gesis
Leibniz Institute for the Social Sciences

Identifying of and Dealing with Item Nonresponse in Open-ended Questions in a Cross-national Context

Katharina Meitinger
(GESIS - Leibniz Institute for the Social Sciences)
Lars Kaczmirek
(University of Vienna, Australian National University)

Member of the
Leibniz Association

CSDI, Limerick. 28 March, 2018

Overview

- Open-ended questions and the issue of nonresponse

- Detecting and reducing nonresponse in open-ended questions

- Results:

  ▶ Initial nonresponse

  ▶ Nonresponse after conversion

- Conclusion

Open-ended Questions

- Important source of information to:
  - Ask about exploratory topics
  - Receive answers when the list of possible response categories is not known or is too long (Fowler 1995)
  - Assess validity and comparability with probing (e.g., Schuman 1966; Behr et al. 2017)

- Open-ended questions are cognitively more demanding for respondents than closed items and increase response burden (Bradburn 1978)
→ Risk of increased item-nonresponse at open-ended questions

How to reduce nonresponse?

Prevent occurrence of nonresponse by optimizing visual design:

▶ Size of answer boxes (e.g., Christian and Dillman 2004; Smyth et al. 2009; Behr et al. 2014)

▶ Examples (Tourangeau et al. 2014)

▶ Use of interactive elements (e.g., Emde and Fuchs 2012)

Convert nonrespondents to provide substantive responses:

▪ Forced answer **→ Risk of break-off**

▪ Repetition of open-ended question

▪ Motivational sentences (see also Oudejans & Christian 2011; Zuell et al. 2014)

Please consider the question again. Your answer is very important for this research project.

Please answer in a bit more detail. This is important so that we can understand your answer better.

You seem to be in a hurry! Please take another moment to answer the question in as much detail as possible.

Types of Nonresponse

We can distinguish between different types of nonresponse that differ regarding the respondents' strength of intention to avoid answering a question:

Willingness to provide substantive response

One word only	economy
Don't knows	I have no idea, I can't make up my mind
Other nonresponse	my personal experience, it depends
Too fast response (Response took less than 2 seconds)	
Refusals	no comment
Complete nonresponse	
No useful answer	dfgjh frtrt

How to detect & convert nonresponse?

Precondition of nonresponse conversion: Detect instances of NR:

- Complete nonresponse: Easy to detect (empty answer box)
- Remaining nonresponse types: Harder to detect, needs coding

EvalAnswer:

- Tool for automatic detection of different types of nonresponse
- Conversion attempts: Repetition of open-ended question & motivational sentence
- Languages: German, English, Spanish
- Free tool available (https://git.gesis.org/surveymethods/evalanswer)
- Kaczmirek, Meitinger, & Behr (2017): Working paper on technical implementation

Higher data quality in web probing with EvalAnswer: a tool for identifying and reducing nonresponse in openended questions

Kaczmirek, Lars; Meitinger, Katharina; Behr, Dorothée

EvalAnswer: Automatic Nonresponse Detection and Conversion

How much do you agree or disagree with the following statement?

I feel more like a citizen of the world than of any country.

○ agree strongly
○ agree
● neither agree nor disagree
○ disagree
○ disagree strongly
_____
○ *can't choose*

Closed question

Please explain why you selected "neither agree nor disagree".

The statement was: "I feel more like a citizen of the world than of any country."

stupid question

Open-ended probe

Automatic nonresponse detection

We would like to understand what you had in mind when you answered the original question. Please try to answer this follow-up question:

Please explain why you selected "neither agree nor disagree".
The statement was: "I feel more like a citizen of the world than of any country."

NR conversion attempt: Repetition of probe & motivational sentence

RESEARCH QUESTIONS

RQ1:   How prevalent is nonresponse in our data?

RQ2:   How prevalent are the different NR types and are there cross-national differences?

RQ3:   Can we convert nonrespondents?

RQ4:   Do countries and NR types differ regarding their conversion level?

