

# gesis

Leibniz-Institut  
für Sozialwissenschaften



## Die Datenqualität von Proxy-Interviews

### Inkonsistente Bildungsangaben im Mikrozensus-Panel

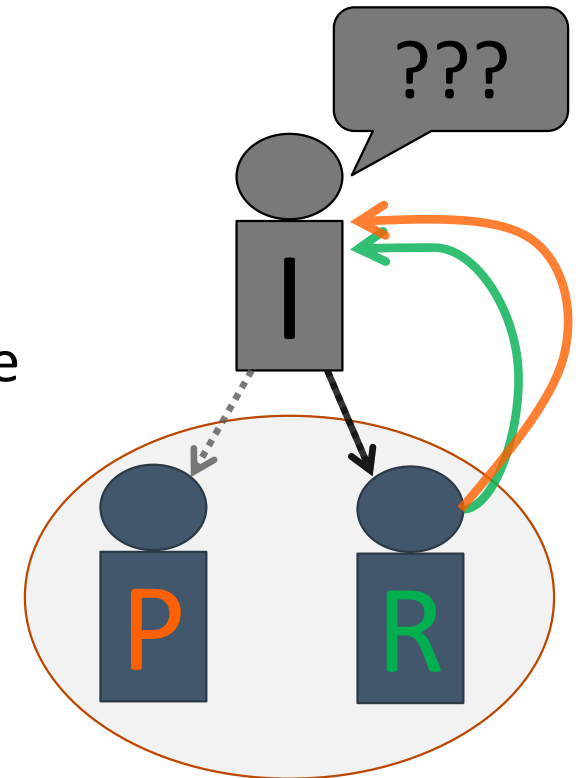
*Simon Börlin, 28. November 2018*

Vortrag auf der **9. Nutzerkonferenz „Forschen mit dem Mikrozensus“: Analysen zur Sozialstruktur und zum sozialen Wandel**  
vom 27. bis 28.11.2018 in Mannheim

# Übersicht

- 1) Einführung & Fragestellung
- 2) Theoretische Überlegungen
- 3) Analysestrategie
- 4) Daten: Mikrozensus-Panel ab 2012
- 5) Ergebnisse
- 6) Fazit

- Was sind Proxy-Interviews?
  - ▶ Zielperson (P) nimmt nicht direkt an Umfrage teil
  - ▶ Eine dritte Person (R) erteilt Auskünfte
- Vorteile:
  - ▶ Erhöhung der Ausschöpfungsquote
  - ▶ Verringerung der Erhebungskosten
  - ▶ Verkürzung der Feldzeit
- Nachteile:
  - ▶ Verringerung der Datenqualität?  
(Moore 1988, Cobb 2018a, Cobb 2018b)



- Problematisch bei Untersuchung von Datenqualität:
  - ▶ Trennung von *Selection Effect & Measurement Effect* (bspw. Moore 1988; Stark 2006)
  - ▶ Lösungsmöglichkeiten:
    1. Erhebungen mit zufälliger Proxyauswahl (bspw. Lee, Mathiowetz & Tourangeau 2004)
    2. Paneldaten: Untersuchung (relativ) zeitkonstanter Merkmale über die Zeit (Test-Rest-Methode) (bspw. mit Mikrozensus-NRW: Zühlke 2008)
- Da meist keine externen Angaben vorliegen, Annahme: Selbstauskunft = „wahrer“ Wert (vgl. Moore 1988)

- Anwendung der Test-Retest-Methode
- Überprüfung der Datenqualität am Beispiel der Bildungsangaben:
  - ▶ Bildungsabschluss ist zentrale Variable in sozialwissenschaftlicher Forschung
  - ▶ Bildungsabschluss & Jahr des Abschlusses ist ab gewissem Alter relativ zeitkonstant
- **Fragestellung:**  
In welchem Ausmaß beeinflusst die Erhebungsart (Selbst- oder Fremdauskunft) inkonsistente Bildungsangaben?

