriented	0%	0%	0%	0%	100%
rather market-oriented	0%	0%	0%	100%	0%
market-oriented	0%	0%	33%	67%	0%
underspecified	2%	2%	0%	0%	97%

state-oriented
rather state-oriented
rather market-oriented
market-oriented
underspecified

A5

state-oriented	79%	17%	0%	0%	4%
rather state-oriented	40%	30%	15%	5%	10%
rather market-oriented	0%	75%	0%	25%	0%
market-oriented	0%	9%	9%	73%	9%
underspecified	10%	10%	0%	5%	76%

state-oriented
rather state-oriented
rather market-oriented
market-oriented
underspecified

A6

state-oriented	50%	50%	0%	0%	0%
rather state-oriented	22%	67%	0%	0%	11%
rather market-oriented	0%	0%	100%	0%	0%
market-oriented	0%	0%	0%	0%	0%
underspecified	0%	100%	0%	0%	0%

state-oriented
rather state-oriented
rather market-oriented
market-oriented
underspecified

A7

state-oriented	60%	30%	5%	0%	5%
rather state-oriented	21%	29%	18%	0%	32%
rather market-oriented	4%	19%	62%	12%	4%
market-oriented	0%	0%	60%	40%	0%
underspecified	2%	16%	2%	0%	81%

state-oriented
rather state-oriented
rather market-oriented
market-oriented
underspecified

A8

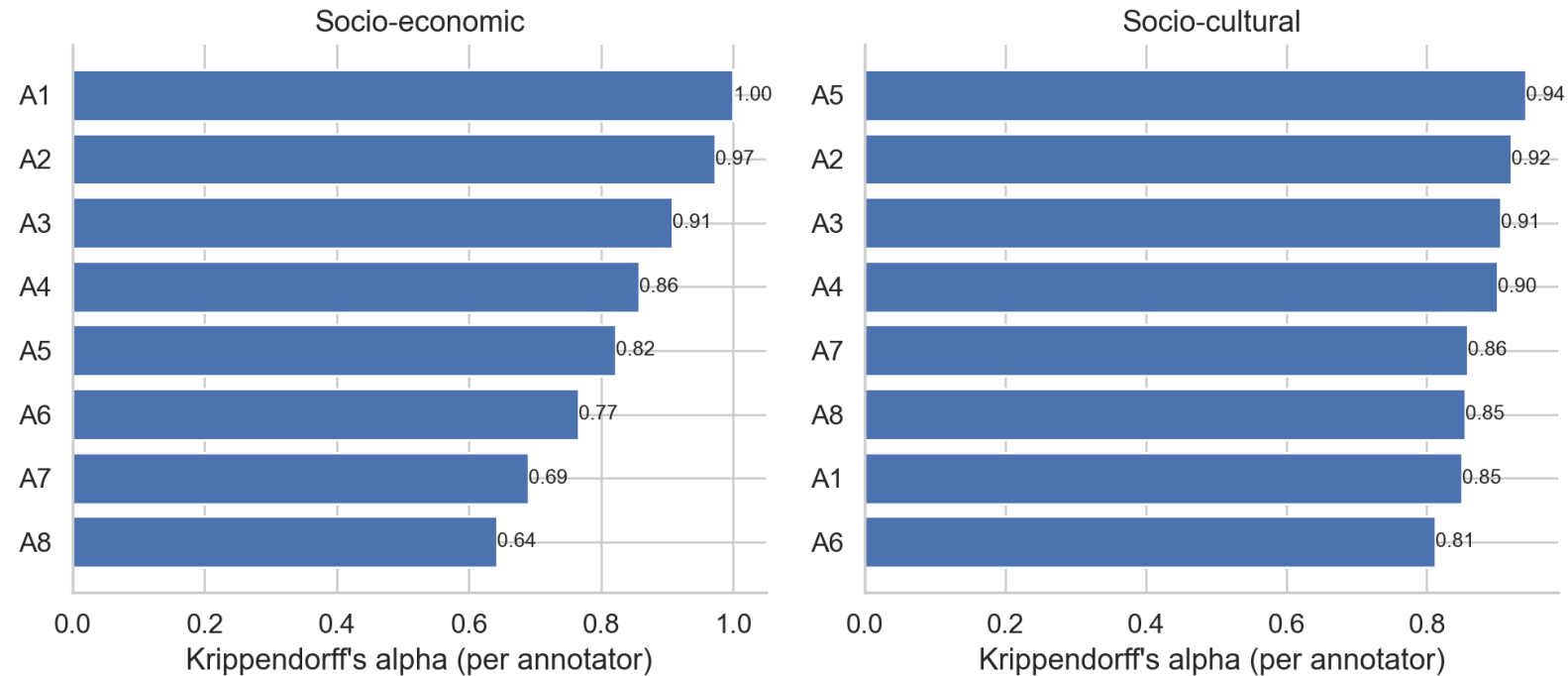
state-oriented	48%	52%	0%	0%	0%
rather state-oriented	18%	65%	6%	4%	7%
rather market-oriented	0%	27%	27%	40%	7%
market-oriented	0%	21%	43%	29%	7%
underspecified	0%	11%	2%	2%	85%

state-oriented
rather state-oriented
rather market-oriented
market-oriented
underspecified

Research Question 4

Intra-Annotator Consistency (IAC)

