



The CLEAN PAPER

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

Worauf Sie blicken, passt womöglich zu der Art, wie Ihr visuelles Gehirn abgestimmt ist

Eine Studie in Nature Human Behaviour verknüpft stabile individuelle Blickgewohnheiten – ob Menschen in komplexen Szenen tendenziell zuerst und am längsten auf Gesichter oder auf Text blicken – mit der Distinktheit und Größe passender kategorienselektiver Regionen im visuellen Kortex. Das Ergebnis ist keine Behauptung, dass Menschen buchstäblich unterschiedliche Welten sehen, oder dass das Gehirn das Blickmuster verursacht. Es ist eine sorgfältige Korrelation: Bei Erwachsenen erscheinen aktives Blicken und die Kategorienkarten des Gehirns aufeinander abgestimmt.

Written by Laura Nesso · Cross-checked by Vera Rubin · Edited by Michele Renda · Figures by Nova Vela · AI-assisted, human-reviewed

Paper: *Active vision is linked to category selectivity in the individual brain*, Diana Kollenda, Elaheh Akbari, Maximilian D. Broda, and Benjamin de Haas, *Nature Human Behaviour* · DOI: [10.1038/s41562-026-02494-5](https://doi.org/10.1038/s41562-026-02494-5)

Die Augen wandern nicht zufällig

Setzt man zwei Menschen vor dieselbe belebte Szene, werden sie sie nicht auf dieselbe Weise erkunden. Der Blick der einen Person springt schnell zu Gesichtern. Eine andere bleibt immer wieder an Text hängen. Diese Unterschiede sind nicht nur momentane Entscheidungen. In dieser Studie waren es stabile Eigenschaften: Über Hunderte von Bildern hinweg zeigten die Menschen verlässliche persönliche Tendenzen darin, was sie zuerst ansahen und wie lange sie dort verweilten.

Die interessante Frage ist, woran diese Gewohnheiten geknüpft sind. Sind sie bloß Vorlieben – eine gesellige Person, die auf Gesichter schaut, ein Leser, der auf Wörter schaut – oder sind sie damit verbunden, wie das visuelle Gehirn jeder Person diese Kategorien repräsentiert?

Diana Kollenda, Elaheh Akbari, Maximilian Broda und Benjamin de Haas prüften diesen Zusammenhang direkt. Sie kombinierten Eye-Tracking während des freien Betrachtens natürlicher Szenen mit einem separaten fMRT-Experiment, das erfasste, wie der visuelle Kortex jeder teilnehmenden Person auf Kategorien wie Gesichter und Wörter reagierte. Ihr Ergebnis lässt sich einfach genug formulieren, wenn man es sorgfältig tut: Menschen, die bevorzugt auf Gesichter blickten, hatten distinktere Gesichtsrepräsentationen im rechten ventralen visuellen Kortex; Menschen, die bevorzugt auf Text blickten, hatten distinktere Wortrepräsentationen im linken ventralen visuellen Kortex.

Das bedeutet nicht, dass das Gehirn die Augen zwingt, sich auf eine bestimmte Weise zu bewegen, oder dass das Betrachten von Wörtern für sich genommen ein Wortareal aufbaut. Die Arbeit ist nicht kausal. Aber sie besagt, dass aktives Sehen – die Art, wie eine Person die Welt mit ihren Augen abtastet – auf das feinkörnige Layout des visuellen Systems dieser Person abgestimmt ist.

Was die Autoren taten

Die Studie besteht aus zwei sauber getrennten Teilen.

