

gesis

Leibniz-Institut
für Sozialwissenschaften

German Longitudinal
Election Study



*Durchgeführt von der
Deutschen Gesellschaft für Wahlforschung und GESIS*

GLES Tracking 2009-2023, Kumulation

ZA6832, Version 2.0.0

Studienbeschreibung

Inhalt

Vorbemerkungen	4
Arbeiten mit GLES-Daten	4
Meldung von Veröffentlichungen.....	4
Zitationsvorschlag.....	5
1 Studieninformationen	6
1.1 Studiennummer.....	6
1.2 Titel der Studie	6
1.3 Version	6
1.4 Erhebungszeitraum.....	6
1.5 Primärforscher*innen	6
1.6 Mitwirkende	7
1.7 Finanzierende Stelle.....	7
2 Studienkonzeption.....	8
2.1 Hintergrund.....	8
2.2 Studienerhebungen	9
3 Hinweise zur Arbeit mit der Kumulation.....	12
3.1 Kumulationsstrategie	12
3.2 Identifikation der Teildatensätze	12
3.3 Variablenbenennung.....	12
3.4 Verschiedene Variablenversionen bei relevanten Frageunterschieden.....	13
3.5 Nicht-Aufnahme von Variablen.....	13
3.6 Erstellung zusätzlicher Variablen	13
3.7 Vereinheitlichung von Missing-Codierungen und Item-Kennzeichnungen	13
4 Stichprobenziehung.....	19
4.1 Untersuchungsgebiet.....	19
4.2 Grundgesamtheit.....	19
4.3 Auswahlgesamtheit.....	19
4.4 Auswahlverfahren.....	19
5 Erhebung.....	20
5.1 Erhebungsverfahren.....	20
5.2 Datenerhebung.....	20
6 Datenaufbereitung	21
6.1 Vorbemerkung	21
6.2 Datenkontrollen	21
6.3 Einheiten im Datensatz.....	21

6.4	Variablen im Datensatz.....	21
6.5	Metadaten.....	22
6.6	Paradaten.....	22
6.7	Laufende Nummer	23
6.8	Zeitunterschreiter*innen.....	24
6.9	Substanzielle Daten	25
6.10	Codierung von offenen Fragen	25
6.11	Kontextvariablen.....	26
6.12	Ost/West.....	27
6.13	Zuordnung von Wahlkreisen.....	27
6.14	Gewichte	28
6.15	Unveröffentlichte Variablen.....	30
6.16	Codierung fehlender Werte	30
7	Hinweise und Anmerkungen.....	31
7.1	Errata.....	31
7.2	Versionshistorie.....	31
	Links.....	32
	Literaturverzeichnis	33

Vorbemerkungen

Erstellung des kumulierten Datensatzes

Der vorliegende Datensatz ist eine Kumulation der Langfrist-Online-Trackings 6 und 8 bis 54 aus den Jahren 2009 bis zum Frühjahr 2023 und ermöglicht damit Wissenschaftlerinnen und Wissenschaftlern die Analyse langfristiger Prozesse der Formierung und des Wandels der öffentlichen Meinung zwischen den Bundestagswahlen 2009, 2013, 2017 und 2021. Dafür wurde die vorherige Kumulation bis T38 um die nachfolgenden Erhebungen bis T54 ergänzt.

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 in der Errata-Liste in der GESIS-Suche (<https://search.gesis.org/>) dokumentiert 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 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 dem aktuellsten Datenrelease zu arbeiten. Zugang zu den Daten erhalten Sie über die GLES Website (www.gesis.org/gles) oder direkt über die Suchfunktion im GESIS-Web (<https://search.gesis.org/>). Neuigkeiten zur Datenveröffentlichung können zudem auf der [GLES Website](#), über den [GLES Newsletter](#) und über [X \(vormals Twitter\)](#) sowie [Bluesky](#) bezogen werden.

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
gles@gesis.org

Zitationsvorschlag

Deutsch (DE)

GLES (2023): GLES Tracking 2009-2023, Kumulation. GESIS, Köln. ZA6832 Datenfile Version 2.0.0, doi: [10.4232/1.14232](https://doi.org/10.4232/1.14232).

Englisch (EN)

GLES (2023): GLES Tracking 2009-2023, Cumulation. GESIS, Cologne. ZA6832 Data file Version 2.0.0, doi: [10.4232/1.14232](https://doi.org/10.4232/1.14232).

1 Studieninformationen

1.1 Studiennummer

ZA6832

1.2 Titel der Studie

GLES Tracking 2009-2023, Kumulation

1.3 Version

2.0.0, 2023-12-19, doi: [10.4232/1.14232](https://doi.org/10.4232/1.14232)

1.4 Erhebungszeitraum

2009-09-18 – 2023-04-21

1.5 Primärforscher*innen

Prof. Dr. Marc Debus – Universität Mannheim

Prof. Dr. Thorsten Faas – Freie Universität Berlin

Prof. Dr. Sigrid Roßteutscher – Goethe-Universität Frankfurt am Main

Prof. Dr. Rüdiger Schmitt-Beck - Universität Mannheim

Prof. Dr. Harald Schoen - Universität Mannheim

Prof. Dr. Bernhard Weißels - Wissenschaftszentrum Berlin für Sozialforschung

Prof. Dr. Christof Wolf - GESIS – Leibniz-Institut für Sozialwissenschaften

1.6 Mitwirkende

Tabelle 1: Rolle, Name und Institution

Rolle	Name	Institution
Project Leader	Roßteutscher, Sigrid	Goethe-Universität Frankfurt am Main
	Blumenberg, Manuela	GESIS – Leibniz-Institut für Sozialwissenschaften
Project Manager	Riebe, Frauke	GESIS – Leibniz-Institut für Sozialwissenschaften
	Schick, Lukas	GESIS – Leibniz-Institut für Sozialwissenschaften
Project Member	Bühler, Pascal	Goethe-Universität Frankfurt am Main
	Dietz, Melanie	Goethe-Universität Frankfurt am Main
	Ehmes, Sven	Goethe-Universität Frankfurt am Main
	Scherer, Philipp	Goethe-Universität Frankfurt am Main
	Stövsand, Lars	Goethe-Universität Frankfurt am Main
	Walke, Lian	GESIS – Leibniz-Institut für Sozialwissenschaften
	Wurthmann, Constantin	GESIS – Leibniz-Institut für Sozialwissenschaften
Data Curator	Jungmann, Nils	GESIS – Leibniz-Institut für Sozialwissenschaften
	Schick, Lukas	GESIS – Leibniz-Institut für Sozialwissenschaften

1.7 Finanzierende Stelle

GESIS – Leibniz-Institut für Sozialwissenschaften

2 Studienkonzeption

2.1 Hintergrund

Die German Longitudinal Election Study (GLES) ist das zentrale wissenschaftliche Umfrageprogramm in Deutschland zur kontinuierlichen Erhebung und Bereitstellung qualitativ hochwertiger Daten für die nationale und internationale Wahlforschung. Die methodisch vielfältigen Erhebungen der GLES ermöglichen es, die politischen Einstellungen und Verhaltensweisen von Wähler*innen und Kandidierenden zu untersuchen. Die GLES wird in enger Zusammenarbeit zwischen der Deutschen Gesellschaft für Wahlforschung (DGfW) und GESIS – Leibniz-Institut für Sozialwissenschaften durchgeführt.

