

# Soziale Erwünschtheit und Randomized Response: Ein experimenteller CATI-Survey

UNIVERSITÄT LEIPZIG  
INSTITUT FÜR SOZIOLOGIE



Ivar Krumpal  
Institut für Soziologie, Universität Leipzig  
krumpal@sozio.uni-leipzig.de

**Workshop: Rational Choice Sociology - Theory and Applications,  
Venice International University, San Servolo, 2009**

---

# Gliederung des Vortrags

- Einleitung
  - Heikle Fragen und Soziale Erwünschtheit
  - Randomized Response Technik (RRT)
- Forschungsdesign
  - Hypothesen und Sampling
  - CATI-Implementation von RRT
- Empirische Ergebnisse
- Schlussfolgerungen und Ausblick

## Was sind unangenehme/ heikle/sensitive Fragen?

- Fowler (1995: 29): *“Questions tend to be categorized as ‘sensitive’ if a ‘yes’ answer is likely to be judged by society as undesirable behaviour. However, for those for whom the answer is ‘no’, questions about any particular behaviour are not sensitive.”*
- Tourangeau/Yan (2007: 860): *“A question is sensitive when it asks for a socially undesirable answer, when it asks in effect, that the respondent admits he or she has violated a social norm.”*

# Sozial erwünschtes Antwortverhalten und Messfehler

- Systematische Antworttendenzen aufgrund sozialer Erwünschtheit gefährden die Validität von Prävalenzschätzungen eines sensitiven Merkmals („social desirability bias“)
- Systematisches „Underreporting“ von negativ konnotierten Eigenschaften
- Systematisches „Overreporting“ von positiv konnotierten Eigenschaften

# Item-Nonresponse

	ALLBUS (1990)	ALLBUS (2000)	ALLBUS (2006)
Einkommen	<b>26.2 %</b>	<b>23.5 %</b>	<b>20.7 %</b>
Wahlabsicht	<b>14.4 %</b>	<b>22.7 %</b>	<b>14.2 %</b>
Religion	0.4 %	0.7 %	0.5 %
Bildung	0.8 %	0.3 %	0.2 %
Erwerbsstatus	0.1 %	0.2 %	0.1 %
Alter	0.4 %	0.0 %	0.3 %

# Randomized Response Technik (RRT)

- RRT schützt die Privatsphäre von Befragten (Warner 1965; Boruch 1971)
- Ein Zufallsgenerator (z.B. Würfel, Münzwurf) stellt eine probabilistische Beziehung zwischen der heiklen Frage und einer gegebenen Antwort her.
- Da die Anonymität des Befragten gewahrt bleibt, werden mehr ehrliche Antworten erwartet.

# Randomized Response Technik (RRT)

- Beispielsweise wirft jeder Befragte unbeobachtet einen Würfel. Bei einem Ergebnis von 2-5 soll der Befragte die heikle Frage wahrheitsgemäß beantworten.
- Wirft der Befragte dagegen eine 1 (eine 6), so soll er, unabhängig vom Frageinhalt, das Vorliegen des heiklen Merkmals bejahen (bzw. verneinen).
- Da die Wahrscheinlichkeitsverteilung durch das Design vorgegeben ist, können Prävalenzraten geschätzt und multivariate Zusammenhänge untersucht werden.

# Randomized Response Technik (RRT)

- Die wenigen empirischen Studien zur Effektivität der RRT im telefonischen Modus, auch CATI („Computer-Assisted-Telephone-Interviewing“) Telefonumfragen genannt, belegen bisher keinen Vorteil gegenüber der konventionellen, direkten Befragungsmethode (Stem & Steinhorst 1984; Weissman et al. 1986; McAuliffe et al. 1991).

# DFG Schwerpunktprogramm 1292

## "Survey Methodology"

- Kooperationsprojekt: „Asking Sensitive Questions: Possibilities and Limits of Randomized Response and Other Techniques in Different Survey Modes“
- Projektgruppe Leipzig: Ivar Krumpal, Karl-Dieter Opp, Thomas Voss
- Projektgruppe Mainz: Peter Preisendörfer, Felix Wolter
- Projektgruppe Zürich: Elisabeth Coutts, Andreas Diekmann, Ben Jann

# Hypothesen und Forschungsdesign

- Bundesweiter, experimenteller CATI-Survey zu Fremdenfeindlichkeit und Antisemitismus
- Offene Äußerung von fremdenfeindlichen und antisemitischen Einstellungen verstößt in der heutigen BRD gegen soziale Normen (Bergmann 1997).
- Eine Teilmenge der fremdenfeindlich bzw. antisemitisch eingestellten Befragten wird ihre Einstellungen im Rahmen einer Bevölkerungsbefragung nicht offenbaren.

