

# Asking Sensitive Questions: Possibilities and Limits of Randomized Response and Other Techniques in Different Survey Modes

*Marc Höglinger (ETH Zurich)*

*Ben Jann (University of Bern)*

*Ivar Krumpal (University of Leipzig)*

*Felix Wolter (University of Mainz)*

PPSM Meeting 'Survey Methods in Future Research', Bremen, September 12-13

# Outline

- Misreporting of sensitive issues in surveys
- Indirect methods to reduce misreporting: RRT, CM, and ICT
  - Randomized Response Technique (RRT)
  - Crosswise Model (CM)
  - Item Count Technique (ICT)
- An overview of our studies
- Conclusions and future research

# Misreporting of sensitive issues in surveys

- Survey respondents might not tell the truth if asked questions on sensitive issues such as norm violations, deviant behavior or stigmatizing traits. Or they might not respond at all.
- As a consequence, prevalence estimates are biased (e.g., social desirability bias) and apparent correlations might be an artifact.
- Validation studies reveal a considerable share of ‘liars’ (respondents incorrectly denying a sensitive characteristic) in surveys that use direct questioning:
  - Penal conviction: 42.5% (F2F, Wolter & Preisendörfer 2013)
  - Welfare and unemployment benefit fraud: 75% (F2F, van der Heijden et al. 2000)
  - Driving under influence: 54% (P&P, Locander et al. 1976)
  - Bankruptcy: 32% (Ibid.)

# Strategies to reduce misreporting of sensitive issues

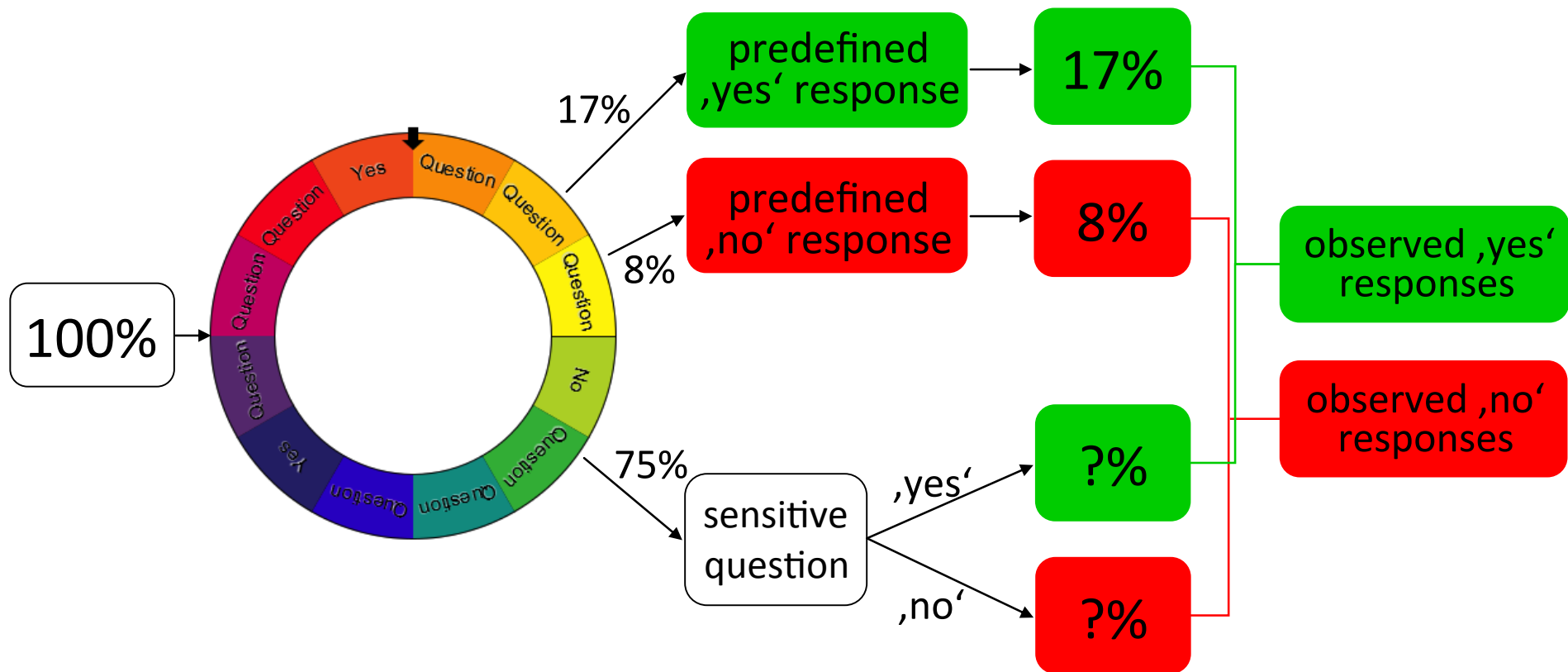
- Choice of a more anonymous survey mode, e.g. without interviewer being present:
  - Compared to F2F, misreporting seems to be lower in CATI mode and – even more – when no interviewer at all is present as in P&P, online or IVR mode (e.g. Tourangeau & Yan 2007, Kreuter, Presser, Tourangeau 2008).
- Question wording to attenuate perceived sensitivity (e.g., forgiving wording)
- Interview situation (e.g., no bystanders)
- Confidentiality/anonymity assurance
- Sealed envelope technique
- Indirect methods

