

Kalibrierung

Wie man gewichtet, was man weiß. Eine Erweiterung des Buchs über die Wahrheit

Kai Froeb

Autor
Kai Froeb
https://kai.froeb.net

Verlag
Reflexivity Press
São Martinho do Porto
Edition Kai Froeb

Neueste Fassung und Hinweise auf Korrekturen:
https://reflexivitypress.com/de/buecher/froeb-kalibrierung
Rückmeldungen: feedback+de.kalibrierung.v1.0@reflexivitypress.com
Zum Cover: https://reflexivitypress.com/de/ueber-uns/bildsprache/

Veröffentlichung
2026-11-17
Version 1.0
Typeset with LaTeX (LuaHBTeX, Pandoc + Eisvogel)

Hinweis zur Entstehung
Geschrieben mit dem Sprachmodell Claude (Anthropic). Anlage, Inhalt und Verantwortung liegen beim Autor. Wie dabei gearbeitet wurde, beschreibt der Aufsatz „Philosophie im Zeitalter der Maschinen, die schreiben“:
https://reflexivitypress.com/de/ueber-uns/maschinen-die-schreiben/

© Copyright 2026 by Reflexivity Press, São Martinho do Porto.
Alle Rechte vorbehalten.

Zero Marginal Cost License 1.3 (ZMC 1.3)

Kurzfassung für Leserinnen und Leser

Was keine Grenzkosten hat, soll keine Barrieren haben.

Sie dürfen

- Das Werk **kostenlos lesen, herunterladen und kopieren**.
- Das **inhaltlich unveränderte Werk nichtkommerziell digital weitergeben**, mit Autor-, Quellen- und Lizenzangabe.
- Das Dateiformat technisch ändern und notwendige Anpassungen zur Barrierefreiheit vornehmen, ohne den Inhalt zu verändern.
- Für den persönlichen, nichtkommerziellen Gebrauch **einzelne Exemplare drucken oder drucken lassen und verschenken**; unmittelbare Kosten dürfen ohne Gewinnaufschlag erstattet werden.
- Für den eigenen persönlichen, nichtöffentlichen und nichtkommerziellen Gebrauch **Übersetzungen anfertigen**.
- Das Werk individuell für Bildung und Forschung nutzen und bei nichtkommerzieller Nutzung nach wissenschaftlichen Standards zitieren, ohne es durch eine Zitatsammlung zu ersetzen.

Bedingungen

Die Nutzung muss **nichtkommerziell** bleiben. Bei öffentlicher Weitergabe sind, soweit vernünftigerweise möglich, **Autor, Werkquelle und Lizenz** anzugeben. Sie dürfen keine zusätzlichen rechtlichen oder technischen Beschränkungen auferlegen und keine Unterstützung durch Autor oder Verlag vortäuschen.

Alle übrigen Rechte bleiben vorbehalten

Diese Lizenz erlaubt ausschließlich die ausdrücklich genannten Nutzungen. Jede darüber hinausgehende Nutzung bedarf einer gesonderten Erlaubnis des Rechteinhabers, soweit sie nicht bereits gesetzlich erlaubt ist.

Gesetzliche Nutzungsbefugnisse und Vergütungsansprüche bleiben unberührt. Rechte Dritter können gesonderten Bedingungen unterliegen.

Wichtiger Hinweis: Diese Kurzfassung dient nur der Orientierung und ist kein eigenständiger Lizenzvertrag. Maßgeblich ist die vollständige **ZMC 1.3**:

https://reflexivitypress.com/de/zmc/



Über den Autor

Kai Froeb (*1963) lebt in São Martinho do Porto an der portugiesischen Atlantikküste.

Wenn Sie dieses Buch lesen, wissen Sie schon das Wichtigste über den Autor: Er bittet Sie nicht, ihm zu glauben. Alle seine Bücher sind so gebaut, dass sich am eigenen Denken prüfen lässt, was in ihnen steht. Wo Ihr Befund vom Text abweicht, ist entweder eine Grenze des Buchs erreicht oder eine Eigenheit Ihres Denkens sichtbar geworden. Beides ist mehr wert als Zustimmung.

Der Weg dahin ging über zwei Spuren, die lange nebeneinanderher liefen. Die eine begann in der Oberstufe in Lüneburg und in den Sommerschulen der Heimvolkshochschule Göhrde, wo drei Wochen lang ein einziger Text durchgearbeitet wurde; sie führte über Mathematik, Informatik, Betriebswirtschaft, Kulturwissenschaft und Jura, bewusst ohne Abschlüsse, um sich einen Weg quer zu den Fächern offenzuhalten. Die andere begann 1980 an einem Commodore CBM 8032 mit einer Nachbildung von Weizenbaums ELIZA und wurde ein Beruf: Programmierung, Anwendungsbetreuung, Projektleitung und Schulung, erst in kleinen Softwarehäusern, dann in großen IT-Konzernen, zuletzt über zehn Jahre Betriebsrat und Gewerkschaft.

Dass beide Spuren zusammengehören, zeigte sich spät. Was in vierzig Jahren Softwarearbeit zur Gewohnheit wird — Strukturen zerlegen, Abhängigkeiten sichtbar machen, prüfen, ob etwas trägt, wenn man daran zieht —, ist dasselbe Handwerk, das ein philosophisches Argument braucht.

Seit dem Erscheinen von ChatGPT Ende 2022 arbeitet er regelmäßig und intensiv mit Sprachmodellen, beim Programmieren wie an seinen Büchern; vorher war eine solche Arbeit nicht möglich. Die Arbeit ist in beiden Fällen gleich geteilt: Anlage, Fragestellung und das Urteil in der Sache liegen bei ihm, ebenso die Regeln, nach denen geprüft wird; das Modell arbeitet ihm zu, sucht Einwände und formuliert aus. Verantwortlich ist er für jeden Satz. Wie ein Text entstanden ist, entscheidet nicht darüber, ob er gilt; das prüft man am Text selbst, und dafür sind seine Bücher gebaut.

Seit 1998 gibt Froeb **hegel.net** heraus, eine Hegel-Ressource in zwölf Sprachen, samt einem öffentlichen Archiv von über fünfhundert digitalisierten Originalausgaben Hegels und der Hegelianer des 19. Jahrhunderts.

Sein Leitmotiv stammt aus jener Schulzeit und hat sich seither als brauchbar erwiesen: „Das Wahre ist das Ganze." Der Satz verlangt eine Integration, die nichts verdrängen muss — und ist deshalb schwerer einzulösen, als er klingt.

Wo Sie in diesem Buch etwas finden, das nicht stimmt, ist Ihre Kritik der wertvollste Beitrag, den Sie ihm leisten können.

INHALTSVERZEICHNIS

Vorwort 1

Teil I: Was fehlt 3
Kapitel 1: Die Karte und das Land . . . 3
1.1 Der Platz in Bagdad . . . 3
1.2 Was jede Darstellung weglässt . . . 4
1.3 Der Elefant, von dem ein Teil fehlt . . . 5
1.4 Was Kalibrieren heißt . . . 5
Zusammenfassung . . . 6
Kapitel 2: Das Fehlende und das falsche Gewicht . . . 6
2.1 Eine Lücke meldet sich, wenn man sie braucht . . . 6
2.2 Ein Gewicht steckt in keinem Teil . . . 7
2.3 Wo man wenig weiß, wiegt der Zufall . . . 8
2.4 Zwei Prüfungen . . . 9
Zusammenfassung . . . 9

Teil II: Zahlen 11
Kapitel 3: Die Grundrate . . . 11
3.1 Ein Test . . . 11
3.2 Die Beschreibung verdrängt die Zahl . . . 12
3.3 Welche Gruppe? . . . 12
3.4 Woher man eine Grundrate bekommt . . . 13
Zusammenfassung . . . 13
Kapitel 4: Nenner und Maßstab . . . 14
4.1 Doppelt so hoch . . . 14
4.2 Zusammengezählt . . . 15
4.3 Wo die Schwelle liegt . . . 15
4.4 Wer im Nenner steht . . . 16
Zusammenfassung . . . 17
Kapitel 5: Was gezählt wird . . . 18
5.1 Zwei Millionen Antworten . . . 18
5.2 Was veröffentlicht wird . . . 18
5.3 Zwei Zählweisen im Krieg . . . 19
5.4 Die Frage an jede Zahl . . . 20
Zusammenfassung . . . 21

Teil III: Fremde Länder **23**
Kapitel 6: Was man aus der Ferne weiß . . . 23
6.1 Jeder Kanal hat seinen Filter . . . 23
6.2 Erzählungen sind gefiltert . . . 23
6.3 Was sich nicht zeigen lässt . . . 24
6.4 Was einmal galt . . . 24
6.5 Aus der Ferne kalibrieren . . . 25
Zusammenfassung . . . 25
Kapitel 7: Vor Ort und die Regionen . . . 26
7.1 Was nur vor Ort mitkommt . . . 26
7.2 Land, Region, Stadt, Viertel . . . 27
7.3 Einen Aufenthalt kalibrieren . . . 28
Zusammenfassung . . . 28

Teil IV: Gewicht als Mittel **31**
Kapitel 8: Propaganda . . . 31
8.1 Die Lüge und ihr Preis . . . 31
8.2 Die Regel: Auswahl . . . 32
8.3 Ausgewogenheit als Inszenierung . . . 33
8.4 Worüber man nachdenkt . . . 33
Zusammenfassung . . . 34
Kapitel 9: Überflutung . . . 35
9.1 Menge statt Überzeugung . . . 35
9.2 Was die Menge mit dem Maß macht . . . 35
9.3 Was das Maß stabil hält . . . 36
Zusammenfassung . . . 37
Kapitel 10: Kognitive Kriegsführung . . . 37
10.1 Das Sicherheitsargument . . . 37
10.2 Zwei Wege . . . 39
10.3 Warum Steuerung schwächer verteidigt . . . 39
10.4 Warum Kalibrierung besser verteidigt . . . 40
10.5 Die Grenze im Ernstfall . . . 41
Zusammenfassung . . . 42

Teil V: Gemeinsam gewichten **43**
Kapitel 11: Gremien, Fachleute, Konsens . . . 43
11.1 Achthundert Schätzungen . . . 43
11.2 Das Wissen im Raum . . . 44
11.3 Der bestellte Widerspruch . . . 44
11.4 Konsens worüber . . . 45
Zusammenfassung . . . 46
Kapitel 12: Künstliche Intelligenz . . . 46
12.1 Was geschrieben wurde . . . 46
12.2 Die Neigung, zuzustimmen . . . 47

12.3 Ein Kanal für viele 47
12.4 Wofür es taugt 48
Zusammenfassung 49

Teil VI: Selbst kalibrieren **51**
Kapitel 13: Die eigene Waage 51
13.1 Wann ein Urteil kalibriert ist 51
13.2 Eine Probe 51
13.3 Buch führen 52
13.4 Die Symmetrieprüfung 53
13.5 Die stärkste Gegenseite 53
Zusammenfassung 54
Kapitel 14: Gemeinsam kalibrieren 54
14.1 Warum es mehrere Stimmen braucht 54
14.2 Strittiges selbst gewichten 55
14.3 Keine falsche Mitte 56
14.4 Mit anderen gewichten 56
Zusammenfassung 57
Kapitel 15: Verkehrte Kalibrierung 57
15.1 Wissen, das in beide Richtungen wirkt 57
15.2 Die Fragen dieses Buchs, verkehrt 58
15.3 Woran man die Verkehrung erkennt 59
15.4 Wie man sich und andere schützt 59
Zusammenfassung 60
Kapitel 16: Grenzen 60
16.1 Wo keine Rückmeldung kommt 60
16.2 Das Fehlende, das sich nicht zeigt 61
16.3 Kein Hauptschalter, kein Abschluss 61
16.4 Dieses Buch als Auswahl 62
16.5 Der Schritt zurück 62
Zusammenfassung 63

VORWORT

Wer eine Zeitung liest, eine Reportage sieht oder einem Freund zuhört, der von einer Reise zurückkommt, bekommt meist Wahres zu hören. Trotzdem kann das Bild schief sein, das er daraus gewinnt. Es fehlt etwas, oder etwas wiegt zu viel, und keiner der Sätze, aus denen das Bild entstand, war falsch. Dieses Buch handelt davon, wie man damit umgeht: wie man das Gewicht der Teile prüft, aus denen man sich ein Bild macht, und das Fehlende sucht, das man nicht sieht.

Das Buch erweitert ein früheres über die Wahrheit, das diese Frage in einem Kapitel behandelt. Dort ging es darum, dass wir prüfen, wie wir prüfen und warum das Prüfen frei macht. Die Kalibrierung war darin der schwerste Fall: das Prüfen dessen, was gar nicht dasteht. Hier wird sie ausgeführt, an den Zahlen, an fremden Ländern, an der Propaganda, an Einrichtungen und an der eigenen Urteilskraft. Das frühere Buch braucht man dafür nicht. Was dieses von ihm übernimmt, steht in einem Satz an der Stelle, wo es gebraucht wird; wo das andere Buch einen Gedanken weiter ausführt, sagt das eine Zeile am Ende des Abschnitts.

Die Beispiele kommen aus vielen Ländern, und keines ist nötig, um den Gedanken zu verstehen; sie zeigen ihn an Wirklichem und sollen darum stimmen. Ihre Belege stehen in den Fußnoten. Wo eine Sache strittig ist, steht die stärkste Fassung jeder Seite da und der Weg, auf dem man selbst gewichten kann. Am Ende jedes Kapitels steht eine Zusammenfassung. Wer nur sie liest, kennt den Gang des Buchs; wer an einer Stelle zweifelt, findet davor die Gründe.

Weiterlesen: *Wahrheit* (2025)

TEIL I: WAS FEHLT

Kapitel 1: Die Karte und das Land

1.1 Der Platz in Bagdad

Am 9. April 2003, drei Wochen nach dem Beginn des Irakkriegs, wurde auf dem Firdos-Platz in Bagdad eine große Statue Saddam Husseins vom Sockel geholt. Fernsehsender in aller Welt übertrugen es, viele stundenlang und immer wieder. Auf den Bildern sah man jubelnde Männer um den Sockel, und wer sie sah, verstand: Bagdad feiert seine Befreiung. Die meisten Bilder waren Nahaufnahmen. Auf den seltenen Totalen sah man einen Platz, der zum größten Teil leer war. Ein Reporter, der die Aufnahmen und die Berichte der Beteiligten später durchging, schätzte die Menge auf höchstens einige hundert Menschen, von denen ein Viertel bis die Hälfte Journalisten und Soldaten waren. Heruntergezogen hatte die Statue ein Bergepanzer der amerikanischen Marineinfanterie.[1]

Nichts an den Nahaufnahmen war gefälscht. Die Statue fiel, die Männer jubelten, und manche von ihnen mochten unter dem Gestürzten gelitten haben. Falsch wurde erst, was der Zuschauer daraus machte: dass die Stadt sich auf diesem Platz versammelt habe. Den Schritt hatte die Kamera vorbereitet, ohne etwas zu behaupten. Sie zeigte einen Ausschnitt, und weil das Bild randvoll war, schien es keinen Rand zu haben. Je weiter man sich vom Sockel entfernte, desto mehr sah man und desto weniger eindrucksvoll war, was man sah.

Jede Kamera hat einen Ausschnitt, auch die beste, und wer den Ausschnitt wählt, entscheidet nicht über die Wahrheit des Bildes, sondern über sein Gewicht. Ob es um einen Platz geht oder um ein Land, um eine Krankheit oder um einen Krieg: Was man sieht, ist ein Teil, und der Teil sagt nicht von sich aus, wie groß das Ganze ist, aus dem er stammt.

1 Peter Maass, „The Toppling. How the Media Inflated a Minor Moment in a Long War“, *ProPublica*, 3. Januar 2011, gemeinsam veröffentlicht mit *The New Yorker*: „Very few Iraqis were there“; die Menge „does not seem to have exceeded several hundred at its largest“, davon ein Viertel bis die Hälfte Journalisten oder Marines; Nahaufnahmen erzeugten „an illusion of wall-to-wall enthusiasm“, auf den seltenen Totalen sei der Platz „mostly empty“ gewesen.

1.2 Was jede Darstellung weglässt

Man könnte meinen, das Übel sei der Ausschnitt, und eine ehrliche Darstellung zeige eben alles. Das geht nicht; es wäre auch nicht wünschenswert. Eine Landkarte, die nichts wegließe, müsste so groß sein wie das Land; sie würde es zudecken und nützte niemandem. Eine Straßenkarte lässt die Höhenlinien weg, eine Wanderkarte die Tankstellen, eine Wetterkarte fast alles außer Druck, Wolken und Wind, und gerade deshalb taugt jede für ihren Zweck. Die Verdichtung ist ihre Leistung. Fehler werden daraus erst, wenn das Weggelassene für die Frage, die man an die Karte stellt, gerade das Wichtige gewesen wäre.

Dasselbe gilt für jede Messung. Ein Thermometer misst die Luft an einer Stelle, nicht im Tal und nicht auf dem Berg. Eine Statistik zählt, was erfasst wurde: die gemeldeten Fälle, nicht die vorgekommenen, die Antworten derer, die geantwortet haben, nicht die Meinung aller. Und schon die Frage, mit der man etwas untersucht, legt fest, was man finden kann. Wer fragt, wie viele Unfälle eine Straße hat, erfährt nichts über die Unfälle, die auf der Umgehungsstraße geschehen, weil Autofahrer die gefährliche meiden.

Das Weggelassene liegt also nicht an einem Ort, an dem man nachsehen könnte. Es liegt außerhalb des Bildausschnitts, außerhalb der Messung, außerhalb der Frage, außerhalb dessen, was bisher überhaupt festgestellt wurde. Man sieht es nicht, und darum sieht man auch nicht, dass es fehlt. Ein leerer Fleck auf der Karte fällt auf. Was gar nicht erst eingezeichnet ist, nicht.

Lewis Carroll hat den Gedanken 1893 in einen Scherz gefasst: Ein Gelehrter berichtet, man habe in seinem Land eine Karte im Maßstab eine Meile auf die Meile gezeichnet, sie aber nie ausgebreitet, weil die Bauern fürchteten, sie werde das Land zudecken und das Sonnenlicht aussperren; nun nehme man das Land selbst als seine Karte, und es tue es fast ebenso gut.[2] Jorge Luis Borges hat daraus später eine kurze Prosa gemacht, in der die Kartographen eines Reichs eine Karte zeichnen, die so groß ist wie das Reich.[3]

2 Lewis Carroll, *Sylvie and Bruno Concluded*, London 1893, Kap. 11 („The Man in the Moon"): „We actually made a map of the country, on the scale of a mile to the mile!" – „the farmers objected: they said it would cover the whole country, and shut out the sunlight!" – „So we now use the country itself, as its own map, and I assure you it does nearly as well."

3 Jorge Luis Borges, „Del rigor en la ciencia"; zur Stellung dieses Texts Ivan Almeida, „Borges à la carte (tres citas de Baudrillard)", *Variaciones Borges* 25 (2008).

1.3 Der Elefant, von dem ein Teil fehlt

Eine alte Geschichte erzählt von sechs Blinden, die einen Elefanten betasten. Der eine fasst das Bein und sagt, ein Elefant sei wie eine Säule, der zweite den Rüssel und sagt, er sei wie eine Schlange, der dritte das Ohr und sagt, er sei wie ein Fächer, und so fort. Sie streiten, bis sich zeigt, dass jeder etwas Wahres ertastet hat und erst alle Berichte zusammen das Tier ergeben. Die Geschichte wird meist als Lehre über verschiedene Perspektiven erzählt, die einander ergänzen.

Das Problem dieses Buchs ist dasselbe, mit einem Unterschied: Hier fehlt etwas. Angenommen, nur drei der sechs kommen zurück: die mit dem Stoßzahn, dem Bein und dem Rüssel. Ihre Berichte stimmen, sie widersprechen einander nicht, und wer sie zusammenfügt, erhält ein Tier aus Speer, Säule und Schlange, das bedrohlicher ist als der Elefant. Das Ohr, der Bauch und der Schwanz fehlen, und niemand vermisst sie, weil niemand weiß, dass es sie gibt.

In der Geschichte weiß jeder, dass er nur einen Teil berührt hat; darum kann er sich mit den anderen verständigen. Im Alltag gibt sich der Teil meist als Ganzes. Die Nahaufnahme vom Firdos-Platz sagte nicht, dass sie eine Nahaufnahme war, und die Reportage über ein Viertel sagt nicht, dass die anderen Viertel anders sind. Das macht das Ergänzen schwerer, als die Geschichte es zeigt. Man muss erst darauf kommen, dass es etwas zu ergänzen gibt.