## ■ Forschungsstand:

- ▶ Proxyangaben ≠ Selbstangaben für Bildungsangaben
  - Telefonische Nachbefragung
    - Eltern > PartnerIn > Kinder (Dawe & Knight 1997)
  - Kinder über Eltern (Kreuter et al. 2010)
  - Mikrozensus-NRW: Unterschiede, doch geringes Ausmaß (Zühlke 2008)
- ▶ Metaanalyse zu Proxyangaben:  
Methodik & thematische Bandbreite der Untersuchungen ausbaufähig (Cobb 2018b)

- **4 kognitive Komponenten bei der Generierung einer Antwort** (Tourangeau, Rips & Rasinski 2000; Lee, Mathiowetz & Tourangeau 2004):
  - 1. Comprehension:** Verstehen & Interpretation der Frage
    - ▶ Mikrozensus: Personen beantworten Frage für sich selbst und andere
  - 2. Retrieval:** Abruf relevanter Informationen aus Erinnerung –
    - ▶ Zugriff auf Information (1. Hand, eigene Erfahrungen vs. Proxy)
    - ▶ fehlende oder unsichere Info
    - ▶ Anker-Heuristiken (Tversky & Kahneman 1974; Bickart et al. 1994)
    - ▶ Motivationsunterschiede zwischen Selbst- & Fremdauskunft?
  - 3. Judgment:** Weniger Information → zeitstabilere Info → gleichbleibendere Meinungsbildung (≠ akkurat) (Schwarz & Wellens 1997)
    - ▶ Für abgeschlossene Bildungsabschlüsse keine Unterschiede erwartbar
  - 4. Response:** Angaben über Dritte (Soziale Erwünschtheit)

# Hypothesen

- H1: Item-Nonresponse tritt bei Proxyangaben häufiger auf
- H2: Inkonsistente Bildungsangaben treten bei Proxyangaben häufiger auf

Konstellation	t <sub>0</sub>	t <sub>1</sub>	
1	Selbst	Selbst	Weniger Abweichungen (reliabel)
2	Selbst	Proxy	Mehr Abweichungen (nicht reliabel)
3	Proxy	Proxy	Mehr Abweichungen? (reliabel?)
4	Proxy	Selbst	Mehr Abweichungen (nicht reliabel)



- Test-Retest-Methode
- Inkonsistente Bildungsangaben als konservatives Kriterium:
  - ▶ Relativ objektives Kriterium (vgl. subjektive Fragen)
  - ▶ Höchster Bildungsabschluss sollte sich generell nicht ändern (ab bestimmten Alter) → Wechsel zwischen zwei Jahren allgemein unwahrscheinlich
- Mögliche Bildungswechsel = Wertung als inkonsistent:
  - ▶ Kein Besuch einer Bildungseinrichtung im letzten Jahr
  - ▶ Altersabgrenzung (ab 20 Jahren)

# Messung inkonsistenter Bildungsabschlüsse:

		2013					
2012		0	1	3	4	5	9
0	Entfällt (Person ohne allgemeinbildenden Schulabschluss) / (6) Abschluss nach höchstens 7 Jahren Schulbesuch	++	-	-	-	-	0
1	Haupt-(Volks-)schulabschluss / (2) Polytechnische Oberschule der DDR mit Abschluss der 8. oder 9. Klasse	--	++	-	-	-	0
3	Realschulabschluss (Mittlere Reife) oder gleichwertiger Abschluss / (7) Polytechnische Oberschule der DDR mit Abschluss der 10. Klasse	--	--	++	-	-	0
4	Fachhochschulreife	--	--	--	++	-	0
5	Allgemeine oder fachgebundene Hochschulreife (Abitur)	--	--	--	--	++	0
9	Keine Angabe	0	0	0	0	0	00

Tabelle 1: Vergleich der Bildungsabschlüsse zwischen den Jahren 2012 und 2013

Legende:

- ++ = gleicher Abschluss in beiden Jahren
- = möglicher Wechsel der Abschlüsse
- = niedrigere/unmögliche Kombination von Abschlüssen
- 0/00 = Keine Angabe (Item-Nonresponse)

Proxy-Interviews im Mikrozensus:

180 Falls Sie 15 Jahre oder älter sind:

Haben Sie die Fragen zu Ihrer Person selbst beantwortet ?

