

LA RÉOLUTION DE L'ATTENTION : LE GRAIN SPATIAL ET TEMPOREL DE LA CONSCIENCE VISUELLE

PATRICK CAVANAGH*‡, and RUFIN VANRULLEN§*

*Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA

‡LPP, Université de Paris 5, 45 rue des Sts Pères, 75270 Paris Cedex 06

§CerCo, Université Toulouse 3, Faculté de Médecine Rangueil, 31062 Toulouse Cedex.

RESUME

L'attention sert de portail à la conscience visuelle et elle impose la limite ultime de ce que nous pouvons consciemment percevoir dans le monde visuel. Notre travail démontre que l'attention n'est pas limitée à un foyer unique mais peut avoir une distribution spatiale et temporelle complexe, ayant l'aspect d'une image spatiale dynamique. Nous avons évalué le grain spatial et temporel de cette image et découvert que les plus petites régions et les plus brèves durées qui peuvent être isolées par l'attention sont étonnamment grossières. Les objets plus proches que la limite spatiale de l'attention ne peuvent pas être isolés pour des traitements ultérieurs et ne peuvent être perçus qu'en tant que texture indivisible. De même, des événements se déroulant plus rapidement que la limite temporelle de l'attention (environ 14 événements à la seconde) sont éprouvés comme un flux temporel indivisible. Ces limites indiquent que l'accès conscient aux événements et aux endroits est quantique. Quand des centres d'intérêt multiples doivent être surveillés, leur accès, en plus d'être quantique, peut également être séquentiel.

1. Introduction

Nous avons tous eu l'expérience de « prêter » attention à une pièce de musique, un livre, ou une pensée et de perdre conscience d'autres événements environnant. Cette absence de conscience d'autres événements en dehors de notre foyer d'attention a des conséquences dramatiques et, occasionnellement, dangereuses sur notre capacité à négocier des situations menaçantes. Comment pouvons-nous traiter les dangers inhérents à notre cécité aux événements qui échappent à l'attention ? Nous sommes souvent testés pour nos capacités visuelles et, dans la mesure du possible, on nous prescrit des lentilles ou verres correcteurs pour compenser nos déficits visuels. La qualité de notre vision corrigée est le critère standard pour obtenir un permis d'opération des véhicules et d'autres machines lourdes. Cependant, même si nous voyons tout à fait bien, nos réactions dépendent de ce que nous enregistrons consciemment. Certains ont des capacités attentionnelles fortement développées (par exemple: les athlètes professionnels ou les pilotes d'avions) tandis que d'autres sont très contrariés par des environnements modérément dynamiques ou encombrés. Malheureusement, quoique nos qualités attentionnelles soient presque plus importantes que notre acuité visuelle pour déterminer nos réactions dans des situations dangereuses, nous n'avons aucun test ou critère légal et peu de littérature établissant les limites de l'attention. En outre, peu d'entre nous se rendent même compte de ces limites attentionnelles - nous croyons que si notre vision est correctement corrigée, et si nous avons reçu une formation adéquate, alors nous pouvons être un aussi bon conducteur ou pilote que n'importe qui. C'est cependant loin d'être le cas ; il est extrêmement important de développer des tests normalisés pour les capacités attentionnelles, afin d'identifier les individus déficitaire et de déployer des méthodes éducatives pour hausser ces capacités à des niveaux acceptables. Dans cette courte revue, nous faisons un premier pas dans cette direction en posant une question spécifique : quelles sont les limites de notre attention dans l'espace et le temps ?

Plusieurs auteurs ont utilisé une métaphore de « faisceau lumineux » ou « spotlight » (Eriksen et Hoffman, 1972 ; Jonides, 1980 ; Jonides, 1983) pour donner l'idée que l'attention peut être limitée dans l'espace à une petite région (l'attention illumine seulement une partie restreinte de notre espace visuel), et que ce faisceau de l'attention peut se déplacer au cours du temps. Quelle est la taille de la région de sélection ? Certains auteurs ont proposé que cette région de sélection puisse s'agrandir ou se rétrécir selon les besoins de la tâche (par exemple, le concept de zoom oculaire ou « zoomlens », Eriksen et St.-James, 1986 ; Klein et McCormick, 1989 ; LaBerge, 1983). Cependant, il est évident qu'il existe une taille minimum pour la région de sélection, et que cette taille minimum doit être plus grande que les plus petits détails que nous sommes capables de percevoir. Dans une étude du dix-neuvième siècle, Landolt (1891) a demandé que les observateurs comptent des lignes fines très rapprochées, arrangées directement autour du point de fixation du regard. Il a rapporté qu'« on arrive à un point où l'on ne peut plus les compter d'aucune façon, alors qu'ils demeurent encore parfaitement et distinctement visible » (P. 385). Les résultats de Landolt suggèrent que l'espacement le plus fin qui permet l'accès de l'attention aux éléments dans la fovéa (afin de les compter) est environ 5 fois plus grossier que l'espacement le plus fin qui peut être résolu visuellement, environ 1 arc minute (Campbell et Gubisch, 1968). Nous avons vérifié et avancé ces évaluations initiales pour montrer que la résolution de l'attention diminue, de 5 fois plus basse que l'acuité visuelle dans la fovéa, à 30 fois plus mauvaise à 15 degrés d'excentricité (Intriligator et Cavanagh, 2001) et que cette chute dramatique continue jusqu'aux limites du champ visuel. Si notre vision était aussi limitée que l'acuité de notre attention, nous serions légalement aveugles.

La recherche récente nous force à réviser la notion classique d'un foyer unique de l'attention, indiquant en fait que l'on peut porter l'attention à plus d'une cible simultanément. En particulier, les études utilisant le paradigme de poursuite d'objets multiples ou « multi-object tracking » (Pylyshyn et Storm, 1988 ; Yantis, 1992), de l'interférence de voisinage ou « flanker interference » (Driver et Baylis, 1989), et de la sélection de cibles multiples (Awh et Pashler, 2000 ; Castiello et Umiltà, 1992 ; Kramer et Hahn, 1995) démontrent que plus d'une région peut être sélectionnée à la fois, et que ceci n'est pas accompli simplement avec une seule région de sélection agrandie qui engloberait les multiples cibles. Awh et Pashler (2000), par exemple, prouvent que l'avantage de l'attention portée sur deux cibles séparées n'est pas observé pour un objet

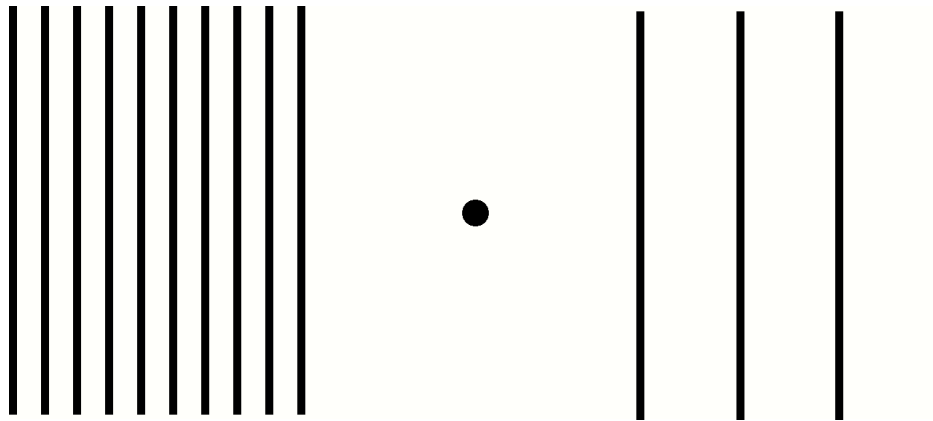


Figure 1. Une démonstration simple qui montre la différence entre l'acuité de la vision et celle de l'attention. En fixant le point central, les lignes à gauche ou à droite sont clairement visibles - elles sont minces, verticales, parallèles, également écartées, noires, et toutes d'une hauteur à peu près identique. Cependant, tout en fixant toujours le centre, il est difficile ou impossible d'isoler une ligne individuelle au milieu du groupe du côté gauche, par exemple, la quatrième. En conséquence, il s'avère impossible de compter les lignes du côté gauche (tout en fixant le point central), en dépit de leur visibilité. Les quatre lignes du côté droit sont plus espacées que la limite de la résolution attentionnelle, et elles peuvent donc être isolées une par une et comptées.

inattendu, présenté entre les deux cibles. Ces multiples régions d'attention forment une image spatiale - une « image d'attention » - et nous pouvons examiner la résolution spatiale de cette image : est-ce que cette résolution de l'attention est déterminée par la capacité d'isoler chaque cible des autres objets qui l'entourent, ou est-elle encore plus restreinte que cela ?

