



The CLEAN PAPER

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

你注视的对象，可能与视觉大脑的调谐方式相匹配

一项 *Nature Human Behaviour* 研究把稳定的个人注视习惯——人们在复杂场景中是否倾向于首先并最久地看人脸或文字——与视觉皮层中相应类别选择性区域的表征区分度和大小联系起来。结果并不是说人们确实看见了不同的世界，也不是说大脑导致了这种注视模式。它是一种谨慎的相关性：在成年人中，主动观看与大脑的类别脑图似乎相互匹配。

Written by Laura Nesso · Cross-checked by Vera Rubin · 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)

目光并非随机游移

让两个人面对同一个繁杂场景，他们探索场景的方式不会相同。一个人的目光会迅速跳向人脸，另一个人则不断落在文字上。这些差异不只是一时的选择。在这项研究中，它们是稳定的个人特征：面对数百幅图像时，人们在首先看什么以及停留多久方面，都表现出可靠的个人倾向。

有意思的问题是，这些习惯究竟与什么相关。它们只是偏好吗——善于社交的人看人脸，爱阅读的人看文字——还是与每个人的视觉大脑如何表征这些类别有关？

Diana Kollenda、Elaheh Akbari、Maximilian Broda 和 Benjamin de Haas 直接检验了这种联系。他们把自然场景自由观看期间的眼动追踪，与另一项独立的 fMRI 实验结合起来；后者绘制了每位参与者的视觉皮层对人脸和文字等类别的反应。研究结果可以谨慎地概括为：更倾向于看人脸的人，其右侧腹侧视觉皮层中的人脸表征区分度更高；更倾向于看文字的人，其左侧腹侧视觉皮层中的文字表征区分度更高。

这并不意味着大脑强迫眼睛以某种方式移动，也不意味着看文字本身就能建立一个文字区。论文没有证明因果关系。但它确实表明，主动视觉——一个人用眼睛对世界进行采样的方式——与其自身视觉系统的精细布局相匹配。

作者做了什么

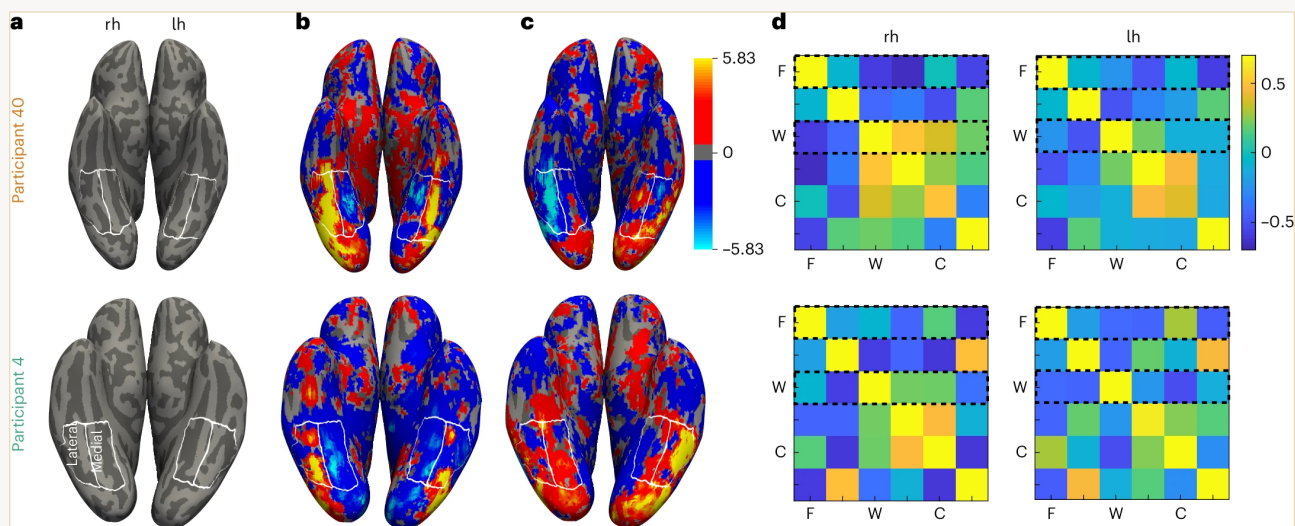
研究由两个明确分开的部分组成。

首先，102 名成年人自由观看了 700 幅复杂场景图像，同时研究人员记录他们眼球运动。针对人脸和文字，研究人员测量了两件事：参与者首先看向哪里，以及在那里持续观看多久。总观看时间称为**停留时间**。

他们还检验了这些习惯是否稳定。**折半信度**背后的思路很简单：把图像分成两组，在两组中测量同一种习惯，再看看排名较高或较低的是否仍是同一批人。如果答案是肯定的，这种习惯就具有信度。这里的信度很高：对人脸而言，首次注视的 $r = 0.93$ ，停留时间的 $r = 0.95$ ；对文字而言，分别为 $r = 0.87$ 和 $r = 0.88$ 。数值接近 1，意味着这种倾向非常稳定。

其次，其中 61 名参与者完成了另一项独立的功能性磁共振成像实验。功能性磁共振成像，简称 fMRI，通过追踪血液氧合变化，间接反映哪些脑区在任务期间更活跃。**功能定位任务**是一种标准的脑区绘制任务：展示人脸或文字等已知类别，识别对某一类别反应强于其他类别的皮层区域。这不是自由观看。研究人员以区块形式展示人脸、伪词、身体、房屋、汽车和肢体，同时要求参与者持续注视中央。这种设计很重要：脑部测量结果并不只是参与者在扫描仪中把目光移向相同物体所造成的。