**i** Als „selbst beantwortet“ gilt auch, wenn Sie die Angaben zu Ihrer Person auf Richtigkeit überprüft und ggf. selbst korrigiert haben.

Ja, selbst beantwortet ..... 1

Nein, durch eine andere Person vertreten ..... 2

Keine Angabe ..... 9

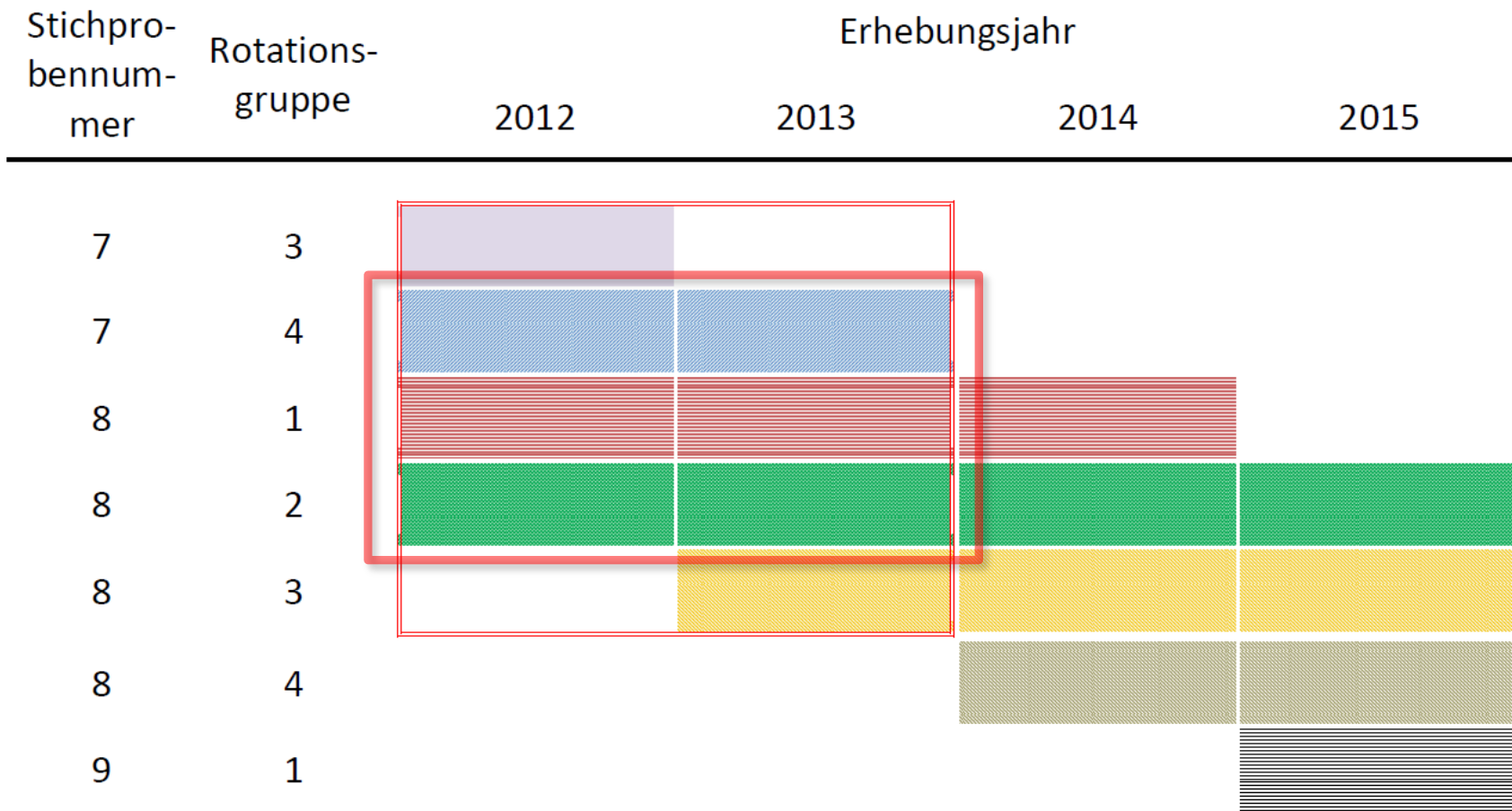
MZ 2012	
69 %	1
24 %	2
7 %	9

freiwillig

Entnommen aus: Erhebungsbogen Mikrozensus 2012 (Frage 180);

Quelle der Berechnungen: FDZ der Statistischen Ämter des Bundes und der Länder, Mikrozensus SUF 2012, eigene Berechnungen.

- Mikrozensus SUF 2012 & 2013
- Zusammenführung des Rotationspanels aus Querschnitten (Herter-Eschweiler & Schimpl-Neimanns 2018)



Entnommen aus: Herter-Eschweiler & Schimpl-Neimanns 2018, Abbildung 1, S.2

- Mikrozensus SUF 2012 & 2013
- Zusammenführung des Rotationspanels aus Querschnitten (Herter-Eschweiler & Schimpl-Neimanns 2018)
- Ausschluss für Analyse:
  - ▶ Fehlerhaftes Merging (Geschlechtswechsel; Geburtsjahrwechsel>6; Geschlechts- & Geburtsjahrwechsel>3)
  - ▶ Missings
  - ▶ Nur Privathaushalte berücksichtigt
  - ▶ Personen auf allgemeinbildenden Schulen
  - ▶ Alter <20 Jahre
- $N \approx 250.000$  Personen

- Anteil an Item-Nonresponse ausgewählter Bildungsangaben:

Variable	2012		2013	
	Selbst	Proxy	Selbst	Proxy
Bildungsabschluss	0,2	0,3	0,1	0,2
Jahr des Ausbildungsabschlusses	3,0	8,1	2,7	8,2
Jahr des Bildungsabschlusses	8,4	11,4	6,5	11,0

