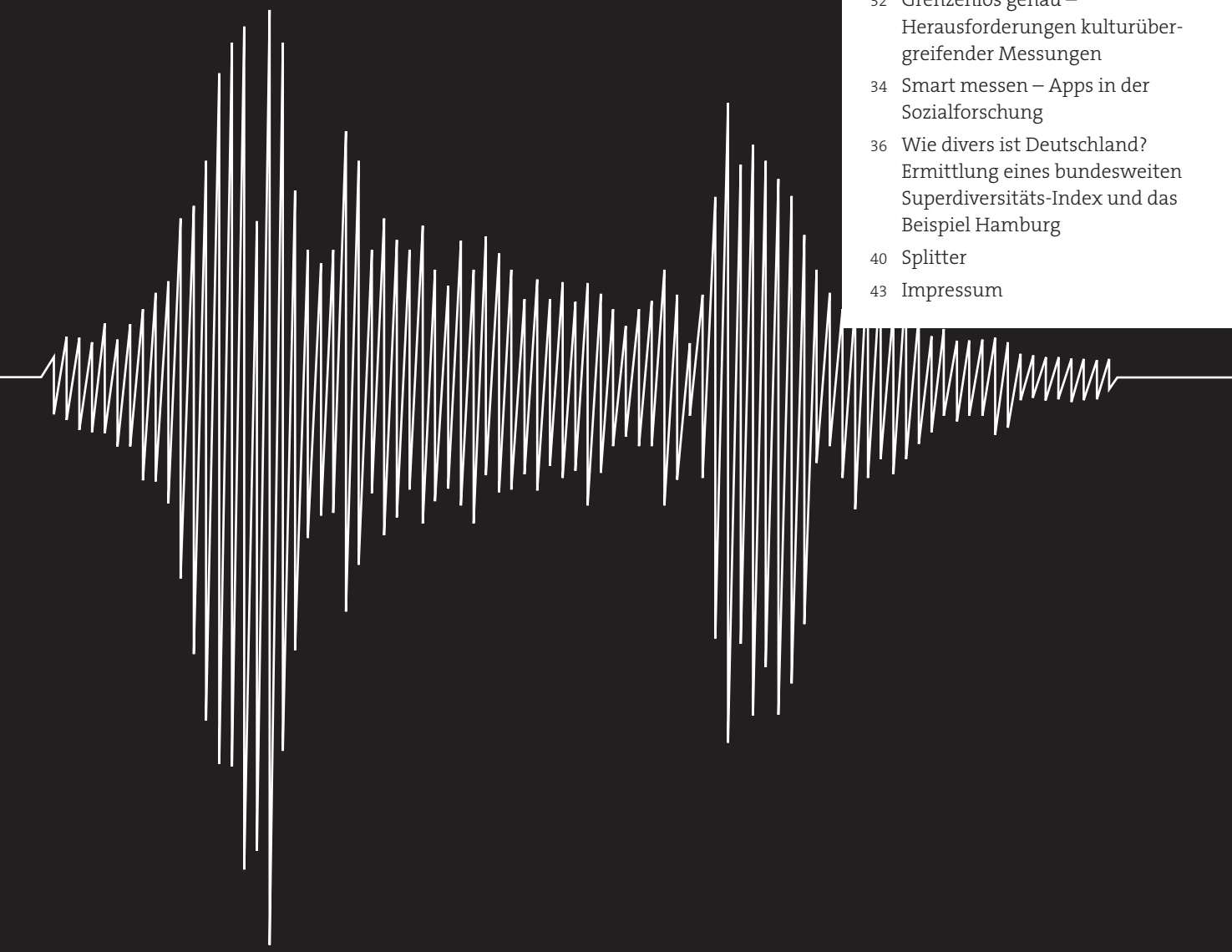
infas

Lagemaß

Messen – Denken – Forschen – Gestalten
Ausgabe 06 | 04.2018

3 Editorial
4 Erbsenzähler!
7 Transparente Dokumentation statt undurchsichtigem Flüssigbeton – ein Zwischenruf anlässlich der Bundestagswahl 2017
10 Vom Unbehagen an der Vielfalt – mit Sorgfalt gemessen
14 Sprechende Zahlen
16 Wege, dem Smartphone entlockt
17 messen
Lynx System Developers
29 Irreführungen, Intervalle, Trump und Brexit
32 Grenzenlos genau – Herausforderungen kulturübergreifender Messungen
34 Smart messen – Apps in der Sozialforschung
36 Wie divers ist Deutschland? Ermittlung eines bundesweiten Superdiversitäts-Index und das Beispiel Hamburg
40 Splitter
43 Impressum



m e s s e n

Editorial

Warum steht das neue Heft unter dem Thema „messen“? Die Antwort ist naheliegend: Messungen sind das Betätigungsfeld eines sozialwissenschaftlichen, empirischen Instituts wie infas. Insofern gibt das Heft einen Einblick in unser tägliches Tun: Wir legen Messsonden unterschiedlichster Art in unterschiedlichen Bereichen. Etwa wenn aktuelle Themen wie die Situation von Flüchtlingen, die Diversität in Deutschland, die Entwicklung unserer Alltagsmobilität oder die Verteilung von Bildungschancen betrachtet werden. Messen ist kompliziert, insbesondere in Zeiten, die von automatischen und quasi unbeabsichtigten Datensammlungen geprägt scheinen. Und nicht zuletzt: Wer misst, kann auch Fehler machen. Auch dies ist eine eigenständige Erörterung wert.

Doch die Vorteile sollten überwiegen. Aktuell bieten vor allem digitale Zugänge neue Chancen. So ermöglichen Apps auf dem allgegenwärtigen Smartphone Messungen, die zuvor nicht denkbar oder sehr aufwendig waren. Durch digitale Quellen werden Daten verfügbar, die bisher entweder nicht erfasst wurden oder in Leitzordnern im Dornröschenschlaf ruhten. Diese Datenquellen in die sozialwissenschaftliche Forschung zu integrieren, fordert vielfältig heraus. Dass dabei klassische Tugenden nicht vernachlässigt werden dürfen, versteht sich von selbst. Denn unverändert erfordern gute Studien viel Know-how, Messgenauigkeit, einen kritischen Blick auf die Details und Transparenz.

Und noch ein Wort zum Heft: Der Mittelteil dieser Ausgabe ist abweichend von vorherigen keine künstlerische Arbeit. Stattdessen geben wir einen Einblick in die Messung in der Welt des Sports, beispielhaft dargestellt am Zieleinlauf bei Wettkämpfen. Lassen Sie sich aber nicht täuschen. Was auf den ersten Blick wie ein schlichtes Foto aussieht, ist es nicht. Die Sache ist wie so oft komplizierter. Auf der letzten Seite des Mittelteils wird das Verfahren erklärt. Und nun viel Vergnügen bei der Lektüre rund um das Messen!

bcm
best of content marketing
Gold 2017

Lesen Sie mehr zum Award im Splitter auf Seite 40.

Ihr

Menno Smid
Geschäftsführer

Erbsenzähler!

Von uns akribischen Empirikern

Von Joachim Scholz

Wenn vom Messen in der empirischen Sozialforschung die Rede ist, rümpfen manche Naturwissenschaftler die Nase. Für sie ist Messen eine geplante Tätigkeit, mit der eine quantitative Aussage für eine Messgröße durch Vergleich mit einer Einheit ermittelt wird. Das Deutsche Institut für Normierung beschreibt Messen in DIN 1319 dann auch als „Ausführen von geplanten Tätigkeiten zu einer quantitativen Aussage über eine Messgröße durch Vergleich mit einer Einheit". So wird Temperatur in Grad Celsius, Kraft in Newton oder Lichtstärke in Candela gemessen.

Vergleichbare Einheiten fehlen der Sozialforschung. Gebremst hat das die empirische Sozialwissenschaft jedoch nicht: Ein zunehmender Bedarf an gesicherten Daten hat dazu geführt, dass quantitative Verfahren vorangetrieben wurden und die Welt auch aus dieser Perspektive vermessen wurde.

Items statt Einheiten

Weiterentwicklungen bei der Erhebung, Auswertung und Nutzung sozialwissenschaftlicher Daten haben für einen praxistauglichen Ersatz für die Einheit der Naturwissenschaft gesorgt: elaborierte Operationalisierungen und daraus abgeleitete Indikatoren. Beide weisen Gütekriterien auf – in erster Linie Objektivität, Validität (Gültigkeit) und Reliabilität (Verlässlichkeit). Im Zusammenspiel stellen sie sicher, dass bei der Messung identischer Sachverhalte oder bei der wiederholten Messung von Sachverhalten gleiche quantitative Werte ausgewiesen werden. Damit hat die Sozialwissenschaft ein Äquivalent geschaffen, das ähnliche Aufgaben wie naturwissenschaftliche Messungen zuverlässig erfüllt.

Noch vor etwa fünfzig Jahren im Rahmen des sogenannten Positivismusstreits war sich die Fachwelt durchaus uneinig, ob Messungen als naturwissenschaftliches Pendant in der Sozialwissenschaft überhaupt möglich und sinnvoll sind. Die Diskussion darüber wurde jedoch von der Praxis eingeholt. Quantitative Erhebungen haben ausgehend von Entwicklungen in der amerikanischen Soziologie der 1930er und 1940er-Jahre weltweit einen Boom erlebt, der unverändert anhält.

Natürlich wird seit vielen Jahren auch in Deutschland sozialwissenschaftlich gemessen. 1959 gegründet, zählt infas zu einem der Pioniere. 2016 erzielten die im Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM) organisierten Unternehmen 92 Prozent ihres Umsatzes mit quantitativer Forschung. Also mit einer Forschung, die sich eben genau auf die Annahme stützt, dass gesellschaftliche Sachverhalte gemessen werden können. Die Öffentlichkeit teilt den Eindruck, dass Sozialwissenschaftler messen und nennt sie gerne auch mal „Erbsenzähler".

Mit dem Anspruch einer exakten Messung in den quantitativ geprägten Erhebungen wuchs ein beträchtlicher Fundus an Forschungsverfahren heran, oft sogar ergänzt durch qualitative Verfahren, die wiederum einer eigenen Logik des Messens gehorchen. Die Stichprobenziehungen wurden optimiert, immer wieder neue Instrumente entwickelt. Beispielsweise Verfahren in der Ereignisanalyse oder zur Selbstpositionierung von Befragten im sozialen Raum oder mehrdimensionale statistische Analysen der häufig nur kategorialen Ergebnisse.

Heute gibt es zu den Erhebungsverfahren und Analysemethoden internationale Standards und für viele sozialwissenschaftliche Fragestellungen empirisch geprüfte Items und Skalen. So bietet als eine von vielen Quellen beispielsweise das GESIS Leibniz-Institut für Sozialwissenschaften mit der „Zusammenstellung sozialwissenschaftlicher Items und Skalen" (ZIS) über 200 Erhebungsinstrumente, etwa zum Messen von politischen Einstellungen oder psychologischen Merkmalen, zu denen die Entwicklung, die Gütekriterien und der

Meilensteine der sozialwissenschaftlichen Messung

1894 Franklin H. Giddings wird als erster Soziologe an die New Yorker Columbia University berufen, um systematisch die quantitative Analyse in die noch junge universitäre Soziologie einzuführen.

1897 Der französische Soziologe David Émile Durkheim veröffentlicht die empirische Studie „Der Selbstmord", für die er unter anderem 26.000 Akten zu Suizidfällen ausgewertet hat. Er untersuchte verschiedene Hypothesen zu den unterschiedlichen Suizidraten von Katholiken und Protestanten.

1906–1916 Mit „The Polish Peasant in Europe and America" von Florian W. Znaniecki und William I. Thomas entstand eine Studie, die die Motive und Einstellungen der polnischen Immigranten in den Vereinigten Staaten untersuchte.

1924–1933 Eine Reihe von Studien basieren auf den „Hawthorne-Experimenten", die in der Hawthorne-Fabrik der Western Electric Company in Chicago federführend von George Elton Mayo durchgeführt wurden. Sie waren Grundlage für die wissenschaftliche Auseinandersetzung mit Methodeneffekten.

1933 Paul Felix Lazarsfeld, seine Frau Marie Jahoda und Hans Zeisel ermittelten mit der Studie „Die Arbeitslosen von Marienthal. Ein soziographischer Versuch über die Wirkungen langandauernder Arbeitslosigkeit" die soziopsychologischen Wirkungen von Arbeitslosigkeit.

1958 Die Studie „The Blackcoated Worker" von David Lockwood analysiert die Identifikation von modernen Angestellten, die beispielsweise im Gegensatz zu Arbeitern nur ein „lauwarmes" Verhältnis zu Gewerkschaften haben.

1979 „Die feinen Unterschiede" ist der Titel des Hauptwerkes des französischen Soziologen Pierre Bourdieu. Anhand zahlreicher empirischer Untersuchungen zeigt er, dass Geschmack nichts Individuelles darstellt, sondern immer etwas von der Gesellschaft Geprägtes ist.

theoretische Hintergrund dokumentiert sind. Kurz, es ist eine umfassende Wissenschaft entstanden, in der es nur wenige weiße Flecken gibt.

Schwarze Schafe

Wo viel Licht ist, gibt es jedoch auch Schatten. Der geflügelte Satz „Traue keiner Statistik, die du nicht selbst gefälscht hast" kommt nicht von ungefähr. Im gleichen Maße, wie empirische Studien boomen und weltweit gefühlt sekündlich eine neue Untersuchung veröffentlicht wird, hat die Zunft mit Humbug zu kämpfen. Da geht es ihr nicht anders als den Naturwissenschaftlern, die ebenfalls mit fragwürdigen Studienergebnissen konfrontiert sind. Fachliches Unvermögen, versuchte politische oder wirtschaftliche Einflussnahme, ein aufgrund mangelnder finanzieller Ausstattung für die Forschungsfrage ungenügendes Studiendesign oder eine falsche, flüsterpostähnliche Darstellung der Studienergebnisse in den Medien sind einige Gründe. Dagegen helfen nur eine gute Praxis, hohe Transparenz und oft auch ein skeptischer Blick. In den nachfolgenden Artikeln dieser Lagemaß-Ausgabe werden einige sozialwissenschaftliche Fragestellungen aufgegriffen und gezeigt, welche besonderen Anforderungen sich für eine empirische Erhebung jeweils ergeben.

Fehler bleiben verborgen

Der Ehrgeiz, mit individuell zugeschnittenen anspruchsvollen Instrumenten die höchste Messgenauigkeit zu erlangen, wird in der sozialwissenschaftlichen Forschung nicht automatisch honoriert. Bedauerlicherweise ist den Ergebniszahlen ihre Entstehungsgeschichte oft nicht auf den ersten Blick anzusehen. Eine handwerklich solide empirische Studie unterscheidet sich hier nur bei genauem Hinsehen und guter Methodenkenntnis von Pfusch. Daher sind Transparenz und Nachvollziehbarkeit unverzichtbar.

Zu einem professionellen Selbstverständnis, dem wir uns bei infas verpflichtet sehen, gehören also Dokumentation und die Offenlegung von Entstehungsbedingungen. Doch auch damit bleibt es für den wissenschaftlichen Laien oft ähnlich schwer wie in den Naturwissenschaften, Zuverlässiges von weniger Gutem und dieses von Scharlatanerie zu unterscheiden. Obwohl in vielen Fällen bereits eine gute Dokumentation hilft, Missstände zu erkennen, ist die Überprüfung einer Studie häufig nur durch Replikation möglich und damit aufwendig und teuer. So können fragwürdige sozialwissenschaftliche Studien nur schwer frühzeitig als solche aufgedeckt werden. Dies gilt insbesondere dann, wenn zuverlässige Vergleichsdaten und Möglichkeiten für eine externe Validierung fehlen, die Hinweise zur Qualität der Ergebnisse liefern würden.

Daneben gibt es Entwicklungen, die seriöse empirische Sozialforschung erschweren. So sinkt die Teilnahmebereitschaft in der Bevölkerung seit einigen Jahren. Quoten von 70 Prozent und mehr, die in den 1980er-Jahren durchaus möglich waren, sind heute nur sehr sel-

ten und unter ganz besonderen Konstellationen zu erreichen. Es ist jedoch schwer zu messen, wenn sich das Forschungsobjekt dem Zugang verweigert.

Diese Entwicklung hat sich nicht plötzlich ergeben, sondern kontinuierlich über einen langen Zeitraum. Um damit umgehen und die geschilderten Gütekriterien trotzdem erfüllen zu können, hat die Sozialwissenschaft zusätzliche Instrumente entwickelt. Dazu rechnen beispielsweise Non-Response-Erhebungen, mit denen inhaltliche Ausfälle identifizierbar werden, oder verbesserte Stichprobenansätze. Der Einsatz dieser Verfahren erhöht jedoch den Aufwand und damit die Kosten. Und auch hier gilt: Auf den ersten Blick bleiben derartige Anstrengungen im Ergebnis kaum sichtbar. Eine Zahl ist eine Zahl. Sie muss aber einer Überprüfung auch auf den zweiten oder dritten Blick standhalten.

Blindes (falsches) Vertrauen

So erfreulich es ist, dass Studienergebnisse in der Bevölkerung eine hohe Aufmerksamkeit genießen, sollte es die Sozialforscher stören, dass das Wissen über die Entstehung von quantitativen Studien, zu möglichen Fallstricken und Fehlerquellen nur wenig entwickelt ist. Seit über 100 Jahren gibt es nun eine in ihren Grundprinzipien stabile, aber ihren Werkzeugkasten immer mehr verfeinernde empirische Sozialforschung. Und doch bleiben die Vorstellungen der Öffentlichkeit darüber vage.

Dies liegt keineswegs immer daran, dass die Vorgänge intransparent blieben – vor allem die Meilensteine und viele bedeutende Studien sind ausführlich dokumentiert. Doch nicht immer wird diesen Regeln entsprochen und nicht immer wird ihre Einhaltung eingefordert. Allerdings ist eine Handvoll knackiger Zahlen leichter zu kommunizieren als die zugehörige mitunter als überflüssig empfundene Entstehungsgeschichte. Nicht nur in Fake-News-Zeiten ist dieser Befund bedenklich.

Neuland mit Chancen

Transparenz ist auch deshalb nötig, weil die Welt der empirischen Sozialforschung nicht einfacher wird: Eine der größten aktuellen Herausforderung beim Messen in der Sozialforschung dürfte im Einbeziehen der Möglichkeiten durch neue digitale Datenzugänge liegen. Methodisch gedacht, rücken diese Quellen neben dem etablierten Befragen wieder die Beobachtung in den Fokus.

Eine eigentlich alte Erhebungsmethode, schließlich ist schon die 1933 von Marie Jahoda und Paul Lazarsfeld durchgeführte Studie „Die Arbeitslosen von Marienthal", die als Meilenstein der empirischen Sozialforschung gilt, keine reine Befragungsstudie, sondern basiert im Wesentlichen auf Beobachtung. Digital sind es heute das Smartphone, zahlreiche Geschäfts- und Telematikdaten, die zum Beobachten einladen. Zudem sind mehr und mehr amtliche Daten online verfügbar. Denn die Behörden stellen ihre Daten in großem Umfang der Öffentlichkeit zur Verfügung. Zusammen mit anderen Quellen ergibt sich ein immenser Datenfundus, den Sozialwissenschaftler bei der Beschreibung gesellschaftlicher Sachverhalte nicht außer Acht lassen können.

Die Analyse von Daten dieser Provenienz verlangt jedoch ein Umdenken. Das induktive Vorgehen, von einer Stichprobe auf die Gesamtheit zu schließen, passt hier zumindest nach dem ersten Eindruck nicht: Die grundsätzliche Vorgehensweise der Sozialforschung, zu einer Fragestellung Hypothesen zu entwickeln und mittels Falsifikation auf Basis empirischer Daten Kausalitäten zu identifizieren, kann nur eingeschränkt funktionieren. Schließlich handelt es sich um sekundäranalytische Auswertungen von Daten, die nicht für die Forschung entstanden sind, sondern im Rahmen von digitalen Alltagsprozessen.

Manche Big-Data-Apologeten schlagen hier daher einen umgekehrten Forschungsweg vor. Erst werden die Daten umfassend auf Korrelationen geprüft und danach werden Ideen zu den zugrunde liegenden Wirkungszusammenhängen entwickelt. Es wird zuerst die Frage nach dem „Was" beantwortet und erst anschließend nach dem „Warum". Ob ein derartiges Vorgehen, das der traditionellen Sozialforschungspraxis entgegenläuft, praktikabel ist, wird sich zeigen müssen. Zumindest sollte die Prüfung der eingangs genannten Kriterien Objektivität, Validität und Reliabilität Bestandteil der guten Praxis bleiben, wenn nicht Artefakte produziert werden sollen.

Zugewinn durch Digitalisierung

Die Einbeziehung dieser umfassenden Datenzugänge in sozialwissenschaftliche Studien ist eine neue Disziplin und bedarf der Entwicklung geeigneter Verfahren und Qualitätskontrollen. Der Zugewinn für die Forschung wird jedoch diese Anstrengungen rechtfertigen. Wie bei Innovationen üblich, werden voraussichtlich Anwendungen entwickelt, die den heutigen Erwartungshorizont übersteigen. Unzweifelhaft ist allerdings, dass die empirische Sozialforschung die Daten, die durch digitale Prozesse neu entstehen und gut verfügbar sind, nicht einfach ignorieren kann. Gesichert ist schon jetzt, dass das Messen in der Sozialforschung auch künftig eine spannende, Kreativität fordernde Aufgabe bleiben wird.

Zum Weiterlesen:
Diaz-Bone, Rainer (2010): Die Performativität der Sozialforschung – Sozialforschung als Sozio-Epistemologie, Workingpaper des Soziologischen Seminars der Universität Luzern, https://core.ac.uk/download/pdf/42106766.pdf
Menno Smid (Hrsg., 2009): Ein Hundert Prozent infas, infas, ISBN: 978-3941991002
Jochen Hirschle (2017): Im Schatten von Big Data? Die Sozialwissenschaften im Wandel, Deutsche Gesellschaft für Soziologie (Bd. 38), http://publikationen.soziologie.de/index.php/kongressband_2016/article/view/503

Transparente Dokumentation statt undurchsichtigem Flüssigbeton

Ein Zwischenruf anlässlich der Bundestagswahl 2017

Von Menno Smid

„Die AfD hat die Stimmen von 1.100.000 CDU-Wählern erhalten!“ Bei der Wahlberichterstattung der ARD zur Bundestagswahl 2017 waren diese und ähnliche Aussagen ein wesentlicher dramaturgischer Effekt des Moderators. Er bemühte sich damit durchaus erfolgreich, die verschiedenen Diskussionsrunden mit Politikern, bereits lange vor der Bekanntgabe des amtlichen Endergebnisses, anzuheizen.