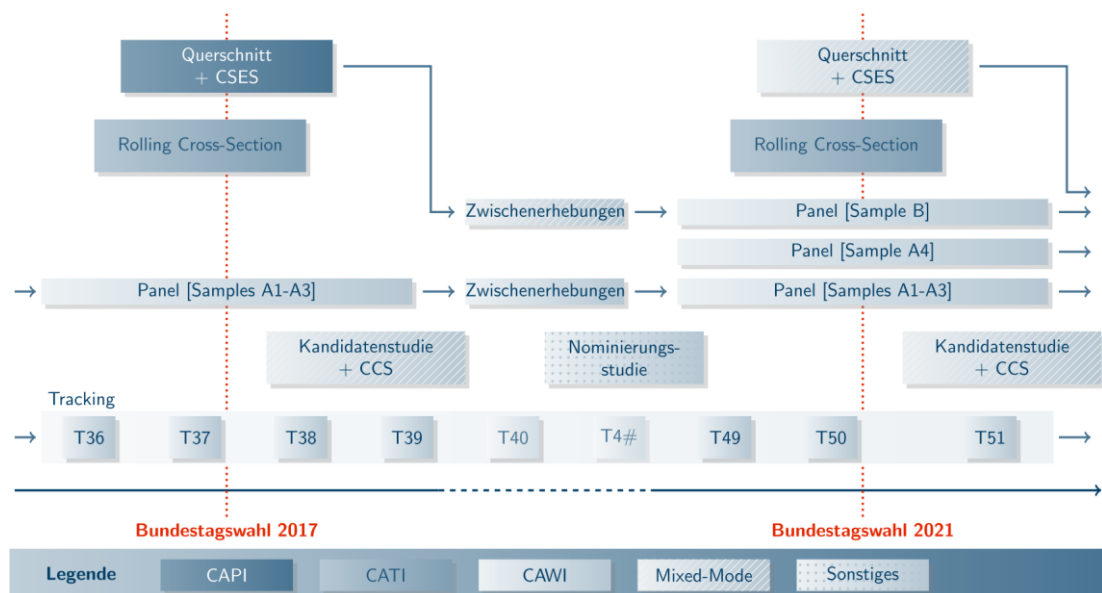


Abbildung 1: Das Design der German Longitudinal Election Study (GLES), 2017-2021

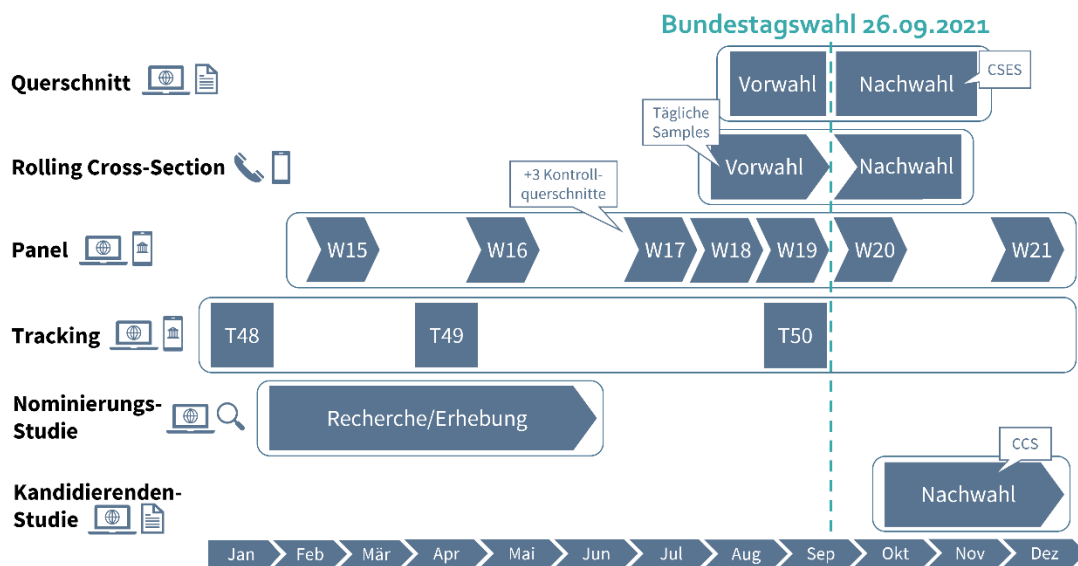


Abbildung 2: Das Design der German Longitudinal Election Study (GLES) 2021

Das GLES Tracking ist dabei auf die Analyse langfristiger Prozesse der Formierung und des Wandels der öffentlichen Meinung ausgerichtet und zielt insbesondere darauf ab, die öffentliche Meinung und ihre Veränderungen zwischen den Bundestagswahlen zu untersuchen (Schmitt-Beck et al. 2010). Hierzu werden je Kalenderjahr in regelmäßigen Abständen zwei bis vier webbasierte Querschnitte mit jeweils etwa 1.000 Befragten aus einem Online-Access-Panel erhoben. Mit Hilfe des GLES Trackings lassen sich die Einstellungen der Wähler*innen zu den wichtigsten politischen und gesellschaftlichen Themen sowie zu den Parteien und ihrem Spitzenpersonal, aber auch zu den Leistungen der Bundesregierung und der Opposition analysieren (Schmitt-Beck et al. 2010).

2.2 Studienerhebungen

Die Durchführung der GLES Trackings begann im Frühjahr vor der Bundestagswahl 2009, damals noch unter dem Studientitel „Langfrist-Online-Tracking“. Seit der 44. Erhebung tragen sie den Titel „GLES Tracking“. Die Grundlage der Studien bilden ein weitgehend konstanter Kernfragebogen zu zentralen Aspekten der Wahl- und politischen Einstellungsforschung und die Abfrage von soziodemographischen Angaben. Bis einschließlich Langfrist-Online-Tracking T38 wurde die Befragung zusätzlich um rotierende thematische Module sowie situationsbezogene variierende Frageblöcke ergänzt. Ab T39 wurde dieses Vorgehen aufgehoben. Eine Übersicht über die Trackings und die jeweiligen thematischen Schwerpunkte bietet die nachfolgende Darstellung (Tabelle 2). In der Kumulation sind T6 und T8 bis T54 enthalten. Wiederholungs- bzw. Nachbefragungen (z. B. T46W und T12NB) wurden nicht kumuliert.

