

Was tun, wenn einem eine Statistik nicht passt?

Zwölf Grundsätze, die nützlich sein könnten, wenn man gegen statistische Daten und Analysen eines Redners argumentieren möchte

von

Peter von der Lippe

09.07.2013

In der Politik und immer dann, wenn wirtschaftliche Interessen im Spiel sind, kommt es vor, dass mit statistischen Betrachtungen untermauerte Thesen vorgetragen werden, die Ihren Vorstellungen zuwider sind. Da aber mit einer stummen Ablehnung nicht viel gewonnen ist und Sie vielleicht auch andere für Ihren Standpunkt gewinnen wollen, sollte es Ihr Ziel sein, überzeugend darzulegen, weshalb die Zahlen das nicht hergeben, was der Redner behauptet und dass die Statistik überhaupt mehr oder weniger stümperhaft ist.

Die Aussichten, hier erfolgreich zu agieren sind durchaus günstig, weil es in der Natur der Sache "Statistik" liegt, dass es hier nichts gibt, was man akzeptieren *muss*. Es gibt schlicht keine Statistik, bei der man nicht mit guten Gründen etwas bemängeln kann, und die nicht zumindest in irgendeiner Weise besser gemacht werden könnte. Wenn man es nur halbwegs geschickt anstellt, steht man also mit einer Kritik nicht auf einem verlorenen Posten. Aber selbst wenn der gewünschte Effekt (trotz der folgenden Ratschläge) nicht erzielt wird, ist das nicht tragisch, denn es kann nicht schaden, im Fall der Statistik die Dinge etwas "sportlich" und als Spiel (andere mit Zahlen als Argumente zu überzeugen) zu sehen.

Zunächst ein paar grundlegende und vielleicht etwas Mut machende Zusammenhänge, bevor wir auf konkrete Tipps zu sprechen kommen:

1. Der Umgang mit Statistik als Instrument, andere zu überzeugen, oder seine Überzeugungen zu bekräftigen, ist sehr viel emotionaler als man dies vielleicht angesichts des mathematischen (und daher für viele abschreckenden) Erscheinungsbilds der dahinter stehenden Wissenschaft "Statistik" vermuten mag.
2. Für die meisten Menschen ist das bei weitem wichtigste Kennzeichen einer guten Statistik, dass es die "richtigen" Leute (oder "unsere" Leute) sind, die sie gemacht haben. Ist das gegeben, sind sie sehr zahlengläubig und sie glauben dann auch die gewagtesten Interpretationen.
3. Aber bei der Statistik der "falschen" Leute sind die gleichen Menschen sehr schnell dabei, Manipulation zu vermuten. Grundsätzlich gilt: Bei Statistiken erwartet man dort eher böse Absichten, wo man sonst

vielleicht nur denkt, dass man es nicht besser gekonnt hat. Es gibt deshalb auch immer jemand, der dankbar ist, wenn ihm Gründe genannt werden, wonach sein Misstrauen berechtigt sein könnte. Wenn also im Auditorium ausreichend Sympathisanten Ihrer Meinung vertreten sind, haben Sie eine ganz gute Chance mit einer Kritik an einer Statistik.

Gleichwohl ist es hilfreich, wenn man sich ein paar praktische Erfahrungen im Interesse eines möglichst wirkungsvollen Einsatzes für seine Überzeugungen zunutze macht.

1. Keine Zeit für komplizierte Gedanken

Auch wenn Sie es mit (aus Ihrer Sicht) noch so unschönen statistischen Ausführungen anderer Leute zu tun haben, sollten Sie nie vergessen, dass es in der Regel nicht angebracht ist, dies durch das Vortragen komplizierter wissenschaftlicher Überlegungen zu Recht rücken zu wollen. Das ergibt sich einmal aus dem erwähnten Umstand, dass das Geschäft, um das es hier geht, mehr von Emotionen bestimmt ist als von tiefschür-

fenden mathematischen Überlegungen. Vor allem aber haben Sie einfach meist gar nicht die Zeit dafür, komplizierte Überlegungen auch nur halbwegs für alle verständlich vorzutragen. Denn die Situation von der wir ausgehen wollen ist die, dass Sie sich als Hörer im Parkett genötigt sehen, etwas zu den "unschönen" statistischen Ausführungen eines Redners sagen zu müssen.¹ Sie sind also nicht der Redner und genießen nicht schon allein deshalb eine gewisse Aufmerksamkeit und Daseinsberechtigung. Sie müssen sich diese vielmehr erst erkämpfen.

Etwas anders hinsichtlich der anzuwendenden Taktiken, ist die Situation als Diskutant bei einer Podiumsdiskussion. Sie könnten hier vielleicht etwas mehr Zeit haben, sich zu entfalten. Sicher ist das aber nicht, weil heutzutage auch solche Veranstaltung immer mehr vom Typ "Talkrunde" sind, wo jederzeit einer dazwischen quatschen und Sie aus dem Konzept bringen kann.²

Wirklich anders mag es vielleicht nur dann sein, wenn Sie zu einem Koreferat eingeladen sind, *und* wenn Ihnen und nicht dem Redner die Sympathie des Publikums gilt.

In jedem Fall aber, egal in welcher Rolle man auftritt, tut man gut daran, davon auszugehen, dass man heutzutage eigentlich nie Zeit für längere Ausführungen hat, schon gar nicht für komplizierte Ausführungen, auch wenn diese noch so interessant sein mögen.

Daraus ergibt sich als erste und oberste Regel, dass man mit einem richtigen, aber schwer zu verstehenden Argument keine Chance hat. Viel besser ist ein falsches Argument, was aber als solches erst *nach* Ihrem Auftritt durchschaut wird, wenn die Leute mehr Zeit haben, sich die Dinge noch einmal durch den Kopf gehen zu lassen. Während Ihres Auftritts ist Kompliziertes tabu, denn

¹ Die ab Punkt 4 behandelten Regeln gelten jedoch auch für eine eventuelle schriftliche Stellungnahme zu entsprechenden Ausführungen, wenn Sie also nicht aktuell eine Rede anhören, die Ihnen nicht gefällt (sondern etwas Entsprechendes lesen).

² Mehr Zeit hat man vielleicht nur dann, wenn man selbst der Redner ist. Aber als Redner besteht Ihre Rolle nicht darin, eine Statistik madig zu machen, sondern sie zu loben. Und das ist ein anderes Thema.

dazu braucht man Zeit, die Hörer meist noch mehr zum Verstehen als Sie zum Darlegen.

Das gilt zumindest beim "normalen" Publikum und dem entsprechenden Veranstaltungstyp (also z.B. nicht in einem Seminar unter Fachleuten), und das führt uns zur zweiten Regel.

2. Die Art des Publikums einschätzen und sich vorbereiten

Was bei einer Veranstaltung zu sagen angebracht ist, und was unangebracht ist, hängt vom Charakter der dort gehaltenen Referate und der Art des Auditoriums ab. Es ist deshalb wichtig, rasch richtig einzuschätzen, von welcher Art die statistischen Darlegungen sind, gegen die Sie ankämpfen wollen und von welcher Art das Publikum ist (genauer: was man ihm zumuten kann und was es gerne hören möchte).³ In beiden Punkten genügt eine Zweiteilung. Hinsichtlich der Art statistischer Ausführungen ist zu unterscheiden

- Berichte über erhobene Daten (datenorientierte oder D Themen) und
- Ausführungen über Berechnungen mit mehr oder weniger komplizierten statistischen Methoden (methodenorientierte oder "M-Gegenstände").

Was das Auditorium betrifft, so sollte man differenzieren in

- "normale" Statistiknutzer, die mehr an den Inhalten als an den Methoden interessiert sind (O-Level⁴) und
- solche, die mehr Kenntnisse über Auswertungsmethoden haben, oder zu haben glauben (A-Level).

Wir konzentrieren uns im Folgenden auf Veranstaltungen in denen es um die Diskussion von D Themen vor einem O-Level Publikum geht. Diese Kombination ist auch die bei weitem häufigste.

³ Es ist klar, dass das Publikum etwas anderes hören möchte wenn Sie vor Gewerkschaftern sprechen als wenn Sie vor Unternehmern sprechen.

⁴ In Analogie zum ordinary und advanced level im angloamerikanischen Bildungswesen.

Wenn M-Themen vor einem O-Publikum vorgetragen werden erlebt man es oft, dass die kritischen Dinge nicht oder nicht wirkungsvoll genug thematisiert werden und dass noch nicht einmal ein einziger Hörer hierzu eine Frage hat. Das spricht dafür, dass die Botschaft, die Sie bekämpfen wollen nicht wirklich "rübergebracht" worden ist. Dann können Sie Ihrem Standpunkt vielleicht auch einfach damit dienen, dass auch Sie sich gar nicht erst zu Wort melden.

Auch die Kombination D und A ist nicht wirklich interessant. Der Redner, der vor einem A Publikum "nur" Daten und läppische deskriptive Betrachtungen bringt und diese nicht wenigstens mit ein paar Signifikanztests "veredelt", wird dort ohnehin keinen Stich machen.

Bei der M-A Kombination entzünden sich seltener Kontroversen über die inhaltliche Interpretation von Statistiken, die Sie zu entsprechenden Angriffen verleiten könnten. Wenn Sie trotzdem meinen, intervenieren zu müssen, sollten Sie anders vorgehen, eher statistisch-fachlich argumentieren und quasi mehr aus der Kiste (wenn es sie denn gibt) Ihres statistischen Wissens herausholen.

Aber genau das sollten Sie bei der uns hier primär interessierenden D-O Kombination auf gar keinen Fall tun. Die meisten Schriften von der Art "Wie lügt man mit Statistik" oder "Lügen mit Zahlen"⁵ sind verdienstvoll aber mehr oder weniger nur verkappte Lehrbücher der Statistik. Wenn man nur das kann, was man dort lernt, dann taugt das nicht unbedingt für eine Wortmeldung. Es wird entweder vom Redner als "bekannt" abgehakt, oder es interessiert niemand (das gilt vor allem dann, wenn das Auditorium mehrheitlich anders denkt als Sie).

Wenn Sie in solchen Büchern z.B. etwas über Signifikanztests gelesen haben und in einer Veranstaltung sitzen, in der über die

⁵ Warum sollte man so etwas lernen? Zu einem wirklichen methodischen Disput reichen die hier erworbenen Kenntnisse nicht aus, und auch wenn das Publikum die Einzelheiten nicht versteht, so spürt es doch, dass hier ein Querulant mit seinem Halbwissen am Werk ist.

Banken geschimpft wird und dies dann auch von den Rednern mit entsprechenden Zahlen untermauert wird, ist es nicht sehr wahrscheinlich, dass viele im Saal ihre Meinung ändern werden, wenn Sie über die notorisch verkannte begrenzte Aussagefähigkeit von Signifikanztests referieren.

