Measuring Income (In)equality: Comparing Survey Questions with Unipolar and Bipolar Scales in a Probability-based Online Panel

Jan Karem Höhne[1,2], Dagmar Krebs[3], Steffen-M. Kühnel[4]

[1]University of Mannheim, Germany

[2]RECSM-Universitat Pompeu Fabra in Barcelona, Spain

[3]University of Gießen, Germany

[4]University of Göttingen, Germany

GOR Conference

Berlin (Germany) – September 10 to 11, 2020

Introduction and Background I

- Rating scale design can affect response behavior.
  - *Scale alignment, direction, labeling, length, polarity etc.*
- Scale polarity has rarely been addressed in survey research.
  - *Unipolar: Agree strongly, (…), agree not at all.*
  - *Bipolar: Agree strongly, (…), disagree strongly.*
- It is controversial whether to use unipolar or bipolar scales.
- Most studies remain on the observational level.
  - *Comparing response distributions.*
  - *Middle attraction in unipolar scales.*
  - *Positivity bias in bipolar scales.*

Introduction and Background II

- Only few studies look at the latent level.
- Multiple-indicator factor level (measurement invariance).
  - *Metric: Equivalence of the latent structure.*
  - *Scalar: Equivalence of the intercepts.*
- Single-question level (focusing on response categories).
  - *Item Response Theory (IRT).*
  - *Modeling latent thresholds.*
  - *Equidistance of response categories.*
- Shedding light on the measurement properties of unipolar and bipolar scales.

Research Hypotheses

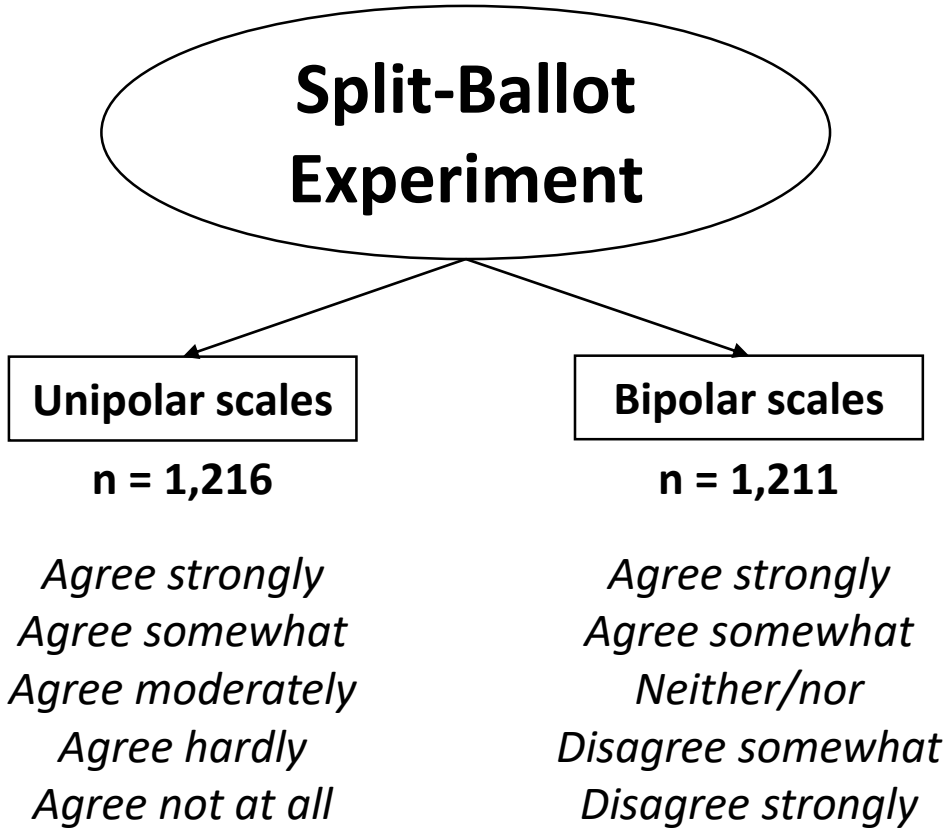
We expect higher proportions of middle responses in unipolar than in bipolar scales (H1a).