Weiterlesen: *Wahrheit* (2025), Kap. 5 (das Elefantengleichnis), *Denken – Ordnen, ohne zu verlieren* (2025), Kap. 2 und 14 (Verdichten und was dabei verloren geht)

1.4 Was Kalibrieren heißt

In der Messtechnik heißt kalibrieren, die Anzeige eines Geräts mit einem bekannten Maß zu vergleichen: Man legt ein Gewicht, das man kennt, auf die Waage und sieht nach, wie weit sie abweicht. Ein Thermometer prüft man an zwei Zuständen, deren Temperatur feststeht, am schmelzenden Eis und am kochenden Wasser, und liest ab, um wie viel es an beiden Stellen danebenliegt. Zwei Dinge gehören dazu. Man braucht ein Maß, dem man mehr traut als dem Gerät. Und das Gerät kann sich nicht an seinen eigenen Anzeigen prüfen; eine Waage, die nur misst, was sie selbst gewogen hat, findet ihre Abweichung nie.

Im Urteilen ist es ebenso. Jeder bringt ein Maß dafür mit, was häufig ist und was selten, was gewöhnlich und was außergewöhnlich, was wichtig und was nebensächlich, was die meisten denken und was nur wenige. Dieses Maß ist kein Satz, den man vertritt, sondern ein Gefühl für Größen, und es hat sich aus allem abgelagert, was man gesehen, gelesen und gehört hat. Kalibrieren heißt hier, dieses Maß an etwas zu prüfen, das nicht aus demselben Strom stammt: an

einer Zählung, an einem zweiten, unabhängigen Weg, an eigener Anschauung, an einem Fall, dessen Ausgang man später erfährt. Wo es abweicht, stellt man es nach.

Mit dem Prüfen einzelner Aussagen ist das nicht erledigt. Wer jede Meldung über den Firdos-Platz einzeln prüfte, fände sie alle bestätigt. Was zu prüfen wäre, ist die Größe, die er ihnen zusammen gibt. Darum richtet sich die Kalibrierung auch auf das, was nicht dasteht, und darum ist sie schwerer als die Prüfung des Gesagten.

Wer fragt, was ihm fehlt, hat aber schon den leichteren Teil der Aufgabe vor sich. Schwerer ist, was man gar nicht als Lücke bemerkt: das falsche Gewicht dessen, was man durchaus vor Augen hat.

Weiterlesen: *Wahrheit*, Kap. 19 (der Kern dieses Buchs in einem Kapitel)

Zusammenfassung

Jede Darstellung lässt etwas weg, und das ist keine Schwäche, sondern die Bedingung ihres Nutzens: Eine Karte, die alles enthielte, wäre so groß wie das Land. Kamera, Karte, Messung und Frage zeigen einen Ausschnitt, und der Ausschnitt sagt nicht von sich aus, wie groß das Ganze ist. Am Firdos-Platz 2003 zeigten die Nahaufnahmen Wahres und ließen die Welt glauben, eine ganze Stadt feiere, wo höchstens einige hundert Menschen standen. Das Problem gleicht dem der Blinden, die einen Elefanten betasten, nur dass hier ein Teil fehlt und niemand weiß, dass es ihn gibt; im Alltag gibt sich der Teil zudem als Ganzes. Kalibrieren heißt, wie in der Messtechnik, das eigene Maß für Häufigkeit, Größe und Gewicht an etwas zu prüfen, das nicht aus demselben Strom kommt, und es nachzustellen, wo es abweicht. Es richtet sich darum auch auf das, was nicht dasteht.

Kapitel 2: Das Fehlende und das falsche Gewicht

2.1 Eine Lücke meldet sich, wenn man sie braucht

Wer gefragt wird, wovon die Menschen in der Mongolei leben, welche Musik man in Ghana hört oder wie in Peru gewählt wird, merkt oft in diesem Augenblick, dass er nichts darüber weiß. Die Lücke war vorher da, aber sie hat sich nicht gemeldet. Jetzt meldet sie sich, weil eine Frage auf sie trifft und keine Antwort findet. Was dann zu tun ist, ist nicht schwer: nachlesen, jemanden fragen, der

dort war, eine Zahl suchen. Eine bemerkte Lücke ist eine Aufgabe, und für die meisten Aufgaben dieser Art gibt es Wege.

Schwerer zu bemerken ist eine Lücke in dem, was man untersucht. Im Zweiten Weltkrieg fragte das amerikanische Militär Statistiker, wie verwundbar seine Bomber an welchen Stellen seien. Die Daten kamen von den Maschinen, die zurückkehrten, mit Einschüssen an Rumpf und Flügeln. Nur zeigen gerade diese Maschinen, wo Treffer zu überstehen sind. Die anderen, die Treffer an den empfindlichen Stellen bekommen hatten, waren nicht zurückgekehrt und kamen in den Daten nicht vor. Wer nur zählt, was vor ihm steht, verstärkt die Stellen, an denen schon jetzt nichts Schlimmes geschieht. Den Ausweg fand, wer fragte, woher die Daten stammen und was es bedeutet, dass ein Teil der Flugzeuge in ihnen fehlt.

Die Geschichte wird meist dem Statistiker Abraham Wald zugeschrieben und in ihrer bekanntesten Form ausgeschmückt. Belegt ist, dass Wald 1943 in einer Reihe von Denkschriften rechnete, wie man aus den Schäden zurückgekehrter Maschinen auf die Verwundbarkeit aller schließen kann. Die Anekdote, er habe die Panzerung dorthin legen lassen, wo die heimkehrenden Flugzeuge keine Löcher hatten, steht in diesen Schriften nicht.[4] Am Gedanken ändert das nichts: Die Lücke wurde erst sichtbar, als jemand nach dem fragte, was im Material fehlte, nicht nach dem, was darin stand.

2.2 Ein Gewicht steckt in keinem Teil

Wer über ein Land fünf Reportagen gesehen hat, eine über Gewalt in einer Großstadt, eine über einen Streik, eine über eine Dürre, eine über ein Fest und eine über ein Technologieunternehmen, kann jede einzelne prüfen. Ob die Gewalt dort geschah, ob der Streik stattfand, lässt sich feststellen. Was sich an keiner der fünf prüfen lässt, ist ihr Verhältnis zueinander und zu allem, worüber nicht berichtet wurde: ob die Gewalt in diesem Land häufig ist oder selten, ob der Streik die Regel ist oder die Ausnahme. Ein Gewicht ist ein Verhältnis. Es steckt in keinem Teil, und darum findet man es nicht, indem man die Teile einzeln prüft.

Daher kommt der Satz, der diesem Buch zugrunde liegt: Dass etwas ganz fehlt, ist leichter zu finden und zu beheben, als dass das Gezeigte falsch gewichtet

4 Abraham Wald, Denkschriften für die Statistical Research Group der Columbia University (1943), veröffentlicht vom Center for Naval Analyses 1980; Marc Mangel, Francisco J. Samaniego, „Abraham Wald's Work on Aircraft Survivability", *Journal of the American Statistical Association* 79 (1984). Zur Unterscheidung von Beleg und Legende Bill Casselman, „The Legend of Abraham Wald", Feature Column der American Mathematical Society (2016): Die Denkschriften enthalten keine Empfehlung zur Panzerung, „there is very little evidence for the best bits".

ist. Die Lücke meldet sich, wenn eine Frage keine Antwort findet. Das falsche Gewicht meldet sich nicht, denn jede Frage nach einem einzelnen Stück findet ihre Antwort – eine, die stimmt. Man müsste das Ganze vor sich haben, um zu sehen, dass die Stücke darin zu groß oder zu klein geraten sind, und das Ganze hat man nicht; sonst bräuchte man die Stücke nicht.

Woher nimmt man dann überhaupt ein Gewicht? Meist aus der Häufigkeit, mit der einem etwas begegnet ist. Was man oft gehört hat, hält man für häufig; was einem leicht einfällt, für verbreitet. Für vieles im Alltag ist das eine gute Faustregel, denn was häufig vorkommt, begegnet einem auch häufig. Sie versagt, sobald zwischen der Welt und dem eigenen Gedächtnis ein Filter liegt, der nach etwas anderem auswählt als nach der Häufigkeit. Nachrichten wählen nach dem, was neu ist und ungewöhnlich; ein Flugzeugabsturz kommt in die Nachrichten, weil er selten ist. Wer sein Gefühl für Häufigkeit aus den Nachrichten bezieht, misst also mit einem Maß, das das Seltene vergrößert. Die Psychologen Amos Tversky und Daniel Kahneman haben diese Neigung, Häufigkeit danach zu schätzen, wie leicht einem Beispiele einfallen, Verfügbarkeit genannt.[5]

2.3 Wo man wenig weiß, wiegt der Zufall

Wo die eigene Erfahrung dünn ist, füllt sich das Maß mit dem, was gerade da ist. In einem Versuch, den dieselben Forscher beschrieben haben, drehten Versuchsleiter vor den Augen der Teilnehmer ein Glücksrad, das eine Zahl zwischen 0 und 100 ergab. Dann fragten sie, ob der Anteil der afrikanischen Staaten unter den Mitgliedern der Vereinten Nationen höher oder niedriger sei als diese Zahl, und danach, wie hoch er sei. Wer die 10 gesehen hatte, schätzte im Mittel 25 Prozent, wer die 65 gesehen hatte, 45 Prozent.[6] Die Zahl vom Glücksrad hatte mit Afrika nichts zu tun. Jeder wusste das. Sie wog trotzdem, weil sonst wenig da war, das hätte wiegen können.

Dass der Versuch ausgerechnet nach Afrika fragt, passt zu ihm. Über einen ganzen Kontinent mit mehr als fünfzig Staaten haben viele Menschen außerhalb nur ein grobes Bild, und ein grobes Bild hat kein eigenes Gewicht, das gegen eine zufällige Zahl ankäme. Was für den Anteil an den Vereinten Nationen gilt, gilt

5 Amos Tversky, Daniel Kahneman, „Judgment under Uncertainty: Heuristics and Biases“, *Science* 185 (1974), S. 1124–1131, hier S. 1127: „one may assess the risk of heart attack among middle-aged people by recalling such occurrences among one's acquaintances … This judgmental heuristic is called availability.“

6 Tversky, Kahneman, „Judgment under Uncertainty“, *Science* 185 (1974), S. 1128: „the median estimates of the percentage of African countries in the United Nations were 25 and 45 for groups that received 10 and 65, respectively, as starting points.“

für die Armut, die Gewalt, die Lebenserwartung, das Wirtschaftswachstum eines fernen Landes ebenso. Die erste Zahl, die einem begegnet, die eindrücklichste Geschichte, das letzte Bild werden zum Maß, bis etwas Besseres an ihre Stelle tritt.

Experiment 1: Woher Ihre Gewichte kommen

Denken Sie an ein Land, in dem Sie nie waren und über das Sie trotzdem etwas zu wissen meinen. Schreiben Sie drei Dinge auf, die dort nach Ihrem Eindruck häufig oder typisch sind. Fragen Sie dann bei jedem: Woher habe ich das, aus einem Film, aus den Nachrichten, von jemandem, der dort war? Wie oft ist es mir begegnet, und auf wie vielen verschiedenen Wegen? Und hätte ich davon erfahren, wenn es dort selten wäre?
Vermutlich stammen die drei aus wenigen Quellen, und bei mancher wird die letzte Frage unbequem. Ein Film zeigt das Ungewöhnliche, weil es eine Geschichte trägt, und eine Nachricht berichtet, was vom Gewöhnlichen abweicht. Was Sie für typisch halten, kann typisch sein. Ihr Grund dafür ist aber oft nur, dass es Ihnen oft begegnet ist, und darüber hat ein Filter entschieden, der nach etwas anderem fragt als danach, wie häufig etwas ist.

2.4 Zwei Prüfungen

Darum gibt es zwei Prüfungen, die verschieden fragen. Die erste fragt bei jeder Aussage, ob sie stimmt: wer sie macht, woher er es weiß, ob unabhängige Quellen dasselbe sagen. Gegen die Lüge ist sie das richtige, lange erprobte Werkzeug. Die zweite fragt nach der Verteilung: ob das Gezeigte repräsentativ ist, was es nicht zeigt und warum gerade dies vor einem liegt. Sie braucht etwas, das die erste nicht braucht, nämlich ein Maß für das Ganze, an dem sich die Teile messen lassen. Woher man ein solches Maß nimmt, wenn man das Ganze nicht vor sich hat, lässt sich am deutlichsten an Zahlen fragen.

Weiterlesen: *Wahrheit* (2025), Kap. 19 (die erste und die zweite Ebene)

Zusammenfassung

Dass etwas ganz fehlt, meldet sich, sobald eine Frage keine Antwort findet; dann kann man nachlesen, fragen, zählen. Schwerer zu finden ist eine Lücke im Material selbst, wie bei den Bombern des Zweiten Weltkriegs, deren Treffer man nur an den

zurückgekehrten Maschinen sah. Am schwersten ist das falsche Gewicht, denn ein Gewicht ist ein Verhältnis: Es steckt in keinem einzelnen Stück, und wer jedes Stück prüft, findet jedes bestätigt. Gewichte entstehen meist aus der Häufigkeit, mit der einem etwas begegnet. Das trägt, solange zwischen Welt und Gedächtnis kein Filter liegt, der nach etwas anderem auswählt; Nachrichten wählen das Ungewöhnliche. Wo man wenig weiß, füllt der Zufall das Maß, wie die Zahl vom Glücksrad, die die Schätzung über Afrika verschob. Darum braucht es neben der Prüfung der einzelnen Aussage eine zweite, die nach der Verteilung fragt, und sie braucht ein Maß für das Ganze.

TEIL II: ZAHLEN

Kapitel 3: Die Grundrate

3.1 Ein Test

Experiment 2: Ein positiver Befund

Eine Krankheit kommt in einer Bevölkerung bei einem von tausend Menschen vor. Ein Test erkennt sie bei 99 von 100 Kranken. Bei Gesunden schlägt er in 5 von 100 Fällen fälschlich an. Jemand aus dieser Bevölkerung wird ohne besonderen Anlass getestet, und der Test ist positiv. Wie wahrscheinlich ist es, dass er die Krankheit hat? Schätzen Sie, bevor Sie weiterlesen.
Rechnen Sie es dann an hunderttausend Menschen nach. Hundert von ihnen sind krank, und bei 99 davon schlägt der Test an. 99.900 sind gesund, und bei fünf von hundert, also bei 4.995, schlägt er ebenfalls an. Positiv sind damit 5.094 Menschen, und krank sind von ihnen 99. Das sind knapp zwei Prozent.

Wer neunzig Prozent oder mehr geschätzt hat, ist in großer Gesellschaft; bei solchen Aufgaben schätzen viele weit zu hoch, solange die Angaben als Prozente gegeben werden. Wer dieselben Angaben als Häufigkeiten bekommt, als „hundert von hunderttausend", kommt öfter auf die richtige Größe.[1] Die Rechnung ist in beiden Fassungen dieselbe. Verschieden ist, ob sie die Grundrate sichtbar lässt.

Die Grundrate ist hier der Anteil der Kranken in der Gruppe, aus der der Getestete stammt, einer von tausend. Sie entscheidet, was der Befund wert ist. Die 99 Prozent des Tests drängen sich auf, weil sie zum Befund gehören; die eins von tausend steht daneben wie eine Nebensache. Dabei wiegt sie mehr. Bei einer seltenen Krankheit sind auch wenige Fehlalarme unter vielen Gesunden zusammen viel mehr als die wenigen richtigen Treffer unter den Kranken. Das Gewicht des Befunds hängt also an einer Zahl, die im Befund selbst nicht vorkommt.

1 Gerd Gigerenzer, Ulrich Hoffrage, „How to Improve Bayesian Reasoning without Instruction: Frequency Formats", *Psychological Review* 102 (1995), S. 684–704.

3.2 Die Beschreibung verdrängt die Zahl

In einem anderen Versuch bekamen Teilnehmer kurze Beschreibungen von Menschen, die angeblich aus einer Gruppe von Ingenieuren und Juristen gezogen waren. Den einen wurde gesagt, die Gruppe bestehe aus 70 Ingenieuren und 30 Juristen, den anderen das Umgekehrte. Sie sollten schätzen, wie wahrscheinlich der Beschriebene Ingenieur sei. Ohne Beschreibung gingen die Teilnehmer richtig von der Zusammensetzung der Gruppe aus. Mit einer Beschreibung kam die Zusammensetzung in ihren Schätzungen fast nicht mehr vor.[2]

Eine Beschreibung ist anschaulich, eine Grundrate ist eine Zahl ohne Gesicht. Im Gedächtnis hat die Beschreibung den Vorzug, und so verdrängt sie die Zahl, auch wo die Zahl mehr sagt. Das ist dieselbe Bewegung wie bei der Reportage und dem Land aus dem vorigen Kapitel, nur genauer messbar: Der einzelne Fall, lebhaft erzählt, nimmt sich das Gewicht, das der Verteilung zustünde.

3.3 Welche Gruppe?

Eine Grundrate ist immer die Rate in einer Gruppe, und welche Gruppe gemeint ist, steht nicht fest. Kommt jemand mit Beschwerden zum Arzt, die zu der Krankheit passen, so stammt er nicht mehr aus der ganzen Bevölkerung, sondern aus der Gruppe derer, die solche Beschwerden haben. In ihr ist die Krankheit häufiger, vielleicht eine von zwanzig. Derselbe Test mit demselben positiven Ergebnis bedeutet dann etwas ganz anderes: Rechnet man wie oben, so ist von den Positiven nun etwa jeder zweite wirklich krank. Am Test hat sich nichts geändert, nur an der Gruppe, auf die man ihn legt.

Die Wahl der Gruppe ist also selbst eine Gewichtung; eine Regel, die sie für jede Frage entscheidet, gibt es nicht. Man nimmt die engste Gruppe, die dem Fall gleicht und für die es eine verlässliche Zählung gibt. Ist sie zu weit, verschwimmt, was den Fall auszeichnet; ist sie zu eng, gibt es keine Zahl mehr, oder sie beruht auf drei Fällen. Wer über ein Verbrechen liest und fragen will, wie häufig so etwas ist, muss entscheiden, ob er die Rate in einem Land meint, in einer Stadt, in einem Viertel oder unter Menschen eines bestimmten Alters. Jede dieser Zahlen sagt etwas anderes. Strittig wird es dort, wo die Wahl der Gruppe ein Urteil nahelegt, das man vorher schon hatte. Dann lohnt es, alle naheliegenden Gruppen nebeneinanderzustellen und anzusehen, wie weit sie auseinanderliegen.

2 Tversky, Kahneman, „Judgment under Uncertainty“, *Science* 185 (1974), S. 1125: Die Gruppe „consisted of 70 engineers and 30 lawyers“ (bzw. umgekehrt); „prior probabilities were effectively ignored when a description was introduced“.

3.4 Woher man eine Grundrate bekommt

Im Alltag hat man selten eine Zahl zur Hand, und das Gefühl für Häufigkeit, das man stattdessen hat, ist genau das, was zu prüfen wäre. Eine Grundrate gewinnt man darum nicht aus Eindrücken, sondern aus Zählungen: aus amtlicher Statistik, aus Erhebungen, die angeben, wen sie gefragt haben, aus den eigenen Aufzeichnungen über Fälle, deren Ausgang man kennt. Auch eine grobe Zählung ist besser als keine. Wer nur weiß, ob etwas bei einem von zehn oder einem von zehntausend vorkommt, weiß schon, ob ein Befund viel oder wenig wiegt.

Gerade bei Meldungen, die etwas Ungewöhnliches berichten, ist die Grundrate doppelt zu fragen. Wie häufig ist das Ereignis selbst, und wie häufig wird über Ereignisse dieser Art berichtet? Eine Meldung beweist, dass es geschehen ist. Ihre Häufigkeit in den Nachrichten sagt aber nicht, wie oft es geschieht, sondern wie oft es für berichtenswert gehalten wird, und das hängt an seiner Seltenheit, an seinem Schrecken und an der Frage, wem es gerade nützt.

Eine niedrige Grundrate widerlegt dabei nichts. Seltenes kommt vor, und ein positiver Test ist auch bei einer seltenen Krankheit ein Grund, weiter zu prüfen. Sie bestimmt, wie viel ein einzelner Befund verschiebt und wie viel man noch braucht, bevor man sich sicher sein darf. Den Gedanken hat im 18. Jahrhundert Thomas Bayes in eine Formel gebracht, die angibt, wie ein neuer Befund eine vorherige Wahrscheinlichkeit verändert. Ihre Mathematik braucht man im Alltag nicht. Es genügt, die Grundrate nicht wegzulassen.