Les limites d'acuité pour l'attention temporelle ressemblent à celles de l'attention spatiale. En particulier, la sélection et l'individuation des événements dans une séquence très rapide s'avèrent fort grossières. Une lumière clignotant à une cadence lente peut être perçue comme une succession de phases lumineuses et sombres, mais quand le rythme est plus élevé, seul le clignotement est perçu. A des cadences encore plus élevées, les phases lumineuses et sombres fusionnent et le clignotement n'est plus visible. La gamme des fréquences pour lesquelles le clignotement est évident mais les phases individuelles ne sont plus accessibles est équivalente à l'exemple de la résolution spatiale de l'attention (Figure 1), où les barres peuvent être clairement vues mais ne peuvent pas être comptées. Plusieurs phénomènes qui dépendent de l'attention temporelle pâtissent de cette limite temporelle très basse d'environ 4 à 8 hertz (voir Verstraten, Cavanagh, et LaBianca, 2000 pour un résumé). Ce taux est équivalent à 8 à 16 scènes à la seconde, compatible avec les fréquences maximales auxquelles des scènes ou des mots peuvent être identifiés dans des expériences visuelles avec présentation séquentielle rapide ou « RSVP » (Potter, 1993).

Finalement, nous passons en revue l'évidence que l'attention est quantique dans l'espace et le temps. Nous montrons que, quand des cibles multiples doivent être surveillées, elles peuvent être sélectionnées par des foyers attentionnels indépendants et non contigus, plutôt que par un seul foyer agrandi. Qui plus est, l'évidence suggère que ces différents foyers attentionnels sont en fait desservis par un seul faisceau qui échantillonne chaque endroit à une cadence régulière.

L'idée d'une limite spatiale et temporelle de l'attention, et de sa nature discrète et quantique, impose des critères importants pour l'établissement de tests normalisés pour les opérateurs humains d'équipement dangereux, et pour la conception des interfaces homme-machine que ces opérateurs doivent employer.

2. Le “grain” spatial de l'attention

Puisque nous allons comparer la résolution attentionnelle à l'acuité visuelle, commençons par une description simple de la résolution visuelle : l'espacement le plus fin qui peut être perçu entre des détails visuels. La résolution visuelle est mesurée, de manière conventionnelle, comme le réseau sinusoïdal le plus fin qui peut être vu à un contraste maximum de 100% (Campbell & Gubisch, 1966). Un observateur normal peut percevoir jusqu'à 55 cycles par degré (55 alternances de lignes noires et blanches) sous une lumière claire, mais « percevoir » dans ce contexte signifie simplement que l'observateur peut distinguer le réseau d'un champ uniforme (une tâche de détection), ou peut deviner son orientation (une tâche de discrimination). Cette limite d'acuité diminue de façon dramatique si l'on déplace le stimulus loin de l'axe du regard.

On peut définir la résolution attentionnelle d'une manière similaire. Quel est l'espacement le plus fin des barres d'un réseau (à contraste maximal) qui permet encore à l'observateur de distinguer chaque barre ? Un simple test de cette capacité à indexer chaque barre consiste à essayer de les compter. La Figure 1 démontre la différence entre l'acuité visuelle conventionnelle, et celle de l'attention. On peut clairement voir le réseau et rapporter qu'il est composé de plusieurs barres fines orientées verticalement. Il est bien plus difficile, tout en fixant du regard la croix centrale, de distinguer et de compter les barres du réseau. On perçoit ce réseau comme une texture, mais on ne peut accéder individuellement à ses éléments.

3. Etudes de “crowding”

Pour le réseau mentionné plus haut, le coût de notre incapacité à en isoler les éléments est relativement faible – on ne sait pas combien de barres sont présentes mais on sait qu'elles sont toutes verticales et effilées. Si chaque barre était différente, par contre, on ne serait peut-être plus capable de rapporter les propriétés d'une

barre individuelle. Cet effet est exploré dans la littérature considérable sur le phénomène de « crowding » (que l'on pourrait traduire par « cohue » ou « pêle-mêle »), typiquement testé avec des lettres présentées en ligne, en tant que mot ou en tant que simple chaîne de caractères (Bouma, 1970; Bouma, 1973). Ces études ont trouvé que les lettres au milieu de la chaîne sont plus difficiles à rapporter que celles qui sont plus proches du point de fixation, et de manière surprenante, aussi plus difficiles à rapporter que les lettres placées le plus à l'extérieur, même avec des temps de présentation prolongés (Townsend, Taylor, & Brown, 1971). En général, si une cible est plongée au milieu de multiples distracteurs, elle devient très difficile à percevoir.

L'explication traditionnelle du crowding est que les interactions latérales suppriment ou masquent les propriétés visuelles de la cible relativement tôt dans la séquence des traitements visuels. Cette idée a changé lorsque He, Cavanagh et Intriligator (1996) ont démontré que des réseaux orientés subissant ce crowding pouvaient néanmoins induire un *aftereffect* spécifique à leur orientation, alors même que les observateurs étaient incapables de rapporter cette orientation. Les stimuli étaient donc clairement enregistrés au moins jusqu'au niveau où émerge l'analyse de l'orientation, soit le cortex visuel primaire V1 (voir encart 1). Encore plus récemment, Parkes et al. (2001) ont montré que l'orientation d'une cible dans un arrangement « pêle-mêle » de multiples réseaux de type Gabor ne pouvait pas être rapportée mais que, par contre, l'estimation de l'orientation moyenne de ces réseaux sur la globalité de l'arrangement était directement influencée par l'orientation de la cible. Les auteurs concluent que, bien que les identités individuelles dans la cohue soient bloquées de l'accès conscient, elles peuvent néanmoins se faufiler jusqu'à des niveaux de traitement assez élevés, sous la forme d'une contribution à la texture globale de l'arrangement.