Tabelle 2: Die GLES Trackings

Erhebung	Studien-nummer	Schwerpunkt	Feldbeginn	Feldende
T1	ZA5334	Grundlegende Tests	2009-04-30	2009-05-05
T2	ZA5335	Wählen auf mehreren Ebenen	2009-05-27	2009-06-05
T3	ZA5336	Wirtschaftliche Lage	2009-07-03	2009-07-13
T4	ZA5337	Koalitionen	2009-07-31	2009-08-11
T5	ZA5338	Skalen- und Reihenfolgeexperimente	2009-08-24	2009-09-01
T6	ZA5339	Wahlkampf	2009-09-18	2009-09-27
T7	ZA5340	Nachwahl	2009-09-29	2009-10-08
T7Exp	ZA5340	Experiment zur Bundestagswahl	2009-10-08	2009-10-25
T8	ZA5341	Nachwahl und Netzwerke	2009-12-10	2009-12-20
T9	ZA5342	Wirtschaftliche Lage	2010-04-15	2010-04-23
T10	ZA5343	Wählen auf mehreren Ebenen	2010-06-24	2010-07-05
T11	ZA5344	Psychologische Konstrukte	2010-09-16	2010-09-26
T12	ZA5345	Positionsissues	2010-12-09	2010-12-19
T12NB	ZA5345	Nachbefragung von Abbrechern	2010-12-20	2010-12-30
T13	ZA5346	Wirtschaftliche Lage	2011-03-09	2011-03-19
T13NB	ZA5346	Nachbefragung von Abbrechern	2011-03-21	2011-03-30
T14	ZA5347	Wählen auf mehreren Ebenen	2011-05-23	2011-06-03
T14NB	ZA5347	Nachbefragung von Abbrechern	2011-06-03	2011-06-13
T15	ZA5348	Psychologische Konstrukte	2011-08-24	2011-09-03
T15NB	ZA5348	Nachbefragung von Abbrechern	2011-09-05	2011-09-14
T16	ZA5349	Positionsissues	2011-12-08	2011-12-18
T17	ZA5350	Wirtschaftliche Lage und Wählen auf mehreren Ebenen	2012-05-02	2012-05-15
T18	ZA5351	Netzwerke, psychologische Konstrukte und Koalitionen	2012-09-17	2012-10-01
T19	ZA5719	EU-Krise und politische Partizipation	2013-01-04	2013-01-19
T20	ZA5720	Positionsissues und Europa	2013-05-24	2013-06-08
T21	ZA5721	Wahlkampf	2013-09-06	2013-09-21
T22	ZA5722	Koalitionen und psychologische Konstrukte	2013-11-29	2014-12-13
T23	ZA5723	Wirtschaftliche Lage und politisches Wissen	2014-02-21	2014-03-07
T24	ZA5724	Europa und Europawahl	2014-05-09	2014-05-23
T25	ZA5725	Psychologische Konstrukte	2014-08-29	2014-09-13
T26	ZA5726	Positionsissues, Netzwerke und Koalitionen	2014-11-21	2014-12-05
T27	ZA5727	Psychologische Konstrukte, Wirtschaftliche Lage, Politisches Wissen, Partnerschaft	2015-02-27	2015-03-13
T28	ZA5728	Europa, Institutionenvertrauen, Positionsissues	2015-06-05	2015-06-19
T32	ZA5732	Europa und Positionsissues	2016-06-03	2016-06-17
T33	ZA5733	Psychologische Konstrukte und Umfragen	2016-08-19	2016-09-02
T34	ZA5734	Positionsissues und Koalitionen	2016-12-02	2016-12-16
T35	ZA6815	Netzwerke, Psychologische Konstrukte, Wirtschaftliche Lage und Politisches Wissen	2017-03-17	2017-03-31
T36	ZA6816	Europa und Positionsissues	2017-06-16	2017-06-30
T37	ZA6817	Wahlkampf, politische Einstellungen und Verhaltensweisen	2017-09-12	2017-09-23
T38	ZA6818	Positionsissues, Koalitionen und psychologische Konstrukte	2017-12-01	2017-12-15

T39	ZA6823	Politische Einstellungen und Wahlverhalten	2018-04-20	2018-05-04
T40	ZA6824	Politische Einstellungen und Wahlverhalten	2018-09-14	2018-09-23
T41	ZA6825	Politische Einstellungen und Wahlverhalten	2018-11-30	2018-12-07
T42	ZA6826	Politische Einstellungen und Wahlverhalten	2019-03-01	2019-03-13
T43	ZA6833	Politische Einstellungen und Wahlverhalten	2019-05-10	2019-05-23
T44	ZA6837	Politische Einstellungen und Wahlverhalten	2019-09-13	2019-09-27
T45	ZA6839	Politische Einstellungen und Wahlverhalten	2020-01-10	2020-01-24
T46	ZA6840	Politische Einstellungen und Wahlverhalten	2020-04-30	2020-05-14
T47	ZA6841	Politische Einstellungen und Wahlverhalten	2020-09-11	2020-09-18
T46W	ZA6842	Politische Einstellungen und Wahlverhalten	2020-09-11	2020-09-25
T48	ZA7705	Politische Einstellungen und Wahlverhalten	2021-01-15	2021-01-22
T46W2T47	ZA7706	Politische Einstellungen und Wahlverhalten	2021-01-15	2021-01-25
T49	ZA7707	Politische Einstellungen und Wahlverhalten	2021-04-22	2021-04-30
T50	ZA7708	Politische Einstellungen und Wahlverhalten	2021-09-15	2021-09-24
T51	ZA7709	Politische Einstellungen und Wahlverhalten	2022-02-03	2022-02-14
T52	ZA7710	Politische Einstellungen und Wahlverhalten	2022-06-03	2022-06-09
T53	ZA7711	Politische Einstellungen und Wahlverhalten	2022-10-13	2022-10-21
T54	ZA7712	Politische Einstellungen und Wahlverhalten	2023-04-17	2023-04-21

3 Hinweise zur Arbeit mit der Kumulation

3.1 Kumulationsstrategie

Um eine bestmögliche und nutzerfreundliche Handhabung des kumulierten Datensatzes zu gewährleisten, wurde eine möglichst hohe Kumulationsdichte angestrebt. Fragen wurden zusammengelegt, wann immer dies inhaltlich sinnvoll war. Nur wenn eine Zusammenlegung nicht zu vertreten war, wurden die entsprechenden Fragen getrennt archiviert. Dies hat den Vorteil, dass viele Daten über die Zeit miteinander verglichen werden können.

Bei kleineren Änderungen im Fragetext, bspw. Änderungen der Satzzeichen, Absätze oder auch kleinere inhaltliche Änderungen wurden die Variablen vereinigt. Auch bei kleineren Anpassungen am Text der Ausprägungen oder Skalen sowie bei Kleinständerungen der Fragefilter erfolgte keine Trennung der Variablen. Sobald aber z. B. der Fragestimulus, die Anzahl der Ausprägungen oder die Skalierung verändert wurde, ist eine neue Version der betreffenden Variablen erstellt worden.

Grundsätzlich sollten möglichst viele Daten miteinander vereint werden, um bestenfalls lückenlose Zeitreihen zu erhalten. Fragetexte, Skalen, Ausprägungen und Filterführung sind in der Fragebogen-dokumentation der jeweiligen Erhebung festgehalten. Eine Übersicht über Variablen in den einzelnen Erhebungen ist in der mitveröffentlichten Korrespondenzliste (als Excel zum Download verfügbar) dargestellt.

3.2 Identifikation der Teildatensätze

Zur Unterscheidung der in der Gesamtkumulation enthaltenen Datensätze existiert die **sample** Variable. Sie gibt die jeweilige Stichprobe an und verweist damit gleichsam auf den jeweiligen Teildatensatz.

3.3 Variablenbenennung

In der vorherigen Kumulation (ZA6832) wurden die Variablen in Modul- und Ereignisfragen unterteilt. Da das System der Modul- und Ereignisfragen seit T39 nicht mehr besteht, wurde die Benennung und die Reihenfolge der Variablen am Inhalt der Fragen orientiert. Die inhaltliche Unterteilung ist in Tabelle 3 dargestellt.

Tabelle 3: Übersicht über Variablenbenennung

Technische Variablen	t0001_1 – t0044
Einstellungen und (Wahl-)verhalten	e0001 – e0452d
Medien	m0001_1 – m0038
Soziodemographie	s0001 – s0076

3.4 Verschiedene Variablenversionen bei relevanten Frageunterschieden

Fragen, die demselben Konstrukt angehören, sind mit derselben Variablennummer versehen. Sind geänderte Fragen als dasselbe Konstrukt, nicht aber als identische Frage identifiziert worden, wurde die betreffende Variable in unterschiedliche Versionen differenziert. Damit der gemeinsame Bezug weiterhin hergestellt werden kann, bleibt der erste Teil der Variablenbezeichnung gleich. Hinter einem Unterstrich werden die verschiedenen Versionen dann durch zusätzliche Nummerierungen kenntlich gemacht (Beispiel: Version 1: s0030_1; Version 2: s0030_2). Wenn möglich, wurden die Versionen dem Erhebungszeitraum entsprechend aufsteigend vergeben. Das bedeutet, dass die Version 1 jeweils die älteste Version ist. Die genaue Zuordnung von Ursprungsvariablen zu kumulierten Variablen ist der Korrespondenztabelle zu entnehmen.

