

# gesis

Leibniz-Institut  
für Sozialwissenschaften

German Longitudinal  
Election Study



## **GLES**

### **Langfrist-Online-Tracking**

### **T28, 2015**

ZA5728, Version 2.1.0

*Studienbeschreibung*

## Inhaltsverzeichnis

Vorbemerkungen .....	3
1. Studieninformationen .....	4
2. Konzeption .....	5
3. Stichprobenziehung.....	8
4. Erhebung .....	11
5. Datenaufbereitung.....	17
6. Hinweise und Anmerkungen.....	31
Literaturverzeichnis.....	33

## 1. Vorbemerkungen

### Arbeiten mit GLES Daten

Bei dieser Publikation und dem zugehörigen Datensatz handelt es sich um Daten der German Longitudinal Election Study (GLES), die von GESIS in Kooperation mit der Deutschen Gesellschaft für Wahlforschung (DGfW) herausgegeben werden. Vor der Veröffentlichung werden die Daten sorgfältig geprüft. Leider kann es trotz gründlicher Überprüfung der Daten passieren, dass Fehler in den Datensätzen unentdeckt bleiben. Fallen diese zu einem späteren Zeitpunkt auf, werden sie dokumentiert (Errata-Liste im Datenbestandskatalog, [www.gesis.org/dbk](http://www.gesis.org/dbk)) und zeitnah behoben.

Um hochqualitative Datensätze zur Verfügung stellen zu können, freuen wir uns über Ihre Mithilfe. Wenn Ihnen bei Ihrer Arbeit mit den GLES-Daten ein Fehler auffällt, helfen Sie uns sehr, wenn Sie eine kurze E-Mail an [gles@gesis.org](mailto:gles@gesis.org) schicken. Bitte schicken Sie uns zusätzlich zu der Beschreibung des Fehlers auch die Studiennummer (ZA-Nummer) sowie die Versionsnummer des Datensatzes.

Wir empfehlen, stets mit der aktuellen Version der GLES-Daten zu arbeiten. Sie können diese unkompliziert über den Datenbestandskatalog herunterladen. Die Links zum direkten Download finden Sie auch auf den Seiten der GLES bei GESIS ([www.gesis.org/gles](http://www.gesis.org/gles)).

### Meldung von Veröffentlichungen

Um einen Überblick über die tatsächliche Nutzung der Daten zu erhalten, bitten wir Sie um eine kurze Mitteilung bei Veröffentlichungen, die Daten der GLES verwenden (bibliographische Angaben, Studiennummer des verwendeten Datensatzes). Veröffentlichungen, die vollständig oder teilweise auf Daten der GLES beruhen, werden in der offiziellen Bibliographie der GLES aufgeführt. Wenn es sich dabei um Konferenzpapiere o.ä. handelt, die nur schwer zugänglich sind, freuen wir uns über die Überlassung eines Exemplars bzw. eines PDF-Dokuments.

### Kontakt

GESIS – Leibniz-Institut für Sozialwissenschaften

Postfach 122155

68072 Mannheim

E-Mail: [gles@gesis.org](mailto:gles@gesis.org)

### Zitation bei Veröffentlichungen

Wird in Publikationen auf Daten der GLES zurückgegriffen, bitten wir diese wie folgt zu zitieren:

Roßteutscher, Sigrid; Schmitt-Beck, Rüdiger; Schoen, Harald; Weißels, Bernhard; Wolf, Christof; Henckel, Simon; Bieber, Ina; Scherer, Philipp (2016): Langfrist-Online-Tracking T28 (GLES). GESIS Datenarchiv, Köln. ZA5728 Datenfile Version 2.1.0, doi: 10.4232/1.12521.

## **2. Studieninformationen**

### **2.1. Studiennummer**

ZA5728

### **2.2. Titel der Studie**

German Longitudinal Election Study, Komponente 8, Langfrist-Online-Tracking, T28: Europa, Institutionenvertrauen und Positionssues.

### **2.3. Version**

2.1.0, 19.04.2016, doi: 10.4232/1.12521

### **2.4. Erhebungszeitraum**

05.06.2015 – 19.06.2015

### **2.5. Primärforscher/innen**

Prof. Dr. Sigrid Roßteutscher	(Universität Frankfurt)
Prof. Dr. Rüdiger Schmitt-Beck	(Universität Mannheim)
Prof. Dr. Harald Schoen	(Universität Mannheim)
Prof. Dr. Bernhard Weßels	(Wissenschaftszentrum Berlin für Sozialforschung)
Prof. Dr. Christof Wolf	(GESIS – Leibniz-Institut für Sozialwissenschaften)

### **2.6. Finanzierende Stelle**

Deutsche Forschungsgemeinschaft e.V. (DFG)

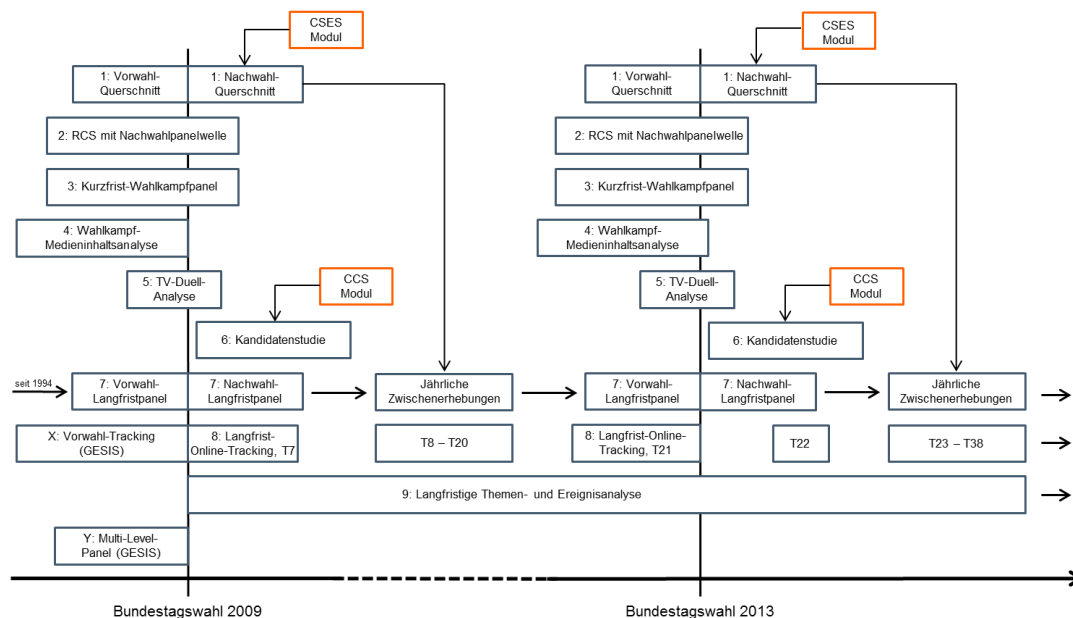
### 3. Konzeption

#### 1.1. Hintergrund und Zielsetzung

Die German Longitudinal Election Study (GLES) ist die bislang größte nationale Wahlstudie in Deutschland. In dem von der DFG geförderten Projekt sollen die politischen Prädispositionen und Einstellungen sowie das politische Verhalten der wahlberechtigten Bürger bei vorerst drei aufeinanderfolgenden Wahlen beobachtet und analysiert werden. Langfristig wird angestrebt, das mit der Bundestagswahl 2009 gestartete Projekt auch nach der Wahl 2017 als institutionalisierte Wahlstudie innerhalb von GESIS weiterzuführen.

Die GLES erlaubt die Analyse des Wahlverhaltens in quer- als auch längsschnittlicher Perspektive sowie in Hinblick auf kurzfristige Dynamiken während des Wahlkampfs und auch langfristige soziale Wandlungsprozesse über den Wahlzyklus hinweg (Schmitt-Beck et al. 2010).

**Abbildung 1: Das Design der German Longitudinal Election Study (GLES)**



Das Langfrist-Online-Tracking (Komponente 8) bildet gemeinsam mit dem Langfristpanel (Komponente 7) und der langfristigen Themen- und Ereignisanalyse (Komponente 9) den auf die Analyse langfristiger Prozesse der Formierung und des Wandels der öffentlichen Meinung ausgerichteten Bestandteil der GLES (siehe Abbildung 1). Das Langfrist-Online-Tracking zielt hierbei insbesondere auf die Beobachtung der öffentlichen Meinung und ihrer Veränderungen zwischen den Bundestagswahlen ab (Schmitt-Beck et al. 2010). Hierzu werden in dreimonatlichen Abständen webbasierte Querschnitte mit jeweils etwa 1.000 Befragten aus einem Online-Access-Panel erhoben. Die Fragebögen des Langfrist-Online-Trackings umfassen neben einem konstanten Kernfragebogen rotierende thematische Module sowie situationsbezogen variierenden Frageblöcke. Mit Hilfe des Langfrist-Online-Trackings lassen sich die Einstellungen der Wähler zu den wichtigsten Issues sowie zu den Parteien und ihrem Spitzenpersonal, aber auch zu den Leistungen der Bundesregierung und der Opposition über die Bundestagswahlzyklen hinweg analysieren (Schmitt-Beck et al. 2010).

### 3.7. Aufbau der Studie

Die Erhebung der Langfrist-Online-Trackings begann mit dem Vorlauf zur Bundestagswahl 2009. Seitdem werden im Rahmen des Langfrist-Online-Trackings Studien mit unterschiedlichen thematischen Schwerpunktsetzungen durchgeführt. Eine Übersicht über die bisher durchgeführten Online-Trackings bietet die nachfolgende Übersicht.

