

## Zwischen Tür und Tab – Die Bedeutung des Befragungsmodus für die Ergebnisse von Wahlumfragen zur Bundestagswahl 2021

Olaf Jandura<sup>1</sup>/Hella de Haas<sup>2</sup>

### 1. Forschungsinteresse

Die Präsidentschaftswahl 2024 in den USA hat erneut gezeigt, dass Wahlumfragen stark vom tatsächlichen Wahlausgang divergieren können. Das Beispiel des „Iowa-Irrtum“ der bekannten Meinungsforscherin J. Ann Selzer illustriert diese Aussage sehr gut. Selzer sah im Swing State Iowa Kamala Harris 3 Prozentpunkte vor Donald Trump, die Auszählung der Wählerstimmen erbrachte hingegen einen Überhang von 13 Prozentpunkten für Donald Trump. Selzer verspricht daraufhin: „We'll look at data to try to understand Iowa Poll miss“ (Pfannenstiel, 2024), allerdings kann dieser Irrtum nicht nur Befragungsinstitute in Verruf bringen, sondern auch die Funktionen von Wahlumfragen für die Demokratie beeinträchtigen: „Wer enttäuscht ist, fühlt sich schnell auch getäuscht“ (Dieckmann, 2024, Herv. i.O.).

Dieses Gefühl der Enttäuschung und die gefühlte Täuschung lassen sich auch aus der Reaktion Sahra Wagenknechts auf das Abschneiden des BSW unmittelbar nach der Bundestagswahl im Februar ablesen. Wagenknecht macht für das Scheitern des BSW unter anderem das Meinungsforschungsinstitut Forsa verantwortlich. Forsa habe durch zu niedrige Umfragewerte (3 % am 21.02.2025, der letzten Veröffentlichung vor der Wahl) den Eindruck vermittelt, eine Stimme für die BSW sei eine verlorene Stimme, und dadurch den Einzug der erst 2024 gegründeten Partei in den Bundestag verhindert (Sieben, 2025).

Angesichts der potentiellen Wirkungen von Wahlumfragen auf Journalisten und Wähler ist eine kritische Auseinandersetzung mit ihrer Güte unerlässlich, um sie als valide Grundlage für die Einschätzung der Meinungsverteilung in der Bevölkerung nutzen zu können.

Unterschiede zwischen Wahlumfragen sowie Abweichungen vom Wahlergebnis sind multikausal erklärbar (Zukin, 2005). Ein relevanter Erklärungsfaktor, der Unterschied zwischen verschiedenen Befragungsmodi, insbesondere zwischen face-to-face Befragungen und Onlinebefragungen, steht im Zentrum dieses Beitrages.

---

<sup>1</sup> Prof. Dr. Olaf Jandura ist Leiter des Forschungsschwerpunkts Kommunikationsforschung am Fachbereich Wirtschaftswissenschaften der Hochschule Düsseldorf und außerplanmäßiger Professor an der Heinrich-Heine-Universität Düsseldorf.

<sup>2</sup> M.A. Hella de Haas ist Wissenschaftliche Mitarbeiterin im hochschulinternen Projekt „Generative AI and Empirical Methods“ am Fachbereich Wirtschaftswissenschaften der Hochschule Düsseldorf.

Face-to-face Befragungen galten lange Zeit als Standard in der Wahlforschung. Seit den 1980er Jahren ist ihr Anteil jedoch mit dem Aufkommen alternativer, häufig zeit- und kosteneffizienterer Befragungsmodi deutlich zurückgegangen. Zunächst wurden sie durch Telefonbefragungen abgelöst, heute stellen Onlinebefragungen den dominierenden Standard dar und in Zukunft könnten KI-generierte synthetische Daten (Argyle et al., 2023; Bisbee et al., 2024) diese Rolle übernehmen.

Die derzeitige Relevanz von Onlinebefragungen lässt sich aus den Statistiken des Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM) herauslesen. Im Jahr 2023 wurden zwei von drei quantitativen Interviews (66 %) online durchgeführt. Im Ranking folgen Telefoninterviews (15 %), face-to-face Interviews (11 %), schriftliche Interviews (4 %) und In-App-Interviews, ebenfalls mit 4 % (ADM e.V., 2023). Die Relevanz von Onlineinterviews ergibt sich aus der Kombination methodischer und forschungsökonomischer Vorteile: Sie ermöglichen eine selbstadministrierte Durchführung, die Einbindung audiovisueller Stimuli, eine klare Filterführung sowie einen geringen Zeitaufwand für die Befragten und Forschenden. Zudem entstehen bei Onlineinterviews im Vergleich zu anderen Befragungsmodi nur ein Bruchteil der Kosten, sodass diese weit günstiger am Markt angeboten werden können (u.a. Baur & Florian, 2009).

In der Etablierungsphase von Onlineinterviews in den frühen 2000er Jahren belegen zahlreiche methodische Vergleichsstudien systematische Unterschiede zwischen den Teilnehmern von Onlinebefragungen und denjenigen klassischer Befragungen (z.B. persönlich-mündlich oder telefonisch) – insbesondere in Hinblick auf soziodemographische Merkmale, politisches Wissen, politische Einstellungen und wahlbezogenes Verhalten (u.a. Duffy et al., 2005; Heerwegh, 2009). Ähnlich wie Telefonbefragungen in den 1980er Jahren haftete ihnen das Image an, sie seien „quick and dirty“ (Maurer & Jandura, 2009, S. 71).

Die zunehmende Durchdringung des Internets in der Gesellschaft sowie eine methodische Grundlagenforschung sollten dazu beitragen, dass Onlinebefragungen die an eine qualitativ hochwertige empirische Sozialforschung gestellten Anforderungen immer besser erfüllen. In ihrem Methodenvergleich zwischen Online-Umfragen und persönlich-mündlichen bzw. telefonischen Umfragen kommen Bieber & Bytzek (2012, S. 212) zu dem Schluss, „[...] dass die Berechnung von Modellen der Wahlbeteiligung und der Wahlentscheidung unabhängig von der Datenbasis sehr ähnliche Ergebnisse liefert. Neben den vielfältigen weiteren Nutzungsmöglichkeiten von Online-Umfragen [...] stellen sie folglich auch für die Untersuchung von Einflussbeziehungen im Bereich des Wahlverhaltens eine nützliche Datenquelle dar“. Folgt man diesem Urteil, wären weitere Methodenvergleiche obsolet.

Jedoch finden sich auch in der aktuellen Literatur Hinweise darauf, dass sich die Ergebnisse verschiedener Befragungsmodi systematisch voneinander unterscheiden (u.a. Bennink et al., 2013). Zur Erklärung dieser Unterschiede werden insbesondere Rekrutierungsbias von Onlinestichproben, etwa aufgrund niedriger Rücklaufquoten von Menschen über 65 Jahren (u.a. Braekman et al., 2022) sowie modusbedingte Antworteffekte herangezogen, wie eine erhöhte Tendenz zu „weiß nicht“- oder fehlenden Antworten und eine reduzierte Skalensvarianz (u.a. Heerwegh & Loosveldt, 2008). Vor diesem Hintergrund wird ein Moduswechsel bei Trendbefragungen hin zu kostengünstigeren Onlineformaten kritisch bewertet. Bennink et al. (2013, S. 336) schlussfolgern: „This research has demonstrated that switching survey modes from one wave to another does not come without the penalty of losing comparability“. Andere Studien versuchen die Art der Fragen zu klassifizieren, bei denen Unterschiede in der Beantwortung festgestellt werden können. Strabac und Aalberg (2011) zeigen, dass modusbedingte Effekte insbesondere bei Fragen mit direktem Onlinebezug auftreten, was auf Differenzen im Grad der digital literacy (Gilster, 1997) der Befragten als potentielle Einflussgröße hinweist.