# Hypothesen und Forschungsdesign

- „More-is-better-Assumption“: Aufgrund der erwarteten systematischen Unterschätzung der Prävalenz fremdenfeindlicher und antisemitischer Einstellungen wird angenommen, dass die validere Messmethode (direkte Befragung versus RRT) zu relativ höheren Prävalenzraten führen wird.

# Hypothesen und Forschungsdesign

- **Hypothese 1:** Da im Vergleich zu konventionellen CATI Interviews die Befragung mit der RRT zu einer höheren Anonymität im Befragungsprozess führt, werden Befragte bei einer Befragung mit der RRT im Durchschnitt mehr fremdenfeindliche und antisemitische Einstellungen offenbaren.
- **Hypothese 2:** Mit zunehmender Sensitivität des Untersuchungsthemas wird sich bei Interviews mit der RRT der Anteil sozial unerwünschter Antworten stärker erhöhen.

# Hypothesen und Forschungsdesign

- Grundgesamtheit: Alle bundesdeutschen Staatsbürger in Privathaushalten mit Festnetzanschluss (>18 Jahre)
- Sampling: Zufallsstichprobe auf Grundlage des ADM-Stichprobensystem für Telefonbefragungen
  - Stufe 1: Ziehung von Telefonnummern proportional zur regionalen Einwohnerstruktur auf Bundesland-Ebene.
  - Stufe 2: Innerhalb des Haushaltes wurde die Zielperson anhand eines Schwedenschlüssels zufällig ermittelt.
- Ausschöpfungsquote (AAPOR): 36.4 %

# Hypothesen und Forschungsdesign

- Datenerhebung in Kooperation mit USUMA; Feldphase: Januar-März 2009
- Vollständige Interviews: N = 2041 (durchgeführt von 38 Interviewern)
- Randomisierung im Verhältnis 1 : 2
  - Split 1: konventionelle CATI Interviews (N = 678)
  - Split 2: RRT (N = 1363)
- 5 sensitive, dichotome Items mit den Antwortoptionen „ja“ oder „nein“

# RRT Interviews mit 3 Münzen

Die Antworten folgen einer bekannten Verteilung :

- 3x Zahl → "Ja" ( $p_1 = 0.125$ )
- 3x Wappen → "Nein" ( $p_2 = 0.125$ )
- Mischung → "Wahrheit" ( $p_3 = 0.750$ )

Erwartungswert für eine "Ja"-Antwort :  $\phi = p_1 + p_3\pi$

$\pi$  ist der unbekannte Populationsparameter von fremdenfeindlichen bzw. antisemitischen Personen, die einer sozial unerwünschten Aussage zustimmen würden.

Da der beobachtete Stichprobenanteil von "Ja"-Antworten  $\hat{\phi}$  ein Schätzer von  $\phi$  ist, ist eine Schätzung des Populationsanteils  $\pi$  bzw. der Stichprobenvarianz  $\text{Var}(\pi)$  leicht möglich.

## Fremdenfeindlichkeit: ‚Es gibt zu viele Ausländer in Deutschland‘

	Direkt	RRT	RRT-Direkt
Zustimmung (in %)	38.8	41.3	50.7
Standard- Fehler	1.9	2.0	3.4
90-% K.I.	35.7 – 41.9	38.1 – 44.6	45.1 – 56.3
95-% K.I.	35.1 – 42.5	37.5 – 45.2	44.0 – 57.4
N	662	1129	215

„Die Bundesrepublik ist durch die vielen Ausländer in einem gefährlichen Maß überfremdet.“

	Direkt	RRT	RRT-Direkt
Zustimmung (in %)	27.0	34.5	35.9
Standard- Fehler	1.7	1.9	3.2
90-% K.I.	24.2 – 29.9	31.3 – 37.7	30.6 – 41.3
95-% K.I.	23.7 – 30.4	30.7 – 38.3	29.5 – 42.3
N	662	1131	220

„Die Ausländer kommen nur nach Deutschland,  
um unseren Sozialstaat auszunutzen.“

	Direkt	RRT	RRT-Direkt
Zustimmung (in %)	31.6	30.1	43.0
Standard- Fehler	1.8	1.9	3.4
90-% K.I.	28.6 – 34.6	26.9 – 33.2	37.4 – 48.6
95-% K.I.	28.0 – 35.2	26.3 – 33.8	36.3 – 49.7
N	655	1130	214