Data

- Two web surveys with panelists from non-random online access panels (Study 1: N=2,685; Study 2: N=2,689)

- Country sample: Germany, Great Britain, the U.S., Spain, and Mexico

- Quotas for age (18-30, 31-50, 51-65), gender, and education (lower education vs. higher education)

- Data collection in May and June 2014

- Replication of questions from International Social Survey Programme Modules (2012,2013, and 2014)

- 29 open-ended questions (probes) with 35,252 responses

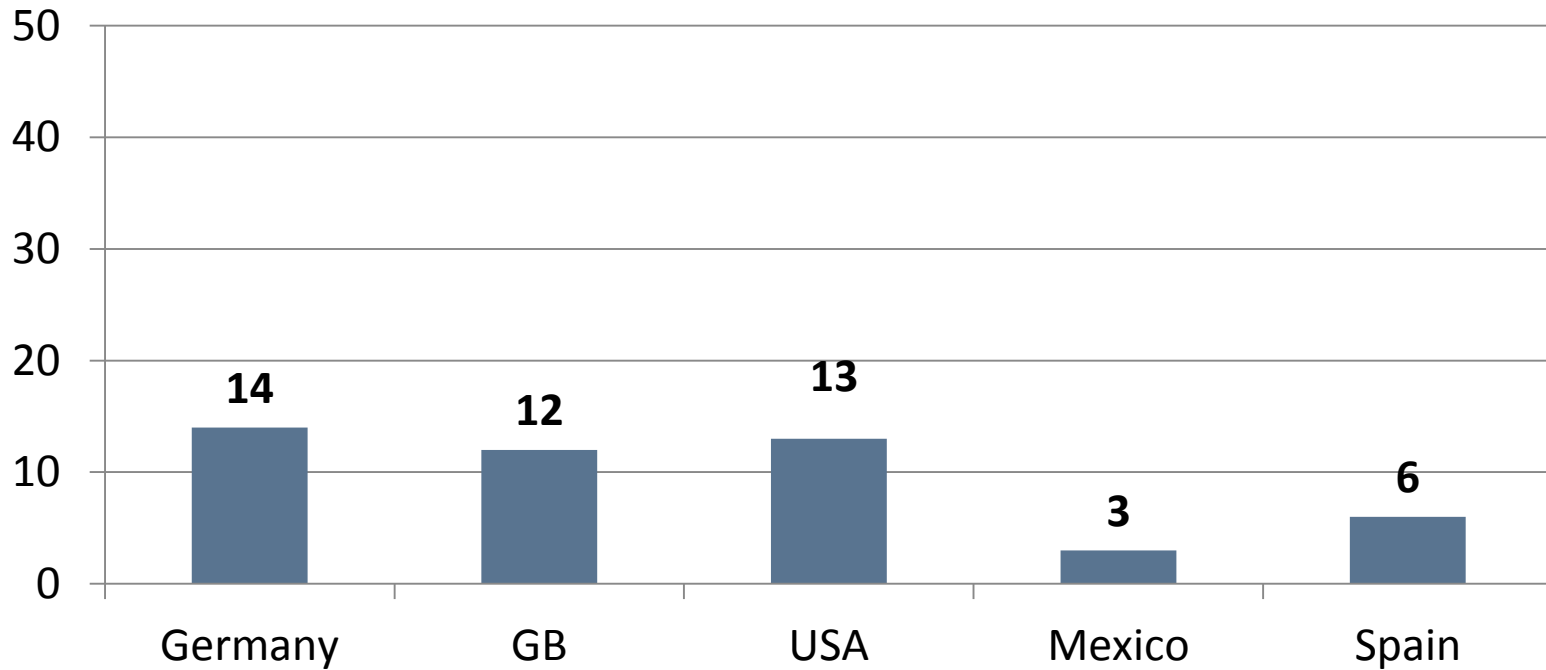- Automatic nonrespose-detection and conversion (EvalAnswer)

RQ 1

How prevalent is nonresponse in the five countries?