**Tabelle 1: Die Langfrist-Online-Trackings der GLES**

Erhebung	Studiennummer	Schwerpunkt	Feldbeginn	Feldende
T1	ZA5334	Grundlegende Tests	30.04.2009	05.05.2009
T2	ZA5335	Wählen auf mehreren Ebenen	27.05.2009	05.06.2009
T3	ZA5336	Wirtschaftliche Lage	03.07.2009	13.07.2009
T4	ZA5337	Koalitionen	31.07.2009	11.08.2009
T5	ZA5338	Skalen- und Reihenfolgeexperimente	24.08.2009	01.09.2009
T6	ZA5339	Wahlkampf	18.09.2009	27.09.2009
T7	ZA5340	Nachwahl	29.09.2009	08.10.2009
T7Exp	ZA5340	Experiment zur Bundestagswahl	08.10.2009	25.10.2009
T8	ZA5341	Nachwahl und Netzwerke	10.12.2009	20.12.2009
T9	ZA5342	Wirtschaftliche Lage	15.04.2010	23.04.2010
T10	ZA5343	Wählen auf mehreren Ebenen	24.06.2010	05.07.2010
T11	ZA5344	Psychologische Konstrukte	16.09.2010	26.09.2010
T12	ZA5345	Positionsissues	09.12.2010	19.12.2010
T12NB	ZA5345	Nachbefragung von Abbrechern	20.12.2010	30.12.2010
T13	ZA5346	Wirtschaftliche Lage	09.03.2011	19.03.2011
T13NB	ZA5346	Nachbefragung von Abbrechern	21.03.2011	30.03.2011
T14	ZA5347	Wählen auf mehreren Ebenen	23.05.2011	03.06.2011
T14NB	ZA5347	Nachbefragung von Abbrechern	03.06.2011	13.06.2011
T15	ZA5348	Psychologische Konstrukte	24.08.2011	03.09.2011
T15NB	ZA5348	Nachbefragung von Abbrechern	05.09.2011	14.09.2011
T16	ZA5349	Positionsissues	08.12.2011	18.12.2011
T17	ZA5350	Wirtschaftliche Lage und Wählen auf mehreren Ebenen	02.05.2012	15.05.2012
T18	ZA5351	Netzwerke, psychologische Konstrukte und Koalitionen	17.09.2012	01.10.2012
T19	ZA5719	EU-Krise und politische Partizipation	04.01.2013	19.01.2013
T20	ZA5720	Positionsissues und Europa	24.05.2013	08.06.2013
T21	ZA5721	Wahlkampf	06.09.2013	21.09.2013
T22	ZA5722	Koalitionen und psychologische Konstrukte	29.11.2013	13.12.2014
T23	ZA5723	Wirtschaftliche Lage und politisches Wissen	21.02.2014	07.03.2014
T24	ZA5724	Europa und Europawahl	09.05.2014	23.05.2014

---

T25	ZA5725	Psychologische Konstrukte	29.08.2014	13.09.2014
T26	ZA5726	Positionsissues, Netzwerke und Koalitionen	21.11.2014	05.12.2014
T27	ZA5727	Psychologische Konstrukte, Wirtschaftliche Lage, Politisches Wissen, Partnerschaft	27.02.2015	13.03.2015
T28	ZA5728	Europa, Institutionenvertrauen, Position-sissues	05.06.2015	19.06.2015

---

### 3.8. Inhalt der Studie

Neben den konstanten Kernfragen und sozio-demografischen Fragen umfasst diese Erhebung des Langfrist-Online-Trackings Fragemodule zu den Themen Europa, Institutionenvertrauen und Positionsissues.

## 4. Stichprobenziehung

### 4.1. Untersuchungsgebiet

Deutschland (DE)

### 4.2. Grund- und Auswahlgesamtheit

Die Grundgesamtheit der Studie bildet die zum Zeitpunkt der Erhebung zur Wahl des Deutschen Bundestags wahlberechtigte Bevölkerung der Bundesrepublik Deutschland. Da die Studie webbasiert durchgeführt wird, haben nicht alle wahlberechtigten Bürger eine Chance, für die Befragung des Online-Trackings ausgewählt zu werden. Die Auswahlgesamtheit umfasst daher ausschließlich die ab 18-jährigen Mitglieder mit deutscher Staatsangehörigkeit des LINK Internet Panel, die in Deutschland leben und mindestens einmal in der Woche zu privaten Zwecken das Internet nutzen. Diese Personengruppe umfasst im LINK Internet Panel zum Zeitpunkt der Erhebung nach Angaben des Instituts etwa 40.000 aktive Panelmitglieder. Als aktive Panelmitglieder werden dabei Personen definiert, die nach der telefonischen Rekrutierung die Stammdatenbefragung (Double-opt-in-Registrierung) ausgefüllt und in den letzten zwölf Monaten an mindestens einer Umfrage teilgenommen haben.

**Tabelle 2: Zusammensetzung des LINK Internet Panels nach sozio-demographischen Merkmalen zum Zeitpunkt der Erhebung**

Merkmal	Anteil in Prozent
<b>Geschlecht</b>	
Weiblich	49,0
Männlich	51,0
<b>Bildung<sup>1</sup></b>	
Niedrig (kein Abschluss, Haupt- und Volksschule)	13,0
Mittel (Realschule, Mittlere Reife)	33,0
Hoch (Abitur, Fachhochschulreife, Studium)	54,0
<b>Altersgruppen</b>	
18-29 Jahre	22,0
30-39 Jahre	22,0
40-49 Jahre	26,0
50-59 Jahre	17,0
60 Jahre und älter	13,0

<sup>1</sup> 17% der Panelmitglieder haben bei der Frage nach ihrer formalen Bildung von der Möglichkeit, keine Angabe zu machen Gebrauch gemacht. Laut LINK kann aufgrund der Ergebnisse anderer Studien davon ausgegangen werden, dass diese Gruppe überproportional viele formal niedrig gebildete Personen umfasst.



Das LINK Internet Panel wird zu 100% aktiv in computergestützten telefonischen Umfragen (CATI) rekrutiert.<sup>2</sup> Laut LINK entfielen somit systematische Verzerrungen durch Selbstselektionsprozesse. Von besonderer Bedeutung sei demnach die weitgehende Reduktion von Stichprobenverzerrenden Einflüssen, die sich bei einer Selbstselektion zum einen aus dem vermeintlich sehr hohen Anteil von Internetnutzern mit einer hohen Nutzungsintensität – sogenannte „heavy Internet user“ - und zum anderen aus der Rekrutierung auf einer mehr oder weniger eingeschränkten Anzahl von Websites ergeben könne.

Laut Angaben von LINK werden standardmäßig die nachfolgenden Methoden zur Panelpflege durchgeführt:

- Eine Zeit- und Antwortenkontrolle soll gewährleisten, dass Panelteilnehmer, die laut LINK „eindeutig falsche“ oder „widersprüchliche“ Antworten geben oder sich „zu schnell“ durch den Fragebogen durchklicken, nicht mehr zu weiteren Befragungen eingeladen werden.
- Bei der Kontrolle der „Stammdatenkonvergenz“ werde ein Abgleich zwischen den Angaben der Befragten in Umfragen - etwa zu Geschlecht und Alter - mit den bekannten Angaben aus der Stammdatenbefragung durchgeführt. Panelteilnehmer mit widersprüchlichen Angaben werden ebenfalls von weiteren Befragungsteilnahmen ausgeschlossen.
- Es erfolge eine Dokumentation der Teilnahmehistorie für alle Panelteilnehmer, so dass eine übermäßig häufige Befragung von individuellen Panelteilnehmern vermieden werden könne. Dies ist LINK zufolge gegeben, wenn jeder Panelteilnehmer maximal einmal pro Monat befragt werde. Zudem erlaube die Aufzeichnung der Teilnahmehistorie, dass Panelteilnehmer nicht mehrmals zu Umfragen zum gleichen Thema eingeladen werden.

Die Zeit- und Antwortenkontrolle sowie die Kontrolle der Stammdatenkonvergenz führe in Verbindung mit inaktiven Panelteilnehmern (Teilnehmer, die mehrfach nicht auf Einladungen reagieren) zu einem jährlichen Ausschluss von ca. 15% der Panelmitglieder. Weiterhin beenden laut LINK jährlich etwa 3-4% der Panelteilnehmer aktiv ihre Mitgliedschaft. Insgesamt liege der Panelausfall bei ca. 20% pro Jahr. Dieser Ausfall werde durch eine kontinuierliche Rekrutierung von neuen Panelmitgliedern in CATI-Umfragen kompensiert. Zudem strebe LINK ein stetiges Wachstum des Internet Panels an.

