



## Entre ver e interpretar: o papel fundamental do Design

*Between seeing and interpreting: fundamental role of Design*

Celia Matsunaga

*visão; movimento dos olhos; eye tracking; design*

*Embora “ver” seja um ato inerente ao ser humano, damos pouca ou nenhuma importância a ele. Abrimos e fechamos, movimentamos os olhos sem perceber o mecanismo por meio do qual os estímulos visuais – que traduzem a realidade em que estamos inseridos – são reconhecidos por nossos sentidos. Estudar a percepção visual como processo de recepção desses estímulos ambientais, remete-nos a um vasto campo do conhecimento. Por essa razão, optou-se por focalizar o estudo do reconhecimento visual e da tecnologia de detecção e avaliação textual durante a leitura. Por meio da investigação sobre os movimentos dos olhos, esta pesquisa buscou compreender as características e a aplicabilidade das tecnologias de interação baseadas na visão: o sistema do eye tracking. O que ocorre quando estamos diante de um estímulo textual? Como se dá o reconhecimento visual de letras, palavras e linhas? Como se dá a leitura? Compreender o que envolve esses processos é amplamente reconhecido como um importante aliado no entendimento do raciocínio visual em design.*

*vision; eye movement; eye tracking; design*

*Although “see” is an inherent act to human beings, we give little or no importance to it. We open and close, we move our eyes paying no attention to the mechanism by which the visual - reflecting the reality in which we operate - are recognized by our senses. Study the visual perception as the reception process of these environmental stimuli, takes us to a vast field of knowledge. Therefore, we focus on the study of visual recognition and detection and textual evaluation technology while reading. Through research on eye movements, this research sought to understand the features and applicability of vision-based interaction technologies: the eye tracking system. What happens when we face a textual stimulus? How is the visual recognition of letters, words and lines? How is reading? What is involved in these processes is widely recognized as an important ally to the understanding of visual thinking in design.*

### 1 A visão

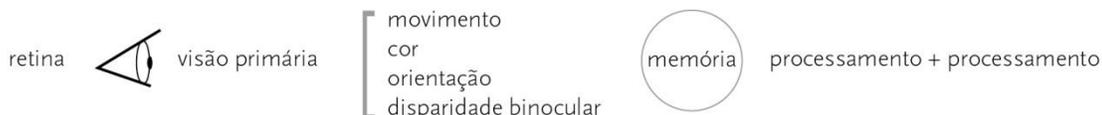
A percepção visual é uma das modalidades mais estudadas dos fenômenos psicológicos (Sternberg, 2010). Entende-se por percepção “o conjunto de processos pelos quais é possível reconhecer, organizar e entender as sensações provenientes dos estímulos ambientais” (Kosslyn, 1995; Sternberg, 2010). Embora “ver” seja um ato inerente ao ser humano, damos pouca ou nenhuma importância a ele. Abrimos e fechamos, movimentamos os olhos sem perceber o mecanismo por meio do qual os estímulos visuais – que traduzem a realidade em que estamos inseridos – são reconhecidos por nossos sentidos. O que ocorre quando estamos diante de um estímulo textual? Como se dá o reconhecimento visual de letras, palavras e linhas? Como se dá a leitura? A pesquisa buscou respostas para estas questões.

O que se conhece é que a visão ocorre por intermédio da luz. Ao abrirmos os olhos, a luz incide sobre sua camada protetora conhecida como córnea, passa pela pupila, pelo cristalino (membrana transparente que se localiza atrás da íris, responsável pela visão de objetos próximos ou distantes) e pelo humor vítreo (que envolve todo o olho). É um processo que resulta na refração dessa luz dentro do olho para a formação da imagem na retina. “A retina é o local em que ocorre a tradução da energia da luz eletromagnética para os impulsos neurais eletroquímicos” (Blake 2000, apud Sternberg, 2010).

É na retina que se encontram os fotorreceptores, responsáveis por converter a luz em energia, possibilitando a sua transmissão pelos neurônios para o cérebro. Gibson (1966, 1979), ao tratar da percepção, introduziu os conceitos de objeto distal, isto é, do mundo externo, meio informacional, estímulo proximal e objeto perceptual. Suponhamos que um determinado objeto do mundo externo, no qual a luz reflete (meio informacional), passe por meio de ondas, atingindo os receptores dos sentidos (os olhos) ou as células neurais especializadas na recepção desse tipo de informação sensorial. Logo que a informação entra em contato com os receptores sensoriais, ocorre a estimulação proximal e assim o objeto perceptual reflete as propriedades da realidade.

Mas qual a substância fundamental da visão? Segundo Edward Andelson e James Bergen (1991), o ponto central para o entendimento do primeiro passo para a extração da informação visual se encontra na visão primária. De uma forma genérica, a “visão primária” envolve uma quantidade de medidas numéricas das propriedades básicas da imagem que incluem a orientação, cor, movimento, entre outros. Vejamos abaixo o modelo que, assim, se tornou popular para humanos e máquinas:

Figura1: Diagrama genérico do processamento paralelo da informação visual baseado em Neisser, 1976. (LANDY; MOVSHON, 1991: 4).



Nesse modelo, o processamento envolve um conjunto de caminhos paralelos, com particularidades em cada uma das propriedades da visão. Nos modelos computacionais para visão, por exemplo, os pesquisadores têm demonstrado que uma quantidade baixa de medidas são necessárias para realizar as tarefas da visão. (MARR, 1982).