Schon an der Grundrate zeigt sich, dass eine Zahl erst etwas sagt, wenn man weiß, worauf sie sich bezieht. Dasselbe gilt für jeden Bruch, für jeden Anteil und jeden Vergleich, und dort entscheidet der Nenner.

Weiterlesen: *Wahrheit* (2025), Kap. 19 (Prüfökonomie: Grundrate, Einsatz, Bilanz der Quelle)

Zusammenfassung

Ein Test, der neunundneunzig von hundert Kranken erkennt und bei fünf von hundert Gesunden anschlägt, ergibt bei einer Krankheit, die einen von tausend trifft, unter den Positiven nur etwa zwei Prozent Kranke. Das Gewicht eines Befunds hängt an der Grundrate, die im Befund selbst nicht vorkommt. In Häufigkeiten statt Prozenten gerechnet, bleibt sie sichtbar. Eine anschauliche Beschreibung verdrängt die Grundrate leicht, auch wo sie wenig sagt. Jede Grundrate ist die Rate in einer Gruppe, und die Wahl der Gruppe ist selbst eine Gewichtung: Für einen Patienten mit Beschwerden bedeutet derselbe Test etwas ganz anderes. Man gewinnt Grundraten aus Zählungen, nicht aus Eindrücken; bei Meldungen ist zu

fragen, wie häufig das Ereignis ist und wie häufig darüber berichtet wird. Eine niedrige Grundrate widerlegt nichts; sie bestimmt, wie viel ein einzelner Befund verschiebt.

Kapitel 4: Nenner und Maßstab

4.1 Doppelt so hoch

Am 19. Oktober 1995 warnte die britische Arzneimittelaufsicht Ärzte und Öffentlichkeit, bestimmte neuere Antibabypillen verdoppelten im Vergleich zu älteren das Risiko einer Venenthrombose. Viele Frauen setzten die Pille ab. In den Monaten danach stieg in England und Wales die Zahl der Schwangerschaften um etwa sieben Prozent, die der Abtreibungen um sechs bis neun Prozent.[3]

Die Warnung war nicht erfunden. Die Studien, auf die sie sich stützte, hatten einen Unterschied gefunden, und eine Behörde, die ihn kennt und verschweigt, versäumt ihre Pflicht. Was die Wirkung bestimmte, war die Form. „Doppelt so hoch" ist ein Verhältnis zweier Zahlen, das nichts darüber sagt, wie groß diese Zahlen sind. Eine spätere, genauere Auswertung kam auf etwa 20 Fälle je hunderttausend Frauen im Jahr bei den älteren Präparaten und etwa 40 bei den neueren.[4] Doppelt so hoch, und in absoluten Zahlen ein Unterschied von zwei Fällen unter zehntausend Frauen. Ob das viel ist, lässt sich nur im Vergleich sagen, etwa mit den Folgen einer ungewollten Schwangerschaft, und die Warnung gab diesen Vergleich nicht mit.

Das relative Risiko ist das Gewicht ohne seinen Gegenstand. Es eignet sich, um zwei Möglichkeiten zu vergleichen; es täuscht, wo jemand wissen will, wie viel auf dem Spiel steht. Wer eine solche Meldung liest, fragt darum nach dem Nenner: Doppelt so hoch wie was, bei wie vielen? Wer sie schreibt, schuldet beides, das Verhältnis und die Zahlen, aus denen es gebildet ist. Das gilt in der Gegenrichtung ebenso. Eine Halbierung, die von zwei Fällen auf einen unter zehntausend führt,

3 Emilia Del Bono, Marco Francesconi, Nicky Best, *Health Information and Health Outcomes: An Application of the Regression Discontinuity Design to the 1995 UK Contraceptive Pill Scare Case*, Discussion Paper 696, Department of Economics, University of Essex 2011: Warnung des Committee on Safety of Medicines vom 19. Oktober 1995 („twice the risk of venous thromboembolism"); Zunahme der Empfängnisse um 5,7 bis 7,4 Prozent, der Abtreibungen um 6,4 bis 9,5 Prozent in den folgenden Monaten.

4 D. Hollander, „Most Refined Analysis to Date Confirms Link Between Third-Generation Pills and Venous Thromboembolism", *Perspectives on Sexual and Reproductive Health* 33 (2001), Nr. 3, S. 133–134: rohe Häufigkeiten von „37–41 cases per 100,000 person-years" bei Pillen der dritten Generation und „20–23 per 100,000" bei levonorgestrelhaltigen Pillen.

ist für den Einzelnen wenig, für ein Land mit fünfzig Millionen Menschen sind es fünftausend.

4.2 Zusammengezählt

1973 bewarben sich an der Universität von Kalifornien in Berkeley 8.442 Männer und 4.321 Frauen um einen Studienplatz nach dem ersten Abschluss. Aufgenommen wurden etwa 44 Prozent der Männer und 35 Prozent der Frauen. Das sah nach Benachteiligung aus, und die Universität ließ es untersuchen. Fachbereich für Fachbereich, also dort, wo tatsächlich entschieden wurde, fand sich keine allgemeine Benachteiligung der Frauen; wenn es eine Schieflage gab, dann eher zu Lasten der Männer. Der Unterschied im Ganzen kam daher, dass Frauen sich häufiger in Fächern beworben hatten, die nur wenige Bewerber aufnahmen, Männer häufiger in solchen, die viele aufnahmen.[5]

Beide Zahlen stimmten, die im Ganzen und die in den Teilen, und sie sagten Verschiedenes. Wer zusammenzählt, gewichtet die Teile nach ihrer Größe, und wenn die Gruppen sich verschieden auf die Teile verteilen, kann das Ganze in die entgegengesetzte Richtung zeigen wie jeder Teil. Welche Zahl die richtige ist, hängt an der Frage. Wer wissen will, ob die Fachbereiche bei der Auswahl benachteiligten, muss die Fachbereiche ansehen, denn dort wurde entschieden. Wer wissen will, warum Frauen sich in die engeren Fächer bewarben und ob dahinter ein anderer Nachteil steckt, früher in der Schule oder später im Beruf, stellt eine neue Frage, und für sie ist die Zahl im Ganzen der Anfang, nicht das Ende.

Zusammengezählte Zahlen gibt es überall: die Sterblichkeit eines Krankenhauses, das schwerere Fälle aufnimmt als ein anderes, den Durchschnittslohn in einem Land, dessen Bevölkerung altert, die Erfolgsquote einer Behörde, die die einfachen Fälle zuerst bearbeitet. In jedem dieser Fälle ist zu fragen, aus welchen Teilen das Ganze besteht und ob sich ihre Anteile verschoben haben.

4.3 Wo die Schwelle liegt

Nach den Zahlen der Weltbank lebten 1990 etwa 38 Prozent der Menschen von weniger als 2,15 Dollar am Tag, umgerechnet auf die Kaufkraft eines Dollars in

5 P. J. Bickel, E. A. Hammel, J. W. O'Connell, „Sex Bias in Graduate Admissions: Data from Berkeley", *Science* 187 (1975), S. 398–404. Zahlen (8.442 Männer, 4.321 Frauen; etwa 44 und 35 Prozent) nach der Wiedergabe bei David Freedman, Robert Pisani, Roger Purves, *Statistics*, 3. Aufl., New York 1998, Kap. 2.

den Vereinigten Staaten. 2019 waren es 8,5 Prozent. Nach denselben Zahlen lebten 2024 immer noch 44 Prozent der Menschheit von weniger als 6,85 Dollar am Tag, und die Zahl der Menschen unter dieser Schwelle hat sich seit 1990 kaum verändert, weil das Wachstum der Bevölkerung den Rückgang ihres Anteils aufwog.[6]

Wer sagt, die Armut der Welt sei in einer Generation dramatisch gesunken, hat recht. Wer sagt, fast die Hälfte der Menschheit lebe weiter in Armut und es seien so viele wie 1990, hat auch recht. Beide nehmen dieselbe Statistik und wählen eine andere Schwelle, und die Wahl der Schwelle entscheidet, welches Bild entsteht. Keine der beiden Schwellen ist falsch. Die niedrige zählt, wer nicht genug zum Überleben hat; die höhere, wer kaum mehr als das hat. Wie man sie gewichtet, hängt daran, was man wissen will: ob das Elend des Hungers weicht, oder ob die Welt ein Ort geworden ist, an dem die meisten Menschen gut leben können.

Wer eine Zahl liest, die sich auf eine Schwelle bezieht, auf eine Armutsgrenze, einen Grenzwert, eine Definition von Erfolg oder Gewalt, fragt darum, wo die Schwelle liegt, wer sie gesetzt hat und was sich zeigt, wenn man sie verschiebt. Bleibt das Bild bei verschiedenen Schwellen dasselbe, ist es fest. Kehrt es sich um, dann sagt die Zahl weniger über die Welt als über die Wahl der Schwelle.

4.4 Wer im Nenner steht

2013 werteten Forscher fast zwölftausend Zusammenfassungen von Fachaufsätzen über die Erwärmung der Erde aus, die zwischen 1991 und 2011 erschienen waren. 66,4 Prozent äußerten sich nicht dazu, ob der Mensch die Erwärmung verursacht. 32,6 Prozent bejahten es, 0,7 Prozent verneinten es, 0,3 Prozent hielten die Ursache für ungewiss. Von denen, die überhaupt eine Position zur Ursache bezogen, bejahten 97,1 Prozent, dass der Mensch die Erwärmung verursacht.[7]

Aus derselben Arbeit lassen sich zwei Sätze bilden. „97 Prozent der Fachleute sind sich einig“ nimmt als Nenner nur die Arbeiten, die eine Position beziehen. „Nur ein Drittel der Arbeiten bejaht die menschliche Ursache“ nimmt alle. Die

6 Daniel Gerszon Mahler, „Progress, Challenges in Ending Extreme Poverty“, World Bank Data Blog, 12. Juli 2023 (Rückgang des Anteils unter 2,15 Dollar „from about 38 percent to 8.5 percent“ zwischen 1990 und 2019); Weltbank, Pressemitteilung zum Bericht *Poverty, Prosperity, and Planet 2024*, 15. Oktober 2024 (2024 fast 700 Millionen Menschen oder 8,5 Prozent unter 2,15 Dollar, 44 Prozent unter 6,85 Dollar; Zahl unter 6,85 Dollar seit 1990 kaum verändert).

7 John Cook u. a., „Quantifying the Consensus on Anthropogenic Global Warming in the Scientific Literature“, *Environmental Research Letters* 8 (2013), 024024, doi:10.1088/1748-9326/8/2/024024 (11.944 Zusammenfassungen; 66,4, 32,6, 0,7 und 0,3 Prozent; 97,1 Prozent der Zusammenfassungen mit einer Position).

erste Lesart lautet in ihrer stärksten Form: Wer in einer Fachzeitschrift über die Folgen der Erwärmung für Gletscher schreibt, wiederholt nicht in jeder Zusammenfassung ihre Ursache, so wenig wie ein Geologe in jedem Aufsatz erwähnt, dass die Kontinente wandern. Die Kritiker halten in ihrer stärksten Form dagegen, dass man aus dem Schweigen nicht auf Zustimmung schließen darf, und dass der Satz mit den 97 Prozent als Zählung von Menschen verstanden wird, während er Texte zählt.

Beide Einwände treffen etwas. Das Mittel ist in beiden Fällen dasselbe: den Nenner hinschreiben. „Unter den Arbeiten, die sich zur Ursache äußern" ist eine genaue Aussage, und man kann sie prüfen. Was sie nicht sagt, ist, worüber genau Einigkeit besteht, wie stark der Mensch die Erde erwärmt und was daraus folgt. Das ist eine eigene Frage; sie gehört zu den Gremien, die einen Konsens feststellen.

Ein Nenner ist eine Antwort auf die Frage, von welchem Ganzen ein Teil gemeint ist. Schwieriger wird es, wenn schon das, was gezählt wird, ausgewählt ist, bevor jemand rechnet.

Weiterlesen: *Denken – Ordnen, ohne zu verlieren* (2025), Kap. 10 (eine Zahl als die falsche Art von Zusammenhang gelesen)

Zusammenfassung

Eine Zahl wiegt erst, wenn man weiß, worauf sie sich bezieht. Die britische Pillenwarnung von 1995 sagte richtig „doppelt so hoch" und verschwieg nichts, aber sie gab das Verhältnis ohne seine Größen: etwa zwanzig gegen vierzig Fälle je hunderttausend Frauen im Jahr. Danach stiegen Schwangerschaften und Abtreibungen. Wer zusammenzählt, gewichtet die Teile nach ihrer Größe; in Berkeley 1973 zeigte das Ganze eine Benachteiligung der Frauen, die in den Fachbereichen, wo entschieden wurde, nicht vorkam, und welche Zahl zählt, hängt an der Frage. Die Wahl einer Schwelle entscheidet, ob die Armut der Welt stark gesunken ist oder kaum: Beides steht in den Zahlen der Weltbank. Und aus derselben Auswertung von Fachaufsätzen über das Klima folgen „97 Prozent" und „ein Drittel", je nachdem, wer im Nenner steht. Das Mittel ist jedes Mal, Nenner, Größen und Schwelle hinzuschreiben und zu sehen, ob das Bild sich hält, wenn man sie ändert.

Kapitel 5: Was gezählt wird

5.1 Zwei Millionen Antworten

Vor der Präsidentenwahl 1936 in den Vereinigten Staaten verschickte die Zeitschrift *Literary Digest* Stimmzettel an zehn Millionen Menschen. 2,4 Millionen schickten sie ausgefüllt zurück. Nach dieser gewaltigen Umfrage würde der Herausforderer Alf Landon den Präsidenten Franklin D. Roosevelt schlagen. Roosevelt gewann mit großem Vorsprung. Die Adressen hatte die Zeitschrift vor allem aus Telefonbüchern und Listen der Autobesitzer genommen, und geantwortet hatte nur knapp jeder Vierte.[8]

Zwei Millionen Antworten sind mehr, als jede heutige Umfrage einholt. Die Menge hat den Fehler nicht verkleinert, sie hat ihn nur genauer gemacht. Wer auf diesen Listen stand, war im Durchschnitt wohlhabender als die übrigen, und wer sich die Mühe des Antwortens machte, unterschied sich noch einmal von denen, die es ließen. Zwei Filter lagen zwischen der Wählerschaft und dem Ergebnis, und keiner von ihnen ließ sich durch mehr Stimmzettel ausgleichen. Eine kleine Stichprobe, zufällig aus allen gezogen, hätte das Ergebnis besser getroffen.

Daraus folgt eine Frage, die man an jede Zahl über Meinungen, Erfahrungen oder Zustände stellen kann: Wer hatte überhaupt die Möglichkeit, gezählt zu werden, und wer hat sie genutzt? Eine Befragung im Netz erreicht, wer im Netz ist; die Kommentare unter einem Artikel stammen von denen, die kommentieren; die Bewertungen eines Hotels von denen, die besonders zufrieden oder besonders verärgert waren. Jede dieser Zahlen kann stimmen und doch nur für die Gezählten gelten.

5.2 Was veröffentlicht wird

Bevor ein Medikament in den Vereinigten Staaten zugelassen wird, muss der Hersteller der Aufsichtsbehörde alle Studien vorlegen, die er dafür angemeldet hat, auch die erfolglosen. Forscher haben 2008 für eine Reihe von Mitteln gegen Depression verglichen, was bei der Behörde lag und was in Fachzeitschriften erschien. Bei der Behörde lagen 74 Studien. Nach ihrer Bewertung waren 38 positiv, und 37 davon wurden veröffentlicht. Von den 36 übrigen, die negativ oder

8 Zehn Millionen verschickte, 2,4 Millionen zurückgeschickte Stimmzettel, Adressen „predominantly from telephone directories and automobile registration records“, Rücklauf 24 Prozent: Matthew J. Salganik, *Bit by Bit. Social Research in the Digital Age*, Princeton 2018, Kap. 3.3.1 („Representation“). Ausführlich Peverill Squire, „Why the 1936 Literary Digest Poll Failed“, *Public Opinion Quarterly* 52 (1988), S. 125–133.

zweifelhaft ausgefallen waren, erschienen 22 nicht und 11 so, dass sie positiv wirkten. In der veröffentlichten Literatur sahen 94 Prozent der Studien positiv aus; nach der Bewertung der Behörde waren es 51 Prozent.[9]

Kein einzelner Aufsatz musste dafür falsch sein. Wer die Literatur las, las Studien, die stattgefunden hatten, mit Ergebnissen, die gemessen worden waren. Die Verzerrung lag in der Auswahl dessen, was erschien, und sie traf gerade den sorgfältigen Leser, der viele Studien liest und ihre Ergebnisse zusammenfasst. Er bekam ein Bild, nach dem die Mittel deutlich besser wirkten, als die vollständigen Daten zeigten.

Entdecken ließ sich das nur, weil es ein zweites Verzeichnis gab, das vor dem Ergebnis angelegt war: die Meldungen bei der Behörde. Darin liegt ein allgemeines Mittel der Kalibrierung. Ein Bestand, der nach dem Ergebnis gefiltert wurde, lässt sich an einem Bestand prüfen, der vorher feststand. Aus demselben Grund verlangen die großen medizinischen Fachzeitschriften seit 2005, dass klinische Studien vor ihrem Beginn öffentlich angemeldet werden, sonst drucken sie die Ergebnisse nicht.[10]

Der Gedanke reicht über die Medizin hinaus. Wer wissen will, ob er das Ganze vor sich hat oder eine Auswahl, sucht nach einer solchen Liste: nach dem Plan einer Untersuchung vor ihrem Ergebnis, nach den Anträgen neben den Bewilligungen, nach allen Fällen neben den berichteten.

5.3 Zwei Zählweisen im Krieg

Wie viele Menschen ein Krieg tötet, ist eine der wichtigsten Zahlen, die es über ihn gibt, und eine der am schwersten zu ermittelnden. Am Irakkrieg nach 2003 lässt es sich sehen, weil er abgeschlossen und gründlich untersucht ist. Eine unabhängige Gruppe zählte die Toten, über die Medien und Behörden berichtet hatten, Fall für Fall. Eine Befragung von 9.345 Haushalten, die die Weltgesundheitsorganisation 2008 vorstellte, schätzte die Zahl der gewaltsam Getöteten von März 2003 bis Juni 2006 auf 151.000, mit einer Spanne von 104.000 bis 223.000. Das war etwa dreimal so viel wie die Zählung aus den Berichten. Eine kleinere Haushaltsbefragung aus

9 Erick H. Turner u. a., „Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy“, *New England Journal of Medicine* 358 (2008), S. 252–260; die Zahlen nach der Nachauswertung von Santiago G. Moreno u. a., „Novel Methods to Deal with Publication Biases“, *BMJ* 339 (2009), b2981.

10 International Committee of Medical Journal Editors, „Clinical Trial Registration“, gleichzeitig in den Mitgliedszeitschriften erschienen, u. a. *Canadian Medical Association Journal* 171 (2004), S. 606 f.: „This policy applies to any clinical trial starting enrollment after July 1, 2005.“

dem Jahr 2006, mit 1.849 Haushalten, war für einen ähnlichen Zeitraum auf rund 600.000 gewaltsame Tote gekommen, etwa viermal so viel wie die größere.[11]

Keine der drei Zahlen ist erfunden. Jede Methode hat in ihrer stärksten Form gute Gründe. Die Zählung aus Berichten enthält nur Tote, die es nachweislich gab, mit Ort und Datum; sie unterschätzt, weil viele Tote in keinem Bericht vorkommen, besonders dort, wo keine Reporter hinkommen. Die Befragungen erreichen auch die Toten, über die niemand berichtet hat; sie schließen von wenigen tausend Haushalten auf ein ganzes Land, und ihr Ergebnis hängt daran, welche Gegenden in die Stichprobe kamen, wie sicher die Befrager dort arbeiten konnten und wie genau sich die Befragten erinnern. Je kleiner die Stichprobe, desto weiter die Spanne.

Wer daraus nur eine Zahl nimmt, die eigene Seite passend, hat nicht gewichtet, sondern gewählt. Kalibriert ist ein Urteil, das die Methoden kennt und die Spanne ernst nimmt: dass die Zählung aus Berichten eine untere Grenze ist, dass die Befragungen weit über ihr liegen und untereinander um ein Vielfaches abweichen, und dass die Frage nach der genauen Zahl darum offen bleibt, die nach der Größenordnung aber nicht ganz. Für Kriege der Gegenwart gilt dasselbe in verschärfter Form. Die Zahlen der Kriegsparteien sind von Interessen bestimmt, die der unabhängigen Zähler von ihren Methoden; zusammen sagen sie mehr als jede für sich.