Das alles impliziert, dass man sich vielleicht doch die Mühe machen sollte, sich auf die (zu erwartenden) Thesen der Redner und die Meinung des Publikums vorzubereiten. Man kann das mit den etwas altertümlich, aber gerade deshalb auch ganz amüsant klingenden Worten des Freiherrn Knigge aus dem Jahr 1788 verdeutlichen:⁶

"Nicht jeder hat so viel Gegenwart des Geistes, um mitten im Getümmel und wenn er durch Fragen und Vorwitz überrascht wird, mit Würde und Bestimmtheit von Gegenständen zu reden, die er vielleicht zu Hause in seinem einsamen Zimmer mit der größten Klarheit durchschaut. Und dann gibt es auch Gesellschaften, in welchen die Leute so gänzlich anders als wir gestimmt sind, die Dinge von so durchaus anderen Seiten ansehen, daß es nicht möglich ist, in dem ersten Augenblicke sich so zu fassen, daß man etwas Gescheutes auf das antworte, was sie uns vortragen."

Wie man sieht, haben offenbar auch schon vor über zweihundert Jahren die Menschen schlechte Erfahrungen mit einem unvorbereiteten Drauflos-Reden gesammelt.⁷

3. Sich nicht lächerlich machen und planvoll Fragen stellen

Das führt uns zum dritten Grundsatz. Es ist klar dass man mit altbekannten Sprüchen wie "ich glaube keiner Statistik, die ich nicht selbst gefälscht habe"⁸ bestenfalls noch bei O-Level Hörern der schlichteren Sorte punkten kann, sich ansonsten aber eher lächerlich

⁶ Adolph Freiherr Knigge, Über den Umgang mit Menschen, Teil III, Kap. 5, Abschn. 2.

⁷ An derselben Stelle heißt es übrigens, was dem Diskussionsredner Mut machen könnte: "Der Mann, der am glänzendsten von einer Kunst schwatzt, ist darum nicht immer der, welcher die gründlichsten Kenntnisse davon besitzt."

⁸ Was angeblich Churchill gesagt haben soll, in Wahrheit aber Joseph Goebbels Churchill in die Schuhe geschoben hatte (mit Erfolg zumindest in Deutschland).

macht, was ganz schlimm ist. Sich nicht lächerlich machen ist nämlich quasi das oberste Gebot überhaupt.

Eine Gefahr sind in diesem Zusammenhang beispielsweise falsche Fragen an den Redner, der über die unliebsame Statistik referiert (im Folgenden kurz "Gegner" genannt).

Andererseits sind Fragen aber auch eine Chance, denn als Hörer eines Vortrags ist Fragen zu stellen häufig die einzige Möglichkeit, seinen Standpunkt überhaupt einbringen zu können. Man muss etwas sagen, was irgendwie in eine Frageform gegossen werden kann, obgleich man im Grunde nicht wirklich eine Frage hat und schon gar nicht von diesem Referenten eine Antwort hören möchte. Und wenn es die "richtige" Frage ist, kann man mit ihr sogar den Referenten gekonnt aushebeln.

Aber – wie gesagt – man kann sich mit Fragen aber auch leicht lächerlich machen; sie sind also ein zweischneidiges Schwert.⁹

Von welcher Art sind nun die Fragen, bei denen man damit rechnen kann, dass sie den Gegner in Verlegenheit bringen können? Bei einem O-Level Publikum, das einem wohlgesonnen ist, kann es durchaus eine platte, eher schlichte Überlegung zur Interpretation der Zahlen sein, die Sie in Gestalt einer Frage vortragen. Sie dürften Erfolg haben wenn Sie dabei beachten: wichtig ist einfach, anschaulich, vielleicht auch etwas lustig und Emotionen ansprechen. Das Ziel muss sein, sympathisch rüber zu kommen.

Das kann natürlich nicht Ihr Ziel sein bei der ganz anderen Situation eines A-Level Publikums, oder wenn Sie in der Höhle des Löwen sind. Das Publikum muss nicht das Gefühl haben, dass Sie ein netter Kerl sind, also einer von ihnen sind, sondern dass Sie eine Art Chef-Experte von der anderen Seite sind. Den Gegner zu attackieren erfordert dann oft mehr als nur eine Frage. Was wann gesagt

wird muss gut geplant sein, darf aber nicht so erscheinen. Der Sinn der ersten Frage, von der – wie gleich deutlich wird – alles Weitere abhängt, ist, sich als Sachkenner auszuweisen und vorzufühlen, wie sicher der Gegner in der Materie ist. Man könnte es vielleicht zunächst mit einer mehr methodenorientierten Frage versuchen, die man mit ernster und wichtiger Miene vorträgt, die aber eher nur ein Randproblem, einen eigentlich uninteressanten Nebenschauplatz betreffen sollte.

Je nebensächlicher das ist, wonach sie "fragen", desto besser, weil Sie damit

- von dem ablenken, was dem Publikum – und auch dem Gegner – eigentlich am Herzen liegt (und beide verbindet), ihnen aber nicht gefallen kann¹⁰ und damit auch
- eher etwas ansprechen, worüber sich der Gegner noch nicht so viele Gedanken gemacht hat, und was ihn deshalb vielleicht eher in Verlegenheit bringen kann.

Erst wenn man den Eindruck haben kann, den Gegner auf den falschen Fuß erwischt zu haben, was dieser natürlich zu vertuschen versucht,¹¹ kann man fortfahren, wobei sich der Charakter der Statements ändern kann und soll: weg von "Fragen" und mehr hin zu Thesen oder gar Angriffen.

Wenn aber der erste Schritt nicht gelungen ist, was daran zu erkennen ist, dass im Publikum während Ihrer "Frage" die Aufmerksamkeit nachließ und gemurmelt wurde, nicht aber während der Widerrede des Gegners, sollte man aufhören oder aber bei einem neuerlichen Versuch kleinere Brötchen backen, also eher Sympathie mit dem Standpunkt des Gegners heucheln. Wer dies nicht beachtet läuft Gefahr, als Querulant wahrgenommen zu werden¹² und aus dieser Rolle heraus andere zu überzeugen ist schwierig.

⁹ Es gilt deshalb die Regel: keine Fragen stellen zu Methoden, die man selbst nicht besonders gut kennt. Es darf nicht (auch nicht beim O-Level Publikum) der Eindruck entstehen, dass man etwas nicht verstanden hat, selbst dann nicht, wenn alle anderen es auch nicht verstanden haben.

¹⁰ Es sind genau die Dinge, weshalb das Publikum *emotional* dem Gegner näher steht und gerade deshalb seiner Sicht eher zustimmen lässt als Ihrer Sicht.

¹¹ Das Startsignal für die nächsten Schritte wird also meist nicht sehr deutlich sein.

¹² Wer es sich zutraut, kann natürlich nach anfänglicher Ablehnung im gleichen Stil weitermachen. Das

Welche ziemlich allgemein und überall verwendbaren Bemerkungen sind es nun, die nützlich sein können für das Ziel, um das es hier geht, eine Statistik (notfalls indirekt über den Referenten) madig zu machen?

4. Die Erhebung ist nicht neutral

Eine erste Frage könnte vielleicht sein: wer hat die Erhebung gemacht, den Auftrag und das Geld dazu gegeben? Und was war das Ziel der Erhebung bzw. der Analyse der erhobenen Daten?

Ein Angriff auf die Seriosität des Auftraggebers, mit dem gleich zu Beginn Neutralität und Unparteilichkeit in Frage gestellt wird, kann funktionieren, denn das Wichtigste bei Statistik ist, dass sie "akzeptiert" wird.

Das ist nicht selbstverständlich, denn Statistik ist aus zwei Gründen notwendig immer Vertrauenssache:

1. Sobald genügend viele Einheiten (Personen, Betriebe etc.) befragt worden sind, ist es praktisch unmöglich, sich durch ein "Nach-Erheben" auf eigene Faust von der Richtigkeit der Daten zu überzeugen.¹³ Man muss also immer irgendwann doch "glauben", dass richtig gefragt, gezählt, gerechnet wurde.
2. Weil viele (oder die meisten) Menschen die Materie "Statistik" und die mit ihr verbundene Mathematik nicht besonders lieben, wenn sie nicht sogar insgeheim das Gefühl haben, die Sache nicht voll zu durchblicken, und dass man mit Statistik "alles beweisen" kann¹⁴, können sie nicht die einzelnen Aussagen eines Redners beurteilen und sie müssen sich

statt dessen daran halten, wie seriös der Redner ist, oder die sind, deren Position er vertritt. Das bedeutet dann aber auch, dass ein Ihnen wohl gesonnenes Publikum entsprechende Andeutungen über zweifelhafte Seriosität dankbar aufgreift.

Dass Statistik oft nicht im Detail beurteilt werden kann und deshalb stets Vertrauen verlangt, kann man sich auch zu Missbräuchen zu Nutze machen. Wer wirklich "ergebnisorientiert" arbeitet, oder sich seiner Sache nicht ganz sicher ist, kann das in der Statistik, mehr als auf anderen Gebieten leicht und eindrucksvoll überspielen; am besten mit der immer häufiger werdenden mathematischen Schaumschlägerei,¹⁵ die insbesondere bei M-Themen in akademischen Kreisen enorm beliebter ist. Das "Vertrauen" erwirbt man hier (i.d.R. zu Unrecht) durch den Respekt, den viele haben vor der Mathematik: wenn jemand so gewandt über so schwierige Sachen reden kann, dann muss er schon viel Ahnung haben.

Sich vor einem O-Publikum auf Schaumschlägerei einzulassen ist aber nicht unbedingt ratsam. Wenn Sie z.B. den Gegner fragen, nach welcher Formel er den Stichprobenfehler berechnet hat, kann es passieren, dass er antwortet: *Wir haben dies mittels des SAS-Makros CLAN von Statistics Sweden berechnet, welches die Taylorlinearisierung der üblichen Formeln der Standardfehler ... nutzt.*¹⁶

Das klingt eher nach heißer Luft, mit der man kaschiert, dass man nicht wirklich weiß, was der Stichprobenfehler ist (und vielleicht auch was "Taylorlinearisierung" ist). Das Problem ist nur, dass *Sie* das vielleicht durchschauen, nicht aber das O-Level Publikum. Es bringt Ihnen also nichts. Denn selbst,

setzt aber viel Selbstbewusstsein, unbekümmerte Konfliktbereitschaft und gute Rhetorik voraus, was sich nicht jeder leisten kann, dann aber vielleicht sogar Bewunderung erzeugen kann.

¹³ Das betrifft vor allem die amtliche Statistik. Es kann nicht oft genug betont werden, dass das Vertrauen in ihre Seriosität ihr eigentliches Kapital ist.

¹⁴ Man kann darüber streiten, ob dies nicht sogar bei dieser Materie bis zu einem gewissen Grade wahr ist. Es scheint zumindest so zu sein, dass man im Falle der Statistik hin und wieder mit Hochstapelei durchaus Erfolg haben kann.