Um aspecto que me chamou bastante atenção, nos estudos sobre os modelos computacionais, foi a ideia da interferência conhecida como “noise” ou ruídos, nos modelos de performance visual de Graham, 1977 (*Computational Models of Visual Processing*, 1991: 147). Essas reações são resultados da inclusão de dados randômicos como teste para que sejam suscetíveis a variações. É sabido que esse ruído pode ocorrer em qualquer estágio do processamento visual humano. Desta forma, um bom sistema de engenharia seria suficiente para evitar que esse tipo de interferência ocorra.

Figura 2: Indicação de uso apenas de um sinal e antes da visão. PELLI, (1990).

“Algumas vezes esses ruídos são partes essenciais do modelo, como resumo de probabilidades, onde a ideia chave é a detecção randômica que possa ocorrer em algum ou muitos canais independentes” (Graham, 1977 apud Pelli, 1991). Segundo Pelli (*Computational Models of Visual Processing*, 1991: 148), há evidências de que a principal característica do ruído que ele considera limítrofes de detecção visual se encontra nas proximidades do estímulo. Há diferentes teorias relacionadas a detecção desses sinais. Em um dos mais importantes, Green e Swets (1974) utilizou um detector de apenas um sinal. A esse ruído, ele somou um simples número randômico. A partir desse experimento, Green e Swets obtiveram as respostas: “SIM, ele está lá” ou “NÃO, ele não está”.

## 2. Reconhecimento de padrões visuais

Segundo Farah (1992), o reconhecimento de padrões está dividido em dois sistemas distintos. Um dos sistemas está relacionado ao reconhecimento de partes do objeto e a reconstituição dessas partes em um todo distinto. Vejamos um caso em que um indivíduo pensa em partes de um objeto (páginas, capa, letras, assim por diante), e logo construirá a ideia de um livro. O outro sistema está relacionado ao reconhecimento de configurações maiores, como o reconhecimento de faces, por exemplo. As faces têm características bastantes complexas, e seu reconhecimento será muito mais difícil.

As teorias ascendentes (bottom-up) e as descendentes (top-down) são relacionadas ao processamento de características de nível elementar. A bottom-up são acionadas pelos estímulos visuais, ao passo que a top-down são conduzidas por processos cognitivos de nível mais elevado, por conhecimentos já existentes ou mesmo por expectativas anteriores à percepção (Clark, 2003, apud Sternberg, 2010). São quatro teorias bottom-up que estudam a percepção da forma: teorias do gabarito, do protótipo, das características e da descrição estrutural.

Teorias do padrão (ou gabarito) – Afirmam que “é possível reconhecer a letra A como A independentemente das variações no tamanho, na orientação e na forma nas quais a letra está escrita”. Nessa teoria acredita-se que possuímos em nossas mentes uma enorme quantidade armazenada de padrões (ou formas) que o reconhecimento se dá por comparação” (Selfridge, Neisser, 1960).

Teorias dos protótipos – Afirmam que “há uma média de uma classe de objetos ou padrões relacionados que integra todos os traços mais as características (mais freqüentemente observados) daquela classe”. Isso pode ser entendido como conjuntos de pontos, formas geométricas, uma letra H ou mesmo configurações aleatórias (Posner, 1967).

Teorias das características – Segundo essas teorias, “o indivíduo tenta estabelecer correspondência entre as características de um padrão e aquelas armazenadas na memória, ao invés de associar um padrão inteiro a um gabarito ou protótipo”(Stankiewicz, 2003).

Teoria da descrição estrutural – Essas teorias “consideram uma forma pela qual se possam formar representações mentais estáveis de objetos em 3-D, com base na manipulação de algumas formas geométricas simples”(Biederman, 1987).

Nas abordagens das teorias top-down, conhecidas como “percepção construtiva”, “quem percebe constrói uma representação cognitiva (percepção) do estímulo, usando informações como base para a estrutura, além de usar outras fontes de informação para construir a percepção” (STERNBERG, 2000). Nessa abordagem, alguns pesquisadores acreditam que “o

mundo não apenas afeta a percepção do indivíduo, como também o mundo experimentado é, na realidade, formado pela própria percepção (Goldstone, 2003).

### 3. O olho e o processo cognitivo

*“Orpheus chamou os olhos de ‘o olhar de vidro da natureza’; Hesichius, ‘as portas para que entrem a luz do sol’; Alexandre, o Peripatecian, ‘as janelas para a mente’, pois pelo olhos é que claramente vemos; nós mergulhamos num pensamento profundo, e entramos nas partes íntimas de sua câmara secreta”* (DU LAURENS, 1596, tradução em 1599: 19, *apud* WADE, 2005).

O mundo, na atualidade midiática, nos envolve entre milhões de estímulos perceptivos que ocorrem de forma simultânea. Desses milhares de estímulos, uma grande variedade de informações visuais, como cor, forma, textura, luz, movimento, chegam aos nossos sentidos, e, por meio de nossos olhos e cérebro, são captados, armazenados e decodificados, ou guardados em nossa memória.

O que realmente vemos? Como ocorre o processamento visual na mente humana? O sistema visual nos seres humanos possibilita aos indivíduos a assimilação da informação do ambiente. Segundo a Psicologia Cognitiva, os seres humanos captam a maior parte da informação visual durante fixações momentâneas, nos instantes em que o olho foca. Entre esses pontos, o olho se move rapidamente para que o cérebro registre a informação.