Diese sich teilweise widersprechenden Befunde sind Anlass genug, Ergebnisse verschiedener Befragungsmodi weiterhin systematisch miteinander zu vergleichen und den Forschungsfragen nachzugehen: (1) Worin unterscheiden sich die Ergebnisse von face-to-face Befragungen und Onlinebefragungen vor der Bundestagswahl 2021 in Deutschland? Und (2) Wie lassen sich diese Unterschiede begründen?

Die Argumentation im Beitrag knüpft an das eingangs skizzierte Wirkungspotential von Wahlumfragen an und unterstreicht deren Relevanz für Journalisten und Wähler. Daran schließt sich eine Analyse der Unterschiede zwischen den Wahlumfragen vor der Bundestagswahl 2021 sowie die Einordnung möglicher Ursachen an. Sodann werden das methodische Vorgehen und die zentralen Ergebnisse dargestellt und ein Fazit und Ausblick beschließen den Beitrag.

## **2. Relevanz von Wahlumfragen für Wähler und Journalisten**

In verschiedenen Erklärungsansätzen des Wahlverhaltens wird der Einfluss von subjektiven Urteilen und demoskopischen Befunden auf die Wahlbeteiligung und das Wahlverhalten untersucht. Die Forschung zu den Wirkungen demoskopischer Befunde geht dabei von folgenden aufeinander aufbauenden Annahmen aus: (1) Die Erwartungen an das Abschneiden der präferierten Partei oder präferierter Koalitionsoptionen haben einen Einfluss auf das eigene Wahlverhalten. (2) Diese Erwartungen resultieren primär aus der Kenntnis demoskopischer Befunde, die über die Massenmedien vermittelt werden. Und (3) Die demoskopi-

schen Befunde sind für die Wähler die valideste und wichtigste Quelle zur Beobachtung des Meinungsklimas (Reinemann et al., 2011).

Aus dieser Kausalitätskette lassen sich Einflüsse auf die Wählerschaft und Einflüsse auf die Journalisten und ihre Berichterstattung ableiten, die nun folgend beginnend mit den Effekten von Wahlumfragen auf das *Wahlverhalten* kurz dargelegt werden. Hierbei wird zwischen Effekten auf die Wahlbeteiligung und Effekten auf die Stimmabgabe unterschieden. Zu den Beteiligungseffekten zählt der Mobilisierungseffekt, der besonders dann auftritt, wenn es sich um eine Richtungswahl handelt oder wenn ein enges Wahlergebnis prognostiziert wird, was die Wahlbeteiligung erhöht (Newton, 1999). Das Gegenteil, eine fehlende Mobilisierung, wird als Lethargieeffekt bezeichnet: Wird ein klares Wahlergebnis erwartet, entsteht bei Teilen der Wählerschaft der Eindruck, dass ihre Stimme keinen entscheidenden Einfluss habe, was in einer niedrigen Wahlbeteiligung, insbesondere in bestimmten Wählergruppen, resultieren kann (Bendor et al., 2003). Die Intensität der Berichterstattung über Wahlumfragen kann dabei die Stärke der Effekte beeinflussen (Bursztyn et al., 2017).

Als Effekte auf die Stimmabgabe werden alle Effekte bezeichnet, die das Wahlverhalten verändern können. Empirisch am besten belegt ist der Bandwagon-Effekt: Er beschreibt die Tendenz von Wählern, Parteien zu unterstützen, die in Umfragen führen. Dieser Effekt verstärkt sich, wenn der Vorsprung führender Parteien in den Umfragen weiter wächst, da dies zusätzliche Unterstützung generieren kann (Dahlgaard et al., 2016). Für den entgegengesetzten Underdog-Effekt, bei dem Rückstände in Umfragen zusätzliche Unterstützung mobilisieren soll, liegt hingegen geringe empirische Evidenz vor (ebenda). In Wahlsystemen mit Verhältniswahlrecht und Sperrklausel sind zudem weitere Effekte der Umfrageberichterstattung auf das Wahlverhalten belegt: Beim sogenannten Fallbeileffekt weichen Wähler von ihrer Erstpräferenz ab, wenn ihre bevorzugte Partei in Umfragen unterhalb der Sperrklausel liegt, und entscheiden sich, um eine „verlorene Stimme“ zu vermeiden, für ihre zweitpräferierte Partei (Hoffmann, 2015). Hinzu kommt strategisches Wahlverhalten, etwa in Form des Koalitionswählens, bei dem Wähler gezielt Parteien unterstützen, um bestimmte Regierungsbündnisse zu fördern (Faas et al., 2008).

Diese Effekte setzen eine umfassende Berichterstattung über Wahlen voraus. Hierbei zeigt sich, dass politische *Journalisten* Meinungsumfragen ein hohes Maß an Vertrauen entgegenbringen (Wichmann & Brettschneider, 2009), was deren breite Verwendung in der Wahlberichterstattung maßgeblich begründet. In den letzten Jahren hat die Skepsis gegenüber der Methodik und Integrität von Umfragen jedoch zugenommen, was zu einem Vertrauensverlust führen kann (Aharoni et al., 2022). Dennoch prägen Wahlumfragen die Medienberichterstattung in verschiedener Art und Weise und bilden ein zentrales Element der Wahlberichter-

stattung. So wird unter dem Begriff des Horse-Race-Journalismus seit den 1980er Jahren die steigende Bedeutung von Wahlumfragen und Popularitätsrankings von Politikern gegenüber der Auseinandersetzung mit politischen Sachthemen beschrieben (Atkin & Gaudino, 1984). Mit der zunehmenden medialen Präsenz von Wahlumfragen steigt auch der Anteil der Wähler, die Kenntnis von deren Ergebnissen haben (Brettschneider, 1996). Die Verwendung von Umfragen kann aber auch negative Effekte mit sich bringen. So zeigt Patterson (2005), dass die Qualität der Berichterstattung leidet, wenn Journalisten Umfragen falsch interpretieren oder methodische Details nicht ausreichend berichten, was Missverständnisse über die Bedeutung von Umfrageergebnissen begünstigen kann (Curtice & Moon, 2018; Louwerse & van Dijk, 2022). Zudem kann die selektive Präsentation vorteilhafter Umfrageergebnisse zugunsten der von einem Medium unterstützten Partei den Eindruck eines einseitigen Medienbias erzeugen (Kuru et al., 2020; Perryman et al., 2020).

