
SABER DÓNDE BUSCAR UN OBJETIVO OCULTO

LEANNE CHUKOSKIE ^a, JOSEPH SNIDER ^a,
MICHAEL C. MOZER ^{a,b,c}, RICHARD J. KRAUZLIS ^d
TERRENCE J. SEJNOWSKI ^{a,e,t,g}

ABSTRACT. Survival depends on successfully foraging for food, for which evolution has selected diverse behaviors in different species. Humans forage not only for food, but also for information. We decide where to look over 170,000 times per day, approximately three times per wakeful second. The frequency of these saccadic eye movements belies the complexity underlying each individual choice. Experience factors into the choice of where to look and can be invoked to rapidly redirect gaze in a context and task-appropriate manner. However, remarkably little is known about how individuals learn to direct their gaze given the current context and task. We designed a task in which participants search a novel scene for a target whose location was drawn stochastically on each trial from a fixed prior distribution. The target was invisible on a blank screen, and the participants were rewarded when they fixated the hidden target location. In just a few trials, participants rapidly found the hidden targets by looking near previously rewarded locations and avoiding previously unrewarded locations. Learning trajectories were well characterized by a simple reinforcement-learning (RL) model that maintained and continually updated a reward map of locations. The RL model made further predictions concerning sensitivity to recent experience that were confirmed by the data. The asymptotic performance of both the participants and the RL model approached optimal performance characterized by an ideal-observer theory. These two complementary levels of explanation show how experience in a novel environment drives visual search in humans and may extend to other forms of search such as animal foraging.

KEYWORDS. Ideal observer, oculomotor, reinforcement learning, saccades.

La influencia de la evolución se puede ver en los comportamientos de forrajeo, que se han estudiado en la ecología del comportamiento. Una suposición de los modelos económicos de forrajeo es que se toman decisiones para maximizar la compensación y minimizar el gasto de energía. Por ejemplo, una abeja que se prepara para salir en busca de flores que están brotando puede viajar kilómetros para encontrar las fuentes de alimento. Buscar información sobre un entorno es una parte importante

^aInstitute for Neural Computation y ^bDivision of Biological Sciences, University of California at San Diego, La Jolla, CA 92093; ^cDepartment of Computer Science e ^dInstitute of Cognitive Science, University of Colorado, Boulder, CO 80309; ^eLaboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD 20892; y ^fHoward Hughes Medical Institute and ^tComputational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 9203. / terry@salk.edu / Este artículo contiene información suplementaria en: www.pnas.org/lookup/suppl/doi:10.1073/pnas.1301210110/-/DCSupplemental.

de la búsqueda de alimento. Las abejas necesitan identificar los objetos a cierta distancia que se asocian con las fuentes de alimento. Los seres humanos también son expertos en la búsqueda de elementos en el mundo y en el aprendizaje de cómo encontrarlos. Este estudio explora el problema de cómo los humanos aprenden dónde mirar bajo el contexto de búsqueda del forraje animal.

Nuestras actividades cotidianas dependen del éxito en las estrategias de búsqueda para encontrar objetos en nuestro ambiente. La búsqueda visual está casi siempre presente en las tareas rutinarias: encontrar nuestro automóvil en un estacionamiento, hallar las llaves de casa en un escritorio desordenado o localizar el botón que se desea oprimir en un teclado de computadora. Al buscar un objeto en los contextos de escenarios comunes, las personas extraen rápidamente información acerca de dónde se encuentran normalmente los objetivos (Castelhana y Heaven, 2010, 2011; Itti y Koch, 2000; Neider y Zelinsky, 2006; Oliva y Torralba, 2006; Potter, 1975; Rayner, Castelhana y Yang, 2009; Torralba, Oliva, Castelhana y Henderson, 2006; Vó y Henderson, 2010). Esta capacidad de utilizar lo "esencial" (*gist*) de una imagen (Oliva y Torralba, 2006; Torralba, et al., 2006) le permite a las personas tener un rendimiento de manera flexible y eficiente en ambientes conocidos. Añádase a eso la secuencia predecible de los movimientos oculares que se produce cuando una persona está ocupada en una tarea manual (Hayhoe y Ballard, 2005) y se hace evidente que, a pesar de la gran cantidad de investigaciones sobre cómo la imagen sobresaliente guía la mirada (Itti y Koch, 2000; Parkhurst y Niebur, 2003), las asociaciones espaciales aprendidas son quizás igual de importantes para la participación efectiva en nuestro entorno visual (Hayhoe y Ballard, 2005; Tatler y Vincent, 2009; Chun y Jiang, 1998). Es de sorprender, sin embargo, que poca investigación se destine a cómo los individuos aprenden a dirigir la mirada en un contexto y modo apropiados a una tarea en ambientes novedosos.

Cierta investigación relevante para aprender dónde buscar proviene de la literatura sobre los movimientos oculares, las recompensas y su valor esperado. Al igual que toda la conducta motora, los movimientos sacádicos se influyen por la recompensa y se producen en una latencia más corta para los objetivos más valorados (Milstein y Dorris, 2007). De hecho, hallar algo que usted está buscando puede ser algo intrínsecamente gratificante (Xu-Wilson, Zee y Shadmehr, 2009). Al perfeccionar la bien conocida relación canónica de "secuencia principal" entre la amplitud de la sacada y la velocidad, el valor del objetivo de una sacada puede alterar los detalles del plan motor que se llevó a cabo, ya sea acelerando o desacelerando la propia sacada, dependiendo del valor de ese objetivo para el sujeto (Shadmehr, 2010; Shadmehr, Orban de Xivry, Xu-Wilson y Shih, 2010). Este resultado es de especial interés a la luz de la investigación que indica que

son diferentes las características de bajo nivel del estímulo (Tatler, Baddeley y Vincent, 2006), las cuales tienen una distribución esperada de fijaciones de atracción (Reinagel y Zador, 1999) y tal vez también valorables de manera diferente (Açyk, Sarwary, Schultze-Kraft, Onat y König, 2010) en función de su distancia a partir de la ubicación actual de fijación. En conjunto, estos resultados ponen de relieve la compleja interacción de la información externa e interna para guiar la elección del movimiento del ojo.

Dos estudios fundacionales de Buswell (1935) y Yarbus (1967) anticiparon los conceptos modernos de mapa de prioridad o de prominencia al mostrar que algunas porciones de una imagen se fijan con mayor probabilidad que otras. Ambos investigadores también proporcionaron las primeras pruebas de que este mapa de prioridad cambia de modo eficaz, dependiendo del tipo de información que se busca. Yarbus observó que eran enormemente diferentes los patrones de la mirada que siguen después de diferentes preguntas basadas en la escena o en tareas que se asignaban al observador, sugiriendo que el observador sabía dónde buscar información en la escena para responder a la pregunta y miraba específicamente a las áreas que contenían dicha información cuando se necesitaba. Henderson y colaboradores (Castelhana, Mack y Henderson, 2009) han replicado este resultado para diferentes tareas de búsqueda visual y memorización de la imagen. Sin embargo, Wolfe y sus colaboradores (Greene, Liu y Wolfe, 2012), usando una pregunta y paradigma ligeramente diferentes, no encontraron evidencia de que los patrones de las sacadas fueran predictivos de estados mentales específicos. Sin importar las réplicas específicas a la demostración de Yarbus, queda claro que lo esencial de la escena —la información de contexto específico acerca de dónde normalmente se encuentran los objetos— surge rápidamente y guía la búsqueda del objetivo de una escena con un contexto conocido (Torralba, et al., 2006). Por ejemplo, cuando a un observador se le muestra una escena de una calle, sabrá de manera inmediata dónde buscar las señalizaciones, automóviles y peatones (Fig. 1A).

Castelhana y Heaven (2011) también han demostrado que además de lo esencial de la escena misma, las asociaciones espaciales aprendidas también guían los movimientos oculares durante la búsqueda. Los sujetos utilizan estas asociaciones aprendidas, al igual que otras experiencias basadas en el contexto, como la probabilidad de los estímulos y las recompensas y sanciones pasadas (Schütz, Trommershäuser y Gegenfurtner, 2012) para perfeccionar el objetivo de un movimiento sacádico del ojo. Una revisión y comentario reciente de Wolfe, et al. (2011) explora la noción de orientación “semántica” en escenas naturales complejas, como proporcionar conocimiento de la probabilidad de hallar un objeto conocido en una parte concreta de una escena. Esta perspectiva relaciona los trabajos sobre

lo esencial de la escena con las tareas de búsqueda visual clásicas, ofreciendo un marco para considerar cómo los individuos podrían utilizar la experiencia pasada para dirigir la mirada tanto a las escenas reales como a los escenarios artificiales de nuestros laboratorios.