3.5 Nicht-Aufnahme von Variablen

Einige Variablen wurden nicht in die Kumulation aufgenommen. Dies betrifft alle Zeitvariablen, einige weitere technische Variablen sowie alle Variablen mit Splits, also solche bei denen die Befragten eines Trackings in mindestens zwei Gruppen aufgeteilt wurden und der Gruppeneinteilung entsprechend unterschiedliche Fragen erhalten haben. In einigen Trackingbefragungen wurde eine Nachbefragung von Abbrechern durchgeführt. Die dazugehörigen Beobachtungen und Variablen wurden nicht in die Kumulation aufgenommen. Von Variablen mit zwei Versionen der Parteicodierung, bei denen in Version A die im Textfeld „andere, und zwar“ genannten Parteien als „andere“ zusammengefasst sind und in Version B die einzelnen Parteien mit einem eigenen Code versehen sind, wurde jeweils nur Version B in den kumulierten Datensatz aufgenommen.

3.6 Erstellung zusätzlicher Variablen

Um die Datenstruktur zu vereinheitlichen, wurden in einigen Fällen neue Variablen erstellt, die in Ursprungsdatensätzen teilweise nicht vorhanden sind. Dies betrifft Fragen mit Mehrfachantworten, d.h. solche, bei denen verschiedene Items durch Markieren einer Checkbox genannt werden können. In manchen Ursprungsdatensätzen besteht eine Variable mit dem Item „keine Angabe“, welche die Ausprägung 1 „genannt“ enthält, wenn keines der Checkboxes markiert wurde, und 0 „nicht genannt“, wenn mindestens eine Checkbox markiert wurde. Für die Ursprungsdatensätze, die diese Variable mit dem „Keine Angabe“-Item nicht enthalten, wurde eine solche erstellt und in die Kumulation aufgenommen.

3.7 Vereinheitlichung von Missing-Codierungen und Item-Kennzeichnungen

Mit dem Ziel, die Nutzung des kumulierten Datensatzes zu vereinfachen, wurden die Codierungen für fehlende Werte weitestgehend vereinheitlicht. Die alten Missing-Codes der GLES 2009 wurden dazu an das seit der GLES 2021 gültige Missing-Schema (siehe Tabelle 16) angepasst.

Außerdem wurden wiederkehrenden Variablen-Items feste Item-Suffixe im Variablennamen zugeordnet. Dies betrifft jene Variablen, bei denen Parteien, Koalitionen, Organisationen, Medien und Politiker*innen die Items darstellen. Die jeweiligen Suffixe sind in Tabelle 4 bis 8 aufgeführt.

Tabelle 4: Übersicht der Suffix-Zuordnung der Parteien

Item-Bezeichnung	Item-Suffix
CDU/CSU	a
CDU	b
CSU	c
SPD	d
FDP	e
Grüne	f
Linke	g
AfD	h
NPD	i
Piraten	j
Freie Wähler	k
Die Partei	l
SSW	m
BIW	n
Trifft nicht zu /nicht wahlberechtigt	v
Keine Partei / Nichtwahl	w
Andere	x
Andere, codiert	xc
Weiß nicht	y
Keine Angabe	z

Tabelle 5: Übersicht der Suffix-Zuordnung der Koalitionen

Item-Bezeichnung	Item-Suffix
Union-Alleinregierung	a
Union, SPD	b
SPD, Grüne	c
Union, FDP	d

SPD, Linke	e
SPD, Linke, Grüne	f
Union, Grüne	g
Union, FDP, Grüne	h
SPD-Alleinregierung	i
SPD, FDP, Grüne	j
Union, AfD	k
SPD, FDP	l
SPD, Grüne, Freie Wähler	m
Union, Freie Wähler	n
Minderheitsregierung Union	o
Union, SPD, FDP	p
Union, AfD, FDP	q
Andere Koalition	s
Keine Angabe	z

Tabelle 6: Übersicht der Suffix-Zuordnung der Politiker*innen

Item-Bezeichnung	Item-Suffix
CDU	
Merkel	ba
Schäuble	bb
de Maizière	bc
von der Leyen	bd
Kristina Köhler	be
Laschet	bf
Merz	bg
Roettgen	bh
CSU	
Seehofer	ca
Friedrich	cb
Guttenberg	cc
Ramsauer	cd

SPD

Steinbrück	da
Gabriel	db
Steinmeier	dc
Müntefering	dd
Nahles	de
Schulz	df
Scholz	dg
Kühnert	dh

FDP

Brüderle	ea
Lindner	eb
Westerwelle	ec
Rösler	ed

Grüne

Trittin	fa
Göring-Eckardt	fb
Künast	fc
Özdemir	fd
Habeck	fe
Baerbock	ff

Linke

Gysi	ga
Lafontaine	gb
Wagenknecht	gc
Bartsch	gd

Piraten

Bernd Schlömer	ja
Thorsten Wirth	jb
Sebastian Nerz	jc

AfD

Bernd Lucke	ha
Frauke Petry	hb

Alexander Gauland	hc
Alice Weidel	hd

Tabelle 7: Übersicht der Suffix-Zuordnung der Sozialen Medien

Item-Bezeichnung	Item-Suffix
Facebook	a
Twitter	b
Instagram	c
Snapchat	d
WhatsApp	e
Xing	f
Threema	g
LinkedIn	h
YouTube	i
Sonstige	j
Gar keine	k
Telegram	l
Signal	m
TikTok	n
Discord	o
StudiVZ	p
Myspace	q
Wer kennt wen	r
Loklisten	s
Google Plus	t
Keine Angabe	z

Tabelle 8: Übersicht der Suffix-Zuordnung der Organisationen

Item-Bezeichnung	Item-Suffix
Gewerkschaft	a
Unternehmer-/Arbeiterverband	b

Berufsvereinigung/-verband	c
Bauern- bzw. Landwirtschaftsverband	d
Religiöse/kirchliche Gruppe	e
Sport-/Hobbyverein	f
Umweltschutzgruppe	g
Globalisierungskritische Gruppe, wie z. B. ATTAC	h
Frauenorganisation	i
Islamkritische Gruppe, wie z. B. PEGIDA	j
Partei	k
CDU/CSU	l
SPD	m
Grüne	n
Linke	o
FDP	p
Andere Partei	q
Andere Organisation	r
Nein, ich bin kein Organisationsmitglied	s
Europafreundliche Gruppe, wie z. B. Pulse of Europe	x
Gegner*innen Corona-Maßnahmen	y
Keine Angabe	z

4 Stichprobenziehung

4.1 Untersuchungsgebiet

Deutschland (DE)

4.2 Grundgesamtheit

Die Grundgesamtheit des GLES Trackings besteht aus Bürger*innen der Bundesrepublik Deutschland, die zum Zeitpunkt der Erhebung für eine Wahl zum Deutschen Bundestag wahlberechtigt waren, d.h. deutsche Staatsangehörige, und die zum Zeitpunkt der Studiendurchführung das 18. Lebensjahr erreicht hatten.

4.3 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*innen eine Chance, für die Befragung des Online-Trackings ausgewählt zu werden. Die Auswahlgesamtheit umfasst daher ausschließlich die ab 18-jährigen aktiven Mitglieder des jeweiligen Access Panels (bis T16: Respondi AG, T17-T32: LINK, T33-T38: forsa omninet, T38-T54: Bilendi&respondi) mit deutscher Staatsangehörigkeit.¹ 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.

4.4 Auswahlverfahren

Die Stichprobe für das Online-Tracking wird durch eine Quotenauswahl aus der Auswahlgesamtheit der im jeweiligen Access Panel (bis T16: Respondi AG, T17-T32: LINK, T33-T38: forsa omninet, T38-T54: Bilendi&respondi) zur Verfügung stehenden aktiven Teilnehmer*innen generiert. Dabei werden nur volljährige Personen mit deutscher Staatsangehörigkeit berücksichtigt. Weiterhin wird darauf geachtet, dass die Panelteilnehmer*innen nur einmal jährlich an einem Online-Tracking der GLES teilnehmen können. Die Teilnehmer*innen werden nach den Merkmalen Geschlecht, Alter und Bildung quotiert.