¹⁵ Auf die Regeln, die hier zu beachten sind, wo jeder unter dem Zwang steht, sich als Statistikexperte aufzuführen, dem nichts zu schwer ist, um es (angeblich) zu verstehen, bin ich in dem satirischen Text "Statistik für Schaumschläger" (ebenfalls auf dieser Website) eingegangen.

¹⁶ Es gibt Leute, die verwechseln eine Frage nach einer Formel mit der Frage danach, mit welchem konkreten Programm die Berechnung der Formel erfolgte. Von wem das Zitat stammt soll höflich verschwiegen werden, aber Fußnote 19 mag man als diskreten Hinweis nehmen.

wenn Ihre Frage als klug empfunden wird, kann dies leider auch für die Antwort des Gegners gelten, selbst wenn diese nur heiße Luft ist.

Es ist deshalb viel sinnvoller, nicht bei Methodenfragen, sondern an einer ganz anderen Stelle anzusetzen. Gerade weil Statistik ganz entscheidend eine Vertrauenssache ist, sollte ein Angriff zuerst auf die Vertrauenswürdigkeit zielen. Die Kompetenz in Frage zu stellen ist demgegenüber weniger erfolgversprechend.¹⁷

Was in den Augen der meisten Menschen vor allem geeignet ist, die Seriosität in Frage zu stellen ist das vermutete Interesse an ganz bestimmten Ergebnissen einer Statistik. Es ist eine ausgemachte Sache, dass Mieten zu hoch sind, wenn der Mieterbund die entsprechende Erhebung gemacht hat und dass sie zu niedrig sind, wenn es die Vermieterseite war. In jedem Fall sind solche (nichtamtliche) Daten quasi a priori "ergebnisorientiert" und schon allein deshalb (!) falsch, bzw. falsch gemacht worden.¹⁸

Hier wird ganz völlig selbstverständlich eine kollektive Schuldvermutung gegenüber den Befragten¹⁹ akzeptiert, wo doch sonst immer die Unschuldsvermutung gilt. Das ist offenbar deshalb so leicht möglich weil viele in Statistik ein eher fragwürdiges Geschäft sehen, das auf Manipulation angelegt ist.

Bei der Unschuldsvermutung heißt es immer "bis zum Beweis des Gegenteils". Aber wie kann der entsprechende "Beweis" bei einer

¹⁷ Bei einer Sache, die mehr Prestige genießt dürfte das anders sein: es mag interessant sein, ob A der bessere (was immer das heißen mag) Physiker ist als B, nicht aber ob A der bessere Statistiker ist als B.

¹⁸ Das verhilft natürlich der amtlichen Statistik zu einem unschätzbaren Wettbewerbsvorteil. Es sei dahin gestellt, ob sie es wirklich immer verdient hat, dass bei ihr (und nur ihr) ein Angriff auf die Seriosität und wissenschaftliche Kompetenz praktisch tabu ist.

¹⁹ Ergebnisorientiertes Agieren kann man eher bei den Befragten als bei den Institutionen, die die Erhebung durchführen insinuierten. Letzteres liefe ja darauf hinaus, dass man die richtigen Antworten der vielen Befragten nachträglich verfälscht hat. Man bekommt mehr Probleme, wenn man schwere Untaten bei wenigen Menschen anprangert als wenn man leichte, "menschliche" Fehler bei vielen vermutet.

Schuldvermutung erfolgen? Wie beweist man als Angegriffener, richtig geantwortet zu haben wenn alle glauben, man müsste doch eigentlich ein Interesse haben, zu lügen?²⁰

Sie als Angreifer müssen nicht beweisen, dass der Mieterbund bei seiner Befragung getrickt hat. Es wäre sogar schlecht, wenn Sie mehr täten, als so etwas nur mal "in den Raum" zu stellen. Andeuten ist genug, "beweisen" wollen ist überflüssig oder sogar schädlich. Denn für viele ist es völlig klar, was passiert (auch wenn es gar nicht passiert), wenn der ADAC Meinungsumfragen zum Tempolimit macht, Greenpeace zum Atomausstieg oder die katholische Kirche zur Abtreibung.²¹

Bei der Manipulationsvermutung ist es genauso, wie bei dem Einwand, um den es im nächsten Punkt geht ganz wichtig, vage zu bleiben, was einem aber auch leicht gemacht wird, weil man darauf bauen kann, dass entsprechende Vorurteile weit verbreitet sind

5. Die Erhebung ist nicht repräsentativ

Dieser Einwand ist zwar altbekannt, aber er wird wohl trotzdem nie obsolet werden, weil er seine Wirkung vor allem dem Umstand verdankt, dass niemand so richtig weiß, was "Repräsentativität" überhaupt bedeutet. Die Kenntnis der Sache beschränkt sich hier meist auf das *Gefühl*, das wohl bei jedem Menschen als gegeben vorauszusetzen ist: eine Erhebung ist gut, wenn sie repräsentativ ist und sie ist schlecht, wenn sie nicht repräsentativ ist.

Leute, denen etwas mehr zur "Repräsentativität" einfällt sind selten. Jeder glaubt zu wissen was das ist, aber hier konkreter zu werden dürfte sehr schwierig sein, weil – wie

²⁰ Es wird auch oft übersehen, dass es für jemand, der kein Interesse daran hat, dass die Wahrheit bekannt wird, viel einfacher und bequemer ist, den Fragebogen einfach gar nicht erst auszufüllen.

²¹ Zu diesen Beispielen vgl. W. Krämer, Lügen auf ökonomischen und wissenschaftlichen Informationsmärkten, in R. Hettlage (Hrsg.) Verleugnen, Vertuschen, Verdrehen, Leben in der Lügengesellschaft, Konstanz 2003, S. 168.

wir gleich sehen werden – "Repräsentativität" auch kein wirklich sinnvoller Begriff ist.

Aber das ist ganz und gar kein Problem, im Gegenteil: man wird Sie "verstehen" gerade weil sie nicht konkreter werden. Sie brauchen auch nicht zu befürchten, dass man Sie auffordert "Repräsentativität" zu definieren oder genauer zu erklären, wie Sie zu Ihrem Urteil gekommen sind. Es wird auch niemand von Ihnen verlangen, etwa den Grad der "Repräsentativität" einzuschätzen, z.B. auf einer Skala von null bis zehn, weil das auch niemand könnte.

Der Einwand "nicht repräsentativ" ist also uneingeschränkt zu empfehlen: man kann ihn auch dann gefahrlos bringen, wenn man von Statistik so gut wie gar nichts versteht und andererseits auch nicht wirklich mitreden kann bei dem inhaltlichen Gegenstand der zur Diskussion stehenden Statistik.

Eine allerdings etwas anspruchsvollere Variante dieses Einwands ist:

- Ihre Daten umfassen ja nur die Jahre bis 2010 und danach hatten wir eine ganz andere Situation (oder gerade 2011 ist ja gewesen), oder
- sie gelten zwar für den Durchschnitt der Bundesländer, sind aber auf Bayern nicht übertragbar (oder auch umgekehrt: sie mögen vielleicht für Bayern gelten, können aber wohl nicht verallgemeinert werden).

Das setzt natürlich voraus, dass Sie im Zweifel etwas dazu sagen können, was nach 2010 oder in Bayern so ganz anders war bzw. ist.

Solche vertieften Kenntnisse sind aber zum Glück nicht nötig, wenn Sie ganz allgemein bemängeln, dass die Stichprobe zu klein ist. Der Vorwurf nicht repräsentativ zu sein wird nämlich meist damit begründet, dass nur so wenige Einheiten (Personen, Betriebe etc.) befragt worden sind.²² Dabei ist die Höhe der

²² Die meisten Menschen kennen nur diesen Bestimmungsfaktor für den Stichprobenfehler. Dass der Stichprobenfehler auch von der Varianz in der Grundgesamtheit abhängt ist oft nicht bekannt. Bei entsprechend geringer Streuung kann z.B. 50 Befragte durchaus sehr viel sein. Es ist überhaupt bemerkenswert,

konkreten Zahl mehr oder weniger Gefühlsache. Man kann vorwurfsvoll sagen: bei der Erhebung sind ja nur 50 Anwälte für Ehe- und Familienrecht gefragt worden. Entscheidend ist allein, dass die Zahl nicht sehr groß ist. Aber ab welcher Grenze wird etwas "repräsentativ"? Das ist eine Frage, die man sich interessanterweise eigentlich nie stellt. Würde man von "nur" 500 Anwälten sprechen wäre das vielleicht nicht so überzeugend. Die Grenze könnte so bei 100 liegen, darunter ist es nicht repräsentativ. Vielleicht gibt es hier gefühlsmäßig Schwellen, so wie z.B. 5 bei einem Preis, weshalb 4,95 € ein akzeptablerer Preis ist als 5,05 €.

Eine weitere Vorstellung, die sehr verbreitet ist im Zusammenhang mit "Repräsentativität" ist, dass in der Stichprobe die Proportionen in etwa so sein müssen, wie sie in der Grundgesamtheit sind. Wenn es plausibel erscheint, dass es etwa gleich viele Fachanwälte für Scheidungsrecht wie für Arbeitsrecht gibt, dann wären 50 Scheidungs- und 200 Arbeitsrechtler nicht akzeptabel.²³

Das hat diese Taktik mit der vorher besprochenen Taktik, die Neutralität in Frage zu stellen gemeinsam, dass man sie

- auch ohne viel statistischen Sachverstand und Einfallsreichtum in der Sache wirkungsvoll einsetzen kann, und dass
- man es dabei am besten bei Andeutungen belässt; Plausibilität reicht, konkreter werden oder gar sich auf Diskussionen einzulassen ist eher schädlich.

Der folgende Einwand lebt ebenfalls davon dass er inexakt ist, aber trotzdem vielen intuitiv einleuchtet. Auch er verlangt nicht Sachverstand in puncto Statistik, wohl aber etwas

dass über die Größe des Stichprobenumfangs offenbar gern gefühlsmäßig geurteilt wird, wobei es für diesen doch ganz klare Berechnungsformeln gibt.

²³ Es fällt wohl niemandem auf, dass sich die beiden Vorstellungen von "Repräsentativität" nicht immer harmonisch ergänzen. Bei gleich vielen Männern wie Frauen wäre eine Stichprobe mit 50 Männern und 50 Frauen nach dem Strukturkonzept besser als eine mit 500 Männern und 480 Frauen. Nach dem Konzept, das auf den Stichprobenumfang n abstellt wäre sie aber schlechter.

mehr Einfallsreichtum hinsichtlich der Inhalte, um die es bei der fraglichen Statistik geht.

6. Die Zahlen sind nicht vergleichbar: hier werden Äpfel mit Birnen verglichen

Wenn es sachlich gesehen auch nur irgendwie angebracht erscheint, wird der Hinweis darauf, dass hier die berühmten Äpfel mit Birnen verglichen werden immer dankbar entgegengenommen. Dabei ist es genaue genommen gar nicht mal so schwer, diesen Einwand zu bringen. Vergleichbarkeit ist nämlich wie Repräsentativität kein exakt definierter Begriff, zugleich wird es aber – und das macht seinen Reiz aus – allgemein als etwas Gutes und Wichtiges empfunden.