Ein mögliches Problem für Befragungen von Teilnehmern aus Online-Panels sind professionelle Befragungsteilnehmer, die auf Grund persönlichen oder finanziellen Interesses an sehr vielen Befragungen teilnehmen. Eine sehr hohe Befragungsintensität kann problematisch sein, wenn die Teilnahme an Befragungen die Erfahrungen, Einstellungen und Verhaltensweisen der Befragten beeinflussen – das sogenannte Panel Conditioning. Einer hohen Befragungsintensität kann durch eine maßvolle Einladungspraxis entgegengewirkt werden. Diese Maßnahme ist jedoch wirkungslos, wenn die Teilnehmer zugleich Mitglieder weiterer Online-Panels sind, über die sie an weiteren Befragungen teilnehmen. Nach eigenen Angaben fragt LINK die Mitglieder seines Internet Panels daher im Rahmen der Stammdatenbe-

---

<sup>2</sup> Die Stichprobenbasis für die CATI-Umfragen bildet nach Angaben von LINK die ADM-Stichprobenbasis. Dies gelte sowohl für deutschlandweite Umfragen, als auch für Umfragen innerhalb einzelner Bundesländer. Für die Stichprobenziehung würden je nach Vorgaben der jeweiligen Auftraggeber entweder nur Telefonanschlüsse im Festnetz verwendet oder es werde ein Dual-Frame-Ansatz gewählt, bei dem sowohl Festnetz-, als auch Mobilfunkanschlüsse in die Stichprobenziehung eingingen. Es liegen der GLES keine Angaben zu durchschnittlichen Response Rates bei diesen CATI-Umfragen vor. Weiterhin ist nicht bekannt, wie viele CATI-Interviews durchschnittlich benötigt werden, um einen Fall für das LINK Internet Panel zu gewinnen.

fragung nach der Mitgliedschaft in weiteren Online-Panels. Demnach seien zum Zeitpunkt dieser Erhebung ca. 10% der Panelmitglieder zugleich auch Mitglieder in mindestens einem weiteren Online-Panel.

### 4.3. Auswahlverfahren

Die Stichprobe für das Online-Tracking wird durch eine Quotenauswahl aus der Auswahlgesamtheit der im LINK Internet Panel zur Verfügung stehenden aktiven Panelteilnehmer generiert. Dabei werden nur volljährige Personen mit deutscher Staatsangehörigkeit berücksichtigt. Weiterhin wird darauf geachtet, dass die Panelteilnehmer nur einmal jährlich an einem Online-Tracking der GLES teilnehmen können.

Die Teilnehmer werden nach den Merkmalen Geschlecht, Alter und Bildung quotiert. Die Soll-Zahlen basieren auf der Verteilung von Best for Planning, gefiltert nach Internetnutzung und deutscher Staatsangehörigkeit. Die Verteilung ist entsprechend repräsentativ für die deutschen Internetnutzer.

**Tabelle 3: Quotierung**

Quotierungsmerkmale	Soll in Prozent	Ist in Prozent
<b>Geschlecht</b>		
Weiblich	47,6	47,8
Männlich	52,4	52,2
<b>Bildung</b>		
Niedrig (kein Abschluss, Haupt- oder Volksschule)	27,0	26,3
Mittel (Realschule, Mittlere Reife)	35,9	36,2
Hoch (Fachhochschulreife, Abitur)	37,1	37,5
<b>Altersgruppe</b>		
18-29 Jahre	21,6	21,2
30-39 Jahre	17,3	17,2
40-49 Jahre	23,7	23,9
50-59 Jahre	19,8	19,8
60 Jahre und älter	17,6	17,9

## 5. Erhebung

### 5.1. Erhebungsverfahren

Online-Befragung mit standardisiertem Fragebogen

### 5.2. Datenerhebung

Das LINK Institut für Markt- und Sozialforschung mit Sitz in Frankfurt am Main übernahm die Programmierung des Fragebogens und die Durchführung der Datenerhebung.

### 5.3. Erhebungssoftware

Die Programmierung des Fragebogens und die Durchführung der Datenerhebung erfolgten mittels der Software IBM SPSS Data Collection Base Professional 6.0.1.

### 5.4. Tests des programmierten Fragebogens

Der programmierte Fragebogen wurde sowohl vom LINK Institut für Markt- und Sozialforschung als auch von den Auftraggebern intensiv geprüft. In den Tests gefundene Probleme oder Programmierfehler wurden bereinigt. Anschließend wurde der Fragebogen in der finalen Version von der Projektgruppe freigegeben.

### 5.5. Incentivierung

Um eine möglichst hohe Ausschöpfung zu erzielen, wurden Incentives eingesetzt. Für die Teilnahme an der ca. 30-minütigen Studie erhielten die Befragten einen Amazon-Gutschein im Wert von 3,50 €.

### 5.6. Einladungen und Reminder

Die Einladung der Teilnehmer erfolgte durch LINK mit folgendem Standardtext:

Liebes Panelmitglied,

hiermit laden wir Sie zu einer interessanten sozialwissenschaftlichen Untersuchung ein.

Diese Studie befasst sich mit derzeitigen Themen in Deutschland, dabei zählt jeder einzelne!

Deshalb möchten wir auch Sie heute herzlich einladen, an unserer aktuellen Umfrage teilzunehmen.

Eine Mitwirkung an dieser Studie dauert ca. 30 Minuten und ist bis zum 19. Juni 2015 möglich.

Für Ihre Teilnahme erhalten Sie einen Amazon-Gutschein in Höhe von 3,50 EUR bis spätestens Ende Kalenderwoche 26.

Wir würden uns sehr über Ihre Teilnahme freuen!

Bitte klicken Sie auf den Link um die Studie zu starten:

[Studie starten](#)

Viele Grüße aus Frankfurt

Ihr LINK Online Team

Wenn ein Panelist die Einladung annahm, wurde er auf die LINK-Seite weitergeleitet und dort im Namen der für die Durchführung der Studie Verantwortlichen der GLES zur Befragung mit folgendem Text eingeladen:



Reagierten Teilnehmer des Online-Access-Panels nicht innerhalb eines bestimmten Zeitraums auf eine Einladung, so wurden Reminder mit folgendem Inhalt an diese Teilnehmer zugesandt:

Liebes Panelmitglied,

zur Zeit läuft eine interessante Studie zu aktuellen Themen in Deutschland, zu der wir Sie kürzlich eingeladen haben. Dabei zählt jede einzelne Meinung!

Sollten Sie den Fragebogen noch nicht beantwortet haben, würden wir uns freuen, wenn Sie sich ein wenig Zeit nehmen könnten. Eine Mitwirkung an dieser Studie dauert ca. 30 Minuten und ist bis zum 19. Juni möglich.

Für Ihre Teilnahme erhalten Sie einen Amazon-Gutschein in Höhe von 3,50 EUR bis spätestens Ende Kalenderwoche 26.

Wir würden uns sehr über Ihre Teilnahme freuen!

Bitte klicken Sie auf den Link um die Studie zu starten:

[Studie starten](#)

Viele Grüße aus Frankfurt

Ihr LINK Online Team

**Tabelle 4: Übersicht über Einladungen und Reminder**

Datum	Gruppe	N	Geschlecht		Alter					Bildung			Keine Angabe <sup>3</sup>
			Männlich	Weiblich	18-29	30-39	40-49	50-59	60+	Niedrig	Mittel	Hoch	
Fr 05.06.2015	E1	1.937	995	942	505	333	439	343	317	905	796	0	236
Sa 06.06.2015													
So 07.06.2015													
Mo 08.06.2015													
Di 09.06.2015	R1	1.495	834	661	413	293	340	290	159	716	565	0	214
Mi 10.06.2015	E2	1.316	767	549	241	317	243	333	182	426	437	453	0
Do 11.06.2015													
Fr 12.06.2015													
Sa 13.06.2015													
So 14.06.2015													
Mo 15.06.2015	R2	1.310	660	650	497	330	282	129	72	494	429	191	196
Di 16.06.2015	E3	907	546	361	174	219	268	114	132	73	120	714	0
Mi 17.06.2015													
Do 18.06.2015	R3	403	338	65	97	149	114	4	39	79	199	225	0
Fr 19.06.2015	E4	285	217	68	110	101	74	0	0	84	49	152	0

Anmerkungen: E = Einladung, R = Reminder. In der Gruppe der Personen, die keine Angabe zu ihrer Bildung machten, wurde von LINK auf Basis von Erfahrungswerten ein überproportionaler Anteil formal niedrig gebildeter Personen vermutet.

## 5.7. Feldverlauf

Die nachfolgende Tabelle illustriert das Teilnahmeverhalten der zur Umfrage eingeladenen Teilnehmer des LINK Internet Panels über den Erhebungszeitraum hinweg. Partielle und vollständige Interviews umfassen alle Interviews, die bis zur letzten Frage des Fragebogens durchgeführt wurden, unabhängig davon, ob die Umfrage unterbrochen wurde oder nicht. Die Anzahl der begonnenen Interviews ergibt sich aus der Anzahl der partiellen und vollständigen Interviews, der abgebrochenen Interviews sowie der auf Grund der Quotierung nicht durchgeführten Interviews.

<sup>3</sup> Überproportionaler Anteil formal niedrig gebildeter Personen in dieser Gruppe auf Basis von Erfahrungswerten vermutet.

**Tabelle 5: Teilnahmestatistik nach Tagen**

Datum	Begonnene Interviews			Partielle und vollständige Interviews		
	N	Prozent	kum. Prozent	N	Prozent	kum. Prozent
Fr 05.06.2015	181	11,6	11,6	156	15,3	15,3
Sa 06.06.2015	58	3,7	15,3	52	5,1	20,4
So 07.06.2015	66	4,2	19,5	59	5,8	26,2
Mo 08.06.2015	48	3,1	22,6	37	3,6	29,8
Di 09.06.2015	84	5,4	28,0	58	5,7	35,5
Mi 10.06.2015	290	18,6	46,6	235	23,1	58,6
Do 11.06.2015	107	6,9	53,5	50	4,9	63,5
Fr 12.06.2015	43	2,8	56,3	13	1,3	64,8
Sa 13.06.2015	39	2,5	58,8	9	0,9	65,7
So 14.06.2015	40	2,6	61,4	7	0,7	66,4
Mo 15.06.2015	80	5,1	66,5	37	3,6	70,0
Di 16.06.2015	260	16,6	83,1	193	18,9	88,9
Mi 17.06.2015	105	6,7	89,8	45	4,4	93,3
Do 18.06.2015	75	4,8	94,6	24	2,4	95,7
Fr 19.06.2015	86	5,5	100,0	44	4,3	100,0
<b>Gesamt</b>	<b>1.562</b>	<b>100,0</b>		<b>1.019</b>	<b>100,0</b>	

## 5.8. Ausschöpfung

Systematischer Unit Nonresponse ist eine Fehlerquelle in Umfragestudien, die die Qualität der erhobenen Daten erheblich mindern kann. Sind die in der Umfrage erhobenen Variablen mit der Teilnahmewahrscheinlichkeit korreliert, so sind Verzerrungen der Umfrageergebnisse („Nonresponse Bias“) die Folge. Die Angabe von Ausschöpfungsquoten („Response Rates“) erlaubt eine Einschätzung darüber, wie stark eine Umfragestudie von möglicherweise systematischer Nichtteilnahme betroffen ist. Die Angabe von standardisierten Indikatoren des Teilnahmeverhaltens der Personen in der Stichprobe sichert die Vergleichbarkeit über Studien hinweg. Für diese Studie werden zwei Indikatoren des Teilnahmeverhaltens nach den 2011 publizierten Standards der American Association for Public Opinion Research (AAPOR 2011) sowie Empfehlungen zur Berechnung der Ausschöpfung von Callegaro und DiSogra (2008) angegeben.