5.4 Die Frage an jede Zahl

Was in diesem Teil an Zahlen geprüft wurde, lässt sich in vier Fragen fassen, die zusammengehören, weil jede an einer anderen Stelle zwischen der Welt und der Zahl ansetzt. Die erste fragt nach der Grundrate: Wie häufig ist das, wovon die Zahl einen Fall zeigt? Die zweite nach dem Nenner: Wovon ist ein Teil gemeint, und wie groß sind die Größen hinter einem Verhältnis? Die dritte nach der Schwelle: Was zählt überhaupt als Fall, und ändert sich das Bild, wenn man die Grenze verschiebt? Die vierte nach der Auswahl: Wer konnte gezählt werden, was wurde veröffentlicht, und gibt es eine Liste, die vor dem Ergebnis feststand? Keine dieser

11 Weltgesundheitsorganisation, „New Study Estimates 151 000 Violent Iraqi Deaths since 2003 Invasion“, Pressemitteilung, 9. Januar 2008 (Iraq Family Health Survey, 9.345 Haushalte; 104.000 bis 223.000; etwa dreimal so viel wie die Zählung von Iraq Body Count aus Medienberichten, etwa viermal weniger als die kleinere Befragung von 2006), veröffentlicht als Iraq Family Health Survey Study Group, „Violence-Related Mortality in Iraq from 2002 to 2006“, *New England Journal of Medicine* 358 (2008), S. 484–493. Die kleinere Befragung: Gilbert Burnham, Riyadh Lafta, Shannon Doocy, Les Roberts, „Mortality after the 2003 Invasion of Iraq: A Cross-Sectional Cluster Sample Survey“, *The Lancet* 368 (2006), S. 1421–1428 (1.849 Haushalte; 601.027 gewaltsame Todesfälle bis Juli 2006, Spanne 426.369 bis 793.663).

Fragen verlangt Mathematik. Sie verlangt, eine Zahl nicht als Antwort zu nehmen, sondern als Ergebnis eines Vorgangs, den man sich ansehen kann.

Zahlen haben den Vorzug, dass man diese Fragen an ihnen stellen kann. Bilder, Erzählungen und Eindrücke haben keinen Nenner, den man nachschlagen könnte, und doch entsteht aus ihnen das meiste dessen, was Menschen über fremde Länder wissen.

Weiterlesen: *Organisation* (2026), Kap. 6 (was Messen in Einrichtungen bewirkt), *Wahrheit* (2025), Kap. 19 (Rückwirkung des Messens)

Zusammenfassung

Was gezählt wird, ist oft schon ausgewählt, bevor jemand rechnet. Der Literary Digest befragte 1936 zehn Millionen Menschen aus Telefonbüchern und Autolisten, bekam 2,4 Millionen Antworten und sagte den falschen Sieger voraus; die Menge macht einen Auswahlfehler nicht kleiner, nur genauer. Von 74 Studien über Mittel gegen Depression, die der amerikanischen Aufsicht vorlagen, war nach ihrer Bewertung die Hälfte positiv, in den Fachzeitschriften aber 94 Prozent; entdecken ließ sich das, weil ein Verzeichnis vor dem Ergebnis feststand. Die Toten des Irakkriegs wurden aus Berichten gezählt und in zwei Befragungen geschätzt, die untereinander um ein Vielfaches abwichen; jede Methode hat gute Gründe, und kalibriert ist das Urteil, das die Methoden kennt und die Spanne ernst nimmt. An jede Zahl lassen sich vier Fragen stellen: nach der Grundrate, dem Nenner, der Schwelle und der Auswahl.

TEIL III: FREMDE LÄNDER

Kapitel 6: Was man aus der Ferne weiß

6.1 Jeder Kanal hat seinen Filter

Von Brasilien kennen viele Menschen, die nie dort waren, den Karneval, die Armenviertel an den Hängen von Rio und den Fußball. Von Indien die Filme aus Bombay, die Armut und die Programmierer, von Japan die überfüllten Züge, die Kirschblüte und die Roboter, von Deutschland das Oktoberfest, die Autos und die Geschichte des zwanzigsten Jahrhunderts. Nichts davon ist erfunden. Jedes dieser Bilder hat aber einen Weg hinter sich, auf dem es ausgewählt wurde – einen Weg, der mit der Häufigkeit dessen, was es zeigt, wenig zu tun hat.

Ein Film braucht eine Geschichte, und eine Geschichte braucht einen Konflikt, eine Ausnahme, etwas, das nicht jeden Tag geschieht. Er spielt an Orten, die man wiedererkennt, in der Hauptstadt, vor den bekannten Gebäuden; das Land außerhalb kommt als Landschaft vor oder gar nicht. Eine Nachricht berichtet, was vom Gewohnten abweicht: den Anschlag, die Wahl, die Flut, den Rekord. Was in den Netzwerken umläuft, ist, was geteilt wird, und geteilt wird, was empört, rührt oder lachen lässt. Reiseberichte zeigen, was sehenswert ist. Jeder dieser Kanäle filtert anders, aber sie filtern in dieselbe Richtung, hin zum Ungewöhnlichen und zu wenigen Orten. Darum heben sich ihre Verzerrungen nicht auf, wenn man alle zusammennimmt. Sie addieren sich.

Wer aus fünf Kanälen dasselbe Bild bekommt, fühlt sich bestätigt, mit Grund, soweit das Bild Tatsachen enthält. Für das Gewicht hat er keinen Grund. Fünf Filter, die nach demselben fragen, sind für die Frage, wie häufig etwas ist, nicht fünf Zeugen, sondern einer.

6.2 Erzählungen sind gefiltert

Was Bekannte erzählen, die ein Land besucht haben, ist näher an der Wirklichkeit als ein Film, aber auch es ist ausgewählt, und zwar nach dem, was den Erzählenden dorthin geführt hat. Wer zu Geschäften reist, erzählt von Geschäften, von

Verhandlungen, Partnern und Hotels; er hat Büros gesehen, Flughäfen und Restaurants, in denen man Gäste ausführt, und von alldem kann er genau berichten. Wie eine Familie dort wohnt, wovon ein Bauer lebt oder wie lange man auf einen Arzttermin wartet, weiß er nicht, und meist erzählt er es auch nicht, weil man ihn danach nicht fragt.

Ähnlich geht es mit anderen Wegen. Ein Korrespondent lebt oft in der Hauptstadt, unter anderen Ausländern und Gebildeten, und sieht das Land von dort. Wer ausgewandert ist und nach Hause schreibt, schreibt eher, was ihm gelungen ist, als was ihm misslang, und wer gescheitert ist, schreibt vielleicht gar nicht. Ein Tourist sieht, was für Touristen eingerichtet ist. Jeder dieser Berichte ist ehrlich gemeint. Jeder ist aber eine Stichprobe, die nach dem Zweck der Reise gezogen wurde. Wer ihn hört, sollte fragen: Was hat er dort getan, wo hat er sich aufgehalten, mit wem hat er gesprochen? Dann weiß er, wofür der Bericht ein Zeugnis ist und wofür nicht.

6.3 Was sich nicht zeigen lässt

Manches erreicht den fernen Betrachter auf keinem Kanal. Ein Geruch lässt sich nicht senden, ein Geschmack nicht, die Hitze am Mittag, die Lautstärke einer Straße, die Feuchtigkeit der Luft auch nicht. Bilder zeigen sie als Andeutung, und der Betrachter ergänzt sie aus seinem eigenen Leben, also falsch. Ebenso wenig zeigt sich das Gewöhnliche, weil es nichts zu berichten gibt: wie ein Werktag verläuft, wie lange man zur Arbeit fährt, was man zu Mittag isst, worüber man sich ärgert, wie man mit Behörden umgeht. Gerade dieses Gewöhnliche ist aber, was das Leben dort zum größten Teil ausmacht. Was in den Bildern fehlt, ist also nicht ein Rand, sondern die Mitte.

Daraus folgt eine einfache Frage, die man an jedes Bild eines fernen Landes stellen kann: Was wäre dort langweilig? Was ist so selbstverständlich, dass niemand es filmen, melden oder erzählen würde? Die Antwort weiß man aus der Ferne meist nicht, aber schon die Frage verschiebt das Gewicht. Sie erinnert daran, dass das, was man kennt, das Erzählenswerte ist, und dass das Erzählenswerte nur einen kleinen Teil eines Lebens ausmacht.

6.4 Was einmal galt

Ein Bild, das einmal entstanden ist, bleibt, auch wenn sich das Land ändert. Wer sein Bild von der Armut der Welt um 1990 gewonnen hat, als nach den Zahlen der Weltbank mehr als ein Drittel der Menschen in äußerster Armut lebte, und

es nicht erneuert hat, sieht noch heute eine Welt, in der das so ist; im Jahr 2019 war es weniger als ein Zehntel. Dasselbe gilt für die Lebenserwartung, die Zahl der Kinder in einer Familie, die Verbreitung von Telefonen oder Schulen in einem Land, das man einmal kennenlernte. Wer zu einer Zeit dort war oder darüber las, kennt das Land zu dieser Zeit. Für jedes Bild ist darum zu fragen, wann es entstand und was sich seitdem geändert haben könnte. Auch die Erinnerung an die eigene Reise hat ein Datum.

6.5 Aus der Ferne kalibrieren

Ganz ohne Reise lässt sich das Bild eines Landes schon ein gutes Stück nachstellen. Das wichtigste Mittel ist, Kanäle nebeneinanderzulegen, die nach Verschiedenem filtern: neben die Bilder die Zahlen, die das Gewöhnliche messen, also wie die Menschen dort wohnen, wovon sie leben, wie alt sie werden und was sie für ihren Lebensunterhalt ausgeben; neben die Berichte von außen die Zeitungen des Landes selbst, die eigene Filter haben, aber andere; neben die Hauptstadt die Provinz. Man kann fragen, wer im Land über das Land schreibt, und ob es Stimmen von denen gibt, die nicht in der Hauptstadt leben und nicht zu den Gebildeten gehören. Und man kann die eigenen Bilder darauf prüfen, aus welcher Zeit sie stammen.

Was aus der Ferne nicht gelingt, ist die Anschauung des Zusammenhangs, in dem all das steht: wie die Zahlen sich in einem Tag, einer Straße, einem Gesicht zeigen. Dafür muss man hinfahren. Aber auch was man dort sieht, ist ein Ausschnitt, und zwar einer, der sich für vollständiger hält als jeder andere.

Weiterlesen: *Wahrheit* (2025), Kap. 19 (was man über fremde Länder weiß)

Zusammenfassung

Was man aus der Ferne über ein Land weiß, kommt aus Kanälen, die jeweils filtern: Filme brauchen Konflikte und bekannte Orte, Nachrichten das Ungewöhnliche, Netzwerke das, was geteilt wird, Reiseberichte das Sehenswerte. Weil sie in dieselbe Richtung filtern, addieren sich ihre Verzerrungen; für die Frage, wie häufig etwas ist, sind fünf gleich gerichtete Filter ein Zeuge. Auch Erzählungen von Bekannten sind nach dem Zweck ihrer Reise gezogene Stichproben. Was sich gar nicht zeigen lässt, sind Gerüche, Geschmäcke, Hitze und das Gewöhnliche, und das Gewöhnliche ist der größte Teil des Lebens dort; die Frage, was dort langweilig wäre, verschiebt das Gewicht zurück. Jedes Bild hat zudem ein Datum. Aus der Ferne kalibriert man,

indem man Kanäle mit verschiedenen Filtern nebeneinanderlegt: Zahlen über das Gewöhnliche, die Presse des Landes, die Provinz neben der Hauptstadt.

Kapitel 7: Vor Ort und die Regionen

7.1 Was nur vor Ort mitkommt

Wer zum ersten Mal in ein Land kommt, das er nur aus Bildern kannte, ist oft von Dingen überrascht, die ihm niemand angekündigt hat. Nicht von den Sehenswürdigkeiten, die kannte er, sondern von Kleinigkeiten, die den Einheimischen so selbstverständlich sind, dass sie keiner erwähnt: Dampf, der aus den Straßen steigt, ein Verkehr, der anders klingt, Pflanzen und Vögel, die er nicht kennt, Gerüche, für die er keine Wörter hat. Diese Überraschungen sind ein Maß dafür, wie gefiltert die Bilder waren, mit denen er kam. Was ihn überrascht, hat ihm gefehlt, ohne dass er es vermisst hätte.

Vor Ort bekommt man auch mit, was sich nur im Ganzen zeigt: wie schnell die Leute gehen, wie nahe sie einander kommen, wie laut sie sprechen, wie sie einander ansehen, wo man einkauft und was in den Auslagen liegt, wie weit die Wege sind. Vor allem sieht man, wie das alles zusammenhängt. Der Film zeigt ein Viertel; vor Ort sieht man, wie es an das nächste grenzt, wer zwischen beiden hin- und hergeht und wer nicht. Die Zahlen aus dem vorigen Kapitel bekommen hier eine Gestalt. Wer weiß, dass in einem Land die meisten Menschen in kleinen Läden einkaufen, weiß es anders, wenn er einen Monat lang dort einkauft.

Und doch ist auch die eigene Anschauung gefiltert, und ihr Filter ist schwerer zu bemerken als der einer Kamera, weil man ihn nicht vor sich hat, sondern selbst ist. Man wohnt in einem bestimmten Viertel, man folgt Warnungen, welche Gegend man meiden soll, man fährt im Bus durch ein Viertel und betritt kein Haus. Man trifft vor allem Menschen, die mit Fremden zu tun haben, im Hotel, im Geschäft, bei der Arbeit. Man ist zu einer Jahreszeit dort, und an Feiertagen sieht eine Stadt anders aus als an Werktagen. Und der erste Eindruck wiegt, wie die Zahl vom Glücksrad, mehr, als er verdient, weil er der erste war. Die eigene Anschauung ist kostbar, gerade weil sie zeigt, was in den Bildern fehlt. Sie ist aber selbst eine Stichprobe und muss gewichtet werden wie alles andere.

7.2 Land, Region, Stadt, Viertel

Wer zwei Wochen in einer großen Stadt war, kennt einen Teil dieser Stadt und noch nicht das Land. Was er dort sah, muss er einordnen: Was davon gehört zum Land, was zur Region, was zur Stadt, was zum Viertel? Die Frage ist schwerer, als sie klingt, weil man die Ebenen am einzelnen Eindruck nicht ablesen kann. Dass die Leute in einer Stadt schnell gehen, kann am Land liegen, an der Großstadt, an der Tageszeit oder daran, dass man im Geschäftsviertel war.

Fast jedes Land zerfällt bei näherem Hinsehen in Regionen, die sich in Sprache, Küche, Glauben, Geschichte und Umgangsformen deutlich unterscheiden, und diese wieder in kleinere. Der Journalist Colin Woodard hat für Nordamerika elf Regionalkulturen beschrieben, die er auf die Siedlungsgeschichte zurückführt und deren Grenzen quer zu denen der Bundesstaaten und sogar der Staaten verlaufen; manche Bundesstaaten haben an mehreren Anteil.[1] Man muss seiner Einteilung nicht in jedem Strich folgen, um den Befund zu sehen: Wer „die Amerikaner“ beschreibt, beschreibt meist eine oder zwei dieser Kulturen und hält sie für das Ganze. Die indische Verfassung erkennt zweiundzwanzig Sprachen an, und wer aus Kerala zurückkommt, bringt ein anderes Indien mit als der, der im Punjab war.[2] In Italien sind der Norden und der Süden, in Nigeria der Norden und der Süden, in China die Küstenstädte und die Provinzen des Westens, in Portugal Lissabon und das Bergland im Nordosten jeweils wie verschiedene Länder für den, der beide kennt. Bayern zerfällt in Franken, Schwaben und Altbayern, und in jedem dieser Landesteile gibt es wieder Unterschiede, die die Einheimischen sofort hören.

Daraus folgt kein Verbot, über ein Land im Ganzen zu sprechen. Manches gilt im ganzen Land, weil es dieselben Gesetze, dieselbe Schule, dieselbe Währung und dieselbe Hauptstadt hat. Es folgt aber, dass man bei jedem Eindruck die Ebene angeben muss, auf der man ihn verallgemeinert, und dass die Ebene sich nicht von selbst versteht. Was man über „die Inder“ sagt, sollte ein Inder aus einer anderen Gegend nicht als Beschreibung eines Nachbarn lesen müssen.

1 Colin Woodard, *American Nations. A History of the Eleven Rival Regional Cultures of North America*, New York 2011; die elf Kulturen (Yankeedom, New Netherland, The Midlands, Tidewater, Greater Appalachia, Deep South, El Norte, The Left Coast, The Far West, New France, First Nation) auch in Woodard, „Up in Arms“, *Tufts Magazine*, Herbst 2013.

2 Verfassung der Republik Indien, Achter Anhang (Eighth Schedule), mit 22 Sprachen in der geltenden Fassung.

7.3 Einen Aufenthalt kalibrieren

Den eigenen Aufenthalt kann man nachkalibrieren, so wie die Bilder aus der Ferne. Man vergleicht ihn mit den Eindrücken anderer, die zur selben Zeit an einem anderen Ort waren oder zu einer anderen Zeit am selben. Man fragt Einheimische, was an dem eigenen Eindruck für ihr Land typisch ist und was nicht; sie kennen die Grundraten ihres Landes, die einem Besucher fehlen, aber auch sie kennen vor allem ihre eigene Gegend und ihre eigene Schicht. Man sieht nach, was die Statistiken über Unterschiede zwischen den Regionen sagen. Und man unterscheidet für sich, was man oft gesehen hat und was einmal, was man selbst gesehen und was man nur gehört hat.

Weggeworfen wird dabei nichts. Was man gesehen hat, bleibt gesehen, auch wenn es nur für ein Viertel gilt. Nur sein Platz im Bild ändert sich: vom Kennzeichen eines Landes zum Kennzeichen einer Gegend, einer Zeit oder einer Lage. Das ist mühsamer, als den Eindruck zu verallgemeinern oder ihn zu verwerfen, aber es ist der einzige Weg, auf dem er zu Wissen wird.

Eine letzte Wendung kehrt das Kapitel um. Wer zurückkommt, wird gefragt, wie es war, und erzählt. Er erzählt, was ihn überrascht hat, was lustig war, was erschreckend war, und nicht, was gewöhnlich war; er wird nun selbst zu einem der gefilterten Kanäle aus dem vorigen Kapitel. Wer das weiß, kann es zum Teil ausgleichen, indem er dazusagt, was er nicht gesehen hat, und wofür sein Bericht nicht steht.

Bis hierher geschah das Gewichten von selbst, durch Kameras, Kanäle, Gedächtnis und Reisewege, ohne dass jemand es so wollte. Dieselben Mittel lassen sich auch mit Absicht einsetzen.

Zusammenfassung

Vor Ort überrascht, was niemand angekündigt hat, weil es den Einheimischen selbstverständlich ist, und an dieser Überraschung misst sich, wie gefiltert die Bilder waren. Dort zeigt sich auch, wie alles zusammenhängt. Die eigene Anschauung ist aber selbst eine Stichprobe: nach Viertel, Warnungen, Gesprächspartnern, Jahreszeit und erstem Eindruck gezogen. Was man sieht, gehört zum Land, zur Region, zur Stadt oder zum Viertel, und die Ebene lässt sich am Eindruck nicht ablesen. Fast jedes Land zerfällt in Regionen, die sich wie Länder unterscheiden: Woodard zählt in Nordamerika elf Regionalkulturen quer zu den Staatsgrenzen, Indiens Verfassung kennt zweiundzwanzig Sprachen. Einen Aufenthalt kalibriert man, indem man ihn mit anderen Orten und Zeiten vergleicht, Einheimische nach dem Typischen fragt und Oftgesehenes von Einmaligem trennt; weggeworfen wird

nichts, nur der Platz im Bild ändert sich. Wer heimkehrt und erzählt, wird selbst zum gefilterten Kanal.

TEIL IV: GEWICHT ALS MITTEL

Kapitel 8: Propaganda

8.1 Die Lüge und ihr Preis

Im Spätsommer und Herbst 1914 töteten deutsche Truppen beim Vormarsch durch Belgien und Nordfrankreich rund 6.500 Zivilisten, viele davon bei Erschießungen in Städten und Dörfern, deren Bewohner sie für Freischärler hielten. Die alliierte Propaganda machte daraus mehr. Sie verbreitete neben den wirklichen Verbrechen Geschichten von Soldaten, die Säuglinge aufgespießt und Frauen verstümmelt hätten, und vieles davon ließ sich nie belegen oder war erfunden. Nach dem Krieg kam das ans Licht, und die Lehre, die viele daraus zogen, war: Gräuelberichte sind Propaganda. Als im Zweiten Weltkrieg Berichte über den planmäßigen Mord an den Juden Europas eintrafen, hielten viele sie deshalb für übertrieben. Sie erwiesen sich im Wesentlichen als zutreffend.[1] Auch die wirklichen Verbrechen von 1914 wurden nach dem Krieg lange bestritten, bis Historiker sie aus den Akten belegten.

