

SICHERHEIT IN ZAHLEN?

Ein Beipackzettel für Meinungsumfragen zu Außen- und Sicherheitspolitik

In außen- und sicherheitspolitischen Debatten verweisen führende Politikerinnen und Politiker immer wieder auf Umfragen, nicht zuletzt bei Waffenlieferungen an die Ukraine. Doch wer Umfragedaten nutzt, sollte wissen, was sie leisten können – und was nicht. Viele Befragte äußern zum Beispiel bei einem komplexen Politikfeld wie der Außen- und Sicherheitspolitik Meinungen zu Themen, zu denen sie wenig Wissen oder keine gefestigte Haltung haben. Deswegen kommt Frageformulierungen und Erhebungsmethoden eine große Bedeutung zu. Dieses Spotlight zeigt, worauf Entscheidungsträger*innen, Think Tanker oder Journalist*innen achten sollten, wenn sie der nächsten Umfrage zu Außen- und Sicherheitspolitik begegnen.



Verteidigungsminister Boris Pistorius und der Generalinspekteur der Bundeswehr Carsten Breuer besuchen ein Ausbildungsbataillon der Luftwaffe in Germersheim, 17. Juli 2025. Foto: © picture alliance / REUTERS | Heiko Becker.

SARAH BROCKMEIER-LARGE, CHRISTINA EDER, ALEXANDER JEDINGER, ALEXIA KATSANIDOU, LEON KRAUSE

„Nur 17 Prozent wären im Ernstfall bereit Deutschland zu verteidigen“, „Bundeswehr als Friedenstruppe in die Ukraine? Die Mehrheit ist dafür.“¹ Schlagzeilen wie diese begegnen uns fast täglich. Meinungsumfragen zu außen- und sicherheitspolitischen Fragen haben insbesondere seit der Vollinvasion der Ukraine durch Russland und der ausgerufenen „Zeitenwende“ in Deutschland an Sichtbarkeit und politischer Relevanz gewonnen. Die Bundeswehr wird massiv aufgerüstet, Politik und Gesellschaft

diskutieren die Wiedereinführung der Wehrpflicht und auch die Verteidigungsbereitschaft der Bevölkerung rückt in den Fokus. In diesem Kontext dienen Meinungsumfragen nicht nur der Einordnung gesellschaftlicher Stimmungen – sie beeinflussen auch politische Entscheidungen und verleihen ihnen Legitimität.² Je größer allerdings die Rolle von Meinungsumfragen in der außen- und sicherheitspolitischen Debatte wird, desto wichtiger ist es, dass deren Nutzer*innen ein Grundverständnis für ihre Aussagekraft und Grenzen haben.

„DIE“ ÖFFENTLICHE MEINUNG

„Öffentliche Meinung“ bezeichnet die kollektiven Einstellungen und Überzeugungen einer Bevölkerung zu politischen, sozialen oder anderen aktuellen Themen. Sie ist keine feste Größe, sondern eine Momentaufnahme, die im Zusammenspiel von Politik, Medien und Öffentlichkeit entsteht. Umfragen können Trends sichtbar machen und Gruppen vergleichen, sind aber stets nur ein Ausschnitt der öffentlichen Meinung. Deshalb ist es gerade bei komplexen Politikfeldern wie der Außen- und Sicherheitspolitik, wichtig, auf die Details zu achten.

GUTE UMFRAGEN ERKENNEN

Umfragen wirken oft eindeutig – doch wie belastbar ihre Ergebnisse sind, hängt entscheidend davon ab, wie die Daten erhoben wurden. Nur wenn zentrale methodische Informationen transparent berichtet werden, lassen sich Umfrageergebnisse sinnvoll einordnen (siehe *Reporting Standards* in Textbox). Folgende Informationen sind besonders relevant: *Wer wurde wie befragt?* Eine Umfrage ist nur dann

REPORTING STANDARDS

Reporting Standards für Umfragedaten

- ✓ Erhebendes Institut
- ✓ Endkunde/Sponsoren
- ✓ Grundgesamtheit/Zielgruppe
- ✓ Fallzahl/Stichprobengröße
- ✓ Erhebungszeitraum
- ✓ Erhebungsverfahren (z. B. mündlich, schriftlich, telefonisch, online, Mixed-Mode)
- ✓ Stichprobenziehung (z. B. Zufall, Quote, Vollerhebung)
- ✓ Stichprobenbasis
- ✓ Statistisches Fehlerintervall
- ✓ Fragen (genauer Wortlaut)
- ✓ Datengewichtung

