

## 1 Monty Hall

Das ist vielleicht ein Gefühl, in Hunderten von Briefen als Spinner oder Dummkopf beschimpft zu werden!

Dabei hatte alles so harmlos angefangen.

An einem Samstag im Sommer saß ich abends spät im Garten, entkorkte eine Flasche und schlug den *Skeptical Inquirer* auf, mein Lieblingsblatt aus den USA : Wissenschaftler und Journalisten gehen darin den Behauptungen von Tischrückern, Gabelbiegern, Geistersehern und anderen Scharlatanen nach. Mich interessierte ein Artikel über die amerikanische Journalistin Marilyn vos Savant. Sie gilt als der Mensch mit dem höchsten Intelligenzquotienten der Welt, was immer das bedeuten mag. Mit der Lösung einer Denksportaufgabe in ihrer Kolumne «Fragen Sie Marilyn » hatte sie eine Lawine hämischer bis empörter Leserbriefe losgetreten. Die Lösung, vorgestellt in der Zeitschrift *Parade*, widersprach nämlich der Intuition ihrer Leserschaft, darunter viele Mathematiker.



FIG. 1: Vos Savant

interessierte ein Artikel über die amerikanische Journalistin Marilyn vos Savant. Sie gilt als der Mensch mit dem höchsten Intelligenzquotienten der Welt, was immer das bedeuten mag. Mit der Lösung einer Denksportaufgabe in ihrer Kolumne «Fragen Sie Marilyn » hatte sie eine

Lawine hämischer bis empörter Leserbriefe losgetreten. Die Lösung, vorgestellt in der Zeitschrift *Parade*, widersprach nämlich der Intuition ihrer Leserschaft, darunter viele Mathematiker.

Ein Leser hatte folgende Aufgabe gestellt :

Sie nehmen an einer Spielshow im Fernsehen teil, bei der Sie eine von drei verschlossenen Türen auswählen sollen. Hinter einer Tür wartet der Preis, ein Auto, hinter den beiden anderen stehen Ziegen. Sie zeigen auf eine Tür, sagen wir Nummer eins. Sie bleibt vorerst geschlossen. Der Moderator weiß, hinter welcher Tür sich das Auto befindet ; mit den Worten «Ich zeige Ihnen mal was » öffnet er eine andere Tür, zum Beispiel Nummer drei, und eine meckernde Ziege schaut ins Publikum. Er sagt : «Bleiben Sie bei Nummer eins, oder wählen Sie Nummer zwei ? »

Zwei Türen, hinter einer steckt der Gewinn. Also bleibt es sich gleich, welche gewählt wird, nicht wahr? *Falsch*, sagt die IQ-Weltmeisterin, *Nummer zwei hat bessere Chancen*.

Da war es : das *Ziegenproblem*.

Irgend jemand spinnt hier, dachte ich beim Lesen. Die schlaue Dame, ihre Leser oder alle zusammen und ich vielleicht auch.

Nun war eine Entscheidung fällig : den geruhsamen Sommerabend mit Riesling fortsetzen oder Kopfzerbrechen mit Schreiber und Papier. Ich wählte das leichtere Vergnügen.

Am nächsten Morgen fiel mich das Ziegenproblem schon wieder an. Anstatt mich aus den Träumen

sanft in den Tag zu leiten, ließ meine Phantasie Türen klappern, Ziegen meckern, Autos blinken. Erst unter der Dusche kam die Eingebung — die Frau hatte recht!

Das konnte ich nicht für mich behalten (Berufskrankheit). Ich setzte mich also hin (am Sonntagmorgen) und schrieb einen kleinen Artikel für die ZEIT, in dem ich das Ziegenproblem und dessen Lösung präsentierte. Am nächsten Tag fuhr ich in Urlaub.

Und so begrüßten mich die Leserzuschriften, als ich zurückkam : der verehrte Herr von Randow sei «wohl ins Sommerloch gestolpert» «jeder normalbegabte Zwölftkläßler» könne schließlich begreifen, daß Frau Savants Rat «typische Laienfehler» enthalte, «haarsträubender Unsinn» «Quatsch» und «Nonsens» «absurd» und «abstrus» sei. Es sei «traurig, daß die ZEIT so etwas überhaupt aufgreift». Die ganze Angelegenheit sei «peinlich» urteilte ein Mathematiker. Bestenfalls ein «Aprilscherz im Juli» schrieb ein Leser mitleidig, eher ein «Ärgernis» meinte ein anderer.

Die alles dies zu Papier brachten, waren zum großen Teil Akademiker, einige mit einschlägiger Ausbildung in Statistik : Prof. Dr.-Ing., Dr. sc. math., Dr. med., Dr. jur. usw. usf. Sie schrieben auf Institutsbriefbögen, legten seitenlange Beweise bei, es kam sogar Post aus den Niederlanden, aus Italien, aus Togo. Zustimmende Briefe blieben rar.