# Indirect methods to reduce misreporting

- Basic principle of indirect methods: Respondents are given full response privacy at the individual level, to make them answer (more) honestly.
- How well these indirect methods perform is far from clear.
- Indirect methods are still ‘exotic’ to most survey researchers and their actual use is very limited.
- The PPSM project «Asking sensitive questions», based at the Universities of Leipzig (Thomas Voss, Karl-Dieter Opp, Ivar Krumpal), Mainz (Peter Preisendörfer, Felix Wolter) and at the ETH Zurich (Andreas Diekmann, Ben Jann, Marc Höglinger) addressed these issues.
- Methods that have been evaluated:
  - Randomized Response Technique (RRT, Warner 1965)
  - Crosswise Model (CM, Yu, Tian, and Tang 2008), a new variant of the RRT.
  - Item Count Technique (ICT, e.g. Droitcour et al. 1991)
  - Item Sum Technique (Trappmann et al. 2013)

# The Randomized Response Technique (RRT)

- RRT originally presented by Warner (1965)
- Idea: anonymize a respondent's answer using a randomization procedure administered by the respondent

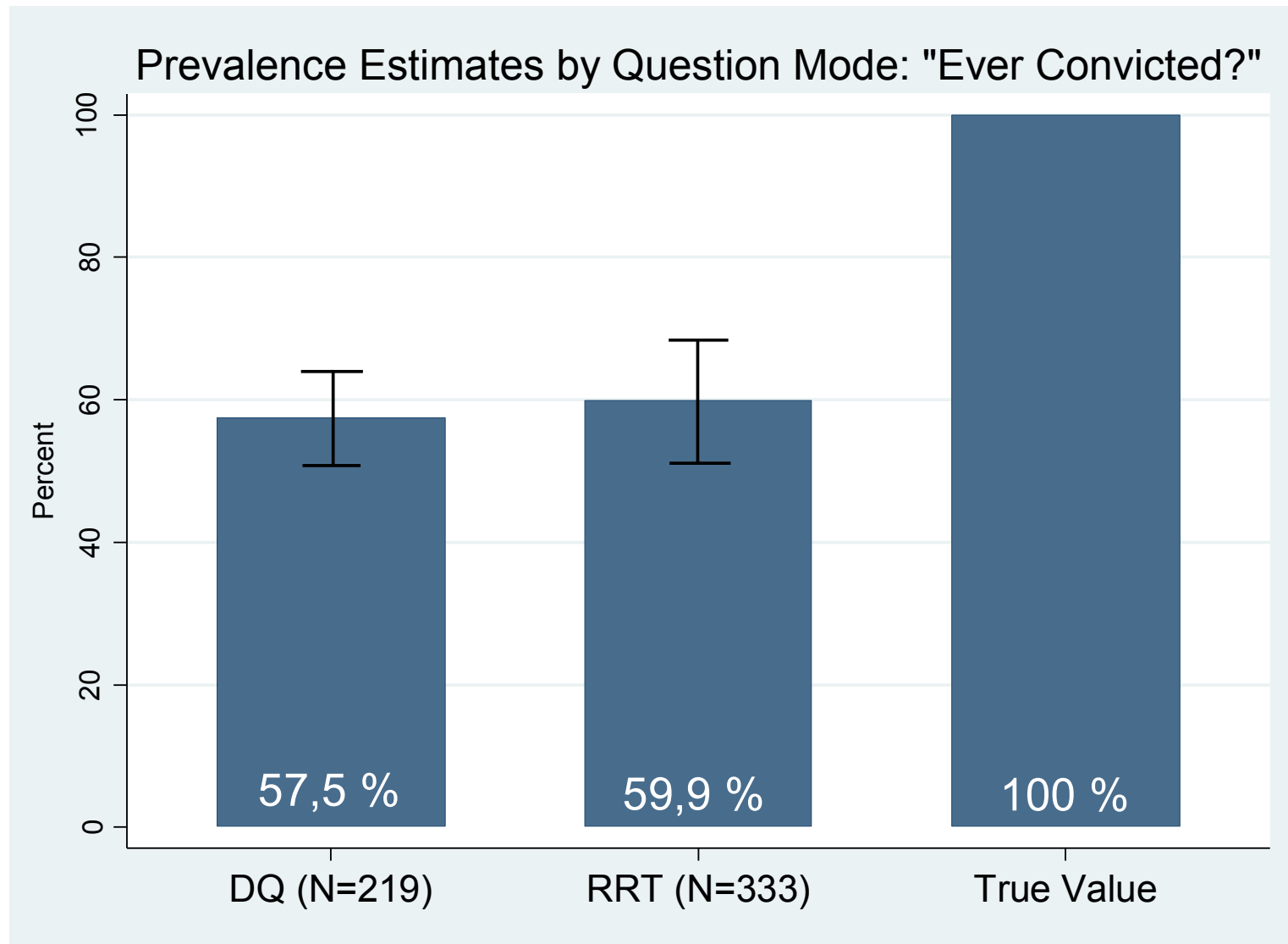


# RRT performance

- RRT seems to work well in F2F and P&P mode: Less average prevalence underestimation than DQ (Meta-Analysis of 23 F2F and 13 P&P RRT studies 1965-2000, Lensvelt-Mulders et al. 2005).
- However, other studies cast doubt whether RRT really produced more valid measurements (Wolter & Preisendörfer 2013, Krumpal, Jann, Auspurg, von Hermanni., fc).
- RRT does not seem to work well in online mode. Results so far showed...
  - no different or even lower prevalence estimates for socially undesirable behavior compared to direct questioning or unrealistically high prevalence estimates for desirable behavior (Coutts et al. 2011 , Coutts & Jann 2011, Peeters 2006, Snijders & Weesie 2008, Holbrook & Krosnick 2010, Höglinger, Jann, Diekmann 2013)