Das verwundert nicht. Bei Wahlen interessieren sich Politiker und die Öffentlichkeit natürlich auch dafür, wie die Wähler zwischen den Parteien gewandert sind. Wer also anders gewählt hat als bei der vorherigen Bundestagswahl, ist von Belang. Oftmals werden weitgehende Schlussfolgerungen aus diesen Wechseln gezogen. Schließlich werden diese Wanderungsströme von Experten interpretiert und daran auch die Ursachen für Verluste und Gewinne von Wählerstimmen dingfest gemacht. Und es stellen sich durchaus interessante Fragen, zum Beispiel nach den Motiven, die den Wanderungen zugrunde liegen: Je nachdem, zu welcher Partei gewechselt wurde, unterscheidet sich die Argumentation. Das eher schlichte amtliche Wahlergebnis wird somit, so wird suggeriert, durch weitere, ebenso unanfechtbare Befunde erweitert. Die Wählerwanderungsanalyse selbst wird dabei nicht hinterfragt, sie wird faktisch akzeptiert wie das amtliche Endergebnis.

Doch wie werden die vorgeblich so objektiven Wanderungsströme zwischen den Parteien ermittelt? Diese Frage wird natürlich nicht am Wahlabend erörtert. Da kommt es eher auf die Wirkung der nackten Zahl an, die manchen Politiker durchaus zu verunsichern vermag, wie manche Statements andeuten. Dabei handelt es sich bei der Ermittlung von Wählerwanderungen um ein durchaus komplexes Geschehen, zu dem sich vertiefende Erläuterungen lohnen.

> An den Wanderungsströmen werden die Ursachen für Gewinne und Verluste von Wählerstimmen festgemacht. Doch wie werden sie eigentlich ermittelt?

Da ist zunächst das Wahlverhalten bei der vorherigen Wahl und natürlich das Ergebnis der neuen Wahl, sozusagen die interessierende Variable, abzubilden. Zwischen den beiden Wahlen haben Veränderungen der Rahmenbedingungen stattgefunden, die auch eine Bedeutung für die Wählerwanderungen zwischen den

Parteien haben: Es gibt Erstwähler, die per definitionem bei der vorherigen Wahl nicht teilgenommen haben, und es gibt Wähler, die an der vorherigen Wahl zwar teilgenommen haben, nicht aber an dieser Wahl, etwa aus Unlust oder auch weil sie verstarben. Des Weiteren sind objektive Veränderungen des Wahlkörpers, insbesondere bei Landtagswahlen, zu berücksichtigen.

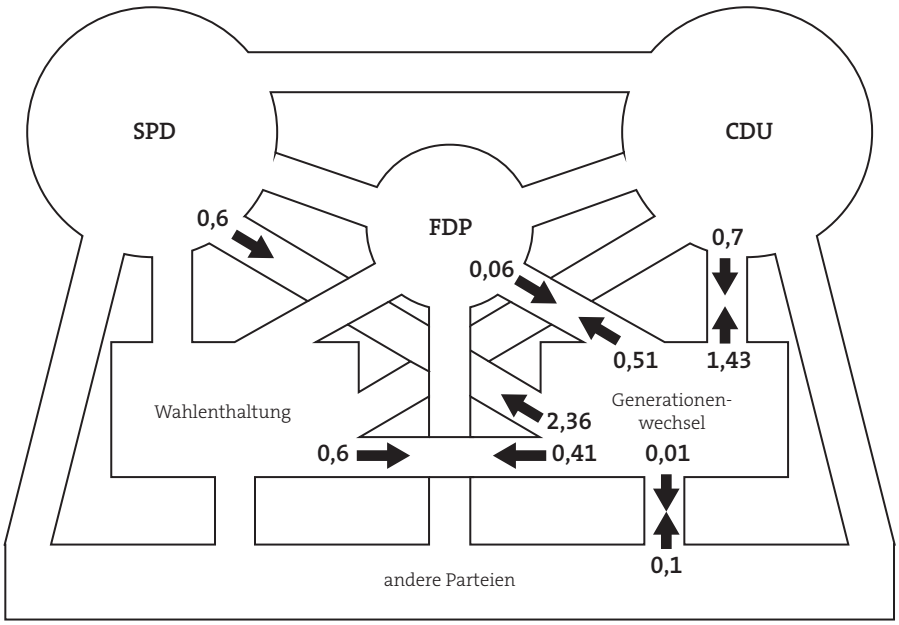
Die Angaben der Befragten bei Erhebungen für Wahlprognosen stimmen hochgerechnet nicht mit dem amtlichen Endergebnis der aktuellen und vorherigen Wahl überein.

Manche Wähler, die sich bei der vorherigen Wahl noch beteiligten, sind verzogen und nicht mehr im Bundesland ansässig. Umgekehrt sind natürlich auch Menschen in das Bundesland zugezogen, die bei der Vergleichswahl für die Analyse der Wählerwanderung fehlten.

Bemerkenswert ist, dass es keine Veröffentlichung der ARD oder des durchführenden Instituts Infratest Dimap gibt, die darstellt, wie und in welcher Weise die Wählerwanderungsanalysen durchgeführt werden. Auf der ARD-Webseite wahl.tagesschau.de finden sich allenfalls spärliche Informationen. Diese deuten darauf hin, dass im Zentrum der Wanderungsanalyse die Wahltagsbefragung steht. Das heißt, Wähler der repräsentativ ausgewählten Stimmbezirke werden am Tag der Wahl nicht nur danach gefragt, welche Partei sie aktuell gewählt haben – was nebenbei bemerkt die Grundlage für die Prognose um 18.00 Uhr ist –, sondern auch, welche Partei sie bei der vorherigen, vergleichbaren Wahl gewählt haben (Rückerinnerung). Damit gerät der Forscher allerdings direkt ins Zentrum eines nicht unerheblichen Messproblems. Kurz gesagt besteht es darin, dass die Angaben der Befragten zu der vorherigen Wahl hochgerechnet nicht mit dem tatsächlichen Wahlergebnis übereinstimmen. Teilweise kommt es hier zu beträchtlichen Abweichungen. Aggregiert müsste im Idealfall die Angabe jeder einzelnen befragten Person zu ihrem Wahlverhalten das vorherige wie eben auch das aktuelle Ergebnis exakt wiedergeben. Dies ist aber aus unterschiedlichen Gründen nicht der Fall, wie in wegweisenden Studien der 70er Jahre unter anderem Max Kaase, Hans-Dieter Klingemann und Franz Urban Pappi nachgewiesen haben.

Das Problem liegt unter anderem in der Rückerinnerung: Entweder wollen oder können sich manche Befragte nicht erinnern, oder sie antworten in einem selbst inter-

Beispiel für eine Wählerwanderungsanalyse von 1972
Wahlenthaltungen und Generationenwechsel sind berücksichtigt



Wahlwanderungs-Index
Quelle: infas 1972

infas

pretierten, sozial erwünschten Sinn und bekennen sich nicht zu ihrem tatsächlichen Wahlverhalten. Dieses Antwortverhalten betrifft auch die Frage, ob sie überhaupt an der vorherigen Wahl teilgenommen haben, also die Ermittlung der überaus wichtigen Gruppe der Nichtwähler. Eigene aktuelle Erhebungen zeigen, dass sich an dem Messproblem nichts geändert hat. Im Gegenteil, es hat sogar den Anschein, dass wie auch immer aufgefasste „sozial erwünschte" Antworten häufiger sind als früher.

Aber das ist nur eines von mehreren Messproblemen: Das Wahlverhalten der Verzogenen kann nicht abgefragt werden, ebenso wenig wie das der Verstorbenen bei der vorherigen Wahl. Die Zugezogenen sind zwar bei der aktuellen Wahl beteiligt, nicht aber bei der vorherigen Vergleichswahl. In allen diesen Fällen muss also mit Annahmen oder Ergebnissen besonderer Analysen gearbeitet werden. Hier wäre es äußerst interessant, welche Annahmen getroffen wurden und wie sie sich auf die Fehlerwahrscheinlichkeit des Modells auswirken. Im Ergebnis müssen die genannten verschiedenen Datenlücken und die abweichenden Befragungsdaten durch einen Anpassungsalgorithmus, eine Kalibrierung, modelliert werden. Der entscheidende Punkt dabei ist: Bei den ermittelten Wählerwanderungen handelt es sich nicht um nachprüfbare harte Ergebnisse. Sie basieren vielmehr auf einem statistischen Erklärungsmodell, dessen Ergebnisse nur mit einer gewissen Fehlerwahrscheinlichkeit stimmen können. Darüber wird jedoch im Rahmen der Berichterstattung nicht aufgeklärt und vielmehr der Eindruck vermittelt, es handle sich um harte Ergebnisse ähnlich der amtlichen Auszählung.

In den 70er Jahren endet die Diskussion

In der Bundesrepublik hat das infas-Institut die Analyse der Wählerwanderungen auf Basis von Befragungen in den 70er Jahren eingeführt. Dazu wurde rund zehn Jahre experimentiert, bis die Grundlagen in einem Artikel von Krauß/Smid offengelegt und kritisch diskutiert wurden. Offenbar war die sich daraufhin entwickelnde Diskussion die bisher letzte, die zu diesem Thema geführt wurde. Es mutet daher einigermaßen eigentümlich an, wenn die ARD in ihrer Wahlberichterstattung ohne Bezug zu bisherigen Diskussionen und ohne die Offenlegung der Grundlagen so tut, als seien ermittelte Wählerwandungen nicht zu hinterfragende Befunde, über die es nur schlicht zu berichten gilt. Vielmehr wäre ein weiterer wissenschaftlicher Diskurs zu erwarten. Denn es ist mehr als unwahrscheinlich, dass und 50 Jahre nach der Einführung von Analysen von Wählerströmen keine Weiterentwicklung möglich und sicherlich auch nötig wäre.

Volatile Ergebnisse

Ein letzter Punkt zum Thema: In den spärlichen Informationen der ARD zur letzten Bundestagswahl 2017 wurde darauf verwiesen, dass sich die Wanderungsströme im Laufe des Wahlabends mehrmals ändern können. Und in der Tat: Am Ende des Wahlabends hieß es dann: „Die AfD hat die Stimmen von 980.000 CDU-Wählern erhalten!" Immerhin ein Unterschied von 120.000 Stimmen zu der anfangs genannten Schätzung von 1.100.000 CDU-Wählern. Vernachlässigbar? Fraglich. Immerhin werden auf Basis dieser Daten politische Interpretationen formuliert. Wäre es nicht sinnvoller, nur über gesicherte Messungen zu berichten, auch wenn das schwerfällt, weil diese erst am späten Wahlabend zu erwarten sind? Zumal Wahlprognosen in der Bevölkerung aufgrund von deutlichen Abweichungen zum Endergebnis zuletzt an Ansehen verloren haben. Wanderungsanalysen, deren Ergebnisse sich im Zeitverlauf mehrfach ändern, tragen da nicht zur Vertrauensbildung bei. Eine methodisch abgesicherte, transparente Hintergrundbetrachtung würde vielleicht mehr Interessenten finden als manche schnelle und nicht immer aufklärende Zahl.

Der WDR-Fernsehdirektor und wesentliche Präsentator der Wahlberichterstattung in der ARD, Jörg Schönenborn, wies in einem Vortrag in der Akademie für politische Bildung in Tutzing auf sein Privileg hin, Fragen, die er als Journalist für interessant halte, beim Umfrageinstitut Infratest Dimap in Auftrag geben zu können. Umfragen seien wie Flüssigbeton. Es komme darauf an, was man daraus macht. Ob das im Sinne einer wissenschaftlich fundierten Messung ist, die eigentlich Grundlage jeder Wahlberichterstattung sein sollte, und zu erhellenden Ergebnissen führt, darf durchaus bezweifelt werden.

Zum Weiterlesen:
Hans-Dieter Klingemann und Franz Urban Pappi: „Die Wählerbewegung bei der Bundestagswahl am 28. September 1969", in: PVS 11 (1970), S.111-138.
Max Kaase: „Determinanten des Wahlverhaltens in der Bundestagswahl 1969", in: PVS 11, (1970) S. 46-110.
Fritz Krauß und Menno Smid: „Wählerwanderungsbilanzen. Ein Vergleich verschiedener Ansätze am Beispiel der Bundestagswahl 1980" in: Zeitschrift für Parlamentsfragen (ZParl) Heft 1/81, S. 83-103

Vom Unbehagen an der Vielfalt – mit Sorgfalt gemessen

Von Robert Follmer, Jette Kellerhoff und Fridolin Wolf

Zu populistischen Politikmustern und ihrer Akzeptanz in der Bevölkerung sind in den vergangenen Monaten verschiedene, auch empirisch ausgerichtete Publikationen erschienen. Dabei werden unterschiedliche Definitionen und Begriffsverständnisse genutzt. Zudem sind einigen der empirischen Annäherungen Grenzen gesetzt, wenn das Segment möglicher „Populisten" genauer unter die Lupe genommen werden soll.

Derartige Begrenzungen sind auch auf forschungsökonomische Aspekte zurückzuführen – etwa bei Stichprobenumfängen von beispielsweise 1.000 oder 2.000 Befragten in bundesweiten repräsentativ angelegten Befragungen. Dies bringt für eine differenzierte Betrachtung schon deshalb Einschränkungen mit sich, weil das Segment der „Populisten" oder ähnlicher Gruppen in solchen Studien nur eine relativ geringe Anzahl von Befragten aufweist.

Ein weiteres Problem besteht möglicherweise in der Erreichbarkeit der zu betrachtenden Gruppe potenzieller „Populisten". Ein kennzeichnendes Merkmal ist in der Regel ihre unterdurchschnittliche Teilnahme am gesellschaftlichen Diskurs. Wenn dazu auch die Beteiligung an einer Befragung gezählt wird – wofür es manche Argumente gibt –, kommt es nicht nur auf eine umfangreiche, sondern auch auf eine sorgfältig gebildete und geprüfte Stichprobe an, um zu belastbaren Aussagen über diese Einstellungsmuster und ihre Hintergründe zu gelangen.

Um hierzu einen Diskussionsbeitrag zu leisten, haben wir auf der Grundlage einer Befragung, die infas bei 5.041 Bürgerinnen und Bürgern im Auftrag der Bertelsmann Stiftung im Frühjahr 2017 durchgeführt hat, eine entsprechende Sekundäranalyse vorgenommen. Diese Studie hatte primär keine politikwissenschaftlichen Fragestellungen. Gleichwohl thematisiert sie Aspekte aus dem Umfeld der Forschung zu populistisch geprägten Entstehungsmustern sowie möglichen Vorbehalten gegenüber einer vielfältigen Gesellschaft, auf die wir uns im weiteren Verlauf genauer konzentrieren werden. Der Zugang zu dieser Befragung erfolgte im Kontakt mit den Befragten themenneutral, ließ also diese Inhalte aufseiten der Befragten nicht unbedingt vermuten. Sie bietet darüber hinaus ein besonders großes Sample mit einer in 80 Teilregionen erfolgten sorgfältigen Stichprobenkontrolle. Diese Prüfung schloss auch Bildungs- und weitere Merkmale ein, die sich in entsprechenden Analysen immer wieder als erklärungskräftig mit Bezug auf den Populismus erwiesen haben.

Was ist überhaupt Populismus und welche Rolle spielt die Vielfalt?

Dass eine sorgfältige Messung hier wichtig ist, hat Gründe. Denn eine angemessene Definition von Populismus sowie die Operationalisierung der damit verbundenen antipluralistischen Einstellungsmuster stehen vor der Herausforderung, dass das Etikett „Populist" in der Öffentlichkeit schnell vergeben und oft nur grob umrissen wird. Im wissenschaftlichen Diskurs besteht indes Konsens darüber, dass Populismus durch drei Dimensionen näher bestimmt ist:

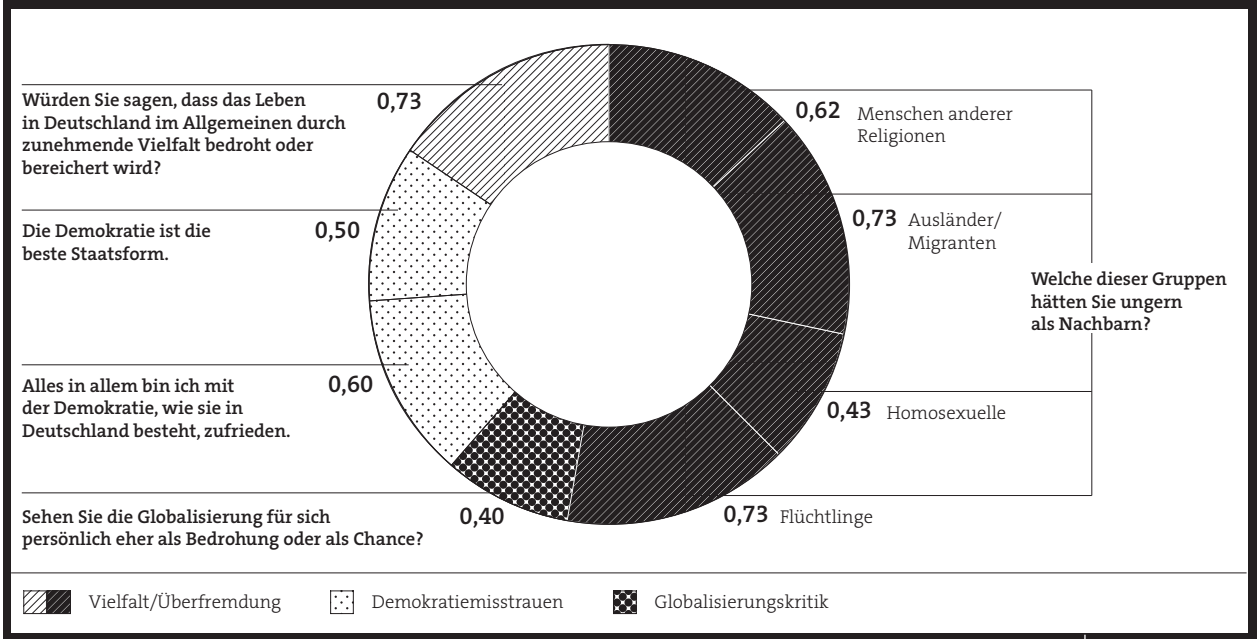
- Vorbehalte gegen Vielfalt und demokratische Orientierungen (Antipluralismus),
- der Anspruch, alleine den „Volkswillen" auszudrücken (Alleinvertretungsanspruch),
- Vorbehalte gegen die etablierten politischen Akteure (Antiestablishment und Skepsis gegenüber Institutionen).

Antipluralismus als ein Bestandteil ist demnach ein allgemeines Kennzeichen populistischer Orientierungen, aber er muss nicht notwendig zu ihnen führen. Andere Merkmale müssen hinzukommen. Während eine demo-

Studie: Vom Unbehagen an der Vielfalt

- Durchführung im Auftrag der Bertelsmann Stiftung
- 5.041 Befragte ab 16 Jahren
- telefonische Interviews (Dual-Frame-Stichprobe, also Festnetz und Mobilfunk, nach ADM-Design)
- Befragungszeitraum: Frühjahr 2017
- Regionalisierung der Stichprobe in eigens für dieses Projekt gebildeten Raumordnungsregionen, entwickelt von infas 360 GmbH
- Interviewlänge etwa 30 Minuten
- Fragebogenentwicklung durch die Bertelsmann Stiftung sowie die Jacobs University Bremen
- Stichprobenkonzeption und Durchführung der Erhebung durch infas
- Bildung eines mehrdimensionalen Indexes zum Antipluralismus auf Basis einer faktoranalytischen Gewichtung der einfließenden Einzelfaktoren
- dargestellt wird die Kurzfassung eines längeren Artikels (siehe auch zum Weiterlesen)

Dimensionen und Indexbildung „Antipluralismus“: Demokratieverständnis, Globalisierungsangst und Vorbehalte gegen Minderheiten



Die Zahlenangabe neben den einzelnen Ringabschnitten stellt die Gewichtung des jeweiligen Items im Gesamtindex auf Basis einer Faktorenanalyse dar. Je näher der Wert an einem Betrag von 1 liegt, desto stärker ist die Bedeutung im Index. Zur weiteren Orientierung geben die Schraffuren den jeweiligen inhaltlichen Schwerpunkt an.
Quelle: Vom Unbehagen an der Vielfalt, Januar 2018

infas

kratische Auseinandersetzung die Anerkennung unterschiedlicher Meinungen und Interessen voraussetzt, legt die populistische Form eine Freund-Feind-Orientierung zugrunde. Diese kann verschiedene Richtungen einnehmen. Verbindet sich die vertikale Ebene („die da oben“ – „wir hier unten“) mit einer horizontalen („wir“ – „die anderen“), lässt sich von rechtspopulistischen Einstellungen sprechen. Konstitutiv ist dabei die Annahme einer homogenen Gruppe („das Volk“), die eine mehr oder minder scharfe Ausgrenzung derer erlaubt, die nicht dazugehören (vgl. van de Wetering 2016). Solche politischen Einstellungen tendieren zum Rechtsradikalismus, aber sie sind so lange nicht rechtsextremistisch, wie sie nicht aktiv gegen die freiheitlich-demokratische Grundordnung gerichtet sind. Unter anderem müsste also eine höhere Gewaltaffinität hinzukommen.

Dabei kristallisieren sich Einstellungsmuster heraus, die zur Operationalisierung antipluralistischer Haltungen herangezogen werden können. Dazu zählen etwa die Wahrnehmung einer Gefährdung der nationalen Identität und Bedrohung durch Vielfalt, Szenarien der Überfremdung, das Negativbild des institutionellen und politischen Establishments sowie die Ablehnung von sozialen und ethnischen Minderheiten. Ebenso gehören demokratieskeptische Überzeugungen dazu.