Sabe-se que o olho humano faz uma série de movimentos pequenos e rápidos conhecidos como saccades ou sacões, que, separados por intervalos de fixação, assimilam uma informação (RAYNER, 1998). O termo *saccade* foi criado por Émile Javal em 1880, diretor do laboratório de oftalmologia da Sorbonne, Paris. Esses movimentos são importantes no mecanismo da visão, pois, por meio deles, a imagem é fixada na retina. Foi Crum Brown (1895), entretanto, que notou a descontinuidade desses movimentos.

### 4. Movimento dos olhos

Nossos olhos estão em constante movimento, deslocando-se em vários sentidos. Cada olho é movimentado por seis músculos (GREGORY, 1966), responsáveis por movê-los ao buscar identificar os objetos (ou cena). Eles se deslocam em uma série de sacões pequenos e rápidos, a fim de dirigir o olhar para uma determinada posição. “Ao identificar a matriz visual, o olho fixa o objeto e “rastrea” toda a sua extensão, enviando, assim, a informação ao nosso cérebro” (GREGORY, 1966).

Desde o século II d.C. já se conhecia a musculatura extra-ocular que envolve nossos olhos (WADE, 2005). Essa descoberta feita por Claudius Galen foi referência para a pesquisa de muitos cientistas na área de anatomia ocular. Desde a antiguidade Babilônica a Egípcia, esses conhecimentos eram utilizado nos cuidados com pacientes e em cirurgias nessa área (Wade, 2005). Foi somente em 1879, entretanto, que Landolt descobriu a descontinuidade dos movimentos dos olhos. Essa descoberta foi publicada no manual de exame dos olhos naquele ano.

Apesar da grande distância entre os cientistas da antiguidade e a descoberta de Landolt, há mais de 2000 anos atrás Aristóteles já estudava os movimentos dos olhos. Naquele período, ele detinha conhecimentos sobre a binocularidade dos movimentos dos olhos, ou seja, que os olhos dirigem-se para a mesma direção e que há a convergência entre eles. “Isso demonstra que Aristóteles acreditava que os olhos operavam como uma unidade, ao invés de serem órgãos independentes” (WADE, 2005).

## 5. A leitura

Keith Rayner (1943 – 2015) foi um dos mais importantes psicologista cognitivo, professor e pesquisador da Universidade da Califórnia em San Diego e pioneiro nos estudos dos atuais dispositivos *eye tracking* e a leitura. Muito do que se conhece hoje sobre esse tema, se deve aos seus longos anos de pesquisa na área. É dele a constatação de que é durante o momento de fixação (que ocorre entre 200-300ms que a nova informação é assimilada pelo nosso sistema de processamento e a representação mental do significado da palavra é construída em nosso cérebro (RAYNER, 1978-1998). O estudo do movimento dos olhos apresenta pontos críticos exatamente por ser esse movimento o mediador dentro dessa complexa seqüência do processo cognitivo (Rayner, 1983). Em seu livro “*Eye movements in reading: perceptual and language processes*” Rayner deixa claro a ligação do movimento dos olhos e sua importância no processo de leitura. Ele acrescenta que fatores psicofísicos, tais como forma e tamanho de letras afetam a performance na leitura.

Um importante estudo proposto por Kevin Larson, apresentado no artigo “*The Science of Word Recognition*”, da Microsoft Corporation, argumenta que, segundo a psicologia cognitiva, para reconhecermos a palavra, nós utilizamos as letras entre a palavra, ou seja, o reconhecimento acontece pela somatória das letras, mais a forma que resulta de seu conjunto. Conhecido como “*Bouma Shape*” pelos teóricos da Psicologia moderna, esse tem sido postulado pelos designers tipográficos, para os quais a palavra é vista e reconhecida pelo contorno criado pela forma da palavra.

Kevin divide em três diferentes categorias o reconhecimento de palavras: por meio de sua forma, reconhecimento em série e reconhecimento em paralelo. O primeiro modelo diz respeito ao reconhecimento da palavra como uma unidade completa, ao invés de letra por letra. Alguns pesquisadores acreditam que esse reconhecimento é determinado pelos caracteres ascendentes e descendentes que se formam por meio de seu contorno como uma imagem única. Essa forma seria facilmente identificada pelos nossos olhos, uma vez que já teríamos visto a palavra tantas outras vezes (CATTEL, 1886, *apud* KEVIN, 2004).

O reconhecimento serial das palavras ocorre letra após letra da esquerda para a direita. Esse argumento já não encontra defensores na atualidade. É um modelo no qual a palavra seria reconhecida em seqüência, podendo ser facilmente deduzido que palavras curtas seriam mais facilmente reconhecida do que as palavras longas (KEVIN, 2004). O reconhecimento das letras em paralelo é o mais aceito pelos psicólogos por ser um modelo mais preciso. Nesse modelo, o reconhecimento das letras dentro da palavra ocorre simultaneamente, argumento que tem sido comprovado pelo desenvolvimento da tecnologia do *eye tracking*. Nele, a informação da letra é utilizada para o reconhecimento da palavra. Esse processo que, em tempo real, identifica o que ocorre durante a leitura não era possível em outros tempos.