Nonresponse

- Overall nonresponse (29 questions): 9.45%
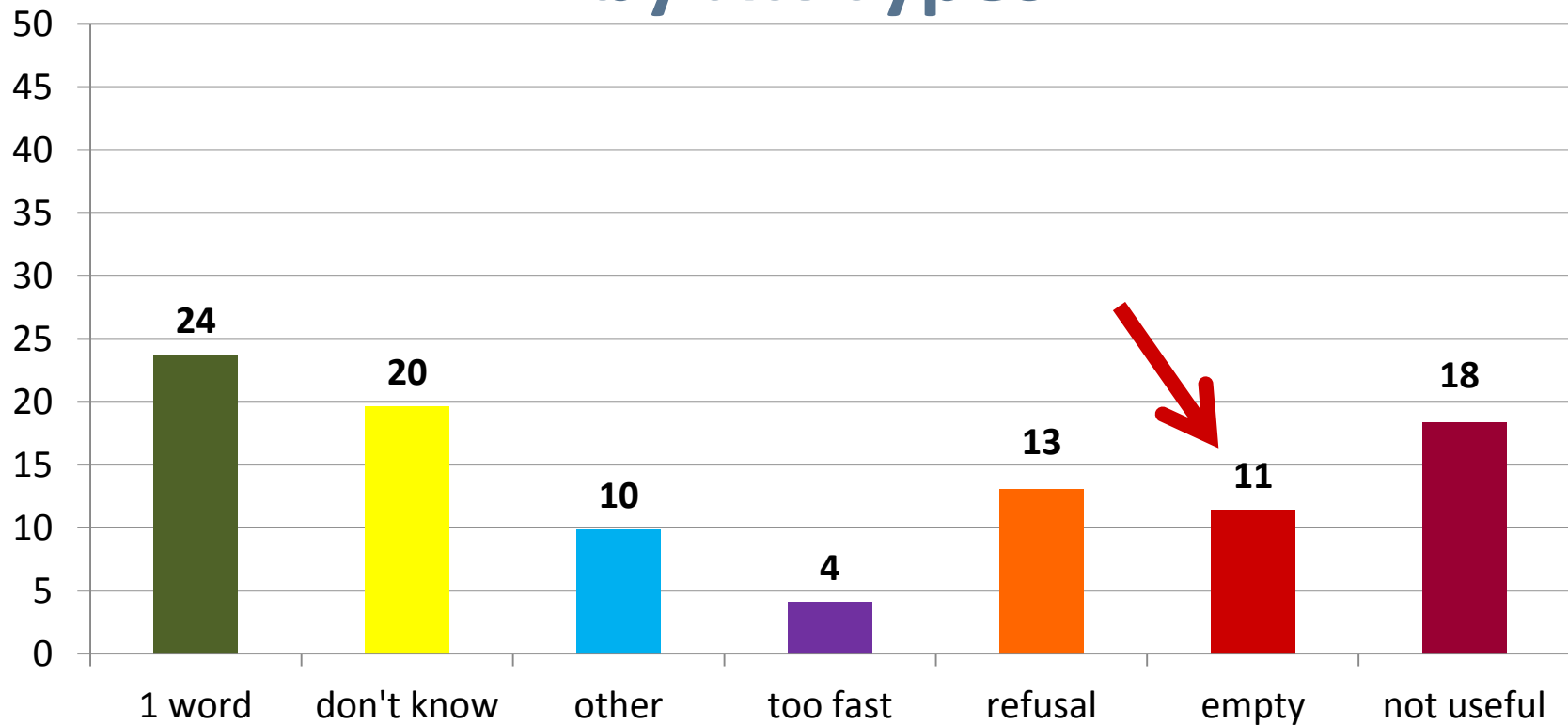
Mean Percent NR By Country