Die Leserbrief-Redaktion wählte drei Briefe aus, die mich kritisierten, und ließ sie unter der Überschrift «Verquere Logik» drucken. Das mochte ich nicht auf mir sitzen lassen und schrieb einen zweiten Artikel. Wieder nahm ich für Frau Savant Partei — und entfachte den zweiten Sturm. Mittlerweile hatte der SPIEGEL die Geschichte aufgegriffen, gab ebenfalls Frau Savant recht und bescherte sich die entsprechende Leserpost.

Das Ziegenproblem hielt offenbar viele Menschen in Atem. Feten platzten, und Ehepaare stritten sich. Professoren setzten ihre Assistenten an das Ziegenproblem, Mathe-Lehrer verwirrten ihre Schüler, Zeitungsredakteure erklärten sich gegenseitig für begriffsstutzig. [...]

In den USA war noch viel mehr los. Ich zitiere von Seite eins der *New York Times* (21.7.1991) :

Die Antwort, wonach die Mitspielerin die Tür wechseln solle, wurde in den Sitzungssälen der CIA und den Baracken der Golfkrieg-Piloten debattiert. Sie wurde von Mathematikern am Massachusetts Institute of Technology und von Programmierern am Los Alamos National Laboratory in New Mexico untersucht und in über tausend Schulklassen des Landes analysiert.

Marilyn vos Savant wurde unterdessen mit Spott überschüttet. «Unsere mathematische Fakultät hat herzlich über Sie gelacht» hänselte eine Professorin. «Es gibt schon genug mathematische Unwissenheit in diesem Land» beschwerte sich ein Akademiker bei der Zeitschrift *Parade*, «wir brauchen nicht den höchsten IQ der Welt, um diese Unwissenheit zu vertiefen. Schämen Sie sich!» Ein weiterer Leser vermerkte höhnisch: «Vielleicht haben Frauen eine andere Sicht mathematischer Probleme als Männer.»

Gero von Randow

## 2 Die fremde Welt der Zahlen

### 2.1 Innumeracy

Der Schock saß tief. Ein Skandal, ein Desaster, eine Katastrophe sei das verheerende Abschneiden der deutschen Schüler in der internationalen Pisa-Studie. In seltener Einigkeit prangert seither die ganze Nation das Versagen der heutigen Lehrer und Schulen an.

Manch einer der Lamentierer allerdings dürfte froh sein, dass er nicht selbst die Aufgaben für Neuntklässler lösen musste, von denen einige auch in der ZEIT veröffentlicht wurden. Denn mit Rechnen und abstraktem Denken sieht es in der Wissensgesellschaft am Standort Deutschland selbst bei Akademikern düster aus.

Hans-Hermann DUBBEN von der Universität Hamburg kann davon ein bitteres Lied singen. «Nicht nur



FIG. 2: Dubben

viele Medizinstudenten», stellte der Statistikdozent in Kursen und auf Kongressen immer wieder verblüfft fest, «auch etliche fertige Ärzte sind schon mit einem Dreisatz schwer am Anschlag.» So hat ein Großteil aller angehenden Mediziner mit Testfragen nach dem Muster: «Von 3000 Menschen haben 1,4 Prozent Krebs. Wie viele Personen sind demnach betroffen?» echte Schwierigkeiten. Und einige davon können solche Aufgaben auch nach minutenlangem Überlegen nicht allein lösen.

### 2.2 Wie verlässlich ist der Aidstest?

Bei angehenden Betriebswirtschaftlern sieht es nicht viel besser aus. Aus purer Neugier hatte ein BWL-Professor der Uni Leipzig seinen Studenten kürzlich zwei der Pisa-Aufgaben vorgelegt. Das Ergebnis war «eine Katastrophe», berichtet er. 43

Prozent der Erst- und Zweitsemester scheiterten an der einen, 34 Prozent an der anderen Aufgabe.

Peinlich ist das offenbar den wenigsten. Im Gegenteil, glaubt DUBBEN. «Ich habe manchmal den Eindruck, dass es zum guten Ton gehört, keine Ahnung von Mathematik zu haben.» Kein Wunder also, dass rund ein Drittel der Deutschen einer Emnid-Umfrage zufolge nicht in der Lage ist, die Angabe «40 Prozent» richtig zu deuten — ein Teil glaubt, dies entspreche einem Viertel, die anderen tippen auf «jeder 40.».

Doch dieser «mathematische Analphabetismus», im Englischen *innumeracy* genannt, kann fatale Folgen haben — für die eigene Gesundheit, die eigenen Rechte oder auch den Geldbeutel. Das zeigen etliche Beispiele von Justizirrtümern sowie Fehlentwicklungen in Medizin und Politik, die



FIG. 3: Gigerenzer

der Psychologieprofessor Gerd GIGERENZER vom Berliner Max-Planck-Institut für Bildungsforschung in seinem neuen Buch *Einmaleins der Skepsis* zusammengetragen hat.

Gerade Ärzte und ihre Patienten, so stellte GIGERENZER fest, unterliegen häufig der verhängnisvollen Illusion von absoluter Gewissheit. Sie sind sich oft nicht bewusst, dass viele moderne und vergleichsweise zuverlässige Diagnose- und Therapieverfahren weitaus weniger sicher sind, als sie auf den ersten Blick erscheinen.

