

Walter KRÄMER, Dortmund

So lügt man mit Statistik

1. Einleitung und Überblick

Mit Statistik und mit Statistiken kann man Lügen. Das ist trivial. Auch mit Bildern, Gesten, Worten kann man lügen. Ich werde mich im weiteren auf eine Auswahl eher indirekter Lügen konzentrieren. Diese indirekten Lügen leben davon, daß der Empfänger einer Nachricht über den Grad der Gewißheit getäuscht wird, über den der Versender der Nachricht verfügt (Abschnitt 2), oder die Botschaft, vom Versender gesteuert, nicht richtig übersetzt (Abschnitte 3 und 4), oder über den Einfluß des Zufalls irreführend unterrichtet wird (Abschnitt 5). In allen diesen Fällen geschieht die Lüge sozusagen durch Mittäterschaft des Opfers: Der Lügner zählt darauf, daß der Belügte ihm bei seinen Lügen hilft, daß er oder sie aus Unwissenheit oder Fahrlässigkeit eine Nachricht anders dechiffriert als es dem wahren Sachverhalt entspricht.

2

3. Damit auch per 30.9. ein Exportüberschuß gemeldet werden kann wird vorgeschlagen, Veränderungen in gleicher Höhe wie für das 1. Halbjahr vorzunehmen..
Für den 30.9. wird vorgeschlagen, ein zu melden
Daraus ergibt sich ein Exportüberschuß von ~~521~~ ⁹¹⁰ Mio VM. *910*
Gleichzeitig wird damit erreicht, daß kein ~~zu~~ starker Rückgang beim Export und Import im Vergleich zu den im Vorjahr gemeldeten Angaben eintritt.

Daraus ergibt sich ein zu meldender

	zum gemeldeten Ist 30.9.1986
Export von 18 803 Mio VM	= 93,6 %
Import von 148 282 Mio VM	= 197,9 %
Umsatz von 137 089 Mio VM	= 195,7 %
Export- überschuß von 521 ⁹¹⁰ Mio VM	

4. Der Exportüberschuß im Handel mit dem sozialistischen Wirtschaftsgebiet beträgt ~~2 044~~ ^{3 654} Mio VM, so daß für den Außenhandel insgesamt ein Saldo von ~~142 564~~ ³⁸⁸ Mio VM per 30.9.1987 an den RGW und UNO-Organen gemeldet wird.

3 654

Wird für den Abj. c 30.9. = 115% bef. abgeleitet

Wird für den Abj. c 30.9. = 134% bef. abgeleitet

Abb. 1: Ein Originaldokument aus den Archiven der amtlichen DDR-Statistik: ein ohnehin schon falscher Außenhandelsüberschuß von 521 Millionen VM wird von Günter Mittag eigenhändig in 910 Millionen VM umgewandelt.

2. Lügen durch die Illusion der Präzision

Laut Auskunft des amerikanischen CIA lebten im Juli des Jahres 2000 in der Volksrepublik China 1.261.482 Menschen, davon 612.855.442 Männer und 648.977.040 Frauen (CIA – The World Factbook 2000). In diesen Zahlen sind alle Ziffern bis auf die ersten beiden falsch. Bzw. sie stimmen allenfalls per Zufall mit den wahren Ziffern überein. Selbst nach aufwendigen Volkszählungen kennt man die Wohnbevölkerung eines so riesigen Landes wie China nur bis auf eine Fehlerquote von +/- 0,5% (siehe etwa Morgenstern 1963), und zu anderen Stichtagen sind akkurate Messungen, so wie durch die obigen Zahlen unterstellt, völlig ausgeschlossen. In diesem Sinn sind – Täuschung des Publikums über die wahre Qualität einer Information – sind diese Zahlen also eine Lüge.

Vielziffrige Zahlen erzeugen eine durch die wahren Informationen nicht immer gedeckte Illusion der Präzision und schützen damit den Verkünder einer Nachricht vor Kritik. Nichts schreckt einen Angreifer so sehr wie eine harte Zahl. Und je mehr Ziffern, desto härter. Die Zahl der Ziffern gleicht sozusagen der Dicke des Stahlrumpfes bei Panzerschiffen: je dicker, desto eher prallten feindliche Attacken ab. Unter anderem auch deshalb datiert der deutsche Reformator Phillip Melancton die Erschaffung der Erde exakt auf das Jahr 3963 vor Christus, und der englische Theologe John Lightfoot ist sogar noch präziser: „Himmel und Erde und alles was dazugehört wurde vom dreifaltigen Gott zusammen und zur gleichen Zeit erschaffen am Sonntag, dem 21. Oktober 4004 vor Christus, 9 Uhr morgens“ (Downs 1987, S. 10, Übersetzung von W.K.). Damit ist jeder Zweifel ausgeräumt.