Vergleichen ist zentral für Statistik: Wir suchen nach signifikanten Unterschieden, berechnen Quoten und Raten, mit Größen in Zähler und Nenner, die miteinander sinnvoll verglichen werden können und ziehen (aus eben diesem Vergleich) Schlüsse.

Oft sind die entsprechenden Schüsse unsinnig, weil auf die Vergleichbarkeit nicht geachtet wird. Ein Beispiel ist die angeblich höhere Lebenserwartung von Dirigenten,²⁴ woraus dann gefolgert werden mag, dass offenbar Musik Balsam für die Seele ist und lebensverlängernd gewirkt hat.

Dabei ist zu beachten, dass die mit ihr verglichene "allgemeine" Lebenserwartung²⁵ schon deshalb geringer sein dürfte, weil sie davon beeinflusst ist, dass Menschen auch schon in jungen Jahren sterben, lange bevor sie überhaupt die Chance haben, einmal Dirigent zu werden.

Das oberste Gebot der Statistik ist also: nur das vergleichen, was auch "vergleichbar" ist. Aber was ist "Vergleichbarkeit"; wann ist

der Vergleich "sinnvoll" und wann "hinkt" er? Und warum sollen überhaupt Äpfel mit Birnen gar nicht vergleichbar sein?

Natürlich kann man Äpfel mit Birnen vergleichen. Man kann sie auch mit Bananen und Pflaumen vergleichen, bzw. statistisch in einen Topf werfen, wenn es um "Obst" geht. Wenn man dagegen mehr Merkmale ins Spiel bringt, in diesem Sinne also genauer hinschaut, kann man Äpfel noch nicht einmal mit Äpfeln vergleichen. *Sehr* genau genommen ist ein Boskop dann eben etwas ganz anderes als ein Braeburn.

Vergleichbarkeit hängt also davon ab, wie genau man hinschauen möchte. Mit genügend vielen Merkmalen im Spiel hat man nur noch Unikate und keine Statistik mehr. Die Dinge müssen in bestimmten Punkten gleich sein, wobei das Schöne ist, dass keiner genau sagen kann, wie viele Punkte es sein müssen, es aber andererseits auch genau darauf ankommt, dass es nicht zu wenige sind (denn ganz verschiedene Dinge, wie Autos und Obst kann man nicht vergleichen). Aber sie müssen aber gleichzeitig auch in einem "entscheidenden" Punkt ungleich sein. Auch hier liegt wieder der Charme darin, dass man endlos darüber streiten kann, ob der Punkt, in dem sich die Dinge unterscheiden auch der entscheidende Punkt ist.

Vergleichbarkeit ist also keine Angelegenheit von ja oder nein,²⁶ sondern eher etwas von mehr oder weniger, wobei keiner so richtig weiß, wie man den "Grad" der Vergleichbarkeit bestimmen sollte (ganz ähnlich wie bei der sog. "Repräsentativität").

Um also das Standardargument bringen zu können, dass hier angeblich wieder einmal Äpfel mit Birnen verglichen werden, muss man nur plausibel machen dass es unbedingt notwendig ist, in bestimmten Punkten weitere Differenzierungen vorzunehmen.

Es macht nicht viel Sinn, sich darüber aufzuregen, dass in einer Statistik eine bestimmte Differenzierung vorgenommen wurde (schon gar nicht, wenn sich herausstellen sollte, dass

²⁴ Das Beispiel wird im Detail kritisch auseinandergenommen in Robert P. Abelson, *Statistics as Principled Argument*, Hillsdale NJ, 1995. Offensichtlich ist hier die Lebenserwartung nicht, wie sonst üblich, mit einer Sterbetafel bestimmt worden (denn es gibt ja keine einjährige Sterbewahrscheinlichkeit von 0, 1, 2, ... jährigen [späteren] Dirigenten), sondern als durchschnittliches Sterbealter von Dirigenten.

²⁵ die in der üblichen Weise aufgrund einer Sterbetafel errechnet wurde.

²⁶ Als so etwas wird sie natürlich hingestellt, wenn man den Einwand der Nichtvergleichbarkeit bringt.

diese gar nicht "signifikant" war), und somit etwas Überflüssiges erhoben wurde (das mag man bedauern, bringt aber keinen auf die Palme). Man muss, im Gegenteil, es als eine absolut dramatische und alles zunichtemachende Fehlleistung hinstellen, dass bestimmte Differenzierungen *nicht* vorgenommen wurden.

Wenn Ihnen nichts Passendes einfällt, was dramatisch genug erscheint, können Sie versuchen, das Argument der fehlenden Vergleichbarkeit in der folgenden Variante vorzubringen: die Zahlen spiegeln nicht nur den Einfluss von x , sondern auch den störenden Einfluss von y und z wider.

Man kann z.B. argumentieren, dass zwar x gestiegen ist, was aber wenig bedeutet und zu relativieren sei, weil in der gleichen Zeit ja auch y gestiegen oder z gefallen sei. Solche Betrachtungen betreffen natürlich in erster Linie den zeitlichen Vergleich.

Aber das Potential der Vergleichbarkeitsmaschine ist sehr viel größer und auch nicht auf den zeitlichen Vergleich beschränkt. Was zunächst fast etwas naiv und laienhaft klingen mag (Äpfel und Birnen!) ist quasi das Einfallstor für eine ganze Reihe von durchaus respektablen methodischen Betrachtungen in der Statistik.

Hierfür nur vier Beispiele:

1. Die Rentabilität in einer Branche kann steigen, ohne dass sich die Rentabilität der Unternehmen verändert hat, nur weil sich die Struktur der erhobenen Gesamtheit der Unternehmen durch Ausscheiden der sog. "Grenzanbieter" zu Gunsten der rentableren Unternehmen verändert hat.

2. Es ist klar, dass man beim Vergleich von "Quoten" oder "Raten"²⁷ auf die Struktur der verglichenen Massen achten muss. Die Sterberate²⁸ der Priester ist höher als die der

Bergarbeiter, ohne dass deswegen Priester ein "gesünderer" Beruf ist als Bergarbeiter. Der Grund ist, dass Priester im Schnitt älter sind als Bergarbeiter und man deshalb eigentlich eine die (Alter-) Strukturabhängigkeit der Sterberate berücksichtigende "standardisierte" Sterberate berechnen müsste.

3. Eine gleich große Arbeitslosenquote (nicht "Rate") kann es geben, wenn Menschen häufig arbeitslos werden, dafür dann aber relativ schnell wieder einen neuen Job finden, oder wenn die Menschen zwar seltener arbeitslos werden, dann aber lange arbeitslos bleiben. In einem solchen Fall von nicht vergleichbaren Situationen, müsste man eigentlich die entsprechende Quote in Komponenten (Betroffenheit und Dauer) zerlegen.

4. Beim zeitlichen Vergleich von Preisen durch Preisindizes ist klar, dass darauf zu achten ist, dass veränderte Qualitäten der Güter oder (was ein extremer Fall hiervon ist) das Aufkommen neuer und Verschwinden alter Güter einen "reinen" Vergleich stören können. In diesem Sinne liefert z.B. ein Kettenindex keinen "reinen" Preisvergleich, weil er auch die laufende Veränderungen des Warenkorb nach Art und Menge der Güter widerspiegelt (und gerade deswegen von vielen geschätzt wird).²⁹

Bei solchen Betrachtungen ist aber das Problem, dass sie zu schnell statistisch-technisch wirken und dass man Sie dann als Schulmeister (wenn das, was Sie sagen, methodisch etwas anspruchsvoller ist) abtut, oder aber als Wichtigtuier belächelt (wenn das, was Sie sagen, jeder sofort versteht).

Das Äpfel-und-Birnen-Argument wirkt oft besser wenn es mit verbalen Ausführungen garniert wird, die aus der Erfahrungswelt des

²⁷ Das sind Begriffe, die erfahrungsgemäß völlig beliebig gebraucht werden. Was einer Quote nennt, kann der andere Rate nennen und umgekehrt. Wenn man schon so wenig Sorgfalt auf die Worte legt, wie unbekümmert geht man dann erst mit der Sache selbst um!

²⁸ Anzahl der Gestorbenen (in einem Jahr) je 1000 Lebende. Die verglichenen Massen sind also "Gestorbene" im Zähler und "Lebende" im Nenner.

²⁹ Aber wenn Sie darauf hinweisen, dass bei Kettenindizes wegen des ständig wechselnden Warenkorbs kein reiner Preisvergleich möglich ist (der Preisindex also nicht isoliert *nur* die Preisbewegung misst), wird man Ihnen entgegenhalten, dass ein konstanter Warenkorb schnell veraltet, also nicht mehr "repräsentativ" ist. Es ist klar, dass beides nicht geht, nur Preise vergleichen bei konstanten Mengen *und gleichzeitig auch* die laufende Veränderung der Mengen berücksichtigen. Aber irgendwie erwartet man immer gerne von der Statistik eine Quadratur des Kreises.

Auditoriums stammen. Beispiele für gekonnt angebrachtes Hintergrundwissen bei entsprechenden Kontroversen kann man fast täglich in der Presse und im Internet finden. Zur Illustration möge nur ein aktuelles (Dez. 2012) Beispiel genügen:

Bei Diskussionen über eine (Re-)Reform der Professorenbesoldung in Deutschland wird z.B. gerne auf höhere Bezahlungen in den USA oder England verwiesen, ohne – wie es in einem Kommentar (eines Lesers) im Internet dazu heißt – zu erwähnen, dass es nur wenige Top-Leute sind, die an den dort üblicherweise privaten Institutionen solche Bezahlungen erhalten, dass so etwas auch nur "auf wenige Forschungsfelder reduziert" zutrifft, und dass sich deshalb die dort meist "zeitlich befristet tätigen Lehrer, die private Vorsorge treffen müssen", in einer ganz anderen Situation befinden, als in Deutschland. Bei uns bekommt sowohl ein Juraprofessor als auch ein Experte in der "Geschichte der Höhlenmalerei" praktisch an jeder Uni die gleiche Besoldung und später auch eine gleich hohe Pension. Weiter sei zu bedenken "dass die Arbeitszeiten von W3 besoldeten Professoren in der Wirtschaft nicht einmal als Halbtagsjob durchgingen".³⁰

Dies alles leitet mühelos über zu der im folgenden Abschnitt behandelten Methode, der impressionistischen, wenn es dann auch in dem Text folgerichtig weiter heißt:

"Wenn es in den USA und in der Wirtschaft so kommod ist, was macht der Autor dann noch an einer für ihn so piffligen deutschen Hochschule?"