Quelle: ADM

aussagekräftig, wenn die Befragten möglichst gut die Bevölkerung bzw. die für die jeweilige Erhebung relevante Gruppe widerspiegeln. Dafür sollte jede Person in der Grundgesamtheit die gleiche oder zumindest bekannte Chance haben, befragt zu werden. Ist das der Fall, spricht man von einer *repräsentativen Stichprobe*. Für eine solche sind *Zufallsstichproben* unumgänglich. Professionelle Umfrageinstitute greifen hierfür auf die Einwohnermeldedaten der Städte und Gemeinden zurück oder benutzen komplexe Auswahlmechanismen wie sie vom Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM) empfohlen werden.³

In der Praxis werden jedoch viele Befragungen inzwischen über sogenannte *opt-in Online-Panels* durchgeführt: Für diese registrieren sich Personen freiwillig und erhalten auf Grundlage festgelegter Kriterien Einladungen zu wissenschaftlichen oder kommerziellen Umfragen. Befragt werden kann hier also nur, wer im entsprechenden Panel registriert ist. Dies sind meist eher online-affine, jüngere und gebildete Menschen. Solche Stichproben können deshalb systematisch von der Gesamtbevölkerung abweichen und zu verzerrten Ergebnissen führen.

Auch klassischere Erhebungsmethoden wie persönliche Interviews („CAPI“), telefonische Befragungen („CATI“) oder postalische Anschreiben bringen jeweils eigene Herausforderungen mit sich – etwa geringe Befragungsbereitschaft oder hohe Erhebungskosten. Um diesen zu begegnen und zugleich eine möglichst gute Stichprobenqualität sicherzustellen, werden heute oft verschiedene Erhebungsmethoden kombiniert.⁴

Weitere Verzerrungsrisiken entstehen durch *Selbstselektion der Teilnehmenden* und *niedrige Rücklaufzeiten*. Wenn nur wenige der eingeladenen Personen antworten, besteht die Gefahr, dass die Ergebnisse nur die Ansichten einer kleinen Gruppe statt der Gesamtbevölkerung zeigen. Die Selbstselektion ist insbesondere bei offenen Online-Umfragen –

etwa auf Medienportalen – problematisch, denn es nehmen typischerweise eher Personen mit starker Meinung oder persönlichem Interesse teil.

Wann wurde befragt (Befragungszeitraum, „Feldzeit“)? Wurden die Daten vor dem interessierenden Ereignis, beispielsweise einer politischen Entscheidung, erhoben, kann aus ihnen keine Erkenntnis über die öffentliche Meinung zu diesem Ereignis gewonnen werden.

Wie wurde gefragt – und was wurde (nicht) verstanden? Nicht nur wer wann gefragt wird, sondern auch wie, beeinflusst die Ergebnisse. Zentral sind hierbei sogenannte *Framing-Effekte*: Je nachdem, welche Begriffe verwendet werden, welcher Aspekt betont oder wie die Frage eingebettet ist, können sich die Antworten unterscheiden – obwohl sich der inhaltliche Kern der Frage nicht ändert. Weitere typische Probleme, die zu Messfehlern führen können, sind:

- Suggestive Fragen, etwa mit Formulierungen wie „Finden Sie nicht auch, dass ...?“
- Unklare Begriffe wie „manchmal“ oder „selten“, unpräzise Zeitangaben oder sehr komplexe und kleinteilige Fragen.
- Fehlende Antwortoptionen oder unverständliche Skalen, die keine passende Antwort ermöglichen.
- Antwortverzerrung durch soziale Erwünschtheit.
- Frage- und Antwortreihenfolge: Früh genannte Möglichkeiten werden öfter gewählt und eine bereits gestellte Frage kann die Folgefragen beeinflussen.
- Ermüdung durch lange oder komplizierte Fragebögen, die zu Abbrüchen oder ungenauen Antworten führt.