Dergleichen scheinpräzise Zahlen finden wir in fast allen Bereichen unseres Lebens. In einer Schätzung der Kosten des 1. Weltkrieges kommt die *Carnegie Endowment for International Peace* auf \$ 186.333.637.097,- (Pitkin 1932, S. 425). Die Olympischen Spiele in Barcelona 1992 dagegen kosteten nach einer Zeitungsmeldung, deren genaue Herkunft ich vergessen habe, 2.409.196.600 Mark. Und im zweiten Weltkrieg kamen auf alliierter Seite 6.548.233 Zivilpersonen um (Central Statistical Office 1995, S. 34). In allen diesen Fällen wird einer u. U. eher zweifelhaften Botschaft durch übertriebene Präzision ein ihr nicht zukommender Glaubwürdigkeitsbonus erschwindelt.

Auf der Grenze zwischen Lüge und Rhetorik liegen dabei Meldungen wie die der *Bildzeitung*, daß im Jahr 1998 genau 3687 deutsche Briefträger von einem Hund gebissen worden seien. Hier war wohl die Absicht, eine ansonsten etwas blasse Meldung wie „Tausende deutscher Briefträger werden jedes Jahr von einem Hund gebissen“ faßbarer und anschaulicher zu machen. Deswegen nutzen auch Literaten gerne diese Stilfigur. Laut Bibel war Goliath nicht riesengroß, sondern sechs Ellen und eine Hand breit groß. In seinem *Hessischen Landboten* klagt Büchner nicht über eine überhohe Steuerlast, sondern er gibt die von den Einwohnern des Großherzogtums Hessen pro Jahr gezahlten Steuern mit exakt 6.363.364 Gulden an. Oder Rabelais: „Mit fröhlichem Lächeln tat Gargantua seinen Hosenlatz auf, zog den Spritzschlauch hervor und ließ einen solchen Erguß auf die Umstehenden los, daß 260.418 Pariser eines bitteren und feuchten Todes starben, Weiber und Kinder gar nicht eingerechnet“ (zitiert nach Schneider 1999, S. 70). In diesen Beispielen kann man eine übertriebene Präzision vielleicht noch als Mittel der Rhetorik gelten lassen.

Auch hinter anderen scheinbar exakten Zahlen steckt nicht immer eine manipulative Absicht. Wenn wir etwa im *World Handbook of Political and Social Indicators* lesen, daß ein Luxemburger im Jahr 1977 genau 3713 Kilokalorien täglich konsumierte, ein Einwohner des Senegal dagegen nur 2261 (Taylor und Jodice 1983, S. 142-143), so kommen diese scheinpräzisen Zahlen vermutlich durch Division eines grob geschätzten nationalen Nahrungsmittelverbrauchs durch die Wohnbevölkerung zustande. Deshalb sagt auch das Statistische Bundesamt nicht: „Im September 1992 wurden 3384 Gartenschirme aus Polen nach Deutschland importiert“, sondern: „Es wurden 3384 importierte Gartenschirme *gezählt*. Aber diese implizite Warnung wird von den Konsumenten solcher scheinpräzisen Zahlen nicht in allen Fällen wahrgenommen.

3. Lügen durch unterschiedliche Definitionen

Wie viele Analphabeten gibt es in Deutschland? Nach dem regierungsamtlichen *Bericht zur Bekämpfung des Analphabetismus in der Bundesrepublik* sind es über vier Millionen (Bundesministerium für Bildung und Wissenschaft 1993). Aber wer ist ein Analphabet? Jemand, der mit einem Kreuz unterschreibt? Oder der oder die Probleme hat, einen Artikel im *Spiegel* ohne Mühe zu verstehen? Je nachdem, wie wir Analphabetismus messen, ist diese Zahl einmal größer und einmal kleiner, und diese Abhängigkeit der Werte gewisser Statistiken von der Art ihrer Definition eröffnet Möglichkeiten zu Lüge und Mißbrauch vielfältigster Art.