“É sabido por mais de 100 anos que, quando lemos, nossos olhos não se movem sem problemas sobre a página, mas que fazem discretos saltos de uma palavra à outra. Nós fixamos a palavra por um período de tempo, 200-250mmms, então movendo-se para a outra palavra. Esses movimentos são conhecidos como sacões e usualmente duram 20-35ms. A maioria dos sacões são movimentos para frente entre 7 a 9 letras, no entanto, 10 a 15% desses sacões são regressivos ou para trás. A maioria dos leitores são completamente desavisados da freqüência dos sacões regressivos quando lêem. A localização das fixações não são randômicas. Fixações nunca ocorrem entre palavras, e usualmente ocorrem exatamente para a esquerda da metade da palavra. Nem todas as palavras são fixas; palavras curtas e particularmente palavras funcionais são freqüentemente ignoradas” (KEVIN 2004).

Um aspecto relevante na abordagem sobre o movimento dos olhos e a leitura, diz

respeito à mudança no tamanho do caractere e o que isso pode influir na dificuldade ou facilidade na sua percepção. No artigo “Tamanho da imagem na retina e o espaço perceptual na leitura”, Robert Morrison (1983) argumenta que há dois caminhos para realizar alterações idênticas na retina do tamanho da letra, diante de um estímulo visual: considerando-se a mudança física da medida do caractere do texto, e outra forma, alterando-se a distância do estímulo do observador. “Com a distância de visualização constante, caracteres maiores resultam em uma imagem da retina maior. Diminuindo a distância de visualização, o tamanho da imagem da retina também será menor” argumenta. Com essa constatação, Morrison (1983) chegou à conclusão de que a mudança no tamanho da letra ou a distância do seu observador, terão efeitos mínimos no espaço perceptivo.

Foram as investigações feitas por Miles Tinker, entretanto, que efetivamente contribuíram para o melhor entendimento do tema. Publicado no livro *Bases for effective reading* (1965), Tinker considerou tanto os efeitos dos fatores psicofísicos quanto dos tipográficos relacionados à leitura de textos. Essas investigações muito contribuíram para a compreensão sobre a velocidade na leitura, movimento dos olhos e o espaço perceptual quando há variações tipográficas, tais como tamanho, largura da linha e da entrelinha. Em seus experimentos, Tinker observou que em média, quando o tamanho do caractere era variado em um comprimento de linha definido em 19 paucas, ou 8 cm, para caracteres de tamanho 10pt e 11pt, o texto era lido mais rapidamente. Além disso, era observadas poucas fixações e longos saltos (MORRISON, 1983). Isso quer dizer que, para esse largura de linha, o leitor conseguirá ler o texto com maior velocidade, movimentando os olhos e fixando-se em poucas palavras, além de percorrer o olhar sobre a linha inteira, sem muitas dificuldades.

Ao compor o texto com letras em tamanhos pequenos, na mesma largura de linha, essas letras eram lidas de forma mais lenta, com uma média de fixação maior e com menores saltos. Tinker concluiu que “a duração da fixação foi mais longa nas letras de tamanho menores, pois os caracteres eram difíceis de serem percebidos”. Esses experimentos mostraram que a relação entre tamanho de letras e o comprimento da linha determinam a velocidade com que o texto pode ser lido.

Tinker descobriu também que a entrelinha tem papel importante, e que ela pode aumentar ou diminuir a velocidade na leitura. Textos compostos sem entrelinhas, ou seja, uma linha sobre a outra sem espaços, dificultam a legibilidade do texto. “Geralmente são necessários dois pontos de entrelinha para otimização da performance na maioria dos tamanhos dos tipos investigados” (TINKER, 1965). Ou seja, para um texto composto no tamanho 10pt, será ótima a entrelinha 12pt.

Foi Jacob (1990), em seu artigo *What You Look At Is What You Get* (WYLIWYG), o primeiro a introduzir a ideia do sistema inteligente de interação baseado na visão. Jacob quis demonstrar o uso do olhar ao invés do mecanismo apontador do mouse. No sistema apresentado nesse artigo, Jacob idealiza uma janela de texto que se abre para mostrar a informação no item selecionado visualmente ((DUCHOWSKI, 2007). Jacob foi o primeiro a utilizar essa tecnologia do *eye tracking* de forma interativa, que se tornou um paradigma, criando com ele um sistema de visão inteligente baseada na informação vista em displays.

Algumas das mais importantes iniciativas na área de acessibilidade, que tem valorizado os estudos com o uso do *eye tracking*, é a da associação dinamarquesa COGAIN (*Communication by Gaze Interaction*), fundada em 2004 e apoiada pela *European Commission*. A COGAIN ([www.cogain.org](http://www.cogain.org)) integra os mais avançados expertises na área de tecnologia da interface para beneficiar os usuários portadores de deficiências. Desta forma, facilita pesquisas ao disponibilizar conhecimentos, conteúdos e aplicativos como *eye communication systems* (Dasher, Gazetalk e UKO II), um conjunto de 3 diferentes softwares ([http://www.cogain.org/wiki/COGAIN\\_Applications](http://www.cogain.org/wiki/COGAIN_Applications)), além de tantos outros.