Um hierzu genauere Erkenntnisse zu gewinnen, wurden aus dem Regionalsurvey Fragen herangezogen, die die Bewertung von Vielfalt, Demokratiemisstrauen, Globalisierungskritik, die Angst vor Überfremdung und die Abwertung von bestimmten gesellschaftlichen Teilgruppen wie Ausländern oder Homosexuellen messen. Indikatoren für Einstellungen gegen das politische Establishment, die für die Populismusdefinition erforderlich wären, waren im Datensatz nicht verfügbar.

Aus den ausgewählten Fragen haben wir ausgehend von den dargestellten Überlegungen einen gewichteten, additiven Gesamtindex zum Themenkomplex „Antipluralismus“ gebildet. Die zugrunde liegenden Fragen fließen mit unterschiedlichen analytischen Gewichten entsprechend ihrer Erklärungskraft ein. Eine vergleichsweise hohe Erklärungskraft erlangen in diesem Zusammenhang die Merkmale zur Akzeptanz von Diversität sowie sozialen und ethnischen Minderheiten. Um im nächsten Schritt trennende und verbindende Faktoren zwischen Befragten mit einem ausgeprägten antipluralistischen Meinungsbild und solchen mit geringeren Vorbehalten aufzuspüren, haben wir ein mehrdimensionales Analysemodell entwickelt, das soziale Vernetzung und Integration, Lebenszufriedenheit, lokale und regionale Verbundenheit, politische Partizipation sowie Einstellungen zu Institutionen, Politik und Gesellschaft misst. Dazu haben wir analog zum dargestellten Antipluralismusindex weitere elf Dimensionsindizes konzipiert. Sie fußen auf einem Merkmalset von insgesamt 73 Fragen, von denen keine in die zuvor gebildete Operationalisierung des Antipluralismusindexes eingeflossen ist. Grau hinterlegt sind die verwendeten Dimensionen, denen stichwortartig die für die Operationalisierung genutzten Fragen zugeordnet sind.

Für die weiteren Betrachtungen stehen nun der Antipluralismusindex sowie die elf erklärenden Dimensionen zur Verfügung. Das Indexergebnis nutzen wir, um die Personengruppe genauer in den Blick zu nehmen, die eine höhere Affinität zu einem antipluralistischen Meinungsbild hat. Zu dieser Gruppe der „Antipluralisten" zählen wir alle Befragten mit einem überdurchschnittlichen Indexwert – insgesamt sind das 39 Prozent. Das heißt: Vier von zehn Befragten haben eine überdurchschnittliche Sympathie für antipluralistische Aussagen. Es handelt sich hier also nicht um ein Minderheitenphänomen, sondern um Haltungen, die in der Mitte der Gesellschaft zu finden sind. Dabei bedeutet ein überdurchschnittlicher Indexwert inhaltlich nicht, dass der oder die Befragte allen im Index verwendeten Aussagen zustimmt. Ausschlaggebend ist lediglich eine im Schnitt höhere Anzahl von Zustimmungen zu den einzelnen Merkmalen – also eine „Affinität" im Sinne einer relativ weit abgesteckten Gruppe.

Signalereignis Flüchtlingswelle

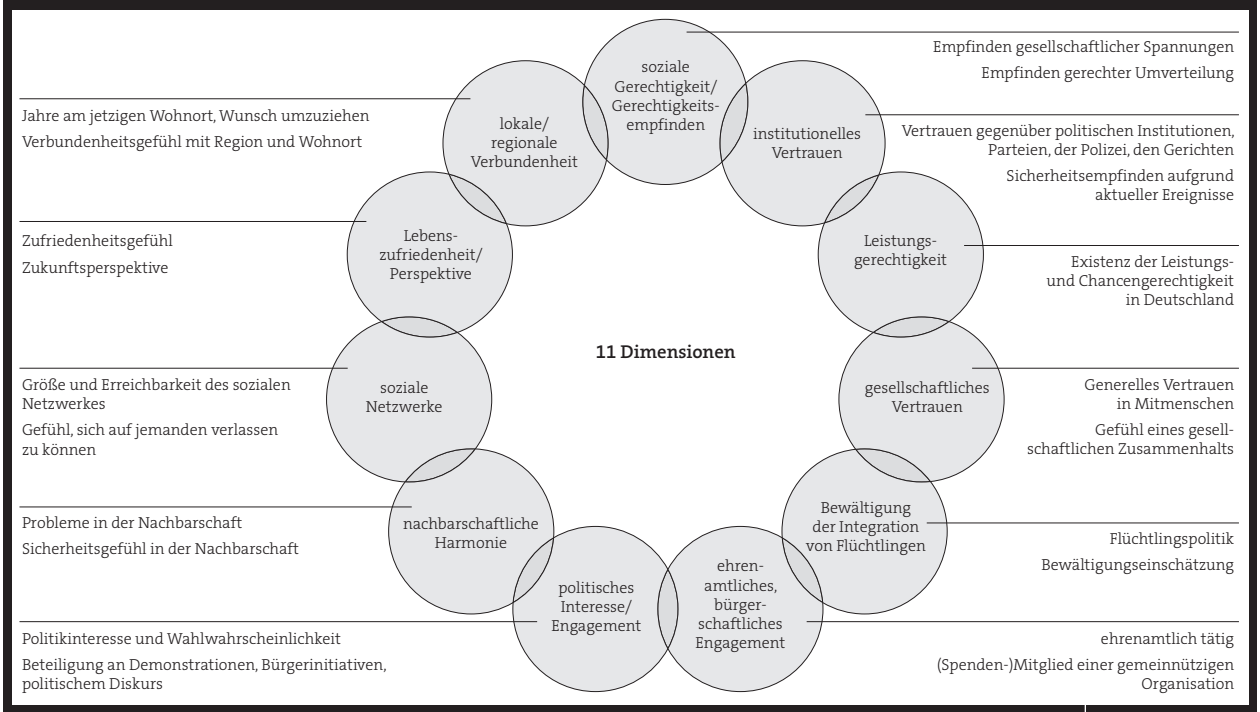
Die Ergebnisse unserer Studie weisen bei einem nicht unerheblichen Teil der Befragten ein tendenzielles Unbehagen gegenüber Vielfalt auf. Eine über dem Durchschnitt liegende Affinität zu antipluralistischen Einstellungen findet sich bei etwa 40 Prozent der Befragten. Ängste vor dem Fremden spielen dabei ebenso eine Rolle wie die Sorge vor wachsender Benachteiligung und Konkurrenz um knappe Ressourcen wie Arbeit und Wohnung. Die jüngste Flüchtlingseinwanderung, die binnen kürzester Zeit eine große Zahl von Menschen vor allem aus muslimisch geprägten Ländern nach Deutschland brachte, war ein Signalereignis. Ängste und Sorgen bekamen einen Namen und die in Teilen der Bevölkerung vorhandene Verunsicherung wurde sichtbar.

Diese Sicht von oben auf die Befragten mit überdurchschnittlich ausgeprägten antipluralistischen Überzeugungen wird allerdings den vielfältigen Wirkzusammenhängen und Konfliktlinien nicht gerecht, die die bei genauem Hinsehen heterogenen Einstellungen innerhalb dieser Gruppe prägen. Hier finden sich einerseits Personen „auf der Schwelle" und andererseits tief überzeugte Vertreter antipluralistischer Ansichten mit oft als benachteiligt empfundenen Lebensbedingungen. Auch das vielfach bemühte Bild der „Modernisierungsverlierer" oder der „Wutbürger" greift an dieser Stelle zu kurz. Vielmehr sind tief sitzende Einstellungsmuster, Irritationen und Ängste zu beobachten, denen es adäquat zu begegnen gilt.

Weitere Segmentierung der Antipluralisten

Zugleich stimmen die Ergebnisse optimistisch über das Ausmaß des engeren, populistisch orientierten Antipluralismus in Deutschland: Dies zeigt eine Segmentierung, die innerhalb der Gruppe der „Antipluralisten" vorgenommen wurde. Dazu wurde eine Clusteranalyse anhand der elf beschriebenen Dimensionen genutzt. So beläuft sich der innerste Kern der „Ausgegrenzten" auf etwa vier Prozent der Befragten.

Zusammensetzung der erklärenden Dimensionen: Integration, Geborgenheit, Zukunft und Chancen



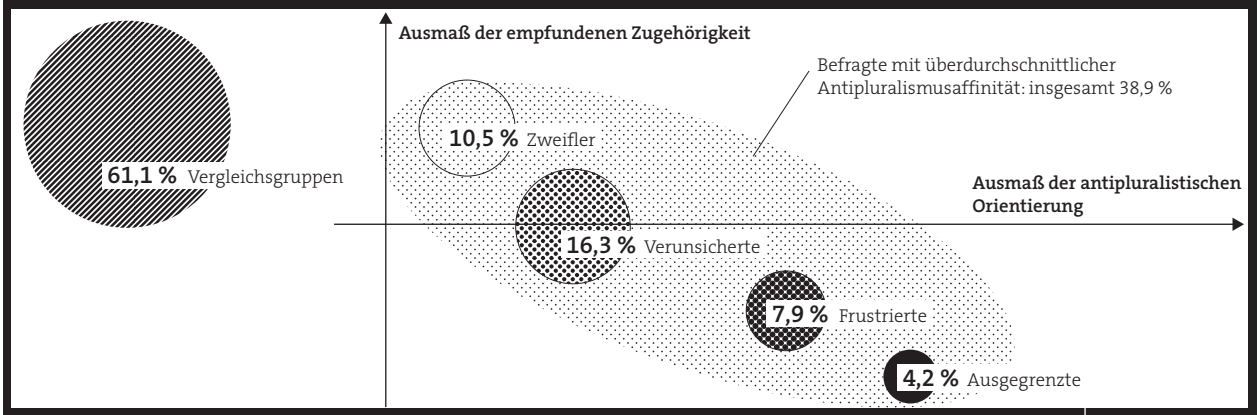
Quelle: Vom Unbehagen an der Vielfalt, Januar 2018

infas

Ihnen müssen eher rechtsorientierte und stabile Überzeugungen gegen gesellschaftliche Vielfalt attestiert werden. Zusammen mit einer weiteren Gruppe der „Frustrierten“ sind es etwa 12 Prozent der Befragten, denen eine verringerte Lebenszufriedenheit, eine pessimistische Lebenseinstellung, persönliche und gesellschaftliche Frustrationserfahrungen, eine fehlende soziale Einbindung sowie unzureichende materielle Ressourcen gemein sind. Dabei entsprechen ihre subjektiven Einschätzungen – etwa in Bezug auf die Flüchtlingseinwanderung – häufig nicht den objektiven Gegebenheiten vor Ort. Der Blick auf die Wohnsituation insbesondere bei den Frustrierten zeigt jedoch, dass es objektiv benachteiligende Faktoren gibt. Dabei gilt es auch zu berücksichtigen, dass vor allem Vertreter der beiden engeren antipluralistischen Gruppen Gesellschaft und Politik für ihre individuelle Lebenssituation verantwortlich machen. Dementsprechend erwarten sie von den gesellschaftlichen und politischen Institutionen Unterstützung. Bleibt diese Unterstützung tatsächlich oder in ihrer subjektiven Wahrnehmung aus, befördert dies weitere Frustrationen.

Was der Blick auf die beiden besonders antipluralistischen Segmente der Frustrierten und Ausgegrenzten auch zeigt, ist, dass es deutliche Zusammenhänge mit einer schwachen regionalen und nachbarschaftlichen Einbindung und einem unterdurchschnittlichen Gefühl des Aufgehobenseins gibt. Unabhängig von aktuellen politischen Herausforderungen wie der Flüchtlingsintegration wirken hier also offenbar auch andere gesellschaftliche Trends, die die Vereinzelung vorantreiben und jedenfalls für manche die Adaption an Veränderungen, wie eine wachsende Vielfalt, erschweren.

Segmente innerhalb der „Antipluralisten“ – Ausmaß des Antipluralismus vs. Zugehörigkeitsgefühl: Nur bei etwa jedem Zehnten ist die antipluralistische Einstellung verhärtet



Quelle: Vom Unbehagen an der Vielfalt, Januar 2018

infas

Viele Antipluralisten zweifeln oder sind verunsichert

Etwas im Gegensatz zu diesen beiden Gruppen stehen die „Verunsicherten“ und die „Zweifler“. Sie sind gesellschaftlich integriert und gut vernetzt – dies ist ihr Schutzpotenzial vor einer Verfestigung ihrer antipluralistischen und in Teilen bereits rechtspopulistischen sowie demokratieskeptischen Haltung. Wirtschaftlich und gesellschaftlich haben insbesondere die Vertreter der Zweifler vergleichsweise häufig das erreicht, was sie angestrebt haben. Aus diesen sozialen Umständen resultiert bei beiden Gruppen eine zufriedene und optimistische Lebenseinstellung, die auch auf die Gesellschaft zurückwirkt. Allerdings fühlen sich Vertreter aus diesen beiden Segmenten durch aktuelle Ereignisse wie die Aufnahme der Flüchtlinge in ihrer Lebenswelt verunsichert. Die tatsächlich oder vermeintlich vergangene Sicherheit gilt es zurückzugewinnen: Denn sie stellen aktuell die Mehrheit der Gruppe der potenziellen rechtspopulistischen Wähler. Im Gegensatz zu den Frustrierten oder Ausgegrenzten handelt es sich bei den Verunsicherten und noch mehr bei den Zweiflern um Personengruppen, die den politischen Diskurs und die Bürgerbeteiligung eigentlich suchen. Falsch wäre es, sie aus der politischen Auseinandersetzung auszuschließen und ihre Verunsicherung und Forderungen als randständig abzuwehren.

Wir danken der Bertelsmann Stiftung und Kai Unzicker für die Unterstützung beim Zustandekommen dieser Publikation.

Zum Weiterlesen
„Vom Unbehagen an der Vielfalt“, (Langfassung): https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/vom-unbehagen-an-der-vielfalt/

Baumann, Zygmunt (2016): Die Angst vor den anderen. Ein Essay über Migration und Panikmache. Edition Suhrkamp, Berlin.
Crouch, Colin (2008): Postdemokratie. Edition Suhrkamp, Frankfurt am Main.
Müller, Jan-Werner (2016): Was ist Populismus? Edition Suhrkamp, Berlin.
van de Wetering, Denis (2016): Rechtspopulistische Gegenwelt: Aktuelle Kommunikationsformen, gesellschaftliche Resonanz und demokratische Antworten, Bertelsmann Stiftung (Hrsg.), Gütersloh.

Sprechende Zahlen: Messen

0,00000000000000001 Sekunden Ungenauigkeit der Atomuhr / **2,9** mg pro 100g Eisen enthält Spinat / **29** mg pro 100g wurde irrtümlich bis 1982 kommuniziert / **57,3** °C ist die höchste auf der Erde gemessene Temperatur (El Asisija, Libyen) / **-89,2** °C ist die niedrigste gemessene Temperatur (Wostok, Antarktis) / **17** Prozent beträgt die Differenz bei Kraftstoffverbrauchsmessungen zwischen ADAC und Hersteller (2015) / **5.970.000.000.000.000.000.000** Tonnen beträgt die Erdmasse,erstmals gemessen 1798 / **20** Matratzen und Eiderdaunendecken legte

die Königin über die Erbse, um herauszufinden, ob das junge Mädchen tatsächlich eine Prinzessin ist / **46** Übersetzungen gibt es von „Die Vermessung der Welt“, einem der erfolgreichsten deutschen Romane / **463.000** Quadratmeter misst das flächenmäßig größte Messegelände der Welt in Hannover / **10,4** Prozent der Katholiken nehmen durchschnittlich an der Feier der Heiligen Messe teil / **2.000.000.000.000** Suchanfragen misst Google jährlich weltweit / **55** Sekunden dauert eine Suche im Schnitt / **1,7** Yoktogramm exakt kann die derzeit genauste Waage der Welt in Barcelona die Masse von einzelnen Atomen bestimmen / **0,000000000000000000000000000017** Kilogramm entsprechen dem Wert

Wege, dem Smartphone entlockt

Von Dana Gruschwitz und Jana Hölscher

Das Smartphone ist für viele ein Begleiter auf allen Wegen im Alltag. Da liegt die Idee nahe, diesen kleinen, schlauen Alltagsbegleiter mit einer App auszustatten, die das Mobilitätsverhalten seines Besitzers mittels GPS-Tracking erfasst. Mithilfe eines lernenden Algorithmus können aus den übermittelten Daten einzelne Etappen und Wege abgeleitet und die jeweiligen Verkehrsmittel zugeordnet werden. Start- und Ankunftszeit sowie der Start- und Zielpunkt des Wegs ergeben sich aus den übermittelten Informationen selbst. Mobilitätsforscher gelangen so an Daten, die sie sonst aufwendig erfragen müssten. Da der Aufwand bei der Messung via App für die Teilnehmer gering ist, werden auch deutlich längere Beobachtungszeiträume denkbar. Aber wie zuverlässig ist eine solche Messung?

Dieser Frage sind wir im multimo-Projekt in einem Feldversuch mit einem Beobachtungszeitraum von 14 Tagen nachgegangen. Als Vergleichsmethode wurde ein Onlinetagebuch eingesetzt, in das die Teilnehmer ihre täglichen Wege eintrugen. Dabei wurden auch Informationen zum Entscheidungshintergrund abgefragt, die bei der späteren Analyse einen vertieften Einblick in die Verkehrsmittelwahl erlauben. Die Tagebuchgruppe und die App-Gruppe bestanden jeweils aus rund 1.000 Teilnehmerinnen und Teilnehmern.

Wie unterscheiden sich die erhobenen Mobilitätsdaten?

Der Vergleich der beiden Messinstrumente für verschiedene Mobilitätsindikatoren zeigt: Während sich die Verkehrsmittelnutzung (Modal Split) in beiden Gruppen nur geringfügig unterscheidet, wird der Vorteil des passiven Trackings bei den Indikatoren „Entfernung“ und „Unterwegszeit“ deutlich. Die Tagebuchnutzer neigen dazu, beide Werte gerundet zu berichten. In der Summe ergeben sich jedoch ähnliche Ergebnisse.

Deutlichere Unterschiede zeigen sich bei den berichteten Ergebnissen eines Tages: Die App-Nutzer legen pro Tag durchschnittlich einen halben Weg und sechs Kilometer weniger zurück und sind 14 Minuten weniger unterwegs als die Tagebuchnutzer. Da die zurückgelegten Entfernungen und Dauern pro Weg in beiden Gruppen jedoch vergleichbar sind, deutet die Differenz auf eine technische Ursache beim Tracking per App hin. Vermutlich schalten manche Nutzer insbesondere älterer Smartphones das GPS und damit die Wegeaufzeichnung wegen des hohen Energieverbrauchs gelegentlich ab. Mit neuen Smartphone-Generationen wird dieses Thema an Relevanz verlieren. Darüber hinaus scheint es auch Aktivitäten zu geben, bei denen das Smartphone nicht mitgenommen wird.

Im Rahmen des multimo-Projektes stellt der Wechsel von der Kommunikation via E-Mail zum Download und der Aktivierung der App eine Herausforderung dar. Lediglich zwei Drittel der Studienteilnehmer haben die App installiert und für mindestens einen Tag Wege aufgezeichnet. Etwas weniger als die Hälfte übermitteln für mindestens zehn Tage Daten. Eventuell spielen hier technische Probleme oder Vorbehalte bei der Installation eine Rolle. Beim Onlinetagebuch waren deutlich weniger Verluste zu verzeichnen: Hier berichten vier von fünf Studienteilnehmern Wege für mindestens einen Tag, etwas mehr als die Hälfte berichten mindestens zehn Tage. Das Wegetagebuch kommt aber bei längeren Beobachtungszeiträumen sicherlich an seine Grenzen und es sind höhere Abbruchquoten zu erwarten.

Das Tracking per App erfordert hingegen im Zeitverlauf keine weitere Interaktion der Teilnehmer. Mithilfe der GPS-Informationen und eines Mapping-Algorithmus können Routen zudem räumlich sehr exakt verortet werden.

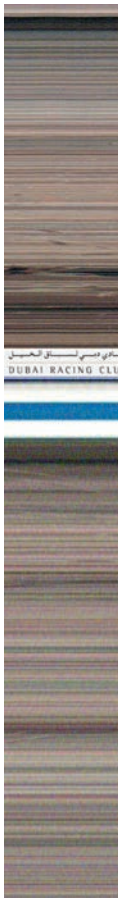
Für die Zukunft ist eine Synthese der App-Messung und des Online-Tagebuchs denkbar. Die App erfasst dann die Bewegung, die durch vertiefende Fragen zu einzelnen Mobilitätsentscheidungen ergänzt wird, um Hintergründe und Motive der Verkehrsteilnehmer zu verstehen und schließlich beeinflussen zu können.