Auch die Selektionsentscheidungen von Journalisten können durch die Kenntnis von Wahlumfragen beeinflusst werden. Jose (2020) bezeichnet dies als Rückkopplungseffekt: Positive Umfrageergebnisse für eine Partei oder einen Kandidaten führen zu einer umfassenderen oder positiveren Berichterstattung und begünstigen positivere Umfrageergebnisse. Eine indirekte Wirkung veröffentlichter Wahlumfragen auf die Bevölkerung zeigt sich, wenn überraschende Umfrageergebnisse dazu führen, dass Journalisten – als zentrale Rezipienten – die politische Lage neu bewerten und infolgedessen den Grundtenor ihrer Berichterstattung ändern (Jandura & Petersen, 2009). Für Deutschland lassen sich verschiedene Fallbeispiele finden, die diese Annahme stützen. So veränderte die Meldung, dass die NPD bei der Landtagswahl 2004 in Sachsen die 5 %-Hürde überspringen könnte, den Medientenor erheblich. War die Partei vor der Meldung in 4 % der Beiträge zur sächsischen Landespolitik präsent, steigt ihr Anteil in den zwei Wochen vor der Wahl auf über 20 %. Vor dem Hintergrund des Sichtbarkeitseffekts der Medienberichterstattung – je umfangreicher über eine Partei berichtet wird, desto stärker wird sie als wählbare Alternative angesehen – erscheint es plausibel, dass der Wahlerfolg der NPD mit über 9 % auch auf die veränderte Berichterstattung zurückzuführen ist. Für die Bundestagswahlen 2005, 2013 und 2017 untersucht Klein (2022) systematisch den Zusammenhang zwischen Umfrageergebnissen und Veränderungen der Berichterstattung über die im Bundestag vertretenen politischen Parteien. Dabei kann er zeigen, dass ein solcher Zusammenhang vorrangig in Sondersituationen für Parteien besteht, deren Umfragewerte nahe der Sperrklausel liegen. So erfährt die AfD mit dem Überschreiten der Sperrklausel eine gesteigerte mediale Relevanz, während die Berichterstattung über FDP und Linkspartei abnimmt, je deutlicher sich deren Umfragewerte von oben der 5 %-Marke annähern. Bei den übrigen Parteien deutet die zeitliche Abfolge darauf hin, dass Veränderungen in der Berichterstattung eher auf vorausgehende Umfrage-

ergebnisse reagieren, sodass ein kausaler Einfluss der Berichterstattung auf die Umfragewerte unwahrscheinlich erscheint.

Resümierend lässt sich festhalten, dass die dargestellten Wirkungsmechanismen sowohl das Verhalten der Wähler als auch die Berichterstattung über Wahlumfragen und Politik insgesamt beeinflussen können. Besonders bei knapp prognostizierten Wahlausgängen oder bei Parteien, deren Umfragewerte nahe der Sperrklausel liegen, können Umfrageergebnisse für die Vergabe politischer Macht mit entscheidend sein. Demokratische Gesellschaften sind daher auf eine hohe Validität von Wahlumfragen angewiesen, wenn diese ihre Orientierungsfunktion für Politiker, Journalisten und Bürger erfüllen sollen.

### 3. Wahlumfragen vor der Bundestagswahl 2021 in Deutschland

Vor der Bundestageswahl 2021 veröffentlicht das Onlineportal Wahlrecht.de regelmäßig aktuelle Umfragen von acht Instituten, die mit divergierenden Befragungsmodi und unterschiedlichen Stichproben die Wahlabsicht der Deutschen erfassen. Das *Institut für Demoskopie Allensbach (IfD)* arbeitet mit face-to-face Befragungen, *Kantar*, *Forsa*, die *Forschungsgruppe Wahlen (FGW)* und *Global Media Services (GMS)* mit Telefonbefragungen, *Infratest dimap* setzt ein Mixed-Mode-Design aus Telefon- und Onlinebefragung ein und *INSA* und *YouGov* arbeiten mit Onlinebefragungen. Die in den Medien publizierten Umfragen unterscheiden sich voneinander, im Folgenden illustriert an den jeweiligen höchsten und niedrigsten Messwerten für die einzelnen Parteien: (1) Für die CDU/CSU differieren die Ergebnisse um 4 Prozentpunkte: Das IfD (face-to-face) sieht die Partei bei 25 %, YouGov (online) bei nur 21 %. (2) Für die SPD liegen keine Unterschiede vor: Das IfD (face-to-face) und Infratest dimap (Mixed-Mode) weisen 26 % aus, die anderen Anbieter 25 %. (3) Für die Grünen differieren die Werte zwischen 14 % und 17 %: Hier hat YouGov (online) den niedrigsten Wert und Forsa (telefonisch) den höchsten. (4) Die FDP variiert zwischen 10,5 % beim IfD (face-to-face) und 13 % bei GMS (telefonisch). (5) Die Linkspartei ist beim IfD (face-to-face) an der 5 %-Grenze, was für den Fallbeileffekt relevant ist, und laut YouGov (online) liegt sie bei 7 %. (6) Die Zahlen für die AfD variieren zwischen 10 % (IfD, face-to-face; Forsa, FGW, telefonisch) und 12 % (YouGov, online) (Tabelle 1). Auch wenn die ausgewiesenen Prozentzahlen keinen Einfluss auf die Koalitionsarithmetik haben, zeigen sich doch auffällige Unterschiede, gerade zwischen den Daten des IfD (face-to-face) und den Ergebnissen von YouGov (online), was die Relevanz eines Vergleichs von Online- und face-to-face Befragungen erneut unterstreicht.

Tabelle 1: „Wenn am nächsten Sonntag Bundestagswahl wäre ...“

Institut	ifD	Kantar	Forsa	FGW	GMS	Infra- test dimap	INSA	YouGov	Wahl- ergeb- nis
Veröffent- licht	24.09.21	23.09.21	24.09.21	23.09.21	14.09.21	16.09.21	20.09.21	23.09.21	26.09.21
CDU/CSU	25 %	21 %	22 %	23 %	23 %	22 %	22 %	21 %	24,1 %
SPD	26 %	25 %	25 %	25 %	25 %	26 %	25 %	25 %	25,7 %
Grüne	16 %	16 %	17 %	16,5 %	16 %	15 %	15 %	14 %	14,8 %
FDP	10,5 %	11 %	12 %	11 %	13 %	11 %	12 %	11 %	11,5 %
Linke	5 %	7 %	6 %	6 %	6 %	6 %	6,5 %	7 %	4,9 %
AfD	10 %	11 %	10 %	10 %	11 %	11 %	11 %	12 %	10,3 %
Sonstige	7,5 %	9 %	8 %	8,5 %	6 %	9 %	8,5 %	9 %	8,7 %
Modus und Stichproben- größe	F • 1.554	T • 1.443	T • 2.002	T • 1.273	T • 1.003	TOM • 1.512	O • 2.054	O • 2.364	
Erhebungs- zeitraum	16.09.– 23.09.	15.09.– 21.09.	20.09.– 23.09.	22.09.– 23.09.	08.09.– 13.09.	13.09.– 15.09.	17.09.– 20.09.	16.09.– 22.09.	

T – Telefon – telefonische Befragung von zufällig ausgewählten Personen

O – Online-Panel – internetbasierte Befragung von nach Quotenvorgaben ausgewählten Mitgliedern eines Befragten-Pools

F – F2F – persönlich-mündliche Befragung von nach Quotenvorgaben ausgewählten Personen

TOM – T-O-Mix – Befragung per Telefon und per Online-Panel

Quelle: In Anlehnung an <https://www.wahlrecht.de/umfragen/archiv/2021.htm>.