Há um aplicativo livre conhecido como *Eye typing*, que tem por base a imersão em VR

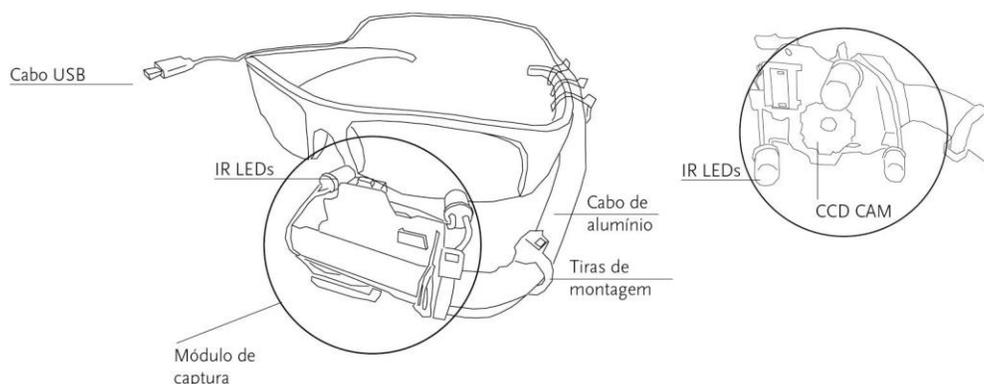
(realidade virtual). Nesse aplicativo, o olho é o apontador que substituiu as funções de digitador. Ele se apresenta como um dispositivo de interação contendo um teclado virtual visível na tela do computador (que pode também ser projetado). Com base no varrimento dos olhos, esse sistema detecta para qual letra o usuário está olhando, e assim ele aciona a letra por meio do clique no teclado virtual. Esse sistema foi desenvolvido pela empresa americana Prentke Romich Company.

Um outro é o *EyeDraw*, desenvolvido pelo laboratório de *Computer and Information Science* da Universidade de Oregon (<http://www.cs.uoregon.edu/Research/cm-hci/EyeDraw/>). *EyeDraw* é um aplicativo para desenho baseado no olhar, que oportuniza portadores de deficiência motora severa. Com uma plataforma simples para desenhos de linhas, círculos, formas etc., o aplicativo permite ao usuário clicar o botão, selecionar, desenhar com um ponto de início e fim, e, por último, salvar ou abrir novamente seus desenhos somente com os movimentos dos olhos.

## 6. A tecnologia do *Eye tracking*

*Eye tracking* é um dispositivo que possibilita mensurar a posição e o movimento dos olhos. Normalmente é utilizado para fins de pesquisa em avaliação do sistema visual. Muitas pesquisas têm sido feitas no campo da percepção visual e processamento cognitivo. Segundo R. L. Gregory (1997), a percepção não é simplesmente determinada pelos padrões de estímulo. Ao contrário, é uma busca dinâmica da melhor interpretação dos dados (sinais sensoriais) disponíveis. “Apesar da realidade dos movimentos dos olhos não estar disponível para o observador, ele pode ser útil para fornecer informações sobre o input no sistema visual, através do qual podemos inferir os processos envolvidos na visão e na percepção” (WADE, 2005).

Figura 3: *Eye tracking* construído durante a pesquisa pelo engenheiro Afonso Yamaguti (2012), com base no projeto TEMPT1 & EyeWriter ([http://ffff.at/eyewriter-or/info\[at\]eyewriter\[dot\]](http://ffff.at/eyewriter-or/info[at]eyewriter[dot])) e [eyewriter.org](http://eyewriter.org), disponível para download gratuito pelos autores.



O *eye tracking* é um equipamento com tecnologia avançada utilizado para medir o movimento dos olhos, durante o processo de obtenção ou identificação de elementos visuais de uma cena (ou objeto). São dois os tipos de monitoramento: um que mede a posição do olho em relação à cabeça, e o outro que mede a orientação dos olhos no espaço (YOUNG; SHENO, 1975 *apud* DUCHOWSKI, 2007). Afim de que tenhamos medidas acuradas em relação à detecção dos movimentos dos olhos, é necessário que a cabeça do indivíduo esteja fixa, e que os olhos estejam em uma posição relativa na qual a cabeça e o ponto de interesse coincidam. Somente desta forma será possível distinguir o movimento da cabeça da rotação dos olhos (DUCHOWSKI, 2007).

A atenção visual humana durante a leitura tem sido um dos mais difundidos campos de aplicabilidade dos estudos dos movimentos dos olhos que utiliza o *eye tracking*. Nos processos atencionais de usuários, esse dispositivo tem sido de grande utilidade para identificação de estímulos visuais, hierarquia de interesses em testes de usabilidade de interfaces interativas ou páginas da web. Essa tecnologia vem sendo aplicada nos campos da psicologia (psicofísica), marketing/publicidade, interface humano-computador e ergonomia.

Rayner sugere que são três as características dos movimentos dos olhos durante a leitura: 1. os movimentos dos olhos diferem de alguma maneira quando a leitura é silenciosa daquela em voz alta; 2. durante a leitura, a fixação dura aproximadamente entre 200-250ms e que a média da medida dos sacões é de 7 a 9 espaços de letras; 3. o movimento dos olhos é influenciado pelas variáveis textuais e tipográficas. Podemos dirigir nosso olhar para uma área de interesse específica de um texto e quando o fazemos, não nos damos conta da rotação de nossos olhos na identificação dessa informação. Para que possamos registrar, armazenar e memorizar esses movimentos, contamos hoje com uma tecnologia de alta precisão.

Com a tecnologia do *eye tracker*, o olhar de um usuário no computador, pode ser gravado em tempo real e os dados registrados e posteriormente analisados, *offline*. Assim, esses dados obtidos podem ser valiosos para conhecer o exato foco de interesse ou atenção do usuário. Dados sobre a fixação do olhar, duração, localização, performance (velocidade e acuidade) são usados para validar e refinar a percepção visual durante múltiplas tarefas apresentadas nas telas dos computadores. "... Movemos nossos olhos para trazer uma parte específica do campo da visão em alta resolução para que possamos ver em detalhes o que está na direção central do nosso olhar" (DUCHOWSKI, 2007). Por meio de uma câmera de infravermelho, posicionada em frente aos olhos, os movimentos são capturados enquanto ocorre um rastreamento da imagem em frente a um computador.