# RRT-Validation, F2F

Wolter & Preisendörfer 2013





# Reasons for RRT failure

- respondents' lack of understanding of RRT's principle 'protection through randomization', no trust in RRT
- reluctance of respondents to give a predefined 'yes' response (Edgell et al. 1982, Lensvelt-Mulders/Boeije 2007)
- self-protective 'no'-bias: to be on the safe side, respondents' dominant strategy is to give always the 'no' response (Ostapczuk, Musch & Moshagen 2009, Jann, Jerke, Krumpal 2011)
- Online: RRT implementations often not well suited to online mode with randomizing devices not at respondents' immediate reach and requiring mode shift (e.g., real coin, banknote, previously sent real dice) or not trustworthy (e.g., virtual coin/dice/random wheel)

# The Crosswise Model (CM)

A new alternative to conventional RRT (Yu, Tian, and Tang 2008)

- Simple idea: Ask a sensitive question and a non-sensitive question.
- Let the respondent indicate whether the answers to the two questions...
  - are **identical** (both 'yes' or both 'no')
  - are **different** (one 'yes', the other 'no')

		<i>non-sensitive question</i>	
		no	yes
<i>sensitive question</i>	no	identical	different
	yes	different	identical

# Performance of CM

- The Crosswise Model seems to be a promising alternative to conventional RRT:
  - higher prevalence estimates than with direct questioning in a p&p survey on plagiarism (Jann, Jerke, Krumpal 2012)
  - and, also, in an online survey on plagiarism and cheating in exam with university students (Höglinger, Jann, Diekmann 2013)
- Advantages of the Crosswise Model over RRT
  - privacy protection is easy to understand, as response options ‘identical’/‘different’ are obviously ambiguous
  - no obvious self-protective answering strategy (such as, e.g., always tick ‘no’)
  - respondents do not feel forced into giving a ‘false’ predefined response as in the case of the forced response RRT