Dargestellt in Prozent; Unterschiede zwischen Selbst- und Proxyauskunft immer signifikant (Pr = 0,000).  
Quelle: FDZ der Statistischen Ämter des Bundes und der Länder, Mikrozensus SUF, 2012-2013, eigene Berechnungen.

- Konstellationen der Erhebungsarten:

Konstellation	t <sub>0</sub>	t <sub>1</sub>	%	Gültige %
1	Selbst	Selbst	62	70
2	Selbst	Proxy	8	9
3	Proxy	Proxy	12	13
4	Proxy	Selbst	7	8
Keine Angabe	--	--	11	

Quelle: FDZ der Statistischen Ämter des Bundes und der Länder, Mikrozensus SUF, 2012-2013, eigene Berechnungen.



Bildungsabschlüsse zwischen den beiden Jahren im Vergleich

2013

2012

0

1

3

4

5

Total

0	Entfällt (Person ohne allgemeinbildenden Schulabschluss) / (6) Abschluss nach höchstens 7 Jahren Schulbesuch	64,86	28,32	4,39	0,59	1,85	100 (7.842)
1	Haupt-(Volks-)schulabschluss / (2) Polytechnische Oberschule der DDR mit Abschluss der 8. oder 9. Klasse	2,11	90,77	6,02	0,64	0,47	100 (100.993)
3	Realschulabschluss (Mittlere Reife) oder gleichwertiger Abschluss / (7) Polytechnische Oberschule der DDR mit Abschluss der 10. Klasse	0,38	7,95	86,36	3,49	1,82	100 (72.236)
4	Fachhochschulreife	0,22	3,43	14,66	66,06	15,62	100 (17.032)
5	Allgemeine oder fachgebundene Hochschulreife (Abitur)	0,20	0,91	2,42	5,37	91,10	100 (50.859)
<b>Total</b>		3,06 (7.630)	40,44 (100.684)	29,13 (72.530)	6,91 (17.194)	20,45 (50.924)	100 (248.962)

Anmerkung: Dargestellt sind Zeilenprozente mit absoluten Zahlen in Klammern.