Um dos mais bem sucedidos projetos que utiliza o *eye tracking* é o projeto colaborativo TEMPT 1, desenvolvido por FAT & OpenFrameworks & Graffiti Research Lab, em Los Angeles CA. O projeto foi criado para atender um paciente do Alhambra Hospital (CA), chamado Tony Quan – um artista grafiteiro americano diagnosticado com uma doença degenerativa (ALS – Amyotrophic Lateral Sclerosis), que afeta as células nervosas no cérebro e espinha dorsal. Essa doença o deixou completamente paralisado. Somente o movimento de seus olhos se manteve em funcionamento.

Para que fosse possível a continuação de seu trabalho como grafiteiro, foi desenvolvido um aparato – *eye tracking* de baixo custo – acoplado a uma plataforma de projeção para ambientes externos, que, em tempo real, projeta as imagens do grafite criado por Tony por toda a cidade. O *software* livre EyeWriter foi desenvolvido para que pacientes como Tony tenham a oportunidade de, por meio de uma tecnologia criativa, manter-se produtivo, participativo apesar da extrema vulnerabilidade de sua condição motora.

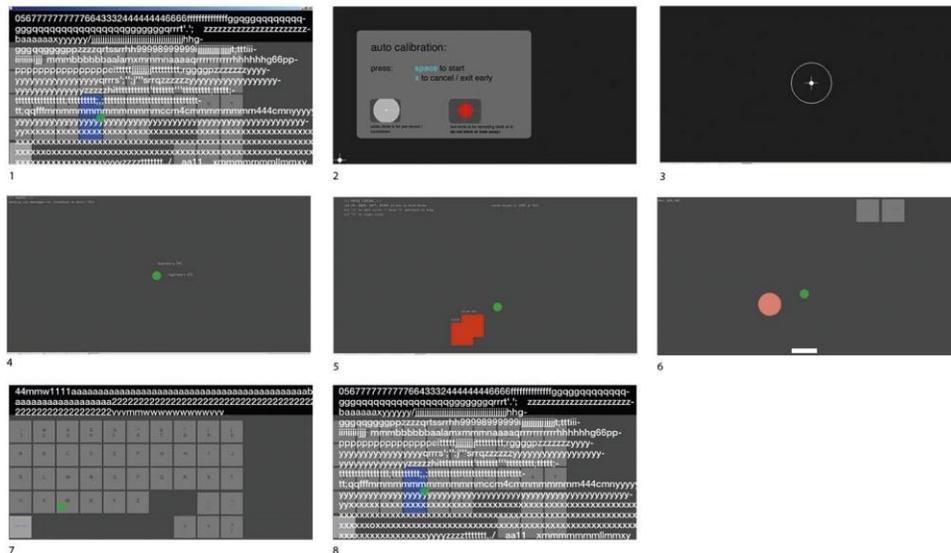
## 7. O sistema do *Eye tracking* e o aplicativo

Ao conectar o cabo usb da câmera do *eye tracking* ao computador, o aplicativo *Eyewriter* reconhece automaticamente o dispositivo. A partir desse primeiro reconhecimento, algumas instruções surgem para que o processo de captura efetivamente aconteça. O sistema busca analisar a recepção da imagem pela câmera do dispositivo para que seja possível capturar o foco de atenção do usuário. Isso só é possível por intermédio da reflexão da córnea pela luz infravermelha da câmara posicionada em frente ao olho do observador, quando o sistema tenta localizar o centro da pupila. Essa reflexão de uma fonte de luz infravermelha na córnea é conhecida por *Purkinje* (CRANE, 1994, apud DUCHOWSKI, 2007). Afim de evitar qualquer erro na localização do ponto central durante esse processo, a cabeça do usuário deve permanecer fixa, para que a posição do olho e o ponto de interesse coincidam.

Todo sistema de *eye tracker* baseado em vídeo necessita de calibragem para que o aplicativo calcule a área onde o usuário estará observando. Normalmente ele se compõe de uma seqüência de estímulos que aparecem em toda a extensão da tela do monitor. No *Eyewriter*, nove pontos de estímulos surgem nos pontos extremos, médios e central da tela. O observador deverá seguir, apenas com o olhar, cada um desses estímulos. Assim que a calibragem é feita, o usuário passa a interagir com o sistema.

A seqüência de imagens ao lado mostram o *eye tracking* sendo avaliado com o aplicativo *Eyewriter* (as imagens foram capturas pelo “*print screen*”). O aplicativo usa algoritmos baseados na imagem obtida pela luz infravermelha. “A webcam captura as imagens de forma similar ao olho humano” (KOWALIK, 2010). Na seqüência de imagens feitas pela autora, no quadro (1), vemos o olho do usuário em quatro diferentes configurações: o olho natural sem modificações; o olho sem o filtro da luz visível; o olho em negativo; o olho com o filtro de infravermelha. O que se observa no último quadro é a reflexão da córnea pela luz infravermelha, quando o sistema busca identificar a localização do centro da pupila (círculo rosa). Logo a seguir (2), vê-se um quadro de instrução com os comandos para dar início à calibragem.