Das ist übrigens ein gutes Beispiel für die Schlussweise, die für die impressionistische Art der Betrachtung typisch ist: Dass die Professoren nicht in die USA gegangen sind

und in Deutschland geblieben sind, ist der Beweis, dass die Besoldung in Deutschland gut war.³¹

7. Mehr Zahlen oder gar keine Zahlen; der impressionistische Gegenschlag

Was auf dem ersten Blick etwas widersprüchlich zu sein scheint ist es nicht, wenn man bedenkt, dass die meisten Menschen ein ambivalentes Verhältnis zu Zahlen haben, die Ihnen von der Statistik geliefert werden.

In einem Lehrbuch der Statistik heißt es:³²

"There is something very sad in the disparity between our passion for figures and our ability to make use of them once they are in our hands."

Es gibt Leute, die es schaffen, den Eindruck zu erwecken, dass wir einer Katastrophe entgegengesehen, weil wir über irgendetwas immer noch keine Statistik haben, oder weil die entsprechende Statistik erst in einer Woche kommt und nicht schon jetzt da ist. Wenn sie dann aber die angeblich so alles entscheidende Zahl "endlich" sehen, dann fällt ihnen kaum mehr dazu ein, als "interessant" oder "hätte nicht gedacht, dass es so groß/klein (oder viel/wenig) ist".

Ähnlich pervers ist das Thema Genauigkeit einer Zahl: viele denken, dass es keinesfalls genügt zu wissen, ob etwas um etwa 3 bis 4% gestiegen ist und dass es sehr viel besser ist, wenn man weiß, dass es 3,147% sind, um die etwas gestiegen ist.³³ Aber sie könnten mit Sicherheit wenig dazu sagen, was anders wäre, wenn es nur 3,146% wären. Manche können es schon deswegen nicht, weil sie Schwierigkeiten mit der Prozentrechnung haben. Ein Schlüsselerlebnis war für mich

³⁰ Diese Bemerkung mag sicher von einigen Professoren als ungerecht und überzogen angesehen werden. Fest steht aber (leider), dass in dieser Hinsicht bei Professoren Unterschiede wie zwischen Tag und Nacht bestehen können und dass es in unserem System in der Tat noch nicht gelungen ist, entsprechenden Auswüchsen zu begegnen. Man kann darüber streiten, wie häufig ein Professor ist, der bestenfalls einen Halbtagsjob macht. Aber, dass es ihn gibt kann man leider nicht bestreiten, und das macht es auch aus, dass man mit den obigen Argumenten viele Nichtprofessoren davon überzeugen kann, dass der Vergleich mit den USA unsinnig ist.

³¹ Ähnlich wurde neulich argumentiert: dass die Bachelor nicht alle arbeitslos sind (zumal es ja auch kaum noch andere Absolventen gibt), ist ein Beweis dafür, dass die Bologna Reform gut war.

³² M. J. Moroney, Facts from Figures, Penguin books. Das Buch war früher einmal recht bekannt. Es wurde erstmals 1951 aufgelegt und erlebte dann in den sechziger Jahren sehr viele Neuauflagen.

³³ Ob das "etwas", um das es hier geht, überhaupt so genau gemessen werden kann verliert sich bei entsprechenden Diskussionen meist ganz aus dem Gesichtsfeld.

eine Anhörung über Preise (insbesondere Mieten) vor dem Bundesverfassungsgericht. Für nicht wenige Journalisten war es schwer zu verstehen, dass mein Ergebnis, dass die Verbraucherpreise in München 20% über denen von Bayern liegen nicht auch bedeutet, dass Bayern um 20% billiger ist als München (sondern um 16,7% denn $1/1,2 - 1 = -0,167$).

Bei Prozentzahlen wird auch oft vergessen, dass auch 100% von nichts eben nichts ist. Wenn es heißt, die Wahrscheinlichkeit für eine unangenehme Sache sei um 60% gestiegen, dann macht das jedem Angst. Wenn man aber weiß, dass die Wahrscheinlichkeit vorher 0,001 war und jetzt auf 0,0016 gestiegen ist, sieht das wieder ganz anders aus.

Gerade weil der Umgang mit Zahlen bei vielen Menschen nicht gerade "sophisticated" ist, können zwei ganz gegensätzliche Taktiken gleichermaßen funktionieren:

- mehr Zahlen: man kann – vor allem wenn es in der Veranstaltung um D-Themen geht – als Kontrastprogramm zum Redner seine alternativen statistischen Zahlen auffahren (je mehr desto besser),³⁴ oder aber
- gar keine Zahlen: Sie kommen mit persönlichen Eindrücken, dass in der Praxis (Realität) alles ganz anders aussieht und deshalb die Zahlen des Redners gar nicht stimmen können. Sie brauchen keine Zahlen (je weniger desto besser), nur eine Position, bei der es glaubhaft ist, dass Sie über entsprechende persönlich Erfahrungen verfügen dürften.

Mit der Taktik "mehr Zahlen" liegen Sie ganz im main stream der Statistikenutzung: sich aus den bekannten Zahlen die aussuchen, die zur Unterstützung der eigenen Meinung etwas taugen. Für viele ist genau das auch der einzige Wert, den Statistik hat: sie bietet Zahlen zur Illustration von Standpunkten. Und weil es genügend viele Statistiken gibt

kann man sich immer etwas aussuchen. Der Redner hat das so gemacht, Sie können es genauso. Er hat ein schwarzes Bild gemalt, Sie malen ein weißes Bild. Für viele ist die Einseitigkeit der Auswahl schon ein erstes Zeichen für die inhärente Unredlichkeit der Statistik. Für andere liegt hierin – wie gesagt – der einzige Wert, den die Statistik im "normalen Leben" überhaupt hat.

Die umgekehrte Taktik "gar keine Zahlen" geht so: der Redner stellt fest, dass die Ärzte im Schnitt in den letzten Jahren erhebliche Einkommenssteigerungen erlebt haben. Sie halten dagegen, dass das nicht stimmen kann, weil Sie gesehen haben, dass in letzter Zeit immer mehr niedergelassene Ärzte ihre Praxis aufgeben und ins Krankenhaus gehen und dass Landarztpraxen praktisch unverkäuflich sind, wenn der Arzt in Ruhestand geht. Das können Sie natürlich dann gut bringen wenn Sie selber Arzt sind. Steuerberater mit vielen Ärzten unter den Mandanten geht auch, Studienrat schon weniger.

Wir wollen dies die "impressionistische Methode" nennen. Es ist wichtig, sich klar zu machen, dass es grundsätzlich zwei Arten gibt, die Dinge zu sehen und zu beschreiben

- die impressionistische und
- die statistische.

Die impressionistische ist anschaulich, emotional ansprechend und unmittelbar einleuchtend. Sie ist zwar nicht sehr systematisch, nicht methodisch reflektiert und ihre Verallgemeinerungsfähigkeit sollte nicht kritisch hinterfragt werden, aber dafür ist sie konkret, bunt und zumindest potentiell erlebbar. Von daher ist sie auch in den Augen vieler die sympathischere und uns näherliegende Herangehensweise. Sie wirkt auch dauerhafter. An eine Tabelle erinnert man sich nicht, aber wenn einem mit einer flammenden Rede über einen bewegenden Einzelfall ein Gefühl suggeriert wird, ist das ganz anders.

Die statistische Art der Betrachtung ist in fast allen Punkten das genaue Gegenteil. Sie ist intuitiv weniger überzeugend, auch wenn sie eher wissenschaftlich fundiert und daher respektabler erscheinen mag und den Nimbus der Exaktheit hat (allein deshalb weil mit

³⁴ Es ist ja in der Tat so, dass man im wirtschaftlichen und sozialen Bereich oft mit etwas Fleiß sowohl lauter "gute" (positive) als auch lauter schlechte Zahlen zusammenstellen kann.

Zahlen operiert wird, "15" ist exakter als "ein paar" und "15,2" ist noch exakter als 15).³⁵

In der Praxis von kontroversen Standpunkten sind dies nicht unbedingt Vorteile für die Statistik. Auch der Nimbus der Exaktheit von Zahlen kann schnell verfliegen, wenn man berücksichtigt, dass hinter jeder Zahl auch ein Mess- und Erhebungsproblem steht; man denke z.B. nur an die endlosen Diskussionen über die evtl. viel zu hohe oder zu niedrige Zahl der "Arbeitslosen".

Sie können mit der nötigen Empörung in der Stimme erklären, dass die soziale Kälte in den letzten Jahren (in der Zeit der Regierung von X) enorm (besser: in einem unerträglichen Maße) zugenommen habe. Keiner von den vielen Menschen, die Ihnen Beifall spenden, wird auch nur eine Sekunde darüber nachdenken, wie man so eine Behauptung überhaupt "beweisen" kann. Es wird auch keiner von Ihnen erwarten, dass Sie selbst über so etwas länger nachgedacht haben. Trotzdem haben alle geklatscht. Niemand wird sie fragen, wie kalt es damals war und wie kalt es jetzt ist und wie Sie so genau wissen wollen, dass es damals "wärmer" war. Bei der impressionistischen Masche verlangt keiner einen Beweis und Sie haben trotzdem ganz ohne jede Mühe gewonnen.

Wenn Sie dagegen sagen, die Ausgaben für X seien um 16% gekürzt worden (wie unerhört!) wird seltener geklatscht, vor allem wenn Sie dies nüchterner und mit weniger Empörung in der Stimme oder mit Verweis auf eine wissenschaftliche Studie sagen. Und es heißt dann gleich: stimmt nicht, es sind nur 3% oder in Wahrheit sind sie sogar gestiegen, und zwar um 9,3% (natürlich kann keiner in dem Augenblick auch nur eine der genannten Zahlen überprüfen), oder es heißt: woher haben Sie diese Zahlen? Oder: das sind doch die Zahlen von Y, die nicht sehr seriös sind (beweisen Sie mal das Gegenteil in der Hitze des Gefechts, wo sich ohnehin keiner für die Methoden von Y interessiert!).

³⁵ Es ist eine Eigentümlichkeit des Rechnens, dass i.d.R. (wenn man nicht gerade durch 0 dividiert) ein Zahlenergebnis dabei herauskommt. Es ist aber nicht gesagt, dass es auch ein "sinnvolles" Ergebnis ist.

Kurz gesagt: bei der statistischen statt impressionistischen Masche verlangt plötzlich jeder einen Beweis. Sie haben die größte Mühe und am Ende mitnichten gewonnen. Bei der impressionistischen Masche können sie, wie gesagt, die tollsten Sachen behaupten ohne sie beweisen zu müssen. Es ist alles viel einfacher und klappt deutlich besser.

8. Was tun, wenn jemand so argumentiert, dass er gegen Statistiken immun ist?

Gerade dort, wo in der Sache selbst enorm viel Sprengstoff stecken könnte, ist es erfolgversprechend, die impressionistische Methode auf die Spitze zu treiben. Die Euro-Rettung ist ein schönes Beispiel hierfür. In der richtigen Annahme, dass die meisten Menschen nicht gerne zahlen, schon gar nicht sehr viel Geld zahlen, oder zusehen wollen, wie sich ihre Ersparnisse allmählich in Nichts auflösen, wird man als Politiker nicht so gerne zugeben, dass wir in der Zwickmühle sind und quasi erpresst werden. Konkrete Zahlen, was uns die "Rettung" kostet, wird man selten hören. Komisch, man sagt nicht, was die Rettung kostet, weiß aber ganz genau, dass die Nicht-Rettung sehr viel mehr kosten würde.