Wir haben hier ein grundsätzliches, in der Statistik auch als „Adäquationsproblem“ bekanntes Dilemma: die Dinge, die man messen will, und die dafür bereitgestellten Zahlen stimmen nicht immer überein. Das fängt bei vergleichsweise einfach zu erhebenden Sachverhalten wie etwa der Zahl der jährlich im Straßenverkehr getöteten Personen an (6962 im Jahr 2001 in der Bundesrepublik). Ein Jugendlicher fährt nachts auf dem Heimweg von einer Diskothek gegen einen Baum und verstirbt vier Tage später im Krankenhaus. Ist er ein Verkehrstoter? In der ehemaligen DDR nein. Hier mußte man binnen dreier Tage sterben, um in die Statistik der Verkehrstoten einzugehen. In der Bundesrepublik Deutschland ja. Hier hat und hatte man mit dem Sterben bis zu 31 Tagen Zeit.

Auch in der Statistik der Säuglingssterblichkeit legten die DDR und die Bundesrepublik verschiedene Maßstäbe an. In der DDR war ein Säugling lebend geboren, und konnte damit sterben, wenn das Herz schlug und die Atmung funktionierte. In der Bundesrepublik war und ist ein Säugling lebend geboren, und kann und konnte damit sterben, wenn das Herz schlägt *oder* die Atmung funktioniert. Mit anderen Worten: Viele der im Westen als verstorben gemeldeten Säuglinge kamen im Osten, da nie offiziell geboren, in keiner Säuglingssterblichkeitsstatistik vor, und eröffneten damit der DDR u. a. die Möglichkeit zu der Lüge, ihr Gesundheitswesen sei dem des Westens überlegen.

Auch Krankheit ist nicht einfach zu erfassen. Bäuerinnen z.B. kommen in der Regel nicht auf den Gedanken, Rückenschmerzen als Krankheit zu bewerten und deshalb zum Arzt zu gehen, während das Stadtfrauen viel öfter tun. Fortgesetztes Husten ist nach für 77 Prozent der Oberschicht ein Grund für einen Arztbesuch, in der Unterschicht dagegen nur für 23 Prozent (Koos 1970). Bei chronischer Müdigkeit ist das Verhältnis 80:19, bei Blut

im Stuhl 98:60 und bei Schmerzen im Brustkorb 80:31. Wer als "krank" gilt, hängt also auch davon ab, wo man wohnt und wieviel Geld man hat.

Vor allem bei der Wahrnehmung und Darstellung des Phänomens der „Armut“ wird diese vielen Begriffen innewohnende Unbestimmtheit auch für politische Propaganda ausgenutzt. (Krämer 2000). Die politische Linke hierzulande nimmt als Maß dafür gern den Anteil der Familien mit einem Einkommen unterhalb der Hälfte des Durchschnitts. Und wundert sich, daß trotz wachsenden Wohlstands die Armut in Deutschland nicht verschwindet. Nach den Maßstäben der Vereinten Nationen dagegen gilt als arm, wer weniger als einen Dollar täglich zum Überleben zur Verfügung hat. Damit ist in Deutschland niemand, weltweit aber jeder fünfte, zusammen über eine Milliarde Menschen, arm.

Daß selbst ein- und dieselben Personen bestimmte Begriffe einmal so und einmal so betrachten, zeigt die letzte Volkszählung in Kanada: die „Ureinwohner“ werden immer zahlreicher; sie haben sich seit der letzten, nur fünf Jahre zurückliegenden Zählung fast verdoppelt. Aber nicht, weil sie wirklich mehr geworden wären, sondern weil immer mehr Kanadier sich als Ureinwohner sehen. Ein gestiegenes Selbstbewußtsein plus verschiedene soziale Vergünstigungen machen diesen Status immer attraktiver, und so stuft sich denn ein Kanadier, der unter seinen 16 Ur-Ur-Urgroßvätern einen Indianer oder Eskimo besitzt, heute gern als Ureinwohner ein.