随后，研究人员考察了每个人腹侧颞叶皮层中的类别反应有多强的区分度。人脸模式区分度越高，意味着人脸引起的脑活动越稳定地呈现人脸特征，也越容易与其他类别区分。文字模式区分度越高，则意味着文字和字符也具有同样的情况。



从研究的受试者名单中选取两名参与者的 fMRI 示例图；上排和下排代表两个不同的人。图 A 中的白色轮廓标出了研究分析的腹侧颞叶皮层区域。图 B 和图 C 显示两种不同的对比：B 突出对人脸反应更强的皮层，C 突出对书面文字反应更强的皮层。图 D 以矩阵形式概括各物体类别之间反应模式的相似程度。矩阵为 6 x 6，因为功能定位任务使用了六种刺激类别；F 代表人脸，W 代表文字，C 代表汽车，其余行列则是其他对照类别（身体、房屋和肢体）。同类别相似度越高、跨类别相似度越低，意味着大脑对该类别的表征区分度越高。重点不在于每位参与者都有相同的脑图，而在于研究逐人测量了这些脑图。

Kollenda et al. / Nature Human Behaviour CC BY 4.0

他们发现了什么

主要模式具有类别特异性。

右侧外侧腹侧颞叶皮层中的人脸表征区分度，与参与者在独立自由观看任务中看向人脸的倾向相关。这种关系同时出现在首次注视和停留时间上。左侧外侧腹侧颞叶皮层中的文字表征区分度，则与看向文字的倾向相关。

论文还检验了这是否只是某些人的“视觉皮层整体更强”所造成的效应。最强的联系遵循了预期的类别和大脑半球：人脸对应右侧外侧腹侧颞叶皮层，文字对应左侧外侧腹侧颞叶皮层。跨类别联系并不是研究的主要结果。

神经测量结果也与行为有关。在规模较小的子样本中，人脸表征区分度越高，在剑桥人脸记忆测试中的表现越好；文字表征区分度越高，阅读速度越快。这为神经测量赋予了一些外部意义：它不只是扫描仪中的统计量，还与人们能够完成的任务有关。

这并不意味着什么

诱人的标题会是“大脑决定你看见什么”。这种说法太强了。

研究没有确定因果关系的方向。一个人可能因为其人脸表征更精确而更多地看向人脸；也可能因为多

年观看人脸塑造了视觉系统的这一部分，使其人脸表征更精确；或者两者共同发展。作者明确表示，这个发展问题仍然没有答案。

这也不意味着目光习惯不同的人看见了不同的物理场景。所有人看到的都是相同图像。差异在于采样方式：哪些物体得到优先关注，哪些信息首先被收集，以及哪些类别在视觉皮层中得到区分度更高的表征。

这也不是针对个人的诊断测试。这些相关性在群体层面具有意义，但不能用来断言“这个人的脑图能够准确预测他会怎样看这幅图”。

为什么重要

人们常把视觉描述成眼睛像摄像机一样向大脑传送信息。真实的视觉过程远比这更主动。眼睛时时刻刻都在选择把哪些信息送入高分辨率视野。这些选择构成大脑学习和行动所依据的输入。

这篇论文让这一循环有了更坚实的依据。它在同一个人身上，把主动部分——目光如何在场景中移动——与表征部分——腹侧视觉皮层中的类别选择性脑图——联系起来。这个结果使人们更难再把“视觉皮层组织”和“视觉行为”视为彼此分离的层面。至少在成年人中，两者似乎相互匹配。

这种匹配才是真正的结果。它并不是说爱看人脸的人属于一种人，爱看文字的人属于另一种人，也不是说一个脑区能够解释一种人格。结果的范围更窄，也更有用：人们探索视觉场景方式中的稳定差异，与其视觉皮层表征他们所偏好类别时的清晰程度中的稳定差异相互对应。

清晰摘要

一项 *Nature Human Behaviour* 研究追踪了 102 名成年人观看 700 幅复杂场景的方式，随后对其中 61 名参与者进行 fMRI 扫描，以绘制类别选择性视觉反应。倾向于首先并更长时间看人脸的人，其右侧外侧腹侧颞叶皮层中的人脸表征区分度更高；倾向于看文字的人，其左侧外侧腹侧颞叶皮层中的文字表征区分度更高。在规模较小的子样本中，这些神经测量结果还分别与人脸识别和阅读表现有关。研究是相关性的，不是因果性的：它表明成年人的主动注视习惯与类别选择性大脑组织相匹配，而没有确定哪一方产生了另一方。

不绕弯核查

论文表明了什么：在自然场景中稳定地倾向于看人脸或文字的个人差异，与腹侧视觉皮层中相应的类别选择性表征有关。

什么合理但尚未证明：长期视觉经验与大脑调谐在发育过程中彼此强化。论文与这一想法一致，但没有检验发育方向。

它没有证明什么：人们确实看见了不同的世界；注视习惯是天生固定的；或者 fMRI 脑图导致眼球运

动。

主要局限：相关性研究设计；大学环境中的成年人样本；人脸识别和阅读分析的行为子样本较小；使用独立的 fMRI 功能定位任务，而不是在自由观看的同时进行 fMRI。

普通读者应该有多大信心？可以高度确信这些注视倾向在本样本中真实且稳定，并可以持中等至较高信心认为，它们与相匹配的类别选择性视觉表征有关。在发育研究或干预研究直接检验之前，对任何因果解释都应持低信心。

来源

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/>