Statt Zahlen zu präsentieren heißt es oft, dass Deutschland vom Euro am meisten profitiert habe (was wohl heißen soll, dass Deutschland deswegen auch für den Erhalt des Euros am meisten Opfer bringen muss). Das ist eine Behauptung, die geradezu gebetsmühlenhaft immer wieder wiederholt wird, obgleich niemand sagen kann, wo sie jemals schlüssig bewiesen worden ist. Das kann man schon daran erkennen, dass es nie heißt: wie dies "die X-Studie", oder "wie es das Y-Institut" gezeigt hat.³⁶ Es ist auch bezeichnend, dass

³⁶ Gemessen daran, dass man das Argument, dass Deutschland angeblich am meisten profitiert hat, schon bis zum Überdruß gehört hat, ist es schon sehr erstaunlich und bezeichnend, dass man bisher nie von jemand gehört hat, der dies bewiesen hat. Man sollte vielleicht in solchen Fällen mal danach fragen, ob es überhaupt eine Studie X oder die Arbeit eines Instituts Y gibt. Profis, wie es Politiker in diesem Punkte ja notwendig sind, können jedoch auch solche kleingeistige Fragen leicht abschütteln. Zur Not haben sie die Frage einfach akustisch irgendwie nicht verstanden.

noch nicht einmal gesagt wird, an welchem Kriterium (Außenhandel, Wachstum, Zinssätze) man das festmachen möchte. Dabei sagt einem schon der gesunde Menschenverstand, dass es sehr unwahrscheinlich ist, dass Deutschland in *jeder* Hinsicht vom Euro mehr als alle anderen Länder profitiert hat.

Es ist bemerkenswert, dass – gerade weil es eine Behauptung ist, die viele Fragen offen lässt – man damit rechnen kann, dass es die meisten Menschen im Publikum auch gar nicht so genau wissen wollen; vielleicht auch weil sie hier eine schwierige Materie vermuten, in der sich nur die Experten auskennen.

Dabei ist eine solche Strategie der totalen Immunisierung einer Behauptung gegen jedes empirisches (statistisches) Gegenargument auch eine Chance, zumindest eine Chance, ein paar Fragen zu stellen, die wohl allerdings nichts ausrichten werden.

Sie können ja z.B. fragen wie viele Milliarden Euro, die Deutschland bekommen hat, per Saldo auf das Konto des Euros gegangen sind. Und wenn dann keine konkrete Zahl kommt, können Sie auch weiterfragen, wie viele Milliarden es demgegenüber in Frankreich oder Griechenland waren. Wenn es heißt, man könne das alles nicht so genau sagen, können Sie den Redner auch danach fragen, woher er trotzdem so genau weiß, dass es mehr ist, als wenn der Euro nicht gekommen wäre, und warum es gerade in Deutschland mehr ist als z.B. in Italien, was man durch den Euro mehr bekommen hat.

Man könnte hier beliebig weiter ins Detail gehen, z.B. darauf hinweisen, dass die Handelsströme innerhalb der Euro-Zone schon vor dem Euro für den deutschen Außenhandel ein ähnlich großes Gewicht hatten und dass inzwischen das Gewicht von Frankreich, Italien, Spanien, Portugal und Irland zurückgegangen ist, während das von China erheblich gestiegen ist und jetzt nur noch knapp unter dem von Frankreich (9%) liegt.³⁷

³⁷ Deutsche Bundesbank, Neuberechnung der Gewichte für die Indikatoren der preislichen Wettbewerbsfähigkeit der deutschen Wirtschaft, Monatsbericht Februar 2012. "Das "Gewicht Griechenlands blieb unverändert, ist mit knapp ½% aber recht ge-

Es ist aber leider sehr die Frage ob sich ein Redner, den Sie so attackieren auf ein solches Spiel einlassen wird. Viel wahrscheinlicher ist es, dass er in puncto Emotionen noch einen drauflegen wird, damit sich jeder, der sich überhaupt noch traut, mit Statistiken zu kommen eigentlich ein schlechtes Gewissen haben muss und sich als kleingeistiger Pfennigfuchser vorkommen muss.

Er wird dann sagen, dass der Euro weit mehr ist als eine Währung und "um jeden Preis" zu retten ist. Solange er "nur" eine Währung war, konnte man noch eine gewisse fachliche Zuständigkeit von Ökonomen vermuten, jetzt aber, wo er weit mehr ist als eine Währung, kann natürlich die Stimme des Ökonomen nicht mehr zählen, als die des Bauingenieurs oder Kardiologen und natürlich die des Theologen.³⁸ In der Sache ist jetzt jeder mindestens so kompetent wie der Professor der Ökonomie, der nun auch spielend zu einer Lachnummer gemacht werden kann, zu einem lästigen, weltfremden Wadenbeißer, der die Tragweite, um die es hier geht, nicht begriffen hat.³⁹ Die Aussage

"Politik ist mehr als Zahlenwerk und Statistik. Europa ist mehr als ein Binnenmarkt und eine gemeinsame Währung. Europa ist ... eine Werte- und Schicksalsgemeinschaft"

von Guido Westerwelle im März 2013⁴⁰ klingt harmlos und trivial (natürlich ist Politik mehr; das bestreitet ja auch niemand), kann aber auch bedeuten: wenn es hart auf

ring." (Genauer gesagt waren die Gewichte bis 1997 0,47% und ab 2007 0,44% während das Gewicht von China von 2,67% auf 8,67% gestiegen ist).

³⁸ Es fragt sich natürlich auch, warum dies als Argument für die Rettung einer *Währung* zu erwähnen noch nötig ist, wenn wir wirklich als Deutsche vom Euro mehr als alle anderen Länder profitieren.

³⁹ Nach dem Motto "der Professor da aus Heidelberg", womit der damalige Bundeskanzler Schröder im Bundestagswahlkampf 2005 den Staats- und Steuerrechtler Paul Kirchhof, der 1987 - 1999 Richter im Bundesverfassungsgericht war, lächerlich machte. Das zeigt, dass in Deutschland, wo gerne geglaubt wird, dass der Praktiker dem Theoretiker überlegen ist, leicht auf diese Art Stimmung gemacht werden kann.

⁴⁰ Interview in Wirtschaftswoche 14/2013, S. 26. Wem obige Ausführungen zur Euro-Rettungspolitik zu pessimistisch erscheinen, der lese mal zum Vergleich "Europa am Abgrund – Deutschland im Abseits?" in dem gleichen Heft der Wirtschaftswoche.

hart kommt, interessieren mich die Zahlen (und die Kosten) nicht.

Man könnte jetzt vielleicht denken, dass man den Spieß umdrehen könnte und sich auch als Diskussionsredner dieser Taktik des Aushebelns der Statistik im Interesse von etwas Höherem bedienen kann. Aber das dürfte kaum gelingen, denn die Taktik steht und fällt damit, dass man als Redner den Bonus eines erfahrenen Politikers in hohen und höchsten Staatsämtern hat. Einem jüngeren Nobody als Diskutant wird man es nicht abnehmen, wenn er versucht, einem Thema eine welt- oder heilsgeschichtliche Perspektive zu verleihen.

Damit ist jetzt ein Punkt erreicht, bei dem genau das eingetreten ist, was zu verhindern (oder zumindest zu erschweren) Aufgabe der Statistik ist (weshalb die Statistik quasi überhaupt erst "erfunden" wurde):

- Statt Zahlen haben wir "Autoritätsbeweise" (es ist etwas richtig, weil es der X, bzw. die Y gesagt hat) und
- bei entsprechender moralischer Überhöhung ("Schicksalsgemeinschaft") ist jeder wissenschaftlichen Einwand (oder gar eine Statistik) kleingeistiges Meckern, das nur lästig und würdelos ist.

Um keine Missverständnisse aufkommen zu lassen: Bei dem Beispiel mit dem Euro ging es uns nicht um Meinungen pro und contra, sondern nur darum, zu zeigen, woran man erkennt, dass etwas in Richtung "Immunität gegen Zahlen" geht (oder gehen könnte).

Fazit: Wenn es um Glaubensbekenntnisse geht, kann man mit Statistik nichts mehr ausrichten. Das zeigt aber auch, dass wer immun ist gegen Zahlen und in einer Auseinandersetzung gegen Statistik auf "impressionistische" Art "siegt", tut dies im Grunde auf zutiefst unwissenschaftliche Art. Solange man noch über Zahlen spricht und nicht über "Schicksalsgemeinschaft" und dergleichen kann man noch rational argumentieren und offen sein für Kritik. So lange muss sich auch niemand schämen, wenn er anderer Meinung ist, oder Angst haben, als kleingeistiger Erbsenzähler verspottet zu werden, der die großen Zusammenhänge nicht begreift.

Immunität gegen Zahlen heißt also letzten Endes, die Ebene der Statistik, und überhaupt die der Wissenschaft, zu verlassen und das ist nicht mehr unser Thema.

9. Scheinkorrelation und Milchmädchenrechnungen

Zurück zu unserem Thema. Ganz anders ist die Situation, wenn ein Redner etwas vorträgt, was sowohl hinsichtlich der Daten, als auch der Raffinesse der angewendeten Methoden in der Tat wenig angreifbar ist. Auch hier hat man keine guten Karten, aber es ist weniger hoffnungslos, als wenn es um Glauben, Moral, und die Weltgeschichte geht.

Wenn man es bei einer solchen Rede auf der statistisch methodischen Schiene versucht, kann es schwierig werden und der Erfolg ist beim Publikum, das bei sehr technischen Diskussionen meist schnell ungeduldig wird und bloße Spitzfindigkeiten vermutet, alles andere als sicher. Die Gefahr ist auch groß, dass man sehr viel Mühe aufwenden muss, was einem dann aber nicht gedankt wird.

Man muss viel Mühe aufwenden, weil der "Gegner", offensichtlich gut bewandert ist (sonst wäre ja auch der Vortrag nicht so gut gewesen) und vielleicht auch auf mögliche Einwände gefasst ist. Mehr noch: Man muss damit rechnen, dass es dem Publikum nicht klar wird, warum Ihr Einwand ein so *gravierendes* Problem betrifft. Das wird dann meist schlecht für Sie ausgehen. Denn der Gegner kann dann nämlich auch sagen, er habe das Problem auch gesehen, wüsste aber nicht wie man ihm beikommen könne, und er kann dann sogar Sie fragen, ob Sie es nicht wüssten. Er kann sich das erlauben, weil er ja auch spürt, dass Ihr "Problem" vom Publikum nicht so richtig verstanden wurde.