Figura 4: Testes de funcionamento do *eye tracking* com uso do aplicativo *eye writer*.



Na calibragem (3), os estímulos visuais surgem nos pontos de coordenadas x e y. São 9 pontos brancos no total, sendo que uma cruz e o círculo maior identificam a fixação do olhar do usuário. Depois de feita a calibragem (4), o aplicativo mostra a localização do ponto exato do olhar (círculo verde) em coordenadas x e y. Ao avançarmos nas etapas do aplicativo (5), surge o jogo interativo. Nesse jogo, o ponto verde é o olho do usuário; o quadrado vermelho indica a interação com a palavra “*click me*”. O jogador deverá fixar o ponto verde no quadrado vermelho. Ao clicar em cima do quadrado com o olho, o objeto se move, indicando o resultado da interação no jogo. Na imagem (6), há outra opção de jogo interativo. Nele, o círculo rosa sobe e desce, de um lado a outro da tela. A barra inferior deverá ser movimentada com o olho (círculo verde) a fim de rebater o círculo rosa, como um bola de *ping-pong*.

Na imagem (7), encontramos o *Eye typing*. O teclado virtual que possibilita a comunicação (interação) do usuário. Por meio da análise do rastreamento do olhar, o sistema identifica a tecla, – ou letra -, para a qual o usuário está olhando. Assim que a identifica, a tecla é iluminada com a cor azul e ao mesmo tempo, a digitação das letras ocorrem. Pacientes

portadores de deficiências motoras graves, além de lesões na fala como Tony (TEMP1 - eyewriter.org), têm utilizado esse teclado virtual para digitarem seus textos apenas com a movimentação de seus olhos. Esse é um dos maiores paradigmas da acessibilidade proporcionada pelo desenvolvimento dessa tecnologia.

## 8. Considerações finais

Com o desenvolvimento da tecnologia do *eye tracking*, tem sido possível a substituição do mouse pelos olhos, criando um novo mecanismo de interação homem-computador. Essa ferramenta tem contribuído para o aprimoramento dos estudos em design ergonômico e de interface voltado para os usuários com limitação de movimentos. Essa tecnologia ainda é de custo muito elevado, o que dificulta as pesquisas nessa área. A alternativa proposta pelo projeto do *eye tracking* de baixo custo abre caminhos para novas experimentações com projetos de interação baseada nos olhos. Porém, é necessário ainda um maior aprofundamento de seus estudos quanto a sua eficácia e também em relação à segurança trazida a seus usuários.

Uma grande promessa para o futuro do *eye tracking* parece ser o que o Google Inc. recentemente lançou nos Estados Unidos. Trata-se do *Google Glass*. O *The Wall Street Journal* publicou um artigo assinado por Clint Boulton que informa sobre esse novo dispositivo. O novo lançamento no segmento de *eyewear* tem funções que incluem: acesso a internet, realidade aumentada, registro do movimento dos olhos, comando de voz, *smartphone*, enfim, tudo que uma grande novidade contém para os aficionados pelo consumo de novas tecnologias. Para Boulton, esse dispositivo será de muita utilidade para profissionais da saúde, por ser de baixo custo, com capacidades de benefício muito grandes. O foco de interesse do Google, no entanto, é o consumidor final. Por estar em fase de teste, a empresa ainda não tem claras as regras que regem o seu uso como dispositivo móvel.

Não há dúvida de que o *Google Glass* revolucionará a forma de comunicação. Esse dispositivo contém comando de voz e *touchpad* (como o *touchscreen*) localizado na armação do óculos que controla uma pequena câmera embutida, e além disso, uma lente que exhibe os dados em frente aos olhos. Pesquisadores estão entusiasmados, pois acreditam que armazenar fotos, vídeos em uma rotina de trabalho em hospitais, por exemplo, poderá reduzir o tempo de trabalho que antes era necessário, quando buscavam essas informações ao consultar o computador. Se todos esses dados podem ser acessados por comando de voz, certamente reduzirá o tempo de busca por essas informações. Há ainda controvérsias quanto a seu uso. Segundo a maioria dos médicos consultados nesse artigo, apesar da facilidade aparente do *Google Glass* em ver o raioX e MRIs (ressonância magnética) durante uma cirurgia, além de checar as informações sobre as condições do paciente, caso tenha alergias a alguma medicação, é preciso cautela quando se trata da privacidade e proteção das informações desses pacientes.

## Referências

- ANDERSON, John R. *Psicologia Cognitiva e Suas Implicações Experimentais*. Rio de Janeiro: LTC, 2004.
- ARNHEIM, Rudolf. *El Pensamiento Visual*. Argentina: Universitária de Buenos Aires, 1976.
- BERKELEY, George. *Tratados Sobre a Visão*. São Paulo: Editora Unicamp, 2010.
- BEST, John. *Cognitive Psychology*. 5. ed. New York: West Publishing, 1999.
- CHOMSKY, Noam. *Linguagem e Mente*. Brasília: Ed. Universidade de Brasília, 1998.
- COREN, Stanley; WARD, Lawrence; ENNS, James. *Sensation and Perception*. Harcourt Brace

