



The CLEAN PAPER

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

你注視的對象，可能與視覺大腦的調諧方式相匹配

一項 *Nature Human Behaviour* 研究把穩定的個人注視習慣——人們在複雜場景中是否傾向於首先並最久地看人臉或文字——與視覺皮質中相應類別選擇性區域的表徵區分度和大小聯繫起來。結果並不是說人們確實看見了不同的世界，也不是說大腦導致了這種注視模式。它是一種謹慎的相關性：在成年人中，主動觀看與大腦的類別腦圖似乎相互匹配。

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

眼睛並不是隨機遊移

把兩個人置於同一個繁忙場景前，他們探索場景的方式不會相同。一個人的眼睛會迅速移向臉部；另一個人的目光則不斷落在文字上。這些差異不只是當下的選擇。在這項研究中，它們是穩定的特質：面對數百張影像，人們首先看什麼，以及在那裡停留多久，都展現出可靠的個人傾向。

有趣的問題是，這些習慣究竟與什麼有關。它們只是偏好嗎——喜歡社交的人看臉，喜歡閱讀的人看文字——還是與每個人的視覺腦如何表徵這些類別有關？

Diana Kollenda、Elaheh Akbari、Maximilian Broda 和 Benjamin de Haas 直接檢驗了這個連結。他們結合自然場景自由觀看時的眼動追蹤，以及另一項獨立的 fMRI 實驗，繪製每位參與者的視覺皮質如何回應臉部和文字等類別。他們的結果可以謹慎地概括如下：偏好看臉的人，在右側腹側視覺皮質中具有區辨性更高的臉部表徵；偏好看文字的人，在左側腹側視覺皮質中具有區辨性更高的文字表徵。

這不表示大腦會迫使眼睛以某種方式移動，也不表示只要看文字就會自行建立文字區域。這篇論文並未證明因果關係。但它確實表示，主動視覺——一個人用眼睛取樣世界的方式——與其視覺系統中細緻的組織相互匹配。

作者做了什麼

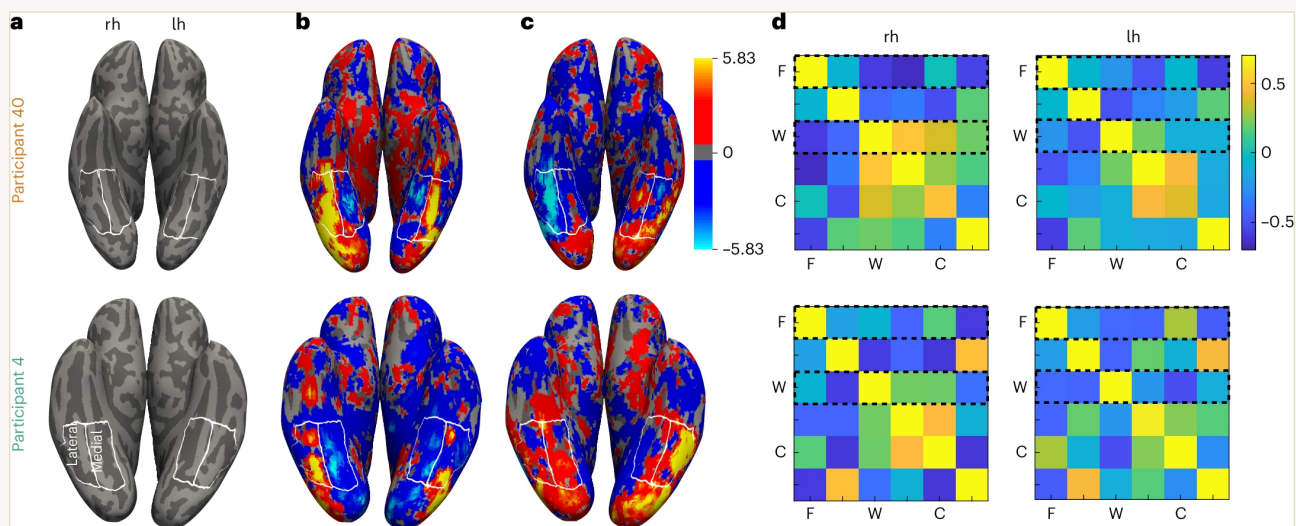
這項研究由兩個清楚分開的部分組成。

首先，102 名成人自由觀看 700 張複雜的場景影像，同時記錄他們的眼球運動。研究人員針對臉部和文字測量兩件事：參與者首先看向哪裡，以及他們在那裡持續看了多久。這段總觀看時間稱為 **停留時間**（dwell time）。

他們也檢查了這些習慣是否穩定。**折半信度**（split-half reliability）背後的簡單概念是：把影像分成兩組，在兩組中測量相同的習慣，看看同一批人的排名是否仍然較高或較低。如果答案是肯定的，這項習慣就具有信度。這裡的信度很高：臉部的首次注視為 $r = 0.93$ ，停留時間為 $r = 0.95$ ；文字則分別為 $r = 0.87$ 和 $r = 0.88$ 。接近 1 的數值表示這種傾向非常穩定。

其次，其中 61 名參與者完成了另一項獨立的功能性磁振造影實驗。功能性磁振造影（functional MRI），或稱 fMRI，會追蹤血液含氧量的變化，將其作為任務期間哪些腦區較活躍的間接訊號。**定位器**（localizer）是一項標準的映射任務：呈現臉部或文字等已知類別，找出對其中一個類別的反應比對其他類別更強的皮質區塊。這不是自由觀看。參與者注視中央位置，同時觀看以區塊方式呈現的臉部、偽詞、身體、房屋、汽車和肢體。這項設計很重要：腦部測量結果不只是參與者在掃描器內把眼睛移向相同物體所造成的結果。