Die AAPOR Response Rate 2 (AAPOR 2011) berechnet sich als

$$\text{Response Rate 2 (RR2)} = \frac{(I + P)}{(I + P) + (R + NC + O) + (UH + UO)}$$

wobei I die Anzahl der vollständigen Interviews, P die Anzahl der partiellen Interviews, R die Anzahl der Verweigerungen und Befragungsabbrüche, NC die Anzahl der Nichtkontakte, O die Anzahl der Ausfälle aus anderen Gründen und UH sowie UO die Anzahl der Fälle ist, bei denen unbekannt ist, ob sie die Befragungseinladung erhalten und gesehen haben, d.h. ob sie teilnahmeberechtigt und teilnahmebereit waren.<sup>4</sup> Die RR2 entspricht der minimalen

<sup>4</sup> Bei Online-Befragungen ist die Interpretation der Response Rate allerdings in vielen Fällen problematisch, da in der Regel keine zufallsbasierten Stichprobenziehungsverfahren angewendet werden. Ein gängiger Ausweg ist daher, die Mitglieder von Online-Panels zu befragen. Bei der Befragung von Teilnehmern aus zufallsbasierten Online-Panels, wie es das LINK Internet Panel ist, empfiehlt die AAPOR die Angabe einer kumulierten Response Rate („cumulative response rate (CUMRR)“),

Response Rate 1 (RR1) mit dem Unterschied, dass bei der RR2 partielle Interviews als Teilnahmen gezählt werden.

Als weiteren informativen Indikator wird der Anteil der Befragungsabbrüche in der Online-Befragung berichtet. Die Breakoff Rate (Callegaro und DiSogra 2008) berechnet sich als

$$\text{Breakoff Rate} = \frac{R_{\text{Breakoff}}}{(I + P) + (R_{\text{Breakoff}})}$$

wobei  $R_{\text{Breakoff}}$  die Anzahl der Befragungsabbrüche,  $I$  die Anzahl der vollständigen Interviews und  $P$  die Anzahl der partiellen Interviews umfasst. Die Breakoff Rate gibt somit den Anteil der Befragungsabbrüche an allen begonnenen Interviews wieder.

Die Brutto-Stichprobe für die Online-Befragung wurde aus dem LINK Internet Panel gezogen. Die ausgewählten Panelmitglieder wurden von LINK zur Teilnahme eingeladen. Panelteilnehmer, die das Interview bis zur abschließenden Seite der Befragung absolvierten, wurden der Gruppe der vollständigen ( $I$ ) und partiell vollständigen Interviews ( $P$ ) zugeordnet. Wurde das Interview begonnen, während der Teilnahme unterbrochen und bis zum Ende der Feldzeit nicht vollständig absolviert, so wird der Panelteilnehmer der Gruppe der Befragungsabbrüche ( $R_{\text{Breakoff}}$ ) zugeordnet. Um den Quotenvorgaben zu genügen, wurden Panelteilnehmer mit bestimmten Ausprägungen bei den Variablen Geschlecht, Alter und höchstem allgemein bildendem Schulabschluss nach Erreichen der einzelnen Vorgaben von der Teilnahme ausgeschlossen. Die betroffenen Teilnehmer wurden „abgewiesen“ (nicht teilnahmeberechtigt). Erfolgte keine Reaktion auf eine Befragungseinladung, so war im Allgemeinen nicht bekannt, ob die Einladung von den Zielpersonen überhaupt erhalten und gesehen und eine Befragungsteilnahme in Erwägung gezogen wurde (UH).

**Tabelle 6: Ausschöpfung**

Kategorie/Indikator	Anzahl/Rate in Prozent	
Anzahl der von LINK versandten Befragungseinladungen		4.445
I & P = Vollständige und partiell vollständige Interviews		1.019
R = Verweigerungen und Befragungsabbrüche		162
Quote erfüllt/abgewiesen = Nicht teilnahmeberechtigt		388
UH = Unbekannt, ob die Befragungseinladung erhalten und gesehen wurde		2.876
AAPOR Response Rate 2 (RR2)	in Prozent	25,1
Breakoff Rate	in Prozent	13,7

die sich wiederum aus der Recruitment Rate (RECR), der Profile Rate (PROR) sowie der Completion Rate (COMR) berechnet. Da sowohl die Recruitment Rate als auch die Profile Rate bei einer kontinuierlichen Rekrutierung von Befragten schwierig zu berechnen sind und von LINK nicht bereitgestellt werden, wird hier lediglich die RR2 angegeben.

## 5.9. Interviewdauer

Die Interviewdauer ist ein Indikator für den Umfang einer Befragung. Sie lässt Rückschlüsse darauf zu, wie groß der Aufwand war, den Befragte auf sich nehmen mussten, um den Fragebogen zu beantworten. Die nachfolgende Tabelle präsentiert deskriptive Statistiken zur Interviewdauer.

**Tabelle 7: Interviewdauer**

	<b>N</b>	<b>Min.</b>	<b>Max.</b>	<b>Arithmetisches Mittel</b>	<b>Median</b>
Interviewdauer	980	00:06:21	26:39:33	00:36:39	00:26:16

Anmerkungen: Die Interviewdauer wird nur für diejenigen Befragten ausgewiesen, die die Umfrage weder unter- noch abgebrochen haben. Wurde die Umfrage unterbrochen, ohne den Fragebogen im Browser zu schließen, wurde dies nicht als Unterbrechung gewertet.



## 6. Datenaufbereitung

### 6.1. Datenkontrollen

Die Daten dieser Studie wurden nach der Erhebung sorgfältig geprüft. Hierunter fiel die Überprüfung auf Vollständigkeit der Variablen im Datensatz, die Überprüfung auf Richtigkeit von Variablennamen und -labels, die Prüfung auf Vollständigkeit von Wertelabels, die Überprüfung der Codierung von Missing Values sowie eine Kontrolle der Randverteilung. Darüber hinaus erfolgte eine technische Kontrolle der Filterführungen. Alle gefundenen Fehler werden während der Datenaufbereitung dokumentiert und korrigiert. Trotz gründlicher Überprüfung der Daten kann es vorkommen, dass Fehler in den Datensätzen unentdeckt bleiben. Fallen diese zu einem späteren Zeitpunkt auf, werden sie in der Errataliste im GESIS Datenbestandskatalog ([www.gesis.org/dbk](http://www.gesis.org/dbk)) dokumentiert und zeitnah behoben.

### 6.2. Einheiten im Datensatz

Der Datensatz umfasst 1.019 realisierte Interviews.

### 6.3. Variablen im Datensatz

Der Datensatz setzt sich aus verschiedenen Typen von Variablen bzw. Daten zusammen. Diese umfassen Meta-, Para- und substantielle Daten sowie Kontext- und GewichtungsvARIABLEN (siehe auch die Abschnitte 6.4 bis 6.8). Die Tabelle 8 gibt einen Überblick über die Anzahl der Variablen im Datensatz nach Variablentyp.

**Tabelle 8: Anzahl der Variablen im Datensatz nach Variablentyp**

Variablentyp	Anzahl
Metadaten	7
Paradaten	125
Substantielle Daten	308
Kontextvariablen	7
GewichtungsvARIABLEN	4
<b>Gesamt</b>	<b>451</b>

### 6.4. Metadaten

Metadaten beschreiben den Datensatz und seine Erstellung (Callegaro 2013). Sie umfassen Variablen, die für die Archivierung und Distribution der Studie benötigt werden. In der GLES umfasst dies die Studiennummer des Datensatzes im vierstelligen Format, unter welcher die Studie in den Archivsystemen von GESIS abgelegt und zu finden ist, die Versionskennung des Datensatzes, Angaben zum Erhebungsjahr und –zeitraum, die Zuordnung der Studie zu den Komponenten der GLES sowie die Kennung der Erhebung innerhalb der Komponente der GLES (siehe Tabelle 9).

**Tabelle 9: Metadaten**

Variable	Label
study	Studiennummer
version	GESIS Archiv Version
doi	Digital Object Identifier
year	Erhebungsjahr
field	Erhebungszeitraum
glescomp	GLES-Komponente
survey	Erhebung/Welle

## 6.5. Paradata

Paradata sind Informationen, die aus dem Erhebungsprozess resultieren (Couper 2000), z.B. das Datum und die Uhrzeit des Beginns und des Endes der Interviewteilnahme, die Interviewdauer, Antwortzeitmessungen oder Informationen zum verwendeten Gerät und dem verwendeten Webbrowser (Callegaro 2013). Auch administrative Informationen wie Identifikationsnummern, Informationen zum Teilnahmestatus von Befragungsteilnehmern und Informationen zur Zuweisung von Teilnehmern zu Gruppen auf Grund von inhaltlichen Angaben oder zufallsbasierten Verfahren werden hier als Paradata behandelt. Die Tabelle 10 gibt einen Überblick über Paradata im Datensatz.