Die Lüge in der Propaganda hat also einen Preis, der zweimal bezahlt wird. Zuerst verschiebt sie das Gewicht nach oben: Das wirkliche Verbrechen wird mit erfundenen Zügen so schwer, dass es zum Beweis für die Bosheit eines ganzen Volkes wird. Ist sie aufgedeckt, verschiebt sie das Gewicht nach unten, und zwar über die Lüge hinaus. Dann wiegt jeder Bericht dieser Art weniger, auch der wahre, auch der nächste, auch der über ein Verbrechen, das größer ist als alles, was die Lüge erfand. Wer lügt, um einem Bericht mehr Gewicht zu geben, beschädigt die Waage, mit der künftig gewogen wird.

Einen Fall dieser Art aus jüngerer Zeit, die Aussage über Säuglinge, die irakische Soldaten 1990 in Kuwait aus Brutkästen genommen hätten, behandelt das Buch über die Wahrheit. Solche Fälle gibt es in vielen Kriegen und auf allen Seiten. Sie

1 John Horne, Alan Kramer, *German Atrocities, 1914. A History of Denial*, New Haven 2001 („some 6500 Belgian and French civilians“, Verlagstext Yale University Press); United States Holocaust Memorial Museum, *Holocaust Encyclopedia*, Artikel „Combating Holocaust Denial: Origins of Holocaust Denial“: Viele britische Berichte über deutsche Gräuel in Belgien hätten sich nach dem Krieg als erfunden herausgestellt; „As a result of that experience, many people were skeptical of reports of mass murder operations during World War II.“

sind aber nicht die Regel, und gerade weil sie sich widerlegen lassen, sind sie für den, der sie verbreitet, gefährlich.

Weiterlesen: *Wahrheit* (2025), Kap. 19 (die Brutkasten-Aussage)

8.2 Die Regel: Auswahl

Die Regel der Propaganda ist bescheidener und darum wirksamer. Vom Gegner zeigt sie das Schlechte, von der eigenen Seite das Gute, wobei jede einzelne Meldung stimmen kann. Die Opfer der eigenen Seite haben Namen, Gesichter und Familien, die der anderen sind Zahlen, wenn sie überhaupt vorkommen. Die eigenen Fehler sind Einzelfälle, die des Gegners Zeichen seines Wesens. Die Gewalt des Gegners kommt ohne Vorgeschichte vor, die eigene als Antwort.

Nichts davon braucht eine Behörde, die es anordnet. Es entsteht auch, wo Redaktionen frei arbeiten, weil Menschen sich leichter in die einfühlen, die ihnen nahestehen, weil die eigenen Opfer leichter zu erreichen sind als die der anderen Seite, und weil ein Publikum lieber liest, was seine Sicht bestätigt. Wo es angeordnet wird, ist es nur gründlicher. Für den Leser macht das wenig Unterschied: Er bekommt in beiden Fällen ein Bild aus wahren Teilen, dessen Gewichte nach Zugehörigkeit verteilt sind und nicht nach dem, was geschah.

Experiment 3: Namen und Zahlen

Nehmen Sie einen Konflikt, über den gerade berichtet wird, und eine Woche Berichterstattung aus einer Quelle, die Sie gewöhnlich lesen. Zählen Sie, wie oft Opfer jeder Seite mit Namen, Bild oder Geschichte vorkommen, und wie oft nur als Zahl. Zählen Sie dann, wie oft die Gewalt jeder Seite mit einer Vorgeschichte erzählt wird, als Antwort auf etwas, und wie oft ohne. Wenn die Verteilung auffällig ungleich ist, wissen Sie noch nicht, wer recht hat; vielleicht hat eine Seite tatsächlich mehr Opfer oder mehr Gründe. Sie wissen aber, mit welchem Gewicht die Teile Sie erreichen, und können dieselbe Zählung an einer Quelle der anderen Seite wiederholen. Fällt sie spiegelbildlich aus, haben Sie zwei gefilterte Bilder und eine Vorstellung davon, wie groß der Filter ist.

8.3 Ausgewogenheit als Inszenierung

Die feinere Form zeigt beide Seiten. Ein Bericht lässt Befürworter und Gegner zu Wort kommen, und keiner kann sagen, eine Seite sei verschwiegen worden. Die Gewichtung liegt dann in der Inszenierung: Die erwünschte Seite bekommt den überzeugenden Sprecher, das stärkere Bild und das letzte Wort, die unerwünschte den schwächsten Vertreter, den man finden konnte. Der Anschein der Ausgewogenheit macht die Gewichtung unsichtbar, denn wer nachzählt, findet beide Seiten vertreten.

Die Gegenform ist ebenso eine Verzerrung. Sie entsteht gerade aus dem Wunsch, nicht zu inszenieren. Wer zu jeder Sache eine Stimme dafür und eine dagegen bringt, gleich lang, gibt einer Position, die in einem Fach kaum jemand vertritt, dasselbe Gewicht wie einer, die fast alle vertreten. Ein Gutachten für die Aufsicht der BBC kam 2011 zu dem Schluss, eine zu starre Anwendung der Regeln zur Unparteilichkeit habe in der Wissenschaftsberichterstattung Randmeinungen zu viel Aufmerksamkeit gegeben, etwa bei der Sicherheit der Masern-Mumps-Röteln-Impfung und bei gentechnisch veränderten Pflanzen.[2]

Zwischen beiden Fehlern liegt keine Mitte, die man ausrechnen könnte, sondern eine Pflicht, die Gewichtung offenzulegen. Ein Bericht kann sagen, wie viele Fachleute welche Position vertreten, worauf sie sich stützen und wo die Minderheit einen Punkt hat, den die Mehrheit noch nicht beantwortet hat. Dann kann der Leser selbst gewichten und prüfen, ob das angegebene Gewicht stimmt. Verdeckt bleibt die Gewichtung, wenn sie in der Auswahl der Sprecher steckt oder in der bloßen Gleichverteilung der Sendezeit. Wie groß eine Mehrheit ist und wie gut sie begründet, ist allerdings selbst eine Frage der Gewichtung, und die Gremien, die es feststellen, haben ihre eigenen Filter.

8.4 Worüber man nachdenkt

Ein amerikanischer Politikwissenschaftler hat die Wirkung der Presse 1963 so beschrieben: Sie habe meist wenig Erfolg damit, den Lesern zu sagen, was sie denken sollen, aber sehr viel damit, ihnen zu sagen, worüber sie nachdenken

2 Steve Jones, Gutachten zur Unparteilichkeit und Genauigkeit der Wissenschaftsberichterstattung der BBC, im Auftrag des BBC Trust, 2011; wiedergegeben in *Chemistry World*, „Report concludes BBC science is good but has weaknesses“ (2011): Eine „over-rigid“ Anwendung der Unparteilichkeitsregeln „gives ‚undue attention to marginal opinion‘“, etwa bei der Sicherheit des MMR-Impfstoffs und bei gentechnisch veränderten Pflanzen.

sollen.[3] Das ist eine Aussage über Gewichte. Welches Thema eine Woche lang auf den ersten Seiten steht, welches einmal kurz erwähnt wird und welches nie, bestimmt, was eine Öffentlichkeit für dringend hält, lange bevor jemand dafür oder dagegen argumentiert.

Auch hier hilft, die Gewichte an etwas zu prüfen, das nicht aus demselben Strom kommt. Was hat in diesem Jahr die meisten Menschen in meinem Land das Leben gekostet, was hat sie das meiste Geld gekostet, was hat ihren Alltag am meisten verändert, und wie viel war darüber in dem zu lesen, was ich gelesen habe? Die Antwort kann ergeben, dass die Aufmerksamkeit gut verteilt war. Sie ergibt oft, dass das Gewöhnliche und Langsame, das viele trifft, weniger Raum hatte als das Plötzliche, das wenige trifft. Der Grund muss keine Absicht sein, denn das Plötzliche ist neu. Die Wirkung auf das Maß, mit dem man das Wichtige vom Unwichtigen trennt, ist dieselbe.

Die Inszenierung und die Themenwahl arbeiten mit der Verteilung des Gezeigten. Eine andere Form arbeitet mit seiner Menge.

Zusammenfassung

Propaganda lügt selten, und wo sie lügt, zahlt sie zweimal: 1914 töteten deutsche Truppen in Belgien und Frankreich rund 6.500 Zivilisten, die alliierte Propaganda fügte erfundene Gräuel hinzu, und als das herauskam, hielten viele die zutreffenden Berichte über den Massenmord im Zweiten Weltkrieg für Propaganda. Wer lügt, um Gewicht zu schaffen, beschädigt die Waage für künftige Fälle. Die Regel ist die Auswahl: die eigenen Opfer mit Namen, die anderen als Zahl, die eigene Gewalt als Antwort, die fremde ohne Vorgeschichte; dafür braucht es keine Anordnung. Die feinere Form zeigt beide Seiten und gewichtet durch die Inszenierung. Ihr Spiegel ist die falsche Balance, die einer Randmeinung so viel Raum gibt wie dem, was fast alle Fachleute vertreten. Zwischen beiden hilft nur, die Gewichtung offenzulegen. Und schon die Wahl der Themen legt fest, worüber eine Öffentlichkeit nachdenkt, lange bevor jemand argumentiert.

3 Bernard C. Cohen, *The Press and Foreign Policy*, Princeton 1963, S. 13: Die Presse „may not be successful much of the time in telling people what to think, but it is stunningly successful in telling its readers what to think about".

Kapitel 9: Überflutung

9.1 Menge statt Überzeugung

Eine Studie der RAND Corporation, eines Forschungsinstituts, das vor allem für die amerikanischen Streitkräfte arbeitet, beschrieb 2016 die russische Staatspropaganda mit dem Bild eines Feuerwehrschlauchs voller Unwahrheiten. Ihre Kennzeichen seien große Menge und viele Kanäle, Schnelligkeit, Dauer und Wiederholung, und ein fehlendes Bekenntnis zur Wirklichkeit wie zur Widerspruchsfreiheit: Was gestern behauptet wurde, darf heute durch das Gegenteil ersetzt werden.[4] Zwei Jahre später beschrieb ein amerikanischer Wahlkampfstratege seinen Umgang mit der Presse im eigenen Land in einem Satz, der sich nicht höflich übersetzen lässt: Man müsse das Feld mit Mist überfluten.[5]

Die beiden Fälle stammen aus verschiedenen Ländern und politischen Lagern, und darum gehören sie hierher nebeneinander. Die Überflutung ist keine Eigenart eines Regimes oder einer Richtung. Sie ist eine Technik, die jedem offensteht, der mehr Behauptungen in Umlauf bringen kann, als andere prüfen können, und ihr Ziel ist nicht, zu überzeugen. Ihr Ziel ist das Maß, mit dem überzeugt wird.

9.2 Was die Menge mit dem Maß macht

Die Menge wirkt auf drei Weisen, die einander verstärken. Die erste ist die Wiederholung. Was einem oft begegnet, wird vertraut, und das Vertraute fühlt sich häufig und glaubhaft an, auch wenn man weiß, woher es kommt. Das Gefühl für Häufigkeit, von dem das zweite Kapitel handelte, lässt sich also unmittelbar füttern: nicht mit Argumenten, sondern mit Wiederholungen.

Die zweite ist die Einebnung. Wo zu jeder Frage zwanzig Behauptungen umlaufen, die einander widersprechen, verliert die gut belegte ihr Gewicht, nicht weil sie widerlegt wäre, sondern weil sie eine unter zwanzig ist. Das Maß, das zwischen gesichert, strittig und unbekannt unterscheidet, wird flach. Am Ende scheint alles ungefähr gleich ungewiss, und wer so weit ist, glaubt nicht mehr

4 Christopher Paul, Miriam Matthews, *The Russian „Firehose of Falsehood" Propaganda Model. Why It Might Work and Options to Counter It*, RAND Corporation, Perspective PE-198, Santa Monica 2016: „high-volume and multichannel", „rapid, continuous, and repetitive", „lacks commitment to objective reality", „lacks commitment to consistency".

5 Steve Bannon im Gespräch mit Michael Lewis, wiedergegeben in Michael Lewis, „Has Anyone Seen the President?", *Bloomberg View*, 9. Februar 2018: „the way to deal with them is to flood the zone with shit".

dem, der die besseren Gründe hat, sondern dem, dem er ohnehin traut, oder niemandem.

Die dritte ist die Erschöpfung. Behauptungen lassen sich schneller aufstellen als prüfen, und wer jede einzelne widerlegt, hat seine Kraft verbraucht, bevor die nächste Welle kommt. Dazu kommt, dass jede Widerlegung die Behauptung noch einmal nennt und ihr damit Aufmerksamkeit gibt. Die Studie über den Feuerwehrschlauch zog daraus selbst den Schluss, man solle nicht erwarten, ihm mit der Wasserpistole der Wahrheit beizukommen. Sie empfahl, die Menschen vorher zu warnen und ihnen zu zeigen, wie die Technik arbeitet, statt jede einzelne Unwahrheit hinterher zu bekämpfen; man müsse denen, auf die der Schlauch gerichtet ist, Regenmäntel anziehen.[6]

Damit verschiebt sich die Frage. Gegen eine einzelne Lüge hilft die Prüfung der einzelnen Aussage. Gegen die Überflutung hilft sie wenig, weil die Menge gerade darauf zielt, sie zu überlasten. Helfen kann nur, was das Maß selbst stabil hält.

9.3 Was das Maß stabil hält

Das erste Mittel ist die Langsamkeit. Die ersten Meldungen über ein großes Ereignis sind oft unvollständig und manchmal falsch, auch ohne jede Absicht, weil niemand schon weiß, was geschehen ist. Wer einen Tag wartet, bevor er sich ein Urteil bildet, verliert wenig und gewinnt viel: Was nach einem Tag noch steht, ist geprüfter als das, was in der ersten Stunde umlief. Was nach einer Woche noch steht, umso mehr. Die Überflutung lebt vom Augenblick. Wer sie an der Zeit misst, sieht, wie viel von ihr nach kurzer Zeit verschwunden ist, ohne je widerrufen worden zu sein.

Halt gibt außerdem eine Bilanz. Wer wenige Quellen hat, deren Trefferquote er kennt, weil er sich gemerkt hat, wie oft sie sich geirrt und ob sie ihre Irrtümer berichtigt haben, kann sich in einer Flut an ihnen festhalten. Das ist keine Treue zu einem Lager. Eine solche Liste umfasst Quellen verschiedener Richtung; jede bleibt auf ihr, solange ihre Bilanz trägt.

Ebenso wichtig ist, die Lagen zu unterscheiden. Die Flut will alles in den Zustand des Strittigen bringen. Dagegen hilft, für sich festzuhalten, was in einer Sache gesichert ist, was wirklich strittig, und was niemand weiß, und diese drei Fächer nicht deshalb zu vermischen, weil viele laut reden. Dass zu einer Frage

6 Paul, Matthews, *The Russian „Firehose of Falsehood" Propaganda Model* (2016): „don't expect to counter the firehose of falsehood with the squirt gun of truth"; Empfehlungen, das Publikum vorzuwarnen („forewarn") und denen, auf die der Schlauch gerichtet ist, „raincoats" anzuziehen.

viele Behauptungen umlaufen, ist ein Befund über den Umlauf, nicht über die Frage.

Schließlich bleibt die Quelle hinter der Quelle. Wo ein Dokument, eine Aufzeichnung, ein Datensatz vorliegt, lohnt der Blick darauf mehr als der auf zehn Berichte darüber. Er kostet Zeit, und darum kann man ihn nicht bei allem tun. Bei dem, worauf es ankommt, ist er die schnellste Art, aus der Flut herauszukommen.

Keines dieser Mittel verlangt, dass jemand die Flut für einen abstellt. Sie verlangen, dass der Einzelne sein Maß pflegt. Ob das genügt, wenn ein Staat oder ein Bündnis sich gegen Überflutung und Täuschung von außen verteidigen will, oder ob er die Gewichte für seine Bürger dann selbst setzen muss, ist inzwischen eine Frage der Sicherheitspolitik.

Weiterlesen: *Wahrheit* (2025), Kap. 18 (die erschlichene Prüfwürdigkeit und die Überflutung)

Zusammenfassung

Die Überflutung zielt nicht darauf, zu überzeugen, sondern auf das Maß, mit dem überzeugt wird. Eine Studie der RAND Corporation beschrieb 2016 die russische Propaganda als Feuerwehrschlauch aus Menge, Wiederholung und Gleichgültigkeit gegen Wirklichkeit und Widerspruchsfreiheit; ein amerikanischer Wahlkampfstratege empfahl 2018 im eigenen Land, das Feld zu überfluten. Die Menge wirkt durch Wiederholung, die Vertrautheit wie Häufigkeit erscheinen lässt, durch Einebnung, die gesichert, strittig und unbekannt ununterscheidbar macht, und durch Erschöpfung derer, die jede Behauptung einzeln widerlegen. Darum hilft gegen sie weniger die Prüfung der einzelnen Aussage als das, was das Maß stabil hält: Langsamkeit, eine Bilanz weniger Quellen verschiedener Richtung, die Trennung von Gesichertem, Strittigem und Unbekanntem und der Blick auf die Quelle hinter der Quelle. Die RAND-Studie selbst empfahl, das Publikum vorzuwarnen, statt jede Unwahrheit hinterher zu bekämpfen.

Kapitel 10: Kognitive Kriegsführung

10.1 Das Sicherheitsargument

Im April 2020, in den ersten Monaten der Pandemie, berichtete der Auswärtige Dienst der Europäischen Union, kremlnahe Quellen und russische Staatsmedien führten eine koordinierte Kampagne, um die Antwort der Union auf die Krise zu

untergraben und Verwirrung über Herkunft und Folgen des Virus zu stiften; amtliche chinesische Stellen versuchten ebenso koordiniert, die Schuld am Ausbruch von sich abzulenken.[7] Mehrere Staaten beschreiben in ihren Lehren die öffentliche Meinung anderer Länder ausdrücklich als ein Feld, auf dem gekämpft wird. Wer in einer offenen Gesellschaft für ihre Sicherheit verantwortlich ist, kann das nicht übergehen. Gerade weil sie offen ist, erreicht man ihre Bürger leicht, und gerade weil sie frei wählen, lohnt es sich für fremde Mächte, auf das einzuwirken, was sie für wahr halten.

Die Bündnisstaaten der NATO haben 2021 erklärt, feindliche Informationsaktivitäten einschließlich Desinformation gehörten zu den Bedrohungen, gegen die sie widerstandsfähig werden müssten, und diese Widerstandsfähigkeit sei eine nationale Verantwortung.[8] Eine Studie, die im Auftrag eines Kommandos des Bündnisses entstand und nach eigener Angabe nicht dessen Haltung wiedergibt, ging weiter. Sie beschrieb den menschlichen Geist als eigenes Feld der Auseinandersetzung, das die ganze menschliche Umgebung umfasse, Freund wie Feind, und jeden Nutzer moderner Informationstechnik als mögliches Ziel.[9] Das Wort dafür ist kognitive Kriegsführung. Die stärkste Fassung des Gedankens, der dahintersteht, lautet: Wenn die Wahrnehmung der eigenen Bevölkerung angegriffen wird, muss man sie verteidigen, und wer verteidigen will, muss das Feld verstehen, auf dem angegriffen wird. Das ist richtig. Die Frage ist, wie man verteidigt.

Weiterlesen: *Wahrheit* (2025), Kap. 17 (wenn der Staat das Prüfen lenkt; die Quellen zur kognitiven Kriegsführung in Anhang D)

7 European External Action Service, *EEAS Special Report Update: Short Assessment of Narratives and Disinformation around the COVID-19/Coronavirus Pandemic (Updated 2–22 April)*, 27. April 2020: „pro-Kremlin sources and Russian state media continue running a coordinated campaign"; zu China dort die koordinierte Verlagerung der Schuld und „covert Chinese operations" in sozialen Netzwerken.

8 *Strengthened Resilience Commitment*, Gipfel der NATO in Brüssel, 14. Juni 2021, Ziff. 4 („Resilience is a national responsibility and a collective commitment") und Ziff. 7 („hostile information activities, including disinformation").

