

NEUROWISSENSCHAFT

Großhirn-Voodoo

Egal, ob Mutterliebe, Trennungsschmerz, Gottesglaube: Über alles gibt der funktionelle Kernspin Auskunft. Nun regen sich Zweifel: Wie groß ist die Aussagekraft der bunten Hirnscans wirklich?

Der wohl ungewöhnlichste Proband, dessen Gehirntätigkeit Craig Bennett jemals untersucht hat, war ein toter Lachs. Seinem leblosen Versuchsobjekt stellte der Neuropsychologe von der University of California in Santa Barbara eine knifflige Aufgabe: Das fast zwei Kilo schwere Tier sollte Bilder von fröhlichen, ängstlichen oder wütenden Menschen betrachten. Währenddessen maß Bennett seine Hirnaktivität.

Das Ergebnis des absurden Experiments sorgte für einen handfesten Skandal. Denn im Gehirn des toten Fisches regte sich etwas.

Müssen wir tote Lachse also mit völlig neuen Augen betrachten? Können Fische, noch während wir sie in die Bratröhre schieben, unsere Gefühle erkennen?

Natürlich nicht, versichert Bennett im neugegründeten „Journal of Serendipitous and Unexpected Results“ (Zeitschrift für zufällige und unerwartete Ergebnisse), wo er sein Lachs-Experiment veröffentlicht hat. Die mittels funktioneller Kernspintomografie (fMRT) im Gehirn des Tiers erkennbaren Signale seien reine Zufallsprodukte, statistische Ausreißer ohne Bedeutung. Würden die passenden Korrekturrechnungen vorgenommen, verschwänden sie von selbst. Doch solche Rechnungen, schreibt Bennett, hätten nach seiner Auswertung selbst in den besten Fachjournals zwischen 25 und 40 Prozent der Wissenschaftler nicht durchgeführt.

Das Experiment sagt also nichts über die emotionale Intelligenz von toten Lachsen aus – umso mehr aber über die Tücken der fMRT.