4. Eine Fallstudie: Wie zählt man die Arbeitslosen?

Der Prototyp einer Statistik, von der fast jeder glaubt, sie sei gefälscht, ist aber die der Arbeitslosigkeit. Hier kommen fast alle Aspekte des wirtschaftsstatistischen Adäquationsproblems zusammen: Der Begriff als solcher läßt verschiedene Ausdeutungen zu, selbst Experten streiten, wie das ganze eigentlich zu messen sei, und die Versuchung, diese Zahlen zum eigenen Vorteil zu manipulieren, und damit das Publikum über den wahren Zustand des Arbeitsmarktes zu täuschen, ist hier ganz besonders groß. Kommt dann noch eine hoffnungslose Überlastung einer einzigen Zahl mit allzuvielen, zum Teil unvereinbaren Informationspflichten hinzu, ist die Katastrophe perfekt. Denn die Arbeitsmarktstatistik „ist eine Dienerin vieler Herren und entsprechend manipulierbar“ (Bröll 2000). „So dient sie einmal der Regierung, um ihre Politik zu rechtfertigen, und gleichzeitig der Opposition, um Versäumnisse der Regierung anzukreiden. Sie ist den Verbandsvertretern zu Diensten, die Reformen anmahnen, und den Gewerkschaften, die sich schon im Vorfeld gegen diese Reformen wehren wollen. Wozu sie allerdings zu wenig dient, ist ihrem eigentlichen Zweck, nämlich ein möglichst exaktes Bild der Wirklichkeit widerzuspiegeln.“

Während diese Zeilen geschrieben werden, zählt die amtliche Statistik in Deutschland rund 4,5 Millionen Menschen als arbeitslos. Dazu gehört, wer (i) mindestens 15 Stunden in der Woche gegen Entgelt arbeiten will, (ii) nicht nur vorübergehend Arbeit sucht, (iii) älter als 15 und jünger als 66 Jahre ist, (iv) dem Arbeitsmarkt unmittelbar zur Verfügung steht, sowie (v) bei einem Arbeitsamt offiziell als arbeitssuchend gemeldet ist. Es ist also gar nicht so einfach, in Deutschland amtlich arbeitslos zu sein. Wer nur eine Teilzeitarbeit von weniger als 15 Stunden oder eine Ferienstelle sucht, wer wegen Krankheit oder Umschulung dem Arbeitsmarkt vorübergehend nicht zur Verfügung steht oder die Suche per Arbeitsamt ganz einfach aufgegeben hat, der oder die ist damit auch nicht arbeitslos. So gesehen sind die amtlichen Zahlen also viel zu klein.

Anderswo ist man sogar noch restriktiver (Sorretino 2000). In Frankreich zählen Teilzeitarbeitslose, die keine Vollzeitstelle suchen, überhaupt nicht mit, auch bei mehr als 15 gewünschten Wochenstunden nicht, und in England zählt man nur solche Menschen als arbeitslos, die auch Arbeitslosenunterstützung erhalten. Mit anderen Worten, wenn die Unterstützung ausläuft, ist auch die offizielle Arbeitslosigkeit vorbei.

Aber auch weitergehende Definitionen kommen vor. In den USA z. B. ist auch ein 70jähriger Rentner, der vergeblich für einige Stunden in der Woche nach einer kleinen Hausmeisterstelle sucht, im Sinne der Statistik arbeitslos. Dort ist die Statistik also viel umfassender.

Ein Begriff, der unserer intuitiven Vorstellung von „arbeitslos“ vermutlich noch am nächsten kommt, ist die sogenannte „Erwerbslosigkeit“. Er liegt z.B. den Zahlen von EUROSTAT (das Statistische Amt der Europäischen Gemeinschaft) oder der Internationalen Arbeitsorganisation ILO zugrunde. Erwerbslos im Sinne dieser Statistik ist jede Person im arbeitsfähigen Alter, die gegen Entgelt arbeiten will, ganz gleich wie lange und unabhängig davon, ob beim Arbeitsamt gemeldet oder nicht, und der oder die bei diesem Bemühen nicht erfolgreich ist.

Auf den ersten Blick scheinen also die Erwerbslosen die Arbeitslosen zu umfassen. Denn für den Status der Erwerbslosigkeit ist es unerheblich, ob man als arbeitssuchend gemeldet ist; auch die Hausfrau/der Hausmann, der oder die gerne wieder ins Berufsleben zurückkehren möchte, Bummelstudenten, die nur mangels Berufsaussichten noch studieren, oder Langzeitarbeitslose, die aus Hoffnungslosigkeit oder wegen Wegfall ihrer Leistungsansprüche die Meldung beim Arbeitsamt zurückgezogen haben, gehören hier dazu. Die Gesamtheit dieser „heimlichen Arbeitslosen“, auch „Stille Reserve“ genannt, umfaßt in Deutschland je nach Zählweise zwischen einer und drei Millionen Menschen.

Auf den zweiten Blick schließen aber die Erwerbslosen auch viele „offizielle“ Arbeitslose aus, die heute noch in der Statistik erscheinen. Nämlich alle diejenigen, die in Wahrheit keine Arbeit suchen und nur zum Abschöpfen verschiedener Vergünstigungen beim Arbeitsamt als arbeitslos gemeldet sind. Auch davon gibt es mehr als genug. Diese Menschen sind zwar amtlich arbeitslos, aber nicht erwerbslos, da an legal bezahlter Arbeit wenig interessiert..