# Antisemitismus: ‚Auch heute noch ist der Einfluss der Juden zu groß.‘

	Direkt	RRT	RRT-Direkt
Zustimmung (in %)	12.1	17.4	21.0
Standard- Fehler	1.3	1.7	2.9
90-% K.I.	10.0 – 14.2	14.5 – 20.2	16.2 – 25.8
95-% K.I.	9.6 – 14.7	14.0 – 20.8	15.3 – 26.7
N	636	1124	200

# „Durch ihr Verhalten sind die Juden an ihren Verfolgungen nicht ganz unschuldig“

	Direkt	RRT	RRT-Direkt
Zustimmung (in %)	15.7	22.5	21.6
Standard- Fehler	1.4	1.8	2.9
90-% K.I.	13.3 – 18.1	19.5 – 25.5	16.8 – 26.4
95-% K.I.	12.9 – 18.5	18.9 – 26.0	15.8 – 27.4
N	638	1124	199

# Item-Nonresponse

	Direkt (%)	RRT (%)	RRT-Direkt (%)
Fremdenfeindlichkeit 1	2.4	0.7	4.9
Fremdenfeindlichkeit 2	2.4	0.5	2.7
Fremdenfeindlichkeit 3	3.4	0.6	5.3
Anti-semitismus 1	6.2	1.1	11.5
Anti-semitismus 2	5.9	1.1	12.0

# Empirische Ergebnisse

- Mehr sozial unerwünschte Antworten in RRT Version als in direkter Version (4 von 5 Anteilswertdifferenzen zeigen in die erwartete Richtung).
- Stärkere Erhöhung des Anteils sozial unerwünschter Antworten bei dem heikleren Thema Antisemitismus.
- Die Gruppe der Wechsler (RRT-Direkt) weist insgesamt das höchste Ausmaß an fremdenfeindlichen bzw. antisemitischen Vorurteilen auf.

# Empirische Ergebnisse

- Die ‚RRT-Direkt‘ Gruppe unterscheidet sich systematisch von den anderen beiden Gruppen (Direkt bzw. RRT).
- Ältere bzw. niedriger gebildete Personen haben eine höhere Wahrscheinlichkeit die RRT zu verweigern und gleichzeitig eine höhere Wahrscheinlichkeit Fremdenfeindlichkeit bzw. Antisemitismus zu offenbaren.
- Kontrolliert man die konfundierenden Drittvariablen im Rahmen eines logistischen Regressionsmodells, dann ist der Vorteil der RRT gegenüber der direkten Befragungsmethode klar erkennbar.

# Maximum-Likelihood Verfahren für RRT

Logistische Regression für RRT Daten schätzt Zusammenhang zwischen Kovariaten und einer dichotomen abhängigen Variable mit den Ausprägungen :1 =" Ja"; 0 =" Nein"  
(Maddala 1983; van der Heijden et al. 2000; Jann 2005)

$$\rightarrow \pi(\mathbf{x}_i) = \frac{\exp(b_0 + \sum_k x_{ik} b_k)}{1 + \exp(b_0 + \sum_k x_{ik} b_k)}$$

$\pi(\mathbf{x}_i)$  ist die Wahrscheinlichkeit einer "Ja"-Antwort als Funktion eines Vektors von Kovariaten  $\mathbf{x}_i$

$1 - \pi(\mathbf{x}_i)$  ist die entsprechende Wahrscheinlichkeit einer "Nein"-Antwort

$x_{ik}$  ist das  $k$  - te Element von  $\mathbf{x}_i$

$b_k$  ist der Regressionskoeffizient von Kovariate  $k$

---

Modifizierte Log - Likelihood Funktion im RRT - Design :

$$\rightarrow \log L = \sum_i n_{i1} \log [P_1 + P_3 \pi(\mathbf{x}_i)] + \sum_i n_{i0} \log [P_2 + P_3 [1 - \pi(\mathbf{x}_i)]]$$

$n_{i1}$  ist Anzahl der Respondenten die dem Item zustimmen  
mit Vektor von Kovariaten  $\mathbf{x}_i$