An dieser Stelle sei darauf hingewiesen, dass Unterschiede zwischen Wahlumfragen multikausal erklärt werden können. In der Auswertung der Wahlumfragen zur US-Präsidentenwahl 2004 differenziert Zukin (2005) zwischen vier Determinanten für diese Unterschiede: (1) Der Zeitpunkt der Datenerhebung ist ungleich dem Zeitpunkt der Veröffentlichung. Gerade in Zeiten eines dynamischen Wählermarktes ist zu bedenken, dass Umfragen immer nur Aussagen über die Wahlabsicht der Bevölkerung zum Zeitpunkt der Datenerhebung treffen können. Verändert sich aufgrund von Schlüsselereignissen im Wahlkampf die Zustimmung für politische Parteien oder Kandidaten, lässt sich dies nur in neuen Umfragen abbilden. (2) Als weitere Ursache werden unterschiedliche Stichprobenansätze genannt. Die Befunde von Zufallsstichproben (z.B. RDD-Telefonstichproben<sup>3</sup>) können sich von Quotenverfahren unterscheiden. Hierbei ist anzumerken, dass Zufallsstichproben in den letzten Jahren zunehmend niedrige Ausschöpfungsquoten aufweisen. Eine Metaanalyse in Bezug auf die Rücklaufquoten von Onlinebefragungen zeigt, dass weder die Zahl der kontaktierten Personen noch Anreize für die Teilnahme diese niedrigen Rücklaufquoten erhöhen können (Wu et al., 2022). (3) Als dritter Punkt wird die unterschiedliche Behandlung der Roh-

<sup>3</sup> Random Digital Dialing (RDD) bezeichnet ein Verfahren der Telefonbefragung, in welchem Teile einer Telefonnummer nach der ortsgebundenen Vorwahl zufällig hinzugefügt werden.

daten genannt, da kein Institut die Befunde in Form von Rohdaten veröffentlicht. Die Datenbereinigung umfasst sowohl Gewichtungsverfahren zum Ausgleich der Über- bzw. Unterrepräsentation bestimmter sozialer Gruppen als auch verschiedene Imputationsverfahren aufgrund fehlender Angaben zur Wahlabsicht (Item-Nonresponse). Folglich können unterschiedliche Gewichtungen und Imputationsverfahren Auswirkungen auf die veröffentlichten Befunde haben (Durand, 2023). (4) Viertens werden modusbedingte Unterschiede diskutiert. Die Bereitschaft, auf die Wahlabsichtsfrage zu antworten, unterscheidet sich in selbstadministrierten Interviews (online) und Interviewer-administrierten Interviews (face-to-face und telefonisch). Die letztgenannten Modi können zudem Interviewereffekte evozieren, wie das Empfinden sozialer Erwünschtheit (Taddicken, 2009).

Da Unterschiede zwischen verschiedenen Wahlumfragen sowie Abweichungen vom tatsächlichen Wahlergebnis multikausale Ursachen haben können, bedarf es Methodenexperimenten, die unter Kontrolle der anderen Determinanten einen spezifischen Faktor isolieren und variieren. Im Folgenden wird ein solches Feldexperiment vorgestellt.

#### 4. Methode und Operationalisierungen

Die vorliegende Studie vergleicht systematisch mehrere vor der Bundestagswahl 2021 durchgeführte persönlich-mündliche und Onlinebefragungen miteinander. Bei den Datensätzen handelt es sich um zwei face-to-face Befragungen und zwei Onlinebefragungen unterschiedlicher Anbieter in Deutschland (Tabelle 2).

Tabelle 2: Überblick Datensätze

Datensatz	Feldzeit	Befragte	Modus
Face-to-face I	26.08.- 09.09.2021	N = 1.514	F2F
Onlinebefragung Anbieter I	08.09.- 13.09.2021	N = 1.015	Online
Face-to-face II	11.09.- 23.09.2021	N = 1.553	F2F
Onlinebefragung Anbieter II	11.09.- 20.09.2021	N = 1.088	Online

Zur Sicherstellung des feldexperimentellen Charakters der Studie werden die zuvor dargestellten Erklärungsfaktoren für Ergebnisunterschiede zwischen verschiedenen Wahlumfragen als Kontrollvariablen konzeptualisiert. (1) Die Befragungszeitpunkte der vier Befragungen liegen allesamt in der heißen Wahlkampfphase des Bundestagswahlkampfes 2021. Da die Realisierung von face-to-face Befragungen mehr Zeit als Onlinebefragungen in Anspruch nimmt, kann eine Parallelität der Erhebungszeitpunkte jedoch nur annähernd erreicht werden. (2) Die Stichprobenbildung ist identisch. Alle Befragten sind nach denselben Quotierungsmerkmalen (u.a. Altersgruppen, Geschlecht, Bildungsgrad, Wohnort, Berufstätigkeit) ausgewählt. (3) Sowohl der Fragebogen als auch der Umgang mit den Rohdaten sind identisch. Dementsprechend stimmen alle Variablen und

deren Ausprägungen überein und die Datensätze sind strukturgleich. Um die Strukturgleichheit herzustellen, erfolgt eine Gewichtung der Daten nach den Variablen Altersgruppe (18-29 Jahre, 30-44 Jahre, 45-59 Jahre, 60 Jahre und älter), Geschlecht (Männer vs. Frauen), Bildung (niedrig, mittel, hoch), Wohnort in Ost- oder Westdeutschland (Ost vs. West), der Ortsgröße (unter 5.000 Einwohner, 5.000 bis unter 20.000 Einwohner, 20.000 bis unter 100.000 Einwohner, über 100.000 Einwohner), der Region, in der man lebt (Norddeutschland, Nordrhein-Westfalen, Südwestdeutschland, Baden-Württemberg, Bayern, Berlin, Nordostdeutschland und Sachsen und Thüringen), Haushaltsgröße (1-Personen-Haushalt, 2-Personen-Haushalt, 3-Personen-Haushalt, 4-Personen-Haushalt, 5- und mehr-Personen-Haushalten) und Erwerbstätigkeit (erwerbstätig vs. nicht erwerbstätig). Basis der Gewichtung bilden die Angaben aus dem Mikrozensus 2019, der zum Zeitpunkt der Bundestagswahl 2021 aktuellsten verfügbaren Bevölkerungsstatistik (Statistisches Bundesamt, 2020). Die Gewichtungskoeffizienten pro Person liegen in angemessenen Wertebereichen: Face-to-face I zwischen ,42 und 3,05; Face-to-face II zwischen ,41 und 3,59; Onlineanbieter I zwischen ,50 und 2,32; Onlineanbieter II zwischen ,56 und 2,31.

Die insgesamt 34 Variablen (acht Variablen aus dem Bereich der Soziodemographie und 24 Variablen, die Einstellungen, Meinungen, Verhalten und Verhaltensabsichten erfassen) bilden einen finalen Gesamtdatensatz (Fragebogen im Anhang A1, <https://doi.org/10.24338/mip-2025142-160>). Die Variablen umfassen folgende Angaben: (1) Soziodemographika: Persönliche Merkmale (Alter, Geschlecht, Bildung, Berufstätigkeit), Haushaltsmerkmale (Anzahl der Personen im Haushalt), räumliche Merkmale (Wohnortgröße, Bundesland) sowie (2) Fragen mit Wahlbezug: Hierzu zählen die Einschätzung der Arbeit der Bundesregierung während der Corona-Pandemie, die Zukunftserwartungen der Befragten, das politische Interesse, die Sympathie gegenüber den verschiedenen Parteien sowie Fragen nach dem vermuteten Wahlsieger und der Kandidatenpräferenz. Zur Ermittlung der Wahlentscheidung dienen außerdem Fragen zur Wahlabsicht, der Sicherheit der Wahlentscheidung, der Regelmäßigkeit der Teilnahme an Wahlen, dem Wählerpotential und der Wahlberechtigung.