We expect higher proportions of positive agreeing responses in bipolar than in unipolar scales (H1b).

We expect to obtain measurement non-invariance between unipolar and bipolar scales on the latent level (H2).

We expect the latent thresholds of unipolar scales to be more equidistantly distributed than those of bipolar scales (H3).

Methods: Research Design



- The two groups received 4 questions on income (in)equality.
  - *The questions were adopted from the ESS.*
- Scale design (unipolar and bipolar):
  - *5-points.*
  - *Fully-labeled.*
  - *Vertical alignment.*
- Each question was presented individually.
- Question order was randomized.

Methods: Participants

The experiment was conducted in the probability-based German Internet Panel (GIP) in March 2019.

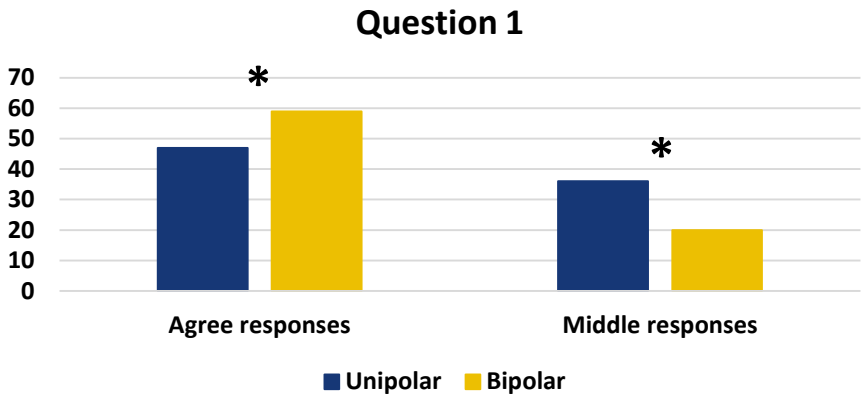
Final sample size:	*N = 2,427*
Gender:	*49% female*
Age (in years):	*Mean = 49*
Education:	*16% lower secondary school*
	32% intermediate secondary school
	52% at least college preparatory secondary school

Note. Chi-square tests revealed no significant differences between the two groups with respect to gender, age, and education.

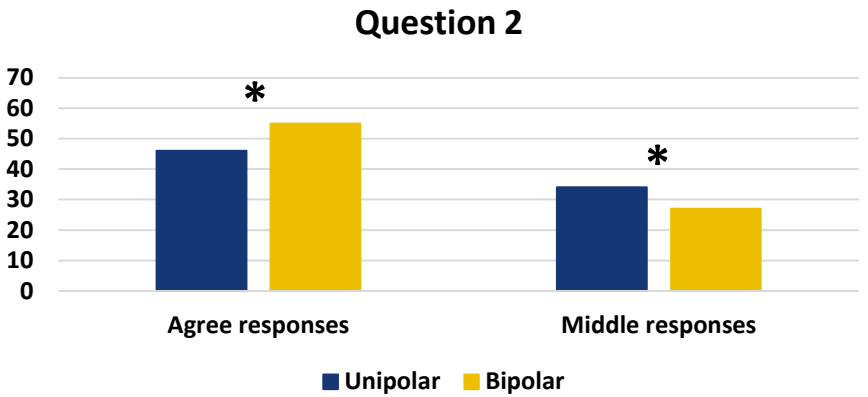
Analytical Strategies

- Comparing response distributions (directed Z-tests).
  - *Middle attraction: Comparing middle responses.*
  - *Positivity bias: Comparing positive agree responses.*
- Testing for measurement invariance.
  - *Multi-Group Confirmatory Factor Analysis (MG-CFA).*
  - *Notion of strong measurement invariance.*
- Modeling latent thresholds of response categories.
  - *Item Response Theory (IRT).*
  - *Unrestricted univariate probit models for each question.*
  - *Computing linear regressions of the estimated unrestricted thresholds.*