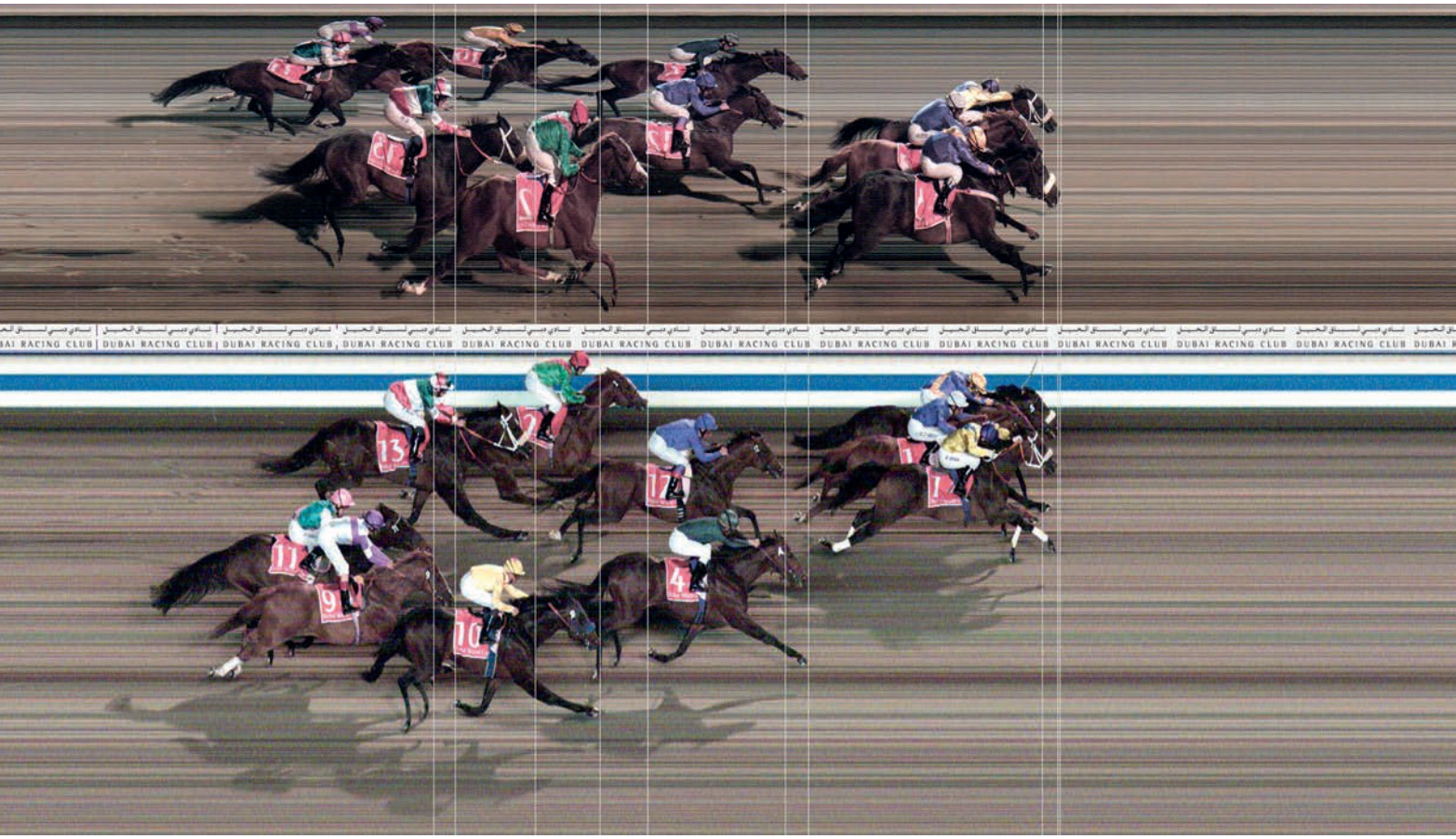
Projekt Multimo

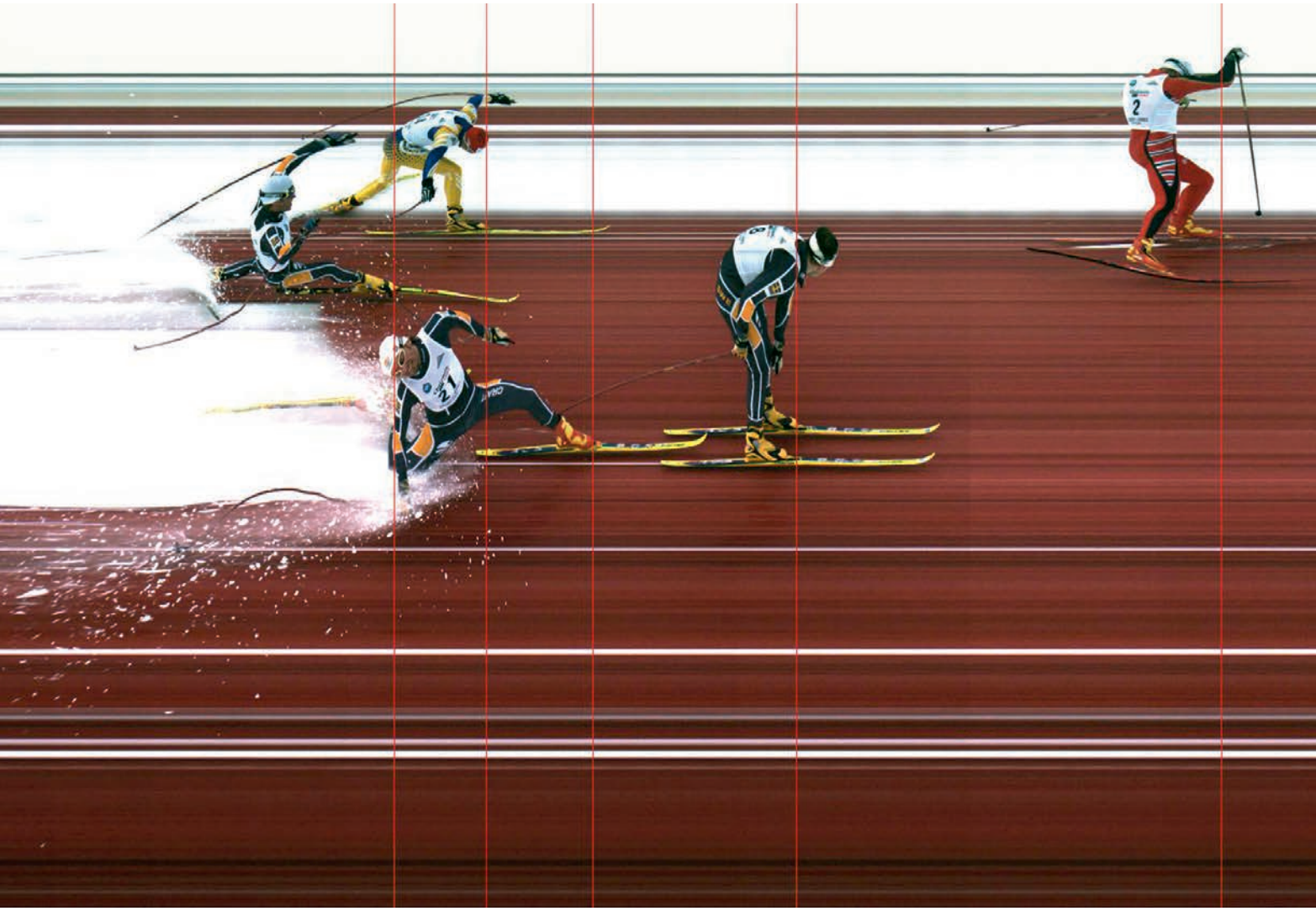
multimo ist ein Gemeinschaftsprojekt von infas, dem Innovationszentrum für Mobilität und gesellschaftlichen Wandel (InnoZ) und GESS, das von verschiedenen Unternehmen aus dem Verkehrsbereich (BVG, DVB, GVH, HVV, KVB, LVB, MVV, RMV, SSB,VBB, VBN, VRR, VRS), dem Verband Deutscher Verkehrsunternehmen (VDV) sowie der Porsche AG gefördert wurde. An dem Projekt haben sich Anfang 2015 mehr als 2.000 Personen in Deutschland beteiligt. Ziel des multimo-Projektes war es, neue Methoden für die Mobilitätsforschung zu entwickeln und zu testen, um neue Mobilitätstrends besser analysieren zu können.

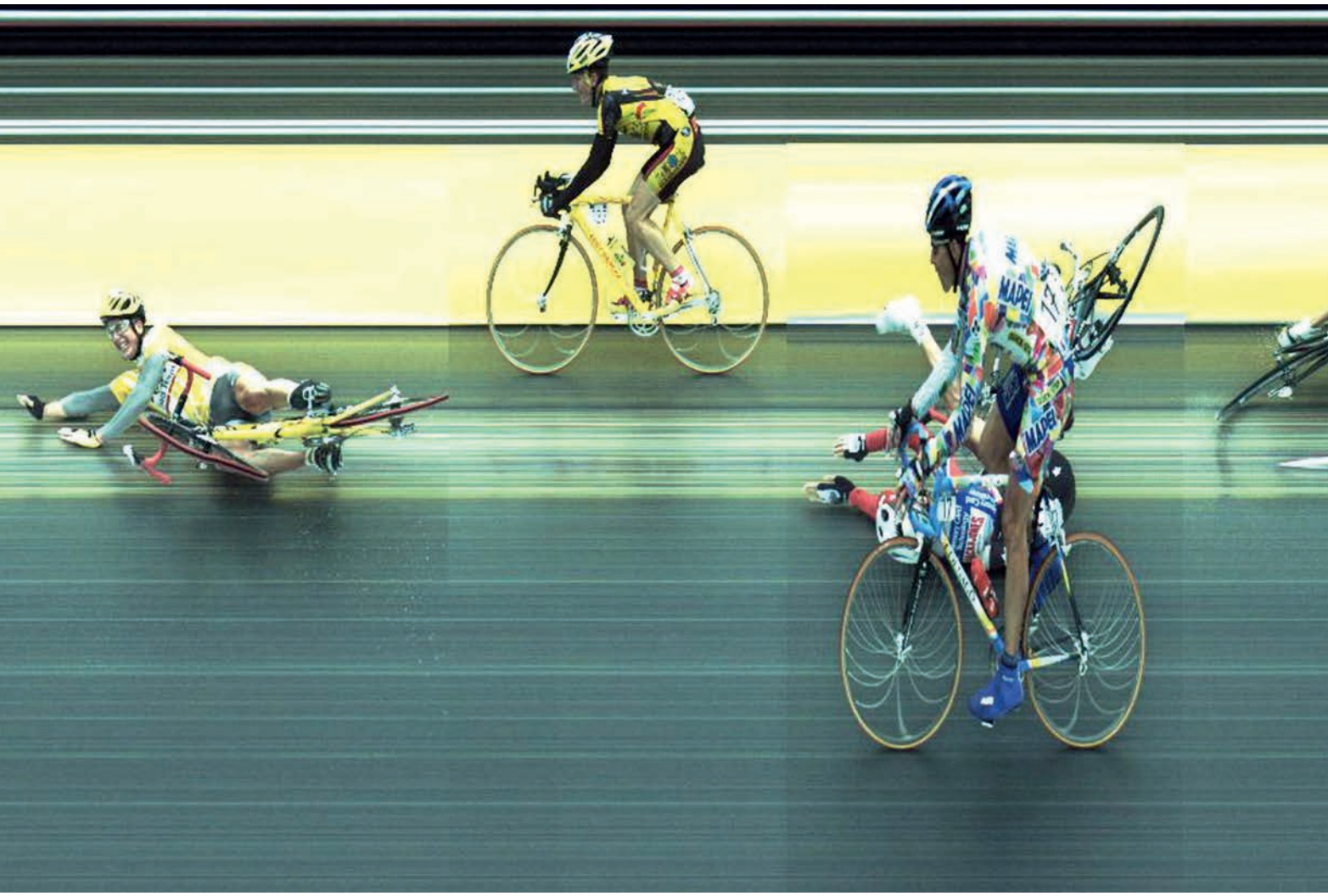
Weitere Informationen auf www.multimo.mobi

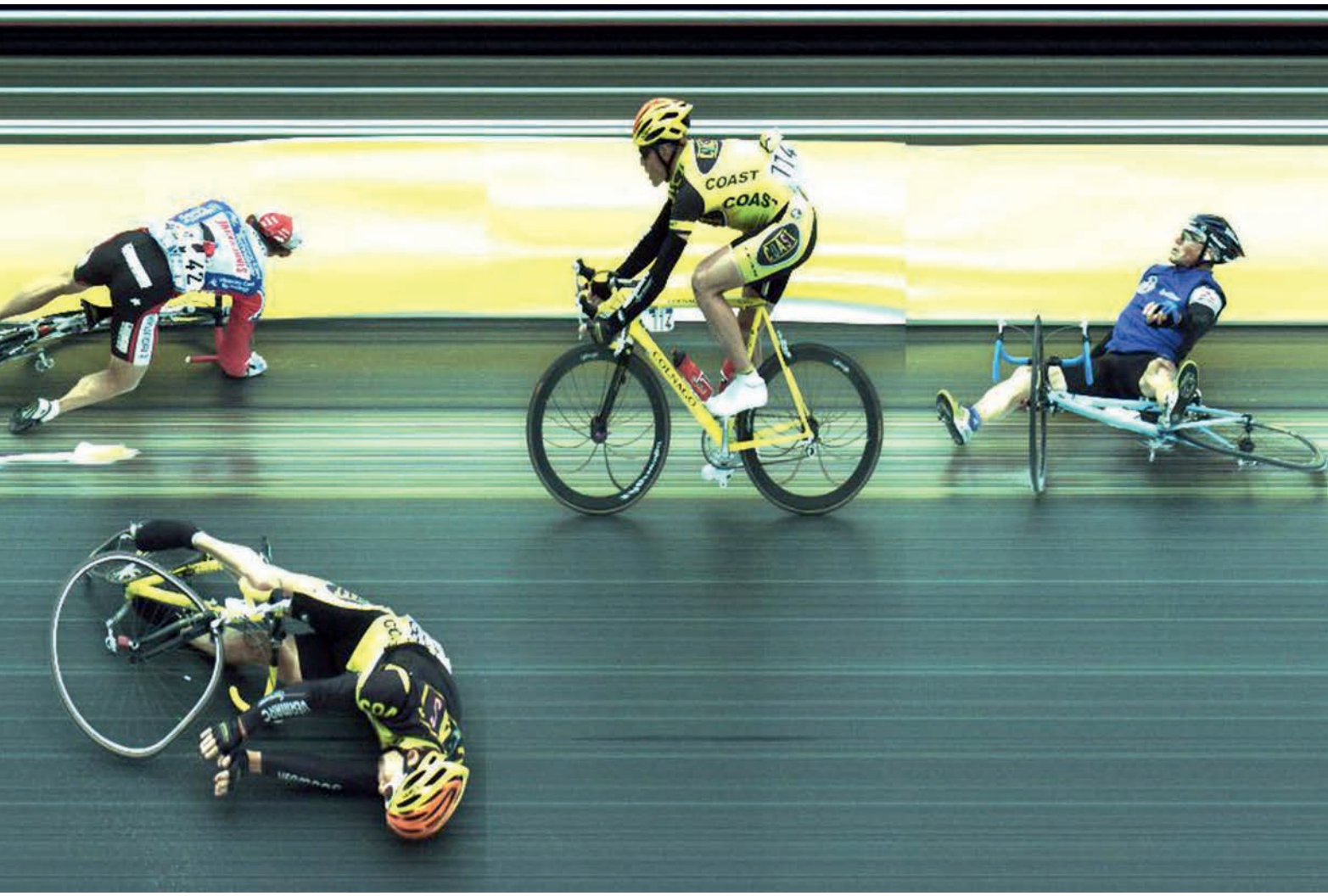
Finish!