**Tabelle 10: Paradata**

Variable	Label
ifdn	Laufende Nummer
starttime	Startzeitpunkt letzter Einstieg
starttime1	Startzeitpunkt erster Einstieg bei Unterbrechung
endtime	Endzeitpunkt letzter Einstieg
interrupt	Unterbrechung der Befragung
duration	Interviewdauer in Sekunden
zstart - zende	Variable: Dauer
speederindex	Zeitunterschreiter-Index
speederflag	Zeitunterschreiter (10% der Befragten mit den niedrigsten Indexwerten)
browser	Genutzter Browser
smartphone	Nutzung Smartphone
javascript	Installation Javascript
hordert122	Reihenfolge Items Frage t122
hordere0111	Reihenfolge Items Frage e0111

### 6.5.1 Laufende Nummer und Teilnahmestatus

Die Variable *ifdn* „Laufende Nummer“ ist eine eindeutige Identifikationsnummer für die Teilnehmer an der Befragung. Eine Identifikation von Personen, die an mehreren Erhebungen des Online-Trackings teilgenommen haben, mittels der laufenden Nummer ist nicht möglich, da diese für jede Erhebung neu generiert wird.

### 6.5.2 Unterbrechung der Befragung

Die Teilnehmer der Studie konnten die Beantwortung des Fragebogens jederzeit unterbrechen und innerhalb der Feldzeit zu einem von ihnen selbst gewählten Zeitpunkt fortsetzen. Hat ein Befragter die Beantwortung unterbrochen und den Browser oder das Browserfenster auf seinem Computer geschlossen, so wurde diese Information in der Variable *interrupt* „Unterbrechung der Befragung“ gespeichert. Die Unterbrechung der Beantwortung der Umfrage ohne Schließen des Fragebogens im Browser kann nicht direkt festgestellt werden.

### 6.5.3 Zeitvariablen

Es wurden zwei Arten von Zeitvariablen Server-seitig und sekundengenau als Paradata erhoben.

Erstens wurde die Interviewdauer (Variable *duration*) gemessen, d.h. die Zeit, die ein Befragter zwischen dem Betrachten und Bearbeiten der Begrüßungsseite bis zum Schließen der letzten Seite der Befragung benötigt hat. Bei der Analyse der Interviewdauer ist zu beachten, dass Befragte den Fragebogen verlassen und die Befragung zu einem späteren Zeitpunkt fortsetzen konnten. Die Information zu Unterbrechungen der Befragung ist im Datensatz (Variable *interrupt*) enthalten.

Zweitens wurde die Zeit erhoben, die Befragte zum Betrachten und Bearbeiten der einzelnen Bildschirmseiten der Befragung (Variablen *zstart* bis *zende*) benötigt haben. Diese Zeitmessungen umfassen neben der reinen Antwortzeit auch diejenigen Zeiten, die für die Datenübertragung zwischen dem Server und dem Computer des Befragten, die Interpretation der Daten und den Bildschirmaufbau durch den Browser sowie das Lesen der Fragen und Antwortoptionen durch den Befragten benötigt werden (siehe Kaczmirek 2008).

### 6.5.4 Zeitunterschreiter

Aufgrund der fehlenden Kontrolle der Befragten durch einen Interviewer in Online-Befragung und der besonderen Anreiz- und Belohnungsstruktur durch die Incentivierung der Befragten ist diese Studie mit dem Problem außergewöhnlich schneller Antwortzeiten („Zeitunterschreitung“) konfrontiert. Zeitunterschreitung bedeutet, dass einige Befragte einzelne Fragen oder auch die gesamte Umfrage erheblich schneller beantworten als der Großteil der Teilnehmer. Eine schnelle Beantwortung der Fragen in der Umfrage ist noch kein Problem an sich, da es aufgrund bestimmter sozialstruktureller und persönlicher Merkmale (z.B. Bildung, Alter, Intelligenz, Reaktionsgeschwindigkeit) deutliche Unterschiede in der Beantwortungsgeschwindigkeit zwischen Befragten geben kann. Dennoch ist davon auszugehen, dass bei einer erheblichen Unterschreitung der mittleren Antwortzeit die Antwortqualität leidet, da sich Befragte mutmaßlich wenig bei der Beantwortung der Fragen bemühen und daher oberflächliche oder willkürliche Antworten abgeben, keine Angabe machen oder „weiß nicht“ antworten, obwohl sie eine substantielle Antworten hätten abgeben können (vgl. Krosnick 1991).

In der Fachliteratur gibt es keine etablierten Standards für die Identifikation von Zeitunterschreitern. In der Regel beziehen die Maße für deren Identifikation den Median bzw. Mittelwert der Verteilung und die Streuung mit ein und wählen auf dieser Basis ein Abschneidekriterium, das nicht unterschritten werden darf (vgl. Mayerl und Urban 2008). Diese Befragten werden dann entweder aus dem Datensatz ausgeschlossen oder durch Markervariablen gekennzeichnet.

In dieser Studie werden Zeitunterschreiter anhand einer modifizierten Routine des von Roßmann (2010) weiterentwickelten Algorithmus identifiziert. Hierfür wird ein Zeitunterschreiter-Index gebildet (Variable *speederindex*), der sowohl die Antwortdauer der Befragten auf allen Bildschirmseiten der Befragung als auch die Interviewdauer je Befragtem einbe-

zieht und Indexwerte zwischen größer 0 und kleiner 2 annimmt. Ein Indexwert von 1 zeigt eine mittlere Antwortgeschwindigkeit an, während Werte gegen 0 im Mittel sehr schnelle und Werte gegen 2 im Mittel sehr langsame Antwortzeiten anzeigen. Als Zeitunterschreiter werden alle Befragten im unteren 10%-Perzentil der Verteilung des Summenindex gekennzeichnet. Die Kennzeichnung von Zeitunterschreitern ist im Datensatz in der Variable *speederflag* enthalten. Diese Markervariable kann zum Ausschluss der Fälle aus Analysen als auch zur Durchführung von Robustheitsanalysen verwendet werden.

In dieser Studie wurden 102 Fälle (10,01%) als Zeitunterschreiter markiert.

**Tabelle 11: Statistiken zur Stichprobe mit und ohne Zeitunterschreiter**

<b>Merkmal</b>	<b>Mit Zeitunterschreitern</b>	<b>Ohne Zeitunterschreiter</b>
<b>Interviewdauer</b>		
<b>(hh:mm:ss)</b>		
Min.	00:06:21	00:16:09
Mittelwert	00:40:58	00:43:52
Median	00:26:27	00:27:29
<b>Geschlecht (in %)</b>		
Männlich	52,2	52,3
Weiblich	47,8	47,7
<b>Altersgruppe (in %)</b>		
18 bis unter 30 Jahre	21,2	18,9
30 bis unter 45 Jahre	27,1	26,1
45 bis unter 60 Jahre	33,9	35,3
60 Jahre und älter	17,9	19,7
<b>Bildung (in %)</b>		
Niedrig	26,3	26,0
Mittel	36,2	36,5
Hoch	37,5	37,5
<b>N</b>	<b>1.019</b>	<b>917</b>

### 6.5.5 Browser, Smartphone und Javascript

Die verwendete Erhebungssoftware liest den sogenannten „user agent header“ aus. Der „user agent“ ist eine Software, die in der Kommunikation im Internet als Interface zwischen den Servern, auf denen die Information gespeichert sind, und dem Nutzer fungiert, z.B. der auf dem Computer des Nutzers verwendete Browser. Es ist somit möglich zu identifizieren, welcher Browser verwendet wurde, ob die Befragung mittels eines Smartphones ausgefüllt wurde und ob der Browser Javascript unterstützt.

## 6.6. Substantielle Daten

Substantielle Daten bilden den Hauptbestandteil des Datensatzes. Sie umfassen inhaltliche und soziodemographische Variablen.