Fachliche Einwände können also Schwierigkeiten bereiten. Einwände, die in einer solchen Situation vielleicht noch am ehesten sinnvoll sein könnten, gehen oft in eine der beiden folgenden Richtungen

- hier könnte eine "Scheinkorrelation" vorliegen und "umgekehrt wird ein Schuh daraus" (was die Kausalität betrifft), oder

- das läuft auf eine "Milchmädchenrechnung" hinaus und die Realität ist komplizierter.⁴¹

Kennzeichen einer Scheinkorrelation zwischen zwei Variablen X und Y ist, dass X und Y korreliert sind, obgleich keine (direkte) Kausalbeziehung besteht. Was demgegenüber eine Milchmädchenrechnung ausmacht ist, dass man etwas als unkorreliert vorausgesetzt, obgleich es kausal verknüpft ist.

Das eine ist also in gewisser Weise das Gegenteil vom anderen.

Scheinkorrelation entsteht, wenn beide Variablen, X und Y jeweils von einer dritten Größe Z abhängen. Oft ist Z nur einfach die Zeit, d.h. die Größen X und Y korrelieren weil sie einen gemeinsamen Trend haben. Störche (X) korrelieren mit Geburten (Y), auch wenn hier natürlich keine Kausalität vorliegt, also X nicht die Ursache für Y ist, weil es (im längeren Zeitraum gesehen) sowohl immer weniger Störche als auch immer weniger Geburten gibt (beide also einen fallenden Trend haben).

Eine Milchmädchenrechnung liegt vor, wenn eine Variable X gedanklich verändert wird (z.B. angenommen wird, sie steigt) und dabei gleichzeitig eine andere Variable Y als konstant angenommen wird, obgleich Y vielleicht sinken könnte, wenn X steigt; wenn also ein Zusammenhang ignoriert wird. Anzunehmen, dass der Umsatz steigt, wenn man den Preis (X) anhebt, könnte eine Milchmädchenrechnung sein, weil es ja möglich ist, dass die Menge (Y) bei steigendem Preis abnimmt und man am Ende sogar weniger Umsatz macht.

Eine Milchmädchenrechnung kann es auch sein, aus einem Zusammenhang zwischen X und Y zu folgern, wie groß X sein muss, damit Y eine bestimmte Größe hat. Es wird argumentiert, dass man bei Tempo 30 weni-

ger Autounfälle und Verletzte hat als bei Tempo 50. Aus dem gleichen Zusammenhang dürfte aber auch folgen, dass man bei Tempo 20 noch weniger und bei Tempo Null gar keine Autounfälle mehr hätte. Mit X = 30 zu operieren wirkt überzeugend, aber bei X = 0 hört irgendwie der Spaß auf,⁴² weil dann ja auch der letzte Hörer merkt, dass das ja eigentlich bedeutet, dass die Autos überhaupt nicht mehr fahren dürfen und es sollte klar sein, dass es dann ja auch keine Autounfälle mehr gibt.

Das alles klingt einfach, aber es ist im Einzelfall nicht immer einfach darzulegen, dass Y nichts mit X zu tun hat (so etwas ist nicht immer so allgemein bekannt, wie im Falle von Störchen X und Geburten Y) oder dass sich X nicht verändern kann ohne dass sich nicht auch Y verändert.⁴³

Wenn einem in dieser Richtung nichts für jeden erkennbar Eklatantes einfällt, sollte man es besser bleiben lassen, zumal man ja auch – wie gesagt – einen solchen Einwand schnell als spitzfindig und mehr oder weniger nebensächlich abtun kann, als etwas, was vielleicht ein Problem sein *könnte*, aber naturgemäß in Kauf zu nehmen ist.

Es ist deshalb vielleicht eine ganz andere Taktik angebracht: man sollte nicht sagen was *falsch* ist (denn das ist ja gerade bei einem annahmegemäß sehr kompetenten Redner schwierig), man sollte stattdessen sagen, was noch *fehlt*, um noch besser zu sein. Das klappt fast immer und wenn man gut reden kann, kann man auch versuchen das Fehlen zu dramatisieren. So kann man z.B. fast immer plausibel machen, dass bei den Merkmalen noch viel weiter differenziert werden müsste und/oder, dass man bei den Befragten weitere wichtige Merkmale erheben müsste, damit das alles erst "aussagefähig" wird.

⁴¹ Dass etwas "graue Theorie" ist und die Realität viel komplizierter ist, fällt in Deutschland fast immer auf fruchtbaren Boden. Aber es ist – wenn überhaupt – nur ein halbes Argument. Denn eigentlich müsste dann nämlich etwas über die angeblich *ganz andere* Realität gesagt werden, aber dazu fällt vielen Debatte-rednern schon nicht mehr viel ein.

⁴² Man könnte ja auch fragen, warum gerade Tempo 30, und nicht Tempo 27, wenn es Tempo 50 nicht sein soll. Hier tritt wieder eine Art Zahlenpsychologie auf, wonach ja auch 4,95€ ein ganz anderer Preis ist als 5,05€ ist auch wenn nur 10 Cent dazwischen liegen.

⁴³ Im Fall von X (Preis) und Y (Menge) kann Y ja auch trotz steigendem X konstant sein (preisunelastische Nachfrage), und dann ist die erfolgreiche Preiserhöhung auch keine Milchmädchenrechnung.

Beides ist geeignet, den Gegner in Verlegenheit zu bringen, weil es darauf hinausläuft von ihm eine Erhebung zu verlangen, die er nicht hat. Und er kann sich auch nicht gut äußern über etwas, was in der Erhebung nicht erfragt wurde.

Dabei interessiert es nur am Rande, dass Fragen, die den Befragten unbedingt noch gestellt werden sollten, von diesen vielleicht gar nicht so einfach beantwortet werden können.

10. Die Kunst, eine Statistik gezielt zu überfordern

Bei einer Vortragsveranstaltung im November 2012, die mir die Inspiration zu den hier vorliegenden Ausführungen gab, ging es um die wirtschaftliche Lage der Kassenärzte in Deutschland. Die Fronten waren damit klar, auf der einen Seite die Ärzte, die klagten, sich bei zu wenig Geld für die Patienten aufzuopfern und auf der anderen Seite die Krankenkassen, die sich (auch wieder nur das Wohl der Patienten, bzw. Beitragszahler im Auge) darum sorgen, dass die Kosten aus dem Ruder laufen (Krankenkassen sind ja auch nicht "systemrelevant", also nicht notfalls vom Steuerzahler zu "retten").

Seitens der Ärzte wurden Berechnungen vorgelegt, bei denen quasi die Mehreinnahmen der Arztpraxen, die durch Behandlung von Privatpatienten entstanden sind "herausgerechnet" wurden. Es wurde errechnet, was die Praxen verdient *hätten*, wenn sie nur Patienten der Gesetzlichen Krankenversicherung (GKV) behandelt *hätten*. Solche "Hätte-Rechnungen" haben naturgemäß ein fiktives Element und sie stehen unter dem Verdacht, dass sich damit die Ärzte künstlich arm rechnen, um höhere Honorare durchzusetzen, und Betrachtungen dieser Art können auf gar keinen Fall Gegenstand der amtlichen Statistik, sondern immer nur der privaten Statistik sein.

Seitens der GKV wurde dann auch gekontert, dass man solche Berechnungen nicht akzeptieren könne, solange dabei nicht auch unterschiedliche Kosten bei der Behandlung von

Privatpatienten im Vergleich zu GKV Patienten berücksichtigen würden.⁴⁴

Aber das ist dann genau die in der Überschrift angesprochene "gezielte Überforderung". Wie kann man erwarten, dass ein Arzt im Rahmen einer Befragung sinnvoll Auskünfte darüber geben kann, wie viel mehr die Privat- als die Kassenpatienten an Miet- und Personalkosten usw. verursachen?

Abgesehen davon müssten sich die Vergleiche der Honorare und Kosten ja streng genommen auch auf wirklich vergleichbare *Behandlungen* beziehen. Hinzu kommt, dass es schwierig sein dürfte, hier bei Kosten und Erträgen die Preis- und Mengenkomponente der Kosten überhaupt begrifflich korrekt fassen.

Die Taktik der Überforderung klappt also in der Regel sehr gut, zumindest solange die Forderungen an eine Befragung nicht ganz offensichtlich unsinnig sind. Genau genommen werden allerdings gar nicht so selten Befragungen veranstaltet, bei denen von den Befragten Antworten verlangt werden, die auch nicht immer sehr fundiert sein *können* (wenn z.B. nach den prominentesten Vertretern eines Faches gefragt wird oder danach, auf welche Hochschule man seine Kinder schicken möchte). Unterschiede in einem auf solche Fragen aufgebauten Ranking werden dann aber gleichwohl nicht selten auch wie auf zwei Nachkommastellen exakte naturwissenschaftliche Messungen behandelt.

Darauf hinzuweisen, dass eine Erhebung nicht alles leistet, was man sich von ihr wünschen könnte ist eine relativ gefahrlose Taktik. Sie müssen auch hier nicht viel eigene Kompetenz und Kreativität aufwenden, weil Sie darauf bauen können, dass sich auch andere noch nicht viele Gedanken darüber gemacht haben, wie man bei einer Erhebung

⁴⁴ Mit dem Argument wurde auch begründet, warum man die amtlichen Zahlen des Statistischen Bundesamts für besser hält als die der Ärzteschaft. Dabei wird in den amtlichen Zahlen der Unterschied PKV/GKV weder auf der Ertrags- noch auf der Kostenseite thematisiert. Die amtlichen Zahlen leisten also erst recht nicht das, was von den Zahlen der Ärzteschaft verlangt wird. Sie können also nicht *mit dem obigen Argument* besser sein als die der Ärzteschaft.

das alles konkret umsetzen sollte, was Sie von ihr fordern.

Bei der Überforderungs-Taktik ist also kaum zu befürchten, dass einem die Unmöglichkeit der Forderung entgegengehalten wird.⁴⁵ oder ihre Unverhältnismäßigkeit moniert, d.h. Sie werden kaum zu hören bekommen, dass Sie mit Ihrer Ablehnung einer Statistik wegen angeblich unzureichender Datenbasis das Kind mit dem Bade ausschütten.

Ebenfalls gefahrlos und universell einsetzbar ist die folgende Taktik, nämlich nach dem "Sinn" einer Zahl zu fragen.

11. Den Zahlen einen Sinn geben oder nehmen

Es ist klar, dass man bei einer Zahl nach deren "Sinn" oder ihrer "Aussagefähigkeit" fragt, aber ebenso klar ist es auch, dass es schwer ist zu sagen, was generell "Sinn" und "Aussagefähigkeit" ausmacht und es ist auch nicht leicht einzusehen, dass eine Zahl grundsätzlich nur den Sinn haben kann, den man ihr durch Interpretation gibt.

Für viele Menschen hat etwas dann "Sinn", wenn es zu Ihren Vorstellungen (oder Vorurteilen) passt und/oder sich "erklären" lässt. Wenn ihr Angriff auf die "Sinnhaftigkeit" letztlich impliziert, dass die Leute ihre Vorurteile aufgeben sollten, kann es schwierig werden. Wenn es aber nicht darum geht, sondern "nur" darum, sich etwas zu "erklären", sind die Menschen oft ganz überraschend anspruchslos.