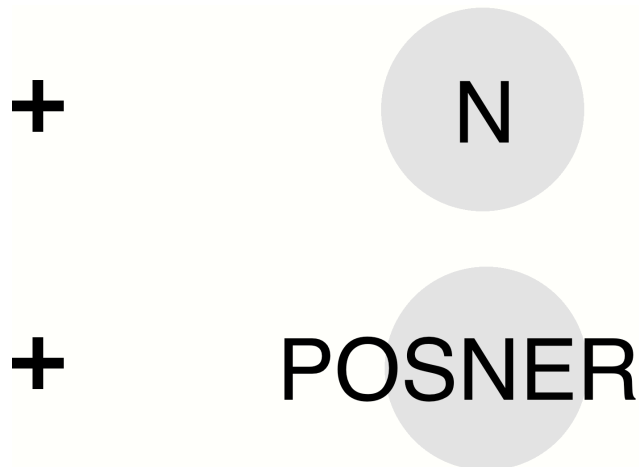


Figure 2. Tout en fixant du regard la croix en haut à gauche, la lettre N à droite peut être aisément identifiée. En fixant la croix juste en dessous, le N semble noyé dans un agrégat de lettres. C'est l'effet de cohue ou « crowding » induit par les éléments alentours, qui empêchent l'accès à l'élément individuel. Notons qu'il ne s'agit pas seulement d'une contrainte d'excentricité puisque la lettre R, encore plus périphérique, est facilement identifiée. Le N n'est en fait pas vraiment perdu, mais peut contribuer efficacement à des jugements de texture globale (Parkes et al, 2001). La région circulaire grisée représente une hypothétique région minimale de sélection à l'excentricité correspondant à la lettre N. Elle permet de sélectionner la lettre N individuellement lorsqu'elle est isolée, mais lorsqu'elle est inscrite dans une chaîne, la même région de sélection ne distingue plus une lettre unique mais plusieurs. Si la sélection était comme un faisceau lumineux permettant de scruter à l'intérieur de la région d'intérêt, le N serait accessible tout comme le S et le E. Ce n'est pas ce qui se produit. Il semble que nous n'ayons accès qu'à une description de « texture » à l'intérieur de la région de sélection.

Les résultats suggèrent donc que les propriétés d'un objet "pêle-mêle" ne sont pas supprimées ou masquées à un niveau précoce. Qu'est-ce qui explique, dans ce cas, notre incapacité à rapporter ces propriétés? Nous avons proposé (He, et al., 1996; Intriligator & Cavanagh, 2001) que cet effet de cohue est dû aux limites qui pèsent sur la résolution spatiale de l'attention. Pour un crowding sévère comme celui de la Figure 2, une cible noyée dans un arrangement dense de distracteurs ne peut pas être sélectionnée isolément, et ne peut donc pas être consciemment identifiée ou rapportée. La cible est perçue uniquement en tant que partie d'un groupe obstrué et indivisible qui conjugue tous les éléments au sein de la région de sélection.

Une deuxième expérience (Intriligator & Cavanagh, 2001) a démontré que le grain de la sélection n'est pas homogène sur le champ visuel. La tâche des sujets consistait à se déplacer mentalement parmi un ensemble de disques arrangés autour d'un cercle, tout en fixant le centre du cercle. Le déplacement avait pour origine l'un des disques qui clignotait brièvement en rouge ; puis les sujets avançaient ou reculaient mentalement autour du cercle en suivant les instructions de l'ordinateur. Après 6 ou 7 déplacements, l'un des disques s'allumait en rouge, et les sujets répondaient si c'était bien celui sur lequel leurs déplacements s'étaient arrêtés. En variant le nombre de disques placés sur la circonférence, nous avons mesuré la densité à partir de laquelle les sujets ne pouvaient plus distinguer les disques individuellement. Les résultats montrent que l'acuité de sélection se gâte rapidement lorsque l'excentricité augmente, et qu'elle est plus grossière dans le champ visuel supérieur qu'inférieur. De plus, en dehors de la fovéa, l'acuité est plus sommaire dans la direction radiale –le long d'un rayon allant de la fovéa à l'endroit considéré– que dans la direction tangentielle. Il s'agissait des premières mesures de la résolution de la sélection attentionnelle qui adressaient purement la sélection spatiale sans impliquer d'identification de la cible. La tâche de déplacement ou « stepping » escortait le foyer attentionnel au sein d'un arrangement dense d'éléments identiques, et la performance n'était donc pas affectée par des interactions préattentives entre éléments qui auraient pu dégrader les mesures de reconnaissance utilisées classiquement dans les conditions de crowding ou de flanker. Pourtant, les résultats en termes d'accès à la position spatiale correspondaient précisément à ceux obtenus pour l'identification de lettre dans les tâches de crowding standard (e.g., Toet & Levi, 1992). Ceci suggère que les erreurs d'identification dans les conditions de crowding peuvent être attribuées à une incapacité de sélectionner l'élément à la position visée. Le facteur déterminant est donc la perte d'accès à la position, et non pas la perte de l'information des propriétés visuelles à cette position.

Encore plus frappant que l'austérité de l'attention visuelle est le déclin rapide de la résolution attentionnelle lorsqu'augmente l'excentricité. A 15 degrés d'excentricité, par exemple, l'espacement critique pour l'attention est approximativement 3 degrés d'angle visuel, équivalent à 0.33 cycles par degré. A la même excentricité, l'acuité pour des réseaux est autour de 10 cycles par degré (Johnston, 1987; Thibos, Walsh, & Cheney, 1987; Virsu, et al., 1987), soit 30 fois plus précise. C'est l'une des raisons pour lesquelles la démonstration de la Figure 1 est si dramatique : à une distance de lecture normale, le milieu du groupe de barres verticales est à peu près à 15 degrés d'excentricité. Par conséquent, l'espacement des barres est bien au-dessus de la limite d'acuité, mais bien en deçà du seuil d'individuation. La différence entre ces deux seuils est bien plus large pour cette démonstration en périphérie (résolution attentionnelle 30 fois inférieure à la résolution visuelle) qu'elle ne le serait à la fovéa (5 fois inférieure).

Bien que la dissociation entre la vision et l'attention puisse paraître contre-intuitive, il existe d'autres exemples de stimuli qui peuvent être distingués par une partie du système visuel, mais pas par une autre. Le travail de He et collaborateurs (He, et al., 1996 – voir Encart 1; He & MacLeod, 2001) a démontré l'existence d'une adaptation à l'orientation à la suite d'une exposition prolongée à des réseaux dont les barres sont si proches que leur orientation ne peut être consciemment rapportée. Pourquoi la résolution de la vision spatiale serait-elle plus fine que le plus petit détail que l'on peut distinguer individuellement? Ou, pour simplifier encore, pourquoi le traitement visuel précoce encode-t-il des détails qui sont inaccessibles à l'attention? Notons d'abord que la limite de la résolution attentionnelle ne fait pas référence à la taille des détails encodés, mais à

l'espace *entre* ces détails. Un élément de la plus petite taille visible est aisément isolé et perçu consciemment s'il est seul au milieu du champ visuel. La résolution attentionnelle ne limite la performance que dans les cas où de multiples éléments sont espacés à une distance moindre que l'espace minimum requis pour leur position spatiale et configuration particulières. Considérons la nature des détails qui sont plus rapprochés que cette limite mais qui néanmoins peuvent être enregistrés par le système visuel. La plupart de ces détails servent probablement en tant que texture de surface, qui favoriserait la classification d'objets. Il est peut-être rare qu'un élément parmi une cohue d'objets semblables ait besoin d'être le sujet d'un intérêt ou d'une poursuite intenses. L'attention n'a pas dû évoluer pour négocier l'individuation et la poursuite de cibles très petites, noyées parmi des objets similaires. Pourtant, dans des cas où l'on doit porter l'attention sur un élément unique d'une scène encombrée d'objets, notre performance visuelle est ultimement limitée par la résolution étonnamment grossière de l'attention. Elle est étonnante non seulement par sa médiocrité, mais aussi par la façon dont nous pouvons ignorer tout de cette imperfection. Notre habileté à sélectionner des détails relativement fins dans des environnements dégagés nous donne l'impression de posséder des compétences visuelles irréprochables en périphérie. Les données montrent cependant que cette confiance n'a pas lieu d'être.