Quelle: FDZ der Statistischen Ämter des Bundes und der Länder, Mikrozensus SUF, 2012-2013, eigene Berechnungen.

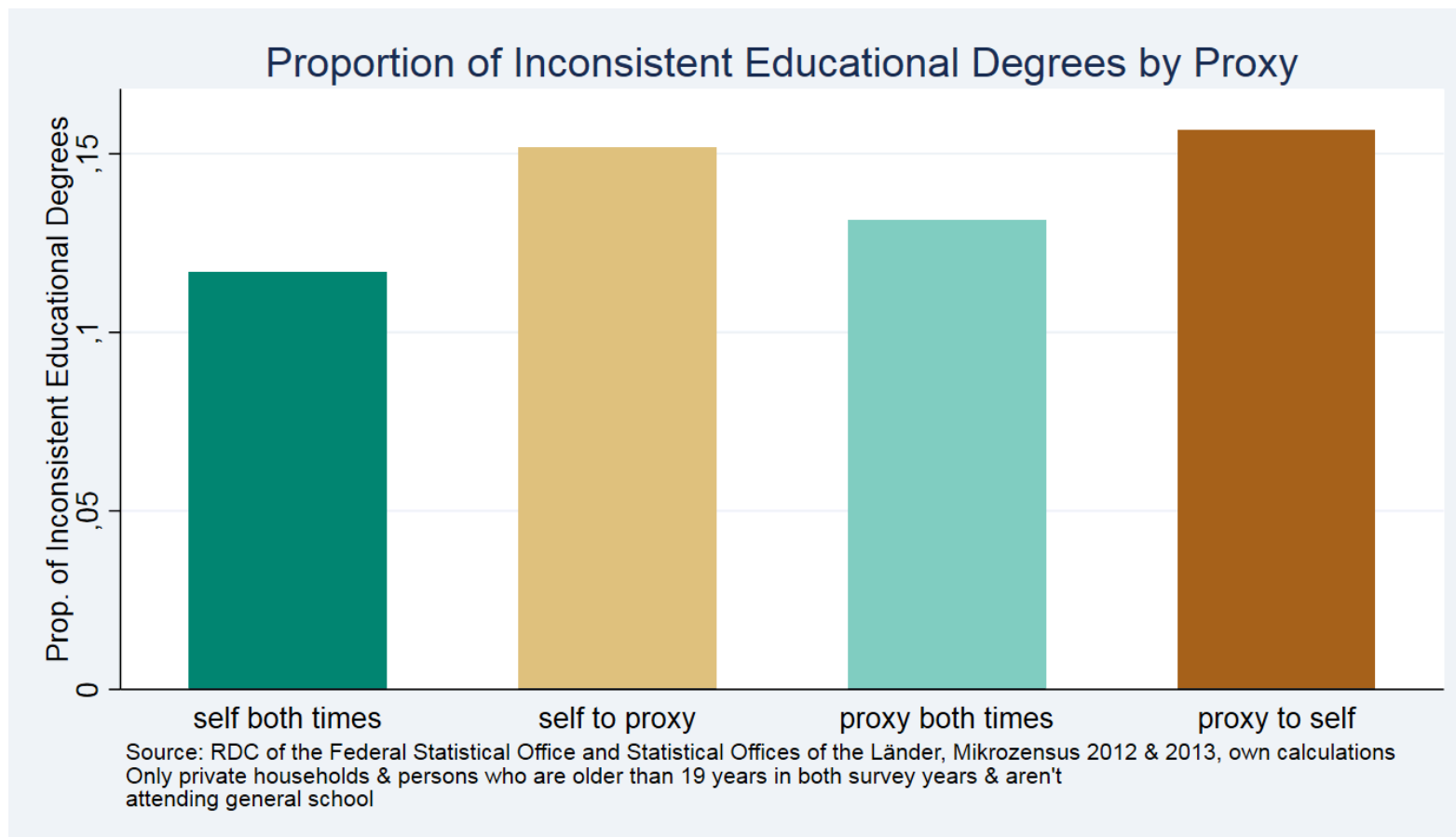
≈ 13% inkonsistente Bildungsabschlüsse

Cohens Kappa für **Wechsel des Bildungsabschlusses**, unterteilt nach Proxy:

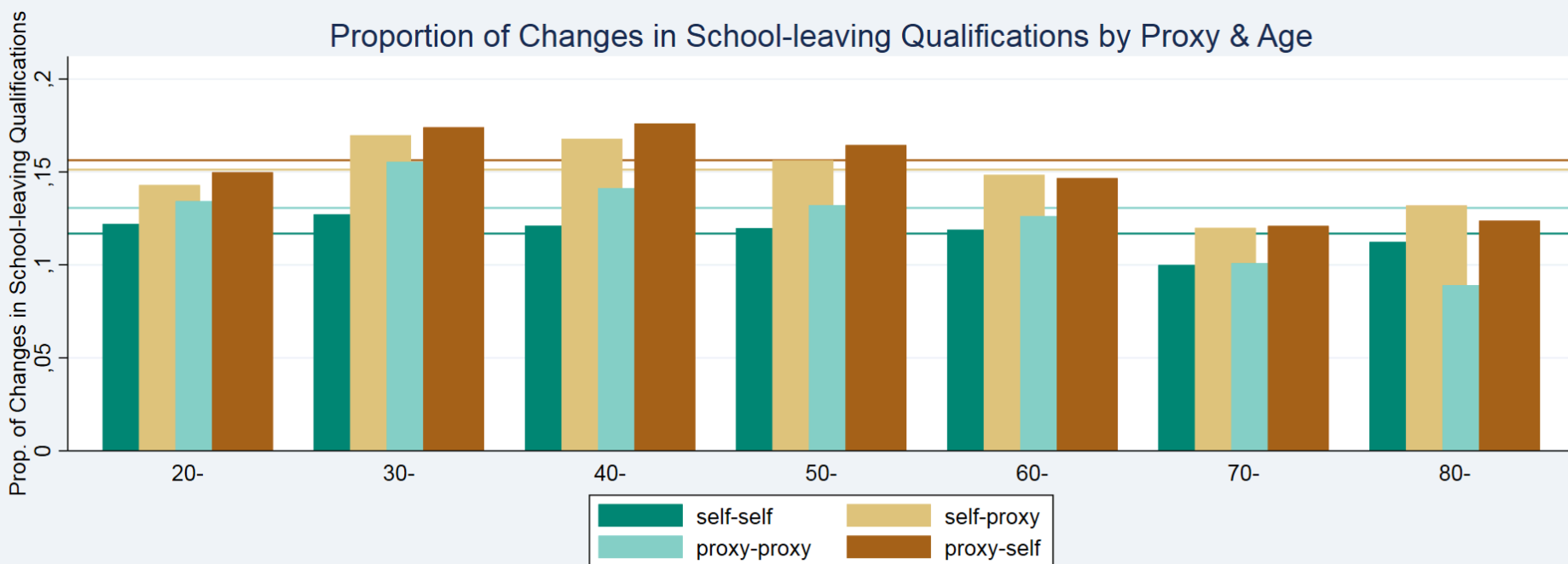
	Prozentuale Übereinstimmung	Cohens Kappa
<b>1. Selbst/Selbst</b>	0,8826	0,8306
<b>2. Selbst/Proxy</b>	0,8470	0,7799
<b>3. Proxy/Proxy</b>	0,8667	0,8140
<b>4. Proxy/Selbst</b>	0,8419	0,7748
<b>2. &amp; 4. Selbst/Proxy oder Proxy/Selbst</b>	0,8446	0,7775
<b>Gesamt</b>	<b>0,8705</b>	<b>0,8161</b>

Quelle: FDZ der Statistischen Ämter des Bundes und der Länder, Mikrozensus SUF, 2012-2013, eigene Berechnungen.

