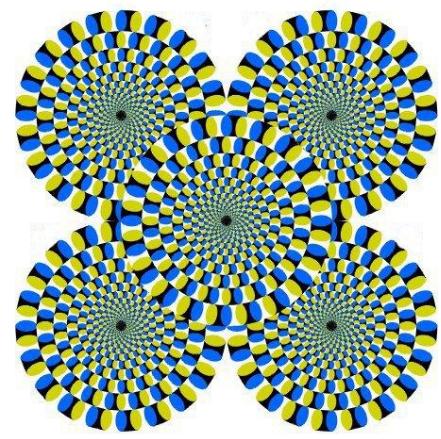
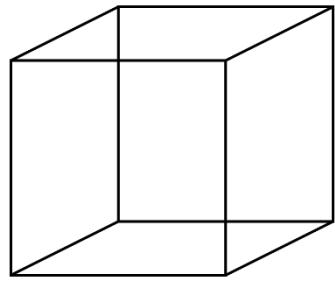
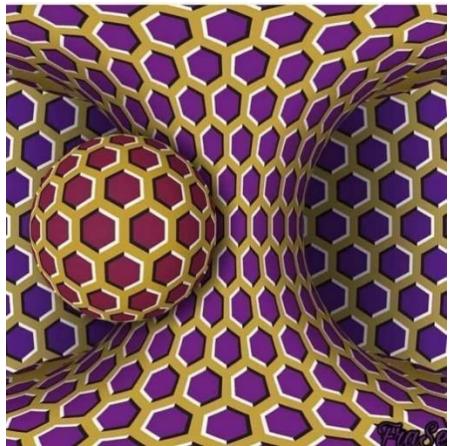


---

## ***PERCEZIONE E PSICOFISICA***

---



## **Lezione 1 – Introduzione – Catena psicofisica – Inizio psicofisica: sensibilità e soglia**

### ***Quanti sono i sensi?***

Sono numerosi i mezzi che adottiamo a livello funzionale per interagire con il mondo esterno: la visione, l'udito, l'olfatto, il gusto, il tocco leggero, la pressione, la temperatura (nello specifico distinguendo tra caldo e freddo), il dolore o il prurito. In questi processi una struttura molto importante è la *vestibula*, situata nell'orecchio interno, che regola l'equilibrio.

Nel complesso, di qualsiasi natura sia lo stimolo esterno crea in noi una **percezione** avente determinate *caratteristiche qualitative*; ad esempio, la qualità di una percezione visiva è completamente diversa da quella acustica. Molto importanti sono anche gli aspetti legati alla somatosensazione; infatti è estremamente difficile stabilire una linea di demarcazione tra i vari stimoli somatosensoriali. La *pressione* o la *vibrazione* sono chiaramente sensazioni diverse, **ma hanno sempre a che fare con uno stimolo meccanico sulla pelle**; lo stiramento e la sensazione di caldo/freddo sembrano invece essere diversi dalla sensazione dello stimolo meccanico. Nello stabilire questa linea di demarcazione, un criterio che può darci svariate spiegazioni è l'organo in sé, per quanto riguarda somatosensazione, temperatura, dolore ecc... possiamo pensare alla **pelle** che è un organo molto complesso che regola meccanismi cutanei e sottocutanei, molto più profondi.

Siamo soliti affermare che esistono **5 sensi** ma, considerando la grande variabilità di stimoli da noi percepiti, potremmo affermare che in verità ne possediamo molti di più.

Noi esseri umani, a differenza di tanti altri animali, siamo capaci di percepire principalmente tre tipi di energia:

- **Energia elettromagnetica** → la luce (che regola anche i ritmi circadiani);
  - **Energia meccanica** → postura, pressione, stiramento, vibrazione e tutti gli stimoli che siamo in grado di codificare con i recettori sotto la pelle;
- P.S.:** Anche la forza di gravità così come il suono, sono forme di energia meccanica; nello specifico, il *suono* è una variazione ciclica di pressione che si propaga nell'aria come se fosse un'onda.
- **Energia chimica** → è la primissima forma di energia dal punto di vista evolutivo.

Un esempio potrebbe essere il girasole che girandosi verso il sole risponde all'energia elettromagnetica con un movimento.

In natura esiste un *continuum di fenomeni* che creano una **costante interazione tra l'organismo e l'ambiente esterno**, che viene percepito proprio per mezzo delle varie forme di energia sopra riportate.

La **sensazione** inizia nel momento in cui ci sono dei neuroni che sono capaci di fare **trasduzione**, ossia trasformare un'energia esterna (elettromagnetica, meccanica o chimica) in un impulso nervoso.

Il meccanismo di trasduzione rappresenta l'**inizio** della percezione, per cui lo stimolo esterno assume una rappresentazione a livello neurale, generando così un'interfaccia tra l'organo e l'ambiente fenomenico. In quest'ottica la “sensazione” nasce dall’interfaccia tra organo e ambiente quindi si può pensare di attribuire ai sensi non più il canonico numero 5, ma un numero **in funzione dei recettori** che rispondono nel momento della percezione di un qualsiasi tipo di stimolo.

#### Esempio

- *Sistema visivo*: nell’occhio abbiamo 4 tipi di fotorecettori che trasducono la luce:
  - tre diversi tipi di coni che, avendo una sensibilità spettrale ad ampio raggio assorbono tre diverse lunghezze d’onda;
  - bastoncelli, utilizzati durante la visione scotopica.

Oltre ai fotorecettori nella retina sono presenti altre cellule, come le **cellule gangliari** che rispondono direttamente alle intensità luminose – ad esempio, rispondono ai ritmi circadiani in quanto intrinsecamente fotosensibili.

- *Sistema somatosensoriale*: prendendo un pezzo di pelle del dito indice, troviamo 4 tipi di meccanocettori, recettori specifici per il dolore, poi ancora recettori per il caldo e per il freddo; se invece ci spostiamo verso l’articolazione della prima falange troviamo recettori che riconoscono altri stimoli, come ad esempio, lo stiramento.

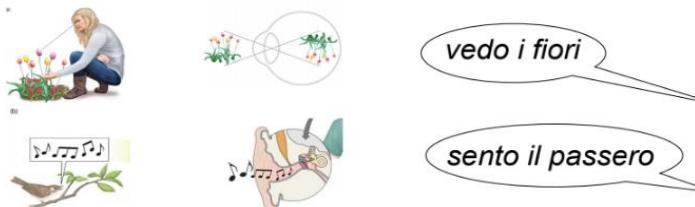
Dovendo tenere conto di tutti i tipi di recettori che esistono per recepire stimoli di natura diversa, il numero che definirebbe le modalità sensoriali tenderebbe all’infinito. **Ciò nonostante siamo soliti distinguere solo 5 sensi in quanto tendiamo erroneamente a vedere la percezione come un processo passivo.**

## La catena psicofisica

### la catena psicofisica

Stimulus → Sensation → Perception

stimolo distale → stimolo prossimale → percepito



La percezione viene intesa come una **catena di eventi**: ad esempio nella visione lo stimolo esterno produce una sensazione tramite l'attivazione dei fotorecettori; poi questa sensazione viene interpretata per diventare l'effettiva percezione.

Visto in quest'ottica il **processo percettivo** comporta un passaggio di trasformazione del mondo esterno (stimoli distali) attraverso un

mezzo (onde luminose o energia meccanica) che produce un secondo stimolo (stimolo prossimale) che si esprime in una costellazione di attivazioni sulla superficie recettoriale. Lo stimolo prossimale diventa poi un percepito tramite le numerose interazioni ed elaborazioni neurali che definiscono ciò che si sente.

**Stimolo distale (fiore) → Stimolo prossimale (informazioni sul fiore cadono sui recettori retinici) → Percetto**

La percezione, così, acquisisce due funzioni diverse ma tra loro collegate:

- una **funzione fenomenologica**, che ci consente di sviluppare un'esperienza multisensoriale dell'ambiente esterno (forme, colori, movimento, spazi, suoni appartenenti alla realtà fenomenica);
- la seconda invece, riguarda l'**interazione con il mondo**; noi utilizziamo l'informazione sensoriale per interagire con l'esterno, per guidare movimenti come afferrare un oggetto, guidare la locomozione nell'ambiente senza perdere l'equilibrio, trovare la direzione, controllare la postura. Queste attività non possono essere definite passive, ma **processi altamente attivi**.