„NON-ATTITUDES“ UND FRAGEFORMULIERUNG

Framing-Effekte und die genaue Frageformulierung sind gerade für Umfragen in der Außen- und Sicherheitspolitik relevant, denn eine zentrale Herausforderung in diesem Politikbereich sind Nicht-Einstellungen (*Non-attitudes*). Außen- und sicherheitspolitische Themen sind für große Teile der Bevölkerung in der Regel keine Fragen des Alltags. Das gilt insbesondere für spezifischere Fragen, wie zum Beispiel zu einzelnen Waffensystemen. Doch auch wenn sie über das Thema wenig wissen und keine gefestigte Meinung haben, drücken viele Befragte in Umfragen eine Meinung aus. Das zeigt ein Beispiel aus der jährlichen Bevölkerungsumfrage des Zentrums für Militärgeschichte und Sozialwissenschaften der Bundeswehr (ZMSBw). 46% der Befragten geben im Jahr 2024 an, „gar nichts“ über den Einsatz der Bundeswehr zur Kontrolle des Luftraums in Rumänien zu wissen, weitere 39% geben an, sie wissen „eher wenig“. Auf die Frage, was sie von diesem Einsatz der Bundeswehr halten, gibt die große Mehrheit der Befragten dennoch eine Meinung an. Lediglich 8% der Befragten antworten mit „ich weiß es nicht“.⁵

Das Beispiel illustriert: Befragte greifen auf andere Heuristiken als das eigene Wissen und gefestigte Meinungen zurück. Dazu gehören außenpolitische Grundeinstellungen – zum Beispiel zur Anwendung von Gewalt oder internationaler Kooperation – und persönliche Grundwerte.⁶ Einfluss-

reich sind aber auch sogenannte *elite cues*, also die Orientierung an der aktuellen medialen und politischen Debatte.⁷ Nicht zuletzt orientieren sich Befragte am Kontext der Umfrage und an der konkreten Frage. Das bedeutet: Das genaue Design der Studie, Frageformulierung und Antwortoptionen, die Reihenfolge der Fragen, der Zeitpunkt der Umfrage und der politische Diskurs zu diesem Zeitpunkt sind essenzieller Kontext für die Interpretation der Antworten.

Die nachfolgende Grafik illustriert die Relevanz der Frageformulierung anhand eines Vergleichs von regelmäßigen Umfragen der Körber-Stiftung, des ZMSBw und des Bundespresseamts (BPA). Alle drei Umfragen enthalten eine Frage, die die grundsätzliche Position der Bevölkerung zum außenpolitischen Verhalten Deutschlands erfragen möchte. Unterschiede in der Frageformulierung zwischen Körber-Stiftung und ZMSBw führten im Jahr 2023 zu einem genau gegenteiligen Ergebnis. Wenn wie in der Umfrage des BPA noch die Option dazu kommt, den Status Quo zu unterstützen, zeigt sich wieder ein anderes Bild. Auch methodische Unterschiede zwischen den Umfragen, die Fragereihenfolge und der genaue Erhebungszeitpunkt im Jahr 2023 können hier eine Rolle gespielt haben. Das Beispiel zeigt jedoch anschaulich: Die spezifische Frage und ihre Schlüsselwörter sind ein zentraler Punkt, den Konsument*innen von Umfragedaten zu Außen- und Sicherheitspolitik beachten sollten. Ob Institute nach „militärischer Unterstützung“ oder „Waffen“ für die Ukraine fragen, macht einen Unterschied – genau wie eine Frage nach der Entsendung von „Soldaten“ in die Ukraine nicht das gleiche ist wie die Frage nach „Friedenstruppen“. Ob der beliebte Verteidigungsminister in der Einleitung der Frage erwähnt wird, kann ebenso die Antwort beeinflussen, wie das Nennen (oder Weglassen) bestimmter Länder oder Krisen.