¹ Ab T17 wurde nochmals präzisiert, dass diese Personen in Deutschland leben und mindestens einmal in der Woche zu privaten Zwecken das Internet nutzen.

5 Erhebung

5.1 Erhebungsverfahren

Selbstaussfüller*innen: CAWI (Computerunterstütztes Web-Interview) mit standardisiertem Fragebogen

5.2 Datenerhebung

Tabelle 9: Überblick Erhebungsinstitute

Erhebung	Erhebungsinstitut
T6, T8-T16	Bamberger Centrum für Empirische Studien (BACES)/Respondi Panel
T17-T31	LINK Institut für Markt- und Sozialforschung, Frankfurt am Main
T32-T38	Forsa.main Marktinformationssysteme GmbH, Frankfurt am Main
T39-T54	GESIS – Leibniz-Institut für Sozialwissenschaften/Bildendi&respondi

6 Datenaufbereitung

6.1 Vorbemerkung

Die Aufbereitung der Daten folgt dem Grundsatz der minimal-invasiven Bearbeitung. Sie erfolgt auf der Grundlage von Syntaxdateien. Es ist daher möglich, alle Schritte zu reproduzieren.

6.2 Datenkontrollen

Alle Daten werden nach Abschluss der Datenerhebung sorgfältig geprüft. Dazu gehört die Überprüfung der Vollständigkeit der Variablen im Datensatz, der Richtigkeit von Variablennamen und -labels, der Vollständigkeit der Wertelabels, der Codierung von Missing Values sowie der Randverteilungen. Die Datensatzkontrolle wird unterstützt durch die Verwendung des frei verfügbaren Stata-Moduls scandata (Kaukal 2016), welches die Variablen in einem Datensatz auf der Grundlage von vorgegebenen Spezifikationen hinsichtlich der Verwendung von Großbuchstaben oder Umlauten, Länge von Variablennamen und Labeln und ungeraden Verteilungen untersucht. Es zeigt Abweichungen von diesen Spezifikationen an und korrigiert diese automatisch – sofern vom Nutzer oder der Nutzerin gewünscht. Darüber hinaus erfolgte eine technische Kontrolle der Filterführungen.

Alle gefundenen Fehler werden während der Datenaufbereitung dokumentiert und korrigiert. Können gefundene Fehler nicht korrigiert werden, wird der Fehlercode -92 „Fehler in Daten“ vergeben.

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 Errata-Liste bei GESIS (https://search.gesis.org/research_data/ZA6832) dokumentiert und, wenn möglich, zeitnah behoben.

6.3 Einheiten im Datensatz

Der Datensatz umfasst 52.341 Einheiten.

6.4 Variablen im Datensatz

Der Datensatz des GLES Trackings beinhaltet verschiedene Typen von Variablen. Sie umfassen Meta-, Para- und substanzielle Daten sowie Kontext- und Gewichtungsvariablen. Tabelle 10 gibt einen Überblick über die Gesamtanzahl der Variablen im Datensatz nach Variablentyp.

Tabelle 10: Detaillierte Übersicht der Anzahl der Variablen im Datensatz

Variablentyp	Anzahl
Metadaten	7
Paradaten	20
Kontextvariablen	38
Gewichte	10
Substanzielle Daten	2111
Gesamt	2186

6.5 Metadaten

Metadaten beschreiben den Datensatz und seine Erstellung und umfassen Variablen, die für die Archivierung und Distribution der Studie erforderlich sind (siehe Tabelle 11).

Tabelle 11: Metadaten

Variable	Erklärung
za_nr	(ZA-)Studiennummer des Datensatzes im vierstelligen Format sowie Studiennummer und -titel als Wertelabel
za_nr_od	(ZA-)Studiennummer des Originaldatensatzes im vierstelligen Format sowie Studiennummer und -titel als Wertelabel
version	Version des Datensatzes, beginnend mit 1.0.0
doi	Digital Object Identifier
field_start	Beginn Erhebungszeitraum
field_end	Ende Erhebungszeitraum
sample	Bezeichnung der Stichprobe

6.6 Paradaten

Paradaten werden als Nebenprodukt des Interviewprozesses erfasst. Sie umfassen z. B. Zeit und Datum der Datenerhebung, Interviewdauer und Informationen über das der befragten Person verwendete Gerät (z. B. Browser-Version, JavaScript-Version). Auch administrative Informationen wie Identifikationsnummern, Informationen zum Teilnahmestatus von Befragungsteilnehmer*innen und Informationen zu deren Zuweisung zu Gruppen auf Grund von inhaltlichen Angaben oder zufallsbasierten Verfahren werden hier als Paradaten behandelt. Die Tabelle 12 gibt einen Überblick über Paradaten im Datensatz.

Tabelle 12: Paradaten

Variable	Erklärung/Label
lfdn	Eine eindeutige fortlaufende Identifikationsnummer für die Teilnehmer*innen an der Befragung. Eine Identifikation mittels der laufenden Nummer von Personen, die an mehreren Erhebungen des Trackings teilgenommen haben, ist nicht möglich, da diese für jede Erhebung neu generiert wird.
lfdn_od	Laufende Nummer Originaldatensätze T6, T8-T54
t0017	Technische Probleme
t0018	quality

t0019	disposition code - Dispositionscode der Befragten, der angibt, ob die Umfrage mit oder ohne Unterbrechung beendet wurde. Hat eine befragte Person die Beantwortung unterbrochen und den Browser oder das Browserfenster des Computers geschlossen, so wurde diese Information gespeichert. Die Unterbrechung der Beantwortung der Umfrage ohne Schließen des Fragebogens im Browser kann nicht direkt festgestellt werden.
t0020	Zeitunterschreiter*innen (10 % der Befragten mit den niedrigsten Indexwerten)
t0021	Zeitunterschreiter*innen-Index
t0022	Status nach 60-Prozent-Regel (Version A)
t0023	Status nach 0,2 quality-Regel (Version B)
t0024	Interviewstatus: Anteil beantworteter Fragen - Anteil der abgeschlossenen Fragen (1 = Vollständig (mehr als 80 Prozent beantwortet), 2 = Partiiell (Zwischen 50 und 80 Prozent beantwortet), 3 = Abbruch (weniger als 50 Prozent beantwortet), 4 = Ausgescreent); Im Tracking werden nur vollständige und partielle Interviews veröffentlicht.
t0025	Interviewstatus: Anteil beantworteter Fragen, prozentual - Anteil der abgeschlossenen Fragen in Prozent
t0026	Interviewdauer (nutzer*innenseitig gemessen)
t0027	Interviewdauer (serverseitig gemessen)
t0028	datetime of survey start - Datum und Uhrzeit des Umfragebeginns, also des Zugriffs auf die erste Seite des Fragebogens
t0029	date of last access - Datum und Uhrzeit des letzten Zugriffs auf die Umfrage
t0030	Nutzung Smartphone - Gibt an, ob die Umfrage an einem Smartphone ausgefüllt wurde, basierend auf den Informationen des User Agents
t0031	Nutzung Tablet - Gibt an, ob die Umfrage an einem Tablet ausgefüllt wurde, basierend auf den Informationen des User Agents
t0032	Genutztes Geraet - Gibt das von der befragten Person genutzte Gerät an, basierend auf den Informationen des User Agents.
t0033	Genutzter Browser - Gibt den von der befragten Person genutzten Webbrowser an, basierend auf den Informationen des User Agents
t0034	Javascript - Enthält das Ergebnis der Prüfung auf JavaScript und gibt ggf. die entsprechende Version an

Anmerkung: Für die Codierung der Paradata-Variablen, die auf Informationen des User Agent zurückgreifen, wurde das frei verfügbare Stata-Modul parseuas eingesetzt (Roßmann & Gummer 2020).