Nach einer Umfrage des Bonner infas-Instituts vom Frühjahr 2000 suchen nur die Hälfte aller beim Arbeitsamt gemeldeten Personen tatsächlich eine Stelle (Brixy et al. 2002). Die wichtigsten wahren Gründe für den Gang zum Arbeitsamt waren (Mehrfachnennungen möglich): „Ich möchte sicherstellen, daß die Zeit der Arbeitslosigkeit später bei der Rentenberechnung berücksichtigt wird“ (83%). „Ich bin auf Leistungen des Arbeitsamtes angewiesen“ (76%). „Ich überbrücke die Zeit bis zum Beginn meiner neuen Stelle bzw. Ausbildung“ (57%). „Ich überbrücke die Zeit bis zum Ruhestand“ (21%), „Das Sozialamt verlangt, daß ich mich arbeitslos melde“ (16%), „Ich muß mich arbeitslos melden, um ausreichend Unterhalt von meinem geschiedenen Ehepartner zu bekommen“ (4%) und „Ich überbrücke die Zeit bis zum Zivil- oder Wehrdienst“ (2%). Mit anderen Worten, das Arbeitsamt ist für viele eine Sozialbehörde und keine Arbeitsplatzvermittlungsstelle.

Seit der Übernahme der Regierung, spätestens seit Gerhard Schröders vielzitiertem Frühjahrs-Interview in *Bild* 2001, in welchem er „kein Recht auf Faulheit in unserer Gesellschaft“ zugestand, sehen das auch solche Kreise in der deutschen Politik, die bislang nicht an Scheinarbeitslose glauben wollten. Ja, das vormalige Abstreiten ihrer Existenz hat sich in eine regelrechte Jagd auf diese umgekehrt. "Wir werden eine international vergleichbare Arbeitsmarktstatistik schaffen, in der nur Personen, die auch tatsächlich dem Arbeitsmarkt zur Verfügung stehen, erfaßt werden", heißt es in der Koalitionsvereinbarung 2002. Wir dürfen gespannt sein, was hier noch auf uns wartet. Denn die Möglichkeiten des Kriteriums „steht dem Arbeitsmarkt unmittelbar zur Verfügung“ sind noch längst nicht ausgereizt. Durch das Verlängern des Erziehungsurlaubs von 6 auf 10 Monate im Jahr 1986 ließ die damalige CDU/FDP Regierung auf einen Schlag 100.000 arbeitslose Frauen aus der Statistik verschwinden. Umschulungsmaßnahmen, ABM-Stellen, Frühpensionierung, Beschäftigungsbrücken: Das alles macht vormalige Arbeitslose nicht mehr kurzfristig verfügbar und läßt damit die Arbeitslosenzahlen schrumpfen.

Vor allem durch das kreative Auslegen dieser Bestimmung, daß nur der als Arbeitsloser zählt, wer dem Arbeitsmarkt auch zur Verfügung steht, haben es etwa die Niederländer geschafft, ihre international bewunderte niedrige Arbeitslosenquote von weniger als 4% zu erzeugen: Von einer Kündigung bedrohte Arbeitnehmer werden kurz vor der Entlassung erwerbsunfähig geschrieben und die Arbeitslosenversicherung wird in „Erwerbsunfähigkeitsversicherung“ umbenannt. Damit ist die Arbeitslosigkeit verschwunden. Mittlerweile ist jeder zehnte Niederländer im erwerbsfähigen Alter als erwerbsunfähig eingestuft und damit gegen Arbeitslosigkeit geimpft.

Oder man steckt die Arbeitslosen ins Gefängnis. In den USA sitzen derzeit zwei Millionen Menschen hinter Gittern, die Hälfte davon jung und schwarz, die ansonsten die Arbeitslosenzahl vergrößern würden.

An Möglichkeiten, dieses gesellschaftlich so brisante Phänomen der Arbeitslosigkeit nach eigenen Vorstellungen gefiltert darzustellen, ist also kein Mangel. Sie resultieren aus der Unmöglichkeit, mit einer einzigen Zahl mehr als eine Facette des Sozialgeschehens abzubilden, und sollten bei der Interpretation der Zahlen stets die Frage auslösen: „Meint der Überbringer den Nachricht damit das gleich wie ich selbst?“

5. Lügen durch Umfragen

Nach einer Umfrage der IG Metall lehnen 95 Prozent der bundesdeutschen Arbeitnehmer das Arbeiten am Samstag ab. Nach einer zeitgleichen Umfrage des unternehmernahen Offenbacher Marplan-Instituts dagegen sind 72 Prozent aller Arbeitnehmer auch zum Arbeiten am Wochenende bereit.