Welcher Schaden dadurch entstehen kann, zeigt GIGERENZER am Schicksal einer jungen Frau in den USA, die sich bei einer Routineuntersuchung einem HIV-Test unterzog. Sie rechnete nicht damit, dass sie sich infiziert haben könnte. Schließlich war die alleinerziehende Mutter weder drogensüchtig, noch hatte sie wechselnde Partner. Doch einige Wochen später kam das Ergebnis: HIV-positiv. Danach ging es mit ihrem Leben nur noch bergab. Als die Diagnose in ihrer Firma bekannt wurde, verlor sie ihre Anstellung. Aus Angst, ihren Sohn anzustecken, küsste sie ihn nicht mehr. Und da nun alles ohnehin zu spät schien, verzichtete sie auch auf ein Kondom, als sie wenig später mit einem HIV-infizierten Mann schlief.

Durch einen Zufall wurde einige Monate später ein zweiter Test gemacht, und es stellte sich heraus: Die Frau war vollkommen gesund, der erste Test hatte schlichtweg versagt. Dass dies überhaupt passieren kann, hatte ihr vorher niemand gesagt.

Dabei sind solche Fehler in der Medizin keine Seltenheit. Und zwar selbst dann, wenn das Verfahren so wie die heutigen HIV-Tests Kranke und Gesunde

mit einer Wahrscheinlichkeit von 99,99 Prozent richtig erkennt. Es gibt also bei einem von 10 000 Gesunden fälschlich einen positiven Befund. Heißt das nun, dass man bei einem positiven HIV-Test mit 99,99-prozentiger Sicherheit davon ausgehen kann, das Aids-Virus zu haben?

### 2.3 Auch die Experten rechnen falsch

Keinesfalls. Entscheidend für die Einschätzung eines «positiven» Befundes ist die Verbreitung der Infektion. Ein Beispiel : Von 10 000 heterosexuellen Männern, die keiner Risikogruppe angehören, ist statistisch gesehen lediglich einer infiziert. Bei 10 000 Tests wird man also zwei HIV-Kranke ermitteln, dabei jedoch einen Mann fälschlich positiv testen — da ja im Durchschnitt mit einer Fehldiagnose zu rechnen ist. Das bedeutet aber : Jeder zweite positive HIV-Test bei Nicht-Risikogruppen ist ein Fehlalarm — die Wahrscheinlichkeit, bei einem positiven Test tatsächlich infiziert zu sein, beträgt nur 50 Prozent!

Aber das hat sich selbst unter denen, die es wissen müssten, noch nicht richtig herumgesprochen. Um aus erster Hand zu erfahren, wie professionelle Aids-Berater über die relevanten Wahrscheinlichkeiten informieren, ging einer von GIGERENZERS Studenten zu mehr als 20 öffentlichen Gesundheitsämtern im ganzen Bundesgebiet. Er ließ einen HIV-Test machen und fragte während der Beratung vor dem Test, was ein Positivbefund in seinem Fall (ohne spezielle Risikofaktoren) bedeuten würde. Fast alle Berater versicherten ihm voller Überzeugung, die Möglichkeit eines Irrtums sei völlig oder nahezu ausgeschlossen, weil der Test «zu 100 Prozent» oder «zu 99,9 Prozent» sicher sei. Kein Einziger nannte den richtigen Wert von rund 50 Prozent.

Die Zahlenblindheit ist ein globales Problem, stellt GIGERENZER fest, sei es bei der jüngsten Diskussion um das Mammografie-Screening, beim Rinderwahn oder beim aktuellen Nitrofen-Skandal. «Nur die Akteure wechseln.» BSE griff im Jahr 2000 in Großbritannien, Irland, Portugal, Frankreich und der Schweiz um sich, während die Bundesregierung Deutschland für BSE-frei erklärte. «Deutsches Rindfleisch ist sicher», wiederholten der Präsident des Bauernverbandes und der Landwirtschaftsminister unentwegt. Als man dann doch BSE-Tests an deutschen Rinderherden vornahm, zeigte sich das Phänomen — welch Wunder! — auch hierzulande. Die Regierung gestand schließlich ein, sich zu lange an die Illusion geklammert zu haben, deutsches Vieh sei von der Krankheit nicht betroffen. Doch jetzt begann die Täuschung von neuem. «Unser Rindfleisch ist getestet», verkündeten Supermärkte und Metz-

ger. Aber niemand erwähnte, dass bei diesen Tests zahlreiche Fehler auftreten.