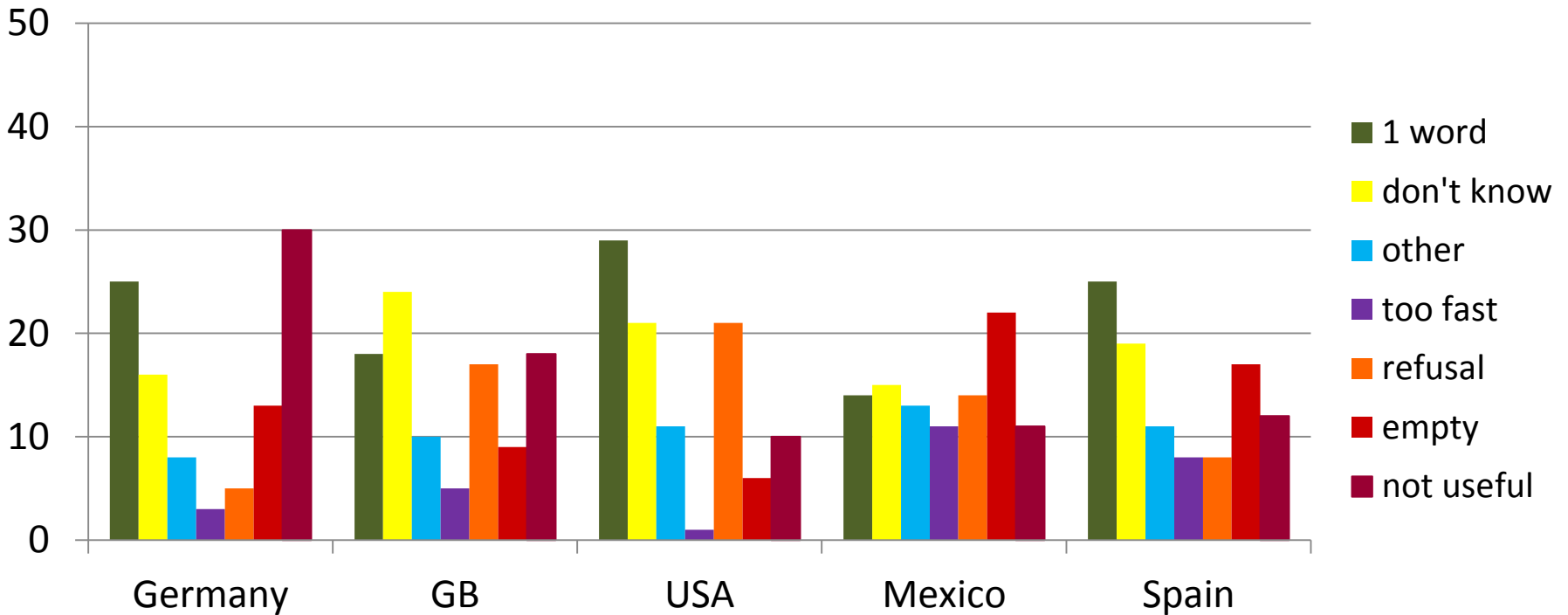→ **NR less prevalent in Mexico and Spain than in other countries**

RQ 2

How prevalent are the different NR types and are there cross-national differences?

Overall Nonresponse Distribution by NR Types

→ **"1 word," "don't know," "not useful" most frequent NR types**
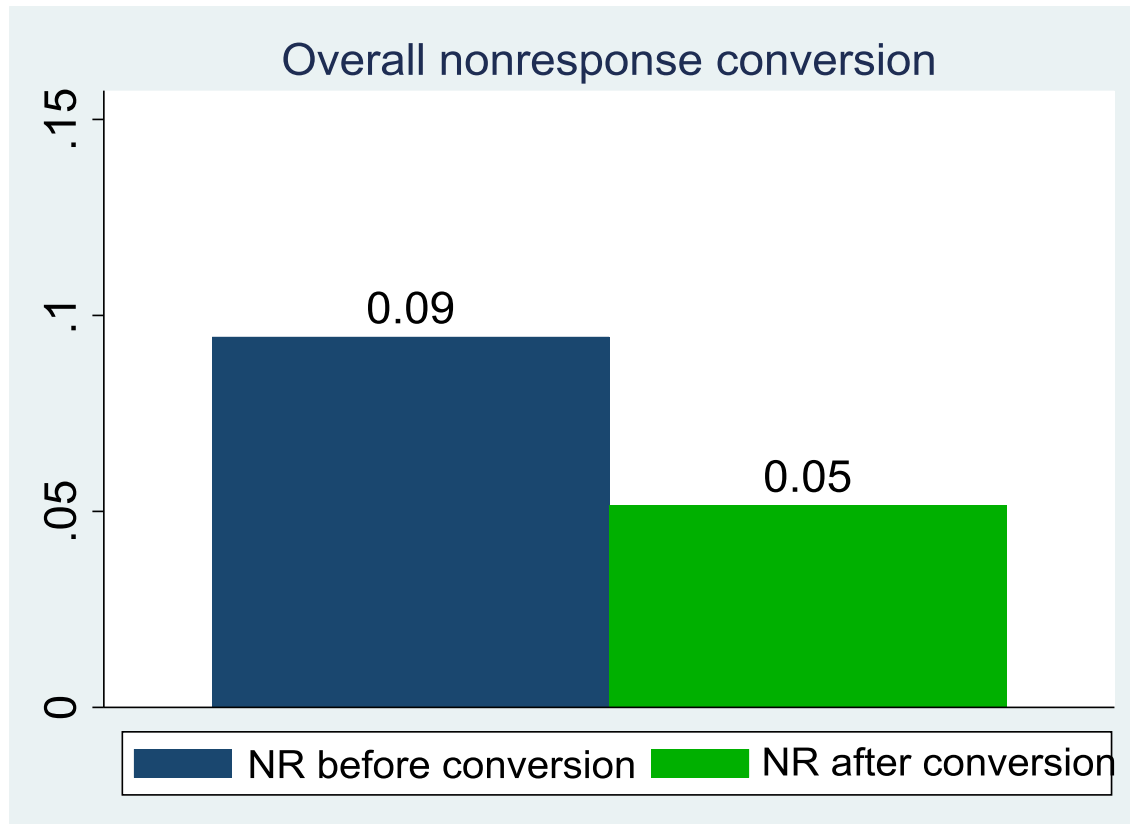
Nonresponse Types by Country