$n_{i0}$  ist Anzahl der Respondenten die das Item ablehnen

$P_1$  ist die Wahrscheinlichkeit eine automatische "Ja"-Antwort zu geben

$P_2$  ist die Wahrscheinlichkeit eine automatische "Nein"-Antwort zu geben

$P_3 = 1 - P_1 - P_2$  ist die Wahrscheinlichkeit eine wahre Antwort zu geben

## 3 Typen von Kovariaten

- **Soziodemografie:** Geschlecht; Alter; Bildung; Erwerbsstatus; Familienstand; Ortgröße; Ost-West
- **Psychologische bzw. situative Variablen:** generalisiertes Vertrauen; Vertrauen in den Datenschutz; wahrgenommene Sensitivität der Fragen; Anwesenheit dritter Personen
- **Interviewerbeobachtungen:** Antwortbereitschaft des Befragten; Empörung

Multivariate Logistische Regressionen für RRT-Daten: Odds-Ratios (Standardfehler in Klammern) – Abhängige Variable: Zustimmung zu einem sensitiven Item (1 = Ja )

	RRT (Ref. Direkt)	RRT (Ref. RRT-Direkt)	RRT-Direkt (Ref. Direkt)
Fremdenfeindl. 1	1.144 (0.151)	0.928 (0.180)	1.232 (0.232)
Fremdenfeindl. 2	<b>1.494***</b> <b>(0.204)</b>	<b>1.474**</b> <b>(0.288)</b>	1.014 (0.194)
Fremdenfeindl. 3	1.017 (0.141)	0.830 (0.163)	1.224 (0.233)
Antisemitismus 1	<b>1.737***</b> <b>(0.315)</b>	1.286 (0.305)	1.350 (0.319)
Antisemitismus 2	<b>1.714***</b> <b>(0.237)</b>	<b>1.519*</b> <b>(0.341)</b>	1.129 (0.254)
* $p < 0.1$ , ** $p < 0.05$ , *** $p < 0.01$ (zweiseitiger Test); Drittvariablen nicht berichtet			

# Empirische Ergebnisse

- Die Chancen eine sozial unerwünschte Einstellung zu offenbaren sind in der RRT Gruppe am höchsten.
- 5 aus 10 Chancenverhältnissen zeigen einen signifikanten Effekt (Referenz: Direkt bzw. RRT-Direkt)
- Keine signifikanten Unterschiede zwischen Direkt und RRT-Direkt nach Kontrolle von konfundierenden Größen
- Mit zunehmender Sensitivität des Themas (Stichwort: Item-Nonresponse), verstärkt sich die Wirkung der Befragungsmethodologie auf das Antwortverhalten.

# Schlussfolgerungen und Ausblick

- Es wurde ein praktikables RRT-Design für CATI-Surveys entwickelt und experimentell evaluiert.
- Vergleich von RRT mit konventionellen CATI Interviews zeigt, dass durch den Einsatz der RRT die Validität der Messung sozial unerwünschter Einstellungen erhöht wird.
- Ausblick: weitere Experimente, insbesondere zu Moduseffekten

# Literatur

- Bergmann, W. (1997) Antisemitism and Xenophobia in Germany since Unification, Pp. 21-38 in *Antisemitism and Xenophobia in Germany after Unification*, edited by H. Kurthen, W. Bergmann, and R. Erb. New York: Oxford University Press.
- Boruch R.F. (1971) Assuring confidentiality of responses in social research: a systematic analysis, *The American Psychologist*, 26, 413-430.
- Fowler, F.J.Jr. (1995) *Improving Survey Questions: Design and Evaluation*. Thousand Oaks, CA: Sage.
- Jann, B. (2005) *RRLOGIT: Stata Module to Estimate Logistic Regression for Randomized Response Data*, Boston: Boston College Department of Economics.
- Maddala, G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.
- McAuliffe W.E., Breer P., Ahmadifar N.W., Spino C. (1991) Assessment of drug abuser treatment needs in Rhode Island, *American Journal of Public Health*, 81, 365-371.

# Literatur

- Stem D.E., Steinhorst R.K. (1984) Telephone interview and mail questionnaire applications of the randomized response model, *Journal of the American Statistical Association*, 79, 555-564.
- Tourangeau R., Yan T. (2007) Sensitive questions in surveys, *Psychological Bulletin*, 133, 859-883.
- van der Heijden, P. G. M./van Gils, G./Bouts, J./Hox, J. J. (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.
- Warner S.L. (1965) Randomized response: a survey technique for eliminating evasive answer bias, *Journal of the American Statistical Association*, 60, 63-69.
- Weissman A.N., Steer R.A., Lipton D.S. (1986) Estimating illicit drug use through telephone interviews and the randomized response technique, *Drug and Alcohol Dependence*, 18, 225-233.