

# Noise – eine Rezension

Andrea Taormina

«Noise» von Daniel Kahnemann, Oliver Sibony und Cass R. Sunstein ist lesenswert. Eines vorweg: Für mich war die Lektüre anstrengend, der Inhalt ist nicht leicht verständlich für Juristen. Und doch lohnt sich die Mühe.

Ich habe mich bei der Lektüre auf die für Strafrechtler relevanten Kapitel konzentriert; diese Kapitel waren für mich auch am Einfachsten lesbar. Ausgangspunkt meiner Rezension ist die im Buch vertretene These, dass es nicht hinnehmbar ist, wenn straffällig gewordene Menschen für genau die gleiche Straftat unter ansonsten gleichen Bedingungen völlig unterschiedliche Strafmasse erhalten. Dass dem so ist, zeigen Studien, die dies empirisch nachweisen (und im Buch zitiert werden). «Strafmasse sollten nicht von der Stimmung des Richters während der Verhandlung oder von der Aussentemperatur abhängen.»

Das Buch geht der Frage nach, warum bei gleichen Fällen völlig unterschiedliche Strafmasse angeordnet werden. Die Antwort darauf ist: Die Urteile der Menschen sind fehleranfällig. Das Buch unterscheidet zwei Arten von Fehlern: den Fehler aufgrund von Verzerrung (*bias*) und den Fehler aufgrund von Streuung (*noise*). *Bias* ist der systematische Fehler («dunkelhäutige beschuldigte Personen werden härter bestraft als nicht dunkelhäutige beschuldigte Personen»). *Noise* ist der Fehler, der nicht systematisch ist («Personen werden für die gleiche Straftat unter den gleichen Umständen anders bestraft, unabhängig, ob sie dunkelhäutig sind oder nicht»). Über *bias* gibt es zahlreiche Studien. *Bias* ist der *star* der *show*, sagt das Buch etwas salopp. *Noise* ist bislang weitgehend unerforschtes Territorium, sagt das Buch. *Noise* stellt unsere Strafjustiz wesentlich grundlegender in Frage als *bias*, sage ich. Denn:

Selbst wenn wir *bias* eliminieren (was wir müssen), bleibt *noise* unser täglicher Begleiter.

Das Buch handelt also von der unerwünschten Streuung von Urteilen. Urteile in diesem Sinne sind eine Form von Messung, wobei der menschliche Intellekt das Messinstrument ist. Ein Urteil in diesem Sinne ist beispielsweise die Einschätzung eines Arztes, ob eine Tumor gutartig oder bösartig ist, eines Versicherungsmathematikers, ob die Höhe einer Versicherungsprämie das Risiko angemessen absichert, eines Vorgesetzten, ob ein Bewerber die am besten geeignete Person für eine Stelle ist, eines Lehrers, ob er einen Aufsatz mit der Note 6 oder mit der Note 5.5 bewertet, ob ein am Tatort gefundener Fingerabdruck eindeutig einem Verdächtigen zugeordnet werden kann und ja, auch, ob die am Tatort gefundene DNA-Probe eindeutig einer beschuldigten Person zugeordnet werden kann. Ein Urteil im vorliegenden Sinne ist zum Beispiel auch der Entscheid eines Richters, ob eine beschuldigte Person eine hohe oder eine tiefe Rückfallwahrscheinlichkeit hat.

Noise tritt überall dort auf, wo geurteilt wird, und zwar in einem höheren Mass als man gemeinhin glaubt. Variabilität ist bei gewissen Urteilen gewollt (zum Beispiel Buchrezensionen (!), Einschätzungen von Wertpapierhändlern etc.), bei anderen Urteilen (Diagnose des Arztes, Prognose des Richters) hingegen nicht. Wenn Urteile ohne triftige Gründe schwanken, dann sind sie fehlerhaft. Die so verstandene Fehlerhaftigkeit wird als störend empfunden. Es ist ungerecht, wenn Menschen in vergleichbarer Lage ungleich behandelt werden. Systeme, bei denen der Eindruck entsteht, sie seien inkonsistent, verlieren an Glaubwürdigkeit.