Country differences regarding prevalence of NR types:
→ Germany: "not useful"
→ GB: "don't know"
→ U.S.: "refusal"
→ Mexico: "empty answer box"

RQ 3

Can we convert nonrespondents?

Conversion rates - Overall

Overall nonresponse conversion

0.09

0.05

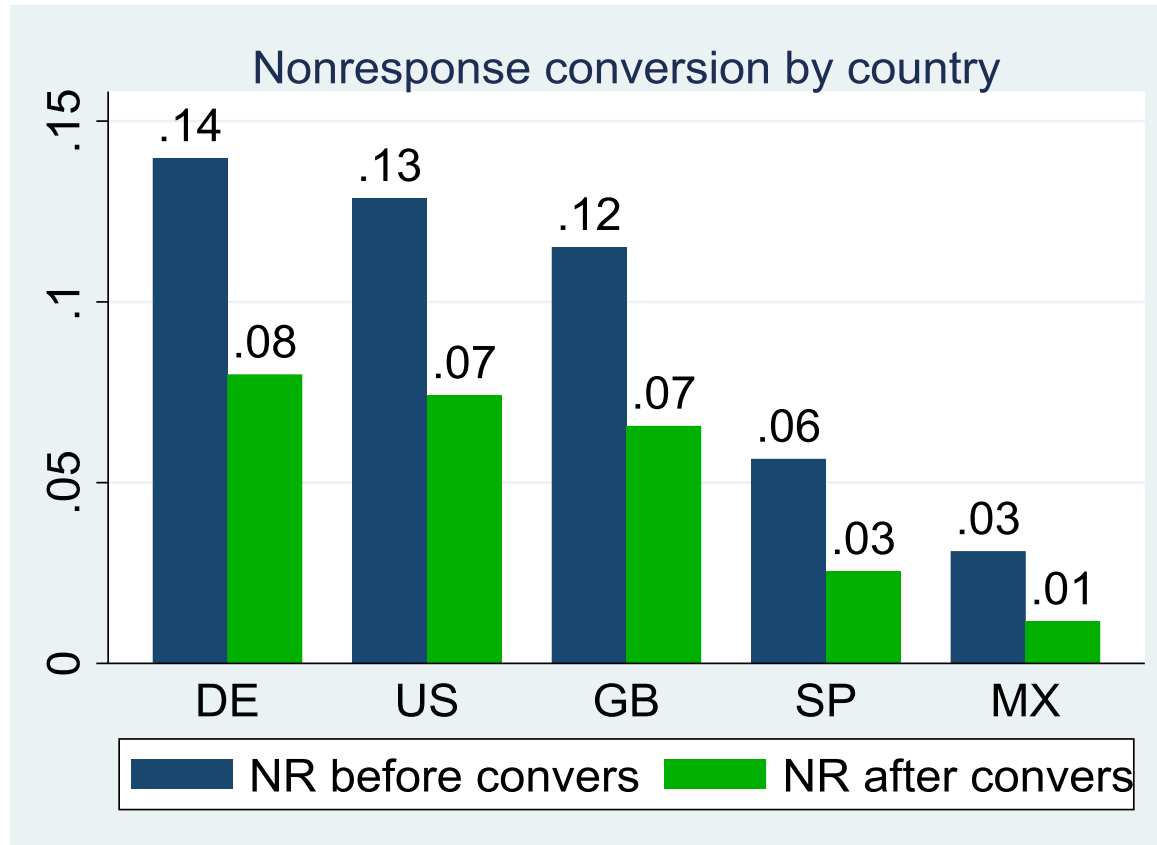NR before conversion   NR after conversion