Hans-Peter BECK-BORNHOLDT, wie DUBBEN Statistikexperte an der Uni Hamburg, sieht so-



FIG. 4: Beck-Bornholdt

gar ein weiteres Problem : Niemand weiß, ob die mehr als 150 «BSE-Fälle» unter den inzwischen fast drei Millionen getesteten Rindern tatsächlich krank waren. Denn um die Verlässlichkeit der Tests zu beziffern, müsste man wie beim Beispiel Aids erst einmal wissen, wie verbreitet BSE in Deutschland ist. «Im Prinzip wird wild drauflos getestet, und etliche Millionen Euro werden für BSE-Tests ausgegeben, ohne dass diese wirklich etwas bringen.» Tatsächlich sei vielleicht alles viel harmloser, als wir befürchten — vielleicht aber auch viel schlimmer. Es gebe vermutlich nicht viele Politiker, die diese Problematik durchschauten. Jedoch : «Wenn man in solch einer Situation einen Test einführt, beruhigt das die Bevölkerung. Und die Regierung kann zeigen, dass sie etwas getan hat.»

Das will offenbar auch die Gesundheitsministerin Ulla Schmidt mit ihrem Vorstoß für ein bundesweites Brustkrebs-Screening : Regelmäßige Röntgenchecks, verkündete ihr Ministerium, würden die Brustkrebstodesrate von Frauen über 50 Jahren um fast ein Drittel reduzieren.

Reine Irreführung, ärgert sich Gerd GIGERENZER. «Natürlich kann man den Nutzen solcher Reihenuntersuchungen so darstellen. Aber man weiß genau, dass dadurch überhöhte Erwartungen geweckt werden.» Will man die Frauen ehrlich informieren, geht es auch anders, wie sich leicht zeigen lässt : Brustkrebs ist eine relativ seltene Krankheit, an der ohne Früherkennungsuntersuchungen jedes Jahr etwa 14 von 1000 Frauen in der gefährdeten Altersgruppe sterben. Mammografie-Screenings können die Zahl der Brustkrebsopfer auf 10 reduzieren. Effektiv haben daher nur 4 von 1 000 Frauen einen Nutzen davon, also 0,4 Prozent — was natürlich viel weniger beeindruckend wirkt.

Was Mammografie-Protagonisten zudem gern verschweigen : Durch falsch-positive Mammogramme werden allein in Deutschland jährlich rund 300 000 Frauen monatelang in Angst und Schrecken versetzt. Schätzungsweise 100 000 von ihnen unterziehen sich infolgedessen zumindest einer kleinen Operation, die nicht nur Blutergüsse und Narben hinterlässt, sondern mitunter auch Wundinfektionen nach sich ziehen kann.

Etliche Ärzte schüren die Angst ihrer Patientinnen zusätzlich, weil sie die Aussagekraft von Mammogrammen völlig überschätzen : Hinter einem hoch verdächtigen Röntgenbild, so sind viele Mediziner überzeugt, stecke fast immer ein Tumor. In Reihenuntersuchungen ist dies tatsächlich aber nur bei einer von zehn Frauen mit Positivbefund der Fall.

## 2.4 Trugschluss bei Cannabis

Pfiffige Strategen nutzen die Zahlenblindheit gezielt aus — sei es zur Vermarktung neuer Diätmittel oder Cholesterinsenker oder aber vor Gericht. Paradebeispiel dafür ist der Mordprozess gegen den US-Football-Star O. J. Simpson Mitte der neunziger Jahre. Sein Verteidiger führte damals weitläufig aus, wie ungerecht es sei, es als belastendes Indiz zu werten, dass Simpson seine Frau früher nachweislich misshandelt hatte. Schließlich würde die Mehrzahl der geschlagenen Frauen ja keineswegs ermordet, sondern «nur» geschlagen. Weniger als einer von 2 500 Männern, die ihre Partnerin schlagen, argumentierte der Advokat, gingen so weit, sie zu ermorden.

Die Zahl stimmte, führte die Geschworenen jedoch an der Nase herum. Entscheidend für deren Urteil war eine ganz andere Frage : nämlich bei wie vielen von allen getöteten Frauen, die zu Lebzeiten von ihrem Mann misshandelt worden waren, dieser auch der Mörder war. Und dies war bei acht von neun aller umgebrachten Frauen der Fall. Die Augenwischerei des Verteidigers hatte wohl Erfolg : O. J. Simpson ist seither auf freiem Fuß.

Der Trick funktioniert auch in der Politik, wie die Rede eines bayerischen Innenministers über die Gefahren des Drogenmissbrauchs zeigte : Weil die meisten Heroinabhängigen zunächst Marihuana geraucht hätten, tönte er, würden die meisten Marihuanaraucher auch zu Junkies. Deshalb müsse Marihuana unter allen Umständen verboten bleiben. Fraglich, ob der Trugschluss Absicht war. Womöglich hatte ja nur der Gedanke an Cannabis den Kopf des Ministers vernebelt : Tatsächlich haben zwar die meisten Heroinsüchtigen am Anfang ihrer Drogenkarriere Cannabis geraucht. Das heißt aber nicht, dass die meisten Marihuanafans auch heroinabhängig werden. Die meisten haben mit Heroin ihr Leben lang nichts am Hut.