Muy distinta de la literatura de búsqueda visual es la literatura de otro tipo de búsqueda que generalmente necesitan los animales y las personas: el forrajeo. Los agentes buscan alimento, el cual a menudo está oculto en el entorno en el que buscan (Fig. 1B). La búsqueda de alimento oculto recompensa los cambios no sólo con la posición de la recompensa, sino también con el tamaño de la distribución de las recompensas (Charnov, 1976). Otros trabajos han concebido la conducta de forrajeo en términos de búsqueda óptima (Bénichou, Coppey, Moreau, Suet y Voituriez, 2005). Lo que distingue el forrajeo de las tareas de búsqueda visual, es que las tareas de búsqueda visual contienen señales visibles que orientan la búsqueda, además de la información contextual que especifica la ubicación probable del objetivo. Para que la búsqueda visual se asemeje más al forrajeo, podemos despojar las señales visibles de esa búsqueda visual. Una tarea de búsqueda visual carente de señales visuales permitiría determinar si existen similitudes subyacentes entre estos dos tipos de búsqueda y si pueden surgir principios generales de búsqueda para tal investigación.

Las personas participantes en la exploración a gran escala destacaron la importancia de buscar objetivos ocultos e incluso invisibles y que se aproxima al forrajeo animal (Gilchrist, North y Hood, 2001; Smith, Gilchrist y Hood, 2005). En uno de tales paradigmas (Smith, Gilchrist y Hood, 2005), se pidió a niños que exploraran una habitación con un suelo compuesto por baldosas tipo caja, una de las cuales contenía un elemento de recompensa. Resulta interesante que los niños exploraron el ambiente de forma diferente cuando se les instruyó que buscaran con su mano no dominante que con su mano dominante. En concreto, se necesitaban más "revisitas" en la condición de la mano no dominante. Este resultado sugiere que el esfuerzo en el aprendizaje y el esfuerzo motor interviene en el rendimiento de tareas que podrían parecer automáticas, lo que sugiere métodos para modelar el comportamiento como el forrajeo. El esfuerzo motor adicional que se requeriría para reducir las revisitas metabólicamente costosas en un escenario de forrajeo, parecen haber implicado sistemas de memoria en un grado mayor de lo que normalmente se observa en las tareas de búsqueda "visual" tradicionales.

El marco de aprendizaje por reforzamiento (AR) ha sido ampliamente aceptado para el modelado del rendimiento en tareas que implican una serie de movimientos que conducen a la recompensa (Montague y Sejnowski, 1994; Sutton, 1988). Además, para los organismos en muchos niveles de complejidad, el AR ha demostrado ser un marco adecuado para considerar el comportamiento adaptativo en ambientes complejos y cam-

biantes (Lee, Seo y Jung, 2012; Niv, Joel, Meilijson y Ruppín, 2002). Aquí describimos el rendimiento en nuestra tarea en términos de una perspectiva AR. Las trayectorias de aprendizaje de los participantes se caracterizaron perfectamente por un modelo de aprendizaje por reforzamiento simple que mantenía y actualizaba continuamente un mapa de ubicaciones de recompensa. El modelo de AR hizo además predicciones adicionales sobre la sensibilidad a la experiencia reciente y que fueron confirmadas por los datos. El rendimiento asintótico, tanto de los participantes como del modelo AR se aproximó a un rendimiento óptimo caracterizado por una teoría de observador ideal, que supone un conocimiento perfecto de la distribución del objetivo estático y de las elecciones de fijación de modo independiente. Estos dos niveles complementarios de explicación muestran cómo la experiencia en un ambiente novedoso dirige la búsqueda visual en los seres humanos.

RESULTADOS

Los seres humanos aprenden rápidamente a buscar objetivos ocultos

Con respecto a la búsqueda visual, experimentos anteriores no pudieron aislar totalmente el aspecto visual de un objetivo de la ubicación de la recompensa que se aprendió. En todos los casos estuvieron a disposición para guiar el movimiento la indicación visual de un objetivo o el recuerdo de un objetivo visible del momento previo (Stritzke y Trommershäuser, 2007) y sus alrededores. Para entender cómo los participantes aprenden dónde buscar en una escena novedosa o en un contexto donde no existe ninguna relación entre los objetivos visuales y las recompensas o sanciones asociadas, diseñamos una tarea de búsqueda en la que los participantes se premiaban por hallar un objetivo oculto, similar al escenario para el forrajeo animal (Fig. 1C).

Los participantes buscaban de manera repetida una sola escena desconocida (contexto) para un objetivo. Sin embargo, para estudiar el papel del conocimiento de la tarea al guiar la búsqueda, además de las señales visuales que normalmente se utilizan para identificar un objetivo, el objetivo se presentó como invisible. La tarea de los participantes consistía en explorar la pantalla con la mirada y encontrar un objetivo oculto que sonaría con un tono de recompensa cuando lo fijaban. Sin el conocimiento de los participantes, la posición del objetivo oculto varió en cada ensayo y se extrajo de una distribución gaussiana con un centroide y dispersión (la media del objetivo y DE [desviación estándar], respectivamente) que se mantuvo constante a lo largo de un periodo de sesiones (Fig. 1C).

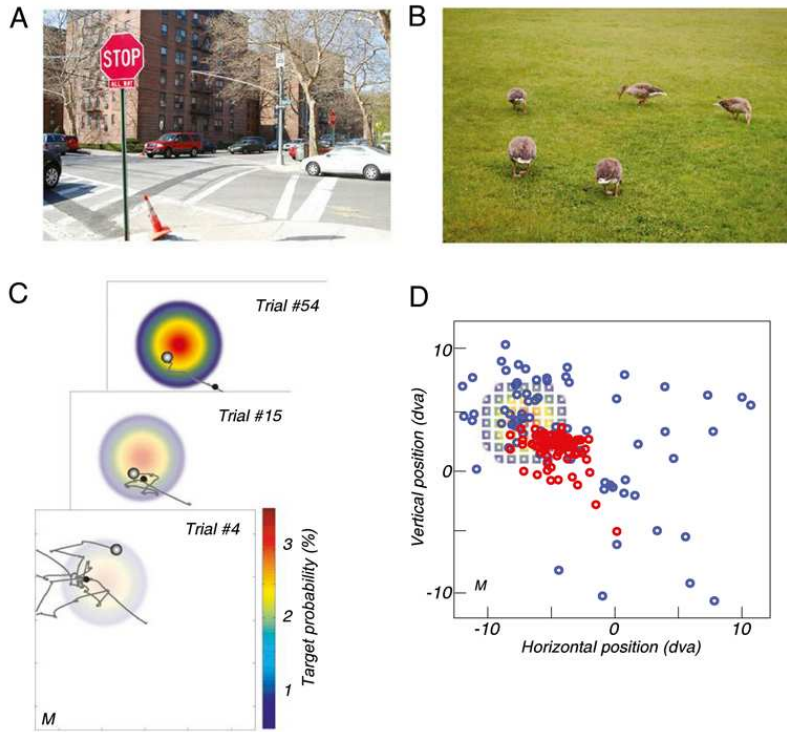


FIGURA 1. Tareas de búsqueda visible y oculta. (A) Un peatón experimentado tiene conocimiento previo de dónde buscar signos, automóviles y aceras en esta escena de calle. (B) Patos forrajeando en una gran extensión de césped. (C) Una representación de la pantalla se superpone con la distribución del objetivo oculto que se aprende durante la sesión, así como muestra de rastros del ojo de los tres ensayos para el participante M. La primera fijación de cada ensayo está marcada con un círculo negro. La fijación final y recompensada está marcada por un círculo sombreado en escala de grises. (D) La región de la pantalla muestreada con fijación se contrae de toda la pantalla en los primeros ensayos (círculos azules; 87 fijaciones sobre los primeros cinco ensayos) hacia una región que se aproxima al tamaño y posición de las ubicaciones del objetivo distribuidos como entero-gaussiano (cuadrados con color proporcional a la probabilidad como aparece en A) en ensayos posteriores (círculos rojos; 85 fijaciones de los ensayos 32–39). Los datos de la posición de fijación son del participante M.