Quando pensiamo alla visione ad esempio, l'occhio riceve un'immagine che viene interpretata, perché la visione rappresenta un processo intrinsecamente **esplorativo**. Quando si guarda qualcosa si muovono continuamente gli occhi tramite i movimenti saccadici per esplorare l'informazione che è disponibile nella luce; è un *continuo campionamento* in cui si captano pezzi di informazione (in un universo di stimoli disponibili) fino ad arrivare a processi in cui l'informazione viene continuamente revisionata, integrata ed aggiornata.

**La percezione è un processo di continua revisione e controllo del mondo esterno e delle azioni che accadono in esso; è una vera e propria attività e non un meccanismo passivo: percepire è qualcosa che noi facciamo, non qualcosa che ci accade.**

## Approccio sistemico

Per spiegare la percezione si utilizza un **approccio sistemico**, in cui va considerata *l'interazione* tra svariati aspetti: l'ambiente, grazie ad una serie di fenomeni fisici, è ricco di informazioni portate da "mezzi" (luce, energia meccanica e chimica) e queste potenziali informazioni sono in grado di avere degli effetti sui nostri sistemi sensoriali. L'effetto di queste informazioni sull'organismo è un *effetto cognitivo* in grado di:

- **avere consapevolezza** di quello che si percepisce all'esterno;
- **guidare i nostri movimenti** nell'interazione con l'ambiente.

In conclusione, piuttosto che chiederci di quanti sensi disponiamo, potremmo dire che la percezione è quindi la creazione di un percepito avviene attraverso **5 modalità di esplorazione**:

- 1) movimenti oculari, del capo, del corpo (esplorazione, campionamento, etc.);
- 2) orientamento delle orecchie (tendere le orecchie per ascoltare, è un'azione attiva e lo facciamo girando il capo, orientando l'attenzione);
- 3) esplorazione aptica (l'esplorazione con le mani, col corpo, quello che chiamiamo tatto attivo);
- 4) succhiare, masticare, deglutire (assaporare);
- 5) olfatto attivo (annusare, sniffare);

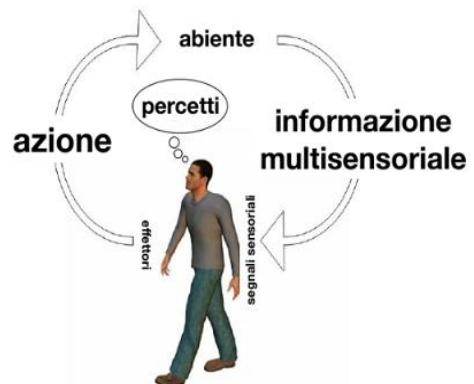
Potremmo poi aggiungere che i nostri movimenti li possiamo utilizzare anche per un *fine comunicativo* quella forma particolare di performance sul mondo che si rivolge ai nostri conspecifici. Le azioni di natura comunicativa rientrano in questo ciclo di percezione.

## Psicofisica 1

### Sensibilità e Soglia

La psicofisica è una disciplina molto articolata nata a metà dell'800 presenta svariati interessi fra cui quello di capire l'**intensità percepita** di uno stimolo. La psicofisica è una disciplina che presenta molte applicazioni pratiche, ad esempio la *psicoacustica* ovvero lo studio della percezione soggettiva umana dei suoni: tramite gli studi di psicoacustica è possibile determinare la quantità di energia meccanica iniziale necessaria a produrre un determinato suono che ha un particolare effetto sul nostro sistema sensoriale acustico.

## ciclo percezione-azione



Un altro esempio riguarda la *psicofisica della visione cromatica*<sup>1</sup>, che sta alla base della colorimetria - riprodurre i colori utilizzando un sistema elettronico/meccanico (fotografia a colori, cinema a colori, laser a colori, lo schermo dei cellulari, televisore, etc). Queste tecnologie sono basate su un'idea semplicissima: se riesco a produrre l'effetto desiderato sulla retina, tutto quello che sarà prodotto sarà uguale a ciò che avviene in una situazione reale e avrò così la stessa percezione. Posso riprodurre anche l'effetto che questo farebbe sulla retina utilizzando altri tipi di tecnologie.

Esempio: un'applicazione della psicofisica è quella di costruire delle scale attraverso le quali si possa dare una misura di “*quanto intensa*” è l’esperienza di una persona. Questo risulta essere un obiettivo molto ambizioso, in quanto è estremamente difficile capire com’è fatta la *percezione dell’esperienza di altre persone*.

**La psicofisica sembra essere un buon metodo per ovviare a questo problema, in quanto consente di trovare delle funzioni che mostrano come il proprio vissuto e la natura della propria esperienza cambi in funzione dello stimolo fisico esterno.**

Esperimento Peperoncini: immaginiamo di avere diversi peperoncini, con concentrazioni diverse di *capsaicina* – componente chimica responsabile della sensazione di bruciore. In base alla sua concentrazione si può osservare come cambia in una persona la sensazione di “*piccantezza*”. Lo scopo è quello di mappare come cambia la sensazione di bruciore, al variare di tale sostanza.

L'esperimento, dunque, si basa sul **modificare la concentrazione di capsaicina**: si prende l'estratto di peperoncino, si aggiunge acqua zuccherata fino a quando non si annulla la sensazione di bruciore.

Supponiamo quindi di avere un millilitro di estratto di peperoncino, che deve essere diluito. È necessario diluirlo in 100 millilitri di acqua zuccherata affinché il bruciore scompaia; 100ml di acqua zuccherata rappresenta, nel caso della *Scala Scoville*, una vera e propria unità di misura. Quindi, maggiore è la quantità di acqua da aggiungere, più potente è quel particolare estratto di peperoncino.

Per costruire questo tipo di scale psicofisiche è necessario **operazionalizzare**, misurando e mettendo in relazione gli stimoli esterni con un dato numerico. In questo esperimento - poco affidabile considerando il fattore soggettivo di tolleranza alla piccantezza - la scala sviluppata è la *Scoville Heat Scale (SHU)*.

---

<sup>1</sup> Psicofisica della visione cromatica: studia la relazione dell'intensità della luce e la sua codifica nella retina.

Ovviamente misurare ad esempio una *lunghezza* è molto più facile che misurare la sensazione di bruciore (soprattutto se si utilizza la quantità di acqua come unità di misura). In questa prospettiva va considerato un aspetto specifico: la relazione tra **sensibilità e soglia**.

***Posso misurare la quantità di peperoncino, ma come si misura l'intensità del bruciore? Come si cattura qualcosa di soggettivo?***

L'uomo è maggiormente predisposto ad essere capace di elaborare certi tipi di informazioni e condizioni, ma risulta esserne meno capace in altri contesti. Questo si applica in molte situazioni, come *nell'ambito visivo* in cui riusciamo a vedere bene le *forme* (se non sono troppo piccole o troppo grandi) e abbiamo una migliore sensibilità per certe *caratteristiche spaziali* dello stimolo visivo.

La stessa cosa vale per i *suoni*: il nostro apparato sensitivo acustico è sensibile solo ad una banda di frequenze sonore, per cui un suono che ricade in questa banda verrà sentito molto bene, mentre tutti i suoni che sono al di fuori di questa banda vengono percepiti male o non percepiti affatto.

➔ Questa è proprio la **sensibilità, intesa come una proprietà della nostra mente che ci permette di discriminare tra vari stimoli**. “*Quanto bene*” un meccanismo percettivo è in grado di elaborare le informazioni negli stimoli (capacità/efficienza discriminativa).

Siamo in grado di discriminare tra colori anche molto simili, oppure distinguere tra pesi poco diversi tra loro. Se si è molto sensibili, si è capaci di fare delle discriminazioni molto fini – anche quando si tratta di oggetti che differiscono poco nel peso.

In questo processo, un ruolo importante in psicofisica viene svolto dalla **soglia**, o anche “**soglia differenziale**”.

➔ La **soglia differenziale** riguarda il *minimo incremento* nell'*energia fisica* di uno *stimolo* (onde elettromagnetiche, frequenze acustiche, etc.) in grado di produrre *sistematicamente un cambiamento* nella percezione. In altre parole è la minima differenza che si è in grado di rilevare in maniera affidabile.