«In einer Demokratie gibt es viele Gruppen, die ihren Zielen mit statistischen Angaben Nachdruck verleihen wollen», sagt GIGERENZER. Welcher Mediziner will schon Autorität oder Einnahmen einbüßen, indem er seine Patienten allzu genau über die Unwägbarkeiten von Diagnosen und Therapien informiert ? Und welcher Richter würde nicht auf das

Gutachten eines Sachverständigen vertrauen, wenn dieser versichert, ein Täter könne anhand eines genetischen Fingerabdrucks mit «100-prozentiger Sicherheit» identifiziert werden ?

Doch es geht dem Berliner Forscher nicht um Ärzte- oder Juristenschelte. Die Misere sei auch Folge einer selbst verschuldeten Unmündigkeit. «Viele Menschen nehmen einfach hin, was man ihnen versichert, ohne es zu überprüfen oder selbst nach Informationen zu suchen, anhand derer sie selbstständig entscheiden könnten.» Mit mangelndem Talent für Statistik habe dies weniger zu tun, betont GIGERENZER. «Das Wichtigste kann man selbst älteren Menschen durch ein paar kleine Übungen beibringen.» Viel wichtiger sei es, mit Unsicherheiten leben zu lernen. «Das ist das Schwierigste. Bei Krisen wollen viele intuitiv am liebsten zurück in den Mutterschoß.»

## 2.5 Dem Täter auf der Spur

In einem Vergewaltigungsfall, der in den Medien für Schlagzeilen sorgte, kommt es zum Prozess. Am Tatort waren DNA-Spuren des Täters gefunden worden, und bei einem genetischen Massentest ist den Strafverfolgern ein Verdächtiger ins Netz gegangen, dessen «genetischer Fingerabdruck» mit dem am Tatort gefundenen übereinstimmt. Es gibt ansonsten keine verwertbaren Zeugen und Indizien — theoretisch kommen zehn Millionen Männer als Täter infrage. Der Inhaber des genetischen Labors wird als Gutachter vor Gericht geladen und erklärt den Richtern und Geschworenen, wie präzise der von ihm verwendete Test ist :

Die DNA eines beliebig aus der Bevölkerung herausgegriffenen Mannes hat nur mit einer Wahrscheinlichkeit von 0,000 1 Prozent der Fälle dasselbe Profil wie das, welches am Tatort gefunden wurde. Und unser Test würde diese Übereinstimmung praktisch mit Sicherheit zeigen. Umgekehrt irrt sich der Test nur in 0,001 Prozent der Fälle und zeigt eine Übereinstimmung an, wenn das Profil eigentlich verschieden ist.

Sie sind Geschworener in diesem Prozess und sollen nun die folgenden drei Fragen beantworten :

- 1° Wie hoch ist die Wahrscheinlichkeit, dass der Angeklagte dasselbe DNA-Profil hat wie die gefundene Spur ?
- 2° Wie hoch ist die Wahrscheinlichkeit, dass die gefundene DNA tatsächlich vom Angeklagten stammt ?

3° Wie lautet Ihr Urteil : schuldig oder unschuldig?

Cornelia Stolze (ZEIT) – 2002

### 3 Ein kahler Kopf macht nicht reich

Der koreanische Wissenschaftler Hwang ist nicht der einzige Schummler. Hans-Hermann DUBBEN sagt : 90 Prozent aller Studien sind falsch. Hans-Hermann DUBBEN sucht nach Irrtümern in der Wissenschaft. Er war selbst enthusiastischer Biophysiker und schreibt nun mit einem Kollegen Bücher über Fehler der Forscher, jüngst „Mit an Sicherheit grenzender Wahrscheinlichkeit“. DUBBEN lehrt am Institut für Allgemeinmedizin an der Uni-Klinik in Hamburg.

*Herr DUBBEN, mal angenommen, ich bin Forscherin bei einem Pharmaunternehmen und habe ein neues Medikament entwickelt, von dem ich befürchte, dass es nicht besser ist als das Medikament der Konkurrenz. Was kann ich da tun ?*

Sie starten eine Studie, in der Sie die beiden Medikamente vergleichen, nur verzichten Sie darauf, den so genannten Endpunkt festzulegen — eine beliebte Methode. Statt genau zu definieren, was Sie untersuchen wollen, sagen Sie pauschal : Ich schaue mal, ob Medikament A besser ist als Medikament B. Dann können Sie an Ihren Patienten eine Reihe von Parametern messen : Blutdruck, Cholesterinspiegel, Lungenvolumen und vieles andere. Und am Schluss machen Sie eine statistische Erhebung für alle diese Endpunkte. Wenn Sie 100 Parameter untersuchen, werden Sie im Mittel fünf Ergebnisse erzielen, bei denen Ihr Medikament zufällig besser abschneidet.

*Und mit denen gehe ich dann in die Öffentlichkeit.*

Genau. Mit der Absicht zu manipulieren oder ohne - viele Forscher wissen gar nicht, dass sie da einen Fehler machen.