Al inicio de una sesión, los participantes no tenían ningún conocimiento previo para informar sobre su búsqueda; su búsqueda inicial era efectivamente “ciega”. Conforme procedía la sesión los participantes acumulaban información de obtener o no obtener recompensa en los puntos de fijación,

además de mejorar su tasa de éxito desarrollando una expectativa para la distribución de los objetivos ocultos y usarla para guiar futuras búsquedas (Fig. 1D).

Después de muy pocos ensayos, los participantes reunieron suficiente información sobre la distribución del objetivo para dirigir la mirada de manera eficiente cercano a la distribución real del objetivo, como lo demuestran los datos de uno de los participantes en Fig. 1 C y D. Observamos un patrón similar de aprendizaje para todos los participantes: las primeras fijaciones se dispersaron en general a través de la pantalla de búsqueda; después de aproximadamente una docena de ensayos, las fijaciones se redujeron a la región con las probabilidades más elevadas del objetivo.

En la Figura 2A se muestra una caracterización de este efecto para todos los participantes. La distancia promedio entre el centroide de la distribución del objetivo hacia las fijaciones individuales en un ensayo, decayó de manera precipitada aproximadamente en la primera docena de ensayos. La Figura 2A muestra esta distancia para todos los participantes en la condición de dispersión del objetivo de 2°. La distancia asintótica del centroide se incrementó de forma monótona con la dispersión del objetivo (Tabla 1).

Una medida de la extensión de la búsqueda es la DE del conjunto de fijaciones en un ensayo. Al inicio, la dispersión de la búsqueda fue amplia y se estrechaba conforme avanzaba la sesión, como se muestra en la Fig. 2B para todos los participantes en la condición de dispersión del objetivo de 2°. La dispersión de la búsqueda asintótica se incrementó de manera monótona con la condición de dispersión del objetivo (Tabla 1). Estos datos sugieren que los participantes estimaron la propagación de la distribución del objetivo oculto y conforme a ella, ajustaron la distribución de su búsqueda. Asimismo, la mediana del número de fijaciones que llevaron a cabo los participantes para encontrar el objetivo (en ensayos de hallar el objetivo), disminuyó rápidamente en una sola sesión para alcanzar una asíntota (Fig. 2C).

Los seres humanos se acercan al rendimiento de observador ideal

Consideramos ahora el comportamiento de los participantes una vez que se estabilizó el rendimiento. Tomando los ensayos 31–60 para reflejar el comportamiento asintótico, examinamos la eficacia de la búsqueda humana en comparación con un óptimo teórico. Se derivó un Observador Ideal para la Tarea de Búsqueda de Objetivo Oculto asumiendo que las fijaciones son mutuamente independientes y que se conoce la distribución del objetivo destino, además de minimizar el número esperado de ensayos (Figs. S1 y Tabla S1). Las líneas discontinuas en la Figura 2 marcan el rendimiento del observador ideal. El rendimiento en una búsqueda ideal requiere una distribución de “conjeturas” (*guesses*) de fijación planeadas y

que es $\sqrt{2}$ más amplia que la propia distribución del objetivo (Snider, 2011). Como se aprecia en la Figura 2 B y C, el rendimiento de los participantes se cernía alrededor de esta distribución ideal de búsqueda después de aproximadamente una docena de ensayos.

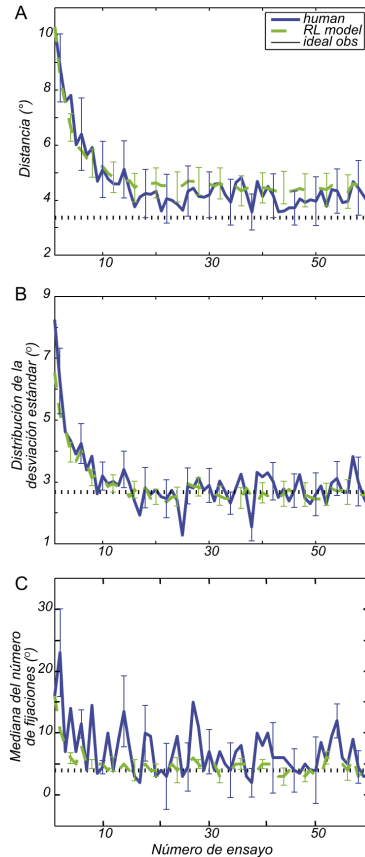


FIGURA 2. Curvas de aprendizaje para la tarea de búsqueda de objetivo oculto. (A) Se muestra en azul y verde la distancia entre la media del conjunto (*cluster*) de fijación para cada ensayo para el centroide del objetivo, promediado entre los participantes, e indica el resultado de las 200 simulaciones del modelo de aprendizaje por reforzamiento para los parámetros de cada participante. El MES [modelos de ecuaciones estructurales] se da para ambos. La predicción de observador ideal se indica por la línea punteada negra. (B) La DE de las distribuciones de la posición del ojo o "extensión de la búsqueda" se muestra para el promedio de todos los participantes (azul) y para el modelo AR (verde) con MES. La línea discontinua es el óptimo teórico del observador ideal en cada caso, asumiendo un conocimiento perfecto de la distribución del objetivo. (C) La mediana del número de fijaciones que se llevan a cabo para encontrar el objetivo en cada ensayo se muestra (azul) junto con la predicción del modelo AR (verde) del número de fijación. El MES se muestra para ambos.

TABLA 1. Rendimiento de la asíntota de aprendizaje para los participantes, la teoría del observador ideal y un modelo de aprendizaje por reforzamiento

Condición de dispersión del objetivo	Distancia media desde el centroide del objetivo hacia las fijaciones en los ensayos 31-60°	Distribución de búsqueda en los ensayos 31-60°
Datos del participante		
0.75	1.97	1.14
2.00	4.08	2.80
2.75	4.39	3.70
Teoría del observador ideal		
0.75	0.70	0.56
2.00	3.36	2.68
2.75	4.47	3.78
Modelo de aprendizaje por reforzamiento		
0.75	3.21	1.56
2.00	4.46	2.61
2.75	6.07	4.29

Datos, teoría y estadística del modelo para la distancia media de fijación y distribución de búsqueda para las condiciones de distribución del objetivo en 0.75°, 2.0° y 2.75°.

Los sujetos mostraron un sesgo de $\sim 1^\circ$ hacia el centro de la pantalla con relación a la distribución del objetivo (Tabla S2), pero el cálculo del comportamiento ideal asumía que los sujetos buscaban de manera simétrica alrededor del centro de la distribución del objetivo. Aunque la adición del sesgo hace analíticamente insostenibles a las matemáticas, un buscador simulado se aproximaba al número esperado de sacadas necesario para encontrar un objetivo con 1° de sesgo sistemático (Fig. 3). En lo esencial no hubo ningún cambio en la predicción del número de movimientos sacádicos o en la propagación de la búsqueda (ubicación del mínimo en la Figura 3), excepto para el caso de 0.75° en la distribución del objetivo, donde el óptimo pasó de una propagación de la búsqueda de 0.56° a 0.85° . Intuitivamente, el efecto del sesgo era pequeño debido a que éste era menor que 2° del radio del objetivo. No obstante, a un nivel de confianza de 95 por ciento en las tres distribuciones del objetivo, se ajustaron de manera cualitativa y cuantitativa con las predicciones el número de pasos, la propagación de búsqueda y el tamaño del paso, suponiendo que el número de movimientos sacádicos se minimizaba.

El modelo de aprendizaje por reforzamiento se ajusta al aprendizaje humano

Además de la teoría del observador ideal y que caracteriza la eficiencia asíntótica de búsqueda humana, desarrollamos una explicación comple-

mentaria y mecanicista que capturó las diferencias individuales y en el aprendizaje, y la dinámica del comportamiento humano. La teoría AR, motivada por el aprendizaje animal y los experimentos conductuales (Yu y Cohen, 2008), sugiere un modelo sencillo e intuitivo que construye un mapeo de ubicaciones de función de valor en el espacio para la recompensa esperada. La función de valor se actualiza después de cada fijación con base en si el objetivo se localiza y se utiliza además para seleccionar los destinos de las sacadas que muy probablemente se recompensarán.