9 François du Cluzel, *Cognitive Warfare*, NATO Innovation Hub, Norfolk, Januar 2021 (gefördert vom Allied Command Transformation, nach eigener Angabe nicht die Haltung des Kommandos oder der Mitgliedstaaten), S. 6: „Any user of modern information technologies is a potential target"; S. 29: „The Human Domain of Operations refers to the whole human environment, whether friend of foe" („of" so im Original).

10.2 Zwei Wege

Grob lassen sich zwei Wege unterscheiden. Der eine will das Bild, das die Bevölkerung sich macht, selbst in die richtige Richtung bringen: erwünschte Darstellungen stärken, unerwünschte schwächen, die eigene Erzählung gegen die fremde setzen, die Aufmerksamkeit lenken. Der andere will die Fähigkeit der Bevölkerung stärken, selbst zu gewichten: zu bemerken, wenn ein Bild aus wahren Teilen schief gebaut ist, die Menge von der Begründung zu unterscheiden, nach Nenner, Auswahl und Quelle zu fragen. Der erste Weg steuert die Wahrnehmung, der zweite stärkt die Kalibrierung.

Beide Wege werden von Menschen gegangen, die ihr Land schützen wollen, und keiner setzt böse Absichten voraus. Der Unterschied liegt in ihrem Bau, und an ihm lässt sich zeigen, warum der zweite die bessere Verteidigung ist, auch gegen fremde Manipulation und auch unter dem Gesichtspunkt der Sicherheit allein.

10.3 Warum Steuerung schwächer verteidigt

Eine Steuerung der Wahrnehmung schafft einen Hauptschalter. Wer festlegt, was die Bevölkerung für wichtig, häufig und wahr halten soll, legt die Gewichte für alle fest, und wenn er sich irrt, irren alle mit ihm. Die Sicherheitslehre kennt das Problem unter einem anderen Namen: Eine Stelle, deren Ausfall das ganze System lahmlegt, ist das erste, was ein Gegner sucht. Wird sie getäuscht, unterwandert oder nur schlecht beraten, verbreitet sie den Fehler mit der ganzen Reichweite, die man ihr zur Abwehr gegeben hat.

Die Steuerung arbeitet zudem mit denselben Mitteln, gegen die sie sich richtet: mit Auswahl, Betonung, Wiederholung und der Inszenierung von Ausgewogenheit. Für den Bürger sind gelenkte Gewichte von außen und von innen nicht zu unterscheiden, solange er nicht weiß, dass gelenkt wird. Eine Bevölkerung, die gewohnt ist, den Signalen einer Autorität zu folgen, statt selbst zu gewichten, ist darum auch für fremde Lenkung leichter zu erreichen. Ein Gegner muss dann nicht überzeugen; er muss nur die Signale nachahmen, denen man zu folgen gewohnt ist.

Steuerung wird früher oder später bekannt, in offenen Gesellschaften meist früher. Dann geschieht, was das achte Kapitel an den Gräuelberichten des Ersten Weltkriegs gezeigt hat: Das Vertrauen sinkt nicht nur für die gelenkte Botschaft, sondern für alles, was aus derselben Quelle kommt, auch für das Wahre. Verteidigung aber lebt von Vertrauen. Im Ernstfall muss eine Bevölkerung einer Warnung glauben, einer Aufforderung, Schutz zu suchen, einer Nachricht darüber, wo

Gefahr ist und wo nicht. Eine Behörde, die in ruhigen Zeiten die Wahrnehmung gesteuert hat, findet in diesem Augenblick weniger Glauben, als sie braucht, und ein Gegner, der das weiß, muss nur an die frühere Steuerung erinnern.

Schließlich wirkt Steuerung auf die zurück, die steuern. Wer täglich gewichtet, was andere sehen sollen, gerät in Gefahr, das eigene Lagebild nach demselben Maß zu gewichten. Einrichtungen, die nach außen dosieren, sehen nach innen schlechter, nicht sofort, aber mit der Zeit.

Weiterlesen: *Wahrheit*, Kap. 17 (nach innen wahr, nach außen falsch)

10.4 Warum Kalibrierung besser verteidigt

Eine Bevölkerung, die selbst gewichten kann, hat keinen Hauptschalter. Ihre Kalibrierung ist auf viele Köpfe verteilt, und wer sie verschieben will, muss viele einzeln erreichen, von denen jeder seine eigenen Proben macht. Das macht sie nicht unverwundbar. Es macht sie aber teuer zu täuschen, und darin liegt eine Abschreckung, die keine zentrale Abwehr bieten kann.

Die Studie über den Feuerwehrschlauch, aus dem Umfeld der amerikanischen Streitkräfte, kam selbst zu diesem Schluss. Gegen die Flut helfe weniger, jede Unwahrheit zu widerlegen, als das Publikum vorzuwarnen und ihm zu zeigen, wie die Technik arbeitet, also denen, auf die der Schlauch gerichtet ist, Regenmäntel anzuziehen. Ein Regenmantel lenkt nicht. Er schützt den, der ihn trägt, gleich wer gerade den Schlauch hält. Ähnlich ist der Auftrag einer schwedischen Behörde gefasst, die 2022 für die psychologische Verteidigung eingerichtet wurde: Sie soll unzulässige Informationsbeeinflussung gegen das Land erkennen, analysieren und bei ihrer Abwehr helfen und dabei die offene, demokratische Gesellschaft und die freie Meinungsbildung schützen.[10] Die Abwehr richtet sich dort auf das Verfahren, mit dem fremde Akteure einwirken, auf verdeckte Herkunft, gefälschte Absender, koordinierte Täuschung, nicht auf die Meinungen im eigenen Land.

Ein Index, der die Widerstandsfähigkeit europäischer Gesellschaften gegen Desinformation vergleicht, misst sie an Pressefreiheit, an Bildung, vor allem an der Lesefähigkeit, und am Vertrauen der Menschen zueinander; 2026 standen

10 Förordning (2021:936) med instruktion för Myndigheten för psykologiskt försvar, beschlossen am 15. Oktober 2021, in Kraft seit 1. Januar 2022, § 1 (Stärkung der „befolkningens motståndskraft"), § 2 Nr. 1 („identifiera, analysera och kunna lämna stöd i bemötandet av otillbörlig informationspåverkan") und § 2 letzter Absatz („värna det öppna och demokratiska samhället och den fria åsiktsbildningen").

Dänemark, Finnland, Irland und die Niederlande gemeinsam vorn.[11] Der Index sagt selbst, dass er nicht die Medienkompetenz misst, sondern Bedingungen, die sie vorhersagen. Keine dieser Bedingungen ist eine Form der Lenkung. Es sind die Bedingungen, unter denen Menschen selbst gewichten können.

Für die Einrichtungen, die ein Land verteidigen, bleibt dabei viel zu tun, und nichts davon verlangt, die Wahrnehmung zu steuern. Sie können aufdecken, wer hinter verdeckten Kampagnen steht, und ihre Belege so vorlegen, dass andere sie prüfen können; die Herkunft einer Botschaft zu klären, ist etwas anderes, als über ihren Inhalt zu entscheiden. Sie können die Bevölkerung vorwarnen, welche Verfahren gerade benutzt werden. Sie können ihre eigenen Lagebilder mit Nenner, Spanne und Unsicherheit veröffentlichen, auch die eigenen Irrtümer, und damit das Maß stärken, an dem fremde Behauptungen sich messen lassen müssen. Und sie können Bildung, freie Presse und unabhängige Zählungen als Teil der Verteidigung behandeln, denn das sind sie.

10.5 Die Grenze im Ernstfall

Ein Einwand bleibt: Im Ernstfall fehlt die Zeit, eine Bevölkerung zu kalibrieren. Wer zum Schutzraum soll, braucht eine Anweisung, keine Abwägung. Das stimmt, und es spricht nicht gegen den zweiten Weg. Eine Anweisung, was zu tun ist, ist etwas anderes als eine Lenkung dessen, was man für wahr und wichtig hält. Sie wird befolgt, wenn die Stelle, die sie gibt, vorher Vertrauen erworben hat, und dieses Vertrauen erwirbt sie in ruhigen Zeiten, indem sie gewichtet, wie die Sache es verlangt, und ihre Gewichte offenlegt. Was im Ernstfall gesagt werden muss, lässt sich danach öffentlich prüfen und begründen. Die Kalibrierung der Bürger ist so keine Alternative zur Führung in der Krise, sondern ihre Voraussetzung.

Die Fähigkeit zu gewichten ist allerdings nicht nur bei Bürgern und Behörden gefragt, sondern überall, wo Menschen gemeinsam entscheiden: in Gremien, in Fachgesellschaften und dort, wo ein Konsens festgestellt wird.

11 Marin Lessenski, *Measuring Vulnerability of Societies to Disinformation: The Media Literacy Index 2026*, Open Society Institute – Sofia, Januar 2026: 41 europäische Länder; Gewichtung Medienfreiheit 40, Bildung 45, Vertrauen 10, elektronische Beteiligung 5 Prozent; „an instrument for measuring if not media literacy itself, then predictors of media literacy“; Dänemark, Finnland, Irland und die Niederlande gemeinsam auf dem ersten Platz.

Zusammenfassung

Fremde Desinformation ist belegt, etwa durch den Auswärtigen Dienst der Europäischen Union in der Pandemie, und offene Gesellschaften sind ihr ausgesetzt, weil sie offen sind. Die NATO-Staaten nennen Widerstandsfähigkeit gegen feindliche Informationsaktivitäten eine nationale Verantwortung; Konzeptstudien beschreiben den menschlichen Geist, Freund wie Feind, als Feld der Auseinandersetzung. Verteidigen lässt er sich auf zwei Wegen: durch Steuerung der Wahrnehmung oder durch Stärkung der Kalibrierung. Steuerung schafft einen Hauptschalter, arbeitet mit den Mitteln des Gegners, gewöhnt die Bevölkerung an Signale, die ein Gegner nachahmen kann, zehrt, sobald sie bekannt wird, das Vertrauen auf, von dem Warnungen im Ernstfall leben, und trübt das Lagebild derer, die steuern. Kalibrierung ist verteilt und darum teuer zu täuschen; die RAND-Studie empfahl, das Publikum vorzuwarnen, die schwedische Behörde für psychologische Verteidigung soll fremde Beeinflussung abwehren und die freie Meinungsbildung schützen. Verteidiger können Herkunft aufdecken, vorwarnen, eigene Lagebilder mit Unsicherheit veröffentlichen und Bildung und freie Presse als Teil der Verteidigung behandeln. Anweisungen im Ernstfall bleiben nötig; befolgt werden sie, wo vorher Vertrauen erworben wurde.

TEIL V: GEMEINSAM GEWICHTEN

Kapitel 11: Gremien, Fachleute, Konsens

11.1 Achthundert Schätzungen

Auf einer Viehausstellung in Plymouth im Südwesten Englands konnte man 1907 das Gewicht eines Ochsen schätzen, nachdem er geschlachtet und zerlegt war. Rund achthundert Besucher gaben einen Zettel ab, und 787 davon waren lesbar. Der Naturforscher Francis Galton wertete sie aus. Die mittlere Schätzung, die, über und unter der gleich viele lagen, betrug 1.207 Pfund. Der Ochse wog 1.198.[1]

Keiner der Besucher war Fachmann für das Gewicht geschlachteter Ochsen, und die meisten lagen weit daneben. Zusammen lagen sie fast richtig, weil ihre Fehler in verschiedene Richtungen gingen und sich gegenseitig aufhoben. Das gelingt unter einer Bedingung: Jeder schätzte für sich, auf seinem Zettel, ohne zu wissen, was die anderen geschrieben hatten. Hätte vorher einer laut gerufen, was er meint, und hätten die anderen sich nach ihm gerichtet, wären die Fehler nicht mehr unabhängig gewesen. Sie hätten sich nicht aufgehoben, sondern addiert, wie die Filter im sechsten Kapitel.

Viele Gremien arbeiten genau umgekehrt. Zuerst spricht, wer den Vorsitz hat oder am meisten gilt, dann die anderen, und ihre Urteile sammeln sich um das erste wie die Schätzungen um die Zahl vom Glücksrad. Das Gremium hat dann nicht das Wissen von zehn Köpfen, sondern das von einem, bestätigt von neun. Eine einfache Abhilfe folgt aus Galtons Befund: Jeder hält sein Urteil schriftlich fest, bevor die Aussprache beginnt, und die Aussprache fängt mit den Abweichungen an, nicht mit der Meinung des Ranghöchsten.

1 Francis Galton, „Vox Populi“, *Nature* 75 (1907), Nr. 1949, S. 450–451 (West of England Fat Stock and Poultry Exhibition, Plymouth; 787 auswertbare Karten; Median 1.207 Pfund, tatsächliches Schlachtgewicht 1.198 Pfund).

11.2 Das Wissen im Raum

Am Abend vor dem Start der Raumfähre Challenger im Januar 1986 warnten Ingenieure des Herstellers der Feststoffraketen, die Dichtungsringe der Raketen seien bei der erwarteten Kälte nicht sicher. Ihr Unternehmen empfahl zunächst, nicht zu starten. Als die Raumfahrtbehörde widersprach, beriet die Leitung des Unternehmens neu. Einer der Beteiligten erinnert sich, wie ein leitender Manager den technischen Leiter aufforderte, seinen Ingenieurshut ab- und den Managerhut aufzusetzen. Der technische Leiter stimmte danach gegen seine Ingenieure. Die Raumfähre zerbrach am 28. Januar kurz nach dem Start, und sieben Menschen starben.[2]

Das Wissen, auf das es ankam, war im Raum. Es fehlte nicht, es wurde gewichtet, und das Gewicht entschied sich nicht an den Daten, sondern an der Rolle: Ein Urteil, das als technisches Urteil schwer gewogen hätte, wog als eines unter mehreren Gesichtspunkten des Managements weniger. Die Aufforderung, den Hut zu wechseln, ist in diesem Sinn ein genaues Bild. Sie änderte nicht das Wissen, sondern den Maßstab, an dem es gewogen wurde.

Die Lehre daraus betrifft jede Einrichtung, in der Fachleute vorbereiten und andere entscheiden. Das Gewicht eines fachlichen Einwands darf nicht davon abhängen, wer im Raum den höheren Rang hat oder welche Rolle gerade gilt. Es muss an dem hängen, was der Einwand zeigt, und wer ihn übergeht, muss sagen, warum er ihn für schwächer hält als die Gründe dagegen.

Weiterlesen: *Organisation* (2026), Kap. 5 (was Vorbereitende und Entscheidende einander schulden), *Ethik – Das Gute, das wir schon tun* (2026), Kap. 15.8 (was man wissen muss)

11.3 Der bestellte Widerspruch

1587 richtete Papst Sixtus V. ein Amt ein, dessen Inhaber in Verfahren zur Selig- und Heiligsprechung alles vortragen sollte, was gegen den Kandidaten sprach: Schwächen in seinem Leben, Zweifel an den berichteten Wundern, Übertreibungen der Fürsprecher. Der Volksmund nannte ihn den Advokaten des Teufels. 1983 wurde das Verfahren reformiert, und das Amt verlor danach an Einfluss.[3]

---

2 National Public Radio, „40 Years after Challenger, Lingering Guilt and Lessons Learned“, Januar 2026, nach der Erinnerung des Ingenieurs Brian Russell: „Bob, it’s time to take off your engineering hat and put on your management hat.“ Zuvor hatte das Unternehmen empfohlen, nicht zu starten, und nach dem Widerspruch der NASA umgeschwenkt.

3 *Encyclopaedia Britannica*, Artikel „devil’s advocate“: Einrichtung des Amts (promotor fidei) durch Sixtus V. 1587; Aufgabe, ungünstige Tatsachen vorzutragen; Neuordnung durch Johannes Paul II.

Der Gedanke dahinter reicht über die Kirche hinaus. Wo alle Beteiligten wünschen, dass eine Sache gelingt, wird das, was dagegen spricht, leicht zu leicht gewogen. Ein bestelltes Gegengewicht gibt ihm einen Platz, an dem es nicht übergangen werden kann, ohne dass es auffällt. Ein Gremium kann jemanden beauftragen, die stärkste Fassung der Gegenposition vorzutragen, oder eine Stelle, die am Ergebnis nichts gewinnt, um eine zweite Sicht bitten.

Das Mittel hat eine Gefahr, die es selbst erzeugt. Wird der Widerspruch zur Rolle, gewöhnen sich alle daran, ihn als Pflichtübung zu hören, und so verliert er gerade das Gewicht, das er schaffen sollte. Ein bestellter Widerspruch trägt nur, solange er tatsächlich gewinnen kann, also solange gelegentlich eine Entscheidung seinetwegen anders ausfällt, und solange man das nachsehen kann.

11.4 Konsens worüber

Der Weltklimarat schrieb 2021 im ersten Satz seiner Hauptaussagen, es sei eindeutig, dass der Einfluss des Menschen Atmosphäre, Ozeane und Land erwärmt hat. Für die Frage, wie stark sich die Erde langfristig erwärmt, wenn sich der Kohlendioxidgehalt der Luft verdoppelt, nannte er als besten Schätzwert drei Grad, mit einer Spanne, die enger sei als im Bericht davor.[4] Zwischen diesen beiden Aussagen liegt ein Unterschied der Gewissheit, und der Bericht drückt ihn aus: Die eine gilt als eindeutig, die andere als Schätzung mit Spanne.

Ein Konsens ist darum kein einzelnes Gewicht, sondern eine Reihe von Gewichten für verschiedene Fragen. Dass die Erde sich erwärmt und der Mensch daran Anteil hat, ist etwas anderes als die Frage, wie stark sie sich weiter erwärmt, und das wieder etwas anderes als die Frage, welche Folgen das an welchem Ort hat, was eine Maßnahme kostet und was sie bringt und welche Kosten man wem zumuten will. Für die erste Frage gibt es eine Übereinstimmung, die man schwer überschätzen kann; für die mittleren Fragen gibt es Spannen, die die Fachleute selbst angeben; die letzten sind nicht allein Fachfragen, sondern auch Fragen der Abwägung, in denen Fachleute keine größere Stimme haben als andere Bürger.

Wer sich auf einen Konsens beruft, schuldet darum die Angabe, worüber er besteht. Wer ihn bestreitet, schuldet dieselbe Angabe, sonst bestreitet er alle

---

1983 mit der Konstitution *Divinus perfectionis Magister*, danach habe der promotor fidei „little sway over the proceedings“.

4 IPCC, *Climate Change 2021: The Physical Science Basis*, Beitrag der Arbeitsgruppe I zum Sechsten Sachstandsbericht, Headline Statements der Zusammenfassung für politische Entscheidungsträger, A.1: „It is unequivocal that human influence has warmed the atmosphere, ocean and land“; A.4: „a best estimate of equilibrium climate sensitivity of 3°C, with a narrower range compared to AR5“.

Stufen auf einmal, auch die, die er gar nicht meint. Dasselbe gilt für jedes Fach, in dem Gremien Übereinstimmung feststellen, in der Medizin, in der Wirtschaft, in der Geschichte. Je genauer die Frage, desto genauer lässt sich sagen, wie viel ein Konsens wiegt, und desto weniger lässt er sich als Hebel für etwas benutzen, das er nicht deckt.

Gremien sind langsam, Fachleute sind teuer, und beide haben ihre eigenen Filter. Seit kurzem gibt es einen Gesprächspartner, der schnell und billig auf alles antwortet und dabei gewichtet, ohne dass man es sieht.

Zusammenfassung

Gemeinsames Urteilen ist besser als einsames, wenn die Urteile unabhängig sind: 787 Besucher einer Viehausstellung schätzten 1907 das Gewicht eines Ochsen im Mittel auf 1.207 Pfund, er wog 1.198. Wer zuerst den Ranghöchsten hört, macht aus zehn Urteilen eines; darum hält jeder sein Urteil fest, bevor die Aussprache beginnt. Vor dem Start der Challenger 1986 war die Warnung der Ingenieure im Raum und wurde untergewichtet, als ein Manager den technischen Leiter bat, den Ingenieurshut gegen den Managerhut zu tauschen; das Gewicht eines Einwands darf nicht an Rang oder Rolle hängen. Ein bestellter Widerspruch, wie das Amt des promotor fidei seit 1587, gibt dem Gegengewicht einen Platz, trägt aber nur, solange er gelegentlich gewinnt. Ein Konsens ist eine Reihe von Gewichten für verschiedene Fragen: Der Weltklimarat nennt den menschlichen Einfluss auf die Erwärmung eindeutig und ihr Ausmaß eine Schätzung mit Spanne. Wer sich auf einen Konsens beruft oder ihn bestreitet, schuldet die Angabe, worüber er besteht.