Anschließend werden alle 34 Variablen auf divergierendes Antwortverhalten hin überprüft (mittels T-Test, Chi-Quadrat-Test und F-Test). Zunächst erfolgt eine Analyse der Randverteilungen. In Anlehnung an die Überlegungen zur Fusion sozialwissenschaftlicher Daten (Kiesel & Rässler, 2006) wird als zweiter Schritt geprüft, inwieweit sich die Korrelationsstrukturen zwischen den Datensätzen unterscheiden. Unterschiedliche Korrelationsstrukturen deuten auf tiefgreifende Unterschiede in den Einstellungsstrukturen der Befragten hin.

## 5. Ergebnisse

Die Ergebnisse sind nach den beiden Forschungsfragen gegliedert. Zunächst interessiert, worin sich die *Ergebnisse von F2F-Befragungen und Onlinebefragungen vor der Bundestagswahl 2021 in Deutschland unterscheiden*.

Alle Fragen (n=19), die nicht in die Gewichtung der Datensätze inkludiert sind, werden mittels Chi-Quadrat-Test und F-Test auf Unterschiede zwischen den Ergebnissen der face-to-face Befragungen und der Onlinebefragungen geprüft. Bei 17 Variablen sind signifikante Unterschiede zwischen beiden Modi evident (Tabelle Anhang A2, <https://doi.org/10.24338/mip-2025142-160>).

Folgende Beispiele illustrieren besagte Unterschiede: Das Wählerpotential der CDU liegt bei beiden face-to-face Befragungen mit 31,4 % bzw. 32,2 % deutlich höher als bei den beiden Onlinebefragungen (22,2 % bzw. 22,4 %,  $p < 0,05$ ). Für die Linkspartei und die AfD weisen die Onlinebefragungen dagegen ein größeres Wählerpotential aus (Linke: 9,5 % bzw. 11,0 % face-to-face und 14,7 % bzw. 13,3 % online; AfD: 9,9 % bzw. 8,7 % face-to-face und 15,4 % bzw. 16,2 % online,  $p < 0,05$ ). Ebenfalls unterscheiden sich das politische Interesse (50,8 % bzw. 55,7 % face-to-face und 62,0 % bzw. 59,1 % online) und die Sicherheit der Wahlentscheidung (61,8 % bzw. 63,5 % face-to-face und 69,3 % bzw. 72,5 % online) stark ( $p < 0,05$ ).

Neben diesen Unterschieden in den Randverteilungen finden sich auch Differenzen in den Korrelationen zwischen den Variablen, die auf unterschiedliche Einstellungsstrukturen in den Datensätzen hinweisen. Diese Unterschiede werden besonders deutlich am Beispiel des Wählerpotentials der CDU im Vergleich zu dem anderer Parteien. Die Parteiensympathien von CDU und SPD korrelieren in den face-to-face Befragungen negativ, während dies in den Onlinebefragungen nicht der Fall ist. Zudem sind die Korrelationen zwischen dem Wählerpotential der CDU und dem anderer Parteien bei Onlinebefragungen und face-to-face Befragungen unterschiedlich stark ausgebildet (Tabelle 3).

**Tabelle 3: Korrelation des Wählerpotentials der CDU mit dem der anderen Parteien**

	Face-to-face I	Face-to-face II	Online I	Online II
Parteiensympathie CDU & SPD	-0,26**	-0,24**	0,08	-0,09
Parteiensympathie CDU & FDP	0,17**	0,26**	0,17**	0,05
Parteiensympathie CDU & Grüne	-0,28**	-0,33**	-0,11**	-0,21**
Parteiensympathie CDU & Linke	-0,58**	-0,61**	-0,31**	-0,56**
Parteiensympathie CDU & AfD	-0,08**	-0,10**	0,05	-0,18**
Politisches Interesse & Sicherheit der Wahlentscheidung	0,091**	0,084**	0,175**	0,143**

\*\*\*  $p < 0,01$ ; \*\*  $p < 0,05$ ; \*  $p < 0,10$

Der Überblick über die Ergebnisse führt zu dem Schluss, dass die face-to-face Befragungen stärker Wählerschichten erreichen, die geringer politisch interessiert sind, eine geringere politische Selbstwirksamkeit empfinden und konservativer sind. Möglicherweise können die segmentären und stratifikatorischen Rekrutierungsunterschiede (Bennett & Iyengar, 2008) nicht durch die soziodemographische Gewichtung ausgeglichen werden. Zudem lassen Unterschiede in der Korrelationsstruktur zwischen den Variablen vermuten, dass face-to-face Befragte weniger stark ausgeprägte Einstellungsstrukturen aufweisen. Als mögliche Ursachen für die beobachteten Unterschiede werden im Weiteren sowohl die Auswirkung unterschiedlicher Grundgesamtheiten als auch die Teilnahmebereitschaft an Onlinebefragungen untersucht.

Eine erste mögliche Erklärung der Unterschiede sind die differierenden *Grundgesamtheiten* der Befragungen. Während die face-to-face Befragungen auch Offliner inkludieren, besteht die Auswahlgesamtheit der Onlinebefragungen ausschließlich aus Onlinern. Die Onlinebefragungen weisen folglich einen Abdeckungsfehler (Groves, 1987) auf. Mit dem Anstieg der Internetnutzung auf 94 % im Wahljahr 2021 hat die Bedeutung von Offlinern für die Beurteilung der Stichprobenqualität deutlich abgenommen (Beisch & Koch, 2021). Evident ist jedoch, dass sich der kleine, aber harte Kern von Offlinern systematisch von den Onlinern unterscheidet. Offliner sind überwiegend älter, formal niedriger gebildet, verfügen über ein unterdurchschnittliches Einkommen und sind aufgrund ihres fortgeschrittenen Alters überdurchschnittlich häufig im Ruhestand (Initiative D21 e.V., 2021). Dieses soziodemographische Profil beeinflusst auch ihre politischen Einstellungen und ihr Wahlverhalten. Im Vergleich zu Onlinern sehen Offliner signifikant (für alle folgenden Unterschiede gilt mindestens  $p < 0,05$ ) skeptischer in die Zukunft (40 % vs. 27 %), bewerten die Arbeit der Bundesregierung positiver (55 % vs. 43 %), interessieren sich weniger für Politik (hohes Interesse: 47 % vs. 60 %), haben eine höhere Affinität gegenüber der CDU (36 % vs. 30 %), bewerten diese auch positiver (Parteiensympathie Mittelwert 2,6 vs. 3,0) und beabsichtigen häufiger, eine der beiden etablierten Volksparteien zu wählen.