# CM, P&P

Jann, Jerke & Krumpal 2012

**Table 2. Prevalence Estimates, in Percent, of Partial and Severe Plagiarism by Experimental Condition** (standard errors in parentheses)

	Direct Questioning (DQ)	Crosswise Model (CM)	Difference (CM – DQ)
Partial Plagiarism	7.3 (2.7)	22.3 (5.5)	15.0 (6.1)
Severe Plagiarism	1.0 (1.0)	1.6 (5.0)	0.6 (5.1)
Observations	96	310	

# The Item Count Technique (ICT)

- Principle: A list of items is asked, respondents indicate the number of “Yes”-responses.
- One group of respondents receives a short item list without the sensitive question, another group receives a long list including the sensitive item.
- Mean difference between the two groups is the prevalence estimate of the sensitive characteristic.
- Double list design to increase efficiency (Droitcour et al.,1991).

## Group SL1 (Short List)

Did you ever keep a book of household accounts?  
Does your household contain more TV's than members?  
Do you own a mobile phone?

## Group LL1 (Long List)

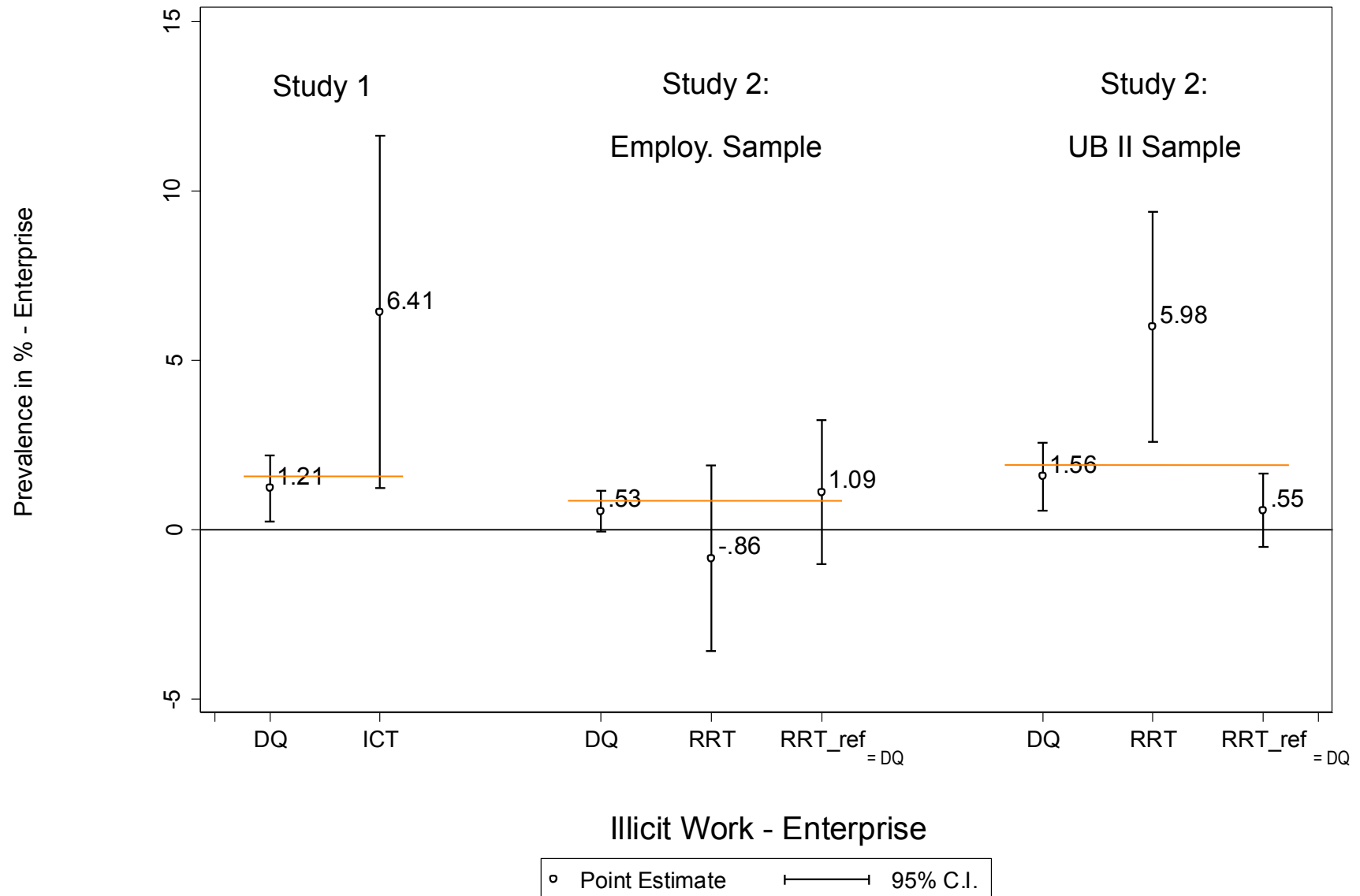
Did you ever keep a book of household accounts?  
Does your household contain more TV's than members?  
Do you own a mobile phone?  
Did you carry out any illicit work for an enterprise this year [2010], that paid you partially in cash, i.e. did not or not fully report this to the authorities?