Esempio: La maggior parte delle persone sa distinguere tra 1 kg e 1 hg. Ciò vuol dire che la differenza è ben al di sopra della soglia differenziale, ossia al di sopra di quella quantità fisica che è in grado di produrre, in maniera molto affidabile, una risposta di *natura discriminativa* sul fronte della sensibilità del soggetto. Questa discriminazione, quindi, ha un valore più o meno oggettivo e ci permette di **misurare determinate variabili tramite un collegamento tra fisico e psichico**.

La psicofisica si regge sull'idea di base che la **sensibilità è il reciproco della soglia**. Questo significa che la sensibilità (fattore psicologico) cresce quando la soglia differenziale diventa più piccola; viceversa, più è grande la soglia differenziale peggio sarà la sensibilità (come nell'esempio di discriminazione tra 1kg e 1 hg).

Capito questo, si può capire come funziona la psicofisica, ovvero come si procede per **costruire delle funzioni che mettono in relazione l'aspetto fisico con quello psichico**.

### 1. Prime teorie della psicofisica

Uno dei primi a parlare di psicofisica fu **Gustav Theodor Fechner**, che era profondamente interessato al problema mente-corpo, e a come si potesse studiare la coscienza mettendola in relazione con proprietà del mondo fisico: “*La psicofisica va intesa come una teoria esatta delle relazioni di dipendenza funzionale tra corpo e anima o più in generale tra materiale e mentale, fra mondo fisico e mondo psicologico*”.

Ciò nonostante il vero padre della psicofisica fu **Ernst Weber**, che si interessò al **rapporto tra sensibilità e soglia**. Nello specifico Weber parla di “*soglia assoluta*”:

→ La **soglia assoluta** sarebbe *la minima quantità* di un certo stimolo fisico tale per cui sono capaci di discriminare quello stimolo dalla sua assenza. La soglia assoluta, quindi, è anche una **soglia differenziale di natura discriminativa**: non c'è nessuna differenza concettuale tra soglia assoluta e soglia differenziale. L'unica differenza è che quella che chiamiamo **soglia assoluta** è l'**inizio di una scala**, il “**grado 0**” della sensazione, per cui prima di quel livello di intensità fisica, non ci può essere sensazione in quanto non c'è **trasduzione**. In molte condizioni non siamo in grado di fare discriminazioni. Le soglie assolute, i gradi 0 della sensazione, se si cerca di stimare a che livello di intensità fisica si collocano, in condizioni ideali si collocano a dei livelli di intensità talmente bassi che non si potrebbe scendere al di sotto.



Ci si può chiedere ad esempio, quant'è la *quantità minima di luce percepibile*; la luce per i fisici si descrive attraverso due unità di misura:

- come un'onda continua, quindi fasci di onde elettromagnetiche che viaggiano nello spazio;
- come dei pacchetti di energia definiti “*quanti di luce*” o “*fotoni*”, così da poterla misurare in numeri interi.

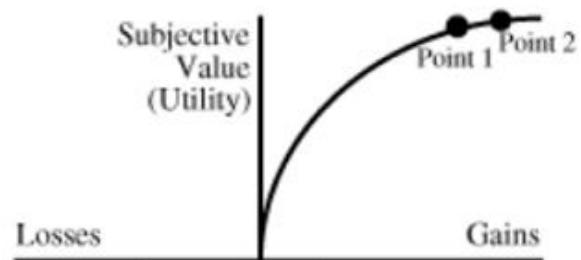
È stato dimostrato che in condizioni ideali (**adattamento al buio**), l'occhio umano è in grado di rilevare le intensità luminose che stanno tra 1 e 10 quanti, l'occhio dunque è in grado di rispondere ad un singolo quanto di luce, una quantità estremamente piccola.

Quello che vale per la visione è vero anche per altri sistemi sensoriali: nel caso dell'*udito*, ad esempio, è stato stimato che risulterebbe disfunzionale avere una sensibilità uditiva migliore di come l'abbiamo – cioè se riuscissimo a rilevare delle energie meccaniche ancora più deboli – in quanto saremmo talmente sensibili che potremmo sentire il pulsare del sangue nelle vene e tutti gli eventi meccanici interni – anche quelli che producono energia meccanica che si può propagare in un mezzo elastico (le nostre ossa e i nostri muscoli).

**Weber**, nella prima metà dell'800 si è dedicato allo studio del *tatto*. Studiando sistematicamente compiti di discriminazione, si è accorto di una cosa fondamentale – una proprietà che vale per la psicofisica e per tutta la nostra psicologia – ovvero che “**la sensibilità è inversamente proporzionale all'intensità dello stimolo di riferimento**”; quindi, più intenso è lo stimolo di riferimento e peggiore è la sensibilità. Infatti facendo il compito di discriminazione con *stimoli deboli* siamo capaci di rilevare differenze molto piccole, ma se gli *stimoli sono molto intensi*, non è necessario essere sensibili alle differenze molto piccole, in quanto sono presenti numerose informazioni.

Un altro esponente della psicofisica fu **Daniel Kahneman**, premio Nobel per l'economia<sup>2</sup>. Kahneman è uno psicologo del pensiero che si occupa di “ragionamento” e di come le persone prendono le decisioni. Il punto di partenza della sua scoperta è che “il *valore soggettivo* (o di utilità) di qualsiasi cosa dipende dal *grado di soddisfazione*”, questo fenomeno è rappresentato da una curva piuttosto che una funzione lineare, ovvero una retta.

**Kahneman & Tversky**



<sup>2</sup> Kahneman è uno psicologo che vinse, nel 2002, il premio Nobel per l'economia; il suo lavoro ha permesso di integrare risultati di ricerca psicologica nella scienza economica, per comprendere come il giudizio umano influenzì le decisioni in condizioni di incertezza.

### ESEMPIO: PERDITA E GUADAGNO

- Sul piano cartesiano l'asse delle X corrisponde alla misura oggettiva di "perdita" e "guadagno";
- Sull'asse delle Y c'è il valore soggettivo di un determinato stimolo.

Prendiamo, sull'asse delle X, un punto 0 che corrisponderà allo "status quo" (es: stipendio).

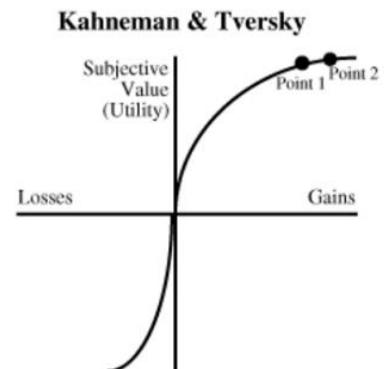
In corrispondenza di questo punto 0 si può immaginare una misura oggettiva del valore, messo in relazione con il valore soggettivo. La funzione che si ottiene è una **curva negativamente accelerata**.

Per intenderci, supponiamo di percepire uno stipendio di 1000 euro al mese; successivamente il datore di lavoro decide di alzare lo stipendio del 10% (100 euro). Naturalmente si tenderà ad associare all'aumento di stipendio un certo grado di contentezza e soddisfazione personale. Se però apportiamo lo stesso aumento (100 euro) ad un soggetto che prende 10mila euro al mese di stipendio, chiaramente il grado di soddisfazione sarà totalmente diverso.

Se lo stimolo di riferimento (il pt 0, lo stipendio) è basso, il grado di contentezza per l'aumento sarà decisamente più significativo rispetto al caso in cui un soggetto parte da uno stimolo di riferimento (stipendio) più alto, in quanto quest'ultimo partendo da una soglia più alta non risulta sensibile ad un cambiamento tanto piccolo (la soglia differenziale dovrebbe essere più grande).

Considerando la **legge di Weber** per cui "*la sensibilità è inversamente proporzionale all'intensità dello stimolo di riferimento*", se lo stimolo di riferimento (il punto di partenza/stipendio) è piccolo allora 100 euro di aumento fanno la differenza e producono un discreto effetto su quanto il soggetto sia soddisfatto e dato che "la soglia differenziale è il reciproco della sensibilità" ciò vuol dire che la soglia è **direttamente proporzionale allo stimolo, quindi è legata ad una costante di proporzionalità**.

Uno degli aspetti della teoria di **Kahneman e Tversky**, è quello di dimostrare che le decisioni economiche si comportano in maniera diversa a seconda del contesto di guadagno, costruendo una curva del valore soggettivo; questa non è solo una curva negativamente accelerata (verso i valori positivi, ossia l'indice di gradimento sul guadagno), ma è anche **una curva negativamente decelerata**: c'è un'altra parte della curva che descrive quanto si è infelici al crescere di una *perdita*. Questa **curva sinusoidale è asimmetrica**. Quindi le persone rispondono diversamente alle perdite e ai guadagni e anche alle prospettive di una perdita e di un guadagno. Questo spiega la gran parte dei comportamenti economici.