Diese Befunde zu Offlinern entsprechen in ihrer Tendenz den Ergebnisunterschieden zwischen den face-to-face Befragungen und den Onlinebefragungen. Um die unterschiedlichen Grundgesamtheiten als Ergebnisunterschiede auszuschließen, werden folgend nur die Antworten der Onliner der face-to-face Befragungen mit den Onlinern der Onlinebefragungen verglichen. Da nur die zweite face-to-face Befragung zwischen Onlinern und Offlinern differenziert, kann folglich nur diese Umfrage mit den beiden Onlinebefragungen verglichen werden. Um eine Strukturgleichheit der Datensätze zu erreichen, werden diese erneut gewichtet. Ebenfalls werden anhand von 19 Fragen mithilfe von Chi-Quadrat-Test und F-Test Unterschiede zwischen Onlinern aus der face-to-face Befragung und Onlinern aus beiden Onlinestichproben ermittelt. Auch hier weisen 17 Fragen mit Wahlbezug

signifikante Unterschiede zwischen beiden Modi auf (Tabelle Anhang A3, <https://doi.org/10.24338/mip-2025142-160>). Die Unterschiede fallen in Relation zum ersten Vergleich jedoch geringer aus. Für die Vergleichbarkeit werden die oben angeführten Beispiele erneut diskutiert. So lag das Wählerpotential der CDU bei den Onlinern der face-to-face Befragung mit 31,3 % deutlich höher als bei den beiden Onlinebefragungen (21,7 % bzw. 22,4 %). Bei der Linkspartei und der AfD weisen die Onlinebefragungen dagegen ein größeres Wählerpotential aus (Linke: 11,8 % F2F und 14,7 % bzw. 13,2 % online; AfD: 9,7 % face-to-face und 15,7 % bzw. 16,3 % online). Darüber hinaus unterscheiden sich das politische Interesse (56,4 % face-to-face und 61,3 % bzw. 58,5 % online) und die Sicherheit der Wahlentscheidung (61,6 % face-to-face und 67,6 % bzw. 70,5 % online) stark. Zudem lassen sich erneut unterschiedliche Korrelationsstrukturen bezüglich der Parteiensympathie von CDU & FDP sowie AfD feststellen (Tabelle 4).

**Tabelle 4: Korrelation des Wählerpotentials der CDU mit dem der anderen Parteien (Onliner)**

	Face-to-face II	Online I	Online II
Parteiensympathie CDU & SPD	0,03	0,08	-0,09**
Parteiensympathie CDU & FDP	0,39**	0,17**	0,05
Parteiensympathie CDU & Grüne	-0,16**	-0,11**	-0,21**
Parteiensympathie CDU & Linke	-0,40**	-0,31**	-0,56**
Parteiensympathie CDU & AfD	-0,00	0,05	-0,18**
Politisches Interesse & Sicherheit der Wahlentscheidung	0,05**	0,175**	0,143**

\*\*\*  $p < 0,01$ ; \*\*  $p < 0,05$ ; \*  $p < 0,10$

Trotz der Einschränkung der Grundgesamtheit auf Onliner und der neuen Gewichtung des Datensatzes auf Basis der Onliner in Deutschland zeigt sich, dass F2F-Befragungen weiterhin stärker Wählerschichten erreichen, die geringer politisch interessiert sind, eine geringere politische Selbstwirksamkeit haben und konservativer sind.

Ein zweiter möglicher Erklärungsfaktor für die Ergebnisunterschiede ist ein *Rekrutierungsbias*, der sich aus unterschiedlicher Teilnahmebereitschaft je nach Befragungsmodus ergeben könnte. Die face-to-face Befragung beinhaltet das Item „Könnten Sie sich vorstellen, einmal an solch einer Befragung im Internet teilzunehmen, oder kommt das für Sie nicht in Frage?“, welches einen möglichen Rekrutierungsbias aufdecken könnte. Es zeigt sich, dass nur 55,4 % der face-to-face befragten Onliner diese Frage mit „ja“ beantworten, was darauf hindeutet, dass eine solche Skepsis gegenüber dem Onlinemodus keine Voraussetzung für die Teilnahme an einer Onlinebefragung ist. Ähnliche Befunde finden sich in den Auswertungen der GLES (German Longitudinal Election Study) zur Bundestags-

wahl 2021. Obwohl sie täglich online sind, bevorzugen 44 % der Befragten im Mixed-Mode-Design die schriftliche gegenüber der Onlinebefragung. Diese Befunde lassen den Schluss zu, dass online sein nicht automatisch die Teilnahme an Onlinebefragungen bedingt.

Zur analytischen Abgrenzung zwischen Onlinern, die zur Teilnahme an Onlinebefragungen bereit sind, und solchen, die diese grundsätzlich ausschließen, wird ein logistisches Regressionsmodell eingesetzt. Die Ergebnisse (Nagelkerkes R-Quadrat: 0,113; Prozentsatz richtiger Zuordnung bei Trennwert 0,5 = 63,5; Codierung abhängige Variable: „1“ Onliner, die teilnehmen & „2“ Onliner, die nicht teilnehmen) bestätigen die Annahme, dass sich beide Gruppen sowohl hinsichtlich der soziodemographischen Angaben als auch hinsichtlich der wahlrelevanten Einstellungen deutlich unterscheiden. Die fehlende Teilnahmebereitschaft online geht einher mit einem geringeren Alter, dem Leben in größeren Haushalten außerhalb von Großstädten, einem niedrigeren politischen Interesse, einem geringeren Wählerpotential für SPD, Grüne und Linke, einer höheren Parteiensympathie für CDU, SPD, Grüne und AfD sowie einer negativeren für die FDP.

**Tabelle 5: Unterschiede zwischen Onlinern, die an Onlinebefragungen teilnehmen bzw. nicht teilnehmen.**

	Regressionskoeffizient	Exp(B)	Sig.
Alter offen	0,026	1,027	0,000
Haushaltsgröße	0,128	1,137	0,002
Ortsgröße	-0,093	0,911	0,003
Politisches Interesse	0,503	1,653	0,000
Wählerpotential SPD	-0,312	0,732	0,013
Wählerpotential Grüne	-0,346	0,707	0,009
Wählerpotential Linke	-0,533	0,587	0,003
Parteiensympathie CDU	-0,103	0,902	0,017
Parteiensympathie SPD	-0,143	0,867	0,002
Parteiensympathie FDP	0,208	1,231	0,000
Parteiensympathie AfD	-0,151	0,860	0,001

Diese Unterschiede schlagen sich auch in der Zweitstimmenwahlabsicht nieder. Diejenigen, die sich vorstellen können, Befragungen online auszufüllen, geben häufiger an, CDU (19,3 % vs. 17,6 %) und SPD (23,1 % zu 19,1 %) zu wählen bzw. seltener Grüne (13,3 % vs. 20,6 %) und Linke (3,4 % vs. 5,1 %) zu wählen. Bei den übrigen Parteien gibt es keine signifikanten Unterschiede. Einzig die Befragten, die unentschieden sind bzw. keine Angabe machen, unterscheiden sich zusätzlich (21,4 % vs. 26,6 %). Auch diese Befunde bestätigen das Muster, dass Befragte, die nicht durch Onlinebefragungen erreichbar sind, politisch weniger interessiert sind und eher konservativere Einstellungen haben.

## 6. Fazit

Ausgangspunkt der Studie ist die Beobachtung, dass sich die Ergebnisse von Wahlumfragen voneinander unterscheiden, wenn sie als face-to-face Befragung oder als Onlinebefragung durchgeführt werden. Ziel ist es, eine Erklärung für diese Unterschiede zu finden. Dazu dient ein Feldexperiment von je zwei face-to-face Befragungen und zwei Onlinebefragungen, bei simultaner Kontrolle des Zeitpunkts der Erhebung, der Stichprobenbildung und des Umgangs mit den Rohdaten. Der Vergleich verschiedener Fragen zu politischen Einstellungen und politischem Verhalten zeigt teilweise deutliche signifikante Unterschiede zwischen den Befragungsmodi. Befragte der face-to-face Befragungen sind geringer politisch interessiert, haben eine geringere politische Selbstwirksamkeit und präferieren stärker konservative Parteien.