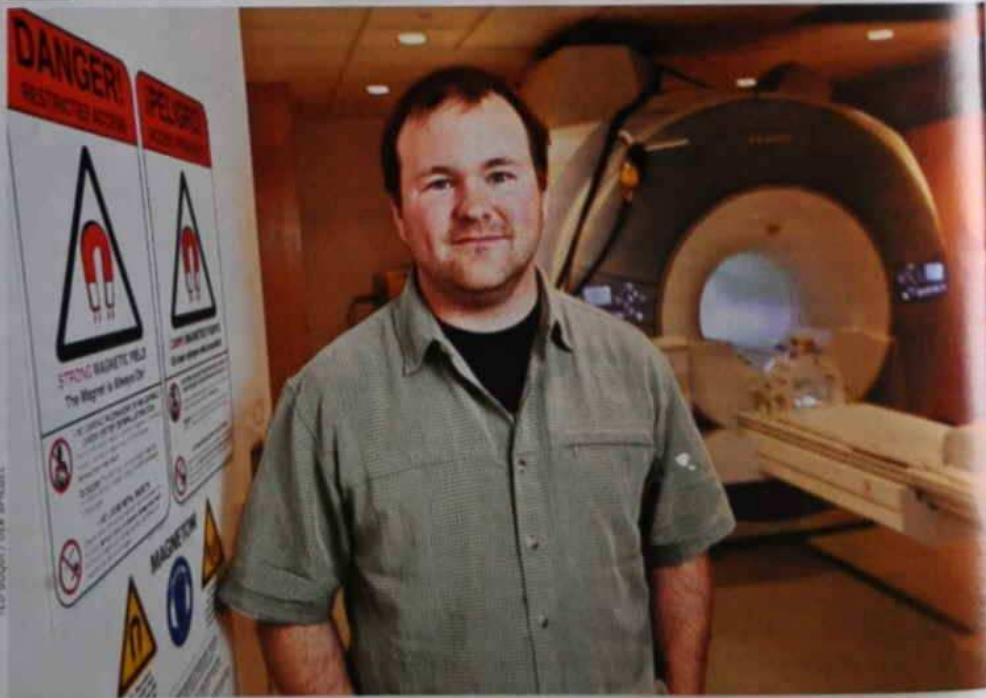
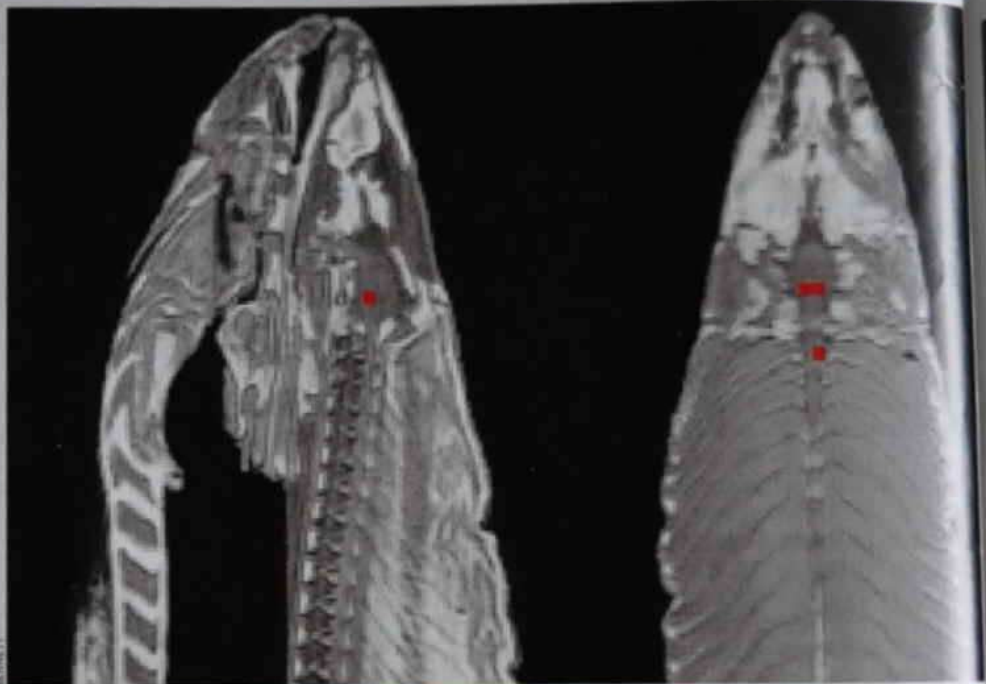
Bennetts Versuch nährt Zweifel an einem Verfahren, das vor 15 Jahren die Neurowissenschaften zu erobern begann. Keine Frage schien zu groß, um sie nicht mit Hilfe der funktionellen Kernspintomografie beantworten zu können: Gibt es einen freien Willen? Warum können Frauen nicht einparken? Wo entsteht die Mutterliebe? Wo sitzt der Trennungsschmerz? Und wo der liebe Gott? Jedes Mal lieferten die fMRT-typischen bunten Aktivitätsmuster im Gehirn der Versuchspersonen wunderbar eindeutige Antworten. Als Lügendetektor, zur Diagnose

psychischer Erkrankungen, in der Medikamentenforschung, sogar zur Wirtschafts- und Marktforschung wird die Methode inzwischen eingesetzt.

Jetzt jedoch melden sich Kritiker zu Wort. Sie warnen eindringlich davor, die

Ergebnisse der Hirnscans überzuinterpretieren. Die funktionelle Kernspintomografie behandle „die Staubflocke der Psyche“, als wäre sie „hart wie Eisen“, schreibt der Neurowissenschaftler Nikolaus Kriegeskorte von der University of Cambridge. Der Methode attestiert er „gefährliche Verführungskraft“, weil sie „das Prestige der harten Wissenschaft mit der breiten Anziehungskraft intuitiv erfassbarer Bilder“ verbinde.

Tatsächlich zeigen Studien, dass jeder wissenschaftliche Artikel sofort an Glaubwürdigkeit gewinnt, wenn ihm ein Hirnscan beigelegt wird – selbst wenn dieses Bild gar nichts mit dem eigentlichen Inhalt des Artikels zu tun hat. Kein Wunder, dass gerade über die fMRT-Studien auch



Toter Lachs in fMRT, Forscher Bennett
Statistische Ausreißer ohne Bedeutung

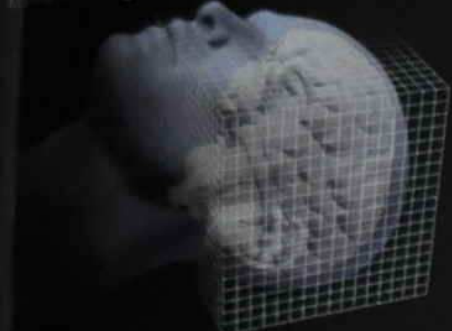
viele Medien gern berichten – je griffiger das Ergebnis, desto lieber.