6.7 Laufende Nummer

Die Variable lfdn „Laufende Nummer“ ist eine fortlaufende eindeutige Identifikationsnummer der Befragten im kumulierten Datensatz. Die Variable lfdn_od „Laufende Nummer Originaldatensätze T6, T8-T54“ enthält wiederum die laufenden Nummern aus den Einzeldatensätzen und ermöglicht die eindeutige Identifikation von Befragten auf dieser Basis. Da sich die Kumulation aus 48 Einzeldatensätzen zusammensetzt kann ein Wert der Variable lfdn_od hier auch theoretisch bis zu 48-mal vorkommen. Eine Identifikation von Personen, die an mehreren Erhebungen des Online-Trackings teilgenommen haben, ist mittels der laufenden Nummern nicht möglich, da diese für jede Erhebung neu generiert wurden.

6.8 Zeitunterschreiter*innen

Aufgrund der fehlenden Kontrolle der Befragten durch einen Interviewer oder eine Interviewerin in Online-Befragungen 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*innen. 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 substantielle Antworten hätten abgeben können (vgl. Krosnick 1991).

In der Fachliteratur gibt es keine etablierten Standards für die Identifikation von Zeitunterschreiter*innen. 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 Abschnidekriterium, das nicht unterschritten werden darf (vgl. Mayerl und Urban 2008). Diese Befragten werden dann entweder aus dem Datensatz ausgeschlossen oder durch Markervariablen gekennzeichnet.

In den Online-Trackings 6 und 8 bis 16 wurden zwei Arten von Zeitunterschreiter*innen identifiziert. Als Zeitunterschreiter*innen der Version A (Variable t0022 in der Kumulation) gelten solche Befragte, deren Bearbeitungsdauer weniger als 60 Prozent des Median der Gesamtbefragungsdauer aller Teilnehmer (ohne Personen, die das Interview unterbrachen) aufwies. Diese werden durch die Markervariable t0022 „Zeitunterschreiter*innen nach 60 Prozent-Regel (Version A)“ gekennzeichnet.

Für die Identifikation der Zeitunterschreiter*innen der Version B wurde ein weiterer Index zur „Qualitätsmessung“ gebildet. Dieser Index gibt Auskunft über die durchschnittliche Bearbeitungszeit je Befragungsseite. Anschließend wurde für jede Seite getrennt ermittelt, wo sich bezüglich der Bearbeitungsdauer der Median aller Befragungsteilnehmer*innen befindet. Die individuelle Bearbeitungsdauer wurde mit der durchschnittlichen Bearbeitungsdauer der gesamten Stichprobe in Beziehung gesetzt. Diese Beziehung wird durch eine Zahl ausgedrückt, die in der Variable t0018 „quality“ festgehalten wurde. Beispielsweise bedeutet ein Wert von 0,5, dass der entsprechende Befragte exakt die durchschnittliche Zeit zur Bearbeitung der Fragebogenseiten benötigt hat. Bei einem Wert von 0,25 hat der Befragte halb so lange wie der Durchschnitt zur Bearbeitung benötigt. Teilnehmer*innen, die einen niedrigeren Wert als 0,2 erreichten, wurden als Zeitunterschreiter*innen der Version B identifiziert. Diese werden durch die Markervariable t0023 „Zeitunterschreiter*innen nach 0,2 quality-Regel (Version B)“ gekennzeichnet.

Ab Online-Tracking 17 werden Zeitunterschreiter*innen anhand des von Roßmann (2015) entwickelten Algorithmus identifiziert (siehe auch Roßmann 2010). Hierfür wird ein Zeitunterschreiter*innen-Index gebildet (Variable t0021 in der Kumulation), der sowohl die Antwortdauer der Befragten auf allen Bildschirmseiten der Befragung als auch die Interviewdauer je Befragtem einbezieht 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*innen werden alle Befragten im unteren 10%-Perzentil der Verteilung des Summenindex gekennzeichnet. Die Kennzeichnung von Zeitunterschreiter*innenn ist in der Kumulation in der Variable t0020 „Zeitunterschreiter*innen (10% der Befragten mit den niedrigsten Indexwerten)“ enthalten. Diese Markervariable kann zum

Ausschluss der Fälle aus Analysen als auch zur Durchführung von Robustheitsanalysen verwendet werden.

6.9 Substanzielle Daten

Eine Übersicht der substanziellen Variablen ist der mitveröffentlichten Korrespondenzliste zu entnehmen. Substanzielle Variablen sind alle Variablen zu Einstellungen und (Wahl-)verhalten, Medien und Soziodemographie.

6.10 Codierung von offenen Fragen

Die Codierung offener und halboffener Fragen 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>). Für die Wellen T39 bis T52 wurden für die Variablen e0087s und e0090s die Codierung des Datensatzes „GLES Codierung des wichtigsten politischen Problems (2018-2022)“ (ZA7957)² herangezogen.

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. Die meisten dieser String-Variablen sind jedoch nicht Teil des Stata oder SPSS-Datensatzes, sondern werden aus Gründen der besseren Lesbarkeit separat als .csv-Datei veröffentlicht. Über die laufende Nummer (lfdn) lassen sich diese Variablen zudem problemlos an den Datensatz anspielen. Eine Übersicht dieser Variablen bietet Tabelle 13.

Tabelle 13: Offene Fragen in separater .csv-Datei

Variable	Label
e0292cs	Regierungserwartung (Version C)
e0296bs	Themenwahrnehmung: CDU
e0296cs	Themenwahrnehmung: CSU
e0296ds	Themenwahrnehmung: SPD
e0296es	Themenwahrnehmung: FDP
e0296gs	Themenwahrnehmung: DIE LINKE
e0296fs	Themenwahrnehmung: GRÜNE
e0087s	Wichtigstes Problem
e0090s	Zweitwichtigstes Problem

² GLES (2023). GLES Codierung des wichtigsten politischen Problems (2018-2022). GESIS, Köln. ZA7957 Datenfile Version 1.0.0, <https://doi.org/10.4232/1.14200>

6.11 Kontextvariablen

Kontextvariablen geben Auskünfte über den regionalen Kontext der Befragten, z. B. den Wahlkreis und die Gemeinde, in welchen Befragte leben. Tabelle 14 listet die in der Kumulation enthaltenen Kontextvariablen auf.