4. Résolution spatiale pour des cibles multiples

Nous avons utilisé l'abondante littérature sur l'effet de crowding (ainsi que les tâches de déplacement mental, Intriligator & Cavanagh, 2001) afin de démontrer que la résolution pour la sélection d'une cible parmi un ensemble de distracteurs qui l'entourent est extrêmement réduite par comparaison à la résolution visuelle. Quand il y a plusieurs cibles, pourtant, les études convergent pour indiquer que la résolution serait encore plus dégradée. La limite pour sélectionner une cible et supprimer ses distracteurs adjacents semble être déterminée par la plus petite taille disponible pour la région de sélection. Si un distracteur et une cible se trouvent tous deux dans la région de sélection, aucun d'eux ne peut être identifié, et le résultat est une description combinée des deux éléments sous forme d'une texture. Si la cible se situe dans la région de sélection, et le distracteur en dehors de cette région, dans un « pourtour suppressif », la sélection peut s'opérer convenablement.

L'idée que l'attention aurait une région centrale de sélection et un pourtour de suppression (Tsotsos et al, 1995; Mounts, 2003) suggère que lorsque deux cibles doivent être sélectionnées, leurs pourtours suppressifs respectifs produiront une interférence mutuelle si elles sont trop proches l'une de l'autre. La preuve de cette interférence entre cibles proches est limpide. Bahcall et Kowler (1999) ont montré que lorsque 2 cibles doivent être rapportées parmi un arrangement circulaire de lettres, la performance diminue quand les cibles se rapprochent, avec un minimum pour des cibles adjacentes. Mounts (2003) a prouvé que si l'attention était portée sur une position spécifique, la performance pour les positions adjacentes était moindre qu'aux positions plus éloignées. Dans des tâches de poursuite d'objets multiples (« Multiple Object Tracking » ou « MOT » ; voir Encart 2) où la tâche est divisée entre deux régions distinctes, la performance se dégrade quand ces régions se rapprochent (Carlson, Alvarez, & Cavanagh, 2007). Cette interférence de cible-à-cible est réduite si les deux régions sont séparées par le méridien vertical (Alvarez & Cavanagh, 2005) ou le méridien horizontal (Carlson et al, 2007).

Une mesure de la résolution spatiale de l'attention dans un paradigme de crowding qui varie l'espace entre une cible unique et un certain nombre de distracteurs passera tout simplement à côté de cette limite supplémentaire due à l'interférence cible-à-cible (liée aux pourtours suppressifs de chaque foyer attentionnel). En d'autres termes, aussi désastreuse que soit l'estimation fournie par les méthodes de crowding, une mesure plus réaliste de la résolution attentionnelle, à l'aide de cibles multiples, s'avèrerait probablement encore pire. Pour l'instant, nous ne disposons que d'estimations initiales sur ces contraintes additionnelles pour l'attention spatiale.

5. Résolution de l'attention temporelle

En éclairage intensif, le clignotement peut être perçu à des fréquences allant jusqu'à 50Hz, ce qui suggère que l'entrée du système visuel est capable de résoudre des changements rapides. Par comparaison, l'attention semble avoir une résolution temporelle bien moins précise. L'idée d'un temps d'escalade attentionnelle ou « attentional dwell time », le temps minimum requis par l'attention pour achever une opération (e.g., Duncan, Ward, & Shapiro, 1994), a derrière elle une longue histoire, et a donné des résultats très divers, de très brefs (40 msec dans des tâches de recherche visuelle) à beaucoup plus longs (200 à 500 ms pour le « clignotement attentionnel » ou « attentional blink », Raymond, Shapiro, & Arnell, 1992). Nous nous sommes concentrés sur la résolution temporelle de l'attention en tant que la fréquence maximale qui permet encore à chaque événement d'une séquence rapide d'être perçu et identifié individuellement, plutôt que d'apparaître comme un flou indistinct. Nous considérons ces tâches comme l'équivalent temporel des paradigmes de crowding spatial, dans le sens que la variation de fréquence rapproche les éléments dans le temps, plutôt que dans l'espace. Nous obtenons systématiquement une valeur de 120 à 150 msec par élément ou 7 à 8 Hz en termes de la cadence à laquelle se répètent les éléments. Au-delà de cette fréquence, la séquence apparaît comme un clignotement indifférencié (Verstraten et al, 2000; He et al, 1998; Battelli et al, 2001, 2003). Le nombre maximal d'événements temporels qui peuvent être énumérés (i.e. dissociés individuellement) atteint une limite autour de 10-12 événements par seconde (White & Harter, 1969). Ce plafond est le même si l'information est présentée alternativement à chaque œil ; la limitation n'est donc pas due à des contraintes périphériques, mais plutôt à la résolution de l'attention (White, Cheatham & Armington, 1953; Andrews, White, Binder & Purves, 1996). Lorsque le stimulus (composé de plusieurs événements visuels distincts à des positions séparées spatialement, comme par exemples 8 flashes lumineux) se répètent en cycles d'une période inférieure à 125 msec, tous les éléments d'un cycle sont rapportés comme ayant eu lieu simultanément, quels que soient les délais spécifiques entre ces éléments (Lichtenstein, 1961). Un phénomène appelé *Gestalt Fusion* démontre aussi une limite autour de 7 Hz (voir van de Grind, Grüsser, & Lunkenheimer, 1973; Grüsser & Landis, 1991). Dans ce cas, une lumière unique s'allume et s'éteint de manière répétée. En dessous de 5-7Hz, la lumière semble alterner entre l'état allumé et l'état éteint. A des fréquences plus élevées, la lumière semble ne jamais s'éteindre, mais reste allumée et clignote. Plusieurs autres phénomènes faisant appel à une sélection ou poursuite d'événements séquentiels montrent également une limite temporelle faible de 4 à 8 Hz. Par exemple, le mouvement apparent à de larges distances entre éléments (spécifiquement, mouvement « phi », Neuhaus, 1930; Caelli & Findlay, 1979, 1981; Tyler, 1973; Verstraten, Cavanagh, & LaBianca, 2000; Battelli et al, 2001), la discrimination de phase de lumières clignotantes (Rogers-Ramachandran & Ramachandran, 1998 ; Battelli et al, 2003), le mouvement de réseaux en translation définis par l'information de stéréo ou de mouvement (Patterson et al, 1992; Lu & Sperling, 1995), la poursuite de mouvement échantillonné (Morgan & Turnbull, 1978), ou la poursuite attentionnelle (Verstraten, Cavanagh, & LaBianca, 2000). Le rythme de 4 à 8 Hz correspond à une fréquence d'alternance de 8 à 16 époques par seconde. Cette valeur est compatible avec les cadences les plus rapides auxquelles des scènes ou mots cibles individuels peuvent être identifiés en présentation visuelle sérielle rapide ou « RSVP » (Potter, 1993).

Cette résolution temporelle étonnamment grossière de l'attention produit-elle des effets néfastes mesurables sur la perception, similaires aux confusions qui se manifestent parfois à cause du crowding spatial (e.g. Figure 2) ? Dans le domaine temporel, ce type de confusion reviendrait à confondre l'ordre de certains événements successifs. C'est précisément ce qui arrive lorsqu'un mouvement périodique (comme une roue en rotation) est présenté à certaines fréquences particulières (entre 8 et 12 Hz) : l'observateur éprouve des impressions fugaces de mouvement inversé (Schouten, 1967; Purves, Paydarfar & Andrews, 1996; VanRullen, Reddy & Koch, 2005). Ce phénomène, appelé « Illusion de la roue continue » ou « continuous Wagon Wheel Illusion », est très certainement dû à une forme d'aliasing temporel : si l'attention échantillonne l'information de mouvement à 10-15Hz, les échantillons successifs pour un mouvement autour de 10Hz vont

s'aligner pour indiquer la mauvaise direction de mouvement (le même phénomène se produit, à des fréquences différentes, au cinéma ou à la télévision, du fait d'un aliasing temporel au niveau de la caméra et/ou du dispositif de projection). Nous avons démontré que ces inversions illusionnelles n'affectent que les processus d'extraction de mouvement qui sont basés sur une poursuite attentionnelle (VanRullen et al, 2005; voir aussi l'Encart 2: "Paradigmes"), alors que le système de mouvement dit « bas niveau » ou « premier ordre » (Cavanagh & Mather, 1989 ; Lu & Sperling, 1995) peut correctement suivre la direction de mouvement réelle jusqu'à des fréquences bien plus élevées. En comparant l'activité électrique du cerveau d'observateurs qui perçoivent le mouvement réel vs. illusoire (avec le même stimulus à l'écran), nous avons pu isoler les corrélats neuronaux de cette illusion (Figure 3) : une activité oscillatoire entre 10-15Hz, localisée sur les électrodes pariétales droites. Cette composante de l'EEG est vraisemblablement le reflet de la cadence à laquelle l'attention échantillonne le monde visuel –au moins pour ce type d'information de mouvement périodique.