- College Publishers. USA: New York, 1994.
- DRUCKER, Johanna. *The Visible Word: Experimental Typography and Modern Art, 1909 – 1923*. USA: The University of Chicago, 1994.
- DU LAURENS, A. *A discourse of the preservation of the sight: of melancholic diseases; of rheumes, and of old age*. London: The Shakespeare Association.
- DUCHOWSKI, Andrew T. *Eye tracking Methodology: Theory and Practice*. Second Edition. London: Springer, 2007.
- GREGORY, R. L. *Olho e Cérebro: Psicologia da Visão*. Rio de Janeiro: Zahar, 1977.
- HABER, R. Norman. *Information-Processing Approaches to Visual Perception*. USA: Holt, Rinehart and Winston, 1969.
- HABER, Ralph Norman. *Information-Processing Approaches to Visual Perception*. USA: Ed University of Rochester, 1969.
- JURY, David. *What is Typography?* UK: RotoVision, 2006.
- LANDY, Michael S. MOVSHON, J. Anthony. *Computational Models of Visual Processing*. Cambridge: MIT Press, 1991.
- LANGTON, G. Richard. *Olho e Cérebro: Psicologia da Visão*. Rio de Janeiro: Zahar, 1979.
- LAUREL, Brenda. *Design research: methods and perspectives*. USA: The MIT Press, 2003.
- LIDWELL, William. *Universal principles of Design*. USA: Rockport Publishers, Inc, 2003.
- LIMA, Manuel. *Visual complexity: mapping patterns of information*. USA: Princeton Architectural Press, 2011.
- LUPTON, Ellen. *A Escrita no Design Gráfico*. Porto Alegre Porto Alegre: Bookman; Artmed, 1996.
- MARR, David. *Vision*. Cambridge, USA: MIT Press, 1982.
- MORRISON, Robert E., RAYNER, K. *Eye Movements in Reading: Perceptual and Language Process*. New York, USA: Academic Press, 1983.
- RAYNER, Keith. *Eye Movements In Reading: Perceptual and Language Processes*. New York: Academic Press, 1983.
- ROGERS, Richard. *Digital Methods*. USA: MIT Press, 2013.
- SCHIFFMAN, Harvey Richard. *Sensation and Perception: an integrated approach*. New York, USA: John Wiley & Sons, 2001.
- STERNBERG, Robert J. *Psicologia Cognitiva*. São Paulo: CENGAGE Learning, 2010.
- TINKER, Miles. *Bases for effective reading*. Minneapolis: University of Minnesota Press, 1965.
- TUFTE, Edward Rolf. *Envisioning Information*. USA: Graphics Press, 1999.
- UNGER, Gerard. *While You're Reading*. New York: Mark Batty Publisher, 2007.
- WADE, Nicholas J. TATLER, Benjamin. *The Moving Tablet of The Eye: The Origins of Modern Eye Movement Research*. New York: Oxford University, 2005.
- WARE, Cole. *Visual Thinking: for Design*. USA: Morgan Kaufmann Series in Interactive Technologies, 2008.
- WILLIAM, James. *The Principles of Psychology*. New York: H. Holt and Co., 1890.

## Artigos

- EDWARDS, Greg. *The Usefulness of Eyetracking*. Center for the Study of Language and Information. USA: Stanford University, 1999.
- EVREINOVA, Tatiana; EVREINOV, Greg. *Pseudo - Graphic Typeface: Design and Evaluation*. Department of Computer and Information Sciences. Finland: University of Tampere, 2005.
- EVREINOVA, Tatiana. *Alternative Visualization of Textual Information for People with Sensory Impairment*. Finland: University of Tampere, 2005.
- GERO, John. *Learning while optimizing a design: a situated agent-based design interaction tool*. Australia: University of Sydney.
- GREEN, Christopher D. (1996). *Where Did the Word "Cognitive" Come From Anyway?* Department of Psychology. Canada: York University, 1996.
- JACOB, Robert J. K. *What you look at is what you get: Eye movement-based interaction techniques*. Naval Research Laboratory. Washington: 1990.
- KOWALIK, Michal. *How to build low cost eye tracking glasses for head mounted . system*. Poland: Faculty of Computer Science and Information Technology, 2010.
- LARSON, Kevin. *The Science of Word Recognition*. Advanced Reading Technology, Microsoft Corporation, 2004.
- LOGAN, Gordon D. *Toward an Instance Theory of Automatization*. University of Illinois, [s. d.].
- RAYNER, Keith. *Eye movements during reading, scene perception, visual search, and while looking at print advertisements*. USA: University of Massachusetts, [s. d.].
- RAYNER, Keith. *Reading Disappearing Text: Cognitive Control of Eye Movements*. France: Université René Descartes Paris 5, 2003.
- REDLINE, Cleo D.; LANKFORD, Christopher P. *Eye-Movement Analysis: A New Tool for Evaluating the Design of Visually Administered Instruments (Paper and Web)*. Montreal: University of Virginia, 2001.
- SIMON, Herbert. *What is an "Explanation" of Behavior?* In: Paul Thogard (Ed.), *Mind Readings*, Massachusetts: MIT, p. 1-28.
- SMITH & KOSSLYN. *How the Brain Gives Rise to the Mind*. Chapter 3, 2008.
- SMITH & KOSSLYN. *How the Brain Gives Rise to the Mind*. Representation and Knowledge in Long-Term Memory, Chapter 4, 2008.
- WOOD, N. and Cowan, Nelson. *The Cocktail Party Phenomenon Revisited: Attention and Memory in the Classic Selective Listening Procedure of Cherry*. Missouri: University of Missouri, 1953.

## Sobre os autores

Celia Matsunaga, PhD, Universidade de Brasília, Brasil <celiamatsunaga@gmail.com>