Tabelle 14: Übersicht Kontextvariablen

Variable	Label
t0001_1	Wahlkreis, 1. Phase
t0001_2	Wahlkreis, 2. Phase
t0001_3	Wahlkreis, 3. Phase
t0001_4	Wahlkreis, 4. Phase
t0001_5	Wahlkreis, 5. Phase
t0002	Ost/West (Zuordnung von Berlinern zu Ostdeutschland)
t0007_1	Wahlkreiszuordnung nicht eindeutig: 1. Möglichkeit, 1. Phase
t0007_2	Wahlkreiszuordnung nicht eindeutig: 1. Möglichkeit, 2. Phase
t0007_3	Wahlkreiszuordnung nicht eindeutig: 1. Möglichkeit, 3. Phase
t0007_4	Wahlkreiszuordnung nicht eindeutig: 1. Möglichkeit, 4. Phase
t0007_5	Wahlkreiszuordnung nicht eindeutig: 1. Möglichkeit, 5. Phase
t0008_1	Wahlkreiszuordnung nicht eindeutig: 2. Möglichkeit, 1. Phase
t0008_2	Wahlkreiszuordnung nicht eindeutig: 2. Möglichkeit, 2. Phase
t0008_3	Wahlkreiszuordnung nicht eindeutig: 2. Möglichkeit, 3. Phase
t0008_4	Wahlkreiszuordnung nicht eindeutig: 2. Möglichkeit, 4. Phase
t0008_5	Wahlkreiszuordnung nicht eindeutig: 2. Möglichkeit, 5. Phase
t0009_1	Wahlkreiszuordnung nicht eindeutig: 3. Möglichkeit, 1. Phase
t0009_2	Wahlkreiszuordnung nicht eindeutig: 3. Möglichkeit, 2. Phase
t0009_3	Wahlkreiszuordnung nicht eindeutig: 3. Möglichkeit, 3. Phase
t0009_4	Wahlkreiszuordnung nicht eindeutig: 3. Möglichkeit, 4. Phase
t0009_5	Wahlkreiszuordnung nicht eindeutig: 3. Möglichkeit, 5. Phase
t0010_2	Wahlkreiszuordnung nicht eindeutig: 4. Möglichkeit, 2. Phase
t0010_3	Wahlkreiszuordnung nicht eindeutig: 4. Möglichkeit, 3. Phase
t0010_4	Wahlkreiszuordnung nicht eindeutig: 4. Möglichkeit, 4. Phase
t0010_5	Wahlkreiszuordnung nicht eindeutig: 4. Möglichkeit, 5. Phase
t0011_2	Wahlkreiszuordnung nicht eindeutig: 5. Möglichkeit, 2. Phase
t0011_3	Wahlkreiszuordnung nicht eindeutig: 5. Möglichkeit, 3. Phase
t0011_4	Wahlkreiszuordnung nicht eindeutig: 5. Möglichkeit, 4. Phase
t0011_5	Wahlkreiszuordnung nicht eindeutig: 5. Möglichkeit, 5. Phase
t0012_4	Wahlkreiszuordnung nicht eindeutig: 6. Möglichkeit, 4. Phase
t0012_5	Wahlkreiszuordnung nicht eindeutig: 6. Möglichkeit, 5. Phase
t0013_4	Wahlkreiszuordnung nicht eindeutig: 7. Möglichkeit, 4. Phase
t0013_5	Wahlkreiszuordnung nicht eindeutig: 7. Möglichkeit, 5. Phase
t0014_4	Wahlkreiszuordnung nicht eindeutig: 8. Möglichkeit, 4. Phase
t0014_5	Wahlkreiszuordnung nicht eindeutig: 8. Möglichkeit, 5. Phase
t0015_4	Wahlkreiszuordnung nicht eindeutig: 9. Möglichkeit, 4. Phase

t0015_5	Wahlkreiszuordnung nicht eindeutig: 9. Möglichkeit, 5. Phase
t0016_4	Wahlkreiszuordnung nicht eindeutig: 10. Möglichkeit, 4. Phase

6.12 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*innen zu Ostdeutschland zugeordnet.

6.13 Zuordnung von Wahlkreisen

Die Zuordnung zu einem Wahlkreis beruht auf der Wahlkreiseinteilung zur entsprechenden Bundestagswahl und wurde unter Verwendung der von den Befragten genannten Postleitzahl ihres Hauptwohnsitzes vorgenommen.

Die Zuordnung der Postleitzahlen zu Wahlkreisen wurde mittels des frei verfügbaren Stata-Ados `plztowknr` (Glinitzer et al. 2021) vorgenommen. Damit kann eine Zuordnung von Postleitzahlen zu Wahlkreisen gemäß den Wahlkreiseinteilungen bei den Bundestagswahlen 1998, 2002, 2005, 2005, 2009, 2013, 2017 und 2021 vorgenommen werden. Da in der Bundesrepublik Deutschland keine offiziellen Listen für die Zuordnung von Postleitzahlen zu Wahlkreisen verfügbar sind, beruhen die Zuordnungstabellen auf verschiedenen administrativen Datenbeständen. Die Zuordnungstabellen für die Wahlkreiseinteilung zu den Bundestagswahlen in den Jahren 2009, 2013, 2017 und 2021 wurden im Rahmen der Datenaufbereitung der GLES erstellt. Sie beruhen auf dem vom Statistischen Bundesamt (www.destatis.de) herausgegebenen Gemeindeverzeichnis sowie der Wahlkreissuche des Deutschen Bundestages (www.bundestag.de/abgeordnete/wahlkreise/).

Bei der Verwendung der Wahlkreisinformationen ist zu beachten, dass einzelnen Postleitzahlen in großstädtischen Gebieten zwischen zwei und vier Wahlkreisen zugeordnet werden können, d.h. eine eindeutige Zuordnung der Postleitzahl zu einem Wahlkreis ist auf Grundlage der vorhandenen Informationen nicht immer möglich. Die Variablen `t0001_1 bis t0001_5` „Wahlkreis, 1.-5. Phase“ enthält die Angabe zum Wahlkreis der befragten Person, wenn eine eindeutige Zuordnung möglich ist. Ist dies nicht der Fall, so sind in den Variablen `t0007_1 bis t0016_4` „Wahlkreis nicht eindeutig zuordenbar: 1.-10. Möglichkeit“ die Nummern der Wahlkreise erfasst, denen die befragte Person anhand der von ihr genannten Postleitzahl potenziell zugeordnet werden kann. In einzelnen Fällen ist eine Zuordnung zu einem oder mehreren Wahlkreisen nicht möglich, da eine Angabe zur Postleitzahl fehlt oder die von den Befragten genannte Postleitzahl nicht existiert.

6.14 Gewichte

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 Einstellungs- und Verhaltensvariablen besteht. Bei der Verwendung von Gewichten ist Vorsicht geboten: Für jede einzelne Analyse müssen Forschende selbst entscheiden, ob eine der angegebenen Gewichtungsvariablen für den Zweck ihrer Analyse angemessen ist.

Bei der Erstellung der Anpassungsgewichte für das GLES Tracking wurde bei jeder Welle jeweils auf die aktuellsten verfügbaren Randverteilungen des Mikrozensus³ und des (N)Onliner Atlas zurückgegriffen. In die 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.

Die Gewichte wurden an soziodemographische und regionalstrukturelle Merkmale angepasst: Geschlecht, Alter, Bildung und Zugehörigkeit zu West- bzw. Ostdeutschland (inkl. Berlin). Tabelle 15 bietet eine Übersicht über die sozial- und regionalstrukturellen Gewichtungsvariablen in der Kumulation.

³ Bei den vom Mikrozensus ausgewiesenen Werten handelt es sich um absolute Personenzahlen nach gebundener Hochrechnung.

Tabelle 15: Übersicht über die sozial- und regionalstrukturellen Gewichtungsvariablen