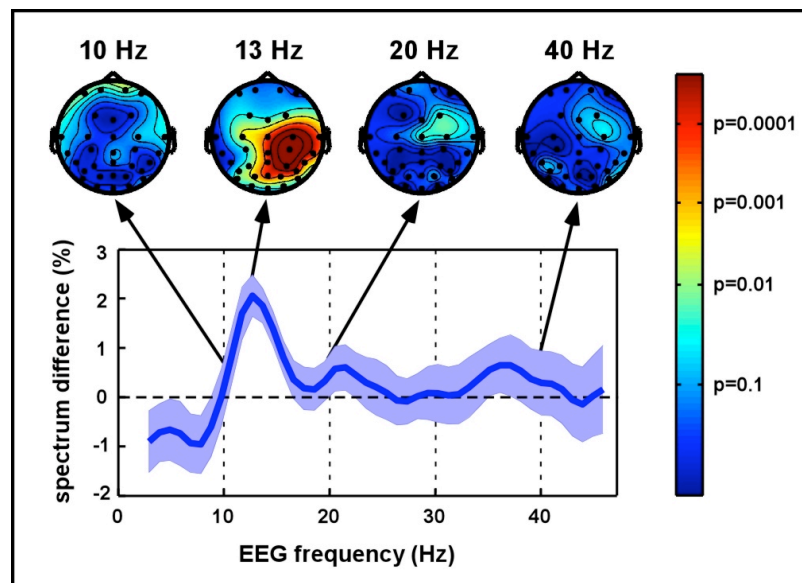


Figure 3. Corrélat neuronal des inversions illusionnelles du mouvement perçu. Nous avons comparé le spectre d'amplitude de l'EEG durant les périodes où les sujets percevaient le mouvement réel ou inversé (illusoire) pour une roue en rotation à une fréquence donnée (les résultats étaient comparables, que cette fréquence soit 10Hz ou 7.5Hz). Bien que la stimulation physique était identique dans ces deux situations, une composante (unique) de l'EEG différait significativement. Cette composante était centrée autour de 13Hz –qui se trouve être la fréquence d'échantillonnage induisant un aliasing maximal pour le mouvement autour de 10Hz, comme observé en psychophysique (voir aussi l'Encart 2: "Paradigmes"). Ces données sont donc compatibles avec l'idée que les inversions illusionnelles du mouvement sont dues à une capture attentionnelle périodique de l'information de mouvement, avec un échantillon pris environ toutes les 75 msec. La localisation de cet effet en regard des régions pariétales droites est compatible avec leur implication connue dans les situations où l'attention doit être portée sur des événements temporels (adapté de VanRullen, Reddy & Koch, 2006).

6. Résolution pour des cibles séquentielles multiples

Si un élément individuel peut être sélectionné dans une séquence rapide à des cadences allant jusqu'à 7 ou 8 Hz, que se passe-t-il lorsque plus d'un élément doit être sélectionné ? C'est le cas du paradigme de « clignement attentionnel » ou « attentional blink » (Raymond, Shapiro, & Arnell, 1992) dans lequel 2 éléments d'une séquence rapide doivent être rapportés. Bien que la vitesse de présentation soit toujours de 7 ou 8 Hz, la performance baisse si les deux cibles sont placées trop près au sein de la séquence temporelle. Les cibles doivent être séparées par plusieurs éléments intermédiaires pour que la performance remonte au niveau observé pour des éléments individuels (notons néanmoins que les cibles immédiatement adjacentes peuvent échapper à cette interférence, Chun & Potter, 1995). Cette interférence entre cibles multiples dans le temps fait écho à celle induite par deux cibles proches spatialement (Bahcall & Kowler, 1999) et suggère l'existence dans le domaine temporel d'un « pourtour suppressif » qui ferait suite à l'accès de la première cible par l'attention. Comme dans le domaine spatial, la résolution temporelle de l'attention pourrait se révéler bien plus imprécise lorsqu'elle est testée par des cibles multiples plutôt que des cibles individuelles.

Les résultats du paradigme de clignement attentionnel suggèrent également que deux traitements qui nécessitent l'attention ne peuvent se dérouler simultanément, puisque sinon l'identification de la première et de la deuxième cible pourraient se chevaucher et il n'y aurait pas d'interférence : une cible qui peut être identifiée dans la séquence pourrait l'être tout aussi bien, même quand une autre cible est déjà en train d'être analysée. L'interférence a lieu, mais une explication possible est que les deux identifications ne peuvent se faire en parallèle parce que les deux cibles sont présentées à la même position spatiale. Après tout, la littérature sur la poursuite d'objets multiples montre que l'on peut porter l'attention sur plusieurs endroits en même temps. Il existe des données montrant que la performance diminue si les sujets doivent identifier les cibles plutôt que seulement poursuivre leurs positions respectives (e.g. Saiki, 2001), mais néanmoins il serait utile de tester si la limite de l'accès séquentiel par l'attention est uniquement valide pour des cibles successives à une même position.

Une expérience de synchronie temporelle (Holcombe & Cavanagh, 2001) adresse cette question. Un élément dans cette étude alternait entre les couleurs rouge et vert, alors que l'élément adjacent alternait entre des orientations obliques vers la gauche et vers la droite. La fréquence d'alternance devait être réduite jusqu'à 3Hz avant que les sujets puissent rapporter quelle couleur correspondait à quelle orientation. Les états individuels de chaque alternance peuvent être identifiés à des cadences de 7 ou 8 Hz (Verstraten et al, 2000) ; donc si l'attention était capable de sélectionner les deux séquences simultanément, les sujets auraient dû pouvoir sélectionner et comparer leur état instantané à cette fréquence. Ce n'est pas ce qui s'est produit. Ceci implique que l'accès attentionnel doit être un mécanisme sériel.

7. Discussions et questions futures

L'information visuelle est traitée en plusieurs étapes. A chaque moment, l'on ne peut être conscient de tout, mais seulement des choses auxquelles on porte l'attention, ou qui capturent cette attention. Dans ce sens, la résolution attentionnelle est l'entonnoir de notre conscience, déterminant les limites de notre expérience spatiale et temporelle. Bien que nous ayons utilisé le mot « attention » comme s'il s'agissait d'un unique mécanisme, ce n'est pas forcément le cas (Allport, 1993; Posner & Dehaene, 1994 ; Driver, 2001). Les études sur la segmentation de texture sont aussi liées à la résolution attentionnelle, puisqu'une texture peut être considérée comme un motif qui peut être visuellement résolu, mais au-delà de la résolution de l'attention. Le concept de résolution attentionnelle pourrait nous aider à mieux comprendre divers phénomènes tels que les effets de contexte, le crowding, le masquage, certaines perceptions subliminales, ou la perception dans les séquences RSVP. Au-delà de ces évaluations scientifiques des limites spatiales et temporelles de la conscience, il reste l'importance de pouvoir mesurer ces limites chez un individu afin de vérifier ses compétences pour des tâches complexes telles que la conduite d'une automobile, le pilotage d'un avion, la

chirurgie, et bien d'autres. Les résultats fondamentaux sur l'attention visuelle, passés en revue dans les pages précédentes, sont une première étape dans le développement de tests pratiques d'évaluation.