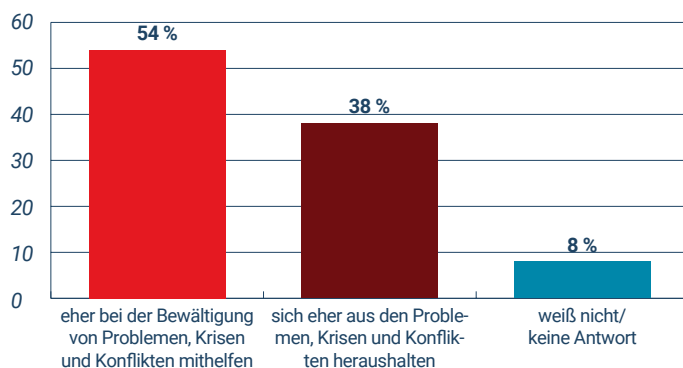
TYPISCHE INTERPRETATIONSFEHLER

Wer sich die Frage genau angeschaut und die Reporting Standards geprüft hat, muss noch mindestens drei typische Fehler bei der Interpretation von Umfragedaten vermeiden. Erstens sind *Fehlerintervalle* zu beachten. Unterschiede zwischen Gruppen oder Veränderungen über die Zeit, die innerhalb des Fehlerintervalls liegen, sind nicht aussagekräftig. Große Institute geben typischerweise Schwankungen von zwei bis drei Prozentpunkten an. Kleine Unterschiede, etwa zwischen Parteianhängern bei der Unterstützung der Ukraine mit Waffen, die in diesem Fehlerbereich liegen, können wie bedeutende Erkenntnisse wirken – sind es aber nicht. Ein zweiter häufiger Fehler ist die *Verwechslung von Korrelation und Kausalität*. Wenn Parteipräferenz etwa mit außenpolitischer Haltung zusammenhängt, heißt das nicht, dass das eine das andere verursacht. Menschen könnten in ihrer Meinung von einer Partei beeinflusst sein, sie könnten aber auch der Partei ihre Sympathie schenken, die ihren Standpunkt vertritt.

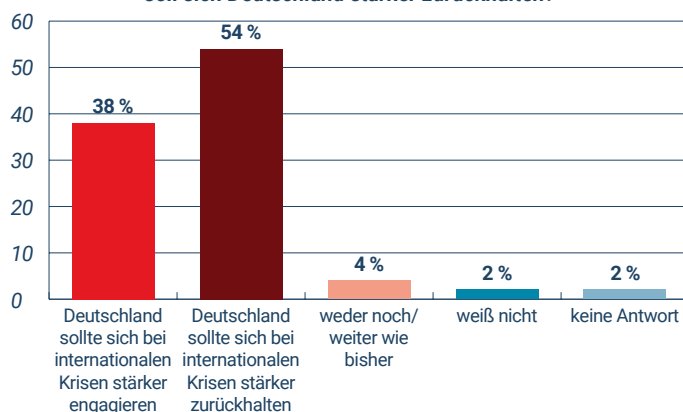
Ein letzter typischer Fehler in der Interpretation von Umfragedaten sind Aussagen über zu *kleine Gruppen*

oder *Subgruppen*. Aussagen über Subgruppen etwa nach Geschlecht, Parteipräferenz, Alter oder Region klingen interessant, basieren aber oft auf zu wenigen Fällen. Werden etwa innerhalb einer Umfrage zu deutschen NATO-Beiträ-

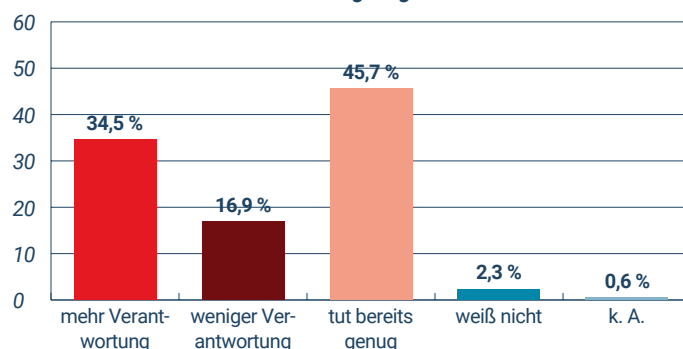
ZMSBw: Bevölkerungsbefragung 2023
Und wie sollte sich Deutschland in der internationalen Politik am ehesten verhalten? Sollte Deutschland...



Körber Stiftung: Deutsche Außenpolitik 2023
Es wird derzeit darüber diskutiert, ob Deutschland in Zukunft international mehr Verantwortung übernehmen soll. Was denken Sie: Sollte sich Deutschland künftig bei internationalen Krisen stärker engagieren oder soll sich Deutschland stärker zurückhalten?



BPA: Deutschland in der Welt 2023
Und wenn Sie an die aktuelle politische Lage denken: Sind Sie da der Meinung, dass Deutschland mehr Verantwortung oder weniger Verantwortung übernehmen sollte, oder sind Sie der Meinung, Deutschland tut bereits genug?