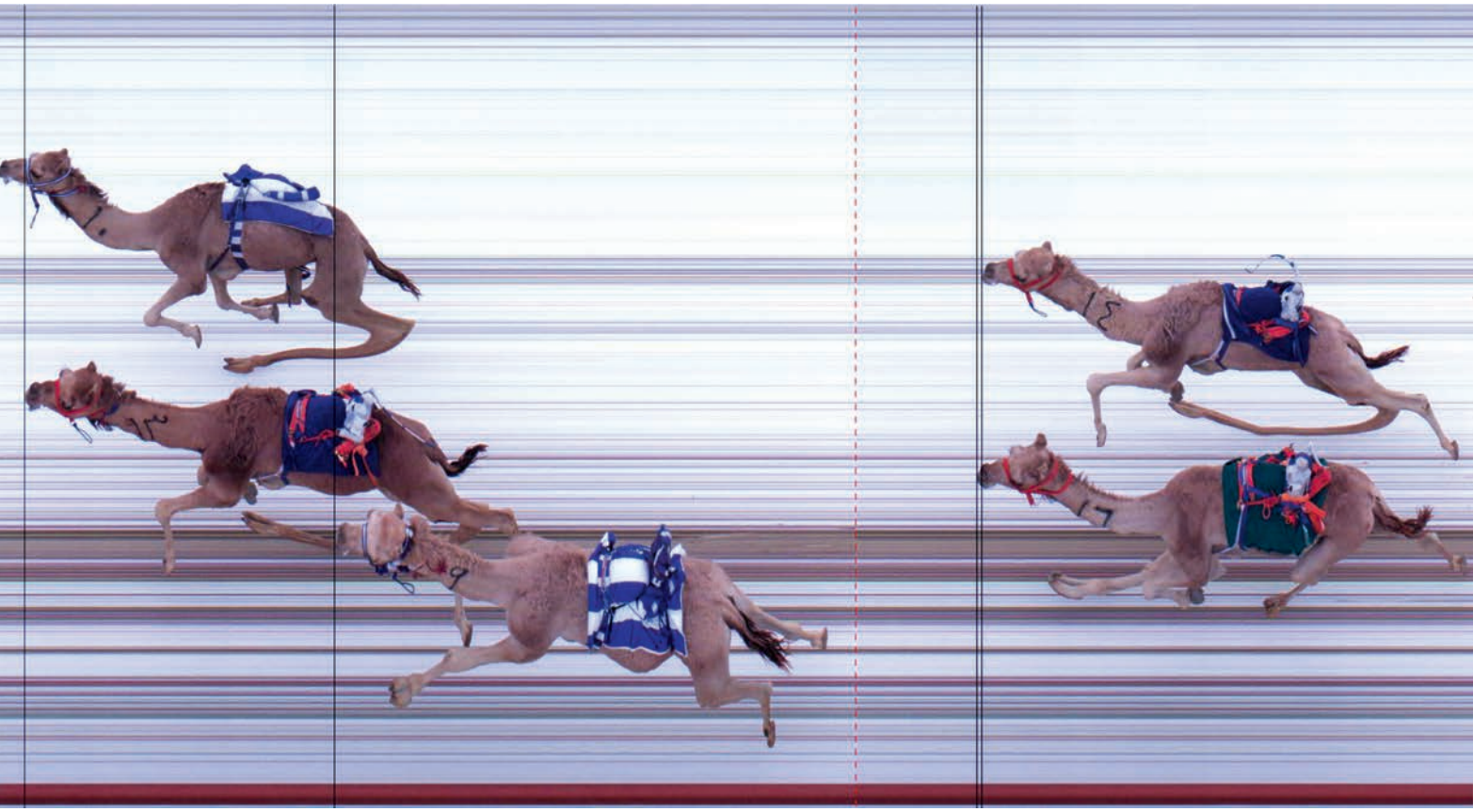


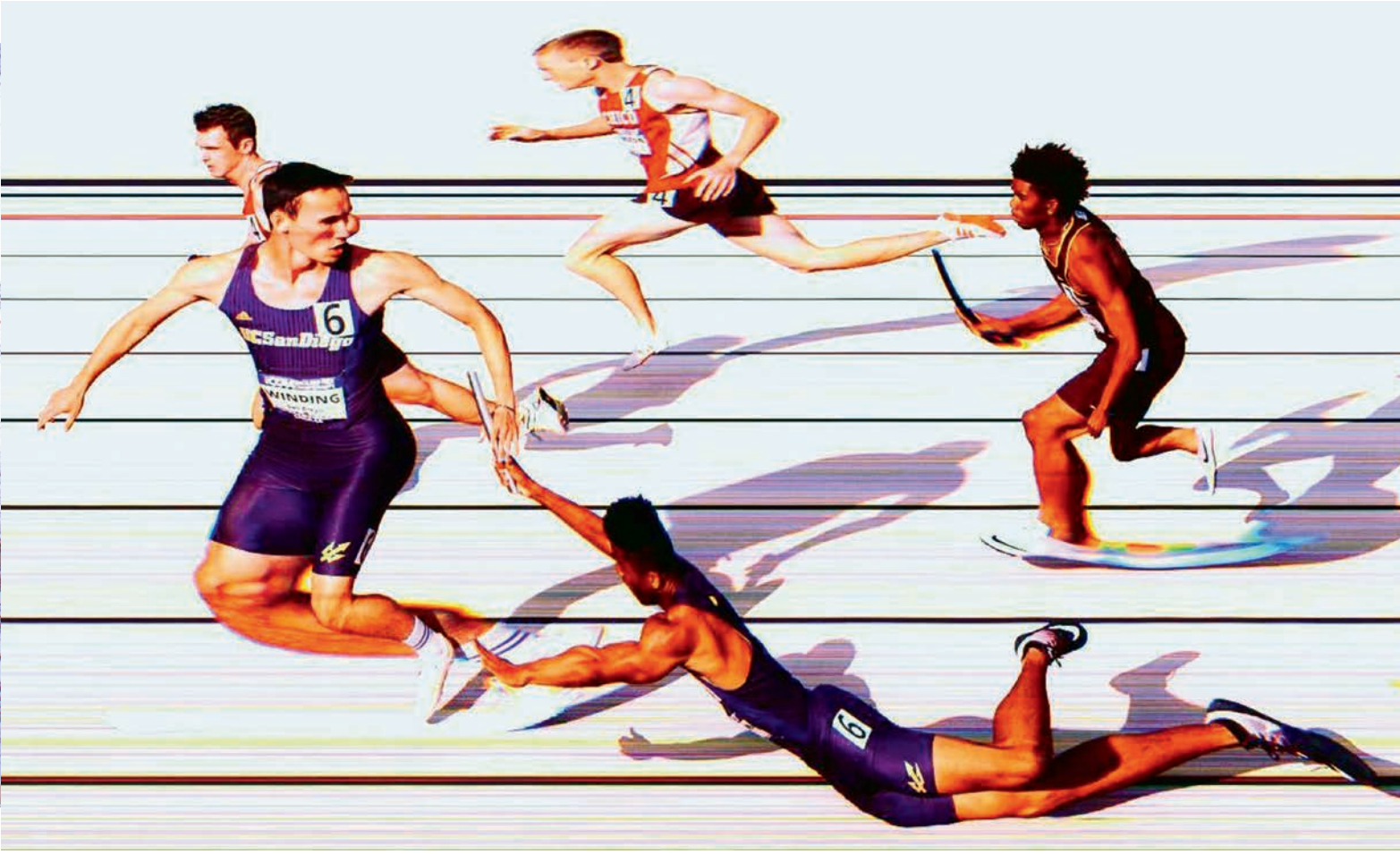

DUBAI RACING CLUB

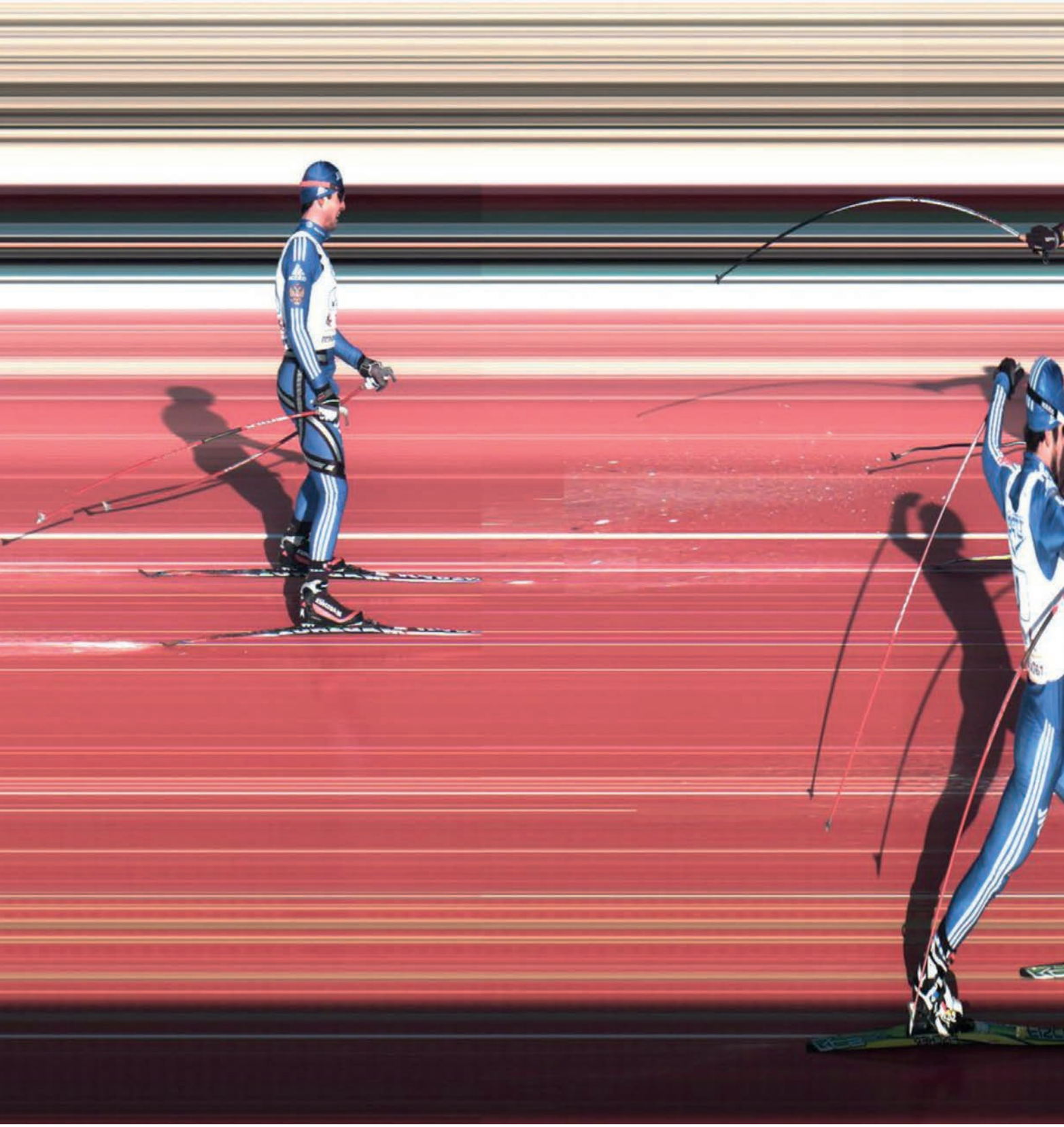


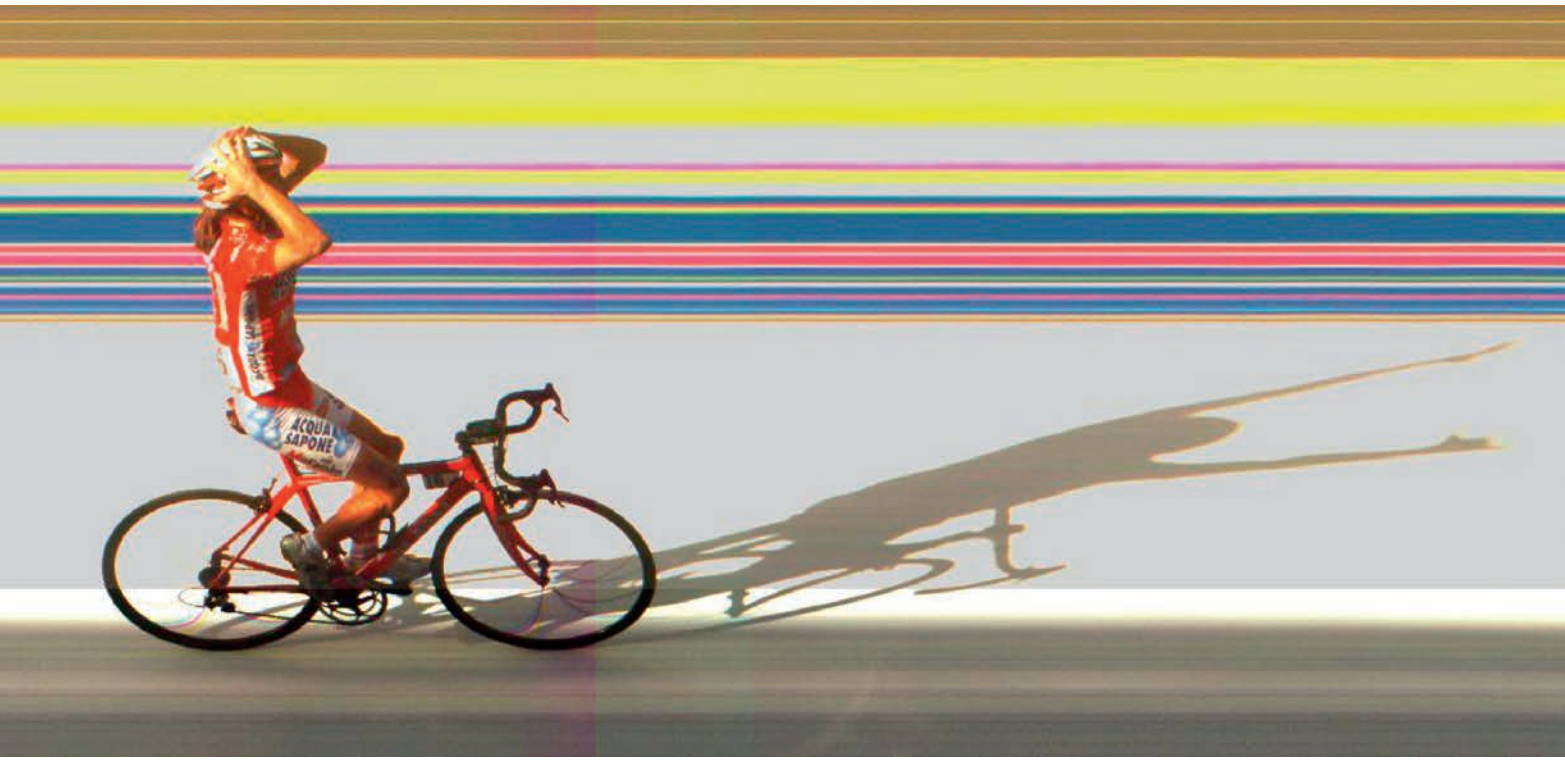

MADEI
MADEI
MADEI


COAST
COAS
774



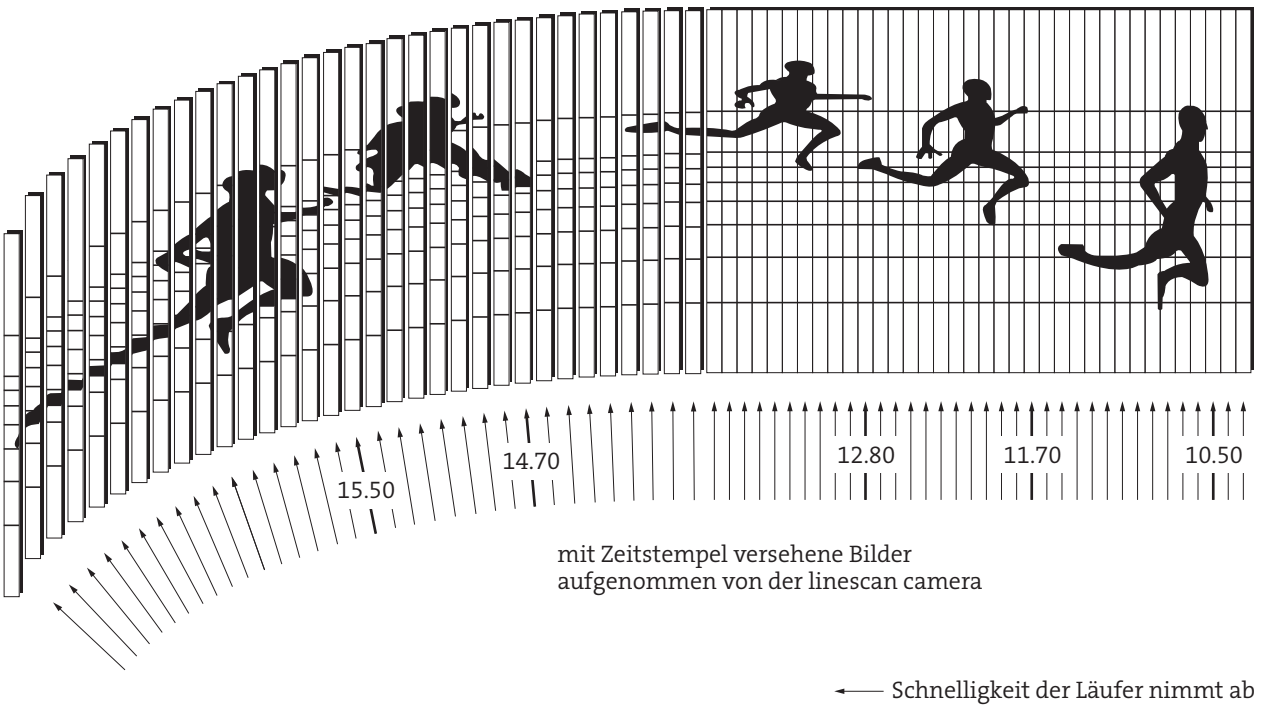

6
WINDING




ACQUA
SAPONE

296
132
85
35

Läufer überqueren die Ziellinie…



Über das Verfahren:
In vielen Sportarten entscheiden heute Tausendstelsekunden über Sieg oder Niederlage. Wie aber werden diese minimalen Unterschiede beim Zieleinlauf gemessen? Es geht um Millimeter. Ein einfaches Foto reicht nicht, um die genauen Zeiten sicher zu ermitteln. Auch eine Videoaufnahme mit ihren 50 oder auch mal 120 Bildern pro Sekunde ist ungeeignet, ein Wimpernschlag-Finish zu dokumentieren.

Die amerikanische Firma Lynx System Developers, die die Bilder für diesen Heftmittelteil zur Verfügung gestellt hat, bietet ein Verfahren an, das die erforderliche Präzision bei der Messung ermöglicht. Das Ergebnis zeigen die Bilder im Mittelteil.

Auf den ersten Blick erscheinen sie wie gewöhnliche Fotografien auf Höhe der Ziellinie. Tatsächlich werden während des Zieleinlaufs aller Sportler Bilder mit einer Breite von einem Pixel genau auf Höhe der Ziellinie fotografiert, und zwar 1.000 Mal pro Sekunde.

Für alle, die die Ziellinie überschritten haben, ist ein vollständiges Abbild – in Streifen – vorhanden. Für jeden vertikalen Bildstreifen ist die Erfassungszeit exakt bekannt und mit einem Zeitstempel dokumentiert. Diese Streifen werden anschließend wie Tapeten nebeneinander gelegt und ergeben die im Mittelteil dargestellten Bilder.

Im Grunde vertauscht das Verfahren von FinishLynx Raum und Zeit. Ereignisse, die alle an derselben Stelle, nämlich der Ziellinie stattfinden, werden räumlich auf einem Bild verteilt. Die zeitliche Abfolge der Ereignisse – die Reihenfolge des Zieleinlaufs – bestimmt die Anordnung auf dem Bild. Die genauen Zeiten der Wettbewerbsteilnehmer können vom Kampfrichter schließlich am Bildschirm eingeblendet werden.

Wir danken der Firma Lynx System Developers Inc. (www.finishlynx.com) für das Bildmaterial.

Irreführungen, Intervalle, Trump und Brexit

Von Reiner Gilberg

Mit dem Brexit und der Trump-Wahl geriet die empirische Umfrageforschung in die Kritik. Schließlich hat keine Prognose im Vorfeld die beiden Ergebnisse sicher vorhergesagt. Kann den Forschern überhaupt getraut werden? Ein detaillierter Blick auf die forschungstheoretischen Grundlagen der empirischen Forschung gibt hier eine klare Antwort.

Aus streng definitorischer Sicht ist Messen in der Statistik bzw. empirischen Sozialforschung definiert als Prozess der Datenerhebung, das heißt der Zuordnung von Zahlen zu Objekten oder Ereignissen nach festgelegten Regeln.

In einer weiter gefassten Bedeutung, bei der die Generierung von Aussagen über gemessene oder messbare Zustände, Ereignisse oder Ähnliches ebenfalls als „Messergebnis" begriffen wird, kommen wir nicht umhin, die in letzter Zeit aufgetretenen und in der Öffentlichkeit als „Falschmessungen" wahrgenommenen Ergebnisse, wie beispielsweise die Prognosen zur letzten US-Wahl oder zur Brexit-Abstimmung in Großbritannien, zur Kenntnis zu nehmen.

Wie aber verhält es sich damit bei genauerer Betrachtung? Lassen sich verschiedene Ereignisse, wie hier der Wahlforschung, mit dem Instrumentarium der Umfrage womöglich gar nicht mehr messen? Sind, um bei den Beispielen zu bleiben, Prognosen bei sehr kleinen Unterschieden (im Falle einer dichotomen Entscheidung für oder gegen den Brexit sehr nahe an 50 Prozent liegende Verteilungen) nicht messbar?

Die Beantwortung dieser Frage führt zu der Beobachtung, dass nicht alles, was bei infas als Standard fest verankert ist, im Bewusstsein der Öffentlichkeit und auch bei Nutzern von Statistiken bekannt bzw. in Vergessenheit geraten ist.

Wesentliche Punkte sind dabei die Transparenz aller Schritte, Informationen über den Datenerhebungsprozess und die Reichweite der gemessenen Befunde, was insbesondere auch die korrekte Angabe über die Intervalle der Schätzungen umfasst.

Am Anfang steht dabei die Stichprobe. Voraussetzung für eine saubere Messung ist eine saubere Stichprobe, das heißt in unserem Verständnis eine Wahrscheinlichkeitsauswahl mit bekannten und berechenbaren Auswahlwahrscheinlichkeiten. Idealerweise deckt die Auswahlgrundlage der Stichprobe die Population dabei vollständig ab. Die Stichprobe, ihr Design und Konzept sind also für die Beurteilung und Einordnung der gemessenen Ergebnisse zentral. Der häufig verwendete pauschale Hinweis auf eine „repräsentative Stichprobe" ist da nicht ausreichend. Zumindest so lange nicht, wie der Begriff keine hinreichende Definition enthält. Unseres Wissens gibt es keine eindeutige Definition des Begriffs „repräsentativ", vermutlich nicht einmal ein allgemeingültiges Verständnis. Als Hilfe zur Beurteilung der Ergebnisse taugt eine solche pauschale Information nicht.

> Unseres Wissens gibt es keine Definition des Begriffs „repräsentativ", vermutlich nicht einmal ein allgemeingültiges Verständnis.

Neben allen Aspekten der Stichprobenziehung, zu denen selbstverständlich auch deren Umfang gehört, sind folgend auch alle weiteren Schritte im Prozess der Datenerhebung relevant. Dazu zählt als ein Punkt auch das Messinstrument. Auch hier können kleine Variationen große Auswirkungen haben.

Ebenfalls relevant ist die Abarbeitung der gezogenen Stichprobe. Hierbei gewinnt in der Umfrageforschung zunehmend das Problem der Erreichbarkeit an Bedeutung. So lässt sich mit einiger Evidenz feststellen, dass nicht (nur) die Teilnahmebereitschaft deutlich abnimmt, sondern mindestens ebenso deutlich sinkt die Erreichbarkeit. Es wird zunehmend aufwendiger, die zu befragenden Personen überhaupt erst zu erreichen. In diesem Zusammenhang bekommt dann auch die Dauer des Datenerhebungsprozesses eine zentrale Bedeutung. Je länger sie dauert und je kontrollierter sie stattfindet, mit dem Ziel möglichst jeden zu erreichen, desto geringer wird eine Verzerrung aufgrund von sozialgruppenspezifisch selektiven Erreichbarkeiten. Beispielsweise sind Einpersonenhaushalte im Mittel schwerer erreichbar als Mehrpersonenhaushalte. Damit sind aber weitere Merkmale verknüpft, die möglicherweise relevant für das zu messende Ereignis sind. Das ist auch einer der Gründe, warum bundesweite Telefonstichproben bei infas seit einiger Zeit standardmäßig nicht nur aus Festnetz-, sondern auch Mobilfunkstichproben bestehen. So wird nicht nur die Erreichbarkeit von Personen ohne einen

Festnetzanschluss überhaupt erst ermöglicht (Coverage-Problem), sondern auch die Erreichbarkeit von hochmobilen Personen und damit auch deren Repräsentanz in der realisierten Stichprobe erhöht.

Am Ende des Datenerhebungsprozesses werden dann die Daten nicht einfach als gemessenes Ergebnis ausgezählt, sondern es erfolgt eine Gewichtung der Daten. Neben der zwingend notwendigen Designgewichtung zum Ausgleich der Auswahlwahrscheinlichkeiten erfolgt üblicherweise eine wie auch immer geartete Anpassung zum Ausgleich von Nonresponse (als Nonresponse-Gewichtung oder als Randanpassung).

Auch hier ist das Vorgehen für die Ergebnisse von erheblicher Bedeutung und mehr als nur eine Randnotiz im gesamten Prozess. In der Regel reduziert die Gewichtung zwar den Bias, vergrößert aber den Stichprobenfehler und ist für die Genauigkeit der Schätzung auf Basis der dann gewichteten Ergebnisse von Bedeutung. In der Regel vergrößert eine Gewichtung das Intervall um das geschätzte Ergebnis. infas versucht diesen Tradeoff anzugehen, indem die Gewichtung nicht als einfaches Standardverfahren implementiert ist, sondern je Untersuchung spezifisch bedacht und bearbeitet wird. Hierbei gilt als grundsätzliches Ziel, so viel wie nötig (zur Reduktion des Bias) und so wenig wie möglich (um den Stichprobenfehler nicht unnötig zu vergrößern).

Der kleine Überblick enthält nicht alle, aber einige zentrale Punkte im Prozess der Datenerhebung zur Messung eines Untersuchungsgegenstandes. Um bei den anfangs zitierten Beispielen zu bleiben, ist es sinnvoll, einen zentralen Aspekt näher zu betrachten, nämlich den Stichprobenumfang. Wie groß müsste dieser sein, um aus den gemessenen Ergebnissen zum Brexit auf den Ausgang der Wahl schließen zu können? Einer Abstimmung also, bei der ein Kopf-an-Kopf-Rennen erwartet wurde. Wie groß müsste die Stichprobe, also die Anzahl der Befragten sein, um das Intervall der gemessenen Schätzung auf Basis einer Befragung im Bereich von +/- 1 Prozentpunkten bzw. noch genauer im Bereich von +/- 0,5 Prozentpunkten zu halten?

Diese Frage führt zu den beiden in der Statistik bekannten Fehlentscheidungen, der Fehler 1. Art und der Fehler 2. Art. Sehr vereinfacht auf das vorliegende Problem bezogen, bedeutet der Fehler 1. Art (Nullhypothese = es gibt weniger Brexit-Befürworter als -Ablehner, die zur Wahl gehen), dass eine Entscheidung für den Brexit erkannt wird, obwohl keine Entscheidung für den Brexit stattfinden wird. Der Fehler 2. Art wäre dann, dass eine Entscheidung für den Brexit nicht erkannt wird, obwohl bei der Wahl für den Brexit gestimmt wird. Beide Fehler hängen mit der Fallzahl zusammen. Vereinfacht ausgedrückt, je höher die Fallzahl, desto geringer die Wahrscheinlichkeit, einen Fehler zu begehen. Die statistische Power bezeichnet die Wahrscheinlichkeit, einen Fehler 2. Art zu vermeiden, und lässt sich für die Berechnung einer notwendigen Fallzahl bei hypothesentestenden Fragestellungen verwenden, die groß genug ist, um mit einem bestimmten Fehler auch sehr kleine Unterschiede erkennen zu können. Steht nicht der Nachweis einer Hypothese im Mittelpunkt, sondern die Schätzung eines Populationsparameters, dann kann eine Strategie der Fallzahlplanung zur Abschätzung der erwarteten Breite von Konfidenzintervallen verfolgt werden.

Allerdings scheint diese Berechnung im Vorfeld ein wenig aus der Mode gekommen zu sein. Bisweilen gibt es die Angabe eines Konfidenzintervalls, also eines Bereiches, in dem die Schätzung mit einer bestimmten Wahrscheinlichkeit liegt. Dieses ist abhängig vom Schätzergebnis, im Falle eines Schätzwertes von 50 Prozent ist es am größten. Zudem ist auch das Stichprobendesign für die Intervallbreite relevant.

> Wie groß müsste der Stichprobenumfang sein, um aus den gemessenen Ergebnissen zum Brexit auf den Ausgang der Wahl schließen zu können?