Bunte Gehirnbilder führen Forscher in Versuchung, Behauptungen aufzustellen, die die soliden wissenschaftlichen Daten hinter sich lassen“, sagt Russell Poldrack, Neurowissenschaftler von der University of Texas. Allzu leicht würden die Grenzen und Widersprüche des Verfahrens außer Acht gelassen. Dies, so Poldrack, führe dazu, „dass die Möglichkeiten der Methode, alles zu erklären, von Liebe und Schönheit bis hin zur Entscheidungsfindung in finanziellen Fragen, allgemein überschätzt werden.“

Nach 15 Jahren des Booms ist jetzt die Euphorie bei vielen Forschern verfliegen. Als ich meine Lachs-Studie auf einem

Bunte Welt der Gedanken

Untersuchung mit funktioneller Kernspintomografie



Die Hirnaktivität kann nur indirekt über die Veränderung der Durchblutung gemessen werden. Dabei wird das Hirn in rund 130 000 würfelförmige Kuben, sogenannte Voxel, unterteilt. Die 130-Millimeter Kantenlänge sind in einem Voxel rund 100 Millionen Neuronen enthalten – eine sehr grobe Auflösung.

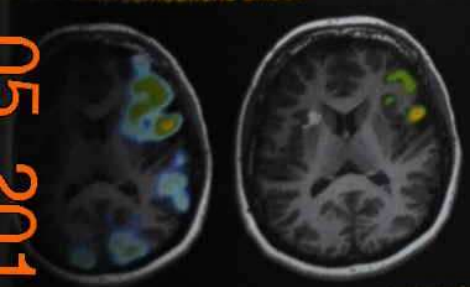
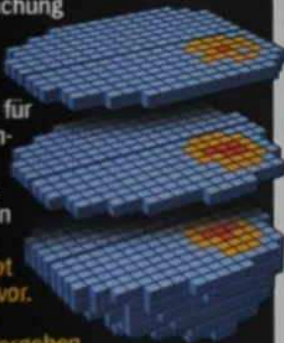
Während der Untersuchung

wechselt die Versuchsperson zwischen Ruhephasen und Aktivität.

Der Computer errechnet für jedes Voxel, ob die Durchblutung sich signifikant verändert hat oder nicht und bildet das Resultat in einem Farbmuster ab.

Die Signifikanzgrenze gibt der Untersucher vor.

Wie er den Schwellenwert festlegt, ergeben sich unterschiedliche Bilder.



Die Kernspinaufnahme errechnete Farbmuster. Links ist so eingestellt, dass die Befunde mit 5% (links) Signifikanzschwelle (rechts) auf bloßem Zufall beruhen.

Kongress vorstellte“, so Bennett, „war die Reaktion zu meiner Überraschung fast einhellig positiv. Manche haben sogar laut ‚Danke!‘ gerufen.“

Erstaunlich lange hat es gedauert, bis sich die Skepsis regte. Anfangs trat die funktionelle Kernspintomografie ihren Siegeszug als die Untersuchungsmethode für das arbeitende Gehirn vor allem deshalb an, weil bei ihr im Vergleich zur Positronen-Emissions-Tomografie (PET) keine Strahlenbelastung entsteht. Zudem sind zeitliche und räumliche Auflösung der fMRT deutlich besser.

Doch ein fMRT-Bild ist kein Foto des arbeitenden Gehirns (siehe Grafik). Anders als oft behauptet wird bei dieser Methode nicht einmal die Gehirnaktivität selbst gemessen, sondern lediglich die vermehrte Durchblutung, die Folge dieser Aktivität ist.

Dafür wird das Gehirn in der Regel in rund 130 000 Würfel von jeweils drei Millimeter Kantenlänge eingeteilt, sogenannte Voxel. Bereits diese Aufteilung setzt der Methode klare Grenzen: Ein Voxel enthält zwischen 500 000 und drei Millionen Neuronen mit mehr als 100 Kilometern an Nervenbahnen und bis zu 27 Milliarden Kontaktstellen (Synapsen). Der Versuch, die Arbeitsweise des Gehirns auf der Basis der gängigen Voxel zu entschlüsseln, gleicht deshalb ein wenig dem Bemühen eines Kurzsichtigen, aus der Ferne das Muster eines Pullovers zu erkennen.