→ We can approximately reduce nonresponse by half

RQ 4

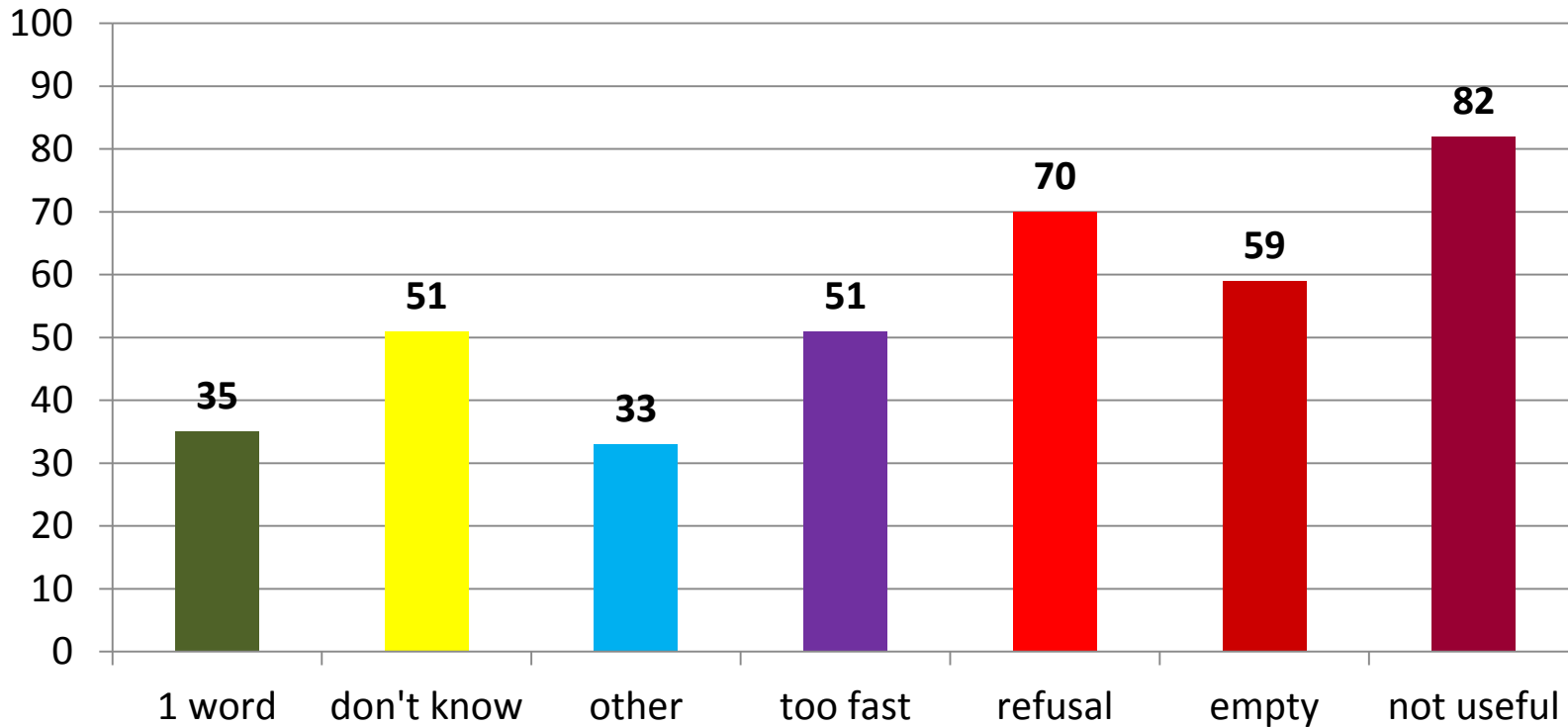## Do countries and NR types differ regarding their conversion level?