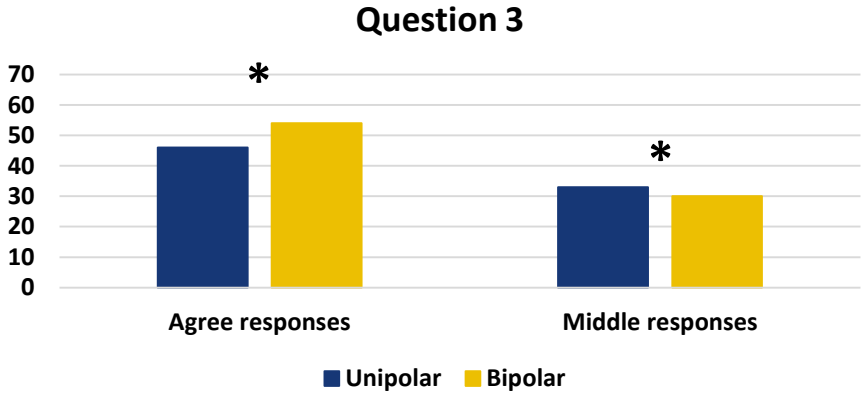
Response Distributions



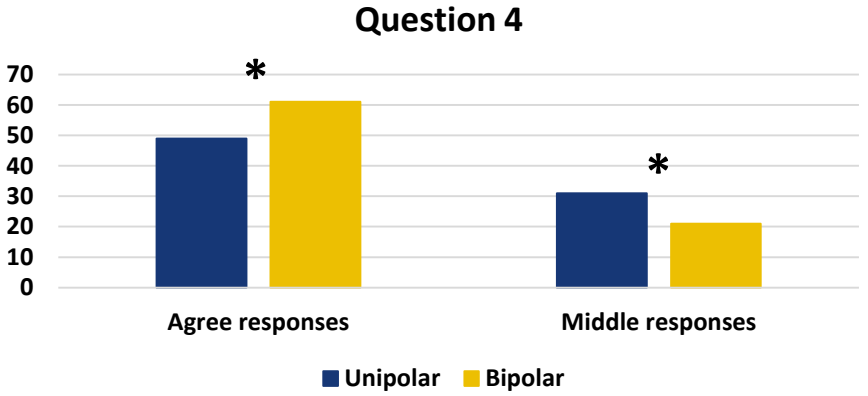
Note. *p < 0.05.



Note. *p < 0.05.



Note. *p < 0.05.



Note. *p < 0.05.

Measurement Invariance

Invariance level	Chi-square value	Df	Chi-square Difference	CFI	RMSEA
Configural	0.45 (1.32)	2		1	0.000
Metric	2.20 (1.19)	5	1.83	1	0.000
Scalar	30.58 (1.18)	8	31.59*	0.981	0.048

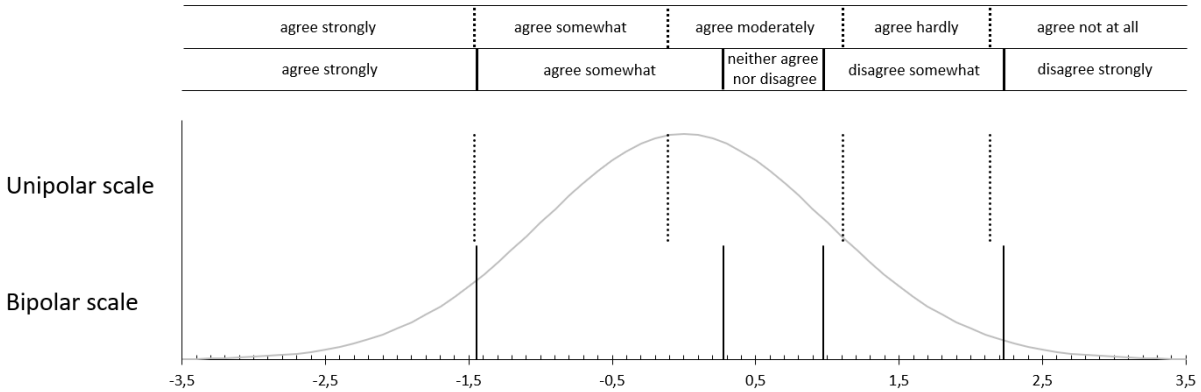
Note. *p < 0.05. The results are based on the MLR discrepancy function. Scale correction factors are in parentheses.