### APPROFONDIMENTO: ESPERIMENTI DI WEBER SULLA DISCRIMINAZIONE DEL PESO TRA DUE OGGETTI

Questo tipo di esperimento è applicabile solo per stimoli ad ampia gamma di intensità fisiche. Al fine di comprendere da che soglia si può iniziare a discriminare in funzione delle intensità degli oggetti, al soggetto venivano dati due pesi in mano e il compito era di riferire quale dei due pesasse di più. Facendo questo tipo di discriminazione, Weber capì che se si parte da un peso basso e si chiede sistematicamente al soggetto “di quanto è necessario aumentare il peso del secondo stimolo” per farlo arrivare alla soglia differenziale (e quindi riscontrare una capacità discriminativa soddisfacente), è più semplice il compito di discriminazione. Infatti su un peso di 100g la **soglia differenziale** era pari a 2g. Se si parte invece da 300g, la soglia differenziale era pari a 6 g, etc.

#### generalizzazione empirica (“legge”)

$$\frac{2}{100} = \frac{6}{300} = \frac{\Delta I}{I} = k$$

frazione di Weber per il peso → 0,02

A partire da questo tipo di osservazioni, **Weber è arrivato a formulare la seguente generalizzazione empirica**. Una *generalizzazione empirica* è un principio generale che vale per un’ampia gamma di osservazioni e può anche essere definita “legge”.

Per un’ampia gamma di intensità fisiche di riferimento, nell’ambito della percezione, questo principio si chiama **legge di Weber** (principio fondamentale che può essere utile per costruire un qualsiasi strumento che deve interagire in qualche misura con l’essere umano, quindi ad esempio per costruire un microfono, una macchina fotografica etc.) in cui si osserva come la nostra sensibilità tende a cambiare in funzione dell’intensità dello stimolo di riferimento:

- **ΔI** è la *soglia differenziale* (minimo incremento discriminabile);
- **I** è lo *stimolo di riferimento*, l’intensità di riferimento (quanto è intensa la luce, quanto è pesante un peso);
- **K** è la *costante di Weber*: la costante di proporzionalità.

(es: la soglia differenziale ( $\Delta I$ ) sulla discriminazione del peso è sempre pari al 2% dello stimolo di riferimento (I); se ci sono 100g ci vogliono 2 grammi per discriminare).

**NB** → Se dovessi farvi la domanda “**Cos’è la costante di Weber?**” non bisogna rispondere “è il 2%”.

La costante di Weber è il k di quella particolare legge che vale 2% nel caso del peso ed in determinate condizioni, ma può assumere un valore diverso se viene coinvolto un altro tipo di canale sensoriale (o per altre condizioni).

#### legge di Weber

$$\frac{\Delta I}{I} = k$$

$\Delta I$  : minimo incremento discriminabile

I : intensità di riferimento

k : costante specifica per la modalità

## Lezione 2 – Legge di Weber – Legge di Weber-Fechner – JND – Soglia ideale e soglia reale

### **Legge di Weber e costante di Weber**

**Ernst Weber** verificò empiricamente la sua legge per un ampio range di intensità di uno stimolo (**I**) (dove lo stimolo può essere *qualsiasi cosa*: l'intensità della luce, la concentrazione di una sostanza in un liquido, il peso di un oggetto). Durante le sue misurazioni riscontrò che per un range abbastanza ampio di intensità (corrispondente alle nostre *potenzialità sensoriali*) il nostro apparato percettivo segue la sua legge, mentre quando cominciamo ad andare verso i *limiti* di una scala, cioè con pesi molto grandi o molto piccoli e usciamo dalla gamma di quello con cui siamo fisicamente in grado di interagire, allora la legge non è più valida. ➔ La **Legge di Weber**:  $\frac{\Delta I}{I} = k$  è vera per tutti i sistemi percettivi all'interno di una determinata gamma di valori (che corrispondono ai **limiti del nostro sistema percettivo/recettoriale**), inoltre è importante ricordare che  $k$  è una costante che vale solo per un particolare canale sensoriale.

Ribadiamo che  $\Delta I$  è il minimo incremento discriminabile, una quantità fisica e anche quella che noi chiamiamo *soglia differenziale*. La **soglia** è una *quantità fisica* che sta in un rapporto molto chiaro con una *quantità psicologica* che è la **sensibilità**: **maggior è la soglia, minore è la sensibilità e viceversa**. **I** è un'intensità di riferimento qualsiasi, e **k** è la costante di Weber, specifica per la modalità percettiva di cui stiamo parlando, le condizioni in cui lo stiamo misurando.

La Legge di Weber si potrebbe trascrivere in italiano così: “*La sensibilità di un sistema sensoriale è inversamente proporzionale all'intensità di uno stimolo di riferimento*”. Se parliamo di proporzionalità inversa però, non dovremmo scrivere la formula in questo modo  $\frac{\Delta I}{I} = k$ , ma come  $\Delta I = kI$ , che si legge  $\Delta I = k$  volte **I**: esprimendo in tal modo la relazione utilizzando la *proporzionalità indiretta*.

#### **APPROFONDIMENTO: PROPORZIONALITÀ DIRETTA E INVERSA**

Due **grandezze** sono **in proporzione** quando esiste una certa relazione tra di esse: possiamo scrivere una in funzione dell'altra e riconoscere qualche caratteristica che resta invariata e riconoscere una **proporzionalità diretta o inversa**.

**Cosa sono la proporzionalità diretta e inversa?** Due **grandezze** sono **direttamente proporzionali** se il loro **rapporto** è sempre **costante**, cioè calcolando la divisione troviamo sempre lo stesso numero. Due **grandezze** sono **inversamente proporzionali** quando il loro **prodotto** è sempre **costante**, cioè calcolando la moltiplicazione troviamo sempre lo stesso numero. Le grandezze direttamente proporzionali e quelle inversamente proporzionali sono molto utili per capire la matematica! Le proporzioni sono utili anche per studiare una **retta** oppure un **iperbole**. Infatti si possono rappresentare le leggi di proporzionalità diretta e inversa rispettivamente con una **retta** o con un'**iperbole**.

Come detto precedentemente questo principio vale in maniera approssimativa in una gamma molto ampia di intensità e tende a non valere se andiamo alle estremità, ma possiamo anche non preoccuparci troppo di questo, anche se qualche volta in psicofisica si cerca di capire cosa succede quando si arriva proprio ai limiti della sensibilità. Ricordiamo infatti il concetto di **soglia assoluta**, che non è altro che una *soglia differenziale* in cui la discriminazione non è tra due stimoli che sono entrambi percepibili, ma **è la soglia differenziale fra uno stimolo percepibile e la sua assenza**. La capacità quindi di arrivare al “grado zero” della sensazione, il punto di partenza della scala psicologica ➔ **prima della soglia assoluta, non c’è percezione anche se potrebbe esserci uno stimolo.**

La legge di Weber è un principio che vale per tantissime situazioni sensoriali, ragionando in termini evoluzionistici, l’evoluzione ha esercitato una pressione nei confronti di organismi a **sviluppare dei sistemi sensoriali che hanno la capacità di modulare la sensibilità in funzione dell’intensità dello stimolo**: *quando lo stimolo è molto debole, l’organismo è molto sensibile mentre quando lo stimolo è molto più intenso, l’organismo evita di sprecare energie non investendo in sensibilità*. Questa caratteristica è fondamentale per un sistema biologico che deve acquisire informazioni dall’ambiente perché gli consente di funzionare in una ampia gamma di situazioni ambientali.