Aumentamos este modelo intuitivo con dos supuestos adicionales. En primer lugar, cada vez que se hace una sacada para una ubicación, la retroalimentación que se obtiene se generalizaba a las ubicaciones espaciales cercanas; en segundo lugar, incorporamos un sesgo de proximidad que favorecía las sacadas más breves. En los datos se presentó una preferencia por las sacadas más breves (Fig. S2) que también observaron otros investigadores (Yarbus, 1967; Sutton, y Barto, 1998, p. xviii), algunos de los cuales mostraron que dicha preferencia puede remplazar el conocimiento que tienen los participantes acerca de la ubicación esperada de un objetivo (Araujo, Kowler y Pavel, 2001). La incorporación de un sesgo de proximidad en el modelo cambió la naturaleza de la tarea debido a que la elección de la fijación siguiente se hizo dependiente de la fijación actual. En consecuencia, los participantes deben planear las secuencias de fijación en vez de elegir fijaciones independientes.

Modelamos la tarea mediante métodos de diferencia temporal (Sutton, 1988), los cuales son particularmente apropiados para las tareas markovianas en las que las secuencias de acciones conducen a la recompensa (*modelo de aprendizaje por reforzamiento* y Figs. S2 y S3 para detalles). Los parámetros libres del modelo se ajustaron a cada secuencia de fijaciones en los primeros veinte ensayos. Teniendo en cuenta estos parámetros, el modelo se corrió en un modo generativo a partir de un estado *de novo* para simular al sujeto llevando a cabo la tarea.

La Figura 2 muestra el rendimiento promedio del modelo en paralelo con el rendimiento promedio humano. El modelo también predijo una dispersión de búsqueda asintótica que incrementó con la dispersión del objetivo (Tabla 1), consistente con el rendimiento total de los participantes. Similar al rendimiento humano que se observa en la Fig. 2A, el modelo AR se aproxima, pero no alcanza la asíntota teórica. Como los participantes humanos, el modelo AR es sensible a lo no estacionario en la distribución, en tanto que la teoría de observador ideal supone que la distribución es estática. Además, el modelo dio cuenta de las diferencias individuales (*modelo de aprendizaje por reforzamiento*). Existen varias razones de por qué la consistencia observada entre los participantes y las simulaciones puede ser más que una prueba de la existencia y podría proporcionar la penetración en los mecanismos biológicos de aprendizaje. El modelo AR en sí tuvo

una dinámica emergente que se reflejaba en el comportamiento humano (Fig. 4 y efectos secuenciales discutidos más abajo). También el criterio utilizado para entrenar el modelo era la verosimilitud de una secuencia específica de fijación. Se utilizó una amplia gama de medidas estadísticas muy distintas a partir del criterio de entrenamiento para comparar el rendimiento humano y del modelo: distancia media desde el centroide del objetivo, DE la distribución de los movimientos oculares y mediana del número de fijaciones (Fig. 2). Por último, sólo los veinte primeros ensayos se utilizaron para entrenar el modelo, pero todas las comparaciones que se muestran en la Tabla 1 se obtuvieron de los ensayos 31–60.

La Figura 2 sugiere que los participantes adquieren la distribución del objetivo en alrededor de una docena de ensayos, después de lo cual su desempeño se vuelve estático. Sin embargo, en el modelo AR la función de valor se ajusta después de cada fijación, sin disminuir con el tiempo. Una señal de este ajuste continuo consiste en una dependencia secuencial en los ensayos —en específico, una dependencia entre la fijación final de un ensayo y la fijación inicial del siguiente. De hecho, se observaron las dependencias en los datos a lo largo de un periodo de sesiones (Fig. 4A), como lo predijo el modelo (Fig. 4B) y éstas explicaron alguna variabilidad de ensayo a ensayo en el rendimiento (Fig. 2 y *modelo de aprendizaje por reforzamiento*). Se predisponía a los participantes para iniciar la siguiente búsqueda de ensayo cerca de las ubicaciones del objetivo que se encontraron en los ensayos recientes. La influencia de los ensayos previos disminuye de forma exponencial; los dos o probablemente tres ensayos previos, influyeron en la elección de sacada del ensayo actual (Fig. 4C). Este amortiguamiento exponencial de la influencia de los ensayos anteriores se aproxima por el caso sin memoria (Snider, 2011), permitiendo que coexistan de manera asintótica tanto el modelo AR como el planeador ideal.

Distribución bimodal de las longitudes de la sacada

Nuestra motivación al diseñar la tarea de búsqueda de un objetivo oculto era vincular la búsqueda visual y la literatura sobre el forrajeo. El rendimiento en nuestra tarea tenía características análogas a las que se encuentran en el contexto más amplio del forrajeo animal (Fig. 5). Aunque los ensayos individuales se asemejaban a los vuelos de Lévy —una mezcla de excursiones grandes de fijación y esporádicas y que se conocen por ser óptimas en algunos casos de la conducta de forrajeo (Schultz, Dayan y Montague, 1997; Humphries, et al., 2010; James, Plank y Edwards, 2011)— la distribución de la longitud de todos los segmentos de línea recta no es tipo Lévy, sino que se separa en dos escalas de longitud diferentes como la búsqueda intermitente que popularizaron Bénichou, et al. (2005). La escala de longitud más corta, las fijaciones de menos de 1°, corresponde a

una ley de búsqueda de poder local con un exponente muy pronunciado, convirtiéndolo en un recorrido aleatorio clásico que sondea (*samples*) el espacio local de manera densa. Tal búsqueda local se combina con los movimientos sacádicos más grandes, pero más raros representados por la joroba puntiaguda (*peaked hump*) en tamaños de paso más grandes que 1° . Estas son las distintas opciones de la distribución planeada que ya se describió (es decir, la distribución de la conjetura o función de valor). La forma distintiva como rodilla en la Figura 5 es similar a la que se halla en otras tareas de búsqueda visual demandantes (Snider, 2011), así como el forrajeo intermitente por una amplia gama de animales (Bénichou, et al., 2005; Humphries, et al., 2010).

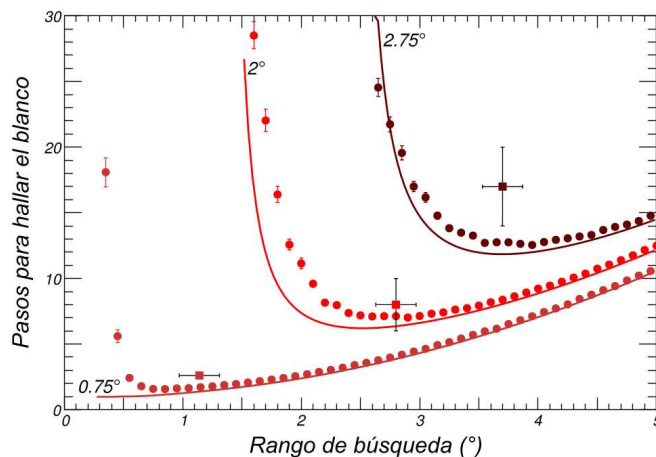


FIGURA 3. Modelo de búsqueda óptima. Se estimó mediante simulación (círculos con media y EEs [errores estándar] a partir de 100 000 ensayos por punto) y cálculo teórico (líneas sólidas) el número teórico de pasos de búsqueda para encontrar el objetivo para las distribuciones del objetivo de tamaño 0.75° (naranja), 2° (rojo) y 2.75° (marrón) como se detalla en la Tabla S1 y la Información de Apoyo. La simulación incluye el sesgo de 1° que se observó en los sujetos, pero no las líneas de la teoría. Las cajas sólidas indican los valores observados para los sujetos (media y EE). Con la adición del sesgo, el mínimo se movió ligeramente a la derecha pero sólo fue significativo para la distribución del objetivo en 0.75° . El costo en términos de movimientos sacádicos extras para dispersiones de búsqueda no óptimas (lejos del mínimo) fue mayor para las distribuciones del objetivo más grandes, y el ascenso comparativamente poco profundo para dispersiones de búsqueda por encima de lo óptimo significaba que si los sujetos habrían de equivocarse, entonces debían dirigirse hacia distribuciones más grandes. De hecho, la tendencia para dispersiones más grandes era evidente cuando los sujetos comenzaban con dispersiones grandes y disminuían hacia el mínimo (Fig. 2). Los pasos adicionales que los sujetos tomaban para encontrar el objetivo para la distribución de 2.75° (parte superior derecha) fue consistente con la tendencia hacia pequeños movimientos sacádicos a pesar de que estaban muy cerca de la mínima correcta (Fig. S2): los movimientos sacádicos más grandes se pudieron haber dividido en múltiples movimientos sacádicos cortos.