Conversion rates – By Country

Nonresponse conversion by country

→ Nonresponse conversion works in all countries

Conversion rates: By NR Type
% of Initial Nonresponse

Bar chart showing conversion rates by NR type:
- 1 word: 35
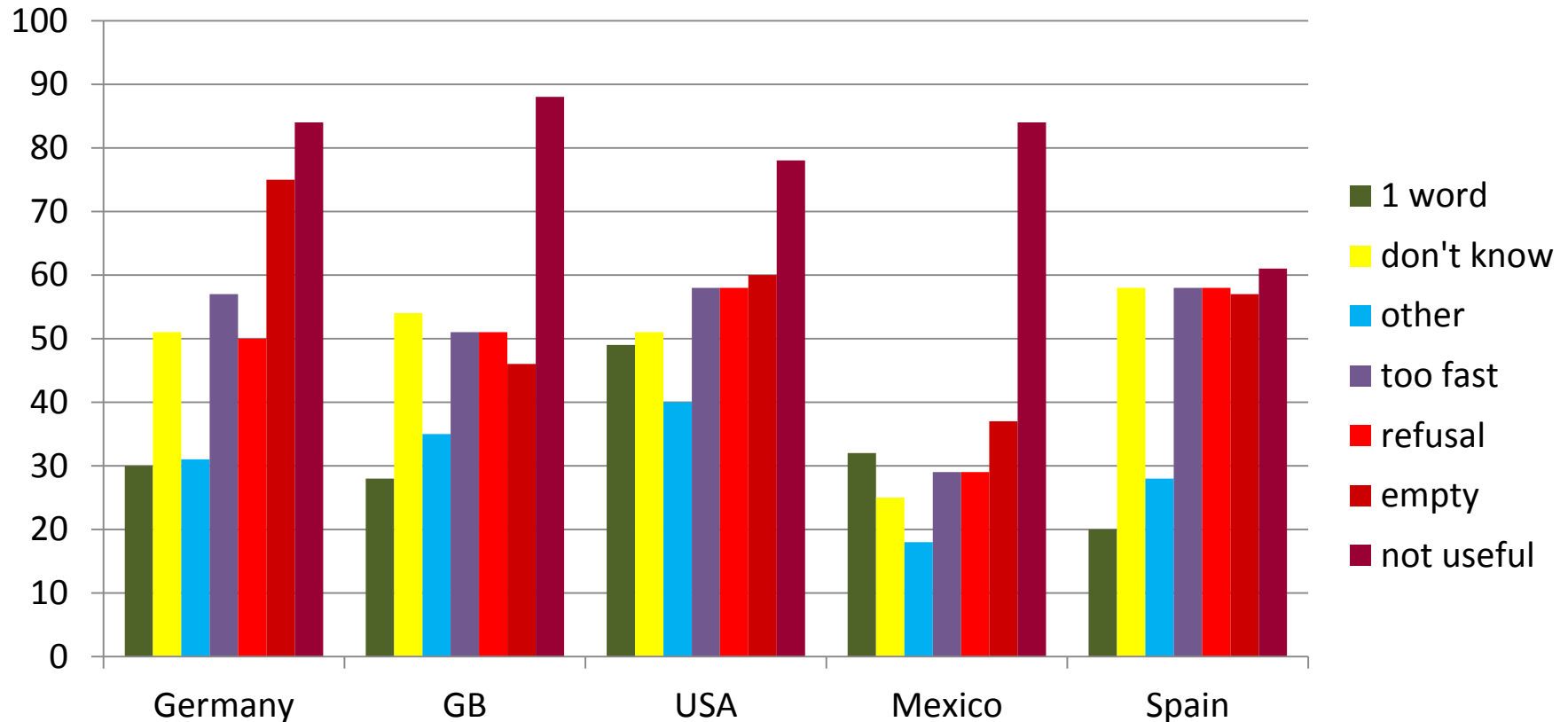- don't know: 51
- other: 33
- too fast: 51
- refusal: 70
- empty: 59
- not useful: 82

→ There are differences in conversion rates along NR Types
→ "Unwilling" NR types more difficult to convert

Conversion rates: NR Type & Country % of Initial Nonresponse

→ "1 word" & "other" most successfully converted in all countries
→ "not useful": lowest conversion rates
→ "empty answer box": Germany lowest, Mexico highest conversion
→ "Don't know": conversion works best in Mexico

CONCLUSION

Nonresponse...