Latent Thresholds I

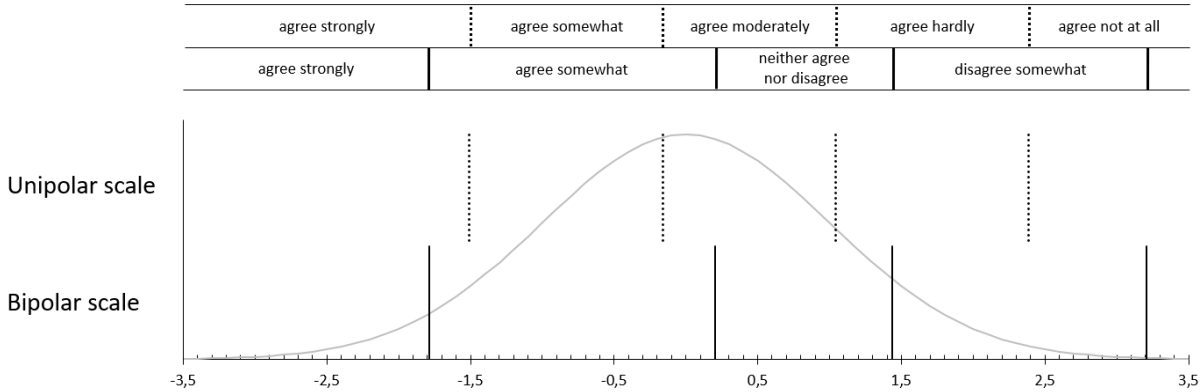
Questions	Unipolar R^2	Corrected R^2	Bipolar R^2	Corrected R^2
1	0.996	0.995	0.975	0.962
2	1	0.999	0.993	0.989
3	0.997	0.995	0.994	0.990
4	0.998	0.997	0.988	0.982

Note. R2 and corrected R2 values of linear regressions of estimated unrestricted latent thresholds (Y) on ascending integers (X = 1 to 4) for unipolar and bipolar scales.

Latent Thresholds II

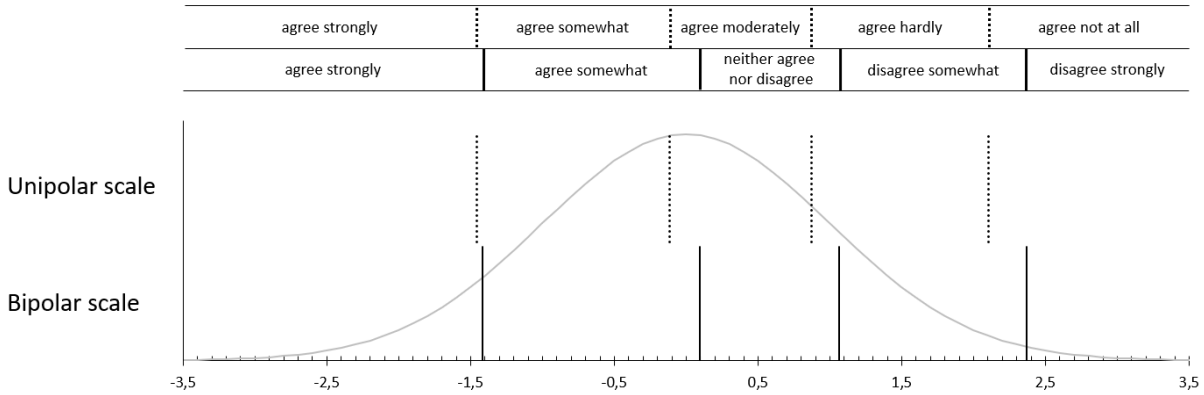


Question 1: Thresholds and latent distributions of unipolar and bipolar scales.

