



Einsichten eines Wissenschaftsnarren (15)

Mit schlichten Wetten die Wissenschaft retten?

Könnte man mit Wetten statt Peer Review entscheiden, welche Projekte gefördert werden sollen? Die Bilanz würde womöglich besser werden.

„Durch Wetten die Wissenschaft retten?“ Diese Frage stellte der US-Ökonom Robin Hanson mit einem Artikel des Jahres 1995. Er schlug darin vor, den klassischen Review-Prozess durch eine Markt-basierte Alternative zu ersetzen: Statt *Peer Review* könnten Wetten darü-



Foto: BIH/Thomas Rafalzyk

Ulrich Dirnagl

leitet die Experimentelle Neurologie an der Berliner Charité und ist Gründungsdirektor des QUEST Center for Transforming Biomedical Research am Berlin Institute of Health. Für seine Kolumne schlüpft er in die Rolle eines „Wissenschaftsnarren“ – um mit Lust und Laune dem Forschungsbetrieb so manche Nase zu drehen.

ber entscheiden, welche Projekte gefördert werden oder welche wissenschaftlichen Fragestellungen priorisiert werden.

In diesen sogenannten *Prediction Markets* (Prognosemärkten) handeln Individuen mit „Wetten“, die auf ein bestimmtes Resultat oder Ergebnis setzen. Je mehr Leute auf einem solchen Marktplatz handeln, umso genauer wird die Vorhersage des Ergebnisses, welche letztlich auf der aggregierten Information der Teilnehmer basiert. Der Prognosemarkt bedient sich also der Schwarmintelligenz.

Bei Sportwetten und Wahlprognosen kennt man das schon. Aber taugt es für die Wissenschaft? Klingt total verrückt, ist es aber nicht! Tatsächlich finden solche Konzepte gerade Eingang in einige Sparten der Wissenschaft. Wie funktioniert's also, und was ist dran?

Ausgangspunkt von Hansons Überlegungen ist die gängige und ernstzunehmende Kritik an der Art und Weise, wie wir Forschungsarbeiten beurteilen und Förderentscheidungen treffen – nämlich, dass *Peer Review* zu Risikoaverser Mainstream-Forschung führe und daher innovationsfeindlich sei, dass er zum *Story Telling* anrege, und dass er Außenseiter diskriminiere wie auch bereits Erfolgreiche favorisiere („Matthäus-Effekt“).

Für Hanson stellt sich daher die Frage, wie Konsens über die Wertigkeit oder Innovation von Forschung gefunden werden kann, ohne sich auf ein paar Meinungen von „Peers“ mit all deren Bias und persönlichen Interessen verlassen zu müssen. Seine Antwort: Analog zu Sportwetten könnte man doch eine viel breitere Basis von Experten auf den Erfolg eines geplanten Projektes oder die Richtigkeit eines Befundes wetten lassen. Je nach Fragestellung könnte man auch Nicht-Experten mitspielen lassen – und damit für Partizipation und gesellschaftlichen Konsens sorgen. Letztlich schlägt er damit eine Art *Crowdsourcing* von Konsensbildung vor – und dies mit einem spielerischen Element!

Dass so etwas in der Wissenschaft tatsächlich funktionieren kann, zeigen einige neu-

ere Arbeiten, in denen die erwähnten *Prediction Markets* (Prognosemärkte) eingesetzt wurden. Dabei ging es zum Beispiel um die Vorhersage, ob eine Studie replizierbar sein werde oder nicht.

Konkret geht das so: Die teilnehmenden Wissenschaftler erhalten hundert Spielmarken (manchmal auch echtes Geld), mit denen sie dann auf den Replikationserfolg setzen. Sie würden verstärkt auf Studien setzen (das heißt: Lose kaufen), von denen sie glauben, dass sie erfolgreich wiederholt werden können. Umgekehrt würden sie diejenigen Stu-

»„Exzellenz“ könnte man auch gut durch Wetten selektieren.«

dien meiden, denen sie das nicht zutrauen. Durch Kauf und Verkauf von Losen stellt sich auf diesem Markt ein Preis für die Lose ein (in Spielmarken oder echter Währung). Dieser Preis reflektiert die Wahrscheinlichkeit, mit der die Marktteilnehmer an die Replizierbarkeit der Studie glauben.

Dieses Verfahren wurde in einer eben veröffentlichten Publikation eingesetzt. Darin wurden 21 Studien aus der Psychologie wiederholt, die zwischen 2010 und 2015 in *Nature* oder *Science* erschienen waren. Zuvor konnte ein Gruppe von Studenten und Wissenschaftlern, die nicht an den Replikationsstudien beteiligt und nicht einmal notwendigerweise als Experten in den jeweiligen Feldern ausgewiesen waren, durch Setzen ihrer Spielmarken Lose kaufen und damit auf die Replizierbarkeit jeder der 21 Studien wetten. Das Ergebnis: Die „Spieler“ sagten nahezu perfekt voraus, welche Studien nachfolgend erfolgreich repliziert wurden und welche nicht wiederholbar waren.

Mit einer Vergleichsgruppe wurde eine klassische Befragung durchgeführt: „Glauben Sie, dass Studie X oder Y wiederholbar

sein wird?“ Die Ergebnisse dieser Umfrage waren jedoch nicht besser als das, was man auch durch pures Raten hätte erreichen können.