Damit sind wir bei einer weiteren weit verbreiteten Methode, das Publikum über wahre Sachverhalte zu täuschen, angekommen: suggestive Fragestellungen. Denn obiger Widerspruch erklärt sich durch die jeweiligen Fragebögen. "Votum für das freie Wochenende" steht bei der IG Metall in großen Lettern obenan. Es folgt eine lange Erläuterung der Mühen, die das Durchsetzen der 5-Tage-Woche den Gewerkschaften gekostet habe, und eine Aufzählung aller Vorteile, die der freie Samstag für die Familie, die Gesellschaft, den Frieden und die Menschheitszukunft bringe, die dann zu der eigentlichen Frage überleitet: "Was entspricht Deiner/Ihrer Meinung?"
(i) Nach meiner Ansicht wäre die Abschaffung des freien Wochenendes ein

schwerer Schlag für Familie, Freundschaften, Partnerschaften, für Geselligkeit, Vereine, den Sport und das Kulturleben (ii) Ich halte den gemeinsamen Freizeitraum des Wochenendes für nicht so wichtig (iii) Weiß nicht/keine Angabe."

Daß hier fast alle wie gewünscht die erste Antwort wählen, sollte niemanden erstaunen.

Genauso suggestiv, wenn auch mit umgekehrter Absicht, fragt auch das Marplan –Institut. Auf die Frage: "Inwieweit wären Sie bereit, samstags zu arbeiten, wenn es für die wirtschaftliche Situation Ihres Unternehmens gut wäre?" bietet es folgende Auswahlmöglichkeiten an: (i) gelegentlich, wenn dafür an einem anderen Tag arbeitsfrei ist, (ii) häufiger, wenn dafür ein Zusatzurlaub herauskommt, (iii) abwechselnd und (iv) nicht bereit.

Auch hier waren die wenigen Kreuze bei "nicht bereit" schon im Fragebogen und in der Art der Fragen vorgefertigt. Dergleichen Umfragen, ob vom ADAC zum Thema Tempolimit, ob von Greenpeace zum Atomausstieg oder von der katholischen Kirche zur Frage der Abtreibung, geben oft vor allem die Meinung der Befrager, weniger die Meinung der Befragten wider und verbreiten in diesem Sinne eine Lüge.

6. Lügen mit Hilfe des Signifikanzbegriffs

Für viele Zeitgenossen ist "Signifikanz" ein neuer Adelstitel: wissenschaftlich untermauert, empirisch unangreifbar, jenseits allen Zweifels abgesichert, sozusagen die TÜV-Plakette für Seriosität und Faktentreue.

In der mathematischen Statistik heißen Beobachtungen oder Daten "signifikant", wenn sie einer Ausgangs- oder Nullhypothese so sehr widersprechen, daß dieser Widerspruch nicht mehr allein durch Zufall zu erklären ist. Ein neues Medikament wird an hundert Probanden getestet, sechzig davon erfahren eine Besserung. In einer mit einem Placebo behandelten Kontrollgruppe dagegen verzeichnen nur fünfzig von hundert Personen eine Besserung. Wirkt das Medikament systematisch besser als Placebo, oder ist seine Wirkung auch durch Zufall zu erklären?

Angenommen, das Medikament wirke wirklich nur wie ein Placebo, würde also im Durchschnitt bei der Hälfte der Kranken eine Besserung bewirken und bei der anderen Hälfte nicht. Unterschreitet die unter dieser Annahme berechnete Wahrscheinlichkeit für 60 oder mehr Besserungen ein vorher festgelegtes Signifikanzniveau - traditionellerweise fünf Prozent - , schließt man den Zufall als Alleinursache aus, die Wirkung ist „signifikant“.

Diese Vorgangsweise ist heute Standard quer durch viele Wissenschaften, und eine konstante Quelle von Irrtümern und Lügen vielfältiger Art. Z. B. ist selbst vielen Wissenschaftlern nicht bewußt, daß "signifikant" nur heißt, daß ein Ergebnis unter der Ausgangshypothese äußerst unwahrscheinlich ist. Ob diese Ausgangshypothese falsch ist oder nicht, ob sie eklatant verletzt ist oder nur ein wenig, oder ob die Abweichung von der Ausgangs-

hypothese in irgendeiner Weise praktisch von Bedeutung ist, dazu sagt eine signifikante Statistik für sich allein genommen noch nichts aus.