Pensiamo ad esempio all’*uomo* e alla sua *nicchia ecologica*, ovvero l’ambiente in cui l’uomo vive e attua i suoi comportamenti, questo è un ambiente con determinate caratteristiche e dove valgono determinate leggi fisiche più di altre: l’uomo non ha ad a che fare con l’infinitamente piccolo o grande, ha ad a che fare con oggetti, eventi, forze fisiche che si collocano in una **zona intermedia, proporzionata con le sue possibilità d’azione**. Pensiamo a questo ambiente dal punto di vista della luce, focalizzandoci sulla visione ad esempio e quindi sulla luce solare: la *luce solare diurna* varia in una gamma sorprendentemente ampia, varia da tanti punti di vista come nell’intensità (più intensa a mezzogiorno di quanto non sia all’alba o al tramonto), varia nella sua composizione spettrale (per motivi legati all’angolo di incidenza dei raggi rispetto all’atmosfera, la composizione spettrale dell’illuminazione solare si modifica durante il giorno) e poi varia in funzione di molte altre cose come la nuvolosità, le stagioni e anche in funzione del fatto che noi non siamo sempre direttamente esposti alla luce del sole eppure abbiamo una certa costanza del e nel mondo. Il fatto che l’uomo è in grado di “funzionare” in queste diverse condizioni si manifesta nella Legge di Weber, in cui la sensibilità si modifica in funzione di questo.

Quello che conta in percezione non è il valore assoluto di uno stimolo, ma le **relazioni** → Questo viene anche chiamato **principio di determinazione relazionale** ed è l'idea secondo la quale la maniera giusta di pensare a uno stimolo non è come se fossimo dotati degli strumenti fisici di misura che ci forniscono valori assoluti, ma come un rapporto rispetto ad uno stimolo di riferimento.

## La Funzione di Weber-Fechner



Il vero inventore della **psicofisica** fu **Gustav Theodor Fechner**, il quale fu il primo a porsi esplicitamente il problema della psicofisica, ovvero quello di trovare e descrivere una **relazione funzionale** tra lo *psichico* e il *fisico*, cercando di definire come la “sfera psichica” vari in funzione della “sfera fisica”.

**È possibile trovare un modello matematico che descriva questa relazione?**

Questo lo possiamo pensare mettendo in grafico due quantità, una *dimensione psicologica* (**S**) e una *dimensione fisica* (**I**) e chiederci se è possibile descrivere con una legge matematica come cresce la dimensione psicologica al crescere dell'intensità dello stimolo.

La dimensione psicologica viene chiamata da Fechner **sensazione**, ma noi dobbiamo pensare a questa **S** in termini più astratti: è una *qualsiasi dimensione psicologica* e il principale problema risiede proprio nel **quantificare una dimensione psicologica**. Infatti l'uomo è in grado di **misurare gli stimoli fisici** (ad esempio in grammi, in metri o in qualche tipo di unità di misura illuminotecnica come il lux, il watt; in psicofisica di solito l'intensità luminosa si misura in candele per metro quadrato ( $\text{cd}/\text{m}^2$ ), che è una misura che fa riferimento all'effetto che ha un'intensità luminosa sull'occhio, l'equivalente di una certa intensità standard rispetto a un'unità di spazio, la cosiddetta “luminanza”), ma è molto meno ovvio misurare la dimensione psicologica per diversi motivi, come ad esempio l'esperienza percettiva cosciente.

Fechner partì dalla Legge di Weber la quale segue il principio generale **che la sensibilità cresce rapidamente quando l'intensità dello stimolo è piccola, mentre diventa molto meno buona/fine quando lo stimolo è molto grande** e affermò che proprio per questo motivo la forma della funzione sarebbe dovuta essere quella di una **curva logaritmica**.

### il problema di Fechner



Siccome la sensibilità è l'inverso della soglia, questo significa che se io penso a quanto mi dovrò spostare sull'asse delle X per passare da una intensità di riferimento a quella che è discriminabile, cioè quanto sarà grande un  $\Delta I$  sull'asse delle x, allora questo  $\Delta I$  sarà piccolo (sensibilità alta) quando sono sulla parte di sinistra del grafico, ma diventerà sempre più grande quanto più mi sposto verso destra.

Uno dei fondamenti sui cui si regge la psicofisica è l'idea che ha avuto Fechner, cioè che si possa creare un grafico mappando come i diversi livelli di  $\Delta I$  su X, corrispondono a uguali step di discriminabilità a partire dal grado 0 della sensazione.

Partendo dalla soglia assoluta dobbiamo chiederci di quanto incrementare  $I$ , per far sì che ci sia l'inizio della sensazione, l'inizio del fatto che ci accorgiamo che c'è qualcosa.

Questo è il primo step della sensibilità, a partire da quello dobbiamo chiederci "***Di quanto devo ancora incrementare?***" passando così al secondo step e così via. Questi step vengono anche chiamati **differenze appena percepibili, JND (Just-Noticeable Differences)**. Nella funzione psicofisica si mettono in relazione i diversi step psicologici in unità **JND** con i diversi incrementi  $\Delta I$ , la soglia differenziale.

Il punto fondamentale di cui si rese conto Fechner è che questa funzione non poteva essere una retta, non poteva essere un rapporto di proporzionalità diretta fra **JND** e  $\Delta I$  perché quanto più mi sposto verso destra sull'asse delle x (cioè quanto più mi sposto verso l'alto sull'asse delle y) tanto più aumenta  $\Delta I$  però gli incrementi psicologici sono sempre gli stessi. L'incremento sull'asse delle Y è un incremento in cui le distanze sono sempre uguali e la differenza è sempre di una JND (ad esempio per andare da 1 a 2 o da 20 a 21 è sempre una JND di differenza). Da un punto di vista fisico è completamente diverso perché per andare da 1 a 2 basta un incremento molto piccolo, ma per andare da 20 a 21, che è il limite estremo della sensibilità di molti canali sensoriali, è necessario un incremento molto grande.

*La funzione psicofisica quindi deve essere una curva con una certa forma.*

È importante non confondere una *funzione psicofisica* con una *funzione psicometrica*:

- La **funzione psicofisica** solitamente è una legge, una funzione che cerca di rispondere alla domanda su come lo psicologico varia in funzione del fisico;
- La **funzione psicometrica** si usa per descrivere delle soglie.

In questa situazione entra in gioco il concetto di **soglia**, quindi potremmo dire che **la funzione psicometrica serve per creare delle funzioni psicofisiche**.

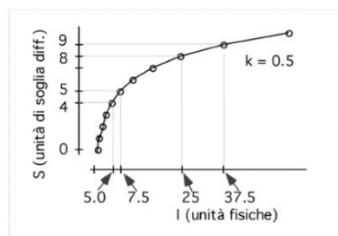
Quanto detto precedentemente può essere rappresentato in questo grafico: vi sono le **unità fisiche** che corrispondono a **JND** sempre più grandi sull'asse delle X mentre sull'asse delle Y vi sono gli incrementi psicologici in termini di **S** e in quei riquadri non c'è altro che la legge di Weber.

Se noi partiamo da una intensità fisica che chiamiamo **I<sub>1</sub>**, che è di qualche livello sopra **I<sub>0</sub>** (in questo grafico è l'intensità fisica che corrisponde ad assenza di sensazione) e andiamo da **I<sub>0</sub>** a **I<sub>1</sub>** noi troviamo la soglia assoluta.

Per andare da **I<sub>1</sub>** al livello successivo dobbiamo moltiplicare **I<sub>1</sub>** per **k** (la costante di Weber  $\Delta I = k$  volte **I**); di quanto dobbiamo spostarci sull'asse delle X sarà **k** volte **I<sub>1</sub>**. Se parto da un altro livello verso destra che chiamiamo **I<sub>2</sub>**, qui sarà **k** volte **I<sub>2</sub>** che è molto più grande di **k** volte **I<sub>1</sub>**.

Da questi pochi punti possiamo già vedere che questa funzione non è una retta, **ma è una curva che all'inizio si impenna rapidamente e poi si appiattisce, ovvero una curva logaritmica**.

**funzione di Weber-Fechner  $k = 0.5$**



In questo grafico è stato immaginato un sistema sensoriale in cui  $k = 0.5$  (che non esiste, ma ci permette di seguire il ragionamento).

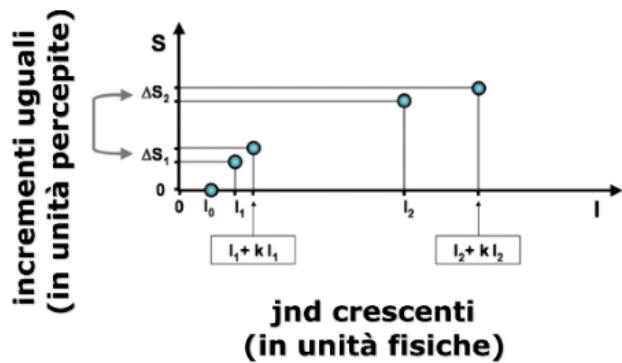
È stato individuato un punto di partenza sull'asse delle X per poi aggiungere la metà di quello, **k** volte **I** di partenza: se parto da 5, viene aggiunto 2.5 si giunge a 7.5. Questo corrisponde al passaggio dalla quarta alla quinta **JND**.