# ICT performance

- Research, suggests that the ICT is better understood and more trusted than the RRT (e.g. Coutts and Jann 2011).
- In some studies, the ICT outperformed the RRT (Holbrook & Krosnick 2010).
- However, the ICT has some drawbacks:
  - low statistical efficiency, even relative to the RRT
  - choice of the non-sensitive items can be tricky
  - mentally counting the items is cognitively demanding (particularly in CATI mode)

# ICT

Kirchner, Krumpal, Trappmann, von Hermann 2013



# Studies of the PPSM project “Asking sensitive questions”

- Comparative studies (under the “more is better” assumption):
  - F2F survey on deviant behavior (RRT, Wolter 2008)
  - CATI survey on xenophobia and anti-Semitism (RRT, Krumpal 2012)
  - CATI survey on illicit work/unemployment benefits (RRT & ICT, Kirchner et al. 2013)
  - Online survey on various sensitive behaviors (RRT & ICT, Coutts & Jann 2011)
  - Online survey on exam cheating and plagiarism (RRT & CM, Höglinger, Jann, Diekmann 2013)
  - P&P survey on plagiarism (CM, Jann, Jerke, Krumpal 2012)
- Validation studies:
  - P&P & F2F survey of convicted persons, validation with court records (RRT, Wolter & Preisendörfer 2013)
  - CATI survey on illicit work/unemployment benefits (RRT & ICT, Kirchner, Krumpal, Trapmann, von Hermanni 2013)
  - Online survey on cheating in an experimental task (RRT, CM, Höglinger, Diekmann, Jann, ongoing)
- Mode differences:
  - P&P vs. F2F survey of convicted persons, validation with court records (RRT, Preisendörfer & Wolter fc)



# Results overview

Sensitive question technique	Results per item		
	DQ < Special Technique	DQ = Special Technique	DQ > Special Technique
Randomized Response (RRT)	10	17	5
Crosswise Model (CM)	5	2	-
Item Count Technique (ICT)	3	7	-

# Conclusions

- Sensitive questions in surveys are a critical issue: There is a non-ignorable amount of misreporting.
- Summary of the RRT-Evaluation: No convincing evidence for better results than DQ in general
  - In most studies, prevalence estimates from classic RRT variants not systematically higher or more valid than direct questioning estimates
  - Implementation details are important: High variance in RRT-performance depending on randomizing device and other design features
  - Success of RRT also depends on respondent and situational characteristics
- However, there are solutions:
  - Item-count-technique: mostly higher prevalence estimates, but big samples necessary
  - Crosswise model: clearly higher prevalence estimates in P&P and online mode
  - Statistical tools for improving RRT-estimates: Cheating detection models (e.g., Clark & Desharnais 1998, Moshagen, Musch, Erdfelder 2011)

# Future research

- Further tests of the promising RRT-alternatives: Crosswise Model, Item Count Technique
- Evaluation of different RRT implementations (e.g. different randomizing devices, framing)
- More validation studies!
- Impact of the degree of question sensitivity and, also, the type of sensitivity on misreporting:
  - social desirability (e.g. shameful behavior)
  - threat of disclosure (e.g. illegal activities)
  - offensive questions (e.g. political participation, attitudes, income)