So hängt etwa die Signifikanz einer gegebenen Abweichung von der Ausgangshypothese entscheidend von der Stichprobengröße ab. In obigem Placebo-Beispiel etwa braucht man bei zehn Probanden eine Erfolgsquote von achtzig Prozent, bei hundert Probanden eine Erfolgsquote von 58 Prozent, bei tausend Probanden eine Erfolgsquote von 52,6 Prozent, und bei zehntausend Probanden eine Erfolgsquote von nur noch 50,8 Prozent für ein signifikantes Resultat. Mit anderen Worten, bei großen Stichproben sind selbst kleinste Abweichungen von der Ausgangshypothese schon "signifikant", bei kleinen Stichproben dagegen kann man selbst grobe Verletzungen der Ausgangshypothese noch dem Zufall in die Schuhe schieben.

Deshalb ist signifikant auch etwas anderes als wichtig. Wenn das neue Medikament die Besserungswahrscheinlichkeit von 50 Prozent auf 51 Prozent erhöht, so kann man das wohl kaum als revolutionären Fortschritt in der Medizin bezeichnen. Aber in einer hinreichend großen Stichprobe wird die Verbesserung trotzdem "signifikant".

Dergleichen Mißverständnisse gehen auf ein falsches Verständnis des Signifikanzbegriffs zurück und sind noch keine Lüge. Das direkte Lügen mit Signifikanzen funktioniert wie folgt: Angenommen, obige Pharmafirma testet ihr neues Medikament an 100 Patienten, mit 55 Besserungen. Damit ist diese Erfolgsquote nicht signifikant (korrekter: nicht signifikant besser als fünfzig Prozent). Die Gesamtstichprobe von 100 Patienten setze sich aus zehn Teilstichproben à zehn Patienten zusammen, mit den folgenden Resultaten:

Stichprobe Nr.	Besserungen	Erfolgsquote
1	4	40%
2	6	60%
3	9	90%
4	7	70%
5	4	40%
6	5	50%
7	6	60%
8	3	30%
9	5	50%
10	6	60%

Trotz einer globalen Erfolgsquote von nur 55% ist in einigen Teilstichproben diese Quote sehr viel höher - in der dritten etwa neun von zehn. Ein häufiger Betrug besteht nun darin, nur diese Teilstichprobe auszuwerten,

und zwar so, als wäre es die einzige gewesen. Dann betrüge, wenn die Ausgangshypothese "Medikament nicht besser als Placebo" zuträfe, die Wahrscheinlichkeit von neun Besserungen bei zehn zufällig ausgewählten Patienten nur 1,07 %; sie läge damit beträchtlich unter 5% - ein hochsignifikantes Resultat.

Was diese Rechnung aber verschweigt, und was viele andere, die so gern das Wort "signifikant" im Munde führen, in ähnlichen Fällen ebenfalls verschweigen, ist die alles andere als zufällige Auswahl der Stichprobe, und die damit unzulässige Berechnung der Wahrscheinlichkeit von 1,07%. Denn in Wahrheit ist die Wahrscheinlichkeit für das Ergebnis zehnmal größer: daß in mindestens einer von zehn Stichproben neun oder mehr Patienten rein zufällig genesen, geschieht mit Wahrscheinlichkeit

$$1 - 0,9893^{10} = 0,1020 = 10,2\%.$$

Mit anderen Worten, dieses Resultat ist auch durch Zufall zu erklären, es ist nicht signifikant.

Solche scheinsignifikanten Resultate plagen die empirische Statistik, seit es empirische Statistik gibt (Sahner 1979). Sie sind besonders dann gefährlich, wenn die Adressaten der Botschaft von diesem Umfeld und der Vorgeschichte nichts erfahren; wir lesen, daß neun Monate nach einem Stromausfall in X dort die Geburten angestiegen sind, daß Katholiken dümmere sind als Protestanten, daß Knoblauchesser länger leben, daß leitende Manager lieber Fluglinie A als B benutzen, daß die Todesstrafe abschreckt oder auch nicht (je nach Weltanschauung), daß Schwarze krimineller sind als Weiße, daß Chemiefabriken (Ziegelwerke, Starkstromleitungen, Mülldeponien etc.) Leukämie erzeugen, und das alles wissenschaftlich abgesichert und hoch signifikant. Wir lesen nicht, wie viele andere Studien und Stichproben ohne signifikante Resultate es außerdem gegeben hat. Wir lesen nicht, in wieviel Studien Katholiken genauso reich oder reicher sind als Protestanten, oder leitende Manager lieber Linie B als Linie A benutzen, oder die Todesstrafe nicht abschreckt, oder Industriebetriebe keine Leukämie erzeugen, und ehe wir das nicht wissen, läßt sich auch die "Signifikanz" der ausgewählten Resultate nicht ermesen.