*Wo genau liegt der Fehler ? Im Verschweigen der restlichen Ergebnisse ?*

Das allein kann schon ein grober Fehler sein. Aber zudem bekommen Sie eine Reihe falsch positiver Untersuchungsergebnisse. Das heißt, es ist gut möglich, dass Ihr Medikament in diesen fünf Punkten nicht wirklich besser ist als das Konkurrenzprodukt, sondern nur zufällig bessere Ergebnisse hatte. Stellen Sie sich vor, Sie würfeln zehn Mal mit einem roten Würfel und im Mittel kommt vielleicht vier heraus, und dann würfeln Sie zehn Mal mit einem grünen Würfel und da kommt im Mittel 3,1 heraus. Dann könnten Sie fälschlicherweise denken, der rote sei grundsätzlich besser als der grüne.

*So etwas passiert Wissenschaftlern tatsächlich ?*

Das ist gang und gäbe. Ich habe gerade erst eine Untersuchung in die Hand bekommen, die international publiziert wurde. Da haben Forscher untersucht, welchen Einfluss das Alter des Fahrers auf einen Verkehrsunfall hat, ob er Alkohol getrunken hat, wie das Wetter war, welche Farbe das Auto hatte und zig andere Dinge. Natürlich haben sie ein statistisch signifikantes Ergebnis gefunden : Silberfarbene Autos haben weniger Unfälle. Da kann ich nur sagen : Schön, dass wir für so etwas Geld ausgeben. Das Problem liegt unter anderem darin, dass die angehenden Forscher an den Universitäten viel zu wenig über Methoden lernen.

*Sie selbst sind eigentlich kein Statistiker, sondern Biophysiker.*

Und ich habe früher genau dieselben Fehler gemacht. Ich habe Experimente und klinische Studien ausgewertet und immer etwas Signifikantes darin gefunden. Ein paar Jahre lang habe ich geglaubt, das läge daran, dass ich ein so guter Wissenschaftler bin. Bis mir auffiel, dass meine Herangehensweise von Grund auf nicht stimmt.

*Dann haben Sie damit begonnen, sämtlichen Statistiken zu misstrauen ?*

Ja. Der Unmut wurde immer größer. Meinem Kollegen Hans-Peter BECK-BORNHOLDT ging es genauso : Wir haben mit hehren wissenschaftlichen Zielen angefangen und wurden immer frustrierter.

*Was hat Sie so geärgert ?*

Da werden zum Beispiel Untersuchungen an einer Hand voll Patienten gemacht und verallgemeinert. Die meisten Untersuchungen in den Fachzeitschriften strotzen vor logischen und methodischen Fehlern. Gut 90 Prozent aller Untersuchungen sind unbrauchbar. Und damit habe ich mich noch vorsichtig ausgedrückt. Mein Kollege und ich haben dann irgendwann begonnen, all diese Untersuchungen zu sammeln.

*Statistiken begegnen uns überall im Alltag. Kein Fußballspiel, das nicht bis zum letzten Ballkontakt ausgezählt wird.*

Eine sehr merkwürdige Entwicklung. Sie können auch in ein Lottogeschäft gehen, da finden Sie Statistiken darüber, wie lange meinetwegen die Zahl 13 nicht mehr gezogen wurde und dass sie an diesem Samstag überfällig sei. Als hätte die Lottotrommel ein Gedächtnis!

*Wie sehr bestimmt die Statistik Ihren Alltag ? Steigen Sie gelassen in ein Flugzeug, weil Sie wissen, dass es statistisch ein sicheres Verkehrsmittel ist ?*

Ich treffe solche Entscheidungen eher aus dem Bauch heraus und nicht mit Hilfe der Statistik. Was

nutzt es mir, wenn ich in einem abstürzenden Flugzeug sitze, zu wissen, dass dies ein sehr unwahrscheinlicher Fall ist?

*Wie finden Sie die Studien über Rotwein, bei denen mal herauskommt, dass ein wenig Wein gesund ist und manchmal das Gegenteil?*

Da ziehe ich natürlich diejenigen vor, die für gelegentlichen Weinkonsum plädieren. Bei diesen Studien ist entscheidend, welche Probanden man einbezieht. Schließe ich zum Beispiel einerseits Personen mit ein, die gar nichts trinken, weil sie schwer krank sind und andererseits schwere Trinker, dann ist es kein Wunder, dass die mäßigen Genießer die höchste Lebenserwartung haben.

*Lassen Sie uns noch ein bisschen weiter an meiner Studie manipulieren. Was haben Sie noch für Vorschläge?*

Sie gehen vor wie bei der Wahl in einer Bananenrepublik : Sie schließen bestimmte Bevölkerungsgruppen von vornherein aus, dann werden noch irgendwelche Zettel im Nachhinein als ungültig definiert, weil das Kreuz nicht genau im Kästchen war, und am Ende gewinnt der eigene Kandidat mit drei Stimmen Vorsprung. Ich bespreche in meiner Vorlesung eine Studie, die in „Lancet“ stand, einer der international wichtigsten medizinischen Zeitschriften. Es ging um die Frage, ob man die Schlaganfallhäufigkeit mit Acetylsalicylsäure, also dem Wirkstoff von Aspirin, reduzieren kann. Und dabei wurde genau so verfahren. Am Ende machten die Wissenschaftler das Ergebnis an zwei oder drei Schlaganfällen fest — bei 20 000 Patienten!