# Thank you!

Marc Höglinger (ETH Zurich)  
marc.hoeglinger@soz.gess.ethz.ch

Ben Jann (University of Bern)  
ben.jann@soz.unibe.ch

Ivar Krumpal (University of Leipzig)  
krumpal@sozio.uni-leipzig.de

Felix Wolter (University of Mainz)  
Felix Wolter felix.wolter@uni-mainz.de

# References I

- Coutts, E. und B. Jann (2011): "Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT)." *Sociological Methods & Research* 40:169-193.
- Droitcour, J., R. A. Caspar, M. L. Hubbard, T. L. Parsely, W. Vischer, and T. M. Ezzati. 1991. "The item count technique as a method of indirect questioning: A review of its development and a case study application." Pp. 185-210 in *Measurement errors in surveys*, edited by P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman. New York: Wiley.
- Fox, J. A. und P. E. Tracy. 1986. *Randomized response: A method for sensitive surveys*, Newbury Park, CA: Sage.
- van der Heijden, P. G. M., G. van Gils, J. Bouts, and J. J. Hox. 2000. "A Comparison of Randomized Response, Computer-Assisted Self-Interview, and Face-to-Face Direct Questioning. Eliciting Sensitive Information in the Context of Welfare and Unemployment Benefit." *Sociological Methods and Research* 28:505–537.
- Höglinger, M., B. Jann, and A. Diekmann. 2013. "Sensitive questions in online surveys: An experimental comparison of the Randomized Response Technique and the Crosswise Model." in Working Paper. ETH Zurich and University of Bern.
- Jann, B., J. Jerke und I. Krumpal. 2011. *Asking Sensitive Questions Using the Crosswise Model: Some Experimental Results*, *Public Opinion Quarterly*.
- Holbrook, A. L., and J. A. Krosnick. 2010. "Measuring Voter Turnout By Using The Randomized Response Technique: Evidence Calling Into Question The Method's Validity." *Public Opin Q* 74:328-343.
- —. 2010. "Social desirability bias in voter turnout reports: Tests using the item count technique." *Public Opinion Quarterly* 74:37-67.

## References II

- Kreuter, F., S. Presser, and R. Tourangeau. 2008. "Social Desirability Bias in CATI, IVR, and Web Surveys." *Public Opinion Quarterly* 72:847-865.
- Krumpal, I. 2012. "Estimating the prevalence of xenophobia and anti-Semitism in Germany: A comparison of randomized response and direct questioning." *Social Science Research* 41:1387-1403.
- Krumpal, I., B. Jann, K. Auspurg und H. von Hermanni (im Erscheinen): "Asking Sensitive Questions: A Critical Account of the Randomized Response Technique and Related Methods." in *Improving Survey Methods: Lessons from Recent Research*, herausgegeben von Uwe Engel, Ben Jann, Peter Lynn, Annette Scherpenzeel, und Patrick Sturgis. New York: Routledge/Taylor & Francis
- Lensvelt-Mulders, G. J. L. M. und H. R. Boeije. 2007. Evaluating compliance with a computer assisted randomized response technique: a qualitative study into the origins of lying and cheating, *Computers in Human Behavior*, Vol. 23, 591-608.
- Locander, W., S. Sudman, and N. Bradburn. 1976. "An Investigation of Interview Method, Threat and Response Distortion." *Journal of the American Statistical Association* 71:269-275.
- Moshagen, M., J. Musch, and E. Erdfelder. 2011. "A stochastic lie detector." *Behavior Research Methods*:1-10.
- Tourangeau, R., and T. Yan. 2007. "Sensitive Questions in Surveys." *Psychological Bulletin* 133:859-883.
- Trappmann, M., I. Krumpal, A. Kirchner und Ben Jann (2013): "Item Sum - A New Technique for Asking Quantitative Sensitive Questions." IAB Nürnberg / Universität Leipzig / Universität Bern, unveröffentlichtes Arbeitspapier.
- Warner, S. L. 1965. "Randomized-response: A survey technique for eliminating evasive answer bias." *Journal of the American Statistical Association* 60:63-69.

## References III

- Wolter, F., and P. Preisendörfer. 2013. "Asking Sensitive Questions: An Evaluation of the Randomized Response Technique vs. Direct Questioning Using Individual Validation Data." *Sociological Methods & Research*.
- Yu, J.-W., G.-L. Tian, and M.-L. Tang. 2008. "Two new models for survey sampling with sensitive characteristic: design and analysis." *Metrika* 67:251-263.