Questa è una curva in cui a uguali incrementi sull'asse delle Y corrispondono incrementi sempre maggiori sull'asse delle X, **ma di quanto maggiori?**

La regola è che a uguali incrementi sull'asse delle Y corrispondano **uguali rapporti** sull'asse delle X. Questa è la definizione di una **funzione logaritmica**, in cui una curva cresce molto rapidamente quando X è piccolo e poi sempre più lentamente. È una curva che tende asintoticamente (tende all'infinito) a diventare piatta; non diventa mai piatta però al crescere di **I**, queste  $\Delta I$  diventano sempre più grandi all'infinito, quindi la curva si appiattisce e la sensibilità si riduce fino a diventare assente.

Si dice che questa curva è “**accelerata negativamente**”, perché cresce, accelera molto all'inizio e poi sempre di meno.

## la soluzione



Questo è il significato della **Legge di Weber-Fechner** che si chiama così perché è una esplicitazione di ciò che era già implicito nel lavoro di Weber.

Nella legge di Weber-Fechner:

- **S** è la sensazione, qualsiasi cosa che riguarda la dimensione percepita psicologica;
- **c** è una costante di proporzionalità. Essa non è la costante di Weber ma *dipende* da quest'ultima, infatti **c** è il reciproco di  $\log(1+k)$  dove **k** è la costante di Weber.

Qui vediamo come si arriva a scrivere la legge di Weber-Fechner a partire dalla legge di Weber.

Fechner scrisse la Legge di Weber come una equazione differenziale in cui:

## da Weber a Fechner

Questo si può riscrivere integrando i membri di questa relazione, in questo modo:

$$ds = c dI / I$$

**"a incrementi uguali in unità di sensazione corrispondono incrementi sempre maggiori in unità fisiche"**

Qui troviamo **c**, la costante di proporzionalità e anche **C** che è una costante di integrazione di cui possiamo trovare il valore con questa formula, in cui invece di scrivere l'equazione per un qualsiasi valore di **I**, la scriviamo per la soglia assoluta, ponendo **I** = soglia assoluta; questo perché la soglia assoluta è

l'inizio della scala psicologica quindi la poniamo uguale a 1.

$$C = -c \log(I_0)$$

Questo valore va a sostituire **C** nell'equazione precedente, viene inserito a destra e quindi scriviamo la differenza tra due logaritmi, dove compare la stessa costante di proporzionalità **c**.

$$s = c \log(I) - c \log(I_0)$$

## "legge" di Weber-Fechner

$$s = c \log(I)$$

**I'intensità percepita è proporzionale al logaritmo dell'intensità fisica dello stimolo**

**c è la costante di proporzionalità**  
 $c = 1 / \log(1 + k)$

A questo punto possiamo riscrivere questa formula portando  $c$  fuori dalla differenza tra i due logaritmi, scrivendo che la sensazione è uguale a  $c$  che moltiplica la differenza tra i due logaritmi.

La differenza fra il logaritmo di due numeri è uguale al logaritmo del rapporto tra quei due numeri:  $\log(I) - \log(I_0)$  è la stessa cosa di  $\log(\frac{I}{I_0})$ .

L'espressione abituale della legge di Weber-Fechner è

$s = c \log(I)$ , ma la vera espressione formale della legge di Weber-Fechner è  $s = c \log(\frac{I}{I_0})$ , un modo di scriverla in cui per convenzione si dice che la soglia assoluta è pari a 1 così che il rapporto fra un numero e 1 non è altro che quel numero. Questo spiega come a partire dalla legge di Weber io posso scrivere una funzione psicofisica come una funzione logaritmica. In conclusione:

$$\cdot \frac{\Delta I}{I} = k \quad \text{Weber}$$

$$s = c \log(I/I_0) \quad \text{Fechner}$$

$$s = c \log(I) \quad \text{Fechner*}$$

## Applicazioni pratiche

Immaginiamo di dover costruire uno *stereo*, uno stereo deve essere dotato di un qualche tipo di controllo per il volume, come ad esempio una manopola che serve appunto a regolare l'intensità del suono percepita. Questo è un problema di psicofisica perché girando la manopola andiamo a cambiare i setting di un amplificatore, modifichiamo l'intensità fisica del suono che esce da questo strumento.

Ma agli psicologi della percezione interessa poco l'**intensità fisica** perché questo strumento serve in quanto si **interfaccia con un utente umano** che ha un sistema uditivo fatto in una certa maniera, che segue alla legge di Weber. Il problema deve tener conto di questo perché se noi avessimo un rapporto lineare tra quanto giriamo la manopola e l'intensità del suono, succederebbe che se l'audio fosse molto basso, girando di due gradi la manopola, questo produrrebbe un incremento abbastanza rilevante, accorgendosi subito che il volume è aumentato. Mentre se stiamo già tenendo l'audio molto alto, lo stesso giro di due gradi magari non produce nessun effetto percepibile.

## da Weber a Fechner

$$s = c \log(I) - c \log(I_0)$$

$$s = c [\log(I) - \log(I_0)]$$

$$s = c \log(I/I_0)$$

↑  
espressione formale della legge di W-F

Costruire uno stereo che non tiene conto della costituzione del nostro apparato uditivo rende quell'oggetto poco funzionale in quanto essendo uno strumento realizzato *per l'uomo* quello che è importante è che il **controllo del volume corrisponda alla nostra sensazione, non a quello che succede da un punto di vista fisico**. Un costruttore vuole avere la sensazione che la manopola ci fa spostare di uguali incrementi sulla scala delle **JND**, sulla scala delle S. A questo punto dobbiamo applicare la Legge di Weber-Fechner, ossia dobbiamo far sì che quando giriamo la manopola di due gradi e l'intensità è bassa questo produca un incremento piccolo, mentre se l'intensità è alta l'incremento deve essere maggiore.

**Utilizzando R**, la costante di Weber della dimensione psicologica chiamata **loudness** (intensità percepita di un suono) è pari a  $k = 0.05$ . Non c'è una vera traduzione in italiano del termine “*loudness*”, potremmo usare il termine “volume”, ma in realtà il volume non c'entra molto con questo concetto, in quanto il concetto di volume fa riferimento, ad esempio, allo spazio occupato da un solido in termini di tre dimensioni. **Il suono** che noi percepiamo da un punto di vista fisico corrisponde a un'onda che si propaga in un mezzo elastico, questa onda può essere più o meno complicata.

Quando si parla di un **tono** si parla di un'onda semplice, che possiamo descrivere con una funzione sinusoidale che dipende da pochissimi parametri uno dei quali è la sua ampiezza  $\lambda$  (più è ampia più l'energia dell'onda è alta) e l'altra è la sua frequenza  $f$  (4kHz), quanto velocemente l'aria passa da compressa a espansa. I **suoni** vanno da molto **bassi**, i quali hanno una frequenza piccola, a quelli più **acuti**, che hanno un'alta frequenza. È importante specificare che un tono ha una particolare frequenza in quanto la sensibilità del nostro apparato acustico cambia in funzione di questa: siamo più sensibili alle frequenze intermedie (quelle a cui parliamo, di circa 4-10 kHz) e meno sensibili ai suoi molto acuti o molto bassi.

Fissato questo possiamo dare una legge di Weber-Fechner per questo particolare suono.

Dopo aver fissato **k**, possiamo trovare **c** che è il reciproco del  $\log(1+k)$ , in questo caso però qui è specificato il log in base 10, si potrebbe usare un'altra base perché non c'è nessuna differenza da un punto di vista concettuale; **x** è una sequenza di numeri arbitrari che rappresenteranno **I** che in questo caso corrisponde all'ampiezza di quell'onda sonora, che va da 1 decimo a 100 a step di 1 decimo. Nel caso di **y** non si fa altro che applicare la legge di Weber-Fechner così che  $y = c * \log_{10}(x)$ .