Für jedes Voxel wird das sogenannte BOLD*-Signal zweimal gemessen: zunächst im Ruhezustand, also während die Versuchsperson möglichst nichts tut, dann in Aktivität, also während die Versuchsperson eine bestimmte Aufgabe lösen muss. Das Signal ist umso stärker, je mehr Sauerstoff das Blut transportiert. Dies ermöglicht einen Rückschluss auf die Hirntätigkeit, denn nach vermehrter Aktivität steigt die Durchblutung und damit auch der Sauerstoffgehalt im Blut des aktiven Hirnareals an.

Jetzt beginnt der Computer des Kernspingeräts zu rechnen: Für jedes Voxel einzeln ermittelt er, ob sich die BOLD-Signalstärken in der Ruhephase und der Phase des Aufgabenlösens signifikant voneinander unterscheiden. Wenn ja, wird dieses Voxel eingefärbt. Anschließend wird das Muster der eingefärbten Voxel auf ein Standardgehirn projiziert: Es entsteht ein buntes Hirnscan-Bild.

Allerdings lauern in dieser rechnerischen Mammutaufgabe etliche mögliche Fehlerquellen. Vor allem muss der Untersucher selbst festlegen, was er unter „signifikant“ versteht. Je großzügiger er die Einstellung wählt, desto mehr Aktivität wird das Gerät zeigen. „Bunte Flecken kann das fMRT, wenn der Untersucher es will, immer zeigen“, sagt deshalb

* Abkürzung für „blood oxygen level dependent“.

Timo Krings, Neuroradiologe von der University of Toronto. „Die Frage ist nur, ab wann die Aussagekraft dieser Flecken gleich null ist.“

Wie beim Lautstärkereglern eines Radios kann ein Forscher bei der Auswertung der Versuchsdaten die Signifikanzgrenzen nach oben oder unten regeln und so bunte Flecken entstehen und wieder verschwinden lassen, bis er sich schließlich für die Grenze entscheidet, die er für die beste hält: ein hoch subjektiver Vorgang und, wie Bennett es nennt, ein „prinzipienloser Prozess“.

Die fertigen Hirnscans sind von einem naturalistischen Abbild der Wirklichkeit so weit entfernt wie das berühmte Warhol-Porträt von der wahren Marilyn Monroe. Wie viel die hochartifizialen, subjektiv gestalteten Rechenkonstrukte am Ende noch mit der untersuchten Fragestellung zu tun haben, wird von vielen Forschern offenbar deutlich überschätzt.

„Voodoo-Korrelationen in der sozialen Neurowissenschaft“: Unter diesem provokanten Titel reichte der MIT-Wissenschaftler Edward Vul vor wenigen Jahren einen Artikel ein, in dem er fMRT-Arbeiten seiner Kollegen genauer analysiert hatte. Selbst hochrangige Wissenschaftler, so Vuls Befund, hatten nicht nur den Zusammenhang zwischen Wirklichkeit und Hirnscan viel zu hoch angegeben, sondern ihn auch völlig falsch berechnet.

Zu den zahlreichen Fallstricken der Statistik kommt noch die Unberechenbarkeit äußerer Einflüsse hinzu. Der Neuroradiologe Krings etwa konnte zeigen, wie ein scheinbar faszinierendes Ergebnis als Trugschluss entlarvt werden kann, wenn man es genauer untersucht.

Mit Hilfe der funktionellen Kernspintomografie hatte seine Arbeitsgruppe das räumliche Vorstellungsvermögen von Männern und Frauen verglichen. Frauen, so zeigte sich schnell, brauchen für die gleiche Aufgabe einen viel größeren Teil ihres Gehirns als Männer – vermutlich, so die naheliegende Deutung, um ihre Schwächen auf diesem Gebiet auszugleichen: Der Beweis, warum Frauen nicht einparken können, schien (wieder einmal) erbracht.

Dann jedoch wiederholte Krings seinen Versuch mit Frauen in verschiedenen Phasen ihres Menstruationszyklus. Nun stellte er fest: Die größten Unterschiede zu Männern bestanden zur Zeit des Eisprungs, wenn das Hormon Östrogen Spitzenwerte erreicht. In den Östrogen-Niedrigphasen hingegen war kein Unterschied zwischen den Geschlechtern mehr feststellbar.