# Anteil inkonsistenter Bildungsabschlüsse: (unterteilt nach Proxy-Konstellation)



# Anteil inkonsistenter Bildungsabschlüsse: (unterteilt nach Proxy-Konstellation & Altersgruppen)



Source: RDC of the Federal Statistical Office and Statistical Offices of the Länder, Mikrozensus 2012 & 2013, own calculations  
Only private households & persons who are older than 19 years & aren't attending general school

- Zusammenfassung:
  - ▶ Item-Nonresponse steigt bei Proxyangaben
  - ▶ Inkonsistente Bildungswechsel treten häufiger bei Wechsel der Erhebungsart auf
  - ▶ Selbstauskünfte zu beiden Zeitpunkten sind am zuverlässigsten
- Ausblick:
  - ▶ Erweiterung des Themenspektrums nötig (bspw. Ausbildungsabschlüsse)
  - ▶ Identifikation des Proxy-Respondenten: Beziehungsart
- MZ-Panel:
  - ▶ Kennung der Rotationsgruppen für sichere Ermittlung der Ausfallgründe (Panelmortalität)

Vielen Dank für Ihre Aufmerksamkeit!

Die Datenqualität von Proxy-Interviews:  
Inkonsistente Bildungsangaben im  
Mikrozensus-Panel

*Simon Börlin, 28. November 2018*

[Simon.Boerlin@gesis.org](mailto:Simon.Boerlin@gesis.org)

- **Bickart**, B., Geeta **Menon**, Norbert **Schwarz** and Johnny **Blair** (1994). The Use of Anchoring Strategies in the Construction of Proxy-Reports of Attitudes. *International Journal of Public Opinion Research*, 6, 375-379.
- **Cobb**, C. (2018a). Answering for Someone Else: Proxy Reports in Survey Research. In David L. Vannette and Jon A. Krosnick (Eds.). *The Palgrave Handbook of Survey Research*, pp. 87–93.
- **Cobb**, C. (2018b). Proxy Reporting. In David L. Vannette and Jon A. Krosnick (Eds.). *The Palgrave Handbook of Survey Research*, pp. 427–437.
- **Dawe**, F. and **Knight**, I. (1997). A Study of Proxy Response on the Labour Force Survey. *Survey Methodology Bulletin*, 40, 30–36.
- **Herter-Eschweiler**, R. and **Schimpl-Neimanns**, B. (2018). *Möglichkeiten der Verknüpfung von Mikrozensus-Querschnitterhebungen ab 2012 zu Panels*.
- **Kreuter**, F., **Eckman**, S., **Maaz**, K. and **Watermann**, R. (2010). Children’s Reports of Parents’ Education Level. Does it Matter Whom You Ask and What You Ask About? *Survey Research Methods*, 4, 127–138.
- **Lee**, S., **Mathiowetz**, N. A. and **Tourangeau**, R. (2004). Perceptions of Disability: The Effect of Self- and Proxy Response. *Journal of Official Statistics*, 20, 671–686.
- **Moore**, J. C. (1988). Self/Proxy Response Status and Survey Response Quality. A Review of the Literature. *Journal of Official Statistics*, 4, 155–172.
- **Schwarz**, N. and **Wellens**, T. (1997). Cognitive Dynamics of Proxy Responding: The Diverging Perspective of Actors and Observers. *Journal of Official Statistics*, 13, 159–179.
- **Statistisches Bundesamt (Destatis)** (2013). *Qualitätsbericht Mikrozensus 2012*. Wiesbaden, Germany.
- **Tourangeau**, R., **Rips**, L. J. and **Rasinski**, K. (2000). *The Psychology of Survey Response*: Cambridge University Press.
- **Tversky**, A. and **Kahneman**, D. (1974). Judgment under Uncertainty. Heuristics and Biases. *Science (New York, N.Y.)*, 185, 1124–1131.
- **Wolf**, C. and **Lüttinger**, P. (2009). Verteilung von Proxy-Interviews im deutschen Mikrozensus. In Weichbold, M., Bacher, J. and Wolf, C. (Eds.). *Umfrageforschung. Herausforderungen und Grenzen*. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden, 395-412.
- **Zühlke**, S. (2008). Auswirkungen von Proxy-Interviews auf die Datenqualität des Mikrozensus. *Statistische Analysen und Studien NRW*, Band 53, 3–10.