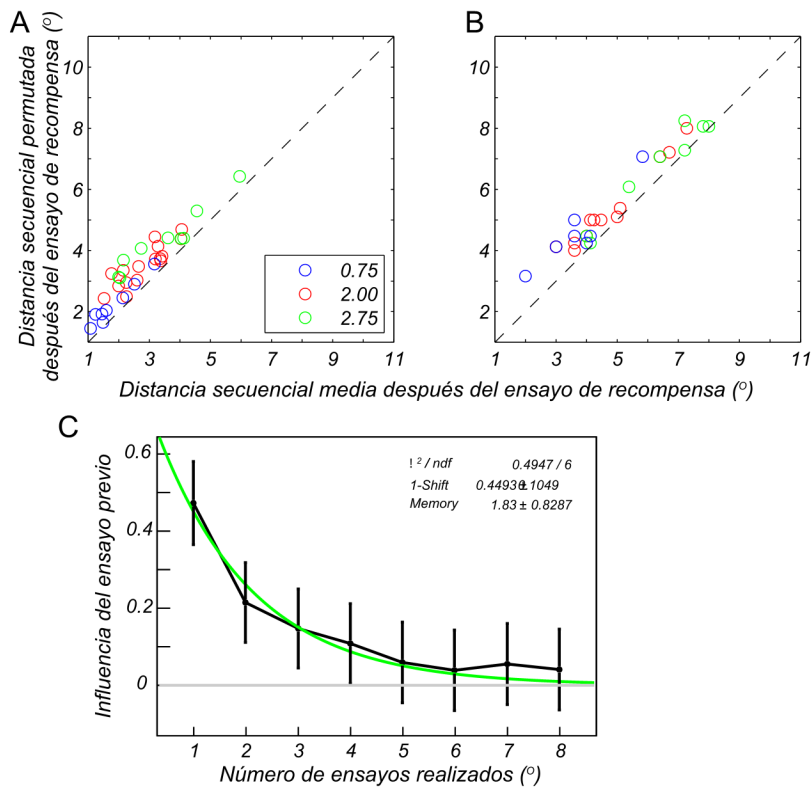


FIGURA 4. Efectos secuenciales en los datos humanos y las predicciones del modelo AR. (A) Para cada sujeto, dispusimos la distancia secuencial media entre ensayos (la distancia entre la fijación final en el ensayo n y la primera fijación en el ensayo $n+1$ cuando el ensayo n produce una recompensa) contra la distancia permutada entre ensayos (la distancia entre la fijación final en un ensayo y la primera fijación de otro ensayo al azar). Cada círculo denota un sujeto y el círculo en color indica la condición de la dispersión del objetivo (azul, $\sigma = 0.75$; rojo, $\sigma = 2.00$; verde, $\sigma = 2.75$). En correspondencia con la predicción del modelo (B), la distancia secuencial media entre ensayos es confiablemente menor que la distancia permutada entre ensayos, como lo indican los puntos que yacen sobre la diagonal. Todas las distancias entre ensayos son más grandes en el modelo, lo que refleja un mayor grado de exploración que en los participantes, pero este desfase es ortogonal a los efectos secuenciales. (C) El efecto de los ensayos previos al buscar en el ensayo actual se representa como una función del número de ensayos hacia atrás. Un ajuste exponencial de los datos se muestran en verde.

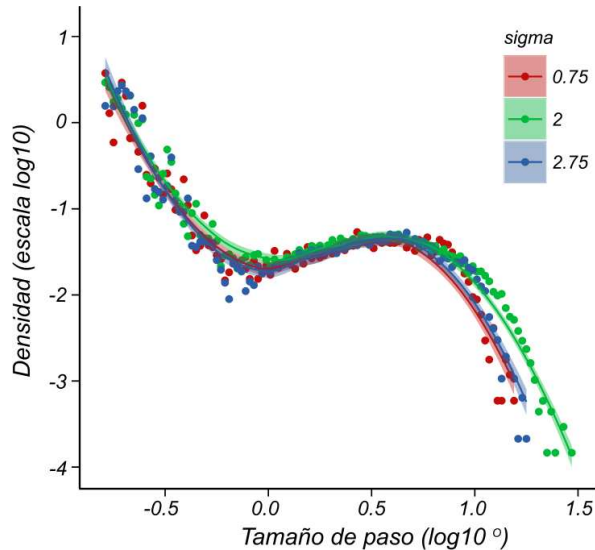


FIGURA 5. Distribuciones de longitud de los movimientos sacádicos en la tarea de objetivo oculto. Un algoritmo de punto de inflexión aplicado a los datos de movimiento crudo del ojo produce una distribución de tamaños de paso para todos los participantes (los detalles se dan en el *modelo de aprendizaje por reforzamiento*). Muy pequeños movimientos oculares de "fijación" constituyen el lado izquierdo de la representación gráfica y grandes saltos sacádicos mayores a la derecha para tres tamaños diferentes de distribución del objetivo. Todos los puntos y líneas (ajustes Loess [regresión local] con 95 por ciento de confianza para intervalo de sombreado) para cada tamaño de distribución de búsqueda comparten una forma similar, en particular, una curva en tamaños de paso cercana a 1° de ángulo visual.

DISCUSIÓN

El rendimiento de la búsqueda en los humanos se puede ubicar en el contexto más general del forrajeo animal, el cual posee estrechas relaciones con los modelos AR (Niv, Joel, Meilijson y Ruppin, 2002) y la teoría de la búsqueda óptima (Charnov, 1976). La tarea de búsqueda del objetivo oculto que se introdujo aquí nos ha permitido separar la influencia de los estímulos externos de la información interna previa para buscar recompensas en un ambiente novedoso (Viswanathan, Buldyrev, Havlin, da Luz, Raposo y Stanley, 1999). En nuestra tarea de búsqueda de objetivo oculto, los participantes exploraron un ambiente novedoso y aprendieron de modo rápido a alinear sus fijaciones con la región del espacio en el que se distribuyeron los objetivos invisibles de manera probabilística. Después de una docena de ensayos, la estadística de fijación estuvo cerca de ajustarse con aquella que se obtuvo por la teoría del observador ideal. Esta

cercanía de ajuste nos permitió considerar al rendimiento humano de búsqueda como una búsqueda óptima libre de memoria con un conocimiento perfecto de la distribución del objetivo. Como complemento a la teoría del observador ideal que se ocupa del rendimiento asintótico, desarrollamos una explicación mecanicista del aprendizaje por reforzamiento ensayo a ensayo. Nuestro modelo AR caracterizó el curso del tiempo del aprendizaje, logró un rendimiento asintótico cercano al observador ideal y vinculó el problema de búsqueda visual a una teoría más amplia de aprendizaje motivado.

Ambientes naturales

Los marcos de observador ideal y aprendizaje por reforzamiento proporcionan la base para una perspectiva teórica más amplia sobre la elección de sacada durante la visión natural, en el que las personas aprenden a buscar en contextos diversos objetivos visibles y en donde son claramente esenciales los rasgos visuales de la escena. En un marco bayesiano, los sujetos en nuestra tarea aprendieron la distribución previa de los objetivos ocultos. En un entorno natural, la distribución previa se combinaría con la información visual para determinar la distribución posterior, de la cual se generan los objetivos sacádicos.