Um diese Unterschiede weiter erklären zu können, wird das feldexperimentelle Design erweitert und die Grundgesamtheit der Stichproben kontrolliert. Zu diesem Zweck werden Offliner aus der face-to-face Erhebung aus dem Datensatz ausgeschlossen, sodass ein Vergleich ausschließlich zwischen Onlinern beider Erhebungsmodi möglich ist. Selbst die Extraktion der Onliner führt jedoch die gleichen Ergebnisse zutage, wenn auch in abgeschwächter Form.

Im dritten Schritt wird die Teilnahmebereitschaft an Onlinebefragungen als Erklärungsvariable für die beobachteten Ergebnisunterschiede untersucht. Dabei zeigt sich, dass sich lediglich gut jeder zweite Onliner aus der F2F-Befragung (55 %) vorstellen kann, denselben Fragebogen auch online zu beantworten. Die Unterschiede zwischen den potentiell teilnahmebereiten und nicht teilnahmebereiten Personen bestätigt das bereits in der Studie beobachtete Muster: Personen, die eine Teilnahme an Onlinebefragungen ausschließen, zeigen ein geringeres politisches Interesse und tendieren stärker zu etablierten Großparteien sowie konservativen Parteien.

In der Gesamtschau lassen die Ergebnisse vermuten, dass Wahlumfragen, die auf dem Modus der Onlinebefragung beruhen und mit Quotenstichproben arbeiten, (1) die Wahlbeteiligung überschätzen, (2) die Stimmanteile der etablierten großen Parteien CDU und SPD unterschätzen und (3) das Wahlergebnis der Grünen sowie der Linken und der AfD überschätzen. Diese verzerrte Abbildung der Meinungsverteilung in der Bevölkerung kann über die direkten und indirekten Einflüsse von publizierten Wahlumfragen auf die Wähler und die Journalisten mittelbar die Verteilung politischer Macht in liberalen Demokratien beeinflussen. Dies ist ein nicht zu unterschätzender Einfluss auf das Wahlverhalten, der auch in aktuellen Wahlen debattiert wird und durch die zunehmende Digitalisierung noch verschärft werden könnte.

Zudem beleuchtet die vorliegende Untersuchung ein Phänomen, das insbesondere in zukünftigen Mixed-Method-Ansätzen stärker berücksichtigt werden

sollte: Die bloße Internetnutzung und jüngeres Alter sind keine hinreichende Voraussetzung für die Teilnahme an Onlinebefragungen.

Der Befragungsmodus bestimmt damit sowohl die Zusammensetzung der befragten Stichprobe als auch die daraus abgeleiteten Wahlprognosen und wirkt so unmittelbar auf die Wahlberichterstattung, die öffentliche Meinungsbildung und potentiell auf das Wahlergebnis selbst. Dies betont die Verantwortung methodischer Entscheidungen in der Wahlforschung.

## Literatur

- ADM e.V. (2023). Jahresbericht 2023: Quantitative Interviews nach Befragungsart.
- Aharoni, T., Tenenboim-Weinblatt, K., Baden, C. & Overbeck, M. (2022). Dynamics of (dis)trust between the news media and their audience: The case of the April 2019 Israeli exit polls. *Journalism*, 23(2), 337–353. <https://doi.org/10.1177/1464884920978105>
- Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C. & Wingate, D. (2023). Out of One, Many: Using Language Models to Simulate Human Samples. *Political Analysis*, 31(3), 337–351. <https://doi.org/10.1017/pan.2023.2>
- Atkin, C. K. & Gaudino, J. (1984). The Impact of Polling on the Mass Media. *The ANNALS of the American Academy of Political and Social Science*, 472(1), 119–128. <https://doi.org/10.1177/0002716284472001011>
- Baur, N. & Florian, M. J. (2009). Stichprobenprobleme bei Online-Umfragen. In N. Jakob, H. Schoen & T. Zerback (Hrsg.), *Sozialforschung im Internet: Methodologie und Praxis der Online-Befragung* (S. 109–128). VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH, Wiesbaden. [https://doi.org/10.1007/978-3-531-91791-7\\_7](https://doi.org/10.1007/978-3-531-91791-7_7)
- Beisch, N. & Koch, W. (2021). 25 Jahre ARD/ZDF-Onlinestudie: Unterwegsnutzung steigt wieder und Streaming/Mediatheken sind weiterhin Treiber des medialen Internets. *Media Perspektiven*(10), 486–503.
- Bendor, J., Diermeier, D. & Ting, M. (2003). A Behavioral Model of Turnout. *American Political Science Review*, 97(02), 261–280. <https://doi.org/10.1017/S0003055403000662>
- Bennett, W. L. & Iyengar, S. (2008). A New Era of Minimal Effects? The Changing Foundations of Political Communication. *Journal of Communication*, 58(4), 707–731. <https://doi.org/10.1111/j.1460-2466.2008.00410.x>
- Bennink, M., Moors, G. & Gelissen, J. (2013). Exploring Response Differences between Face-to-Face and Web Surveys. *Field Methods*, 25(4), 319–338. <https://doi.org/10.1177/1525822X12472875>

- Bieber, I. E. & Bytzek, E. (2012). Online-Umfragen: Eine geeignete Erhebungsmethode für die Wahlforschung? Ein Vergleich unterschiedlicher Befragungsmodi am Beispiel der Bundestagswahl 2009. *Methoden — Daten — Analysen*, 6(2), 185–211.
- Bisbee, J., Clinton, J. D., Dorff, C., Kenkel, B. & Larson, J. M. (2024). Synthetic Replacements for Human Survey Data? The Perils of Large Language Models. *Political Analysis*, 32(4), 401–416. <https://doi.org/10.1017/pan.2024.5>
- Braekman, E., Demarest, S., Charafeddine, R., Drieskens, S., Berete, F., Gisle, L., van der Heyden, J. & van Hal, G. (2022). Unit Response and Costs in Web Versus Face-To-Face Data Collection: Comparison of Two Cross-sectional Health Surveys. *Journal of medical Internet research*, 24(1), e26299. <https://doi.org/10.2196/26299>
- Brettschneider, F. (1996). Wahlumfragen und Medien – Eine empirische Untersuchung der Presseberichterstattung über Meinungsumfragen vor den Bundestagswahlen 1980 bis 1994. *Politische Vierteljahresschrift*, 37(3), 475–493. <http://www.jstor.org/stable/24198261>
- Bursztyjn, L., Cantoni, D., Funk, P. & Yuchtman, N. (2017). Polls, the Press, and Political Participation: The Effects of Anticipated Election Closeness on Voter Turnout. *Political Behavior: Voting & Public Opinion eJournal*.
- Curtice, J. & Moon, N. (2018). Foreword. *International Journal of Market Research*, 60(2), 129–131. <https://doi.org/10.1177/1470785318755935>
- Dahlgaard, J. O., Hansen, J. H., Hansen, K. M. & Larsen, M. V. (2016). How are Voters Influenced by Opinion Polls? The Effect of Polls on Voting Behavior and Party Sympathy. *World Political Science*, 12(2), 283–300. <https://doi.org/10.1515/wps-2016-0012>
- Dieckmann, C. (6. November 2024). Wahlsieg von Donald Trump: Wie zuverlässig waren die Umfragen? DER SPIEGEL. <https://www.spiegel.de/ausland/us-wahl-2024-wie-zuverlaessig-waren-die-umfragen-zum-wahlsieg-von-donald-trump-a-b476cd6c-5f53-4b05-9b44-7dcb68509e55>
- Duffy, B., Smith, K., Terhanian, G. & Bremer, J. (2005). Comparing data from online and face-to-face surveys. *International Journal of Market Research*, 47(6).
- Durand, C. (2023). Analyzing the polls of the US 2020 Presidential Election Campaign: A New Perspective. <https://doi.org/10.2139/ssrn.4550801>
- Faas, T., Mackenrodt, C. & Schmitt-Beck, R. (2008). Polls that Mattered: Effects of Media Polls on Voters' Coalition Expectations and Party Preferences in the 2005 German Parliamentary Election. *International Journal of Public Opinion Research*, 20(3), 299–325. <https://doi.org/10.1093/ijpor/edn034>