Kapitel 12: Künstliche Intelligenz

12.1 Was geschrieben wurde

Ein Sprachmodell hat gelernt, was Menschen geschrieben haben. Wenn es sagt, was gewöhnlich ist, was die meisten Fachleute meinen oder wie es in einem Land zugeht, gibt es wieder, wie häufig etwas in diesen Texten vorkam. Das ist eine andere Häufigkeit als die in der Welt, und sie ist auf dieselbe Weise gefiltert wie die Kanäle aus dem sechsten Kapitel. Geschrieben wird über das Ungewöhnliche mehr als über das Gewöhnliche, über die Hauptstädte mehr als über die Dörfer, in den Sprachen, in denen viel veröffentlicht und digitalisiert wurde, mehr als in den anderen, und von denen, die schreiben, mehr als von denen, die es nicht

tun. Was ein Modell über ein Land weiß, in dem wenig geschrieben wird, weiß es zum großen Teil aus dem, was andere über dieses Land geschrieben haben.

Ein Modell kann darum ein sehr belesener Gesprächspartner sein und doch dasselbe schiefe Maß haben wie ein Mensch, der viel gelesen und wenig gesehen hat. Anders als bei dem Menschen sieht man seiner Antwort nicht an, woher sie ihre Gewichte hat. Sie kommt in ruhigem, sicherem Ton, ob sie auf tausend Quellen beruht oder auf dreien, und dieser Ton ist kein Gewicht, sondern eine Eigenschaft der Sprache, die das Modell gelernt hat.

12.2 Die Neigung, zuzustimmen

Dazu kommt eine Neigung, die aus der Art stammt, wie solche Modelle verfeinert werden. Menschen bewerten ihre Antworten; das Modell lernt, Antworten zu geben, die gut bewertet werden. Untersuchungen haben gezeigt, dass verbreitete Assistenten dieser Art dazu neigen, dem Fragenden nach dem Mund zu reden, und dass menschliche Bewerter solche Antworten oft bevorzugen.[5] Für die Kalibrierung heißt das: Wer eine Frage so stellt, dass seine eigene Vermutung darin steckt, bekommt sie leicht bestätigt, mit Gründen, die gut klingen. Das Modell wird zum Spiegel, der das Maß des Fragenden zurückwirft, statt es zu prüfen.

Gegen die Neigung hilft, was auch bei Menschen hilft, nur gezielter. Man fragt, ohne die eigene Vermutung zu verraten. Man lässt das Modell die stärkste Fassung der Gegenposition vortragen, und dann die stärkste Fassung der eigenen, und vergleicht. Man stellt dieselbe Frage einmal von der einen und einmal von der anderen Seite her und sieht nach, ob die Antwort mitwandert. Wandert sie, so hat man über die Frage wenig erfahren und über das Modell einiges.

12.3 Ein Kanal für viele

Zeitungen gibt es viele, und jede hat ihren eigenen Filter; ihre Fehler gehen in verschiedene Richtungen und heben sich zum Teil auf, wie die Schätzungen in Plymouth. Sprachmodelle, die von sehr vielen Menschen gefragt werden, gibt es wenige. Wenn Millionen ihr Maß für das Gewöhnliche und das Strittige an denselben wenigen Modellen nachstellen, werden ihre Urteile einander ähnlicher, und ihre Fehler laufen in dieselbe Richtung. Das ist die Lage des Gremiums, in

5 Mrinank Sharma u. a., „Towards Understanding Sycophancy in Language Models“, arXiv:2310.13548 (2023): fünf KI-Assistenten zeigten das Verhalten durchgehend in vier Aufgaben; „sycophancy is a general behavior of RLHF models“, wohl auch, weil menschliche Bewertungen gefällige Antworten bevorzugen.

dem alle nach dem Ranghöchsten sprechen, und die des Hauptschalters aus dem zehnten Kapitel, nur ohne dass jemand ihn absichtlich eingerichtet hätte. Wer die Gewichtungen eines verbreiteten Modells verändert, sei es durch die Auswahl der Texte, aus denen es lernt, sei es durch die Regeln, nach denen es antworten soll, verändert das Maß aller, die ihm vertrauen.

Daraus folgt kein Rat, Modelle zu meiden, sondern einer, sie wie jede andere Quelle zu gebrauchen: als eine Stimme unter mehreren, deren Bilanz man kennt. Wer etwas Wichtiges entscheidet, fragt mehr als ein Modell, sieht bei Zahlen und Zitaten an der ursprünglichen Quelle nach und behandelt die Angabe „die meisten Fachleute meinen“ aus einem Modell wie dieselbe Angabe aus dem Mund eines Bekannten: als Hinweis, den man prüfen kann, nicht als Zählung.

12.4 Wofür es taugt

Als Werkzeug der zweiten Prüfung, der nach der Verteilung, kann ein Modell gerade das, was dem Einzelnen schwerfällt. Man kann es fragen, was in einer Darstellung fehlt, welche Gruppen, Zeiträume oder Gegenden sie übergeht. Man kann es eine Prozentangabe in Häufigkeiten umrechnen lassen, nach dem Nenner fragen, nach der Grundrate, nach anderen Schwellen, nach der Methode, mit der eine Zahl gewonnen wurde. Man kann es bitten, die stärkste Kritik an einer Studie zu nennen, die einem gefällt. Alle diese Fragen verlangen keine Entscheidung vom Modell, sondern Material, das man selbst prüft.

Dieselbe Fähigkeit hat eine Kehrseite. Ein Modell kann in Sekunden einen Text schreiben, der beide Seiten nennt und doch eine bevorzugt, oder hundert Fassungen derselben Behauptung, die verschieden klingen. Die Inszenierung von Ausgewogenheit und die Überflutung werden dadurch billig. Das Gegenmittel ist dasselbe wie in den beiden Kapiteln, die sie beschrieben haben: nach der Gewichtung fragen statt nach der Zahl der Stimmen, auf die Quelle hinter der Quelle sehen und die Menge nicht für ein Gewicht nehmen.

Bis hierher ging es um Gewichte, die andere setzen: Kameras, Kanäle, Redaktionen, Gremien, Staaten, Modelle. Am Ende hängt jedes dieser Gewichte aber an dem Maß, mit dem der Einzelne liest, und dieses Maß lässt sich selbst prüfen.

Weiterlesen: *Wahrheit* (2025), Kap. 19 (Sprachmodelle als Sediment ihrer Texte), *Organisation* (2026), Kap. 11 (Sprachmodelle in Einrichtungen)

Zusammenfassung

Ein Sprachmodell gibt wieder, wie häufig etwas in den Texten vorkam, aus denen es gelernt hat, und diese Häufigkeit ist gefiltert wie jeder Kanal: zum Ungewöhnlichen, zu den Hauptstädten, zu den Sprachen, in denen viel veröffentlicht wurde. Sein sicherer Ton ist kein Gewicht. Weil Menschen seine Antworten bewerten, neigen verbreitete Modelle dazu, dem Fragenden zuzustimmen; dagegen hilft, ohne die eigene Vermutung zu fragen, beide Seiten in ihrer stärksten Fassung vortragen zu lassen und zu prüfen, ob die Antwort mit der Frage wandert. Weil wenige Modelle sehr viele Menschen beraten, werden deren Urteile einander ähnlicher, und wer die Gewichte eines Modells ändert, ändert das Maß aller, die ihm vertrauen. Man gebraucht es darum als eine Stimme unter mehreren und prüft Zahlen an der Quelle. Als Werkzeug der zweiten Prüfung taugt es gut: Es kann Fehlendes, Nenner, Grundraten und Gegenkritik suchen. Dass es Inszenierung und Überflutung billig macht, verlangt dieselben Gegenmittel wie zuvor.

TEIL VI: SELBST KALIBRIEREN

Kapitel 13: Die eigene Waage

13.1 Wann ein Urteil kalibriert ist

Wer im Wetterbericht hört, die Regenwahrscheinlichkeit morgen liege bei dreißig Prozent, kann dem Satz am nächsten Tag weder recht noch unrecht geben; es regnet oder es regnet nicht. Prüfen lässt er sich erst an vielen Tagen. Ein Vorhersager ist in einem genauen Sinn kalibriert, wenn es an den Tagen, für die er dreißig Prozent angibt, ungefähr an drei von zehn regnet, an den Tagen mit siebzig Prozent an sieben von zehn. Untersuchungen haben gezeigt, dass Meteorologen solche Wahrscheinlichkeiten zuverlässig angeben können.[1]

Dass gerade sie es können, hat Gründe, die sich übertragen lassen. Sie urteilen jeden Tag über dieselbe Art von Frage, sie erfahren am nächsten Tag, wie es ausging, und ihre Treffer werden gezählt. Die meisten Urteile im Leben haben nichts davon. Man urteilt über verschiedene Dinge, erfährt den Ausgang spät oder nie, und niemand zählt. Ein Maß, das nie an Ausgängen geprüft wird, kann jahrzehntelang schief stehen, ohne dass sein Besitzer es merkt.

13.2 Eine Probe

Experiment 4: Neun von zehn

Geben Sie für jede der folgenden Größen eine Spanne an, in der der wahre Wert nach Ihrer Überzeugung mit neunzig Prozent Wahrscheinlichkeit liegt. Die Spanne soll weder so eng sein, dass Sie sich nur wenig zutrauen müssten, noch so weit, dass sie nichts mehr sagt, sondern so, dass Sie bei zehn solchen Spannen neun Treffer erwarten. Schlagen Sie nichts nach.

1 Allan H. Murphy, Robert L. Winkler, „Reliability of Subjective Probability Forecasts of Precipitation and Temperature“, *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 26 (1977), S. 41–47: „These results indicate that weather forecasters can formulate such forecasts in a reliable manner.“

Die mittlere Entfernung von der Erde zum Mond. Die Höhe des Mount Everest. Der Umfang der Erde am Äquator. Die Zahl der Knochen im Skelett eines erwachsenen Menschen. Die Geschwindigkeit des Lichts im Vakuum. Die Werte: etwa 384.400 Kilometer; 8.849 Meter, nach der gemeinsamen Vermessung Chinas und Nepals von 2020; etwa 40.075 Kilometer; 206; etwa 299.792 Kilometer in der Sekunde.[a] Zählen Sie, wie viele Werte in Ihren Spannen liegen.

Bei fünf Größen ist das Ergebnis eine kleine Stichprobe, und ein Fehlgriff kann Zufall sein. Wer aber nur zwei oder drei Treffer hat, ist nicht allein: In Versuchen dieser Art sind die Spannen, die Menschen für neunzig Prozent sicher halten, meist zu eng.[b] Man traut seinem Maß mehr, als es verdient. Wer dagegen fünf von fünf Treffern hat, mag Spannen gewählt haben, die weiter sind, als er bei klarem Kopf vertreten würde, und auch das ist eine Abweichung, nur in die andere Richtung.

---

a Mond: NASA, Moon Fact Sheet (große Halbachse 0,3844 × 10^6 km); Everest: 8.848,86 m, gemeinsam bekanntgegeben von China und Nepal am 8. Dezember 2020 (u. a. *Al Jazeera*, 8. Dezember 2020); Äquatorumfang aus dem Äquatorradius 6.378,137 km (NASA, Earth Fact Sheet); Knochen: National Cancer Institute, SEER Training Modules, „Divisions of the Skeleton"; Lichtgeschwindigkeit: 299.792,458 km/s, seit 1983 durch die Definition des Meters festgelegt.

b Marc Alpert, Howard Raiffa, „A Progress Report on the Training of Probability Assessors", in: Daniel Kahneman, Paul Slovic, Amos Tversky (Hg.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge 1982, S. 294–305; zusammenfassend David R. Mandel u. a., „Effect of Confidence Interval Construction on Judgment Accuracy", *Judgment and Decision Making* 15 (2020), S. 783–797: „the confidence intervals people generate are often too narrow or overly precise"; nach Budescu und Du (2007) seien neunzigprozentige Spannen zu eng gewesen, siebzigprozentige gut kalibriert.

Das Experiment prüft nicht, was Sie wissen, sondern wie gut Ihre Gewissheit zu Ihrem Wissen passt. Das ist die Kalibrierung im genauesten Sinn. Sie lässt sich üben wie jede Fertigkeit.

13.3 Buch führen

Was die Meteorologen von Berufs wegen haben, kann sich jeder für einen Teil seiner Urteile selbst einrichten. Man schreibt sich auf, was man erwartet, mit einer Wahrscheinlichkeit dazu: dass ein Vorhaben bis zum Herbst fertig wird, dass eine Wahl so ausgeht, dass ein Streit sich beilegen lässt, dass eine Zahl im nächsten Jahr steigt. Nach einer Weile sieht man nach. Wenige Dutzend solcher

Notizen reichen, um zu sehen, ob man seine Gewissheit eher zu hoch oder zu niedrig ansetzt, und auf welchen Gebieten. Es kann sich zeigen, dass man bei technischen Fragen gut liegt und bei politischen schlecht, oder umgekehrt.

Ein solches Buch der Erwartungen ist die Kalibrierung der Kalibrierung: nicht nur das Urteil über die Welt prüfen, sondern das Urteil über die eigene Verlässlichkeit. Sie hat einen Nebennutzen, der vielleicht wichtiger ist als der Hauptnutzen. Wer weiß, dass er seine Erwartungen aufschreibt, formuliert sie genauer und kann sie hinterher nicht mehr so umdeuten, dass er schon immer recht hatte.

13.4 Die Symmetrieprüfung

Die schwerste Verzerrung des eigenen Maßes ist die, die mit den eigenen Wünschen läuft. Eine Studie, die zur eigenen Überzeugung passt, wiegt leicht schwerer als eine, die ihr widerspricht, auch wenn beide gleich gut gemacht sind. Ein einzelner Fall auf der Gegenseite gilt als Ausnahme, einer auf der eigenen als Zeichen. Nachgeprüft wird vor allem, was stört.

Dagegen hilft eine Probe, die die Etiketten vertauscht. Man nimmt eine Meldung, die einen empört oder erfreut, und ersetzt in Gedanken die Beteiligten durch die der anderen Seite: die Regierung, die man ablehnt, durch die, die man wählt, das fremde Land durch das eigene, die Partei durch ihren Gegner. Dann fragt man, ob man die Meldung genauso gewichten würde, ob man genauso schnell glauben, genauso streng prüfen, genauso laut empört sein würde. Wo die Antwort nein ist, hat man nicht entdeckt, wer recht hat, aber man hat entdeckt, wo das eigene Maß an der Zugehörigkeit hängt statt an der Sache. Dieselbe Probe gilt für Studien: Würde ich diese Methode akzeptieren, wenn sie das Gegenteil ergeben hätte?

Die Probe hat eine Grenze. Nicht alle Lagen sind symmetrisch; manchmal hat eine Seite tatsächlich mehr Unrecht getan oder bessere Gründe. Die Symmetrieprüfung verlangt nicht, beide gleich zu beurteilen, sondern dieselben Maßstäbe an beide anzulegen. Was dann verschieden ausfällt, ist verschieden, weil die Sache es ist.

13.5 Die stärkste Gegenseite

Ein Maß für die eigene Überzeugung lässt sich nur gewinnen, wenn man weiß, wogegen sie sich behauptet. Wer die Gegenseite nur aus ihren schwächsten Vertretern kennt, aus den lautesten, den ungenauesten, denen, über die die eigene Seite lacht, hält die eigene Position für stärker, als sie ist, weil er sie nie

an ihrem eigentlichen Gegengewicht gewogen hat. Kalibrieren heißt hier, die stärkste Fassung der Gegenposition zu suchen, in ihren eigenen Worten und von ihren besten Vertretern, und erst dann zu sehen, wie schwer die eigenen Gründe wiegen. Oft bleibt man bei seiner Überzeugung. Aber man weiß dann, wie sicher man sein darf, und man kann sagen, woran man es messen würde, wenn man sich irrte.

John Stuart Mill hat das 1859 in einen Satz gefasst: Wer nur seine eigene Seite einer Sache kennt, weiß wenig von ihr.[2]

Die stärkste Gegenseite zu kennen, verlangt freilich, dass es sie gibt und dass sie sprechen kann. Das ist keine Sache des Einzelnen mehr, sondern des Raums, in dem viele gemeinsam kalibrieren.

Weiterlesen: *Wahrheit* (2025), Kap. 19 und 11 (der gekippte Filter; wann man zuletzt umgedreht wurde)

Zusammenfassung

Kalibriert ist ein Urteil, wenn die angegebene Gewissheit zu den Ausgängen passt: An den Tagen, für die ein Meteorologe dreißig Prozent Regen angibt, regnet es an etwa drei von zehn, und Untersuchungen zeigen, dass Wettervorhersager das zuverlässig können, weil sie täglich urteilen, den Ausgang erfahren und gezählt werden. Die meisten Urteile haben das nicht. In Versuchen sind Spannen, die Menschen für neunzig Prozent sicher halten, meist zu eng; man traut dem eigenen Maß mehr, als es verdient. Wer seine Erwartungen mit Wahrscheinlichkeiten aufschreibt und später nachsieht, prüft nicht nur sein Urteil über die Welt, sondern das über seine Verlässlichkeit. Die Symmetrieprüfung vertauscht die Etiketten und fragt, ob man dieselbe Meldung über die andere Seite gleich gewichten würde; sie verlangt nicht gleiche Urteile, sondern gleiche Maßstäbe. Und das Gewicht der eigenen Überzeugung erfährt nur, wer die stärkste Fassung der Gegenseite kennt.

Kapitel 14: Gemeinsam kalibrieren

14.1 Warum es mehrere Stimmen braucht

Die Schätzungen in Plymouth trafen, weil viele unabhängig voneinander schätzten. Die Fehler der Kanäle, aus denen man ein fernes Land kennt, addierten sich,

2 John Stuart Mill, *On Liberty*, London 1859, Kap. II („Of the Liberty of Thought and Discussion"): „He who knows only his own side of the case, knows little of that."

weil sie in dieselbe Richtung filterten. Beides zusammen ergibt eine Bedingung, ohne die Kalibrierung nicht gelingt: Man braucht Stimmen, die verschieden filtern und nicht voneinander abschreiben. Eine Waage lässt sich nur an einem Gewicht prüfen, das nicht auf ihr selbst gewogen wurde.

Daraus folgt, dass die Vielfalt der Zeitungen, Schulen, Fachrichtungen und Meinungen in einer Gesellschaft mehr ist als eine Höflichkeit gegenüber Andersdenkenden. Sie ist das Instrument, mit dem eine Gesellschaft ihre Gewichte prüft. Wo nur ein Kanal spricht, oder viele, die alle dasselbe sagen, weil abweichende Stimmen Nachteile fürchten, kann niemand mehr feststellen, ob das gemeinsame Maß stimmt, auch die nicht, die es setzen. Pluralismus ist in diesem Sinn keine Haltung, die man haben kann oder nicht, sondern eine Bedingung des Wissens über Gewichte.

Pluralismus ist aber nicht Beliebigkeit. Vielfalt dient der Kalibrierung, solange die Stimmen sich auf dieselbe Sache beziehen und sich an ihr prüfen lassen, wie die Blinden am selben Elefanten. Eine Stimme, die nicht sagen kann, was sie widerlegen würde, und keine Gründe gibt, die andere prüfen können, trägt zur gemeinsamen Waage nichts bei, so laut sie ist.

Weiterlesen: *Wahrheit* (2025), Kap. 20 (orientierter Pluralismus), *Wahrheit*, Kap. 16 (gemeinsam prüfen braucht Verfahren)

14.2 Strittiges selbst gewichten

In heftigen Streitfragen der Gegenwart, ob über das Klima, über Impfungen oder über einen Krieg, gibt es meist ein vorherrschendes Bild und ein Gegenbild. Beide treten als Ganzes auf. Wer dem einen folgt, übernimmt alle seine Gewichte, wer dem anderen folgt, ebenso. Selbst gewichten kann nur, wer das Ganze in Fragen zerlegt, die sich einzeln prüfen lassen, denn die Teilfragen haben verschiedene Gewissheit, und die beiden Bilder verdecken das gleichermaßen.