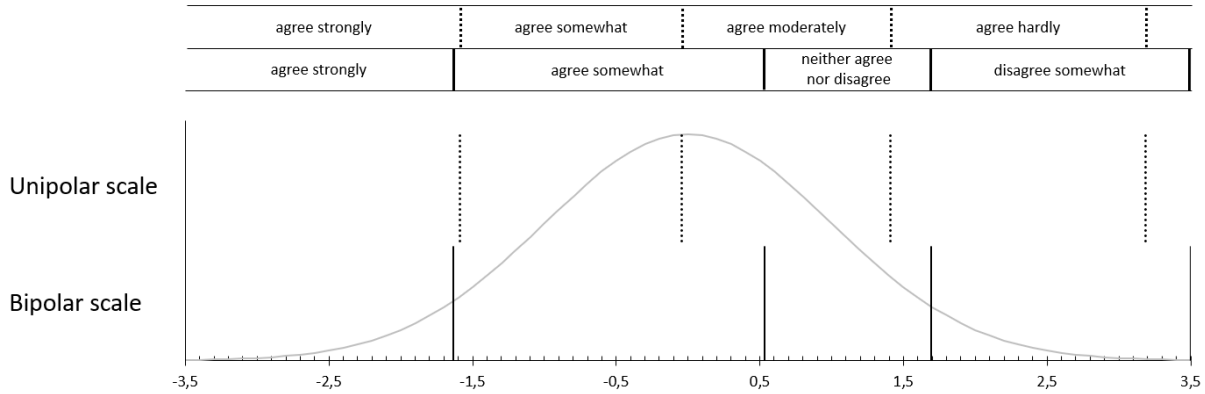


Question 2: Thresholds and latent distributions of unipolar and bipolar scales.

Latent Thresholds III



Question 3: Thresholds and latent distributions of unipolar and bipolar scales.



Question 4: Thresholds and latent distributions of unipolar and bipolar scales.

Discussion and Conclusion

- Successful replication of previous findings.
  - *Middle attraction in unipolar scales.*
  - *Positivity bias in bipolar scales.*
- New evidence on measurement non-invariance.
  - *Intercepts of unipolar and bipolar scales differ (only metric invariance).*
  - *Points to the presence of systematic measurement error.*
- New evidence on equidistance of response categories.
  - *Unipolar and bipolar scales differ in measurement properties.*
  - *Latent thresholds of unipolar scales are more equidistant.*

Conclusion: Unipolar and bipolar scales are not interchangeable.

This talk is based on:

Höhne, J. K., Krebs, D., & Kühnel, S.-M. (2020). Measuring Income (In)equality: Comparing Survey Questions With Unipolar and Bipolar Scales in a Probability-Based Online Panel. Social Science Computer Review. DOI: 10.1177/0894439320902461

Many thanks for your attention!

Contact: hoehne@uni-mannheim.de

upf. RECSM

UNIVERSITY OF MANNHEIM

Appendix I

- Employees need strong unions to protect their working conditions and wages (Question 1).
- Large income differences are acceptable to adequately acknowledge different talents and achievements (Question 2).
- To ensure a fair society differences in people's living standards should be small (Question 3).
- Social benefits lead to more equality in society (Question 4).

Appendix II

MEASUREMENT INVARIANCE (Mplus commands)

```
VARIABLE:
  NAMES ARE scale v1 v2 v3 v4;
  USEVARIABLES ARE scale v1 v2 v3 v4;
  GROUPING scale (1 = unipolar 2 = bipolar);
ANALYSIS:
  ESTIMATOR IS MLR;
MODEL:
  F1 BY v1 v2 v3 v4;
  v1 WITH v3;
  [F1@0];
```

Appendix III

LATENT THRESHOLDS

```
VARIABLE:
  NAMES ARE scale v1;
  CATEGORICAL IS v1;
  USEVARIABLES ARE v1;
  USEOBSERVATIONS ARE scale EQ 1;
ANALYSIS:
  ESTIMATOR IS WLS;
  PARAMETERIZATION IS THETA;
MODEL:
  v1@1;
  [v1$1] (t1);
  [v1$2] (t2);
  [v1$3] (t3);
  [v1$4] (t4);
   F1 BY v1@1; [F1@0]; F1@0;
MODEL CONSTRAINT:
  NEW (d*1.0);
  t2=t1+d;
  t3=t2+d;
  t4=t3+d;
```