Variable	Anpassung an Randverteilungen des ...	Berücksichtigung Zeitunterschreiter*innen
t0003_1	<ul style="list-style-type: none"> • Mikrozensus 2006 (T6, T8, T9, T11, T12) 	ja
t0003_2	<ul style="list-style-type: none"> • Mikrozensus 2009 (T10, T13, T14, T15, T16) • Mikrozensus 2009 (T17, T18, T19, T20) • Mikrozensus 2012 (T21, T22, T23, T24, T25, T26, T27) • Mikrozensus 2013 (T28, T29, T30, T31, T32, T33, T34, T35, T36, T38) • Mikrozensus 2017 (T37, T39 – T47) • Mikrozensus 2019 (T48 – T54) 	ja
t0004_1	<ul style="list-style-type: none"> • Mikrozensus 2006 (T8, T9, T11, T12) • Mikrozensus 2009 (T10, T13, T14, T15, T16) 	nein (Ausschluss Version A)
t0004_2	<ul style="list-style-type: none"> • Mikrozensus 2006 (T8, T9, T11, T12) • Mikrozensus 2009 (T10, T13, T14, T15, T16) 	nein (Ausschluss Version B)
t0004_3	<ul style="list-style-type: none"> • Mikrozensus 2009 (T17, T18, T19, T20) • Mikrozensus 2012 (T21, T22, T23, T24, T25, T26, T27) • Mikrozensus 2013 (T28, T29, T30, T31, T32, T33, T34, T35, T36, T38) • Mikrozensus 2017 (T37 – T47) • Mikrozensus 2019 (T48, T49, T51, T52, T53, T54) 	nein
t0005_1	<ul style="list-style-type: none"> • (N)Onliner 2008 (T6, T8, T9, T11, T12) • (N)Onliner 2009 (T10, T13) • (N)Onliner 2010 (T14, T15, T16) 	ja
t0005_2	<ul style="list-style-type: none"> • (N)Onliner 2012 (T17, T18, T19, T20, T21, T22) • (N)Onliner 2014 (T23, T24, T25, T26, T27, T28, T29, T30, T31, T32, T33, T34, T35, T36, T38) • (N)Onliner 2016 (T37, T39 – T49) 	ja
t0006_1	<ul style="list-style-type: none"> • (N)Onliner 2008 (T8, T9, T11, T12) • (N)Onliner 2009 (T10, T13) • (N)Onliner 2010 (T14, T15, T16) 	nein (Ausschluss Version A)
t0006_2	<ul style="list-style-type: none"> • (N)Onliner 2008 (T8, T9, T11, T12) • (N)Onliner 2009 (T10, T13) • (N)Onliner 2010 (T14, T15, T16) 	nein (Ausschluss Version B)
t0006_3	<ul style="list-style-type: none"> • (N)Onliner 2012 (T17, T18, T19, T20, T21, T22) • (N)Onliner 2014 (T23, T24, T25, T26, T27, T28, T29, T30, T31, T32, T33, T34, T35, T36, T38) • (N)Onliner 2016 (T37, T39 – T49) 	nein

6.15 Unveröffentlichte Variablen

Bei der Durchführung webbasierter Umfragen werden Informationen gesammelt, die für die Durchführung der Umfrage relevant sind, aber keine inhaltliche Aussagekraft haben. Offen erhobene Antworten werden im Normalfall nur als vercodete Variablen veröffentlicht. Diese Variablen werden nicht veröffentlicht, können aber angefordert werden. Einige erhobene Daten können zudem aufgrund von Datenschutzbestimmungen nicht zur Verfügung gestellt werden (z. B. Postleitzahl). Diese Variablen können unter bestimmten Bedingungen genutzt werden. Diese sensiblen Daten der Einzeldatensätze wurden jedoch nicht kumuliert. Eine Übersicht über die unveröffentlichten Variablen der Einzeldatensätze, sind in den jeweiligen Studienbeschreibungen zu finden. Wenden Sie sich für Anfragen bitte an gles@gesis.org.

6.16 Codierung fehlender 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 .r vorsieht. Den Datensätzen liegen Syntaxfiles bei, anhand denen die fehlenden Werte in die vorgesehenen Missing-Value-Codes recodiert werden können.

Tabelle 16: 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
-91	.i	Modus
-86	.j	nicht wahlberechtigt
-85	.k	nicht wählen
-84	.l	keine Erst-/Zweitstimme abgeben
-83	.m	ungültig wählen
-82	.n	keine andere Partei wählen
-81	.o	noch nicht entschieden
-73	.r	Mehrfachnennungen
-72	.p	nicht einzuschätzen
-71	.q	nicht bekannt

7 Hinweise und Anmerkungen

7.1 Errata

- Vereinzelt waren Systemmissings in Variablen vorhanden, die nun auf -94 „nicht in Auswahlgesamtheit“ rekodiert wurden.
- Bei Variable e0005b_1 (Variablenname v1-1-0, k0090bb_2) wurde eine Angabe mit dem Code -87 auf den Code 800 rekodiert.
- Bei Variablen s0012ks (Variablenname v1-1-0, k0560ks) und s0051 (Variablenname v1-1-0, m4954) wurden Systemmissings rekodiert, entsprechend der Filterführung im Fragebogen.
- Bei e0297e (Name v1-1-0, e2571d) wurden Systemmissings auf -92 „Fehler in Daten“ rekodiert.
- Für Sample T45 wurde die GLES Codierung des wichtigsten politischen Problems (ZA7957, Version 1-0-0) auf -92 „Fehler in Daten“ recodiert, da die Codierung den Datensatz Version 1-0-0 des Samples T45 beinhaltet und die Kumulation den Datensatz Version 2-0-0. Zwischen Version 1-0-0 und 2-0-0 unterscheiden sich die Anzahl der Fälle und entsprechend die Variable lfdn.
- Bei einem Fall in Sample T49 wurde e0090_c2 und e0090_c3 auf -99 „keine Angabe“ recodiert.

Aktuelle Erratalisten der Einzeldatensätze können über den GESIS Datenkatalog (<https://search.gesis.org>) abgerufen werden. Dort werden auch alle Änderungen zwischen den einzelnen Versionen der Datensätze dokumentiert.

7.2 Versionshistorie

Änderungen zwischen Version 1.1.0 und Version 2.0.0

- Variablennamen wurden geändert (siehe Tabelle 2)
- Anpassungen von Variablenlabels als auch vereinzelt Wertelabels
- Anpassungen von Versionierungen von Variablen
- Überarbeitung von Codierungen für Parteien, Organisationen, Medien, Koalitionen, Politiker*innen (siehe Tabelle 4 bis Tabelle 8)

Veränderung von Version 1.0.0 zu Version 1.1.0

- Redaktionelle Anpassungen in Variablenlabels
- Rekodierung der Systemmissings in k0570_1 bis k0570_4 in -94 „nicht in Auswahlgesamtheit“
- Veröffentlichung der Korrespondenzliste

Links

Projektseite der GLES: www.gles.eu

Deutsche Gesellschaft für Wahlforschung e. V. (DGfW): www.dgfw.info

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

Universität Frankfurt: www.uni-frankfurt.de

Universität Mannheim: www.uni-mannheim.de

Statistisches Bundesamt: www.destatis.de

Wahlkreissuche des Deutschen Bundestages: www.bundestag.de/abgeordnete/wahlkreise/

Literaturverzeichnis

Glinitzer, K., Gummer, T., Kaukal, M., Roßmann, J., & Jungmann, N. (2021). PLZTOWKNR: Stata module to translate German zip codes into electoral districts (Version: 2.0). Chestnut Hill, MA: Boston College Department of Economics. Retrieved from: <https://econpapers.repec.org/software/boc-bocode/s458480.htm>

Kaukal, M. (2016). SCANDATA: Stata Module to Scan a Dataset for Specified Characteristics. Chestnut Hill, MA: Boston College Department of Economics. Retrieved from <https://ideas.repec.org/c/boc/bocode/s458186.html>

Krosnick, J. (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *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, J. (2010). Data Quality in Web Surveys of the German Longitudinal Election Study 2009. Paper presented at the 3rdECPR Graduate Conference, Dublin.

Roßmann, J. (2015). SPEEDERGLES: Stata Module to Compute the GLES Response Speed Index (Version 1.0). Chestnut Hill, MA: Boston College Department of Economics. Retrieved from <https://econpapers.repec.org/software/bocbocode/s457996.htm>

Roßmann, J., & Gummer, T. (2020). PARSEUAS: Stata module to extract detailed information from user agent strings (Version: 1.4). Chestnut Hill, MA: Boston College Department of Economics. Retrieved from <https://econpapers.repec.org/software/bocbocode/s457937.htm>

Schmitt-Beck, R., Rattinger, H., Roßteutscher, S., & Weißels, B. (2010). Die deutsche Wahlforschung und die German Longitudinal Election Study (GLES). In F. Faulbaum & C. Wolf (Hrsg.), *Gesellschaftliche Entwicklungen im Spiegel der empirischen Sozialforschung* (S. 141-172). Wiesbaden: VS Verlag für Sozialwissenschaften.