- differed in prevalence across countries (Mexico and Spain lowest NR)
- differed in prevalence across NR types:
  - ▸ "1 word," "don't know," "not useful" most frequent NR types
  - ▸ "empty answer box" account for only 11% of all NR types
    → likely there is an underestimation of NR in classical methodological studies
- differed in prevalence across NR types and countries:
  - ▸ Germany: "not useful"; GB: "don't know"; U.S: "refusal"; Mexico: "empty answer box"

- NR conversion:
  - ▸ reduces NR by half and works in all countries
  - ▸ NR Types: "Unwilling" respondents more difficult to convert
  - ▸ Similarities in conversion across countries: easiest converted were "1 word" & "other" successfully; most difficult to convert: "not useful"
  - ▸ Differences in conversion:
    - "empty answer box": Germany lowest, Mexico highest conversion
    - "Don't know": Mexico highest conversion

Next steps & things to consider

- Develop more targeted approaches to reduce nonresponse that adress specific nonresponse types

- Necessary to avoid pushing respondents too far

- Understanding the cultural specifities of nonresponse

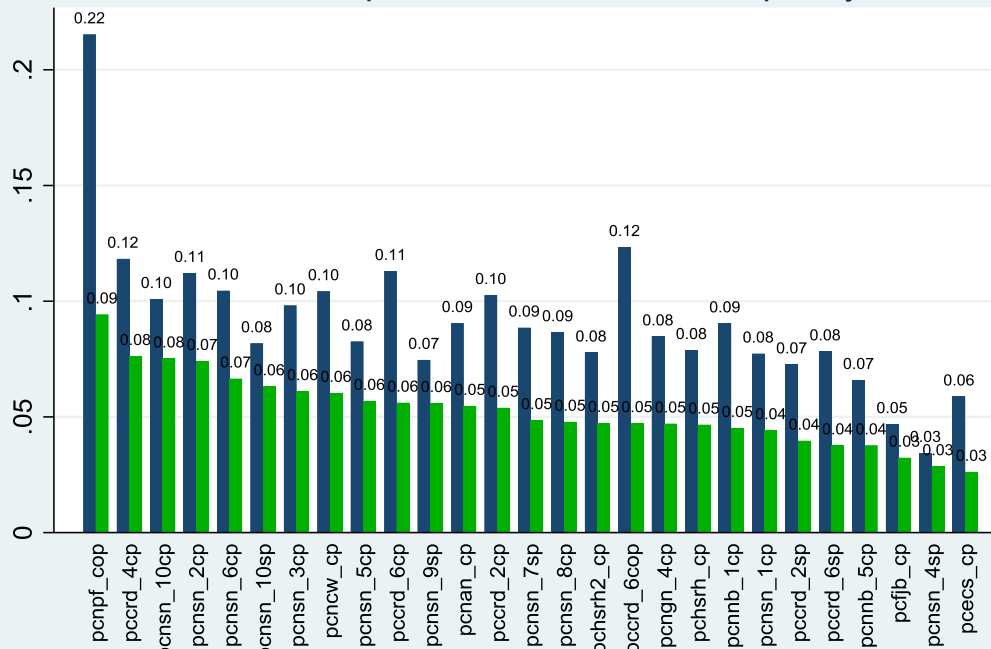- Extending tool by adding additional languages

Thank you!

Contact: katharina.meitinger@gesis.org

gesis   Leibniz Institute
          for the Social Sciences

Success of nonresponse conversion attempts by variable

```
Logistic regression                                    Number of obs    =      35,252
                                                       Wald chi2(7)     =      501.37
                                                       Prob > chi2      =      0.0000
Log pseudolikelihood = -10360.978                      Pseudo R2        =      0.0608
```

(Std. Err. adjusted for **5,374** clusters in case_id)

nonresponse1	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf. Interval]	
age	-.0245711	.0020299	-12.10	0.000	-.0285497	-.0205925
men	.1188542	.0608898	1.95	0.051	-.0004876	.238196
edu_high	-.4528201	.0618897	-7.32	0.000	-.5741217	-.3315185
DE	1.67613	.1125621	14.89	0.000	1.455512	1.896748
SP	.6440151	.1230056	5.24	0.000	.4029284	.8851017
GB	1.444733	.1137312	12.70	0.000	1.221824	1.667642
US	1.552755	.1153124	13.47	0.000	1.326747	1.778763
_cons	-2.360749	.1293158	-18.26	0.000	-2.614204	-2.107295