Die Ursachen von *noise* bei Urteilen sind vielfältig, die wichtigsten Gründe sind kognitive Schwächen von Urteilenden und objektive Unwissenheit. Interessanterweise veranlasst das überzogene Vertrauen in die Richtigkeit ihrer Urteile viele Menschen dazu, ihre kognitiven Schwächen, ihre objektive Unwissenheit und ihre *bias* zu unterschätzen. Seltsamerweise sind wir mit unseren Urteilen oftmals zufrieden. Das Buch liefert eine Erklärung, wieso das so ist: Sobald wir die Fakten und das Urteil zu einer stimmigen Geschichte zusammengefügt haben, geben wir uns eine Belohnung, was uns das befriedigende Gefühl verschafft, richtig zu liegen. Noise entdecken ist schwierig, weil hierfür eine statistische Betrachtungsweise («sind alle meine Strafzumessungen abhängig von meiner Tagesform?») notwendig ist, wir aber den kausalen Erklärungen den Vorzug geben («drei Jahre sind angemessen, weil das Vorleben des Beschuldigten bereits von Gewalt geprägt war»).

Warum finde ich das Buch lesenswert? Es betrachtet juristische Probleme grundlegend anders als ich es aufgrund meiner Ausbildung im Strafrecht und Strafprozessrecht gewohnt bin. Meine Herangehensweise ist – vielleicht etwas verallgemeinernd gesagt – dogmatisch. Das Buch hingegen schärft den empirischen (und nicht dogmatischen) Blick auf juristische Probleme. Ich versuche das zu erläutern.

Nehmen wir die Strafzumessung. Ich war vor Kurzem an einer Hauptverhandlung, in der sich das Gericht bei der Urteilseröffnung ausführlich zur Strafzumessung äusserte. Es erklärte dogmatisch korrekt und mit gleichsam mathematischer Genauigkeit, aus welchen Gründen die ausgesprochene Strafe angemessen sei. Die Richter schienen tatsächlich überzeugt zu sein, dass sie die richtige Strafe ausgesprochen haben (Originalzitat: «Wir sind überzeugt und haben keine Zweifel, dass die heute ausgefallte Freiheitsstrafe von 24 Monaten angemessen ist.»).

Wir wissen nun aber aus den erwähnten empirischen Studien, dass andere Richter oder eine andere Zusammensetzung des Gerichts oder sogar dieselben Richter an verschiedenen Tagen oder verschiedenen Tageszeiten zu einem völlig anderen Strafmass gelangen würden. Das heisst: Die in meinem Beispiel ausgefallte Strafe ist – jedenfalls, wenn man der Theorie der Punktstrafe folgt – offensichtlich nicht die richtige Strafe. Die Erkenntnis aus den empirischen Studien lässt uns den Streit zwischen der eben erwähnten Theorie der Punktstrafe und der

Die Richter schienen tatsächlich überzeugt zu sein, dass sie die richtige Strafe ausgesprochen haben.

Spielraumtheorie als Nebenschauplatz erscheinen. Viel wichtiger als die metaphysische Frage zu klären, ob es eine auf den Punkt richtige Strafe gibt oder ob der Richter einen Spielraum zwischen der nach oben noch angemessenen und nach unten schon angemessenen Strafe zur Verfügung hat, dünkt mich, die Rechtswirklichkeit anzuschauen und Daten darüber zu erfassen, welche Ungleichheiten in der Strafzumessung tatsächlich bestehen. Insofern ist **Claus Roxin** zuzustimmen, der schon früh darauf hinwies, dass allein die Spielraumtheorie der Realität des Strafzumessungsaktes gerecht wird (Strafzumessung im Lichte der Strafzwecke, in: Hans Walder / Stefan Trechsel, Hrsg., Lebendiges Strafrecht, Festgabe zum 65. Geburtstag von Hans Schultz, Schweizerische Zeitschrift für Strafrecht, 94. Jahrgang, Bern: Stämpfli 1977, 466). Allerdings – und dies ist das spannende an *noise* – sollte es nicht bei

dieser allgemeinen Erkenntnis bleiben. Vielmehr sollte die Forschung ihr Augenmerk darauf richten, die tatsächlichen Ungleichheiten bei der Zumessung systematisch zu erfassen und die richtigen Schlüsse daraus ziehen.

Schauen wir ein weiteres Beispiel an, weswegen ich die empirische Betrachtungsweise von juristischen Problemen für beachtenswert halte. Entscheide von Zwangsmassnahmengerechten und Rechtsmittelinstanzen enthalten scheinbar exakte Begründungen für das Vorliegen der Haftgründe. Wir lernen diese Begründungen im Studium und für die Anwaltsprüfung und wenden sie in der Praxis an. Dabei befassen wir

uns – mit ganz wenigen Ausnahmen – schlichtweg nicht mit der Realität («99% der Haftanträge werden gutgeheissen» «Entsiegelungsbegehren werden fast immer gutgeheissen»). Wir blenden den empirischen Blick fast vollständig aus.

Wie gesagt, ist die Lektüre nicht ganz einfach. Aber sie lohnt sich.