# Anhang



Table: Prop. of Proxy Reports & Types of Respondents by Household Size

	Household Size													
	1		2		3		4		5		6 or more		Total	
	Obs	Col %	Obs	Col %	Obs	Col %	Obs	Col %	Obs	Col %	Obs	Col %	Obs	Col %
<b>Proxy in 2012</b>														
<b>0 Self Report</b>	49,808	<b>94.0</b>	80,293	<b>79.9</b>	27,905	<b>65.8</b>	20,765	<b>58.9</b>	5,847	<b>54.5</b>	1,845	<b>48.5</b>	186,463	<b>75.9</b>
<b>1 Proxy Report</b>	3,174	<b>6.0</b>	20,213	<b>20.1</b>	14,475	<b>34.2</b>	14,463	<b>41.1</b>	4,883	<b>45.5</b>	1,957	<b>51.5</b>	59,165	<b>24.1</b>
<b>Total</b>	52,982	<b>100.0</b>	100,506	<b>100.0</b>	42,380	<b>100.0</b>	35,228	<b>100.0</b>	10,730	<b>100.0</b>	3,802	<b>100.0</b>	245,628	<b>100.0</b>
<b>Type of Respondent</b>														
<b>0 Proxy Report (known Resp.)</b>	0	<b>0.0</b>	18,217	<b>18.1</b>	10,429	<b>24.6</b>	9,772	<b>27.7</b>	2,778	<b>25.9</b>	984	<b>25.9</b>	42,180	<b>17.2</b>
<b>1 Self Report (only)</b>	49,808	<b>94.0</b>	59,993	<b>59.7</b>	6,762	<b>16.0</b>	2,124	<b>6.0</b>	361	<b>3.4</b>	100	<b>2.6</b>	119,148	<b>48.5</b>
<b>2 Respondent for Self &amp; Proxy</b>	0	<b>0.0</b>	20,300	<b>20.2</b>	7,746	<b>18.3</b>	6,075	<b>17.2</b>	1,455	<b>13.6</b>	419	<b>11.0</b>	35,995	<b>14.7</b>
<b>3 Proxy Report (unknown Resp.)</b>	3,174	<b>6.0</b>	1,996	<b>2.0</b>	4,046	<b>9.5</b>	4,691	<b>13.3</b>	2,105	<b>19.6</b>	973	<b>25.6</b>	16,985	<b>6.9</b>
<b>4 Self Report (unknown if add. Proxy Resp.)</b>	0	<b>0.0</b>	0	<b>0.0</b>	13,397	<b>31.6</b>	12,566	<b>35.7</b>	4,031	<b>37.6</b>	1,326	<b>34.9</b>	31,320	<b>12.8</b>
<b>Total</b>	52,982	<b>100.0</b>	100,506	<b>100.0</b>	42,380	<b>100.0</b>	35,228	<b>100.0</b>	10,730	<b>100.0</b>	3,802	<b>100.0</b>	245,628	<b>100.0</b>

Source: RDC of the Federal Statistical Office and Statistical Offices of the Laender, Mikrozensus 2012; own calculations. Only private households and persons older than 14 years.