Dafür, dass es ein Bedürfnis gibt nach dem "Sinn" einer Zahl, nach einer "Erklärung" – auch wenn diese bei genauem Hinsehen schon geradezu peinlich ist, weil sie so läppisch ist – scheint mir das Phänomen der "Gestern-Erklärer" des Börsengeschehens

⁴⁵ Eine eher peinliche Ausnahme ist allerdings, wenn Sie etwas verlangen, was schon aus rein rechtlichen Gründen gar nicht möglich ist. Bei der genannten Vortragsveranstaltung, wurde einem Redner aus der "Fraktion" der Krankenkassen u.a. eine (nach wie vor nichtamtliche) Befragung von Ärzten gefordert, die nicht mehr freiwillig ist. Man hätte wissen müssen, dass nur die amtliche Statistik eine Erhebung mit Auskunftspflicht durchführen darf.

ein Zeichen zu sein. Dieselben Leute, die es jetzt auf einmal so genau wissen, dass es so kommen musste, weil man z.B. eine "Blase", Gewinnmitnahmen oder eine "Seitwärtsbewegung" hatte, haben gestern noch nicht gesehen, dass es so kommen wird.

Dass sie für ihre intellektuell eher denkbar ärmlichen Auslassungen (im Nachhinein ist man ja immer schlauer) überhaupt so viel Gehör finden, spricht dafür, dass man sich aus Zahlen gerne "einen Reim machen" möchte und dass uns quasi instinktiv Zahlen ohne "Reim" (oder anders gesagt, ohne "Sinn") suspekt erscheinen.⁴⁶

Sinn geben und Sinn nehmen liegt erstaunlich eng beieinander. Mit einer "Geschichte" (oder "story") wird klar, was die Zahlen "bedeuten". Die story kann aber auch irgendwie zu klar sein, und eine innere Notwendigkeit heraufbeschwören, so dass sie eine "neue" Erkenntnis gänzlich entwertet: "Das sagt einem ja schon der gesunde Menschenverstand" ist ähnlich tödlich, wie "Das hat ja bereits der X in seiner Arbeit Y gezeigt".

Natürlich ist es letztlich ein entscheidendes Kriterium (wenn nicht *das* Kriterium), ob das Ergebnis einer statistischen Erhebung oder Analyse "sinnvoll" und "interessant" ist und man kann Abelson⁴⁷ nur zustimmen, wenn er schreibt

"Data analysis should not be pointlessly formal. It should make an interesting claim; it should tell a story that an informed audience will care about, and it should do so by intelligent interpretation of appropriate evidence from empirical measurements or observations"

und wenn er klagt, dass die Lehre in Statistik an den Hochschulen leider ganz anders abläuft, allein schon deshalb weil die meisten

⁴⁶ Das gilt natürlich nicht für Zahlen aus der Naturwissenschaft, die uns gefühlsmäßig nicht tangieren. Dass wir 23 Chromosomen haben, und nicht 25, wie der Gorilla, dürfte nicht ganz einfach zu erklären sein. Es macht uns aber die 23 nicht suspekt. Es ist sicher auch schwer, Gründe dafür anzugeben, warum die Erdachse in einem Winkel von $66,56^\circ$ ($90 - 23,44$) und nicht $68,72^\circ$ zur Erdumlaufbahn steht. Es ist eben so, dass es Zahlen mit und Zahlen ohne Emotionen gibt, und das sollte man als Statistiker berücksichtigen.
⁴⁷ Robert P. Abelson (siehe Fußnote 25), S. 70.

Studenten mit den Formeln und dem Nachvollziehen der bloßen Rechnerei schon so vollauf beschäftigt sind, dass für "intelligent interpretation" nicht mehr viel übrig bleibt. Kein Wunder, dass für sie dann Statistik "pointless" ist und auch schnell wieder vergessen wird, nachdem man sich widerwillig damit beschäftigen musste.

Aber wann ist die Interpretation "intelligent", die story interessant und die Zahl "sinnvoll"? Das Problem ist, dass alle Versuche, so etwas zu definieren, unbefriedigend sind und mehr oder weniger auf eine Auflistung von ähnlichen Begriffen hinauslaufen,⁴⁸ etwa anschaulich, überzeugend, einleuchtend, plausibel (dem "gesunden Menschenverstand" nicht widersprechend), aussagefähig usw.

Man könnte vielleicht auch sagen eine statistische Berechnung ist dann "sinnvoll", wenn sie unter gleichen Umständen gleiche Ergebnisse liefert und unter ungleichen nicht gleiche (was i.d.R. mit "Reliabilität" bezeichnet wird). Das mag bei psychologischen Versuchen (Tests) funktionieren, aber wo haben wir bei den typischen großen Befragungen (von vielen Tausend Haushalten bzw. Unternehmen) im Rahmen der Wirtschaftsstatistik zwei Befragungen zur gleichen Sache und wann sind dabei die Bedingungen wirklich "gleich"?

Die Taktik, dem Gegner vorzuwerfen, dass seine statistischen Ausführungen keinen "Sinn" machen (oder wenig "interessant" sind) ist viel vager und damit keineswegs so narrensicher⁴⁹ wie die Einwände "nicht objektiv", "nicht repräsentativ" oder "nicht vergleichbar". Es kann sein, dass Sie ein zustimmendes Nicken ernten, es kann aber auch sein, dass Ihre Worte verpuffen.

⁴⁸ Bei Abelson findet man z.B. Begriffe wie: degree of comprehensible detail in which conclusions are phrased; methodological soundness, importance of the issue (with potential to change beliefs), effective research narrative, theoretical coherence etc.

⁴⁹ Es dürfte hier kaum möglich sein, die Reaktion des Publikums sicher voraussehen zu können. Man kann wohl nur darauf vertrauen, dass andere ähnlich denken (besser: dass sie die Dinge etwas ähnlich *empfinden*) wie man selbst, wenn man nicht gerade ein Sonderling ist und mit beiden Beinen im Leben steht.

12. Statistik als Streit mit rationalen Argumenten und Fakten verstehen

Der unsichere Ausgang von Diskussionen über den "Sinn" macht nur noch einmal deutlich, was Statistik eigentlich ist, nämlich ein Versuch, mit Vernunft und Fakten zu überzeugen. Es ist ein Spiel, das für Sie auch trotz guter Vorbereitung und trotz guter Argumente schiefgehen kann oder auch so ausgehen kann, dass es keinen Gewinner gibt.

Es ist ein Spiel, bei dem man verschiedene Taktiken versuchen kann, aber wenn man Pech hat können auch alle schiefgehen.

Wenn die Zahlen, die in einem Vortrag präsentiert werden, so sind, dass sie schwer angreifbar sein dürften, können Sie versuchen, die Methode der Zahlengewinnung zu kritisieren.⁵⁰ Wenn auch die Methode kaum zu kritisieren ist, können Sie versuchen, zu argumentieren, dass die Ergebnisse, also die Zahlen, im Lichte der bisher üblichen Vorstellungen doch etwas eigenartig, also "unplausibel" (z.B. zu hoch/niedrig) zu sein scheinen. Und wenn alles nicht so richtig fruchtet, kann man es damit versuchen, dass man die "Sinnhaftigkeit" des ganzen Unternehmens in Frage stellt, dass z.B.

- alles doch recht wenig sei, was man damit so viel Aufwand herausgefunden hat (oder glaubt, herausgefunden zu haben),
- nicht so richtig zu erklären sei angesichts dessen, was man bisher zu denken gewohnt war, oder umgekehrt, dass das
- alles genau genommen ja nicht wirklich neu sei.

Es sollte nicht zu schwierig sein, einen dieser Punkte zu kommunizieren, wenn es Ihnen denn gelingt, dabei auch die Aufmerksamkeit des Publikums zu bekommen. Denn nicht der Kritiker einer Statistik, sondern der investigator hat die Beweislast (burden of proof); der Kritiker kann vage und "free

⁵⁰ Auf so etwas kann der Gegner aber meist mit Recht und mit allgemeiner Zustimmung des Auditoriums (das so etwas oft gar nicht beurteilen kann) antworten, dass er/sie es in puncto Methoden so gemacht hat, wie alle Anwender der Statistik es auch machen. Damit gehen ihre Bemühungen dann auch schnell ins Leere.

wheeling" sein, aber der investigator nicht.⁵¹ Man kann als Kritiker sogar damit Erfolg haben, dass man dem Gegner (investigator) vorwirft, nur eine "rein formale" Betrachtung abgeliefert zu haben.⁵² Das kann v.a. bei Wissenschaftlern (insbesondere Statistikern) funktionieren, wenn der Vortrag nicht weit über das rein Statistische hinausgegangen ist. Solche Leute können lang (und langweilig) über Methoden und auch über "Ergebnisse" reden, aber sie sind oft nicht gut darin, deutlich zu machen, was uns das alles konkret sagen soll.

Je mehr Statistik von immer mehr Leuten angewendet wird (weil es quasi zum guten Ton gehört, in einem Vortrag in dieser Hinsicht etwas zu bieten), desto häufiger dürfte so etwas vorkommen und das Arbeiten mit Statistik zu einem Ritual erstarren. Typisch für ein Ritual (wie es die obligatorischen Signifikanztests inzwischen sind) ist, dass es durch beständige Übung nicht mehr "hinterfragt" wird. Also kann auch immer mehr Unsinn und Unverstandenes auf diese Weise produziert werden, was vom Produzenten (und nicht nur von diesem) auch nicht mehr wirklich interpretiert werden kann. Der Vorwurf "rein formal" könnte also immer öfter berechtigt sein.

Ob es aber gelingt, mit einer Kritik auf dieser Schiene zu überzeugen, ist ganz offen. Sie können es ja auch nicht verhindern, dass es immer genug Menschen gibt, die auch einen mit oder ohne Zahlen vorgetragenen "Unsinn" glauben.

So bleibt als Quintessenz, dass Statistik ein "principled argument" (Abelson) ist, also ein rational geführter Streit darüber, was die "Realität" ist, und damit letztlich ein "Spiel", das man nicht zu verbissen spielen sollte, auch nicht überall spielen kann (siehe Abschn. 8) und notfalls auch mit Würde verlieren können muss,⁵³ zumal es auch nicht

immer mit Würde (oder Wissenschaft) zu tun hat, wenn man es "gewinnt" (siehe ebenfalls Abschn. 8).

⁵¹ Abelson S. 14.

⁵² Das läuft dann quasi auf eine Variante von "der Professor da aus Heidelberg" hinaus, die in Deutschland ja zu funktionieren scheint.

⁵³ Auch beim Thema von Abschn. 8, der Euro-Rettung, könnte es ja auch der Skeptiker sein, der am

Ende, wenn es "noch einmal gut gegangen ist" der Verlierer ist (was zu hoffen wäre).