### **loudness di un tono a 4 kHz**

```

k <- 0.05
c <- 1/log10(1 + k)
x <- seq(0.1, 100, by = 0.1)
y <- c * log10(x)
plot(x, y, type = "l", xlab =
  "I (physical sound pressure
  level)", ylab = "s (loudness,
  logI)")
text(60, 0, "k = 0.05")

```

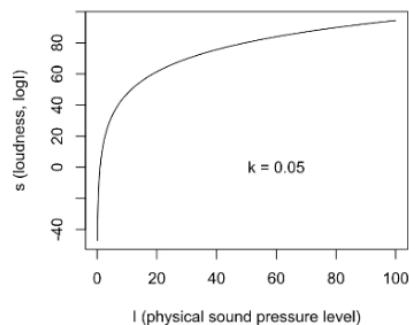
Plotando questo script su R, compare questo:

### **Questa è la Legge di Weber-Fechner per la loudness di un tono a 4 kHz.**

La quantità sull'asse delle y è proporzionale a quello che chiamiamo **decibel**, [l decibel (simbolo dB) è la decima parte del bel (simbolo B):  $10 \text{ dB} = 1 \text{ B}$  ed è un'unità di misura logaritmica del rapporto fra due grandezze omogenee (es. due potenze, due pressioni, due potenziali elettrici)], ovvero una scala logaritmica di intensità percepita che è basata sulla Legge W-F.

Alcuni valori di S sono negativi, ma questo è dovuto al fatto che è il logaritmo di un numero compreso tra 0 e 1, come un numero frazionario, è un valore negativo. In unità di sensazione si potrebbe riscrivere come corrispondente a sensazioni in cui il rapporto è molto piccolo, compreso tra 0 e 1.

### **loudness di un tono a 4 kHz**



### **Misurare una soglia**

Per ricavare empiricamente una di queste funzioni psicofisiche bisogna **misurare una soglia**: si deve trovare quanto vale quel  $\Delta I$  in funzione di diversi livelli di  $I$  e poi metterli in grafico e vedere se i dati corrispondono alla previsione teorica.

In psicofisica infatti spesso si misurano delle soglie non per cercare quali siano i limiti della sensibilità, ma perché questo è un modo di valutare come la sensibilità varia in funzione dell'intensità dello stimolo di riferimento e andare a verificare per esempio se è proprio come prevede la legge di Weber-Fechner. In realtà la legge di W-F è valida per molte situazioni pratiche, come ad esempio la loudness, ma non è così per altre situazioni.

## Come si misurano le soglie?

È utile ragionare su due modelli diversi di come si potrebbe immaginare una soglia.

Questi due modelli sono rappresentati come dei grafici, due modi di disegnare una funzione sigmoide, due sigmoidi che sono tra l'altro due funzioni psicometriche.

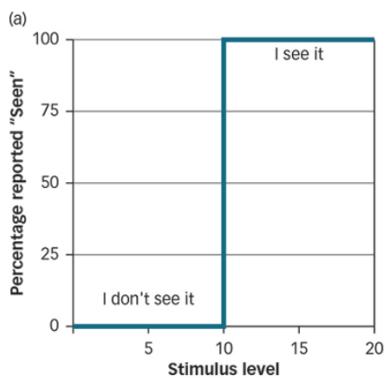
## Cosa rappresentano queste funzioni psicometriche?

Sull'asse delle X c'è il livello di *intensità dello stimolo* (proprio come la funzione psicofisica di Weber-Fechner) però sull'asse delle Y c'è una misura molto diversa, ovvero la *misura di prestazione*.

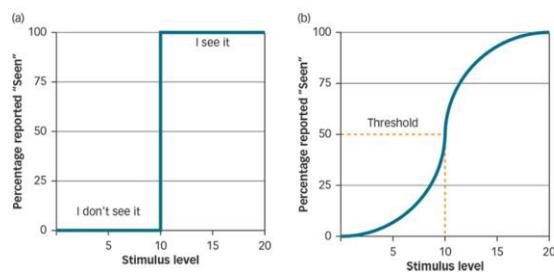
Bisogna immaginare di chiedere al soggetto qualche tipo di risposta in funzione dell'intensità di **I** su X: in questo caso per esempio potremmo immaginare di cercare una soglia assoluta e la cosa più facile che potremmo fare, anche se non migliore dal punto di vista metodologico, consiste nel mostrare una serie di stimoli a diversi livelli di **I**, molto bassi oppure più alti e chiedere al soggetto se ha visto qualcosa oppure no.

Questo è il **compito yes/no** in cui c'è sempre uno stimolo che viene presentato e ci si "fida" del soggetto che risponde di averlo visto o meno. Il limite di questo metodo è la **soggettività**, in quanto quello che una persona considera *evidenza* può essere diverso da quello di un'altra persona e dipende dalle differenze interindividuali tra soggetti differenze. Possiamo quindi ottenere una misura in funzione dei diversi livelli di **I** della percentuale di volte che, in funzione di quel particolare **I**, il soggetto risponde di "aver visto" una differenza.

Immaginiamo come ad esempio il livello 1 di **I** viene presentato 50 volte e di quelle 50 vado a vedere quante volte il soggetto ha risposto con yes, siccome 1 è molto basso probabilmente non molto spesso ma al crescere di **I** ovviamente le cose cambieranno. Questi grafici riflettono delle misure empiriche e queste sono due possibili ipotesi su come potrebbero formare una *funziona a forma di S*.



## soglia ideale e soglia reale



Se la soglia fosse una vera soglia cioè un *confine netto*, allora dovrei osservare una sigmoidea che in realtà è uno scalino: in una situazione ideale, la **soglia ideale sarebbe una situazione in cui finché non arrivo al valore critico di soglia ideale io non vedo mai niente e poi improvvisamente superando la soglia, a quel punto vedrò sempre qualcosa**. Questa situazione non esiste in natura, non c'è un salto netto.

In realtà quello che noi osserviamo in una situazione come questa variando **I** e andando a vedere la proporzione di “yes”, è trovare una funzione che ha effettivamente la forma di S, che significa che al crescere di **I** varia la probabilità che io dica “yes”, varia nel senso che quando **I** è molto piccolo la probabilità è bassa, quando **I** cresce la probabilità aumenta.

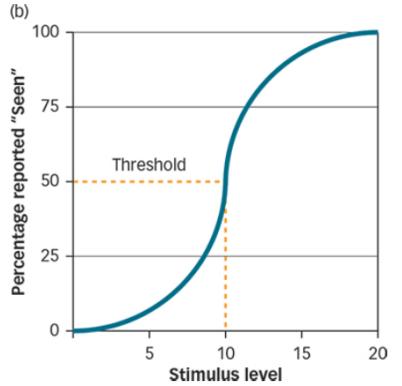
Questo andamento naturalmente dipenderà dalla **propria sensibilità**. Se si è molto sensibili vuol dire che la S tenderà a somigliare a uno scalino smussato, ma dato che essere molto sensibile significa percepire differenze piccole, vuol dire che è come se ci fosse una zona di incertezza.

Quando **I** è molto grande o molto piccola non c’è incertezza, dico sempre no o sempre sì, mentre nella *parte centrale* c’è una **zona di incertezza** la cui ampiezza dipende dalla sensibilità: più è alta la sensibilità e più la sigmoidea tenderà a essere piatta e a somigliare sempre più a una retta e quindi la sigmoidea tenderà a somigliare a uno scalino.

Lo stesso concetto di soglia quindi non va pensato come un singolo valore, un numero che una volta superato c’è la sensazione, ma va pensato come una **funzione di probabilità**, ovvero una funzione psicométrica. Dal punto di vista pratico ciò che determiniamo è una funzione di probabilità che descrive anche un certo livello di incertezza che non è un singolo numero.

### **Da cosa dipende l’incertezza?**

- potrebbe dipendere da *come viene misurato* un parametro, infatti ogni misura è soggetta all’**errore casuale**, quindi ogni misura è diversa dalla precedente e quindi la probabilità che il soggetto dica di “sì” ad ogni prova dipenderà anche da questo errore casuale, motivo per cui non sarà mai effettivamente uguale di prova in prova;
- potrebbe dipendere dal fatto che all’interno di quello che sto misurando nel sistema sensoriale ci sono anche delle variazioni sensoriali, il cosiddetto **rumore neurale**, perché le frequenze di scarica dei neuroni sono soggette anch’esse ad una certa quota di variazione sensoriale;
- potrebbe dipendere dal fatto che la soglia stessa sia variabile, cioè che il famoso **limen** ovvero questo limite che di volta in volta devo attraversare per scatenare una sequenza di eventi che poi produce una risposta cosciente, magari non sia fisso nel tempo, ma cambia in funzione di altre cose interne al sistema.