Gehen wir bei einem gemessenen Wert von 50 Prozent für den Brexit und damit der Vorhersage eines exakten Gleichstands von einer einfachen uneingeschränkten Zufallsauswahl (simple random sampling) aus. Unterstellen wir zudem eine intensiv bearbeitete Stichprobe mit einer ausreichenden Feldzeit und vernachlässigen die Effekte einer aufgrund von selektivem Nonresponse möglicherweise notwendigen Gewichtung. Dann ergeben sich bei einem gewählten Konfidenzniveau von 95 Prozent, entsprechend einem Signifikanzniveau von 5 Prozent (vereinfacht, die Wahrscheinlichkeit, eine Entscheidung gegen den Brexit fälschlicherweise zu verwerfen) folgende Vertrauensintervalle:

+/- 1 Prozentpunkte Intervall: rund 10.000 Fälle.

+/- 0,5 Prozentpunkte Intervall: rund 40.000 Fälle.

Bei einem Signifikanzniveau von 1 Prozent, d.h. bei einer deutlich geringeren Wahrscheinlichkeit, einen Fehler zu begehen, ergibt sich eine notwendige Fallzahl von:

+/- 1 Prozentpunkte Intervall: 17.000 Fälle.

+/- 0,5 Prozentpunkte Intervall: rund 65.000 Fälle.

Diese Vertrauens- oder Konfidenzintervalle geben den Bereich an, der bei unendlicher Wiederholung mit einer bestimmten Wahrscheinlichkeit, dem Konfidenzniveau, den wahren Wert einschließt. Das Konfidenzniveau gibt an, wie viel Prozent aller auf Grundlage von gemessenen Daten berechneten Vertrauensintervalle (bei der Berechnung ein und desselben Vertrauensintervalls) den wahren Wert in der Population umfassen.

Am Rande bemerkt sei, dass aufgrund des US-amerikanischen Wahlsystems eine solche Fallzahl in allen 52 Bundesstaaten der USA notwendig wäre, um für jeden Staat die Entscheidung für oder gegen Trump/Clinton bei einer Kopf-an-Kopf-Entscheidung in jedem Bundesstaat mit einer solchen Genauigkeit bzw. in einem derart ge-

ringen Intervall zu messen.

Das reale Ergebnis der Abstimmung über den Brexit war dann auch mit 51,9 Prozent für den Brexit und 48,1 Prozent gegen den Brexit in der Tat sehr eng. Die korrekte Vorhersage entsprechend schwierig. Die üblichen Fallzahlen in der Wahlforschung liegen bei n = 1.000, n = 1.500 und bisweilen auch bei n = 2.000 Befragten. Mit diesen Fallzahlen lässt sich aber das Ergebnis einer derart knappen Entscheidung nicht abschätzen, wie ein Blick auf die „naiven" Vertrauensintervalle zeigt:

Vertrauensintervalle:

n = 1.500, 95 % Konfidenzniveau: +/- 2,5 %
n = 1.500, 99 % Konfidenzniveau: +/- 3,2 %
n = 2.000, 95 % Konfidenzniveau: +/- 2,2 %
n = 2.000, 99 % Konfidenzniveau: +/- 2,9 %

Demnach liegt selbst bei einer je nach Konfidenzniveau und Fallzahl durchaus deutlich über 50 Prozent liegenden Messung gegen den Brexit das reale Ergebnis von nur 48,1 Prozent bei diesen Fallzahlen (selbst bei einer erlaubten Fehlertoleranz von 5 Prozent) noch innerhalb des Vertrauensintervalls.

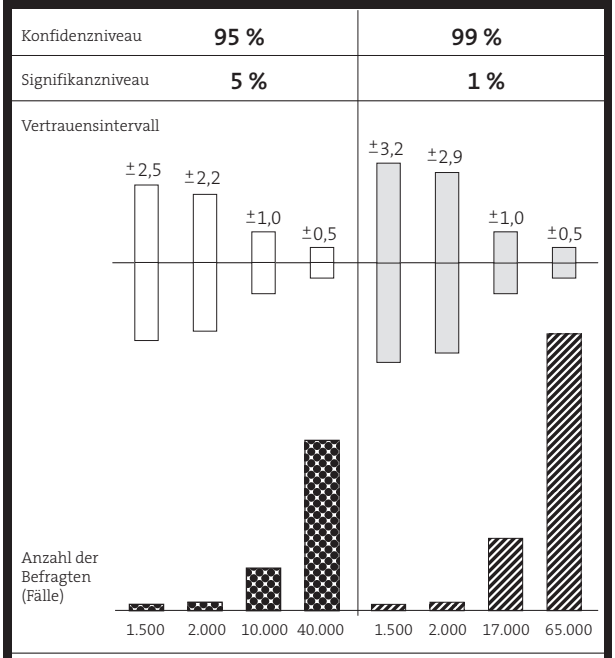
Angegeben ist in den obigen Beispielen zudem stets das „naive" Vertrauensintervall, also das Konfidenzintervall unter den genannten Bedingungen. In der Umfrage- bzw. Wahlforschung kann allerdings eher selten auf eine uneingeschränkte Zufallsauswahl zurückgegriffen werden und zudem erfolgt in der Wahlforschung (und den meisten Umfrageforschungen) die Gewichtung der Daten standardmäßig. Beides vergrößert den Stichprobenfehler und erhöht damit das Vertrauensintervall unter Umständen erheblich. Auch dies kann und sollte berücksichtigt werden. Die Angabe eines Vertrauensintervalls ohne Berücksichtigung des tatsächlichen Stichprobendesigns inkl. Gewichtung ist andernfalls ebenfalls irreführend.

Betrachten wir die Fallzahlen, wird klar, dass viele der in der Öffentlichkeit als „Falschmessungen" wahrgenommenen Ergebnisse streng genommen nicht wirklich falsch waren. Sie waren lediglich zu ungenau, was aber nicht immer erkennbar wurde. Knappe dichotome Entscheidungen sind also durchaus schwierig zu messen und stellen die Forschung (hier die Wahlforschung) vor Probleme, die nur mit einigem Aufwand bearbeitet werden können. Dennoch können sie bearbeitet werden.

Um die anfangs gestellte Frage zu beantworten: Aus unserer Sicht lassen sich mit dem Instrumentarium der Umfrageforschung auch Ereignisse mit sehr kleinen vermuteten Unterschieden messen. Allerdings ist das mit erheblichem Aufwand, mit erheblichen Kosten und auch mit Unsicherheiten verbunden.

Zusammengefasst lässt sich sagen, dass es lohnenswert ist, den gesamten Prozess der Datenerhebung sorgfältig zu planen. Im Sinne der Gültigkeit und Zuverlässigkeit der Ergebnisse lohnt es sich, darüber nachzudenken und die Probleme zu benennen. Sicher können nicht immer alle Probleme gelöst werden, aber deren Auswirkungen auf das gemessene Ergebnis lassen sich häufig benennen und sollten dann auch benannt werden. Am Ende bleibt nicht nur die Frage, was habe ich gemessen, sondern auch die Frage, bei wem habe ich gemessen und wie genau habe ich gemessen bzw. erhoben. Eine transparente Darstellung ist dabei sicher ein Schritt, das Vertrauen in die gemessenen Ergebnisse der empirischen Sozialforschung zu stärken.

Konfidenzintervalle, Vertrauensintervalle und Fallzahlen: Bei knappen Entscheidungen sind hohe Fallzahlen erforderlich



Lesebeispiel: Für eine sichere Prognose eines sehr knappen Ergebnisses (z.B. Brexit) wird ein genaues Studienergebnis benötigt. Wenn beispielsweise zu 99 Prozent sichergestellt sein soll, dass der tatsächliche Anteil in der Bevölkerung in einem schmalen Intervall von ± 0,5 Prozent um den in der Studie gemessenen Wert liegt (die Irrtumswahrscheinlichkeit also nur bei 1 Prozent liegen soll), müssten 65.000 Personen befragt werden.

Quelle: infas, März 2018

infas

Zum Weiterlesen:
Schnell, Rainer, Noack, M., 2014, The Accuracy of Pre-Election Polling of German General Elections. In: MDA – Methods, Data, Analysis 8 (1) 5-24, http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/mda/Vol.8_Heft_1/MDA_Vol8_2014-1_Schnell_Noack.pdf
Schnell, Rainer, Hill, P. B., Esser, E., 2013, Methoden der empirischen Sozialforschung (10. überarb. Aufl.), München, Wien: R. Oldenbourg Verlag.
Weichbold, Martin, Bacher, J., Wolf, C. (Hrsg.), 2009, Umfrageforschung, Herausforderung und Grenzen, Österreichische Zeitschrift für Soziologie, Sonderheft 9, VS Verlag für Sozialwissenschaften
Cochran, William G., 1977, Sampling Techniques (3rd Edition), New York: John Wiley & Sons.
Leslie Kish, 1965, Survey Sampling, New York: John Wiley & Sons.

Grenzenlos genau – Herausforderungen kulturübergreifender Messungen

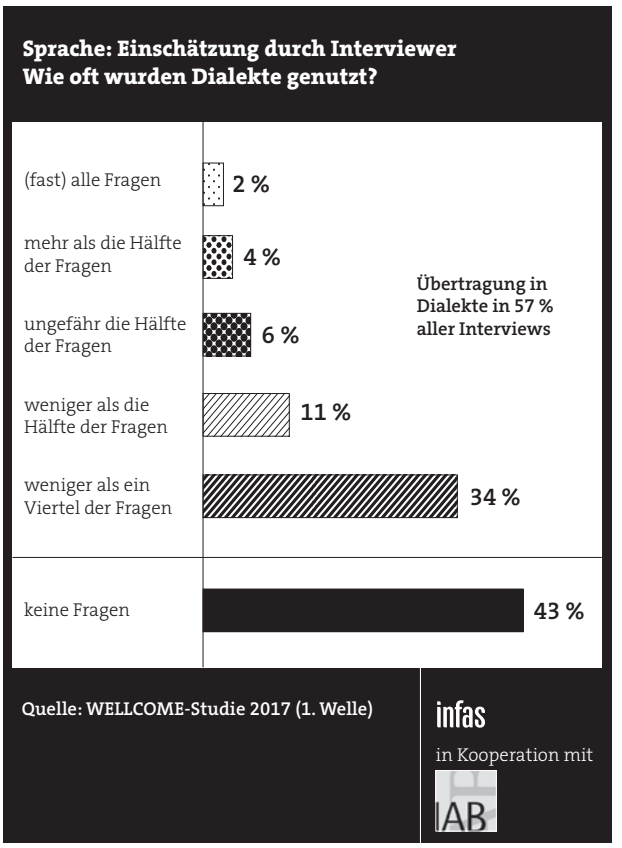
Von Angelika Steinwede, Michael Ruland und Hans Dietrich

Eine globale Welt erfordert sozialwissenschaftliche Messungen, die auch kulturübergreifend eine valide Operationalisierung und Messung von Inhalten erlauben. Daraus ergibt sich die Herausforderung, dass das Studiendesign für internationale Befragungen oder Studien mit Befragten aus unterschiedlichen Kulturen so konzipiert werden, dass Inhalte kulturübergreifend vergleichbar gemessen werden können. Anhand der Studie WELLCOME, die infas gemeinsam mit dem Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg, durchführt, werden im Folgenden die Strategien dargestellt, die im Rahmen dieser Studie ergriffen wurden, um eine valide kulturübergreifende Messung zu erreichen.

Hintergrund der WELLCOME-Studie

Im Jahr 2015 wurde bei infas die Studie „Jugendarbeitslosigkeit, psychische Gesundheit und Arbeitsmarktergebnisse“ für das IAB begonnen. Inhaltlich befasst sich diese Studie mit dem Zusammenhang zwischen psychischen Belastungen und Arbeitslosigkeit bei jungen Erwachsenen. Bei Sichtung der ersten Stichprobenpopulation wurde deutlich, dass der verstärkte Eintritt junger (Flucht-)Migranten ab Herbst 2015 die angestrebte Analysepopulation (junge Erwerbslose, die erstmalig im Arbeitslosenregister der BA erfasst werden) erheblich verzerrt. Das IAB entschied daher, eine Zusatzstichprobe junger syrischer (Flucht-)Migranten getrennt zu befragen. Die WELLCOME-Studie wurde mit dem Ziel konzipiert, den Einfluss psychischer Belastung und posttraumatischer Stresserfahrung auf den Integrationsverlauf in Bildung und Beschäftigung zu analysieren. Dabei wurde wesentlich auf das Erhebungsinstrument der Studie „Jugendarbeitslosigkeit, psychische Gesundheit und Arbeitsmarktergebnisse“ zurückgegriffen. Jedoch waren zielgruppenspezifische Ergänzungen erforderlich, um ein kulturübergreifendes Studiendesign zu erreichen. Insbesondere – das zeigte der Pretest – sind spezifische Anforderungen an die Qualifikation der Interviewer und die Gesprächsführung mit den jungen Flüchtlingen zu beachten.

Die WELLCOME-Studie wurde in einem Multi-Method-Design von telefonischer und Online-Befragung angelegt. Dies war notwendig, da nur für einen Teil der Stichprobe Telefonnummern vorlagen. Um Non-Response-Verzerrungen zu reduzieren, wurden Personen ohne verfügbare Telefonnummern gebeten, online teilzunehmen. Zu Beginn erhielten alle Zielpersonen ein Anschreiben mit zweisprachigen Studieninformationen (Arabisch, Deutsch). In der telefonischen Befragung wurden nur Interviewer mit arabischer Muttersprache ein-



gesetzt. Diese Entscheidung wurde zum einen aufgrund der vorhandenen Sprachkompetenz getroffen, außerdem auch aufgrund der Tatsache, dass sich gerade muttersprachliche Interviewer gut in den kulturellen Kontext dieser Befragten hineinversetzen können. Es bestand für alle Befragten die Möglichkeit, das Interview entweder auf Deutsch oder auf Arabisch zu geben.

Ein zentrales Ergebnis des vorgelagerten Pretests der WELLCOME-Studie lag darin, dass viele Befragte Verständnisprobleme mit dem Hocharabischen hatten und den Fragebogen deutlich besser verstanden, wenn er im Dialekt präsentiert wurde. Dies hat den Hintergrund, dass das Hocharabische in der Regel als Schriftsprache und nicht als gesprochene Alltagssprache verwendet wird. Um das Verständnis bei den Befragten sicherzustellen, wurde entschieden, dass die Interviewer die Fragen in einen alltagssprachlichen Dialekt übertragen durften. Dabei waren zunächst alle Fragen grundsätzlich in Hocharabisch vorzulesen und bei Bedarf noch einmal erneut im exakten Wortlaut zu wiederholen. Gab es dann noch Verständnisprobleme, so durften die muttersprachlichen arabischen Interviewer die Frage oder die Antworten anschließend in den entsprechenden Dialekt übertragen. Diese Abfolge war streng vorgegeben, um einerseits möglichst standardisiert vorzugehen, andererseits aber ein maximales Verständnis der Fragebogeninhalte sicherzustellen. Diese Übertragung in den Dialekt wurde bei einer deutlichen Mehrheit (knapp 60 Prozent der Befragten) in unterschiedlicher Ausprägung genutzt (siehe Abbildung). Statistische Überprüfungen der Skalenqualität (Reliabilität der Antworten zu Itembatterien) ergaben, dass die Nutzung der Dialekte keinen nachteiligen Einfluss auf das Antwortverhalten und die Datenqualität hat.

Frage zur Arbeitsmotivation
Sozial erwünschtes Antwortverhalten

	Gesamt	Männer	Frauen
Geschlecht			
Weiblich (ref.)	--		
Männlich	0,23***		
Alter (linear)	0,01**	0,02***	-0,01
Skala zu Geschlechterrollen	0,04***	0,03***	0,17***
Bildung			
Abschluss der Primarstufe	0,17***	0,14**	0,45*
Abschluss sekundäre/ intermediate Stufe	0,07	0,06	0,24
Abitur/Hochschulreife	0,02	-0,01	0,32
Abschluss Postsecundary Institute/ berufliche Ausbildung	0,09	0,08	0,26
Universitätsabschluss (ref.)	--	--	--
Bildung der Eltern			
kein Universitätsabschluss (ref.)	--	--	
Universitätsabschluss	0,04*	0,01	0,19**
Geschlecht Interviewer/in			
Weiblich (ref.)	--	--	--
Männlich	0,10***	0,01***	0,11(*)
Konstante	3,96***	4,20***	3,80***
Anzahl an Fällen	2.275	1.935	340
R^2	0,055	0,032	0,105

Befragte mit syrischer Staatsangehörigkeit
Signifikanzniveau: * p<.05; ** p<.01; *** p<.001
Lineare Regression mit der Arbeitsmotivation als abhängige Variable

Quelle: WELLCOME-Studie 2017 (1. Welle)

infas in Kooperation mit IAB

Sozial erwünschte Antworten vermeiden

Die beschriebenen Maßnahmen zur besseren kulturübergreifenden Verständlichkeit des Fragebogens lösen jedoch nicht ein zentrales Phänomen im Rahmen von Messungen. Das Problem der sozialen Erwünschtheit. Die Analysen der Befragungsdaten zeigten einen signifikanten Effekt der Flüchtlinge beim Antwortverhalten: Telefonisch Befragte gaben im Vergleich zu online Befragten eine signifikant höhere Arbeitsorientierung an, die sich inhaltlich schwer erklären lässt. Darüber hinaus zeigte sich ein Effekt des Interviewergeschlechts auf das Antwortverhalten.

Die Annahme war, dass Befragte die vermuteten Erwartungen der deutschen Bevölkerung an Flüchtlinge erfüllen wollen. Dies setzt natürlich voraus, dass den Befragten die gesellschaftlichen Erwartungen klar sind – hinsichtlich der Arbeitsorientierung scheint es für Flüchtlinge recht klar, da sie Deutschland als ein Land wahrnehmen, in dem viel gearbeitet wird. Und die Arbeitswelt wird für diese Befragten vor allem durch Männer, in diesem Fall durch männliche Interviewer repräsentiert. Dieser Effekt der sozialen Erwünschtheit konnte sowohl bei männlichen als auch bei weiblichen Flüchtlingen nachgewiesen werden. Beide Geschlechter nannten bei männlichen Interviewern eine deutlich höhere Arbeitsmotivation als bei weiblichen Interviewern.

Kulturelle Unterschiede berücksichtigen

Die WELLCOME-Studie ist ein Beispiel dafür, welche besonderen Herausforderungen sozialwissenschaftliches Messen im internationalen Kontext an Forscher stellt. Das sehen wir auch bei anderen infas-Studien, wie zum Beispiel bei der Untersuchung „Refugees in the German Educational System“ ReGES im Auftrag des Leibniz-Instituts für Bildungsverläufe e.V., Bamberg, oder beim „Religionsmonitor“ der Bertelsmann Stiftung, Gütersloh. Auch bei diesen Erhebungen sind zahlreiche methodische Anpassungen erforderlich, um kulturübergreifend adäquat zu messen und so valide Ergebnisse zu erzielen.

Zum Weiterlesen
Behr, Dorothée, Braun, Michael, Dorer, Brita (2015): Messinstrumente in internationalen Studien, Mannheim, GESIS – Leibniz-Institut für Sozialwissenschaften (GESIS Survey Guidelines). DOI:10.15465/gesis-sg_006

Smart messen – Apps in der Sozialforschung

Von Jana Hölscher und Thomas Weiß

Der Erfolg von Smartphones und Tablets ist unweigerlich mit dem Erfolg von mobiler Anwendungssoftware, kurz Apps, verbunden. Gerade einmal sieben Minuten telefonieren Smartphone-Nutzer durchschnittlich mit ihrem Gerät, während sie im Schnitt rund zwei Stunden mit der Verwendung von Apps verbringen. Mag dies auf den ersten Blick ziemlich viel erscheinen, so ist es doch angesichts der zahlreichen Einsatzmöglichkeiten von Programmen gar nicht mehr so erstaunlich. Inzwischen scheint es für nahezu jede Anforderung die passende App zu geben: Sie möchten wissen, mit welchem Verkehrsmittel man am schnellsten von A nach B kommt, von unterwegs eine Banküberweisung tätigen oder etwas spielen? Bei rund 5,5 Millionen Apps, die die Marktführer Google und Apple in ihren App-Stores zusammen anbieten, haben Smartphone-Nutzer die Qual der Wahl, die passende für die Lösung ihres „Alltagsproblems" zu finden.

Seit dem Start der App-Stores von Apple und Google 2008 wächst der Markt rasant. 2017 stieg die Anzahl der weltweiten App-Downloads auf 175 Milliarden. Dies entspricht einem Zuwachs von 60 Prozent im Vergleich zu 2015. Die Apps sind jedoch nicht allein für ihre Nutzer attraktiv. Auch immer mehr Unternehmen nutzen sie, um Kunden zu gewinnen und Umsätze zu erzielen. In verschiedenen Arbeitsumfeldern tragen Apps außerdem einen entscheidenden Anteil zur Digitalisierung bei.

Vorreiter Medizin

Auch in der wissenschaftlichen Forschung finden Apps zunehmend Anwendung. Vor allem in der Medizin werden sie schon seit mehreren Jahren als Erhebungsinstrument eingesetzt. So ermöglichen in Smartphones verbaute Messinstrumente, wie beispielsweise Barometer, Beschleunigungs- oder Bewegungssensoren, das Erfassen von unabhängigen oder objektiven Daten direkt mithilfe der Hardware der Probanden. Ebenso hat Apple mit der Lancierung des sogenannten ResearchKit Bewegung in diesen Bereich gebracht. Die von Apple entwickelte Software ermöglicht es Forschungseinrichtungen, einfach und schnell Apps für ihre Studien zu entwickeln und zu verbreiten. Durch die verfügbaren Apps kann dann in wenigen Wochen eine Vielzahl an Studienteilnehmern rekrutiert werden, wodurch die Forschung gegenüber konventionellen Methoden der Teilnehmerrekrutierung deutlich beschleunigt wird.