Per annotator consistency



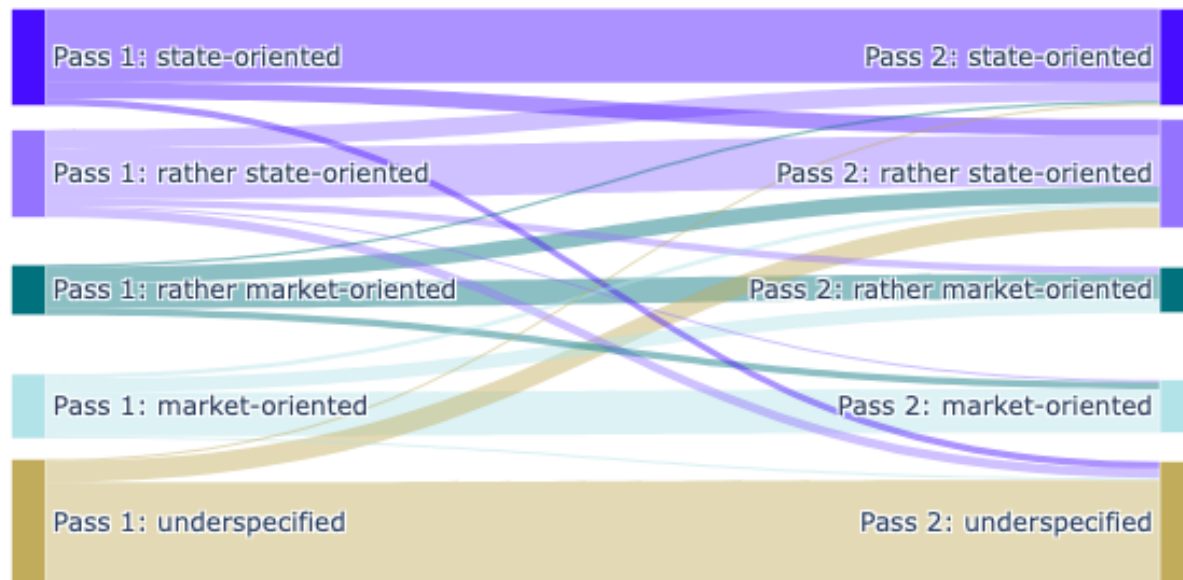
IAC Label Transitions

What kind of revisions do annotators make?

Socio-economic

1st Pass

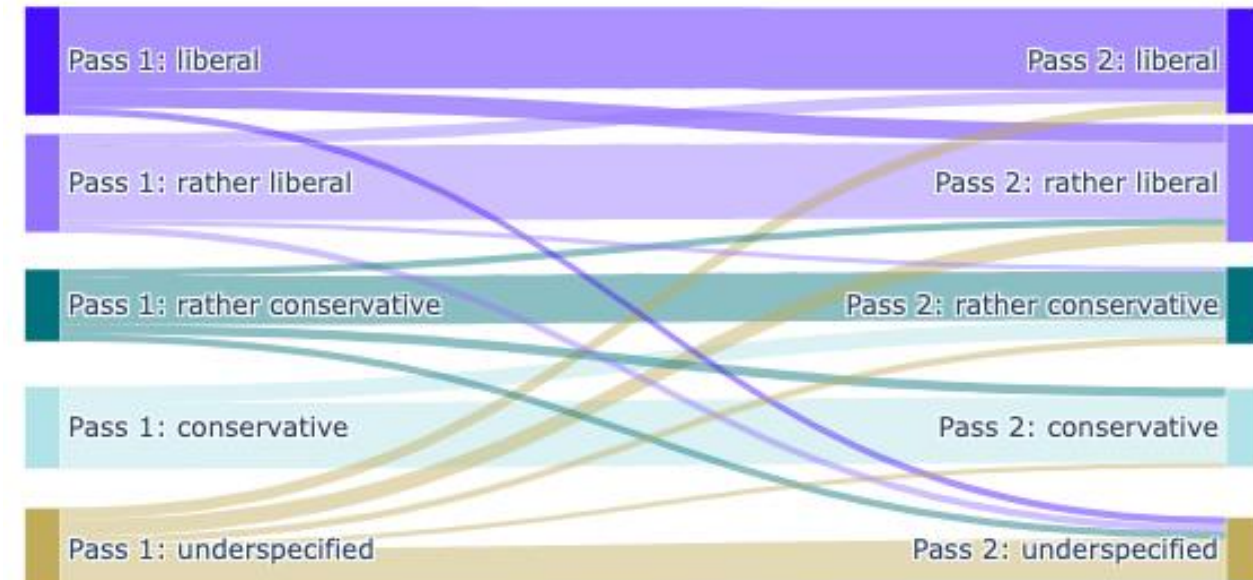
2nd Pass



Socio-cultural

1st Pass

2nd Pass



Corpus

METRIC	VALUE
Total anchor instances	279
- Climate	59
- Pension	134
- Immigration	87
Total political positioning annotation decisions	970
Number of annotators	8 (from political/social science)
Qualification	Test annotation before participation

More Statistics on Dataset

- Speeches from 15 debates (2017-2021)

Parties distribution:

- 'AFD': 67
- 'FDP': 55
- 'LINKE': 54
- 'SPD': 46
- 'CDU': 45
- 'GRUENE': 18