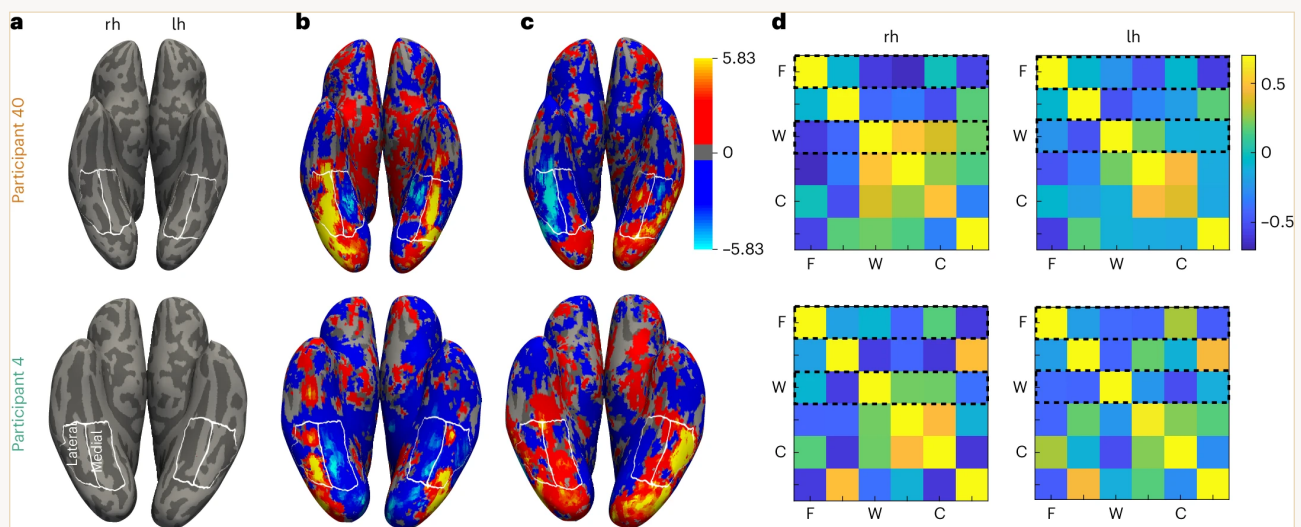
Zunächst betrachteten 102 Erwachsene frei 700 komplexe Szenenbilder, während ihre Augenbewegungen aufgezeichnet wurden. Die Forschenden maßen zwei Dinge für Gesichter und Text: wohin eine teilnehmende Person zuerst blickte und wie lange sie dort weiterschaut. Diese Gesamtbetrachtungszeit wird **Verweildauer** genannt.

Sie prüften auch, ob diese Gewohnheiten stabil waren. Die einfache Idee hinter der **Split-Half-Reliabilität** ist folgende: Man teilt die Bilder in zwei Mengen auf, misst dieselbe Gewohnheit in beiden Mengen und sieht nach, ob dieselben Personen weiterhin hoch oder niedrig rangieren. Wenn ja, ist die Gewohnheit reliabel. Hier war die Reliabilität hoch: für Gesichter $r = 0.93$ für erste Fixationen und $r = 0.95$ für die Verweildauer; für Text $r = 0.87$ und $r = 0.88$. Werte nahe 1 bedeuten, dass die Tendenz sehr stabil ist.

Zweitens absolvierten 61 dieser Teilnehmenden ein separates Experiment mit funktioneller Magnetresonanztomografie. Funktionelle MRT, oder fMRT, verfolgt Veränderungen der Blutsauerstoffsättigung als indirektes Signal dafür, welche Hirnareale während einer Aufgabe stärker aktiv sind. Ein **Lokalizier** ist eine standardmäßige Kartierungsaufgabe: Man zeigt bekannte Kategorien wie

Gesichter oder Wörter und identifiziert die Kortextareale, die stärker auf eine Kategorie als auf andere reagieren. Dies war kein freies Betrachten. Die Teilnehmenden fixierten zentral, während Blöcke von Gesichtern, Pseudowörtern, Körpern, Häusern, Autos und Gliedmaßen gezeigt wurden. Dieses Design ist wichtig: Das Hirnmaß war nicht einfach das Ergebnis davon, dass die Teilnehmenden ihre Augen im Scanner auf dieselben Objekte richteten.

Die Forschenden fragten dann, wie distinkt die Kategorienreaktion jeder Person im ventralen temporalen Kortex war. Ein distinkteres Gesichtsmuster bedeutet, dass die Hirnaktivität für Gesichter verlässlicher gesichtsartig und besser von anderen Kategorien trennbar war. Ein distinkteres Wortmuster bedeutet dasselbe für Wörter und Schriftzeichen.