**Tabelle 12: Inhaltliche und soziodemographische Variablen**

Variable	Label
t1	Geschlecht
t2	Alter
t3	Schulabschluss
t4	Bundesland
t5	Politisches Interesse
t6	Demokratiezufriedenheit
t7	Wahlbeteiligungsabsicht
t8a-b	Beabsichtigte Stimmabgabe
t9a-b	Hypothetische Stimmabgabe
t10	Wichtigstes Problem
t11	Wichtigstes Problem, Lösungskompetenz
t12	Zweitwichtigstes Problem
t13	Zweitwichtigstes Problem, Lösungskompetenz
t14a-h	Skalometer Parteien (CDU, CSU, SPD, DIE LINKE, GRÜNE, FDP, AfD)
t15a-j	Skalometer Politiker
t16	Leistung Bundesregierung
t17a-c	Leistung Regierungsparteien (CDU, CSU, SPD)
t18a-b	Leistung Oppositionsparteien (DIE LINKE, GRÜNE)
t21	Eigene wirtschaftliche Lage, aktuell
t19	Eigene wirtschaftliche Lage, retrospektiv
t20	Verantwortlichkeit eigene wirtschaftliche Lage
t22	Eigene wirtschaftliche Lage, prospektiv
t23a-h	Links-Rechts-Einstufung Parteien (CDU, CSU, SPD, DIE LINKE, GRÜNE, FDP, AfD)
t24	Links-Rechts-Selbsteinstufung
t25	Unterschied Regierung
t26	Unterschiede Parteien
t29	Allgemeine wirtschaftliche Lage, aktuell
t27	Allgemeine wirtschaftliche Lage, retrospektiv
t28	Verantwortlichkeit allg. wirt. Lage

---

t110	Verantwortlichkeit allg. wirt. Lage (Länder)
t109	Verantwortlichkeit allg. wirt. Lage (EU)
t30	Allgemeine wirtschaftliche Lage, prospektiv
t145	Regionale wirt. Lage, aktuell
t31	Recall, Wahlberechtigung BTW 2013
t32	Recall, Wahlteilnahme, BTW 2013
t33a-b	Recall, Wahlentscheidung, BTW 2013
t45	Gespräche über Politik
t304	Netzwerkgröße
t444	Wahlbeteiligung, Nachbarschaft
t443	Wahlbeteiligung, Bekanntenkreis
t362	Recall, Wahlberechtigung, LTW
t363	Recall, Wahlteilnahme, LTW
t35	Recall, LTW
t308a-m	Issuebatterie
e0113	Issuebatterie, Tarifeinheit
t36	Wichtigste Informationsquelle
t37a-h	Nutzung Tageszeitung
t38a-d	Printmedien, Nutzung, Wochenzeitschriften
t39a-e	Nutzung TV-Nachrichten
t80	Internetnutzung, allgemein
t40	Internetnutzung, politische Information
t41a-h	Internetnutzung, Seiten
t42	Sonntagsfrage, LTW (1 Stimme)
t43	Sonntagsfrage, LTW (2 Stimmen)
t44a-l	Sonntagsfrage, LTW (5 und 10 Stimmen)
t119	EU, Mitgliedschaft
t131	EU, Meinung
t305	Volksentscheid EU-Beitritt
t121	Einfluss Europa auf dt. Politik
t122a-i	EU, Aussagen
t132a-c	Aussagen zum Euro
e0111a-b	EU-Austritt, Maßnahmen
t124a-p	Vertrauen, Institutionen
t128a-g	Europa, Parteien
t129	Europa, Ego

---

---

t130	Europa, Salienz
t113a-h	Positionsissue: Sozioökonomische Dimension, Parteien
t114	Positionsissue: Sozioökonomische Dimension, Ego
t115	Positionsissue: Sozioökonomische Dimension, Salienz
t317a-h	Positionsissue: Klimawandel, Parteien
t318	Positionsissue: Klimawandel, Ego
t319	Positionsissue: Klimawandel, Salienz
t153a-h	Positionsissue: Libertär-autoritäre Dimension, Parteien
t154	Positionsissue: Libertär-autoritäre Dimension, Ego
t155	Positionsissue: Libertär-autoritäre Dimension, Salienz
e0112	Flüchtlinge
t156a-e	Einstellungen, Politik und Gesellschaft
t46	Parteiidentifikation
t47	Parteiidentifikation, Stärke
t48	Parteiidentifikation, Dauer
t126a-h	Verbundenheit Batterie
t49a-j	Organisationsmitgliedschaft
t57	Familienstand
t50	Haushaltsgröße
t348a-g	Haushaltsmitglieder, Alter
t303a-m	Berufliche Bildung
t52	Erwerbstätigkeit
t53	Angst Stellenverlust
t54	Erwerbstätigkeit, früher
t55	Beruf
t56	Beruf, früher
t351	Beruf, Nachfrage I
t352	Beruf, Nachfrage II
t353	Beruf, Nachfrage III
t354	Beruf, Nachfrage IV
t58	Beschäftigungssektor
t59	Beschäftigungssektor, früher
t60	Religionszugehörigkeit
t320	Häufigkeit Gottesdienst
t64	Deutsche Staatsbürgerschaft Geburt

---

---

t65	Deutsche Staatsbürgerschaft, Dauer
t66	Geburtsland (V2)
t72	Geburtsland, Bundesland
t73	Geburtsland, Ausland
t68	Geburtsland, Mutter (V2)
t67	Geburtsland, Vater (V2)
t454	Erfassung Migration in 3. Generation
t69	Schichtzugehörigkeit, subjektiv
t70	Nettoeinkommen HH
t71	Postleitzahl

---

### 6.6.1 Codierung von offenen Fragen

Offene Fragen wurden durch LINK codiert. Die Codierung erfolgte unter Verwendung von Codierschemata, die von der GLES-Projektgruppe entwickelt und bereitgestellt wurden. Diese können auf den Seiten der GLES bei GESIS abgerufen werden (<http://www.gesis.org/gles>).

Darüber hinaus werden die offenen Fragen als String-Variablen bereitgestellt. Hierfür werden die offen erhobenen Antworten intensiv auf datenschutzrechtlich bedenkliche Nennungen geprüft und nötigenfalls zum Schutz der Befragten zensiert.

Für die Arbeit mit den offen erhobenen Antworten empfehlen wir die Nutzung des SPSS-Datensatzes, da aufgrund von Programmrestriktionen im Stata-Datensatz alle offenen Angaben bei 244 Zeichen abgeschnitten werden.

### 6.6.2 Codierung von halboffenen Fragen

Halboffene Fragen wurden durch LINK codiert. Die Codierung erfolgte in der Regel nach Vorgaben und anhand von Codierschemata der GLES-Projektgruppe. Im Datensatz werden in der Regel nur die vercodeten Variablen veröffentlicht. Bei Interesse an den nicht vercodeten offenen Antworten melden Sie sich bitte unter [gles@gesis.org](mailto:gles@gesis.org).

### 6.6.3 Codierung der Parteien

Die Codierung der Parteien erfolgte bei allen geschlossenen, halboffenen und offenen Fragen nach dem einheitlichen Codierschema der GLES. Das Codierschema kann auf den Seiten der GLES bei GESIS abgerufen werden (<http://www.gesis.org/gles>).

In der Regel werden die Angaben der Befragten zu Parteien in zwei Variablen veröffentlicht, die mit Version A und Version B bezeichnet werden. Version A enthält für Parteienfragen stets die Parteien, die auf dem verwendeten Musterstimmzettel aufgeführt sind (inklusive der NPD), sowie eine Kategorie für „andere“ Parteien. Version B unterscheidet sich von Version A ausschließlich durch eine detaillierte Ausweisung der von den Befragten genannten „anderen“ Parteien.

## 6.7. Kontextvariablen

Kontextvariablen geben Auskünfte über den regionalen Kontext der Befragten, z.B. den Wahlkreis und die Gemeinde, in welchen Befragte leben.



**Tabelle 13: Kontextvariablen**

Variable	Label
ostwest	Ost/West (Zuordnung von Berlinern zu Ostdeutschland)
elecdist	Wahlkreis (Electoral District)
elecdist1	Wahlkreis nicht eindeutig zuordenbar: 1. Möglichkeit
elecdist2	Wahlkreis nicht eindeutig zuordenbar: 2. Möglichkeit
elecdist3	Wahlkreis nicht eindeutig zuordenbar: 3. Möglichkeit
elecdist4	Wahlkreis nicht eindeutig zuordenbar: 4. Möglichkeit
bik10	BIK10

### 6.7.1 Ost/West

Die Zuordnung zu Ost- und Westdeutschland erfolgte auf Grundlage der Angabe zum Bundesland, in dem die Befragten ihren Hauptwohnsitz haben. Da bei der Abfrage des Bundeslands keine Differenzierung zwischen Ost- und Westberlin vorgenommen wurde, werden Berliner zu Ostdeutschland zugeordnet.

### 6.7.2 Wahlkreise

Die Zuordnung zu einem Wahlkreis beruht auf der Wahlkreiseinteilung zur Bundestagswahl 2013 und wurde unter Verwendung der von den Befragten genannten Postleitzahl an ihrem Hauptwohnsitz vorgenommen. Die Zuordnung der Postleitzahlen zu Wahlkreisen erfolgte anhand einer Zuordnungstabelle, die vom Bundeswahlleiter bereitgestellt und durch eigene Recherche ergänzt wurde.

Bei der Verwendung der Wahlkreisinformationen ist zu beachten, dass einzelne Postleitzahlen in großstädtischen Gebieten zwischen zwei und fünf Wahlkreisen zugeordnet werden können, d.h. eine eindeutige Zuordnung der Postleitzahl zu einem Wahlkreis ist auf Grundlage der vorhandenen Informationen nicht immer möglich. Die Variable *elecdist* „Wahlkreis“ enthält die Angabe zum Wahlkreis des Befragten, wenn eine eindeutige Zuordnung möglich ist. Ist dies nicht der Fall, so sind in den Variablen *elecdist1-elecdist5* „Wahlkreis nicht eindeutig zuordenbar: 1.-5. Möglichkeit“ die Nummern der Wahlkreise erfasst, denen der Befragte anhand der von ihm genannten Postleitzahl potenziell zugeordnet werden kann. In einzelnen Fällen ist eine Zuordnung zu einem oder mehreren Wahlkreisen nicht möglich, da eine Angaben zur Postleitzahl fehlt oder die vom Befragten genannte Postleitzahl nicht existiert.