APP Tracy

Bei einem derzeit noch laufenden Projekt von infas zeigt sich, dass eine App sowohl Messinstrument als auch selbst Gegenstand der Forschung sein kann. Die im Projekt eingesetzte App „Tracy" ging den digitalen Spuren der Projektteilnehmer an einem Ort nach, an dem die meisten Menschen inzwischen eine Vielzahl solcher Spuren hinterlassen: auf dem eigenen Smartphone. Die Teilnehmer installierten die App für einen Erhebungszeitraum von drei Monaten. In dieser Zeit sammelte Tracy verschiedene Informationen über sie. Dabei konnte die Art und Menge der gesammelten Daten je nach Betriebssystem, Smartphone und Nutzer sehr unterschiedlich ausfallen. Zu den erhobenen Daten gehörten verschiedene IDs des Smartphones (wie z. B. der IMEI oder Ad-ID), GPS-Koordinaten sowie die Anzahl und Nutzungshäufigkeit anderer Apps auf dem Smartphone.

In der empirischen Sozialforschung steht der Einsatz von Apps trotz ihrer vielfältigen Anwendungsmöglichkeiten noch relativ am Anfang. Als Vorreiter in diesem Bereich werden Apps seit ein paar Jahren in der Mobilitätsforschung als GPS-Tracker eingesetzt. Dabei greifen sie auf die GPS-Daten des Smartphones zurück, um beispielsweise Wege und genutzte Verkehrsmittel der Smartphone-Besitzer zu erfassen (siehe Artikel Seite 16). Darüber hinaus eröffnen Apps eine Vielzahl weiterer Anwendungsmöglichkeiten im Rahmen der empirischen Sozialforschung. Die Potenziale liegen dabei nicht nur in den erweiterten technischen Möglichkeiten, wie der Erfassung von objektiven Daten (z. B. GPS-Daten oder anderes Smartphone-basiertes Nutzungsverhalten), sondern auch in der Möglichkeit, zusätzliche (Para-)Daten (z. B. Gerätekennzeichnung oder Betriebssystem) erheben zu können, die bislang für den Forschungsprozess nicht oder nur schwer nutzbar gemacht werden konnten. So ist es beispielsweise möglich, mithilfe einer App die Nutzung der auf dem jeweiligen Smartphone installierten Apps zu erfassen.

Ein Vorteil mobiler Anwendungssoftware in der Empirie ist der schnelle und von einem festen Standort bzw. Adresse losgelöste Zugang zu den Zielpersonen, da Smartphones mittlerweile zum Alltagsbegleiter geworden sind.. Auch der Anteil an Smartphone- oder Tabletnutzern steigt kontinuierlich an, sodass gleichzeitig die potenzielle Zielgruppe für den Einsatz von

Apps wächst. Gerade bei jungen oder technikaffinen Personen kann der Einsatz einer App von Vorteil sein. In der Altersgruppe der 14- bis 29-Jährigen nutzten 2017 in Deutschland bereits 95 Prozent ein Smartphone. In der Altersgruppe zwischen 30 und 49 Jahren lag die Zahl mit 97 Prozent noch höher. Bei den 50- bis 64-Jährigen waren es 88 Prozent und immerhin jeder dritte Deutsche von mindestens 65 Jahren zählte 2017 bereits zu den Smartphone-Nutzern. Im Kontext sozialwissenschaftlicher Panelerhebungen kann eine App im Rahmen des Tracking und für die Panelpflege eingesetzt werden. Unter Tracking versteht man in diesem Zusammenhang das Ermitteln von Befragten, die ansonsten für eine Befragung im Rahmen einer Panelstudie verloren wären bzw. nicht interviewt werden könnten. In der Literatur finden sich vermehrt Verweise auf die Möglichkeit, im Rahmen von Panelstudien auch transferierbare Kontaktinformationen zu erheben, die bei Adressänderungen erhalten bleiben, wie beispielsweise die E-Mail-Adresse und Kontaktinformationen aus sozialen Netzwerken. Allerdings sprechen zahlreiche Argumente gegen einen Einsatz von sozialen Netzwerken, wie beispielsweise datenschutzrechtliche Aspekte. Der Einsatz einer für die Studienteilnehmer kostenfreien Panel-App, die auf mobilen Endgeräten leicht installiert werden kann, hat den Vorteil, dass alle datenschutzrechtlichen Zusicherungen auch eingehalten werden können.

APP my infas

Die infas-Panel-App „my infas" kann sowohl einen direkten Kontakt zur Zielperson als auch einen direkten Kontakt der Zielperson zu infas aufbauen. Die App stellt jedoch keine eigenständige Trackingmaßnahme dar, sondern unterstützt das Tracking. Hauptziel der App ist es, mit den Zielpersonen in Kontakt zu bleiben („to stay in contact with respondents"). Ebenso sollen die Zielpersonen durch die App motiviert werden, weiterhin an der Studie teilzunehmen. In diesem Zusammenhang kann die infas-Panel-App ebenfalls als transferierbare Kontaktinformation verstanden werden.

Die Empirie muss es zeigen

Es ist davon auszugehen, dass nicht alle Teilnehmer einer Studie eine Panel-App nutzen werden. Jedoch stellt sie einen Versuch dar, die besonders App-affinen Personengruppen adäquater zu berücksichtigen und dabei die Vorteile der Technologie nutzbar zu machen. Da der Einsatz solch einer App bislang relativ neu ist, liegen zu deren Nutzen nur wenige ungesicherte Informationen vor. Die Herausforderung für die empirische Sozialforschung liegt vor allem darin, methodisch abgesicherte Erkenntnisse zum Einsatz von Apps mit verschiedenen Anwendungsbereichen zu sammeln. Im Kontext einer Panel-App stellt sich beispielsweise die methodische Frage, ob ihr Einsatz die Panelmortalität senkt und wie viele Studienteilnehmer unter welchen Bedingungen die App überhaupt nutzen. Auch im Bereich der Datenaufbereitung und -auswertung müssen neue Verfahren entwickelt und Erkenntnisse aus (Pilot-)Projekten gesammelt werden. Anders als bei Befragungsdaten bietet die technische Messung zwar den Zugang zu objektiven und verlässlichen, aber für viele Fragestellungen auch überflüssigen Daten(mengen). Hier gilt es, je nach Projekt und Fragestellung theoriegeleitet abzuwägen.

Zum Weiterlesen:

Albrecht, Urs-Vito et.al. (2016), Chancen und Risiken von Gesundheits-Apps (CHARISMHA). Medizinische Hochschule Hannover, http://charismha.weebly.com
Schmeißer, Daniel R. und Oberg, Nicole (2011): Die kleinen Helfer im Alltag. Zur Bedeutung von Apps für Marken und Märkte, www.phaydon.com/fileadmin/Bilder/Fachartikel_PDF-Versionen/phaydon_Helfer_im_Alltag.pdf
Schickler, Marc et al. (2015): Entwicklung mobiler Apps: Konzepte, Anwendungsbausteine und Werkzeuge im Business und E-Health. Springer: Berlin, Heidelberg.

Wie divers ist Deutschland?

Ermittlung eines bundesweiten Superdiversitäts-Index und das Beispiel Hamburg

Von Robert Follmer und Jonathan Petzold

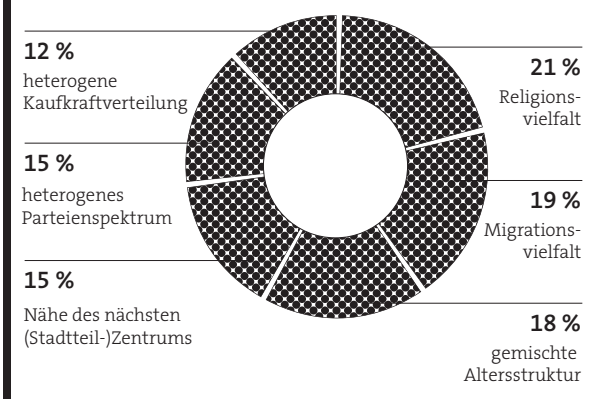
Superdiversität – was steckt denn da dahinter? Das Thema „Superdiversität“ wird seit einigen Jahren in der Stadtforschung diskutiert. Ihren Ausgangspunkt hatte diese Diskussion bei der Lebenssituation in internationalen Metropolen wie etwa London. Dort sind durch Zuwanderung und die Arbeitsplatzsituation vielfältig geprägte Stadtquartiere entstanden. Diese Vielfalt wird als kennzeichnend für kulturübergreifende Entwicklungen gesehen. Oft verbindet sich damit eine positive Erwartung, ausgelöst durch die potenzielle wechselseitige Bereicherung der in den Quartieren vorhandenen „Lebensentwürfe“.

Operationalisiert wird dies in der Regel über die Dimensionen „Zuwanderung und Migration“. Hinzu kommen – etwa in den USA – Ansätze, die Vielfalt über die Zugehörigkeit und die Verteilung ethnischer Gruppen in den dortigen Metropolen zu definieren. Entsprechende Ansätze gibt es beispielsweise für Chicago. Darüber hinaus gibt es jedoch kaum Versuche, die Dimension „Superdiversität“ genauer zu beschreiben oder weiterzuentwickeln. Dabei liegt es auf der Hand, dass ein „superdiverser“ Stadtteil nicht alleine durch eine in ihren Herkunftsnationalitäten heterogene Bewohnerschaft definiert wird, sondern weitere Merkmale von Bedeutung sind.

„Superdiversität“ in Deutschland und am Beispiel Hamburgs

Für Deutschland liegen Auswertungen in systematischer Form weder für ein auf Nationalitäten begrenztes Konzept noch eine weiter gefasste Definition vor. Ausgehend von einer Initiative der Körber-Stiftung und in deren Auftrag hat infas in Zusammenarbeit mit infas 360 als Spezialist für mikrogeografische Daten einen „Superdiversitäts-Index“ erstellt. Er wurde deutschlandweit berechnet. Die Berechnung erfolgte kleinräumig unterhalb der Ebene des Gemeindeaggregats. Damit wird auch eine Betrachtung einzelner Stadtquartiere

Der Diversitätsindex: Vielfältige Einflussfaktoren



Vorgehen im Detail: Berechnung der Ausgangswerte auf Basis der infas 360-Datenbank, Ermittlung von Heterogenitätsmaßen nach Gini für etwa 70.000 Ortsteile bundesweit, Transformation in einen einheitlichen Zahlenwert und statistisch basierte gewichtete Bepunktung je Dimension, Summenbildung für Gesamtindex über die gewichteten Einzelwerte.

Quelle: Kommunale Vielfalt. Verunsicherung oder Chance? November 2017

infas

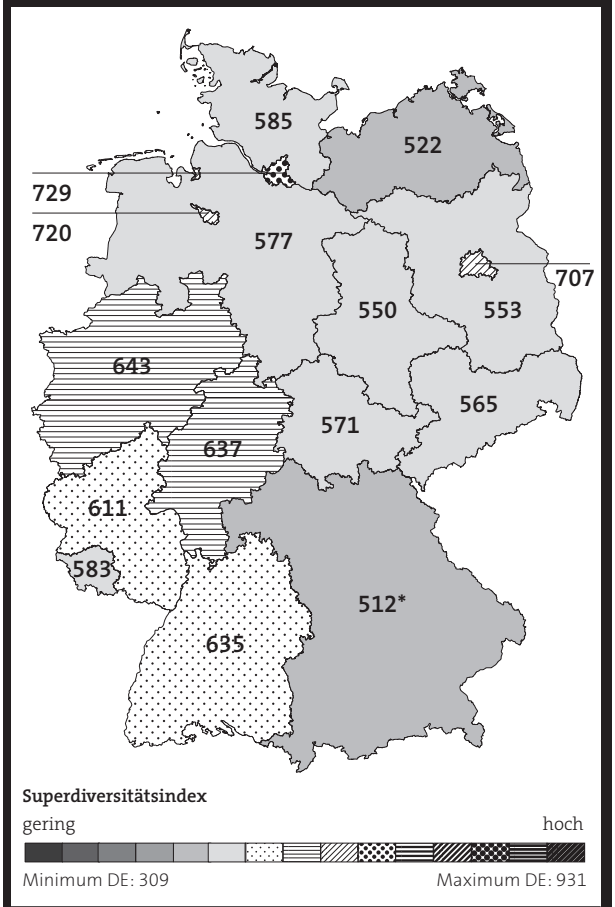
möglich. Der Index wurde für über 70.000 Ortsteile ermittelt und liegt für die gesamte Bundesrepublik vor. Dabei wurden sowohl amtliche, öffentlich verfügbare Daten wie auch mit dem Know-how von infas 360 entstandene Berechnungen spezieller Merkmale genutzt. Die wesentliche Innovation dieses Vorgehens besteht in der Zusammenführung dieser Daten und ihrer Verdichtung zu einem in dieser Form bisher nicht verfügbaren Index. Es entstand ein Vorschlag zur Messung von „Superdiversität" und dieser soll nach seiner Vorstellung auf dem Symposium 2017 der Körber-Stiftung hier ebenfalls zur Diskussion gestellt werden.

Entscheidend bei der Konzeption war die Überlegung, die bisher oft eindimensionale Betrachtung der Ebene „Migrationsanteil" um weitere Kennziffern zu ergänzen, „Superdiversität" setzt sich aus einer ganzen Reihe von Faktoren zusammen. Ausgehend von der verfügbaren Datenlage wurden zusätzlich zu den Migrantenanteilen in den jeweiligen Ortsteilen weitere Kennwerte einbezogen. Dabei kam es darauf an, dass diese Merkmale kleinräumig und vollständig für die gesamte Bundesrepublik verfügbar sind. Dadurch entstehen inhaltliche Begrenzungen. Aber ein entscheidender Vorteil liegt darin, dass der Index flächendeckend berechnet werden kann. Auf diese Weise wurden fünf zusätzliche Merkmale ausgewählt. Dies sind die Kaufkraftverteilung als Schätzer für die wirtschaftliche Situation der Bewohner in einem Quartier, die Altersverteilung, das Verhältnis der konfessionellen Orientierungen sowie das Parteienspektrum, ausgedrückt an den lokalen Wahlergebnissen bei der Bundestagswahl 2013. Als letztes Kriterium wurde schließlich noch die Erreichbarkeit des nächsten Subzentrums als Zentralitätsmerkmal und Schätzer für eine Versorgungsvielfalt berücksichtigt.

Für alle Merkmale wurde die Heterogenität im jeweiligen Ortsteil ermittelt. Dazu wurden jeweils nach dem Konzept des Gini-Koeffizienten Heterogenitätsmaße berechnet. Sie drücken aus, in welchem Ausmaß ein Merkmal in einer Gebietseinheit möglichst breit gestreut oder nur geklumpt beobachtet werden kann. Der Bezug lag dabei auf sogenannten Siedlungsblöcken innerhalb der Ortsteile. Sie stellten also die kleinste betrachtete Einheit dar und umfassen vereinfacht ausgedrückt jeweils einzelne Häuserblöcke. Nur die Entfernung des Subzentrums wurde als einfacher Distanzwert genutzt. So drückt der Index nach der Zusammenführung aller Dimensionen die „Vielfalt" bzw. „Superdiversität" eines Ortsteils aus. Er basiert auf folgenden Eingangsdaten mit daraus abgeleiteten Indexwerten und Gewichtungen innerhalb des Gesamtindex, die auf Basis einer Faktorenanalyse über alle Merkmale und den sich ergebenden Faktorladungen festgelegt wurden:

- Migrantenanteil (Indexwerte 0 – 234, Gewicht 20,5 %),
- Kaufkraftverteilung als Schätzer für die wirtschaftliche Situation der Bewohner in einem Quartier (Indexwerte 0 – 144, Gewicht 13,0 %),
- Altersverteilung (Indexwerte 0 – 214, Gewicht 18,8 %),
- Verhältnis der konfessionellen Orientierungen (Indexwerte 0 – 251, Gewicht 15,9 %),
- Parteienspektrum, ausgedrückt an den lokalen Wahlergebnissen bei der Bundestagswahl 2013 (Indexwerte 0 – 150, Gewicht 16,0 %),
- Erreichbarkeit des nächsten Subzentrums, ausgedrückt in der Entfernung (Indexwerte 0 – 179, Gewicht 15,9 %).

Diversität nach Bundesland:
Der Osten ist wenig divers – außer Berlin



Basis: 71.435 klassifizierbare Orte
* Bayern ist nur eingeschränkt vergleichbar, da dort administrativ bedingt die Ortsteile sehr klein sind (geringe Einwohnerzahl) und nicht 1:1 den Zuschnitten in den anderen Bundesländern entsprechen.

Quelle: Kommunale Vielfalt. Verunsicherung oder Chance? November 2017

infas

Mit der Zusammenführung aller Indikatoren entsteht ein Gesamtindex, der erstmals sehr unterschiedliche Dimensionen für Vielfalt berücksichtigt. In der höchsten Ebene können seine Ergebnisse auf Bundeslandebene und nach Regionstypen dargestellt werden. Dabei zeigt sich das plausible Resultat besonders hoher Indexwerte in den Metropolen und geringer Werte in den ländlichen Räumen, vor allem in einigen Regionen Ostdeutschlands. Aufschlussreicher als bei dem Blick auf diese relativ hohen Aggregate werden die Ergebnisse bei einer weiteren regionalen Differenzierung. Die Ergebnisse sind daher online in Form einer interaktiven Karte für ganz Deutschland einsehbar. In der Karte werden die

Ergebnisse zunächst nach Gemeinden dargestellt. Bei einem Zoom auf einzelne Gemeinden wechselt die Darstellung automatisch auf die Ortsteilebene. Dabei wird aufgrund der höheren Diversität in den größeren Städten auch automatisch die Kategorienabstufung angepasst. Dies kann jeweils der Legende entnommen werden. Ortsteile mit weniger als 30 Einwohnern werden in der Ortsteildarstellung nicht mit einem Wert dargestellt, da die Datengrundlage zur Berechnung des Superdiversitätsindex in diesen Fällen zu klein ist. Dies betrifft aufgrund der anderen raumstrukturellen Gliederung vor allem Ortsteile in Bayern.

Um das Ergebnis auch hier im Heft anhand einer kleinräumigen Betrachtung zu veranschaulichen, wurde wie schon auf der Veranstaltung der Körber-Stiftung im November 2017 das Beispiel Hamburg gewählt. Dort kann der Index den 104 Hamburger Stadtteilen entsprechend differenziert werden. Für die Stadtteile liegen jeweils ein Gesamtergebnis sowie die Einzelwerte der Indikatoren vor. Für das Stadtgebiet zeigt sich eine Spannweite von „superdiversen" Stadtteilen bis hin zu Stadtregionen, die im Verhältnis deutlich weniger Vielfalt aufweisen. Mit an der Spitze liegt Wilhelmsburg, geprägt durch einen relativ hohen Migrantenanteil, aber zunehmend auch durch Zuzüge vor allem jüngerer Bewohner ohne einen Migrationshintergrund. Ebenso noch relativ „superdivers" sind innenstadtnahe Stadtteile. Am anderen Ende der Hamburger „Superdiversitätsskala" liegen vor allem Regionen im Norden und Südosten der Stadt Hamburg mit relativ großer Entfernung zum Stadtzentrum.

In anderen deutschen Großstädten zeigen sich teilweise ähnliche, teilweise etwas anders gelagerte Ergebnisse. Gemeinsam ist allen Metropolen, dass sie ein deutlich höheres „Superdiversitätsniveau" aufweisen als die suburbanen oder ländlichen Räume.

Messung als Grundlage – Stadtplanung sollte „Superdiversität" ermöglichen

Das Ausmaß der Superdiversität vor allem in den Großstädten kann dabei sowohl durch einen hohen Migrantenanteil als auch durch einen oder mehrere der übrigen Faktoren geprägt sein. Superdiversität entsteht nach der gewählten Operationalisierung überall dort, wo möglichst alle Indikatoren auf eine breite Mischung im Stadtteil hindeuten. Homogene Stadtteile mit einer oft geringen Altersspannbreite der Bewohner, sei es in der arrivierten bürgerlichen Form oder ausschließlich als „Problembezirk" geprägt, werden dagegen nicht als „superdivers" klassifiziert.

Inhaltlich zeigen sich vor allem Struktureffekte. In den „diverseren" Stadtteilen ist der Anteil jüngerer Menschen und im Idealfall auch der der Familien höher als im Durchschnitt. Es gibt ebenso Stadtteile, die auf der Schwelle zur „Superdiversität" stehen, aber diese Bedingung für einzelne Dimensionen nicht erfüllen. Ein Beispiel hierfür ist der Hamburger Stadtteil Eimsbüttel, wo der Anteil von Jugendlichen – und damit der Familienanteil – unter dem Durchschnitt liegt.

Für die Stadt- und Wohnungsplanung heißt dies, dass auf eine „Vielfalt" auch bei der Quartiersgestaltung geachtet werden muss. Dazu zählen als wichtige Kriterien die Finanzierbarkeit, aber auch das Spektrum des vorgehaltenen Wohnungsbestands ebenso wie die Gestaltung städtischer Freiräume, um Familien das innerstädtische Wohnen zu ermöglichen. Unser Vorschlag zur Messung von „Superdiversität" kann dabei helfen, Stärken und Schwächen innerhalb der Städte besser als bisher und mit neuem Blickwinkel zu erkennen. Er stellt eine Anregung zur Nutzung dieser Perspektive und zur weiteren, auch methodisch ausgerichteten Diskussion dar. In der Diskussion mit externen Experten und dem Team der Körber-Stiftung sind dazu Empfehlungen erarbeitet worden, die auf der Veranstaltung vorgestellt wurden.

„Superdiversität" für alle das Richtige?

In einem weiteren Analyseschritt wurden die Indexwerte aller Ortsteile Deutschlands mit Befragungsergebnissen aus dem infas Lebenslagenindex ilex kombiniert. Hierzu liegen über 5.000 Interviews bundesweit vor. Damit wird es möglich, nicht nur Struktureffekte, sondern auch subjektive Urteile zur Lebenssituation und bestimmte „Einstellungsmilieus" abhängig von der „Superdiversität" der Wohnumgebung zu betrachten. Diese Analyse zeigt wie bereits das Hamburger Ergebnis anhand der Strukturdaten, dass vor allem lebensphasenabhängige Zusammenhänge beobachtet werden können. Die gebildeten „Einstellungsmilieus", aber auch weitere Merkmale – etwa zur Lebenszufriedenheit – hängen dagegen kaum mit der Diversitätseinstufung des Wohnquartiers zusammen. Im Umkehrschluss heißt dies jedoch, dass die subjektiv empfundene Lebenssituation ebenfalls vielfältig ist. Zufrieden lässt es sich also in superdiversen wie auch in homogenen Wohnumgebungen sein. Dahinter mag eine andere Beobachtung stecken: „Superdiversität" ist nichts für alle und je nach Einstellung und Lebensphase kann auch eine nicht-diverse Lebens- und Wohnumgebung die erste Wahl sein.