Lässt Östrogen also das räumliche Vorstellungsvermögen verkümmern? Mitnichten, meint Krings. Für den beobachteten Effekt gebe es eine viel einfachere Erklärung: „Östrogen erhöht die Gehirndurchblutung und beeinflusst deshalb das

BOLD-Signal. Man muss mit voreiligen Schlüssen wirklich sehr vorsichtig sein.“

Auch Kaffee trinken am Morgen vor der Untersuchung, Treppen steigen, Fernsehen gucken, ja sogar schon die Aufregung angesichts der lautstark hämmern- den Riesenröhre können die Daten maßgeblich verzerren. Die Zuverlässigkeit, mit der fMRT-Ergebnisse reproduziert werden können, ist entsprechend niedrig. Bennett gibt sie in einer umfangreichen Untersuchung mit 33 bis 66 Prozent an.

Welchen Wert, fragt man sich, kann da noch ein fMRT-Lügendetektor haben? Welchen Wert hat eine psychiatrische Diagnose per funktionellen Kernspin oder eine Medikamentenstudie, bei der

Je komplizierter die Fragestellung, desto fragwürdiger ist oft die Interpretation des Ergebnisses.

die Wirkung eines Arzneimittels mit Hilfe von Hirnscan-Veränderungen nachgewiesen werden soll?

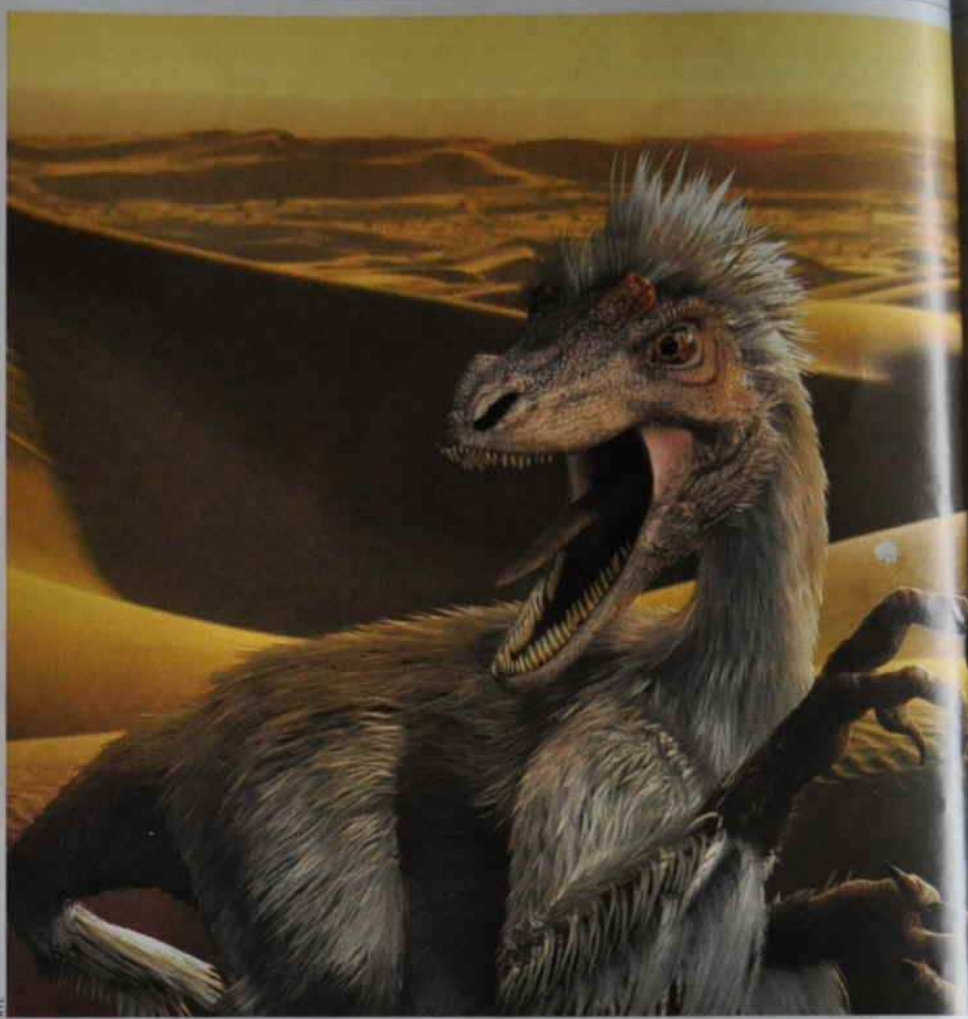
Je komplizierter die Fragestellung eines Experiments, desto fragwürdiger ist oft die Interpretation des Ergebnisses. Mitunter bekommt man den Eindruck, um die bunten Flecken im Hirnscan herum werde schlicht mit viel Phantasie eine passende Geschichte gestrickt. „Bewiesen: Liebe wirkt wie Kokain“, dichtete beispielsweise die Berliner „B.Z.“. Die „Welt“ schrieb: „Forscher beweisen: Nach 20 Jahren haben manche Paare Gefühle wie am ersten Tag.“ „Sex findet bei Frauen zwischen den Ohren statt“ und „Natureszenen sind Balsam fürs Gehirn“ lauteten weitere Schlagzeilen.