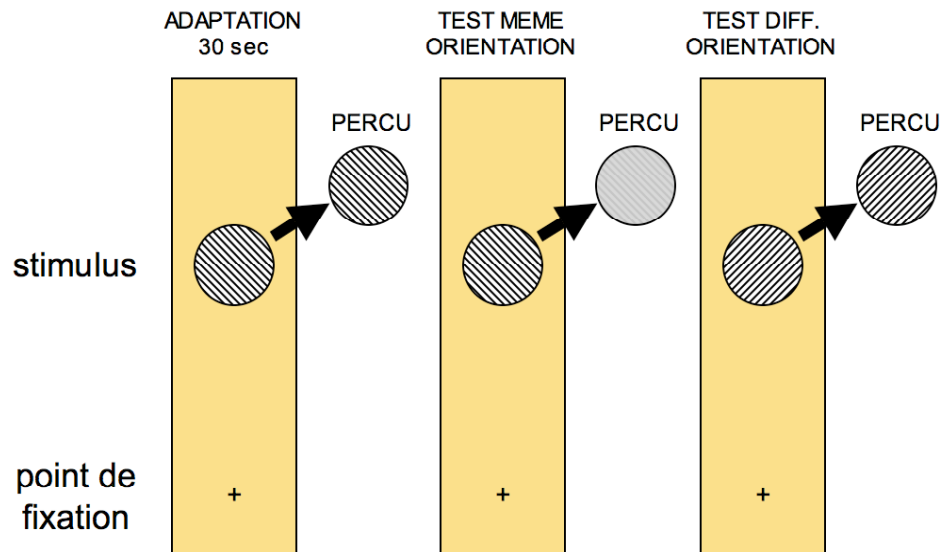
References

- Allport, A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In S.K. David E. Meyer (Ed.), *Attention and performance 14: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*. (pp. 183-218): MIT Press, Cambridge, MA, US.
- Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual fields. *Psychological Science*, **15**, 106-111.
- Andrews, T. J., White, L. E., Binder, D., & Purves, D. (1996). Temporal events in cyclopean vision. *Proc Natl Acad Sci U S A*, *93*(8), 3689-3692.
- Awh E., & Pashler, H. (2000). Evidence for split attentional foci. *J Exp Psychol Hum Percept Perform.* *26*, 834-46.
- Bahcall, D. O., & Kowler, E. (1999). Attentional interference at small spatial separations. *Vision Res.* **39**, 71-86.
- Battelli, L., Cavanagh, P., Intriligator, J., Tramo, M. J., Hénaff, M.-A., Michèl, F., & Barton, J. J. S. (2001) Unilateral right parietal damage leads to bilateral deficit for high-level motion. *Neuron*, *32*, 985-995.
- Battelli, L., Martini, P., Barton, J. S. S., & Cavanagh, P. (2003). Bilateral deficits of transient visual attention in right parietal patients. *Brain*, *126*, 2164-2174.
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(4), 177-178.
- Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, *13*(4), 767-782.
- Caelli, T., & Finlay, D. (1979). Frequency, phase, and colour coding in apparent motion. *Perception*, *8*, 59-68.
- Caelli, T., & Finlay, D. (1981). Intensity, spatial frequency, and temporal frequency determinants of apparent motion. *Perception*, *10*, 183-189.
- Campbell, F.W., & Gubisch, R.W. (1966). Optical quality of the human eye. *Journal of Physiology*, *186*, 558-578.
- Carlson, T. A., Alvarez, G. A., & Cavanagh, P. (2007). Attentional interference reveals anatomical constraints in attentional tracking. Submitted.
- Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental Psychology: Human Perception & Performance*, *18*(3), 837-848.
- Cavanagh, P. (1992). Attention based motion perception. *Science*, *257*, 1563-1565.
- Cavanagh, P., & Alvarez, G. (2005). Tracking multiple targets with multifocal attention. *TICS*, **9**, 349-354.
- Cavanagh, P., & Mather, G. (1989). Motion: the long and short of it. *Spatial Vision* **4**, 103-129.
- Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *J Exp Psychol Hum Percept Perform.*, *21*, 109-27.
- Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*, 53-78.
- Driver, J., & Baylis, G.C. (1989). Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception & Performance*, *15*, 448-456.
- Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, *369*, 313-315.

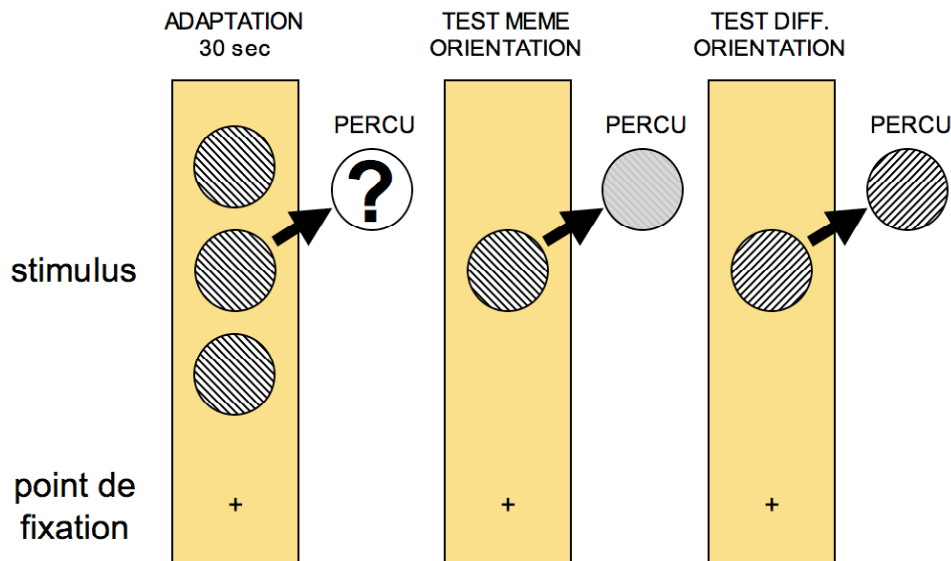
- Eriksen, C.W., & St. James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, 40, 225-240.
- Eriksen, C.W.W., & Hoffman, J.E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics*, 12(2-B), 201-204.
- Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, 423(6939), 534-537.
- Green, C. S., & Bavelier, D. (2006). Enumeration versus multiple object tracking: the case of action video game players. *Cognition*, 101(1), 217-245.
- Grüsser, O.-J., & Landis, T. (1991). *Visual Agnosias and other disturbances of visual perception and cognition*. London: The Macmillan Press.
- Hamker, F. H. (2004). A dynamic model of how feature cues guide spatial attention. *Vision Res*, 44(5), 501-521.
- He, S., & MacLeod, D. I. (2001). Orientation-selective adaptation and tilt after-effect from invisible patterns. *Nature*, 411, 473-6.
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, 383, 334-337.
- He, S., Intriligator, J., Verstraten, F.A.J. & Cavanagh, P. (1998). Slow mechanism for phase discrimination of both luminance and color flicker. *Investigative Ophthalmology & Visual Science*, 39, 1110 (suppl).
- Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nature Neuroscience*, 4, 127-128.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, 43, 171-216.
- Johnston A. (1987). Spatial scaling of central and peripheral contrast-sensitivity functions. *J Opt Soc Am A*, 4, 1583-93.
- Jonides, J. (1980). Towards a model of the mind's eye's movement. *Canadian Journal of Psychology*, 34(2), 103-112.
- Jonides, J. (1983). Further toward a model of the mind's eye's movement. *Bulletin of the Psychonomic Society*, 21(4), 247-250.
- Klein, R., & McCormick, P. (1989). Covert visual orienting: Hemifield-activation can be mimicked by zoom lens and midlocation placement strategies. *Acta Psychologica*, 70, 235-250.
- Kramer, A.F., & Hahn, S. (1995). Splitting the beam: Distribution of attention over noncontiguous regions of the visual field. *Psychological Science*, 6, 381-386.
- LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception & Performance*, 9(3), 371-379.
- Landolt, E. (1891). Nouvelles recherches sur la physiologie des mouvements des yeux. *Archives d'ophtalmologie*, 11, 385-395.
- Lichtenstein, M. (1961). Phenomenal simultaneity with irregular timing of components of the visual stimulus. *Percept Mot Skills*, 12, 47-60.
- Lu, Z.-L. & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, 35, 2697-2722.
- Morgan, M. J., & Turnbull, D. F. (1978). Smooth eye tracking and the perception of motion in the absence of real movement. *Vision Research*, 18, 1053-1059.
- Mounts, J. R. (2000). Evidence for suppressive mechanisms in attentional selection: feature singletons produce inhibitory surrounds. *Percept Psychophys*, 62, 969-83.
- Neuhaus, W. (1930). Experimentelle Studien über das Schein von Bewegung. *Pflügers Archiv für die Gesamte Psychologie des Menschen und der Tiere*, 75, 315-458.

- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nat Neurosci.*, 4, 739-44.
- Patterson R, Ricker C, McGary J, Rose D (1992). Properties of cyclopean motion perception. *Vision Res* 32, 149-56
- Posner, M.I., & Dehaene, S. (1994). Attentional networks. *Trends Neurosci*, 17(2), 75-9.
- Potter, M. C. (1993). Very short-term conceptual memory. *Mem Cognit.*, 21, 156-61.
- Purves, D., Paydarfar, J. A., & Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proc Natl Acad Sci U S A*, 93(8), 3693-3697.
- Pylyshyn, Z.W., & Storm, R.W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3), 179-197.
- Rizzolatti, G., Riggio, L., Dascola, I., & Umiltà, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, 25(1A), 31-40.
- Rogers-Ramachandran, D. C., & Ramachandran, V. S. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Res.*, 38, 71-77.
- Saiki, J. (2003). Feature binding in object-file representations of multiple moving items. *J Vis.*, 3, 6-21.
- Schouten, J.F. (1967). Subjective stroboscopy and a model of visual movement detectors. In: Wathen-Dunn, W. (ed), *Models for the perception of speech and visual form* (pp. 44-45). Cambridge Mass: MIT Press.
- Thibos, L. N., Cheney, F. E., & Walsh, D.J. (1987). Retinal limits to the detection and resolution of gratings. *J Opt Soc Am A.*, 4, 1524-9.
- Toet, A., & Levi, D.M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32, 1349-1357.
- Townsend, J.T., Taylor, S.G., & Brown, D.R. (1971). Lateral masking for letters with unlimited viewing time. *Perception & Psychophysics*, 10(5), 375-378.
- Tsotsos, J. K., Culhane, S. M., Wai, W. Y., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*. 78, 507-545.
- Tyler, C. W. (1973). Temporal characteristics in apparent movement: omega movement vs. phi movement. *Quarterly Journal of Experimental Psychology*, 25, 182-192.
- van de Grind, W.A., Grüsser, O.-J. and Lunkenheimer, H.U (1973). Temporal transfer properties of the afferent visual system. Psychophysical, neurophysiological and theoretical investigations. In Jung, R. (ed.), *Handbook of Sensory Physiology*, Vol. VII/3, Chapter 7, Berlin: Springer [pp431-573].
- VanRullen, R., Reddy, L., & Koch, C. (2005). Attention-driven discrete sampling of motion perception. *Proc Natl Acad Sci U S A*, 102(14), 5291-5296.
- VanRullen, R., Reddy, L., & Koch, C. (2006). The continuous wagon wheel illusion is associated with changes in electroencephalogram power at approximately 13 Hz. *J Neurosci*, 26(2), 502-507.
- Verstraten, F. A.J., Cavanagh, P. & Labianca, A. T. (2000). Limits of attentive tracking reveal temporal properties of attention. *Vision Research*, 40, 3651-64.
- Virsu, V., Naesaenen, R., & Osmoviita, K. (1987). Cortical magnification and peripheral vision. *Journal of the Optical Society of America*, 4, 1568-1578.
- White, C. T., Cheatham, P. G., & Armington, J. C. (1953). Temporal numerosity: II. Evidence for central factors influencing perceived number. *J Exp Psychol*, 46, 283-287.
- White, C., & Harter, M. R. (1969). Intermittency in reaction time and perception, and evoked response correlates of image quality. *Acta Psychol (Amst)*, 30, 368-377.
- Yantis, S. (1992). Multi-element visual tracking: attention and perceptual organization. *Cognitive Psychology*, 3, 295-340.

Encart 1. “Points de Recherche”: La résolution attentionnelle et le site de la perception consciente.



a) Adaptation à une grille visible réduit le contraste perçu pour un test ayant la même orientation



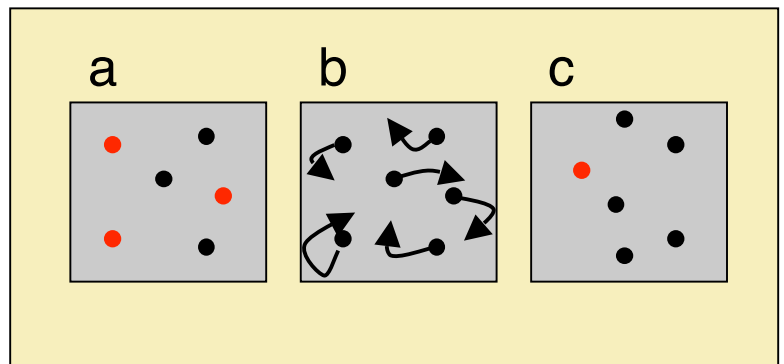
b) Adaptation à une grille rendue invisible par crowding produit autant d'effet d'adaptation comparable

Les propriétés de la résolution attentionnelle suggèrent que la sélection et la perception consciente ont lieu au-delà du cortex visuel primaire (He, Cavanagh, & Intriligator, 1996). Dans cette étude, nous avons bloqué l'accès conscient d'un fragment du champ visuel composé de barres orientées, en le soumettant à des conditions de « crowding » — c'est-à-dire en le plongeant parmi d'autres fragments similaires, plus rapprochés que la résolution de l'attention. Avec cet espacement, les observateurs ne pouvaient pas rapporter

l'orientation des barres. Nous avons exposé nos sujets à ces stimuli pendant 30 secondes, et avons ensuite testé la présence d'un aftereffect d'orientation à l'endroit adapté. Dans une condition de comparaison, l'observateur s'adaptait au même type de fragment, mais sans autre fragment alentours ; dans ce cas, le sujet pouvait facilement rapporter l'orientation des barres de ce stimulus « adaptateur ». De manière surprenante, l'aftereffect était le même avec ou sans crowding — c'est-à-dire, indépendamment de la perception consciente de l'adaptateur. Ces résultats indiquent que le stimulus doit avoir activé des neurones au moins au niveau de V1 (le premier endroit dans le cortex visuel où l'on peut enregistrer une sélectivité à l'orientation), que le sujet ait ou n'ait pas perçu le stimulus adaptateur. Le site de la sélection attentionnelle, et des limitations spatiales associées, ainsi que de la perception consciente, doit donc se trouver au-delà de V1. Cette étude montrait également un avantage marqué pour le champ visuel inférieur dans deux tâches impliquant l'attention. Puisque cette asymétrie se retrouve dans le cortex pariétal, mais pas dans V1, nous avons proposé un locus pariétal pour la sélection attentionnelle.