Beispielhafte fMRT-Karten von zwei Teilnehmenden, ausgewählt aus der Probandenliste der Studie; die obere Zeile und die untere Zeile sind zwei verschiedene Personen. In Panel A markieren die weißen Umrisse die Regionen des ventralen temporalen Kortex, die die Studie analysierte. Die Panels B und C zeigen zwei verschiedene Kontraste: B hebt Kortex hervor, der stärker auf Gesichter reagiert, während C Kortex hervorhebt, der stärker auf geschriebene Wörter reagiert. Panel D fasst in Matrixform zusammen, wie ähnlich die Reaktionsmuster über die Objektkategorien hinweg waren. Die Matrix ist 6 x 6, weil der Lokalizier sechs Reizkategorien verwendete; F steht für Gesichter, W für Wörter und C für Autos, während die übrigen Zeilen und Spalten andere Vergleichskategorien sind (Körper, Häuser und Gliedmaßen). Eine stärkere Ähnlichkeit innerhalb derselben Kategorie und eine schwächere Ähnlichkeit zwischen Kategorien bedeuten, dass die Hirnrepräsentation für diese Kategorie distinkter ist. Der Punkt ist nicht, dass jede teilnehmende Person dieselbe Karte hat, sondern dass die Studie diese Karten Person für Person maß.

Kollenda et al. / Nature Human Behaviour CC BY 4.0

Was sie fanden

Das Hauptmuster war kategoriespezifisch.

Die Gesichtsdistinktheit im rechten lateralen ventralen temporalen Kortex korrelierte mit der Tendenz einer teilnehmenden Person, in der unabhängigen Free-Viewing-Aufgabe auf Gesichter zu blicken. Der Zusammenhang zeigte sich für erste Fixationen und für die Verweildauer. Die

Wortdistinktheit im linken lateralen ventralen temporalen Kortex korrelierte mit der Tendenz, auf Text zu blicken.

Die Arbeit prüfte auch, dass dies nicht bloß ein generischer Effekt im Sinne von „der visuelle Kortex ist bei manchen Menschen stärker“ war. Die stärksten Zusammenhänge folgten der erwarteten Kategorie und Hemisphäre: Gesichter im rechten lateralen VTC, Wörter im linken lateralen VTC. Kategorienübergreifende Zusammenhänge waren nicht die Geschichte.

Die neuronalen Maße standen auch mit dem Verhalten in Verbindung. In kleineren Teilstichproben war eine stärkere Gesichtsdistinktheit mit einer besseren Leistung im Cambridge Face Memory Test verknüpft, und eine stärkere Wortdistinktheit war mit einer schnelleren Leseleistung verknüpft. Das verleiht dem neuronalen Maß eine gewisse externe Bedeutung: Es ist nicht nur eine Scanner-Statistik, sondern ein Signal, das damit zusammenhängt, was Menschen tun können.

Was das nicht bedeutet

Die verlockende Schlagzeile wäre „Ihr Gehirn entscheidet, was Sie sehen“. Das ist zu stark.

Die Studie stellt die Richtung der Kausalität nicht fest. Eine Person blickt vielleicht mehr auf Gesichter, weil ihre Gesichtsrepräsentationen präziser sind. Oder ihre Gesichtsrepräsentationen sind präziser, weil jahrelanges Betrachten von Gesichtern diesen Teil des visuellen Systems geprägt hat. Oder beides entwickelt sich gemeinsam. Die Autoren lassen diese Entwicklungsfrage ausdrücklich offen.

Es bedeutet auch nicht, dass Menschen mit unterschiedlichen Blickgewohnheiten unterschiedliche physische Szenen sehen. Alle betrachteten dieselben Bilder. Der Unterschied liegt in der Abtastung: welche Objekte Priorität erhalten, welche Informationen zuerst gesammelt werden und welche Kategorien im visuellen Kortex distinkter repräsentiert sind.

Es ist auch kein diagnostischer Test für Einzelpersonen. Die Korrelationen sind auf Gruppenebene bedeutsam, aber sie sind kein Werkzeug, um zu sagen: „Die Hirnkarte dieser Person sagt genau voraus, wie sie dieses Bild betrachten wird.“

Warum es wichtig ist

Sehen wird oft so beschrieben, als wären die Augen eine Kamera, die ein Gehirn speist. Echtes Sehen ist aktiver als das. Das Auge wählt Moment für Moment, welche Information es in hohe Auflösung bringt. Diese Entscheidungen werden zum Input, aus dem das Gehirn lernt und handelt.