**Tabelle 14: Zuordnung von Postleitzahlen zu Wahlkreisen**

Art der Zuordnung	N	Prozent
Eindeutig einem Wahlkreis zuordenbar	844	82,9
Nicht eindeutig einem Wahlkreis zuordenbar	145	14,2
Fehlende Angabe oder nicht zuordenbare Postleitzahl	30	2,9
<b>Gesamt</b>	<b>1.019</b>	<b>100,0</b>

### 6.7.3 BIK-Regionsgrößenklassen

Die BIK-Regionsgrößenklassen klassifizieren die BIK-Regionen gemäß ihrer Bevölkerungszahl. Dabei wird nicht die Größe der Gemeinde selbst angegeben, sondern die Menge der

Bevölkerung des Raumes, in den die Gemeinde funktional eingebunden ist. Weiterhin wird zwischen vier Strukturtypen (SGTYP) der Gemeinde unterschieden:

- SGTYP1: Kernbereich der Stadtregion,
- SGTYP2: Verdichtungsbereich,
- SGTYP3: Übergangsbereich,
- SGTYP4: peripherer Bereich (siehe BIK Aschpurwis+Behrens 2001).

Aus Gründen des Datenschutzes liegt die Information zur BIK-Regionsgrößenklasse in der 10er-Systematik (siehe Tabelle 15) im veröffentlichten Datensatz nur in recodierter Form vor. Hierbei wurden die untersten drei Kategorien zur Regionsgrößenklasse „bis unter 20.000 Einwohner“ zusammengefasst (siehe auch Abschnitt 6.9).

**Tabelle 15: BIK-Regionsgrößenklassen in der 10er-Systematik**

Code	BIK-Regionsgrößenklasse
1	bis unter 2.000 Einwohner
2	2.000 bis unter 5.000 Einwohner
3	5.000 bis unter 20.000 Einwohner
4	20.000 bis unter 50.000 Einwohner
5	50.000 bis unter 100.000 Einwohner (peripherer Bereich, Übergangs-, Verdichtungsbereich)
6	50.000 bis unter 100.000 Einwohner (Kernbereich)
7	100.000 bis unter 500.000 Einwohner (peripherer Bereich, Übergangs-, Verdichtungsbereich)
8	100.000 bis unter 500.000 Einwohner (Kernbereich)
9	500.000 Einwohner oder mehr (peripherer Bereich, Übergangs-, Verdichtungsbereich)
10	500.000 Einwohner oder mehr (Kernbereich)

## 6.8. Gewichtungsvariablen

Mit Hilfe von Anpassungsgewichten kann die Verteilung von Variablen eines Datensatzes an bekannte Verteilungen in der Grundgesamtheit angepasst werden. Dabei wird davon ausgegangen, dass zumindest ein schwacher Zusammenhang zwischen den Anpassungsvariablen und den inhaltlichen Untersuchungsvariablen besteht. Die für die GLES-Daten erstellten Gewichte werden mittels einer stets gleichen Vorgehensweise berechnet, so dass sie über die verschiedenen GLES-Komponenten hinweg vergleichbar sind.

Zur Berechnung der Anpassungsgewichte der GLES wurden sowohl sozial- als auch regionalstrukturelle Merkmale ausgewählt. Bei der gleichzeitigen Anpassung an mehrere Merkmale ergeben sich häufig Nullzellen oder sehr kleine Zellbesetzungen, was bei einer einfachen Soll/Ist-Gewichtung zu Problemen führt. Daher wurde bei der Berechnung der Anpassungsgewichte für die GLES auf das iterative Anpassungsverfahren („iterative proportional fitting“, IPF)<sup>5</sup> zurückgegriffen. Bei der IPF-Gewichtung wird die Ist-Verteilung der einzelnen Zellen schrittweise an die jeweilige Soll-Verteilung der Gewichtungsvariablen angepasst. Bei diesem schrittweisen Prozess der Anpassung (Iteration) bilden die nach jedem Anpassungsschritt berechneten Gewichtungsfaktoren die Ausgangsbasis für die Anpassung des nächsten Merkmals. Der Anpassungsprozess endet, wenn die Differenz zwischen der gewichteten

<sup>5</sup> Siehe zum IPF-Verfahren Deming und Stephan (1940).

Randverteilung aller Faktoren und der Soll-Verteilung das Abbruchkriterium von 0,05 unterschreitet.<sup>6</sup> Um sehr große Gewichtungsfaktoren zu vermeiden wurden die Faktoren (nach jedem Iterationsschritt) auf den fünffachen Mittelwert der Gewichtungsvariable (also fünf) getrimmt.<sup>7</sup>

Bei der Erstellung der Anpassungsgewichte für das Tracking wurde auf die aktuellsten verfügbaren Randverteilungen des Mikrozensus<sup>8</sup> und des (N)Onliner Atlas<sup>9</sup> zurückgegriffen. Bei der Erstellung der Sollverteilungen des Mikrozensus ging nur die Bevölkerung über 18 Jahren mit deutscher Staatsbürgerschaft in Privathaushalten am Ort der Hauptwohnung ein. Die Sollverteilungen des (N)Onliner Atlas beruhen auf den Angaben zu Personen mit deutscher Staatsbürgerschaft ab einem Alter von 18 Jahren.

Für diese Studie wurden insgesamt vier Anpassungsgewichte erstellt, wobei die Gewichte einmal mit und einmal ohne Zeitunterschreiter berechnet wurden.

**Tabelle 16: Übersicht über die sozial- und regionalstrukturellen Gewichtungsvariablen**

Gewichtungsvariable	Anpassung an Randverteilungen des ...	Berücksichtigung von Zeitunterschreitern
wei_mzz	Mikrozensus 2013	ja
wei_mzoz	Mikrozensus 2013	nein
wei_onz	(N)Onliner Atlas 2014	ja
wei_onoz	(N)Onliner Atlas 2014	nein

Angepasst wurde an sozio-demographische und regionalstrukturelle Merkmale: Geschlecht, Alter, Bildung, BIK-Regionsgrößenklassen und Zugehörigkeit zu West- bzw. Ostdeutschland (inkl. Berlin).

Die Variable Alter wurde kategorisiert und umfasst nun vier Gruppen: „18 bis unter 30 Jahre“, „30 bis unter 45 Jahre“, „45 bis unter 60 Jahre“ und „60 Jahre und älter“.

Bei der Kategorisierung der Variable Bildung wurden die drei folgenden Gruppen gebildet:

- niedrige Bildung: Schule beendet ohne Abschluss, Hauptschulabschluss, Volksschulabschluss, bin noch Schüler/in;
- mittlere Bildung: Realschulabschluss, Mittlere Reife, Fachschulreife oder Abschluss der polytechnischen Oberschule 10. Klasse;
- hohe Bildung: Fachhochschulreife (Abschluss einer Fachoberschule etc.), Abitur bzw. erweiterte Oberschule mit Abschluss 12. Klasse (Hochschulreife).

Die BIK-Regionsgrößenklassen wurden zu zwei bzw. drei Kategorien zusammengefasst (siehe nachfolgende Tabelle). Die unterschiedliche Kategorisierung ist den verfügbaren Randverteilungen geschuldet. Für die Anpassung an den Mikrozensus wurde an drei Kate-

<sup>6</sup> Die Anpassungsgewichte wurden mit dem Statistikprogramm Stata berechnet, wobei auf das frei verfügbare Ado „ipfweight“ (Bergmann 2011) zurückgegriffen wurde. Zur Installation muss lediglich „ssc install ipfweight“ in die Kommandozeile von Stata eingegeben werden.

<sup>7</sup> Dieses Vorgehen findet auch bei der Berechnung der Gewichte der American National Election Study (ANES) Anwendung (siehe DeBell et al. 2009).

<sup>8</sup> Bei den vom Mikrozensus ausgewiesenen Werten handelt es sich um absolute Personenzahlen nach gebundener Hochrechnung.

<sup>9</sup> Die Randverteilungen (für Personen ab 18 Jahren) aus dem (N)Onliner Atlas wurden für die Gewichtung freundlicherweise von TNS Infratest zur Verfügung gestellt.

gorien angepasst, während beim (N)Onliner Atlas lediglich die Anpassung an zwei Kategorien möglich war.<sup>10</sup> Fehlende Werte wurden für die Erstellung der GewichtungsvARIABLEN auf den Modalwert der Verteilung gesetzt.

**Tabelle 17: Soll- und Ist-Verteilungen der zur Erstellung der Gewichte verwendeten Variablen**

Merkmal	Ist in Prozent		Soll in Prozent	
	Mit Zeitunter- schreibern (N=1.019)	Ohne Zeitun- terschreiber (N=917)	Mikrozensus	(N)Onliner Atlas
<b>Geschlecht</b>				
Männlich	52,2	52,3	48,4	52,6
Weiblich	47,8	47,7	51,6	47,4
<b>Altersgruppe</b>				
18 bis unter 30 Jahre	21,2	18,9	16,0	21,3
30 bis unter 45 Jahre	27,1	26,1	21,7	27,3
45 bis unter 60 Jahre	33,9	35,3	28,4	32,9
60 Jahre und älter	17,9	19,7	33,9	18,5
<b>Bildung</b>				
Niedrig	26,3	26,0	41,2	34,9
Mittel	36,2	36,5	29,7	31,3
Hoch	37,5	37,5	29,1	33,8
<b>BIK</b>				
Unter 20.000 Einwohner	11,6	11,3	-	12,6
Über 20.000 Einwohner	88,4	88,7	-	87,4
<b>BIK</b>				
Unter 50.000 Einwohner	21,5	20,9	24,4	-
Über 50.000 Einwohner SGTYP2/3/4	28,5	28,7	33,4	-
Über 50.000 Einwohner +SGTYP1	50,1	50,4	42,2	-
<b>Bundesland</b>				
Alte Bundesländer	80,5	80,8	78,9	80,4
Neue Bundesländer (inkl. Berlin)	19,5	19,2	21,1	19,6

Die Tabelle 18 gibt die Anzahl der zur Berechnung notwendigen Iterationen sowie deskriptive Statistiken zu den resultierenden Gewichtungsfaktoren wieder.