*Und zogen daraus den Schluss, dass Acetylsalicylsäure die Schlaganfallhäufigkeit senken kann?*

So war es. Für die Autoren der Studie war das sicher ein Meilenstein auf der Karriereleiter.

*Bei einer so groß angelegten Untersuchung ist der Druck groß, ein positives Ergebnis zu bekommen.*

Ja. Da geht es oft um viel Geld auf einem heiß umkämpften Markt. Es gibt Medikamente, deren Entwicklung Milliarden kostet.

*An der Universität Kopenhagen hat der Wissenschaftler Bodil Als-Nielsen den Einfluss von Sponsoren auf Forschungsergebnisse untersucht. Dabei kam heraus, dass das Ergebnis einer Studie in 51 Prozent der Fälle positiv ausfällt, wenn ein Pharmaunternehmen diese finanziert, und lediglich in 16 Prozent bei einem neutralen Geldgeber.*

Es gibt Fälle, da haben Firmen missliebige Studien lange zu verschweigen versucht. Die sollten sich Gedanken über ihre Glaubwürdigkeit machen.

*Die Firma Schering gab im letzten Jahr bekannt, dass ihr neues Darmkrebsmittel das Tumorstadium nicht mehr bremst als andere Medikamente. Prompt brachen die Aktien um 15 Prozent ein.*

Eine sehr kluge Reaktion der Firma, die damit viel für ihr Ansehen und damit für ihre Zukunft tut. Ich hoffe, dass auch die Aktionäre so etwas eines Tages verstehen.

*In Ihrem Buch kritisieren Sie den Umgang der Wissenschaftler mit einer möglichen Klimakatastrophe — ein Thema, das für Laien undurchsichtig ist.*

Ich bin auch kein Klimafachmann. Aber wenn ich lese, dass uns demnächst in Braunschweig das Wasser bis zur Nase stehen könnte, wenn der Nordpol schmilzt und somit der Meeresspiegel steigt, dann muss ich den Kopf schütteln. Vorhersagen sind schwierig, vor allem, wenn sie die Zukunft betreffen. Und wenn es dabei um die Zukunft eines komplexen Gebildes wie einen ganzen Planeten geht, wird die Sache auch nicht einfacher. Eis ist leichter als Wasser, und da das Eis am Nordpol im Meer schwimmt, ändert sich am Wasserspiegel überhaupt nichts, wenn es schmilzt. Das können Sie leicht ausprobieren : Werfen Sie einen Eiswürfel in ein Wasserglas und markieren Sie den Wasserpegel mit einem Strich. Wenn der Eiswürfel geschmolzen ist, ist der Wasserspiegel noch genauso hoch.

*Klimaforscher warnen aber vor allem vor dem Abschmelzen des Grönlandeises, das im Moment ja noch auf dem Festland gebunden ist : Die Schmelze des Eisschildes würde den Meeresspiegel weltweit um fast sieben Meter steigen lassen, heißt es.*

Das mag sein, aber die Erklärung mit den schmelzenden Polen finden Sie permanent in der Presse. Das ist wie mit dem Spinat, von dem bis heute viele Menschen glauben, dass er besonders viel Eisen enthält. Dabei beruht diese vermeintliche Erkenntnis auf einem Tippfehler. Ein falsch gesetztes Komma hat den Eisenwert um ein Zehnfaches erhöht. Tatsächlich enthält Spinat nicht mehr Eisen als anderes Gemüse.

*Ihr zweiter Kritikpunkt am Umgang mit einer Klimakatastrophe ist der grafische Ausschnitt, der als Beleg für die Erderwärmung herangezogen wird, und der Ihrer Meinung nach zu eng gefasst ist.*

Ich wähle ganz einfach den Ausschnitt, der mir am besten passt ... Mit der seit 1880 steigenden Temperaturkurve kann man gut suggerieren, dass es auch in Zukunft so weitergeht. Nehme ich aber die Daten der letzten 110 000 Jahre, komme ich zu einem ganz anderen Bild : Es ergibt sich ein kontinuierlicher Wechsel zwischen Eis- und Warmzeiten.

*Sie stützen sich bei diesen Daten auf Messungen an Bohrkernen im grönländischen Eisschild. Aber*

*wenn Sie alles hinterfragen, dann müssten Sie hier weiterfragen : Wer sagt, dass diese Technik tatsächlich funktioniert ?*

Zumindest diejenigen, die sie anwenden. Da wird aus der Zusammensetzung des Eises auf die Temperatur geschlossen. Ich habe keine Ahnung, wie valide diese Methode ist. Ich muss mich, wie in der gesamten Forschung, darauf verlassen, dass hier keiner schummelt.