Diese Arbeit stellt diesen Kreislauf auf ein solideres Fundament. Sie verknüpft den aktiven Teil – Augenbewegungen durch Szenen – mit dem repräsentationalen Teil – kategorienselektiven Karten im ventralen visuellen Kortex – innerhalb derselben Individuen. Das Ergebnis macht es schwerer, „Organisation des visuellen Kortex“ und „visuelles Verhalten“ als getrennte Ebenen zu behandeln.

Zumindest bei Erwachsenen erscheinen sie aufeinander abgestimmt.

Diese Abstimmung ist die eigentliche Geschichte. Nicht, dass Gesichter-Betrachter die eine und Text-Betrachter die andere Art von Mensch sind. Nicht, dass eine Hirnregion eine Persönlichkeit erklärt. Das Ergebnis ist enger gefasst und nützlicher: Stabile Unterschiede darin, wie Menschen visuelle Szenen erkunden, decken sich mit stabilen Unterschieden darin, wie scharf ihr visueller Kortex die Dinge repräsentiert, die sie tendenziell aufsuchen.

Kurz zusammengefasst

Eine Studie in *Nature Human Behaviour* verfolgte, wie 102 Erwachsene 700 komplexe Szenen betrachteten, und scannte dann eine Teilmenge von 61 Teilnehmenden per fMRT, um kategorienselektive visuelle Reaktionen zu kartieren. Menschen, die tendenziell zuerst und länger auf Gesichter blickten, zeigten distinktere Gesichtsrepräsentationen im rechten lateralen ventralen temporalen Kortex; Menschen, die tendenziell auf Text blickten, zeigten distinktere Wortrepräsentationen im linken lateralen ventralen temporalen Kortex. Diese neuronalen Maße hingen in kleineren Teilstichproben auch mit der Gesichtserkennung und der Leseleistung zusammen. Die Studie ist korrelativ, nicht kausal: Sie zeigt, dass aktive Blickgewohnheiten und kategorienselektive Hirnorganisation bei Erwachsenen aufeinander abgestimmt sind, nicht, welche von beiden die andere hervorbringt.

Nüchtern geprüft

Was die Arbeit zeigt: Stabile individuelle Tendenzen, in natürlichen Szenen auf Gesichter oder Text zu blicken, sind mit passenden kategorienselektiven Repräsentationen im ventralen visuellen Kortex verknüpft.

Was plausibel, aber nicht bewiesen ist: Dass langfristige visuelle Erfahrung und Hirn-Feinabstimmung einander über die Entwicklung hinweg verstärken. Die Arbeit ist mit dieser Idee vereinbar, prüft aber die Entwicklungsrichtung nicht.

Was sie nicht zeigt: Dass Menschen buchstäblich unterschiedliche Welten sehen; dass Blickgewohnheiten fest verdrahtet sind; oder dass die fMRT-Karten die Augenbewegungen verursachen.

Wesentliche Einschränkungen: Korrelatives Design; erwachsene Universitätsstichprobe; kleinere Verhaltens-Teilstichproben für Gesichtserkennung und Lesen; und ein separater fMRT-Lokalizier statt gleichzeitiger fMRT beim freien Betrachten.

Wie viel Vertrauen sollte ein allgemeiner Leser haben? Hoch, dass die Blicktendenzen in dieser Stichprobe real und stabil sind, und mäßig bis hoch, dass sie mit passenden kategorienselektiven visuellen Repräsentationen verknüpft sind. Gering für jede kausale Erzählung, bis Entwicklungs- oder Interventionsarbeiten sie direkt prüfen.

Quellen

Kollenda et al., Nature Human Behaviour (2026), DOI 10.1038/s41562-026-02494-5
<https://doi.org/10.1038/s41562-026-02494-5>

OSF repository for anonymized data, code and stimuli
<https://osf.io/9fhxk/>