<sup>10</sup> Die Zusammenfassung der Variablen erfolgte entsprechend dem Vorgehen in anderen GLES-Komponenten.

**Tabelle 18: Deskriptive Statistiken zu den Gewichtungsfaktoren**

Statistik	wei_mzz	wei_mzoz	wei_onz	wei_onoz
Anzahl Iterationen <sup>11</sup>	4	4	4	4
Mittelwert	1,00	1,00	1,00	1,00
Standardabweichung	0,56	0,54	0,21	0,25
Minimum	0,47	0,43	0,78	0,69
Maximum	4,34	4,12	1,68	1,98
1. Quartil	0,63	0,65	0,87	0,80
Median	0,81	0,83	0,91	0,94
3. Quartil	1,15	1,22	1,25	1,18
Max./Min. <sup>12</sup>	9,23	9,58	2,15	2,87
N	1.019	917	1.019	917

## 6.9. Unveröffentlichte Variablen

Aus datenschutzrechtlichen Gründen können nicht alle erhobenen Variablen zum freien Download zur Verfügung gestellt werden. Generell müssen frei verfügbare Datensätze so aufgebaut sein, dass ein unverhältnismäßiger Aufwand zur De-Anonymisierung betrieben werden müsste. Das heißt, die Datensätze müssen faktisch anonym sein. Um dies zu gewährleisten, wurden einige wenige Variablen aus den Datensätzen entfernt bzw. Ausprägungen zusammengefasst. Die Tabelle 19 gibt eine Übersicht über die unveröffentlichten Variablen in dieser Studie.

**Tabelle 19: Übersicht unveröffentlicher Variablen**

Variable	Label	Verfügbarkeit
t67s	Geburtsland, Vater (offene Angabe)	Im veröffentlichten Datensatz nicht enthalten.
t68s	Geburtsland, Mutter (offene Angabe)	Im veröffentlichten Datensatz nicht enthalten.
t71	Postleitzahl	Im veröffentlichten Datensatz nicht enthalten.
t73s	Geburtsland, Ausland (offene Angabe)	Im veröffentlichten Datensatz nicht enthalten.
bik10	BIK10	Liegt im veröffentlichten Datensatz in recodierter Form vor.

<sup>11</sup> Die Anzahl der Iterationen gibt an, wie viele Anpassungsschritte bei der Gewichtung vorgenommen wurden bis die Abweichung zwischen den gewichteten Istwerten in der Stichprobe und den Sollwerten aus der Referenzverteilung weniger als 0,05 betrug.

<sup>12</sup> Der Wert Max/Min gibt das Verhältnis zwischen dem höchsten und niedrigsten Gewichtungsfaktor an. Im Idealfall werden die Gewichte weder besonders groß noch sehr klein, so dass ein niedriger Wert positiv zu betrachten ist.

Diese Variablen stehen interessierten Nutzer zur Verfügung, je nach Variable in einem Secure Data Center (SDC) bei GESIS (Köln, Mannheim) oder nach Abschluss eines Nutzungsvertrags. Bei Interesse melden Sie sich bitte unter [gles@gesis.org](mailto:gles@gesis.org).

Weiterhin werden Systemvariablen aus dem veröffentlichten Datensatz entfernt, die lediglich zur Durchführung der Online-Befragung relevant sind, aber keine inhaltliche Aussagekraft haben. Darüber hinaus werden offen erhobene Antworten im Normalfall nur als vercodete Variablen veröffentlicht. Bei Interesse an den entfernten Systemvariablen oder den nicht vercodeten offenen Antworten, melden Sie sich bitte ebenfalls unter [gles@gesis.org](mailto:gles@gesis.org).

## 6.10. Fehlende Werte

Fehlende Werte wurden bei allen Variablen nach dem einheitlichen Codierschema der GLES vergeben, welches negative Werte im Bereich -71 bis -99 und in Stata Codes von .a bis .p vorsieht. In den SPSS-Datensätzen sind diese Werte standardmäßig als fehlende Werte definiert. Den Stata -Datensätzen hingegen liegt ein Do-File bei, anhand dessen die fehlenden Werte in die vorgesehenen Missing-Value-Codes recodiert werden können.

**Tabelle 20: Übersicht über die Missing-Value-Codes der GLES**

Code	Stata Missing-Value-Code	Label
-99	.a	keine Angabe
-98	.b	weiß nicht
-97	.c	trifft nicht zu
-96	.d	Split
-95	.e	nicht teilgenommen
-94	.f	nicht in Auswahlgesamtheit
-93	.g	Interview abgebrochen
-92	.h	Fehler in Daten
-86	.i	nicht wahlberechtigt
-85	.j	nicht wählen
-84	.k	keine Erst-/Zweitstimme abgeben
-83	.l	ungültig wählen
-82	.m	keine andere Partei wählen
-81	.n	noch nicht entschieden
-72	.o	nicht einzuschätzen
-71	.p	nicht bekannt

## 6.11. Analyse-Systeme

Der Datensatz ist für die Analyse-Systeme SPSS (IBM) und Stata (StataCorp LP) verfügbar.

**Tabelle 21: Analysesysteme**

Analyse-System	Dateiformat	Versionsbeschränkungen
SPSS	.sav	Version 18 oder höher
Stata	.dta	Version 12 oder höher

---

## 7. Hinweise und Anmerkungen

### 7.1. Errata

Eine aktuelle Errataliste kann über den GESIS Datenkatalog ([www.gesis.org/dbk](http://www.gesis.org/dbk)) abgerufen werden. Dort werden auch alle Änderungen zwischen den einzelnen Versionen der Datensätze dokumentiert.

### 7.2. Versionshistorie

#### Veränderung von Version 1.0.0 zu Version 2.0.0

Die Variablen e0111a-b (EU-Austritt, Maßnahmen), e0112 (Flüchtlinge) und e0113 (Issu-ebatterie, Tarifeinheit) fehlten fälschlicherweise und wurden nun hinzugefügt.

#### Veränderung von Version 2.0.0 zu Version 2.1.0

In der Variable t60 „Religionszugehörigkeit“ gab es eine Fehlcodierung. Die Codes für „keine Angabe“ (-99) und „keine Religionsgemeinschaft“ (9) waren vertauscht. Dies ist nun korrigiert.

Die Berechnung der Gewichte zur Anpassung an die Gesamtbevölkerung wurde auf Grundlage des nun vorliegenden Mikrozensus 2013 aktualisiert.

## Links

Projektseite der GLES: [www.gles.eu](http://www.gles.eu)

Deutsche Gesellschaft für Wahlforschung e. V. (DGfW): <http://www.dgfw.info>

GESIS – Leibniz-Institut für Sozialwissenschaften: <http://www.gesis.org/gles>

Universität Frankfurt: <http://www.uni-frankfurt.de>

Universität Mannheim: <http://www.uni-mannheim.de>

Wissenschaftszentrum Berlin für Sozialforschung gGmbH (WZB): <http://www.wzb.eu>

LINK Institut für Markt- und Sozialforschung: [www.link-institut.de](http://www.link-institut.de)



## Literaturverzeichnis

- AAPOR (2011): Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7. Aufl. The American Association for Public Opinion Research.
- Bergmann, Michael (2011): IPFWEIGHT: Stata module to create adjustment weights for surveys. Statistical Software Components S457353, Boston: Boston College Department of Economics.
- BIK Aschpurwis+Behrens (2001): BIK Regionen: Ballungsräume, Stadtregionen, Mittel-/Unterezentrengebiete. Methodenbeschreibung zur Aktualisierung 2000, [www.bik-gmbh.de/texte/BIK-Regionen2000.pdf](http://www.bik-gmbh.de/texte/BIK-Regionen2000.pdf) [05.11.2011].
- Callegaro, Mario (2013): Paradata in Web Surveys, in: Kreuter, Frauke (Hrsg.), Improving Surveys with Paradata: Analytic Uses of Process Information. (Aufl.) Hoboken, NJ: Wiley, 261-280.
- Callegaro, Mario und Charles Disogra (2008): Computing Response Metrics for Online Panels, in: Public Opinion Quarterly 72, 1008-1032.
- Couper, Mick P. (2000): Usability Evaluation of Computer-Assisted Survey Instruments, in: Social Science Computer Review 18, 384-396.
- DeBell, Matthew, Jon A. Krosnick, Arthur Lupia und Caroline Roberts (2009): User's Guide to the Advance Release of the 2008-2009 ANES Panel Study. Palo Alto, CA und Ann Arbor, MI: Stanford University and University of Michigan.
- Deming, Edwards W. und Frederick F. Stephan (1940): On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known, in: The Annals of Mathematical Statistics 11: 427-444.
- Kaczmirek, Lars (2008): Human-Survey Interaction. Usability and Nonresponse in Online Surveys. Universität Mannheim, Mannheim.
- Krosnick, Jon A. (1991): Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys, in: Applied Cognitive Psychology 5, 213-236.
- Mayerl, Jochen und Dieter Urban (2008): Antwortreaktionszeiten in Survey-Analysen. Messung, Auswertung und Anwendung. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Roßmann, Joss (2010): Data Quality in Web Surveys of the German Longitudinal Election Study 2009. 3rd ECPR Graduate Conference. Dublin.
- Schmitt-Beck, Rüdiger, Hans Rattinger, Sigrid Roßteutscher und Bernhard Weßels (2010): Die deutsche Wahlforschung und die German Longitudinal Election Study (GLES), in: Faulbaum, Frank & Wolf, Christof (Hrsg.), Gesellschaftliche Entwicklungen im Spiegel der empirischen Sozialforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 141-172.