Am Klima hat das elfte Kapitel die Zerlegung schon vorgeführt: von der Frage, ob der Mensch die Erde erwärmt, über die, wie stark, bis zu der, was man tun soll und wer die Kosten trägt. Bei Impfungen lässt sich ebenso zerlegen. Wie gut schützt ein Impfstoff vor schwerer Krankheit, und in welcher Gruppe? Wie häufig sind welche Nebenwirkungen, wieder in welcher Gruppe? Wie hoch ist das Risiko der Krankheit selbst für Menschen dieser Gruppe? Und was folgt daraus für eine Empfehlung, eine Pflicht oder eine freie Entscheidung? Die ersten drei Fragen sind Fragen nach Häufigkeiten, und für sie gilt alles, was der zweite Teil über Grundraten, Nenner und Auswahl gesagt hat. Die vierte ist eine Abwägung, in der auch die Freiheit des Einzelnen und die Lage der anderen zählen. Die

stärkste Fassung der einen Seite stützt sich meist auf den Nutzen für die, die am stärksten gefährdet sind, und auf das, was eine Krankheit anrichtet, die sich ungebremst ausbreitet. Die stärkste Fassung der anderen stützt sich darauf, dass der Nutzen für Menschen mit geringem Risiko klein sein kann, so dass auch seltene Nebenwirkungen ins Gewicht fallen, und darauf, dass eine Pflicht mehr begründen muss als eine Empfehlung. Beide Fassungen können je nach Gruppe recht haben, und wer sie nebeneinanderstellt, sieht, dass ihr Streit zu einem großen Teil ein Streit darüber ist, in welcher Gruppe man rechnet.

Bei Kriegen kommt hinzu, dass beide Seiten ein Interesse an ihrem Bild haben, das über die Wahrheit hinausgeht. Auch hier hilft das Zerlegen. Was ist durch Quellen belegt, die keiner der Parteien gehören: Bilder, deren Ort und Zeit sich prüfen lassen, Dokumente, unabhängige Zählungen? Was sagen die Zählungen der Opfer, nach welcher Methode, mit welcher Spanne, wie es das fünfte Kapitel am Irakkrieg gezeigt hat? Welche Vorgeschichte erzählt jede Seite, und welche lässt sie weg? Und was ist schlicht noch nicht bekannt? Wer diese Fragen für einen Krieg beantwortet, kommt selten zu einem Bild, das einer der beiden Seiten ganz gefällt. Das ist kein Mangel des Verfahrens, sondern oft sein Ergebnis.

14.3 Keine falsche Mitte

Selbst gewichten heißt nicht, in jeder Frage die Mitte zwischen zwei Lagern zu suchen. Ergibt die Prüfung, dass eine Teilfrage fast eindeutig beantwortet ist, dann ist die kalibrierte Antwort fast eindeutig, auch wenn ein Lager sie bestreitet. Ergibt sie, dass eine Frage offen ist, dann ist die kalibrierte Antwort „offen“, auch wenn beide Lager Gewissheit verkünden. Die falsche Balance, die einer Randmeinung so viel Gewicht gibt wie dem gut Belegten, ist so wenig kalibriert wie das Lagerdenken, das die Randmeinung gar nicht erst prüft.

Zur Kalibrierung gehört darum auch, „ich weiß es nicht“ sagen zu können und eine Frage offen zu halten, statt sie mit dem Gewicht des eigenen Lagers zu schließen. Wer das kann, wird in Streitgesprächen seltener gewinnen. Er wird aber seltener auf eine Seite festgelegt, deren Gewichte er nie geprüft hat.

14.4 Mit anderen gewichten

Die stärkste Hilfe beim Kalibrieren sind Menschen, die anders gewichten und es begründen können. Ein Gespräch mit ihnen ist ergiebig, wenn beide vorher sagen, was sie umstimmen würde: welche Zahl, welcher Befund, welche Quelle. Dann streitet man nicht mehr über das Ergebnis, sondern darüber, wie man es prüfen

könnte, und oft stellt sich heraus, dass man sich über die Tatsachen einiger ist als über ihr Gewicht, oder über das Gewicht einiger als über die Werte, an denen man es misst. Beides ist ein Fortschritt, denn erst dann weiß man, worüber man eigentlich uneins ist.

Ein solches Gespräch setzt voraus, dass der andere nicht als Feind gilt, dessen Gründe man nur widerlegen muss, sondern als ein Instrument, das anders geeicht ist als das eigene und darum anzeigt, wo das eigene abweichen könnte. Diese Haltung ist ein Schutz gegen das, was dieses Buch selbst anrichten könnte. Denn wer gelernt hat, wie Gewichte entstehen und verschoben werden, kann dieses Wissen auch gebrauchen, um sie zu verschieben.

Zusammenfassung

Kalibrieren braucht Stimmen, die verschieden filtern und nicht voneinander abschreiben; darum ist die Vielfalt der Medien, Schulen und Meinungen keine Höflichkeit, sondern das Instrument, mit dem eine Gesellschaft ihre Gewichte prüft. Sie dient diesem Zweck, solange die Stimmen sich auf dieselbe Sache beziehen und Gründe geben, die sich prüfen lassen. In Streitfragen der Gegenwart treten Bild und Gegenbild als Ganzes auf; selbst gewichten kann, wer sie in Teilfragen zerlegt, die verschiedene Gewissheit haben. Beim Klima reicht die Zerlegung vom Ob über das Wieviel zum Was-tun, bei Impfungen von Schutz, Nebenwirkung und Krankheitsrisiko in einer bestimmten Gruppe zur Abwägung von Empfehlung und Pflicht, bei Kriegen von unabhängig Belegtem über Zählungen und Vorgeschichten zum noch Unbekannten. Kalibrieren heißt nicht, die Mitte zu suchen; es schließt ein, „ich weiß es nicht“ zu sagen. Am meisten hilft das Gespräch mit Menschen, die anders gewichten, wenn beide vorher sagen, was sie umstimmen würde.

Kapitel 15: Verkehrte Kalibrierung

15.1 Wissen, das in beide Richtungen wirkt

Wer weiß, wie eine Nahaufnahme eine Menge vergrößert, wie ein relatives Risiko erschreckt, wie eine Schwelle ein Bild umkehrt, kann das alles erkennen. Er könnte es auch benutzen. Ein Buch über Gewichte lässt sich darum gegen den Strich lesen; verhindern kann das kein Satz. Es gibt aber einen Grund, warum solches Wissen dem, der sich schützen will, mehr nützt als dem, der täuschen will. Wer Gewichte verschiebt, braucht ein Publikum, das nicht weiß, wie es geschieht;

für ihn sind die Techniken längst bekannt. Wer sie erkennen will, erfährt sie meist erst, wenn er auf sie hereingefallen ist. Je mehr Menschen wissen, wie verzerrte Gewichte entstehen, desto weniger wirken sie. Das war der Rat der Studie über den Feuerwehrschlauch, das Publikum vorzuwarnen, und es ist der Grund, warum dieses Buch die Verfahren beschreibt, immer von der Seite dessen, der sie bemerken und berichtigen will.

Gefährlicher als die Benutzung der beschriebenen Techniken ist ihre Verkehrung: dass die Mittel der Kalibrierung selbst, die Fragen dieses Buchs, so gebraucht werden, dass sie nicht mehr prüfen, sondern ein Ergebnis sichern, das vorher feststand. Das geschieht selten mit Absicht. Meist geschieht es Menschen, die überzeugt sind, besser zu gewichten als andere.

Weiterlesen: *Wahrheit* (2025), Kap. 17 (verkehrte Prinzipien), *Wahrheit*, Kap. 18 (wenn die Methode selbst gefälscht wird)

15.2 Die Fragen dieses Buchs, verkehrt

Die Frage „Was fehlt?“ ist das wichtigste Werkzeug dieses Buchs und am leichtesten zu verkehren. Verkehrt wird sie, wenn jedes Fehlen als Beweis gilt: Darüber wird nicht berichtet, also wird es verschwiegen, also ist es wahr und wichtig. Eine so gestellte Frage kann durch nichts mehr beantwortet werden, denn auch die Antwort, die man bekommt, gilt dann als Teil der Verschleierung. Man erkennt die Verkehrung daran, dass niemand angeben kann, was man sehen müsste, wenn nichts verborgen wäre. Das Gegenmittel ist, diese Angabe zu verlangen und dann nachzusehen.

Die Grundrate lässt sich verkehren, indem man mit ihr jeden einzelnen Fall abweist: Das ist selten, also zählt es nicht. Eine Grundrate verteilt aber Aufmerksamkeit und Beweislast, nicht Wahrheit; der einzelne Fall ist selbst ein Befund, und wenn er sich häuft, ändert er die Rate. Man erkennt die Verkehrung daran, dass auf die Rate verwiesen wird, wo der Fall geprüft werden müsste.

Nenner und Schwelle lassen sich so wählen, dass das gewünschte Bild entsteht, und zwar gerade mit dem Anspruch der Genauigkeit: Ich zeige ja nur die Zahlen. Man erkennt es daran, dass nur eine Schwelle, ein Zeitraum, ein Nenner vorkommt, und das Gegenmittel ist die Frage aus dem vierten Kapitel, ob das Bild sich hält, wenn man sie ändert. Wer die anderen Fassungen nicht zeigen will, hat einen Grund dafür.

Die Symmetrieprüfung lässt sich in zwei Richtungen verkehren. In der einen wird aus gleichen Maßstäben ein gleiches Urteil, und jede Sache bekommt ihre zwei gleich starken Seiten, auch wo die Belege ungleich sind. In der anderen

wird sie zur Abwehr: Auf jede Kritik an der eigenen Seite folgt der Hinweis auf die Vergehen der anderen, und damit ist die Kritik erledigt, ohne geprüft zu sein. Beides verfehlt, was die Probe verlangt, nämlich dieselben Maßstäbe an beide anzulegen und dann zu sehen, was verschieden ausfällt.

Auch die Sprache der Kalibrierung selbst lässt sich verkehren. Wer anderen Verzerrung, Auswahl, verfehlte Grundraten vorhält und es bei sich nie prüft, hat aus einem Werkzeug eine Waffe gemacht. Und wo die Stärkung der Urteilskraft angekündigt und das Einüben bestimmter Urteile betrieben wird, ist aus der Kalibrierung die Steuerung geworden, von der das zehnte Kapitel handelte, nur unter anderem Namen.

15.3 Woran man die Verkehrung erkennt

Die verkehrten Formen haben gemeinsame Zeichen, und sie lassen sich an jedem Text, jeder Schulung und jeder Rede prüfen, auch an diesem Buch. Das erste ist die Einseitigkeit: Das Werkzeug wird nur in eine Richtung angewandt, gegen die anderen, nie gegen die eigene Seite. Das zweite ist die Unwiderlegbarkeit: Es gibt keinen Befund, der die Frage beenden oder das Urteil ändern könnte. Das dritte ist der Abbruch: Die Prüfung hört genau dort auf, wo sie dem eigenen Ergebnis gefährlich würde. Das vierte ist die Richtung, in die der Zuhörer gebracht wird: Wird er unabhängiger, kann er hinterher allein besser gewichten, oder braucht er den, der ihm die Gewichte erklärt, mehr als vorher?

Die vier Zeichen hängen zusammen. Alle zeigen ein Ergebnis an, das feststand, bevor geprüft wurde, und eine Prüfung, die nur noch dazu dient, es zu schützen. Wo sie zusammen auftreten, ist nicht mehr zu entscheiden, wer recht hat, sondern ob überhaupt noch geprüft wird.

15.4 Wie man sich und andere schützt

Den sichersten Schutz gibt die Anwendung auf sich selbst, zuerst und am strengsten. Wer die Fragen dieses Buchs zuerst an die eigenen Überzeugungen stellt, an die Gewichte der eigenen Seite, an die eigenen Lieblingszahlen, kann sie schwer als Waffe gebrauchen, denn er hat erfahren, wie sie schneiden.

Ebenso wirksam ist die Offenlegung. Verzerrte Gewichte wirken, solange sie verborgen sind; gezeigt, verlieren sie ihre Kraft. Wer Zahlen mit Nenner, Spanne und Schwelle gibt, Berichte mit dem, was sie nicht zeigen, und Urteile mit dem, was sie ändern würde, macht es jedem schwer, ihn zu verkehren, auch sich selbst.

Dazu kommt die Verbreitung. Je mehr Menschen die Verfahren kennen, desto weniger lohnen sie sich. Wer anderen dabei hilft, sollte ihnen zeigen, wie man gewichtet, nicht, wie sie gewichten sollen. Wer anderen seine Gewichte als die richtigen übergibt, auch in bester Absicht, macht sich zu dem Hauptschalter, vor dem dieses Buch warnt, und wer seine Fragen weitergibt, verteilt die Kalibrierung auf mehr Köpfe.

Wo viele zusammen entscheiden, helfen Formen, die die Gewichtung sichtbar halten: Vorlagen, die die verworfenen Möglichkeiten nennen, Register, die vor dem Ergebnis angelegt werden, Gremien, in denen jeder zuerst für sich urteilt, Zählungen, die unabhängig von den Parteien sind. Sie verhindern die Verkehrung nicht, aber sie machen sie sichtbar, und was sichtbar ist, lässt sich berichtigen.

Was keiner dieser Schutzmittel leistet, ist eine Sicherheit, dass man selbst richtig gewichtet. Darin liegt die letzte Grenze der Kalibrierung.

Zusammenfassung

Wissen über Gewichte wirkt in beide Richtungen, aber es nützt dem, der sich schützen will, mehr als dem, der täuschen will, denn Täuschung braucht ein Publikum, das die Verfahren nicht kennt; darum beschreibt dieses Buch sie aus der Sicht dessen, der sie bemerken will. Gefährlicher ist die Verkehrung der Kalibrierung selbst: die Frage „Was fehlt?“, nach der jedes Fehlen ein Beweis ist; die Grundrate, mit der man jeden Einzelfall abweist; Nenner und Schwellen, die nur in der passenden Fassung gezeigt werden; die Symmetrie als falsche Gleichheit oder als Verweis auf die Vergehen der anderen; die Sprache der Kalibrierung als Waffe gegen andere und als Name für Steuerung. Erkennbar ist die Verkehrung an Einseitigkeit, Unwiderlegbarkeit, Abbruch der Prüfung, wo sie unbequem würde, und daran, dass der Zuhörer abhängiger wird. Schutz geben die Anwendung auf sich selbst, die Offenlegung der Gewichte, die Verbreitung der Fragen statt der Antworten und Formen, die die Gewichtung sichtbar halten.

Kapitel 16: Grenzen

16.1 Wo keine Rückmeldung kommt

Die Meteorologen konnten ihre Gewissheit an Tausenden von Tagen prüfen. Viele der wichtigsten Urteile lassen sich so nicht prüfen, weil die Lage selten ist oder

einmalig. Ob ein Land einen Krieg beginnen, eine Seuche mit harten Maßnahmen bremsen, eine Währung aufgeben soll, entscheidet man nicht tausendmal, sondern einmal, und was geschehen wäre, wenn man anders entschieden hätte, erfährt man nie. Für solche Urteile gibt es keine Trefferquote, an der sich die eigene Gewissheit messen ließe.

Man kann sich behelfen, indem man nach ähnlichen Lagen sucht, die es früher gab, und nach der Rate, mit der sie so oder anders ausgingen. Das ist die Grundrate aus dem dritten Kapitel, und mit ihr ihr Problem: Je seltener die Lage, desto dünner und unsicherer die Gruppe vergleichbarer Fälle, und desto mehr hängt an der Wahl, welche Fälle man für vergleichbar hält. Wo die Kalibrierung so dünn wird, folgt aus ihr vor allem eine Haltung: Spannen statt Punkte, Entscheidungen, die sich zurücknehmen lassen, wo das möglich ist, und die Bereitschaft, früh nachzusehen, ob die Annahmen tragen.

16.2 Das Fehlende, das sich nicht zeigt

Manches fehlt so, dass keine Frage es findet, weil niemand sie stellen kann, bevor sich etwas zeigt. Erst im Rückblick wird sichtbar, dass eine Voraussetzung, die allen selbstverständlich war, nicht stimmte. Der allgemeine Vorbehalt, dass etwas fehlen könnte, hilft dagegen wenig, denn er gilt für jedes Urteil gleich und sagt nicht, wo. Er ist kein Grund, ein bestimmtes Urteil zu verwerfen. Er ist ein Grund, die offenen Stellen sichtbar zu halten, das Unerklärte nicht wegzulegen, und dorthin zu sehen, wo etwas nicht aufgeht.

16.3 Kein Hauptschalter, kein Abschluss

Kalibrierung ist nie fertig. Die Welt ändert sich, und mit ihr die Raten, an denen das eigene Maß zu prüfen wäre; was vor zwanzig Jahren stimmte, ist heute vielleicht eine Erinnerung. Das eigene Maß verschiebt sich auch ohne Änderung der Welt, durch das, was man gerade liest und wem man gerade zuhört. Was einmal nachgestellt wurde, muss darum wieder geprüft werden, nicht ständig, aber regelmäßig, und vor allem dort, wo viel davon abhängt.

Diese Pflicht zur Wiederholung kann lähmen, wenn man sie falsch versteht. Kalibrieren heißt nicht, so lange zu prüfen, bis keine Unsicherheit mehr bleibt; die bleibt immer. Es heißt, mit dem am besten geprüften Maß zu handeln, das man hat, und zu wissen, wie sicher es ist. Wer wartet, bis alles gewichtet ist, hat auch gewichtet, nur zugunsten dessen, was ohne ihn geschieht.

Eine Gesellschaft, in der viele so urteilen, hat keinen Hauptschalter, an dem sich ihr Maß umstellen ließe. Das macht sie langsamer in der Einigung und schwerer zu lenken, auch von denen, die es gut meinen. Es macht sie aber auch schwer zu täuschen. Ihre Fehler kann sie bemerken, weil nicht alle sie zugleich machen.

16.4 Dieses Buch als Auswahl

Auch dieses Buch ist eine Auswahl. Es muss sich an dem prüfen lassen, was es von anderen verlangt. Viele seiner Fälle stammen aus den Vereinigten Staaten und aus Großbritannien. Das liegt nicht daran, dass dort mehr verzerrt würde, sondern daran, dass diese Fälle gut untersucht und in Quellen zugänglich sind, die ein Autor schnell prüfen kann: in englischer Sprache, frei im Netz, oft zitiert. Das ist genau der Filter, von dem das sechste und das zwölfte Kapitel handeln. Wer in einem anderen Teil der Welt lebt, findet die Fälle seines eigenen Landes vermutlich nicht hier. Er sollte prüfen, ob die Gedanken an ihnen genauso tragen. Tun sie es nicht, hat das Buch an dieser Stelle unrecht.

Dasselbe gilt für die Auswahl der Themen. Ein Buch, das vor allem Fehler der Gewichtung beschreibt, kann den Eindruck erzeugen, alles sei falsch gewichtet. Die meisten Gewichte, mit denen Menschen im Alltag urteilen, sind gut genug für das, wozu sie gebraucht werden; sonst käme niemand durch den Tag. Kalibrieren muss man vor allem dort, wo viel abhängt, wo die Sache fern ist und wo jemand ein Interesse am Bild hat, das man sich macht.

16.5 Der Schritt zurück

Am Firdos-Platz in Bagdad zeigte sich, was die Nahaufnahme verschwieg, wenn man einen Schritt zurücktrat: Je weiter man vom Sockel entfernt war, desto mehr sah man. Darin liegt, was Kalibrieren im Kern ist. Man tritt einen Schritt zurück von dem Bild, das einen erreicht, und fragt, wie groß der Ausschnitt ist, was außerhalb liegt, wie häufig das Gezeigte ist, wer es ausgewählt hat und wie schwer es neben dem Rest wiegt. Man stellt das eigene Maß an etwas nach, das nicht aus demselben Strom kommt. Und man hält es offen für den nächsten Schritt zurück.

Weiterlesen: *Wahrheit* (2025), Kap. 10 und 20 (der Irrtum, der sich noch nicht gezeigt hat; der Wahrheitsprozess)

Zusammenfassung

Wo eine Lage selten oder einmalig ist, kommt keine Rückmeldung, an der sich die eigene Gewissheit prüfen ließe; die Grundrate ähnlicher Fälle hilft, wird aber dünn, und es bleiben Spannen statt Punkte, umkehrbare Entscheidungen und frühes Nachsehen. Manches Fehlende zeigt sich erst im Rückblick; der allgemeine Vorbehalt verwirft kein bestimmtes Urteil, verlangt aber, offene Stellen sichtbar zu halten. Kalibrierung ist nie fertig, weil sich die Welt und das eigene Maß verschieben. Sie heißt nicht, zu prüfen, bis keine Unsicherheit bleibt, sondern mit dem am besten geprüften Maß zu handeln und zu wissen, wie sicher es ist. Eine Gesellschaft, in der viele so urteilen, hat keinen Hauptschalter: Sie ist langsamer zu einigen und schwerer zu täuschen. Auch dieses Buch ist eine Auswahl, mit vielen Fällen aus englischsprachigen Quellen, weil sie leicht zu prüfen waren; Leser anderswo sollten die Gedanken an ihren eigenen Fällen prüfen. Kalibrieren ist im Kern der Schritt zurück vom Bild, wie am Firdos-Platz, und die Frage, wie groß der Ausschnitt war.