Zum Weiterlesen

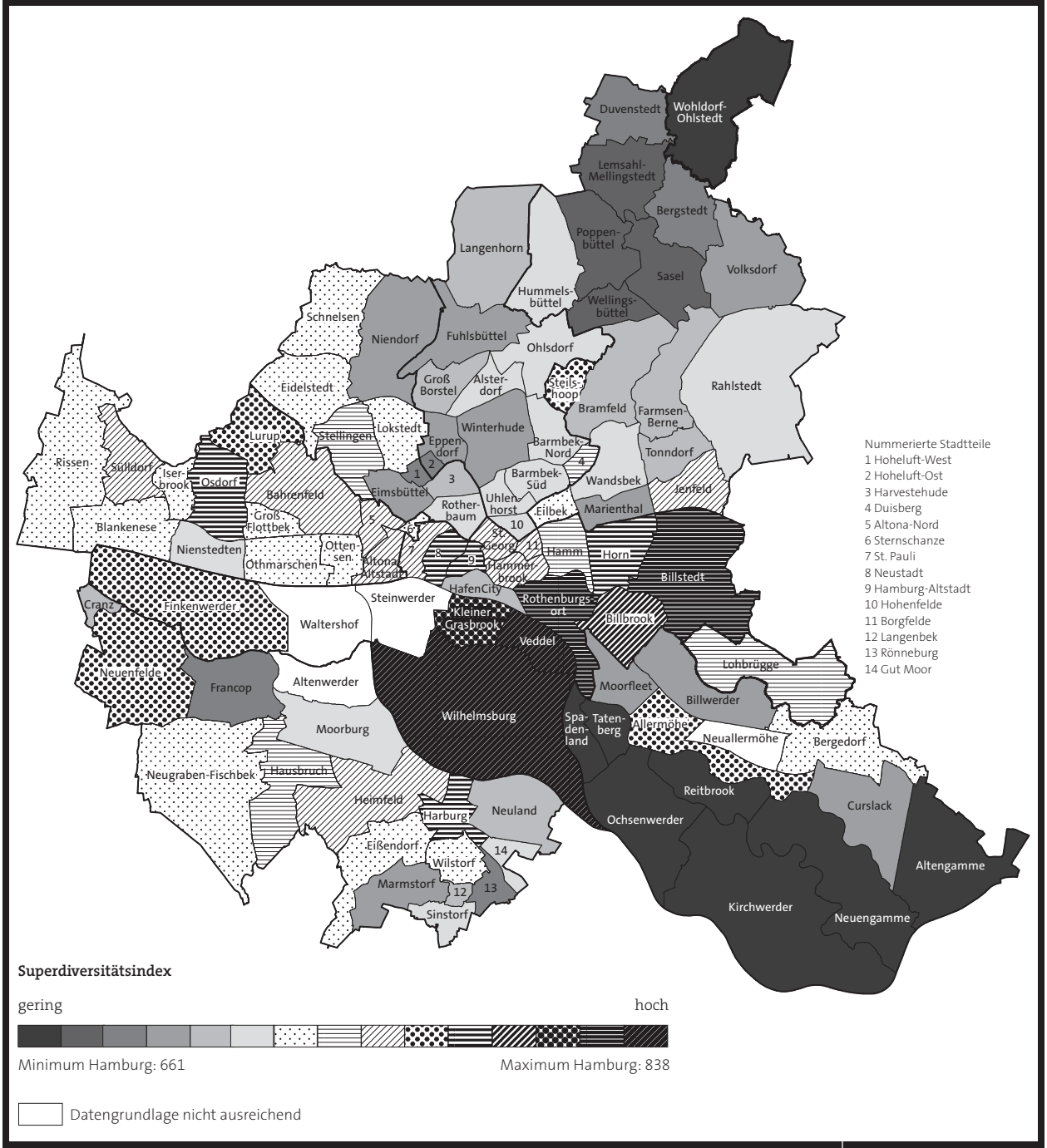
Vertovic, Steven (2012): Superdiversität, online unter https://heimatkunde.boell.de/2012/11/18/superdiversitaet, Stand 2.8.2017

8. Körber Demografie-Symposium, zu dessen Anlass die Studie durchgeführt wurde: https://www.koerber-stiftung.de/koerber-demografie-symposien/symposien/symposium-2017.html

Video mit Studienergebnissen: https://www.koerber-stiftung.de/mediathek/vielfalt-messen-geht-das-hamburgs-superdiversitaet-unter-der-lupe-1514.html

Bundesweite Auswertungen: https://www.koerber-stiftung.de/koerber-demografie-symposien/symposien/symposium-2017/superdiversitaet.

Diversität in Hamburg:
Starke Unterschiede zwischen Zentrum und Peripherie



Quelle: Kommunale Vielfalt. Verunsicherung oder Chance?, November 2017

infas
in Kooperation mit
der Körber-Stiftung

Zum Autor: Jonathan Petzold betreut bei der Körber-Stiftung gesellschaftliche Themen.

Splitter

Jul Interna

Lagemaß gewinnt Gold, 2017

„Ein bestechend klares und schnörkelloses Konzept, das inhaltlich spannend und substanziell perfekt zur angesprochenen Zielgruppe passt und die Seriosität des Absenders transportiert“, so urteilte eine Jury aus anerkannten Experten über das Kundenmagazin Lagemaß von infas und prämierte es beim „Best of Content Marketing“-Award 2017 (BCM-Award) in der Kategorie „Kundenmagazin B2B“ mit dem Hauptpreis, Gold.
Damit erzielt die Publikation eine besondere Auszeichnung, schließlich ist der BCM-Wettbewerb mit jährlich rund 750 Einreichungen Europas größter Contest im Bereich inhaltsgetriebene Unternehmenskommunikation. infas und die Berliner Designagentur kognito haben sich gegen renommierte Unternehmen wie die KfW-Bankengruppe, Ernst & Young, Arvato Infoscore oder Roland Berger und Agenturen wie Axel Springer SE Corporate Solutions, muehlhausmoers oder Sequoia Media durchgesetzt.
Bereits das Buch zum 50-jährigen Jubiläum von infas, das ebenfalls gemeinsam mit der Agentur kognito entstanden ist, hat bei diesem Wettbewerb (damals noch „Best of Corporate Publishing“-Award) im Jahr 2010 einen Award in Silber gewonnen.

bcm
best of
content marketing
Gold 2017

Dez Methoden

DIN Spezifikation für Stichproben verabschiedet, 2017

Zum Jahreswechsel 2017/18 wurde die DIN SPEC 91368 der Öffentlichkeit präsentiert. Sie wurde binnen sechs Monaten vom Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. und sechs Forschungsinstituten, darunter auch infas, erarbeitet. Die Spezifikation beschreibt die qualitätsrelevanten Kriterien und Dokumentationsanforderungen für die Ziehung von Stichproben für Umfragen in der Markt-, Meinungs- und Sozialforschung. Die DIN SPEC 91368 schließt sowohl die vom Auswahlverfahren, dem Modus der Datenerhebung und der Zielgruppe unabhängigen allgemeinen Qualitätskriterien und Dokumentationsanforderungen als auch die spezifischen Anforderungen für bestimmte Auswahlverfahren, Erhebungsmodi und Zielgruppen ein. Anhand der in der Spezifikation definierten Qualitätskriterien lassen sich Stichproben von Umfragen und damit auch die Aussagekraft der Umfrageergebnisse umfassender bewerten. Es hat sich nämlich gezeigt, dass die Fixierung auf die Ausschöpfungsquote als alleiniges Bewertungskriterium bestimmte Selektivitäten einer Stichprobe nicht adäquat abbildet. Ein zentrales Merkmal moderner Gesellschaften, nämlich die mangelnde Erreichbarkeit der Personen, die es zu befragen gilt, erfordert zusätzliche Kriterien für die Beurteilung der Güte einer Stichprobe. Deshalb sind mit der DIN SPEC 91368 nun zusätzliche Kriterien definiert worden, um Stichproben bewerten und nachfolgend über die Generalisierbarkeit der Umfrageergebnisse entscheiden zu können

Dez Studie

Bertelsmann Stiftung zum gesellschaftlichen Zusammenhalt, 2017

Im Dezember 2017 veröffentlichte die Bertelsmann Stiftung die Ergebnisse ihres „Radars gesellschaftlicher Zusammenhalt“. infas hat für die Untersuchung rund 5.000 Personen ab 16 Jahren telefonisch (Dual-Frame) befragt. Gemeinsam mit infas 360 ist ein neuer Stichprobenansatz umgesetzt worden, bei dem auf Basis von Kreisen eine Neuordnung der Gebiete vorgenommen wurde. Im Vergleich zur bisherigen BBSR-Kategorisierung ist die neue Aufteilung proprotionaler zur Bevölkerungszahl. Dadurch sind auch kleinräumigere Auswertungen unterhalb der Bundeslandebene möglich.
Ergänzt wurde die Untersuchung mit einer Zuspielung von mikrogeografischen Variablen auf amtlicher Basis. Insgesamt ergibt sich durch die Studie, an der auch die Jacobs University, Bremen, mitgewirkt hat, ein detaillierter empirischer Überblick über die Einstellungen der Bevölkerung bei Themen wie Nachbarschaft, Vertrauen, Toleranz, Gerechtigkeit, Teilhabe und Demokratie.

Mrz Interna

Luxemburger Telefonstudio erweitert Leistungen, 2018

Als eines von wenigen empirischen Forschungsinstituten in Luxemburg verfügt die dortige infas-Niederlassung über ein eigenes Telefonstudio. Das Angebot dieses Studios ist jetzt deutlich erweitert worden. Ab sofort bietet infas von Luxemburg aus Telefonerhebungen in über zehn Sprachen und in allen Kernländern Europas an. infas nutzt die

hohe Sprachkompetenz in der Region für mehrsprachige Studienprojekte aus einer Hand.
Von Luxemburg aus führt infas sowohl B2C- als auch B2B-Studien durch. Für B2B-Erhebungen verfügt die Niederlassung über ein spezialisiertes Interviewerteam für anspruchsvolle multilinguale Befragungen mit Business-Entscheidern auf Augenhöhe (inkl. A-Level Management). Ab dem 2. Quartal 2018 bietet infas für Luxemburg zudem eine kostengünstige monatliche B2C-Mehrthemenbefragung (Omnibusbefragung) für Einzelfragen oder kleinere Studien an.

Aug Studie

Studie zu Konfliktlinien in der Bundesrepublik, 2017

In Zusammenarbeit mit der Wochenzeitung DIE ZEIT hat infas im Vorfeld der Bundestagswahl (24. September 2017) eine Studie zu politischen Konfliktlinien in der Bundesrepublik durchgeführt. Dabei wurden Aspekte des Zusammengehörigkeitsgefühls der Deutschen (Was macht ein „Wir-Gefühl“ aus?) genauso zum Thema gemacht wie die wichtige Frage, aus welchen Quellen sich das neue Phänomen des Rechtspopulismus in der bundesdeutschen Wählerschaft speist. Im Frühsommer 2017 wurden zu diesen Themen 1.501 wahlberechtigte Bürgerinnen und Bürger auf Basis einer repräsentativen Zufallsstichprobe befragt. Bestimmende Faktoren für das Vorhandensein rechtspopulistischer Einstellungen sind den Befunden zufolge vor allem: eine geringere Bildung und ökonomisch niedrige Lebenslagen bei Bürgerinnen und Bürgern. Als weitere wichtige Ursache kann zudem die – nicht wenig verbreitete – Auffassung gelten, dass man sich nicht mehr richtig von der Politik vertreten fühlt. Diese Gruppe kennzeichnet zum Beispiel eine hohe Zustimmung zur Aussage „Die etablierten Parteien und die Regierung kümmern sich nicht um Leute wie mich“.
Die Studienergebnisse wurden noch vor der Bundestagswahl 2017 exklusiv in der ZEIT veröffentlicht (www.zeit.de/2017/35/bundestagswahl-deutschland-umfrage-toleranz-liberalismus).

Feb Interna

Qualitätssicherung in der Sozialforschung, 2018

Seriöses Messen erfordert immer auch eine Qualitätssicherung. Die ISO-Norm 20252, nach der infas zertifiziert ist, fordert diese auch explizit ein. Und die empirische Sozialforschung verfügt hier über einige Instrumente, beispielsweise die Interviewerkontrollen durch Mitschnitte.
Dieses Verfahren eignet sich sowohl für CATI- als auch für CAPI-Studien. Ein Beispiel ist das Panel Arbeitsmarkt und Soziale Sicherung (PASS). Seit der zehnten Welle werden hier zusätzlich zur standardmäßigen schriftlichen Kontrolle Haushaltsinterviews und Personeninterviews mitgeschnitten, sofern die Einwilligung des Befragten vorliegt. So wurden in der elften Welle beispielsweise 15.689 Interviews durchgeführt, von denen 5.707 aufgezeichnet wurden. Auf jeden der 343 Interviewer entfallen so im Durchschnitt 17 Mitschnitte. Diese werden anschließend bei infas kontrolliert und das Ergebnis anhand eines Bewertungsschemas dokumentiert.
Aus Bewertungen ergeben sich Feedbacks an Interviewer, Nachschulungen, zusätzliche Qualifizierungsmaßnahmen. Dieses Verfahren verursacht offensichtlich zusätzliche Kosten, gewährleistet aber eine äußerst hohe Qualitätssicherheit.

Nov Studie

Kurzfilmwettbewerb zur Vermächtnisstudie, 2017

Inspiriert von den Ergebnissen der Vermächtnisstudie, einer gemeinsamen Studie von DIE ZEIT, dem Wissenschaftszentrum Berlin für Sozialforschung und infas, hat die Deutsche Filmakademie Produktion zu einem Kurzfilmwettbewerb aufgerufen. Unter dem Titel „Die Welt, die wir uns wünschen. Die Filme zur großen Vermächtnisstudie“ waren Nachwuchs-Filmkünstler aufgerufen, Kurzspielfilme oder Dokumentar- und Animationsfilme einzureichen, die sich mit den Ergebnissen der Vermächtnisstudie befassen. Eine Jury hat 22 Projekte in die nähere Auswahl genommen und drei Arbeiten prämiert. In Berlin wurden die drei Gewinner gekürt und ein weiterer Publikumspreis vergeben. Der Wettbewerb wurde vom Auswärtigen Amt gefördert, die eingereichten Projekte mit je 15.000 Euro dotiert. Für ihre Vermächtnisstudie haben die ZEIT, infas und das WZB gemeinsam 3.000 repräsentativ ausgewählte Menschen in Deutschland befragt.
Die Studie liefert eine umfassende Bestandsaufnahme der Befindlichkeiten in Deutschland mit Blick auf die erwartete und gewünschte Zukunft. Weitere Informationen unter www.dieweltdiewirunswuenschen.de

Sep Studie

Zwischenbericht zur Teilhabestudie erschienen, 2017

Im Auftrag des Bundesministeriums für Arbeit und Soziales (BMAS) führt das infas Institut für angewandte Sozialwissenschaft GmbH in Bonn von 2017 bis 2021 die erste umfassende „Repräsentativbefragung zur Teilhabe von Menschen mit Behinderungen“ in der Bundesrepublik Deutschland durch. Soeben ist der erste Zwischenbericht der Studie erschienen und online abrufbar (www.bmas.de, Artikelnummer FB492). Der Bericht vermittelt einen Einblick in die Vorgehensweise und die methodischen Besonderheiten der Studie.
Die Studie untersucht, inwiefern sich Beeinträchtigung und Behinderung auf Möglichkeiten der Teilhabe in verschiedenen Lebensbereichen auswirken. In Privathaushalten werden 16.000 Menschen mit Beeinträchtigungen und Behinderungen, parallel dazu weitere 5.000 Personen ohne Beeinträchtigungen befragt. In stationären Ein-

richtungen werden zudem bundesweit 5.000 Bewohnerinnen und Bewohner für die Studie befragt, ferner 1.000 wohnungslose und schwer erreichbare Personen. Die Teilhabebefragung arbeitet auf Basis eines saturierten Stichprobenkonzepts mit umfassendem Screening in bundesweit 250 Gemeinden. Die Studie operiert multimethodisch und mit barrierefreien Befragungsmethoden. Als partizipatives Forschungsprojekt bezieht sie Menschen mit Beeinträchtigungen sowie entsprechende Fachexpertinnen und Fachexperten in den Forschungsprozess ein.

Nov Studie

BMAS veröffentlicht Studienergebnisse zu Werksverträgen, 2017

Im Auftrag des Bundesministeriums für Arbeit und Soziales (BMAS) hat ein Konsortium bestehend aus dem Zentrum für europäische Wirtschaftsforschung (ZEW, Mannheim) und infas in den vergangenen zwei Jahren eine umfassende Studie zu Verbreitung, Nutzung und möglichen Problemen von Werkverträgen durchgeführt. Jetzt wurden die Ergebnisse vom Ministerium veröffentlicht (www.bmas.de, Artikelnummer FB495 und FB496).
Die Untersuchung beinhaltet eine qualitative und eine quantitative Studie. Das Ziel der qualitativen Untersuchung „Verbreitung, Nutzung und mögliche Probleme von Werkverträgen – Qualitative Betriebsfallstudien“ lag in der empirischen Erfassung des Spektrums und der qualitativen Vielfalt der Werkvertragsnutzung in der Bundesrepublik. Basierend auf einem methodisch kontrollierten mehrstufigen Auswahlprozess wurden im Rahmen der Studie im Verlauf des Jahres 2016 insgesamt 410 qualitative Interviews in 285 Betrieben vor Ort durchgeführt. Mittels systematischer Inhaltsanalysen der Interviews identifiziert die Studie unterschiedliche Einsatzgründe und Nutzungsvarianten von Werkverträgen. Weitere Ergebnisse betreffen das Verhältnis von Werkverträgen und Leiharbeit sowie die Formen rechtlicher Grauzonen, die im Kontext von Werkvertragsnutzungen auftreten. Die quantitative Untersuchung analysiert erstmals auf Basis repräsentativer und teilweise umfangreicher Daten von mehr als 9.500 Unternehmen mit mindestens einem sozialversicherungspflichtig Beschäftigten die Werkvertragslandschaft in Deutschland. Sie trägt dazu bei, wesentliche Informationsdefizite zur Verbreitung, Nutzung, Motivlagen und Ausgestaltung von Werkverträgen (insbesondere Arbeitsbedingungen und rechtliche Aspekte) auf der Grundlage von quantitativen und belastbaren Ergebnissen zu schließen. Überdies werden erste Abschätzungen der gesamten Wertschöpfungs- und Beschäftigungseffekte durch die werkvertragsbedingte Vorleistungsnachfrage für ausgewählte Branchen vorgenommen. Die Vergabe von Werkverträgen erweist sich als ein weit verbreitetes Phänomen mit einer Vielfalt unterschiedlicher Nutzungsmuster.

Aug Studie

MiD-Befragung abgeschlossen, 2017

infas hat die Feldarbeit zur Mobilität in Deutschland (MiD) planungsgemäß abgeschlossen. Damit hat die größte bundesweite Repräsentativbefragung zur Alltagsmobilität der Bevölkerung einen wesentlichen Meilenstein genommen.
Bundesweit wurden 135.000 Haushalte befragt, 30.000 im Auftrag des Bundesministeriums für Mobilität und digitale Infrastruktur (BMVI) und über 100.000 Interviews im Auftrag von regionalen Institutionen wie Bundesländern, Verkehrsverbünden und -unternehmen, Landkreisen, Städten und Gemeinden. Aktuell werden die Daten aufbereitet und analysiert. Im Frühsommer 2018 werden die Auftraggeber der Studie mit Ergebnissen beliefert und die zentralen Resultate der MiD der breiteren Öffentlichkeit präsentiert. Dann wird auch ein Online-Tabellentool zur Verfügung stehen, etwa als Grundlage für wissenschaftliche Arbeiten.

Nächste Ausgabe Winter 2018

Die 7. Ausgabe von Lagemaß wird voraussichtlich Ende 2018 erscheinen. Das Schwerpunktthema wird der „Zusammenhalt“ sein. Er ist grundlegende Voraussetzung für jede Gesellschaft. Er bedarf regelmäßiger Anstrengungen, um auch jene einzubeziehen, die eher am Rand stehen. infas hat zahlreiche Studien durchgeführt, die den Zusammenhalt aus unterschiedlichen Perspektiven beleuchten.

66

Prozent der Bevölkerung finden es wichtig, sozial Benachteiligten und gesellschaftlichen Randgruppen zu helfen.

Impressum
Herausgeber: infas Institut für angewandte Sozialwissenschaft GmbH
Redaktion: Robert Follmer, Joachim Scholz, Menno Smid
Gestaltung: kognito gestaltung, Berlin, Astrid Blome (infas)
Autoren: Robert Follmer, Reiner Gilberg, Dana Gruschwitz, Jana Hölscher, Jette Kellerhoff, Michael Ruland, Joachim Scholz, Menno Smid, Angelika Steinwede, Thomas Weiß,
externe Autoren: Hans Dietrich (Institut für Arbeitsmarkt- und Berufsforschung), Jonathan Petzold (Körber-Stiftung), Fridolin Wolf
Mitarbeit: Kai Unzicker (Bertelsmann Stiftung)
Bildmaterial Innenteil: Lynx System Developers Inc. (www.finishlynx.com)
Anschrift: infas Institut für angewandte Sozialwissenschaft GmbH, Friedrich-Wilhelm-Straße 18, 53113 Bonn

ISSN 2509-3002

Lagemaß ist eine Publikation des infas Instituts für angewandte Sozialwissenschaft GmbH in Bonn, die in loser Folge erscheint. Das Magazin richtet sich an Interessenten und Kunden und beleuchtet gesellschaftliche Themen aus Institutssicht. Lagemaß liefert aktuelle Studienergebnisse und Einblicke in die Arbeit von infas und hat 2017 den Best of Corporate Award (BCM) in Gold gewonnen. www.lagemass.de

ISSN 2509-3002