In einer anderen Studie wurde ein *Prediction Market* benutzt, um die Ergebnisse der sogenannten REF (*Research Excellence Framework*) für das Fach Chemie vorherzusagen. Mit dem hochkomplexen und teuren REF-Verfahren evaluiert der englische Staat seine Universitäten, um auf Basis der Resultate (Exzellenz!) seine Forschungsmittel zu verteilen. In der Studie sagte ein einfacher *Prediction Market* mit nur 13 Chemikern, vom Studenten bis zum Professor, ziemlich akkurat das Ergebnis der REF 2014 für alle 33 Chemie-Departments Englands voraus.

Hätte man die frisch in der deutschen Exzellenzstrategie von Bund und Ländern gekürten Exzellenzcluster womöglich ebenso durch ein einfaches Wettspiel selektieren können? Ohne dass Tausende von Wissenschaftlern Anträge schreiben mussten, statt zu forschen, und ohne dass ein Heer von internationalen Gutachtern durch die Republik reisen musste?

Woran liegt es aber, dass Wetten bessere Vorhersagen liefern können als Umfragen oder *Peer Review*? Dass dies so ist, ist keineswegs neu. Wir wissen das, wie gesagt, beispielsweise aus Sportwetten oder Wetten auf Wahlausgänge. Diese erreichen erstaunliche Vorhersagekraft und liegen fast immer deutlich besser als Umfragen.

Ein Faktor hierbei ist vermutlich, dass der Anreiz größer ist, sich mit dem Umfrage-Gegenstand auseinanderzusetzen, wenn man dabei etwas gewinnen kann. Selbst wenn es nur „Spielmarken“ sind und kein reales Geld, wie bei manchen der *Prediction-Market*-Studien in der Wissenschaft. (In Deutschland könnten diese wohl ohnehin nicht mit echtem Geld

durchgeführt werden, denn das wäre unerlaubtes Glücksspiel!)

Ein weiterer Faktor könnte sein, dass man in einem *Prediction Market* seine Wetten über längere Zeit durch Kauf oder Verkauf von Losen adjustieren kann. Man kann also seine Kan-

»Wichtiger aber ist die Schwarmintelligenz, die sich einstellt.«

didaten wechseln – auch durch den Blick auf den Kurswert der Wette, der sich ja durch Käufe und Verkäufe anderer verändern kann.

Womöglich noch wichtiger ist aber die Schwarmintelligenz, die sich einstellt, wenn viele mit unterschiedlichem Wissen oder Perspektiven sich einer Fragestellung annehmen.

Ist das alles bloße Gedankenspielerei, oder gäbe es für *Prediction Markets* in unserem Wissenschaftssystem tatsächlich ernsthafte Anwendungen? Eine offensichtliche Einschränkung ist natürlich, dass sich auf diese Weise vor allem dichotomisierbare Entweder-Oder-Fragestellungen bearbeiten lassen. Ist eine Studie replizierbar oder nicht? Ist ein bestimmtes Resultat richtig oder falsch? Sollte eine bestimmte Fragestellung untersucht werden oder nicht? Ist der Antrag förderwürdig oder nicht? Die Ergebnisse des Marktes sind dann lediglich Wahrscheinlichkeiten, ein inhaltliches Feedback liefert ein *Prediction Market* nicht.

Dennoch zeigt das oben erwähnte Beispiel aus der Psychologie, dass ein solcher Prognosemarkt eine einfache und zudem offenbar sehr präzise Konsensbildung in einem Forschungsfeld ermöglichen kann. Forschungsförderer könnten dies durchaus bei

ihren Entscheidungen nutzen, welche Forschungsprogramme sie priorisieren und etablieren wollen. Die Entscheidung wäre dann „besser informiert“ – mithilfe der Community, vielleicht aber auch mithilfe anderer Stakeholder (etwa Patienten).

In der translationalen Medizin stellt sich etwa häufig die Frage, ob mit einer in Modellsystemen wirksamen Substanz eine teure und potenziell Patienten-gefährdende klinische Entwicklung gestartet werden soll. Ein Prognosemarkt wäre ein simples Verfahren, die von der Community als vielversprechend angesehenen Substanzen auszuwählen; sie ließen sich sogar nach Erfolgswahrscheinlichkeiten sortieren. Natürlich garantiert das keinen Erfolg, aber mangels objektiver Kriterien wird ja auch heute bereits auf die Expertenmeinung gesetzt. Allerdings nicht auf diejenige des „Schwarms“, sondern auf diejenige weniger ausgewählter Individuen, die überdies häufig eigene Interessen verfolgen.

Nach meinem Vorstoß zur Einführung einer *Peer-To-Peer*-Forschungsgrundförderung (*LJ 6/17*: 22-23) und zur Fortbildung für Gruppenleiter und Professoren (*LJ 6/18*: 28-29) also wieder mal eine völlig närrische Idee: Wetten auf die Wissenschaft! Auch wenn es womöglich nie dazu kommen wird, liefert uns die Beschäftigung mit radikal von der gängigen Praxis abweichenden Lösungen dennoch einen informativen Blick in den Spiegel auf eben diesen Status quo – mit all seinen Stärken und Schwächen. Und womöglich erkennen wir dabei umso mehr: Es läuft nicht optimal, wir sollten etwas anderes probieren!

(Die zitierten Artikel sowie weiterführende Literatur und Links finden sich wie immer unter <http://dirnagl.com/lj/>)

Wir entwickeln und produzieren für die
MIKROBIOLOGIE

Chromogene Substrate

Fluoreszenz-Substrate

Pro-Luziferine (caged luciferins)

STAADIUM™ Enzym-aktivierte Inhibitoren

NEU: AquaSpark™
Chemilumineszenz-Substrate für schnelle und direkte Bakterien-Identifikation *in vitro* und *in vivo*

BIO SYNTH[®]
CHEMISTRY & BIOLOGY