Insbesondere sind auch die immer wieder gern gedruckten Schreckensmeldungen über die „signifikante“ Häufung von Krankheiten aller Art in der Nähe von Atomkraftwerken (von Dewdney (1993) als "numerischer Terrorismus" bezeichnet), bis auf weiteres als statistische Manipulationen abzutun (Becker und Wahrendorf 1991): Unter rund 400 Teilstichproben - so viele Kernkraftwerke gibt es derzeit auf der Welt - wird diejenige herausgesucht, die am besten in das vorgeprägte Weltbild paßt, und daß dann na-

türlich "signifikante" Abweichungen nur so vom Fließband fallen, ist alles andere als überraschend.

6. Fazit und Ausblick

Die hier exemplarisch dargestellten Methoden zur falschen Darstellung wirtschaftlicher und wissenschaftlicher Sachverhalte mit Hilfe der Statistik bilden natürlich nicht das ganze Missbrauchs-Spektrum ab. Weitere Möglichkeiten zur Täuschung des Publikums eröffnen sich durch synthetische Superlative, verzerrende Graphiken, ungerechtfertigt extrapolierte Trends oder kontraindizierte Mittelwerte und sind in Krämer (2003a, 2003b) nachzulesen. Gemeinsam ist allen, daß sie vergleichsweise einfach zu durchschauen sind. Von dem Mißbrauch des Signifikanzbegriffes abgesehen ist kein Expertenwissen nötig, um dergleichen Lügen zu entlarven, so daß die Hoffnung, sie möchten eines Tages aus Mangel an Erfolg dem Aussterben verfallen, nicht ganz von der Hand zu weisen ist.

Literatur

- Brixy, Udo; Gilberg, Rainer; Hess, Doris; Schröder, Helmut (2002): „Wie nah am Arbeitsmarkt sind die Arbeitslosen?“ IAB Kurzbericht, 21. Januar.
- Bröll, Claudia (2000): „Das Wunder in den Aktenordnern ,“ Frankfurter Allgemeine Zeitung vom 6. Juli.
- Bundesministerium für Bildung und Wissenschaft (Hrsg.): (1993): Bericht der Bundesregierung zur Bekämpfung des Analphabetismus in der Bundesrepublik Deutschland. Bonn.
- Central Statistical Office (Hrsg.) (1995): Fighting with Figures, London
- CIA – The World Factbook 2000, im Internet unter <http://www.cia.gov/cia/publications/factbook/geos/ch.html>
- Downs, Robert B. (1987): Scientific Enigmas, Englewood (Libraries Unlimited).
- Dewdney, A. K. (1993): 200% of nothing, New York (Wiley).
- Koos, E. L. (1970): „Krankheit in Regionville,“ in Mitscherlich et al. (Hrsg): Der Kranke in der modernen Gesellschaft, Köln.
- Krämer, Walter (2000): Armut in der Bundesrepublik – Zur Theorie und Praxis eines überforderten Begriffs, Frankfurt (Campus).
- Krämer, Walter (2003a): So lügt man mit Statistik (4. Auflage), München (Piper).
- Krämer, Walter (2003b): Denkste! Trugschlüsse aus der Welt des Zufalls und der Zahlen, (5. Auflage), München (Piper).
- Pitkin, Walter B. (1932): A short introduction to the history of human stupidities, New York.
- Morgenstern, O.: (1963): On the accuracy of economic observations, Princeton.
- Sahner, Heinz (1979): "Veröffentlichte empirische Sozialforschung: eine Kumulation von Artefakten?" Zeitschrift für Soziologie 8, 1979, 267-278.
- Schneider, Wolf (1999): Wörter machen Leute, 8. Auflage, München (Piper)
- Sorrentino, C. (2000): „International unemployment rates: how comparable are they?“ Monthly Labor Review, Juni, 3-20.
- Taylor, C. L. und Jodice, D.: Handbook of Political and Social Indicators, Vol. 1, 3. Auflage, New Haven 1983.