接著，研究人員探討每個人在腹側顳葉皮質中的類別反應有多具區辨性。更具區辨性的臉部模式，表示大腦對臉部的活動模式更可靠地呈現臉部特徵，也更容易與其他類別區分。更具區辨性的文字模式，對文字和字元而言也是如此。



研究受試者名單中兩名參與者的 fMRI 地圖範例；頂列和底列是兩個不同的人。在面板 A 中，白色輪廓標示研究分析的腹側顳葉皮質區域。面板 B 和 C 顯示兩種不同的對比：B 突顯對臉部反應較強的皮質，而 C 突顯對書寫文字反應較強的皮質。面板 D 以矩陣形式總結不同物體類別之間的反應模式有多相似。矩陣是 6 x 6，因為定位器使用了六種刺激類別；F 代表臉部，W 代表文字，C 代表汽車，其餘列和欄則是其他比較類別（身體、房屋和肢體）。同類別相似度更高、跨類別相似度更低，表示該類別的腦部表徵更具區辨性。重點不是每位參與者都有相同的地圖，而是研究逐一測量了每個人的地圖。

Kollenda et al. / Nature Human Behaviour CC BY 4.0

他們發現了什麼

主要模式具有類別特異性。

右側外側腹側顳葉皮質中的臉部表徵區辨性，與參與者在獨立自由觀看任務中看臉的傾向相關。這種關係在首次注視和停留時間兩項指標中都出現。左側外側腹側顳葉皮質中的文字表徵區辨性，則與看文字的傾向相關。

論文也檢查了這是否只是一種一般性的「有些人的視覺皮質較強」效應。最強的連結符合預期的類別和半球：右側外側 VTC 對應臉部，左側外側 VTC 對應文字。跨類別連結並不是這個故事的核心。

神經測量也與行為相連。在較小的子樣本中，更強的臉部表徵區辨性與劍橋臉孔記憶測驗中較佳的表現相關，而更強的文字表徵區辨性則與較快的閱讀表現相關。這讓神經測量具有某種外部意義：它不只是掃描器中的統計數值，也是一個與人們能做什麼有關的訊號。

這不代表什麼

最容易吸引人的標題會是「你的大腦決定你看見什麼」。這說得太過頭了。

這項研究沒有確立因果方向。一個人可能因為其臉部表徵更精確，所以更常看臉。也可能是因為多年看臉的經驗塑造了視覺系統的那一部分，所以其臉部表徵更精確。或者，兩者可能共同發展。作者明

確地把這個發展問題留待日後研究。

這也不表示具有不同注視習慣的人看見不同的物理場景。所有人看的都是相同影像。差異在於取樣：哪些物體獲得優先處理、哪些資訊最先被蒐集，以及哪些類別在視覺皮質中被表徵得更具區辨性。

這也不是針對個人的診斷測驗。這些相關性在群體層級具有意義，但不能用來說「這個人的腦圖能精確預測他會如何看這張圖片」。

為什麼重要

人們常把視覺描述成眼睛如同相機，將影像傳送給大腦。但真實的視覺比這更主動。眼睛會一刻接一刻地選擇要把哪些資訊帶入高解析度。這些選擇成為大腦學習和行動所依據的輸入。

這篇論文為這個迴路提供了更堅實的依據。它把主動的部分——眼睛在場景中的移動——與表徵的部分——腹側視覺皮質中的類別選擇性地圖——連結起來，而且是在同一批個體中進行。這項結果讓人更難把「視覺皮質的組織」和「視覺行為」視為彼此分離的層次。至少在成人中，兩者似乎是相互匹配的。

這種匹配才是故事的核心。重點不是看臉的人是一種人，看文字的人是另一種人。也不是某個腦區能解釋一種人格。結果更狹窄，卻也更有用：人們探索視覺場景的方式存在穩定差異，而這些差異與視覺皮質表徵他們傾向尋找之事物的清晰程度，也呈現穩定的對應。

重點摘要

一項發表於 *Nature Human Behaviour* 的研究追蹤 102 名成人如何觀看 700 個複雜場景，接著以 fMRI 掃描其中 61 名參與者的子樣本，繪製類別選擇性的視覺反應。傾向首先且較久看臉的人，在右側外側腹側顳葉皮質中呈現區辨性更高的臉部表徵；傾向看文字的人，在左側外側腹側顳葉皮質中呈現區辨性更高的文字表徵。這些神經測量也與較小子樣本中的臉部辨識和閱讀表現相關。這項研究是相關性的，而非因果性的：它顯示成人的主動注視習慣與類別選擇性的大腦組織相互匹配，但沒有顯示兩者之中哪一個造成另一個。

務實檢視

論文顯示了什麼：在自然場景中看臉或文字的穩定個人傾向，與腹側視覺皮質中相應的類別選擇性表徵有關。

合理但尚未證實的部分：長期視覺經驗與大腦調諧會在發展過程中彼此強化。這篇論文與這個想法一致，但沒有檢驗發展方向。

它沒有顯示什麼：人們實際上看見不同的世界；注視習慣是天生固定的；或 fMRI 腦圖會造成眼球運動。

主要限制：相關性設計；成人大學樣本；用於臉部辨識和閱讀的行為子樣本較小；以及使用獨立的 fMRI 定位器，而不是同時進行自由觀看的 fMRI。

一般讀者應有多大信心？對於這個樣本中的注視傾向確實存在且穩定，信心應該很高；對於這些傾向與相應的類別選擇性視覺表徵有關，信心則應該是中等至偏高。在發展研究或介入研究直接檢驗之前，對任何因果說法的信心都應該很低。

來源

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/>