„Wenn eine Geschichte zu einfach klingt, um wahr zu sein“, warnt fMRT-Experte Poldrack, „dann ist die Geschichte oft auch nicht wahr.“

Und selbst wenn ein Befund einer kritischen Hinterfragung standhält, bleibt noch die große Frage, was eigentlich die Ursache und was die Wirkung ist: Was beweist es schon, wenn das Gehirn jüngerer Probanden bei klassischer Musik weniger Aktivität zeigt als das älterer Hörer? Langweilt sie Mozart, weil ihr Hirn dafür weniger empfänglich ist, oder bleibt ihr Hirn stumm, weil sie Mozart langweilig finden? „Die Deutung von fMRT-Ergebnissen lässt viel Spielraum“, sagt Krings, „vor allem bei komplexen Aufgaben oder Emotionen.“

Krings und die anderen Kritiker wollen die Methode keineswegs ächten, sondern dringen nur darauf, sie zu verbessern. Auch Bennett schiebt weiterhin Probanden in den Tomografen – wengleich bis auf weiteres nur noch menschliche. Den Lachs haben er und seine Kollegen nach dem Experiment aufgegessen, mit einer schmackhaften Zitronen-Knoblauch-Sauce.

VERONIKA HACKENBROCH



Dinosaurier Velociraptor*: „Schockierend lebendiger Beweis für die Existenz der Evolution.“

PALAONTOLOGIE

Huhn mit Zähnen

Der US-Paläontologe Jack Horner ist eine Legende der Dinosaurierforschung. Als Krönung seiner Karriere möchte er nun einen lebendigen Dino erschaffen – aus einem Huhn.

Jack Horner hat fast alles erreicht, wovon ein Paläontologe nur träumen kann: Er wird von einer Heerschar Sechsjähriger wie ein Gott verehrt. Er hat Tonnen versteinertes Dinosaurierknochen im Keller liegen. Und seit über 40 Jahren buddelt er fast jeden Sommer weiter nach den Überresten der Urzeit-echsen.

Nur einen Wunsch hat sich der Forscher aus Bozeman im US-Bundesstaat Montana bis heute nicht erfüllt. „Ich wollte schon immer einen Dinosaurier als Haustier haben“, sagt er und freut sich diebisch am verblüfften Blick seiner Zuhörer. „Ich habe Enten im Garten; sie lau-

fen hinter mir her und sind ziemlich süß“, sagt er. „Aber cool ist das nicht; cool wäre so etwas wie ein Velociraptor.“

„Die Wissenschaft ist bald an einem Punkt angekommen, an dem wir Dinosaurier erschaffen können“, behauptet Horner und streicht sein langes grauweißes Haar nach hinten. „Chickenosaurus“ hat er das Geschöpf seiner Träume getauft. Ein Huhn soll als Blaupause dienen. Denn Horner will die Uhr der Evolution zurückdrehen und aus Vögeln wieder Dinosaurier machen.

„Es ist ein kühner Plan“, räumt er ein, „aber es ist nur eine Frage der Zeit, bis wir die richtigen Genschalter entdecken.“

Horner, 64, ist eine lebende Legende der Dinosaurierforschung. Steven Spielberg heuerte ihn Anfang der neunziger

* Szene aus dem Dokumentarfilm „Die letzten Jahre der Dinosaurier“ von Pierre De Lespinois, 2003.