Los ambientes naturales no son estacionarios. Por ejemplo, un animal que se esté alimentando puede agotar el suministro en una localidad y tiene que pasar a otra. Un buscador debe ser sensible a tales cambios en el ambiente. Las dependencias secuenciales (Fig. 4) son una señal de esta sensibilidad (Adams, Watson, Pearson y Platt, 2012; Fecteau y Munoz, 2003; Herrnstein, 1961): los objetivos recientes influyen en el comportamiento posterior, incluso después de que el buscador ha aprendido de manera aparente la distribución del objetivo, como se refleja en el rendimiento asintótico. El modelo AR predijo las dependencias secuenciales, las cuales generaron un comportamiento muy cercano al de los participantes como un grupo, además de capturar las idiosincrasias individuales (*modelo de aprendizaje por reforzamiento*). La sensibilidad a los ambientes no estacionarios puede explicar por qué nuestros participantes y el modelo AR lograron una distribución asintótica de búsqueda algo más alejada del centroide del objetivo de lo se predice en una teoría del observador ideal basada en la estacionariedad.

Una de las hazañas más impresionantes del forrajeo de animal es la conducta de igualación (*matching behavior*). La ley de igualación de Herrnstein (1961) describe cómo los animales que forrajean tienden a responder en proporción con el valor esperado de diferentes parches (*patches*). La conducta de igualación se ha estudiado en varias especies, desde las abejas a los seres humanos (Bradshaw, Szabadi y Bevan, 1976; Gallistel, King, Gottlieb, Balci, Papachristos, Szalecki y Carbone, 2007; Greggers y

Mauelshagen, 1997; Lau y Glimcher, 2005). Sin embargo, al buscar diferentes intervalos de reforzamiento sobre dos palancas o botones, muchos de estos estudios de laboratorio en realidad eliminan el elemento espacial del forrajeo de la tarea; en este escenario, los animales detectan rápidamente los cambios en los intervalos de reforzamiento (Mark y Gallistel, 1994); ya se examinó el esfuerzo motor al cambiar entre parches espaciales (Baum, 1982). En la naturaleza, el forrajeo se extiende de manera espacial, y el paradigma de la búsqueda del objetivo oculto podría servir como un ambiente eficaz para examinar una tarea de forrajeo explícitamente espacial en el contexto de la conducta de igualación. Por ejemplo, una versión de nuestro paradigma de búsqueda de objetivo oculto con una distribución bimodal podría explorar la conducta de cambio (*changeover behavior*) y el esfuerzo motor variando los tamaños de las dos distribuciones y la distancia entre ellas (Baum, 1982).

Bases neurales de la búsqueda

La neurobiología del comportamiento del movimiento del ojo ofrece una perspectiva alternativa sobre las similitudes de la conducta de búsqueda visual y el forrajeo. La cuestión de dónde buscar a continuación se ha explorado neurofisiológicamente, y células en diversas regiones del cerebro de macacos parecen llevar las señales de los componentes de la tarea necesarios para una búsqueda visual exitosa. El área lateral interparietal (ALI) y el colículo superior (CS) son dos regiones del cerebro que contienen un mapa de prioridad que representa la ubicación de estímulos relevantes que podrían servir como objetivo de la próxima sacada. Registros en el área ALI y SC del macaco muestran que este mapa de prioridad integra la información de las señales externas ("*bottom-up*") e internas ("*top-down*") en las tareas de búsqueda visual (Bisley y Goldberg, 2010; Fecteau y Munoz, 2006).

Recientemente, Bisley y colaboradores (Mirpour, Arcizet, Ong y Bisley, 2009) han utilizado una tarea de búsqueda visual como forraje para demostrar que células del área ALI distinguen entre objetivos y distractores, y mantienen una estimación corriente de las recompensas meta de la sacada. Las neuronas de ALI integran información de diferentes modalidades relevantes para el forrajeo con el fin de codificar el valor asociado a un movimiento para un objetivo particular (Klein, Deaner y Platt, 2008; Platt y Glimcher, 1999). Los mecanismos neuronales que sirven a las decisiones de forrajeo de parche quedarse-huir se han caracterizado recientemente en una tarea de elección visual simplificada (Hayden, Pearson y Platt, 2011), proporcionando un esquema para las investigaciones de cómo precisamente la información previa y otras exigencias de la tarea se combinan con la información visual disponible en la escena. La microestimulación de bajo umbral en el área ALI (Mirpour, Ong y Bisley, 2010) o en

CS (Carello y Krauzlis, 2004) también predispone a los movimientos sacádicos de selección hacia el objetivo en el campo estimulado. Tomados en conjunto, estos resultados sugieren que las áreas ALI y CS pueden ser los sustratos neurales mediando el mapa de las probables ubicaciones de la siguiente sacada en nuestra tarea, similar a la del mapa de valor en nuestro modelo AR.

Nos cuestionamos cómo los sujetos aprenden a elegir objetivos valiosos en un nuevo entorno. Recientes experimentos neurofisiológicos en los ganglios basales ofrecen algunas sugerencias sobre cómo se codifica la información previa para usarse al elegir el objetivo más valioso para la sacada en un ambiente complejo (Nakahara y Hikosaka, 2012). Hikosaka y colaboradores (Yasuda, Yamamoto y Hikosaka, 2012) han identificado señales que se relacionan con información de valor recientemente aprendida y aún lábil para objetivos de sacada en la parte anterior del núcleo caudado y con información de valor más estable en la región posterior del caudado, de la sustancia negra y de la parte reticulada. Debido a que las células que llevan esta información de valor estable parecen proyectarse de forma preferencial al SC, estas señales son propicias para influir en las decisiones de sacada a través de un circuito rápido y evolutivamente conservado para controlar la conducta de orientación. Estos resultados proporcionan una base neurofisiológica para entender cómo se aprende y consolida la experiencia para servir a los movimientos sacádicos que hacemos para recopilar información sobre nuestro ambiente en aproximadamente tres veces cada segundo.

CONCLUSIONES

En nuestra tarea de búsqueda de movimiento ocular, los sujetos aprendieron a elegir metas de sacada con base en la experiencia de recompensa previa que se separaba de los rasgos visuales específicos en una escena novedosa. El rendimiento de búsqueda resultante se describió perfectamente por un modelo AR similar al que se utilizó anteriormente para examinar tanto el comportamiento de forrajeo animal como el disparo neuronal de las células dopaminérgicas. Además, el rendimiento de búsqueda se aproximó a un óptimo teórico de rendimiento en esta tarea. Caracterizando cómo la experiencia previa guía la elección del movimiento ocular en nuevos contextos e integrándolo tanto en modelo como en teoría, hemos creado un marco para considerar cómo la experiencia previa guía la elección de sacada durante la visión natural. Se ha estudiado bien el sistema oculomotor de los primates, lo que permitirá descubrir los mecanismos neuronales subyacentes al aprendizaje y el rendimiento de la tarea de objetivo oculto y que se puede compartir con otras conductas de búsqueda.

MÉTODOS

Definimos una región espacial de una imagen como prominente al asociarla con la recompensa para examinar cómo los participantes utilizaron su experiencia previa de encontrar objetivos para dirigir futuras sacadas. Aprovechamos el hecho de que la meta de los movimientos oculares sacádicos es obtener información sobre el mundo y pedimos a los participantes humanos “llevar a cabo una búsqueda de movimiento del ojo para encontrar la ubicación de un objetivo recompensado tan pronto como fuera posible”. También se les dijo a los participantes que aprenderían más sobre los objetivos recompensados conforme avanzaba la sesión y que debían tratar de encontrar la ubicación del objetivo recompensado lo más rápido posible. Los objetivos recompensados no tenían ninguna representación visual en la pantalla y eran invisibles al sujeto. La pantalla era la misma en cada ensayo dentro de una sesión y no proporcionaba ninguna información sobre la ubicación del objetivo. La ubicación y la extensión de la distribución del objetivo recompensado variaban con cada sesión.

Cada ensayo comenzaba con una cruz de fijación central en una pantalla gris neutra con una luminancia media de 36.1 cd/m^2 (Fig. 1). La pantalla de búsqueda abarcaba 25.6° del centro de la visión del sujeto mientras estaba sentado con su cabeza inmovilizada por una mentonera.