Encart 2: "Paradigmes"

Poursuite d'objets multiples: La poursuite visuelle chez l'humain dépend de manière critique des ressources attentionnelles; mais classiquement, l'attention visuelle ne sélectionne l'information qu'à un seul endroit à la fois, se comportant comme une seconde fovéa qui parcourrait le champ visuel indépendamment des mouvements oculaires. En 1988, Pylyshyn et Storm ont renversé cette idée du focus attentionnel unique en démontrant qu'il était possible de poursuivre plusieurs cibles

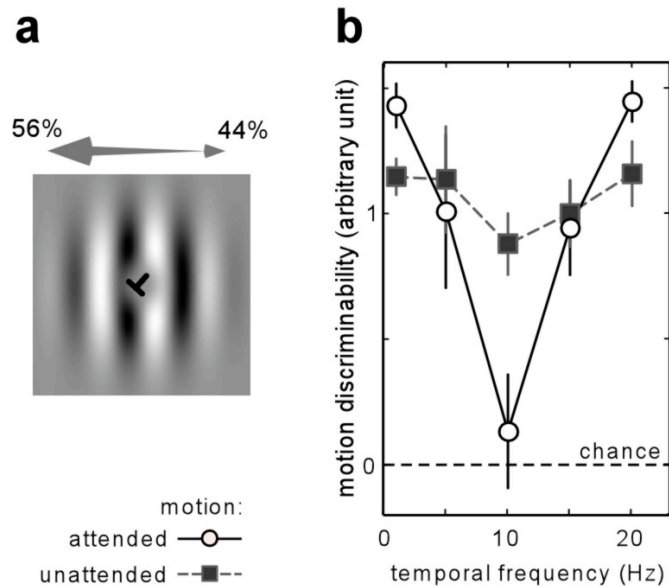


simultanément et de façon prolongée. Ce paradigme a été utilisé dans de nombreuses études depuis cette époque, aidant ainsi à établir les limites de l'attention dans différentes conditions et pour différentes populations de sujets. Dans leurs expériences, des groupes d'éléments identiques se déplaçaient dans une région rectangulaire, en rebondissant les uns sur les autres ainsi que sur les bords. Au début de chaque essai, certains des éléments étaient marqués brièvement comme « cibles » (en rouge dans le panneau de gauche) et redevaient ensuite identiques aux autres éléments. Tous les objets se mettaient ensuite à se déplacer aléatoirement au sein de la zone pendant les 7-15 secondes suivantes (panneau central). Pour mesurer la performance à la fin de l'essai, un élément unique était indiqué, et les sujets rapportaient s'il était ou non l'un des éléments cibles (panneau de droite). La plupart des sujets pouvaient poursuivre ainsi jusqu'à 4 ou 5 cibles pendant plusieurs secondes. Ils ne pouvaient pas se fier à leurs seuls mouvements oculaires, ni à un focus attentionnel unique sélectionnant une large région comprenant toutes les cibles — puisque cette région engloberait aussi nécessairement les distracteurs. Les sujets se comportaient comme s'ils avaient à leur disposition plusieurs foyers attentionnels, un pour poursuivre chacune des cibles (Cavanagh & Alvarez, 2005).

L'attention et les inversions de mouvement illusoirs.

Pour tester la capacité d'un observateur à discriminer la direction de mouvement (gauche ou droite), nous avons utilisé des réseaux sinusoidaux en « contre-phase inégale » : une superposition de deux réseaux sinusoidaux verticaux de fréquences spatiale et temporelle identiques, mais de directions opposées. En manipulant le rapport de contraste entre ces deux réseaux, on peut biaiser le mouvement du stimulus en faveur d'une direction (dans cet exemple, vers la gauche). Quand on demande à des observateurs de rapporter la direction perçue pour ces stimuli, leur performance chute spécifiquement, jusqu'au niveau de la chance, pour

les fréquences temporelles autour de 10Hz (cercles blancs dans le panneau **b.**): à cause de l'aliasing temporel à cette fréquence, le stimulus en mouvement est perçu dans la mauvaise direction, quasiment une fois sur deux. Quand les sujets doivent en même temps faire attention à une séquence de lettres qui changent rapidement, au centre du réseau sinusoidal (L ou T ayant subi une rotation aléatoire, voir **a.**), leurs résultats changent de manière très prononcée (carrés noirs dans le panneau **b.**): l'aliasing temporel à 10Hz semble disparaître. Ceci implique que ce n'est pas le traitement du mouvement lui-même, mais plutôt la *capture attentionnelle* de l'information de mouvement (Cavanagh, 1992) qui dépend d'un échantillonnage périodique, à une fréquence d'environ 13Hz (voir aussi Figure 3).



Encart 3 : Questions en Suspens.

- L'attention repose-t-elle sur le même système que les mouvements oculaires pour représenter les régions d'intérêt ? Le système responsable des mouvements oculaires enregistre les cibles potentielles pour un déplacement du regard (une saccade), et ces cibles potentielles entrent en compétition, de sorte que la cible la plus forte et la plus saillante domine, et supprime toutes les autres. Dès qu'une cible unique émerge de la compétition, le mouvement oculaire peut démarrer. De nombreux auteurs suggèrent que l'attention est une fonction qui pourrait être supportée par le système des mouvements oculaires, lorsque ces mouvements sont activement retenus. C'est ce qu'on nomme parfois la théorie « prémotrice » de l'attention (Rizzolatti et al, 1987). Chaque cible potentielle est comparable à un focus attentionnel; si la compétition entre cibles était atténuée, plusieurs foyers pourraient être maintenus simultanément (Hamker, 2004 ; Cavanagh & Alvarez, 2005).

- Comment l'attention est-elle divisée *dans l'espace* entre des cibles spatiales multiples ? Quelle est la forme exacte du foyer attentionnel, et de son pourtour suppressif ? Cette forme change-t-elle si le nombre de foyers attentionnels augmente, ou s'ils doivent se déplacer pour poursuivre une cible mobile ?

- Comment l'attention est-elle divisée *dans le temps* entre des cibles spatiales multiples ? Échantillonne-t-elle toutes les cibles à la fois, ou chacune à son tour ? Quelle est sa fréquence

d'échantillonnage, et comment celle-ci diffère-t-elle de la fréquence d'échantillonnage pour une cible unique ?

- Les limites temporelles de l'attention sont-elles comparables pour différents types de discrimination visuelle? Existe-t-il une fonction de crowding temporel « universelle », similaire à la règle « étendue du crowding = $\frac{1}{2}$ excentricité » qu'on observe dans le domaine spatial (Bouma, 1970; Toet & Levi, 1992; Intriligator & Cavanagh, 2001), quelque soit l'objet de l'attention ?

- L'acuité spatiale et/ou temporelle de l'attention peuvent-elles être augmentées par un entraînement spécifique ? Nous suggérons qu'il est nécessaire de développer des tests permettant d'évaluer la résolution attentionnelle. Tout aussi important serait le développement de programmes d'entraînement qui pourraient ramener à un niveau de performance standard des pilotes ou des conducteurs souffrant de déficits attentionnels. Des études récentes indiquent, par exemple, que la pratique des jeux vidéo peut améliorer la résolution attentionnelle (Greene et Bavelier, 2003, 2006).