Grafik: Drei Fragen zur internationalen Verantwortung Deutschlands aus dem Jahr 2023 mit unterschiedlichen Formulierungen und Antwortoptionen. Details zu Erhebungszeitraum, Erhebungsmethode, Stichprobengröße und weitere Informationen finden sich in den Fußnoten im Online-Anhang.⁸

Sarah Brockmeier-Large leitet das Berliner Büro von PRIF. Dr. Christina Eder leitet das Forschungsdatenzentrum „Wahlen“ bei GESIS. Dr. Alexander Jedinger ist Senior Researcher bei GESIS. Prof. Dr. Alexia Katsanidou leitet die Abteilung Survey Data Curation bei GESIS. Leon Krause ist Doktorand am FGZ Frankfurt.

KONTAKT brockmeier-large@prif.org

PRIF, Darmstädter Landstr. 110–114,
60329 Frankfurt am Main
PVst, DPAG 43853, Entgelt bezahlt, ISSN-2512-627X

gen mit 1.000 Befragten die Antworten von FDP- oder BSW-Anhängern untersucht – beide Parteien stehen derzeit bei drei Prozent in der Sonntagsfrage – handelt es sich jeweils nur um einige Dutzend Befragte – zu wenig, um belastbare Aussagen über diese Gruppe zu treffen.

EINORDNEN STATT ÜBERBEWERTEN

Meinungsumfragen können wichtige Anhaltspunkte für politische Entscheidungen und die außen- und sicherheitspolitische Debatte liefern. Dafür ist es jedoch wichtig, dass sich Nutzer*innen dieser Daten der hier aufgeführten Punkte bewusst sind. Neben der grundsätzlichen Überprüfung der Qualität der Umfragedaten heißt das auch, die Ergebnisse von einzelnen Umfragen nicht überzubewerten und jedes Mal kritisch zu prüfen: In welchem Kontext wurde welche Frage gestellt? Was waren zu diesem Zeitpunkt die dominierenden Themen in der öffentlichen Debatte? Wie wahrscheinlich sind Non-attitudes und Framing-Effekte?

Des Weiteren sollte es – gerade auch für Journalist*innen – Standard sein, die Ergebnisse mit anderen Umfragedaten und idealerweise auch Zeitreihen zu vergleichen. Anhand verschiedener jährlicher Umfragen zur Außen- und Sicherheitspolitik wie etwa der Umfragen des ZMSBw, des Bundespresseamts, der Körber-Stiftung oder wiederholt gestellter Fragen der German Longitudinal Election Study (GLES) ist es möglich, ad-hoc Umfrageergebnisse in den Kontext von Zeitreihen zu setzen.⁹ Wer das eingangs erwähnte Beispiel, dass nur 17% der Deutschen bereit wären, Deutschland „mit der Waffe“ zu verteidigen, etwa mit den jährlichen Ergebnissen des ZMSBw vergleicht, stellt fest: Bei Männern unter 50 antworteten 2024 19% mit „ja“ und 42% mit „eher ja“ auf diese Frage. Eine informierte außen- und sicherheitspolitische Debatte braucht eine solche Einordnung.

Dieses Spotlight entstand im Rahmen des Leibniz-Lab „Umbrüche und Transformationen“ der Leibniz-Gemeinschaft.

PRIF SPOTLIGHT: Das Peace Research Institute Frankfurt (PRIF) ist das größte Friedensforschungsinstitut in Deutschland. PRIF analysiert die Ursachen gewaltsamer internationaler und innerer Konflikte, erforscht die Bedingungen des Friedens und arbeitet daran, den Friedensgedanken zu verbreiten.

V.i.S.d.P.: Elisabeth Waczek, Öffentlichkeitsarbeit (PRIF), Darmstädter Landstraße 110–114, Frankfurt am Main, Telefon (069) 959104-0, info@prif.org, www.prif.org. Design: Anja Feix · Layout: PRIF · Druck: Druckerei Spiegler

Textlizenz: Creative Commons (Namensnennung/Keine Bearbeitungen/4.0 International).

Die verwendeten Bilder unterliegen eigenen Lizenzbedingungen.



Fußnoten und weiterführende Links:
prif.org/spotlight0725-fn
DOI 10.48809/prifspot2507