*Wie der südkoreanische Forscher Hwang Woo Suk, der mit bahnbrechenden Stammzellstudien zum Star wurde. Jetzt hat eine Kommission bestätigt, dass er sie komplett gefälscht hat. Wie konnte Hwang so weit kommen, ohne aufzufliegen ?*

Zunächst einmal, weil es keinen lückenlosen Kontrollprozess für wissenschaftliche Veröffentlichungen gibt. Zwar ist da die Institution der Peer-Review — wissenschaftliche Publikationen werden vor der Veröffentlichung von unabhängigen Experten geprüft —, doch das ist mehr mit einer Stichprobe vergleichbar. Es bekommen ja auch nicht alle Parksünder einen Strafzettel! Wenn ich als Experte meine Aufgabe ernst nehme, brauche ich viel Zeit dafür — ohne, dass ich finanziell entschädigt würde oder mein Name irgendwo auftaucht. Da machen es sich die meisten Experten lieber einfach und schauen eine Veröffentlichung nur auf grobe Fehler durch.

*2001 war der hoffnungsvolle Nachwuchsphysiker Jan Hendrik Schön mit seinen elektronischen Schaltkreisen aus organischem Material als Fälscher entlarvt worden. Haben derartige Fälle zugenommen ?*

Das ist schwer zu sagen, weil wir ja nicht wissen, wie viele Fälscher zuvor unentdeckt blieben. Vielleicht wird heute nur kritischer hingeguckt, vielleicht aber auch tatsächlich mehr gelogen. Die Zeitschrift „Nature“ hat im Juni 2005 eine quantitative Umfrage veröffentlicht, in der jeder dritte befragte Forscher angab, schon unredlich gearbeitet zu haben. Indem er beispielsweise Ergebnisse verfälscht oder Daten, die eigenen früheren Publikationen widersprechen, gänzlich unterschlägt.

*Was schlagen Sie vor ?*

Zunächst einmal eine bessere Ausbildung. Mein Kollege und ich bieten seit einiger Zeit Forschungsmethodik-Seminare für Wissenschaftler und Journalisten an. Ein anderer Punkt ist, dass die Forschung mehr Freiräume braucht, sie funktioniert nicht, wenn sie komplett den Regeln des Marktes unterworfen ist.

*Sie wollen die Wissenschaftler wieder in den Elfenbeinturm schicken.*

Die Forschung kann sich natürlich nicht völlig vom Markt loslösen, aber die derzeitige Entwicklung

finde ich sehr besorgniserregend. Auch an den Universitäten greift das amerikanische Modell immer mehr um sich, dass Wissenschaftler an den Drittmitteln, die sie einwerben, finanziell beteiligt werden. Der Elfenbeinturm ermöglicht es dem Forscher zumindest, in Ruhe über etwas nachzudenken und seine eigene Arbeit kritisch zu überprüfen und zu wiederholen. Qualitäten, die zurzeit auf der Strecke bleiben. So, wie Forschung momentan betrieben wird, ist sie zum großen Teil Geldverschwendung, Verschleiß an Nachwuchsforschern und Betrug am Patienten.

*Aber was mache ich, wenn ich als Patient auf die aktuellen Studien angewiesen bin, um mich für eine Behandlungsmethode zu entscheiden ?*

Sie sollten sich auf jeden Fall fachlichen Rat einholen und zwar von mehreren Ärzten unterschiedlicher Disziplinen. Ich fände es zumindest sehr beruhigend, wenn alle etwas Ähnliches raten.

*Und wie gehe ich mit den zahlreichen Statistiken um, die mir im Alltag begegnen ? Wenn ich Sie ernst nehme, dürfte ich gar nichts mehr glauben.*

Diese Statistiken sollten Sie als Unterhaltung ansehen, so wie Horoskope. Wenn Sie zum Beispiel lesen : „Gassi gehen mit dem Hund vermindert das Risiko für Brustkrebspatientinnen, erneut zu erkranken und zu sterben.“ Der Artikel stand im „Journal of the American Medical Association“.

*Und ?*

Tatsächlich gibt es einen Zusammenhang zwischen den Patientinnen, die am längsten überlebt haben, und dem Spazierengehen. Die Erklärung : Wem es gut geht, der lebt länger und geht also auch öfter spazieren. Die Studie findet auch einen Zusammenhang zu anderen Todesursachen, egal ob Herzinfarkt oder Autounfall. Ich würde also andersherum formulieren : Tod hindert an Bewegung.

*Es ist doch so, dass viel Bewegung die körperliche Fitness steigert.*

Das heißt noch lange nicht, dass Spaziergänge vor sämtlichen Krankheiten und sogar Autounfällen schützen. Es gibt auch einen Zusammenhang zwischen dem Jahresbruttoeinkommen und dem Haarwuchs : Glatzköpfe verdienen mehr. Der Mechanismus, der dahintersteckt, ist einfach : Glatzköpfe sind erstens häufig älter und kriegen deshalb mehr Geld, und meistens sind es Männer, die ebenfalls oft mehr verdienen. Es wäre trotzdem sinnlos, sich den Kopf kahl zu scheren, um reich zu werden.

Simone Kosog — Der Tagesspiegel