Los participantes iniciaban cada ensayo pulsando un botón que indicaba que estaban fijando la cruz central. La misma pantalla gris neutra servía como la pantalla de búsqueda después de 300 ms de fijación de la cruz. Una vez que la cruz de fijación desaparecía, los participantes tenían 20 s para encontrar la ubicación recompensada para ese ensayo antes de que la pantalla de fijación volviera. En cada ensayo un objetivo invisible se extraía de una distribución predefinida de posibles objetivos. La forma de la distribución fue gaussiana con el centro en un número entero de grados de la región de fijación (generalmente $\pm 6^\circ$ en x e y) y la propagación se mantenía fija sobre cada sesión experimental. Los objetivos se produjeron sólo en valores enteros de la gaussiana. La expansión de la distribución (ED 0.75° , 2° y 2.75°) daba una probabilidad asociada a una ubicación de objetivo recompensada la cual variaba entre 4 y 0.1 por ciento. Cuando la mirada del sujeto se posaba en 2° de la meta en las direcciones x e y , un tono de recompensa marcaba el final exitoso del ensayo. Para que el objetivo pudiera “encontrarse”, la fijación (que se monitoreaba en tiempo real como se detalla a continuación) necesitaba permanecer constante dentro de la ventana del objetivo por al menos 50 s. Esta duración aseguraba que el objetivo nunca se encontrara simplemente escruñiando durante una sacada. Si al finalizar el segundo 20 s el objetivo no se encontraba, el ensayo terminaba sin ningún tono y aparecía una cruz de fijación indicando el comienzo de un nuevo ensayo.

El tiempo de ensayo y la recopilación de datos se administraron mediante el sistema de software TEMPO (Reflective Computing) y se interconectaron con la pantalla de estímulo usando una extensión de la Caja de Herramientas de Psicofísica (Brainard, 1997) en MATLAB (MathWorks). Se obtuvieron los datos del movimiento ocular utilizando un rastreador ocular de video (ISCAN), muestreado a 240 Hz para los seres humanos. Se calibraron los datos oculares con

los participantes mirando los estímulos en lugares conocidos. Los movimientos oculares se analizaron fuera de línea en MATLAB. Detectamos los movimientos sacádicos y los parpadeos utilizando un umbral de velocidad conservadora (40°/s con un hombro 5 ms después de cada sacada) después de distinguir las señales de posición del ojo. Después se marcaron periodos de fijación constante durante cada ensayo y se extrajeron análisis adicionales. Se descartaron del análisis las posiciones del ojo fuera de la pantalla de búsqueda. La inspección visual de los ensayos individuales confirmó que los periodos marcados de fijación eran libres de movimientos sacádicos o parpadeos.

Puntos de inflexión

Además de los movimientos sacádicos identificados por criterios de velocidad, los datos del seguimiento del ojo se procesaron para estimar la distribución del tamaño de paso de todos los movimientos de ojo, incluso dentro de una fijación. Para ello, primero se eliminaron los parpadeos retirando muestras fuera de la pantalla. A continuación, consideramos los puntos de los datos de tres en un tiempo, x_{t-1} , x_t y x_{t+1} , donde x son los puntos de los datos 2D y t indexa las muestras de tiempo, para la construcción de dos tramos del rastreo ocular $a = x_{t-1} - x_t$ y $b = x_t - x_{t+1}$. Entonces probamos si el coseno del ángulo entre los dos fue mayor o menor de 0.95. Si el coseno era superior a 0.95, entonces el punto central, x_t , se marcaba como un punto de "inflexión". Además, algunos de los grandes pasos se curvaban de forma lenta lo que introdujo extraños puntos (es decir, dividir un paso largo en dos pasos). Para superar este problema, aprovechamos el hecho de que dos pasos largos casi nunca ocurren uno tras otro sin una región de fijación densa en medio, y cualquier punto sin ningún vecino dentro de 0.5° se asumió como extraño y se eliminó. Esto dio lugar a puntos en los que el ojo hacía una desviación significativa del movimiento balístico y se utilizó para generar las distribuciones de tamaño de paso en la Figura 5.

AGRADECIMIENTOS. Agradecemos a Krista Kornlyo y Natalie Dill por la asistencia técnica. Este trabajo fue financiado en parte por el National Science Foundation (NSF) Grant SBE 0542013 para el Temporal Dynamics of Learning Center, por una subvención del NSF Science of Learning Center (a L.C., J.S., M.C.M. y T.J.S.), por una subvención Blasker Rose-Miah de la San Diego Foundation (a L.C.), por la Office of Naval Research Multidisciplinary University Research Initiative Award N00014-10-1-0072 (a J.S. y Valdés) y por el Howard Hughes Medical Institute (T.J.S.).

Traducción de Israel Grande-García.

REFERENCIAS

- Açyk, A., Sarwary, A., Schultze-Kraft, R., Onat, S. y König, P. (2010), "Developmental changes in natural viewing behavior: Bottom-up and top-down differences between children, young adults and older adults", *Frontiers in Psychology* 1: 207.
- Adams, G. K., Watson, K. K., Pearson, J. y Platt, M. L. (2012) "Neuroethology of decision-making", *Current Opinion in Neurobiology* 22(6): 982–989.
- Araujo, C., Kowler, E. y Pavel, M. (2001), "Eye movements during visual search: The costs of choosing the optimal path", *Vision Research* 41(25–26): 3613–3625.
- Baum, W. M. (1982), "Choice, changeover, and travel", *Journal of the Experimental Analysis of Behavior* 38(1): 35–49.
- Bénichou, O., Coppey, M., Moreau, M., Suet, P.-H. y Voituriez, R. (2005), "Optimal search strategies for hidden targets", *Physical Review Letters* 94(19): 198101–198104.
- Bisley, J. W. y Goldberg, M. E. (2010), "Attention, intention, and priority in the parietal lobe", *Annual Review of Neuroscience* 33: 1–21.
- Bradshaw, C. M., Szabadi, E. y Bevan, P. (1976), "Behavior of humans in variable-interval schedules of reinforcement", *Journal of the Experimental Analysis of Behavior* 26(2): 135–141.
- Brainard, D. H. (1997), "The Psychophysics Toolbox", *Spatial Vision* 10(4): 433–436.
- Buswell, G. T. (1935), *How People Look at Pictures: A Study of the Psychology of Perception in Art*. Chicago: University of Chicago Press.
- Carello, C. D. y Krauzlis, R. J. (2004), "Manipulating intent: Evidence for a causal role of the superior colliculus in target selection", *Neuron* 43(4): 575–583.
- Castelano, M. S. y Heaven, C. (2010), "The relative contribution of scene context and target features to visual search in scenes", *Attention, Perception, & Psychophysics* 72(5): 1283–1297.
- Castelano, M. S. y Heaven, C. (2011), "Scene context influences without scene gist: Eye movements guided by spatial associations in visual search", *Psychonomic Bulletin & Review* 18(5): 890–896.
- Castelano, M. S., Mack, M. L. y Henderson, J. M. (2009), "Viewing task influences eye movement control during active scene perception", *Journal of Vision* 9(3): 1–15.
- Charnov, E. L. (1976), "Optimal foraging, the marginal value theorem", *Theoretical Population Biology* 9(2): 129–136.
- Chun, M. M. y Jiang, Y. (1998), "Contextual cueing: Implicit learning and memory of visual context guides spatial attention", *Cognitive Psychology* 36(1): 28–71.
- Fecteau, J. H. y Munoz, D. P. (2003), "Exploring the consequences of the previous trial", *Nature Reviews Neuroscience* 4(6): 435–443.
- Fecteau, J. H. y Munoz, D. P. (2006), "Saliency, relevance, and firing: A priority map for target selection", *Trends in Cognitive Sciences* 10(8): 382–390.
- Gallistel, C. R., King, A. P., Gottlieb, D., Balci, F., Papachristos, E. B., Szalecki, M. y Carbone, K. S. (2007), "Is matching innate?" *Journal of the Experimental Analysis of Behavior* 87(2): 161–199.
- Geng, J. J. y Behrmann, M. (2005), "Spatial probability as an attentional cue in visual search", *Perception & Psychophysics* 67(7): 1252–1268.
- Gilchrist, I. D., North, A. y Hood, B. (2001), "Is visual search really like foraging?", *Perception* 30(12): 1459–1464.