- Gilster, P. (1997). *Digital literacy*. Wiley Computer Publishing.
- Groves, R. M. (1987). Research on Survey Data Quality. *Public Opinion Quarterly*, 51(part 2: Supplement: 50<sup>th</sup> Anniversary Issue), S156. <https://doi.org/10.1086/269077>
- Heerwegh, D. (2009). Mode Differences Between Face-to-Face and Web Surveys: An Experimental Investigation of Data Quality and Social Desirability Effects. *International Journal of Public Opinion Research*, 21(1), 111–121. <https://doi.org/10.1093/ijpor/edn054>
- Heerwegh, D. & Loosveldt, G. (2008). Face-to-Face versus Web Surveying in a High-Internet-Coverage Population: Differences in Response Quality. *Public Opinion Quarterly*, 72(5), 836–846. <https://doi.org/10.1093/poq/nfn045>
- Hoffmann, H. (2015). Wahlumfragen und der Willensbildungsprozess der Wähler im Vorfeld der Bundestagswahl 2013. *Zeitschrift für Parlamentsfragen*, 46(4), 708–729.
- Initiative D21 e.V. (2021). D21 Digital Index 2020/2021: Jährliches Lagebild zur Digitalen Gesellschaft. Kantar. [https://initiated21.de/app/uploads/2021/02/d21-digital-index-2020\\_2021.pdf](https://initiated21.de/app/uploads/2021/02/d21-digital-index-2020_2021.pdf)
- Jandura, O. & Petersen, T. (2009). Gibt es eine indirekte Wirkung von Wahlumfragen? *Publizistik*, 54(4), 485–497. <https://doi.org/10.1007/s11616-009-0065-6>
- Jose, A. (2020). Accounting for the Impact of Media Coverage on Polling. *viXra*.
- Kiesl, H. & Rässler, S. (2006). How Valid Can Data Fusion Be? Beiträge zum wissenschaftlichen Dialog aus dem Institut für Arbeitsmarkt- und Berufsforschung. IAB-Discussion Paper, 15.
- Klein, B. (2022). Umfrageeffekte auf Journalisten? Eine neue Perspektive für die Wahlforschung am Beispiel der Berichterstattung zu den Bundestagswahlen 2009, 2013 und 2017 [Dissertationsschrift], Philosophische Fakultät der Heinrich-Heine-Universität Düsseldorf.
- Kuru, O., Pasek, J. & Traugott, M. W. (2020). When Pundits Weigh In: Do Expert and Partisan Critiques in News Reports Shape Ordinary Individuals' Interpretations of Polls? *Mass Communication and Society*, 23(5), 628–655. <https://doi.org/10.1080/15205436.2020.1774780>
- Louwerse, T. & van Dijk, R. E. (2022). Reporting the polls: the quality of media reporting of vote intention polls in the Netherlands. *Acta Politica*, 57(3), 548–570. <https://doi.org/10.1057/s41269-021-00208-5>
- Maurer, M. & Jandura, O. (2009). Masse statt Klasse? Einige kritische Anmerkungen zu Repräsentativität und Validität von Online-Befragungen. In N. Jakob, H. Schoen & T. Zerback (Hrsg.), *Sozialforschung im Internet: Methodologie*

- und Praxis der Online-Befragung (S. 61–73). VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH, Wiesbaden. [https://doi.org/10.1007/978-3-531-91791-7\\_4](https://doi.org/10.1007/978-3-531-91791-7_4)
- Newton, K. (1999). Mass Media Effects: Mobilization or Media Malaise? *British Journal of Political Science*, 29(4), 577–599. <https://doi.org/10.1017/S0007123499000289>
- Patterson, T. E. (2005). Of Polls, Mountains: U.S. Journalists and Their Use of Election Surveys. *Public Opinion Quarterly*, 69(5), 716–724. <https://doi.org/10.1093/poq/nfi065>
- Perryman, M. R., Foley, J. & Wagner, M. W. (2020). Is Bad News Biased? How Poll Reporting Affects Perceptions of Media Bias and Presumed Voter Behavior. *International Journal of Communication*, 14(0), 21. <https://ijoc.org/index.php/ijoc/article/view/13167/3160>
- Pfannenstiel, B. (3. November 2024). Iowa Poll: Kamala Harris leapfrogs Donald Trump to take lead near Election Day. Here's how. *The Des Moines Register*. <https://eu.desmoinesregister.com/story/news/politics/iowa-poll/2024/11/02/iowa-poll-kamala-harris-leads-donald-trump-2024-presidential-race/75354033007/>
- Reinemann, C., Jandura, O., Maurer, M. & Zerback, T. (2011). Wer sind die Spätereisende? Erste Befunde einer Mehr-Methoden-Studie zur Bundestagswahl 2009. In H. Oberreuter (Hrsg.), *Am Ende der Gewissheiten: Wähler, Parteien und Koalitionen in Bewegung* (S. 247–270). Olzog.
- Sieben, P. (25. Februar 2025). BSW will nach Bundestagswahl Strafanzeige stellen – Experte: „Beruht auf Mutmaßungen“. <https://www.fr.de/politik/bsw-will-nach-bundestagswahl-straftanzeige-stellen-experte-beruht-auf-mutmassungen-zr-93591055.html>
- Statistisches Bundesamt. (2020). *Mikrozensus 2019: Qualitätsbericht*. <https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Bevoelkerung/mikrozensus-2019.html>
- Strabac, Z. & Aalberg, T. (2011). Measuring Political Knowledge in Telephone and Web Surveys: A Cross-National Comparison. *Social Science Computer Review*, 29(2), 175–192. <https://doi.org/10.1177/0894439310371340>
- Taddicken, M. (2009). Methodeneffekte von Web-Befragungen: Soziale Erwünschtheit vs. Soziale Entkontextualisierung. In M. Weichbold, J. Bacher & C. Wolf (Hrsg.), *Österreichische Zeitschrift für Soziologie. Sonderheft: 9/2009. Umfrageforschung: Herausforderungen und Grenzen* (S. 107–130). VS Verlag für Sozialwissenschaften.

- Wichmann, W. & Brettschneider, F. (2009). American and German Elite Journalists' Attitudes toward Election Polls. *International Journal of Public Opinion Research*, 21(4), 506–524. <https://doi.org/10.1093/ijpor/edp048>
- Wu, M.-J., Zhao, K. & Fils-Aime, F. (2022). Response rates of online surveys in published research: A meta-analysis. *Computers in Human Behavior Reports*, 7, 100206. <https://doi.org/10.1016/j.chbr.2022.100206>
- Zukin, C. (2005, 26. Mai). Polls and the 2004 Election. ICA conference, New York.