- Greene, M. R., Liu, T. y Wolfe, J. M. (2012), "Reconsidering Yarbus: A failure to predict observers' task from eye movement patterns", *Vision Research* 62: 1–8.
- Greggers, U. y Mauelshagen, J. (1997), "Matching behavior of honeybees in a multiple-choice situation: The differential effect of environmental stimuli on the choice process", *Animal Learning & Behavior* 25(4): 458–472.
- Hayden, B. Y., Pearson, J. M. y Platt, M. L. (2011), "Neuronal basis of sequential foraging decisions in a patchy environment", *Nature Neuroscience* 14(7): 933–939.
- Hayhoe, M. y Ballard, D. (2005). "Eye movements in natural behavior", *Trends in Cognitive Sciences* 9(4): 188–194.
- Herrnstein, R. J. (1961), "Relative and absolute strength of response as a function of frequency of reinforcement", *Journal of the Experimental Analysis of Behavior* 4: 267–272.
- Humphries, N. E., Queiroz, N., Dyer, J. R. M., Pade, N. G., Musyl, M. K., Schaefer, K. M., Fuller, D. W., Brunnschweiler, J. M., Doyle, T. K., Houghton, J. D. R., Hays, G. C., Jones, C. S., Noble, L. R., Wearmouth, V. J., Southall, E. J. y Sims, D. W. (2010), "Environmental context explains Lévy and Brownian movement patterns of marine predators", *Nature* 465 (7301): 1066–1069.
- Itti, L. y Koch, C. (2000), "A saliency-based search mechanism for overt and covert shifts of visual attention", *Vision Research* 40(10-12): 1489–1506.
- James, A., Plank, M. J. y Edwards, A. M. (2011), "Assessing Lévy walks as models of animal foraging", *Journal of the Royal Society Interface* 8(62): 1233–1247.
- Klein, J. T., Deaner, R. O. y Platt, M. L. (2008), "Neural correlates of social target value in macaque parietal cortex", *Current Biology* 18(6): 419–424.
- Lau, B. y Glimcher, P. W. (2005), "Dynamic response-by-response models of matching behavior in rhesus monkeys", *Journal of the Experimental Analysis of Behavior* 84(3): 555–579.
- Lee, D., Seo, H. y Jung, M. W. (2012), "Neural basis of reinforcement learning and decision making", *Annual Review of Neurosciences* 35: 287–308.
- Mark, T. A. y Gallistel, C. R. (1994), "Kinetics of matching", *Journal of Experimental Psychology: Animal Behavior Processes* 20(1): 79–95.
- Milstein, D. M. y Dorris, M. C. (2007), "The influence of expected value on saccadic preparation", *Journal of Neuroscience* 27(18): 4810–4818.
- Mirpour, K., Arcizet, F., Ong, W. S. y Bisley, J. W. (2009), "Been there, seen that: A neural mechanism for performing efficient visual search", *Journal of Neurophysiology* 102(6): 3481–3491.
- Mirpour, K., Ong, W. S. y Bisley, J. W. (2010), "Microstimulation of posterior parietal cortex biases the selection of eye movement goals during search", *Journal of Neurophysiology* 104(6): 3021–3028.
- Montague, P. R. y Sejnowski, T. J. (1994), "The predictive brain: Temporal coincidence and temporal order in synaptic learning mechanisms", *Learning & Memory* 1(1): 1–33.
- Nakahara, H. y Hikosaka, O. (2012), "Learning to represent reward structure: A key to adapting to complex environments", *Neuroscience Research* 74(3-4): 177–183.
- Neider, M. B. y Zelinsky, G. J. (2006), "Scene context guides eye movements during visual search", *Vision Research* 46(5): 614–621.
- Niv, Y., Joel, D., Meilijson, I. y Ruppin, E. (2002), "Evolution of reinforcement learning in uncertain environments: A simple explanation for complex foraging behaviors", *Adaptive Behavior* 10(1): 5–24.

- Oliva, A. y Torralba, A. (2006), "Building the gist of a scene: The role of global image features in recognition", *Progress in Brain Research* 155: 23–36.
- Parkhurst, D. J. y Niebur, E. (2003), "Scene content selected by active vision", *Spatial Vision* 16(2): 125–154.
- Platt, M. L. y Glimcher, P. W. (1999), "Neural correlates of decision variables in parietal cortex", *Nature* 400(6741): 233–238.
- Potter, M. C. (1975), "Meaning in visual search", *Science* 187(4180): 965–966.
- Rayner, K. (1998), "Eye movements in reading and information processing: 20 years of research", *Psychological Bulletin* 124(3): 372–422.
- Rayner, K., Castelano, M. S. y Yang, J. (2009), "Eye movements when looking at unusual/weird scenes: Are there cultural differences?", *Journal of Experimental Psychology: Learning, Memory and Cognition* 35(1): 254–259.
- Reinagel, P. y Zador, A. M. (1999), "Natural scene statistics at the centre of gaze", *Network* 10(4): 341–350.
- Schultz, W., Dayan, P. y Montague, P. R. (1997), "A neural substrate of prediction and reward", *Science* 275(5306): 1593–1599.
- Schütz, A. C., Trommershäuser, J. y Gegenfurtner, K. R. (2012), "Dynamic integration of information about salience and value for saccadic eye movements", *Proceedings of the National Academy of Sciences USA* 109(19): 7547–7552.
- Shadmehr, R. (2010), "Control of movements and temporal discounting of reward", *Current Opinion in Neurobiology* 20(6): 726–730.
- Shadmehr, R., Orban de Xivry, J. J., Xu-Wilson, M. y Shih, T. Y. (2010), "Temporal discounting of reward and the cost of time in motor control", *Journal of Neuroscience* 30(31): 10507–10516.
- Smith, A. D., Gilchrist, I. D. y Hood, B. M. (2005), "Children's search behaviour in large-scale space: Developmental components of exploration", *Perception* 34(10): 1221–1229.
- Snider, J. (2011), "Optimal random search for a single hidden target", *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics* 83(1 Pt 1): 011105.
- Stritzke, M. y Trommershäuser, J. (2007), "Eye movements during rapid pointing under risk", *Vision Research* 47(15): 2000–2009.
- Sutton, R. S. (1988), "Learning to predict by the method of temporal differences", *Machine Learning* 3(1): 9–44.
- Sutton, R. S. y Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Tatler, B. W. y Vincent, B. T. (2009), "The prominence of behavioural biases in eye guidance", *Vision and Cognition* 17(6–7): 1029–1054.
- Tatler, B. W., Baddeley, R. J. y Vincent, B. T. (2006), "The long and the short of it: Spatial statistics at fixation vary with saccade amplitude and task", *Vision Research* 46(12): 1857–1862.
- Torralba, A., Oliva, A., Castelano, M. S. y Henderson, J. M. (2006), "Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search", *Psychological Review* 113(4): 766–786.
- Viswanathan, G. M., Buldyrev, S. V., Havlin, S., da Luz, M. G. E., Raposo, E. P. y Stanley, H. E. (1999), "Optimizing the success of random searches", *Nature* 401 (6756): 911–914.
- Võ, M. L. y Henderson, J. M. (2010), "The time course of initial scene processing for eye movement guidance in natural scene search", *Journal of Vision* 10(3): 11–13.
- Wilder, M. H., Mozer, M. C. y Wickens, C. D. (2011), "An integrative, experience-based theory of attentional control", *Journal of Vision* 11(2), 10.1167/11.2.8.

- Wolfe, J. M., Võ, M. L., Evans, K. K. y Greene, M. R. (2011), "Visual search in scenes involves selective and nonselective pathways", *Trends in Cognitive Sciences* 15(2): 77–84.
- Xu-Wilson, M., Zee, D. S. y Shadmehr, R. (2009), "The intrinsic value of visual information affects saccade velocities", *Experimental Brain Research* 196(4): 475–481.
- Yarbus, A. L. (1967), *Eye Movements and Vision*. Nueva York: Plenum.
- Yasuda, M., Yamamoto, S. y Hikosaka, O. (2012), "Robust representation of stable object values in the oculomotor Basal Ganglia", *Journal of Neuroscience* 32(47): 16917–16932.
- Yu, A. J. y Cohen, J. D. (2008), "Sequential effects: Superstition or rational behavior?", en D. Koller, D. Schuurmans, Y. Bengio y L. Bottou (eds.), *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, Vol. 21, pp 1873–1880.