**La soglia è di fatto sempre una funzione di probabilità, una stima del mio livello di incertezza nel fare una determinata operazione, che dal punto di vista del soggetto è un'operazione di discriminazione nel saper dire se c'è o meno qualcosa in questo particolare esempio. Convenzionalmente si dice che la soglia corrisponde, nella funzione psicometrica, all'intensità dello stimolo (“*stimulus level*”) che corrisponde al 50% delle risposte discriminative. Questo non è sbagliato nel senso che il 50% corrisponde a una certa definizione di quella che è la *massima incertezza*.**

***Come facciamo ad essere sicuri che quella curva sia simmetrica?***

Sopra il 50% e sotto il 50% la forma è la stessa, meno di una deflessione, ma potrebbe anche non esserlo se ricordiamo **Kahneman e Tversky**. Se è simmetrica significa che corrisponde al 50% di volte in cui lo stimolo viene visto e al 50% di volte in cui lo stimolo non viene visto quindi la massima incertezza.

Potremmo dire da un punto di vista pratico se ci serve un numero e non una funzione allora il numero che ci conviene prendere è il numero 10, perché lo stimulus level è pari a 10 che potrei chiamare la soglia.

Bisogna però fare attenzione perché in questo caso è come prendere un campione di misure con la sua distribuzione, con la sua forma, dispersione e poi all'atto pratico ci serve sapere qual è la *tendenza centrale* ovvero la mediana, la moda o la media geometrica che sembra sensata per quel particolare problema.

**La media però non è una distribuzione, la nostra osservazione è la distribuzione, non la media.**

La media cattura certe caratteristiche di quella distribuzione così come la soglia, intesa come singolo numero, cattura un qualcosa di quella sigmoidea e allo stesso tempo trascura altre informazioni.

In letteratura spesso molti autori non prendono come valore di soglia il 50%, ma per esempio il 60% o l'80% questo perché una volta capito che la soglia è una funzione e non un singolo valore, dipende da che idea abbiamo in mente di *ragionevole incertezza*: prendere come valore di soglia l'80% significa dire che la soglia è quel valore in cui i soggetti sono abbastanza sicuri di aver visto qualcosa, il grado 0 della sensazione.

Possiamo ricavare questa curva empiricamente, ma è interessante trovare una legge matematica che riassuma questi dati.

***Come faccio a trovare i parametri di una funzione a forma di S in modo da poter “fittare” la funzione a forma di S con i miei dati così da giungere a una caratterizzazione ancora più generale della soglia?***

La risposta è che ciò dipende dal fatto che i dati siano suggestivi di una funzione simmetrica o non simmetrica: se è una funzione simmetrica è molto facile perché già sappiamo qual è la funzione che per eccellenza ha la forma di S. Di solito questo tipo di dati si possono fittare in modo abbastanza soddisfacente usando una funzione cumulativa di una distribuzione normale.

In R, data una distribuzione è molto facile ricavare la cumulativa, la funzione pnorm che da l'altezza di questa curva in funzione di una x. Naturalmente potrebbe anche non essere una distribuzione normale, potrebbe essere usata anche la funzione cumulativa di una distribuzione logistica che è un po' diversa.

## I metodi di Fechner

Fechner descrisse tre metodi per misurare soglie:

1. **Aggiustamento;**
2. **Limits;**
3. **Stimoli costanti.**

In realtà, sono tutti e tre variazioni su come si può stimare la forma di quella funzione psicometrica.

Vengono definiti come metodi di **psicofisica indiretta** perché la funzione psicofisica viene ricavata indirettamente misurando soglie. Esistono anche metodi di **psicofisica diretta** in cui questo passaggio si salta e si possono ottenere funzioni in modo più rapido, anche se si tratta sempre di stimare la sensibilità.

### **È possibile misurare la sensazione con tecniche di neuroimaging?**

Una cosa è chiedersi se noi possiamo avere accesso al vissuto cosciente, qui e ora, fatto da un insieme di caratteristiche qualitative della nostra esperienza. Questo aspetto qualitativo della coscienza è quello che viene chiamato *il problema difficile* (hard problem) della coscienza o i *qualia*. Guardando una fMRI c'è un salto logico totale fra i colori con cui rappresenti certe statistiche fatte sulle attivazioni di diversi vox esaminati e i qualia dall'altra parte.

Altro è chiedersi se possiamo stabilire delle relazioni fra la struttura della nostra coscienza (ad esempio il fatto che un oggetto ci sembra diverso dall'altro, più pesante, più grande, più intenso questa è una relazione di natura strutturale) e una qualche caratteristica strutturale che misuriamo in una risonanza o in una EEG.

Una cosa che si fa solitamente lavorando con i potenziali evocati con la EEG è quello di mettere in relazione i tempi di risposta. Questo tipo di relazioni si possono osservare.

## **Lezione 3 – Metodi di Fechner – Esperimento di Pirenne – Illusione di Muller-Lyer: PES e PEO**

### **I metodi di Fechner**

Tra i contributi di **Fechner** si riconosce quello di aver provato a formalizzare dei metodi per **misurare le soglie**. Tali metodi di misurazione indiretta (psicofisica indiretta) sono ancora utilizzati, eventualmente con piccole varianti, mantenendo l'idea di base: **tentare di fare una stima della funzione psicometrica\***.

- **Funzione psicofisica:** un esempio di funzione psicofisica è dato da una curva logaritmica – legge che mostra come l'intensità della sensazione, con step uguali di soglie, corrisponde ad un cambiamento di intensità fisica.
- **Funzione psicometrica\*:** utilizzata per costruire la funzione psicofisica – serve per misurare le soglie. **La soglia non è un numero ma una funzione di probabilità, di forma sigmoide, che rappresenta la probabilità che un soggetto emetta una certa risposta discriminativa in funzione del cambiamento dell'intensità fisica.** *Ragionando su una soglia assoluta partiamo da una I molto bassa, con risposta “NO” costante, e passando da una soglia di incertezza, si arriva gradualmente ad una soglia di certezza con risposta “SI” costante.*

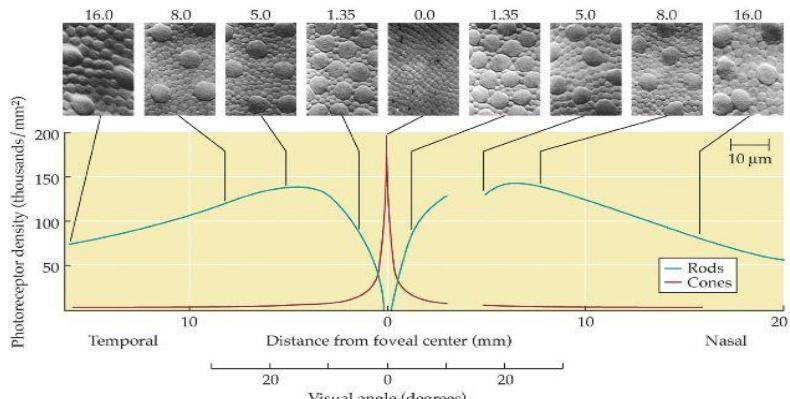
### **Metodi:**

1. Stimoli costanti;
2. Limiti;
3. Aggiustamento.

#### **1. Metodo degli stimoli costanti**

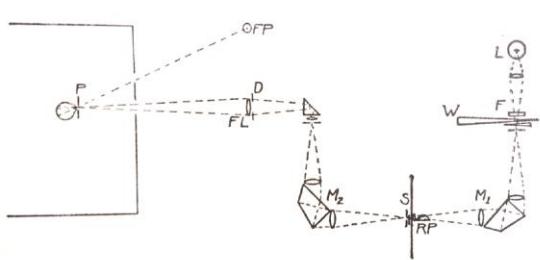
Il metodo degli stimoli costanti **consiste nello scegliere un set di intensità (valori “I”), intorno alla soglia, in un compito yes-no.** Si procede chiedendo al soggetto se vede o meno lo stimolo. Gli stimoli sono presentati in ordine casuale e la probabilità che il soggetto risponda è data dal rapporto tra il numero di risposte affermative (yes) e il numero totale di presentazioni. Questa probabilità è sostanzialmente 0 per un disco di intensità bassa corrispondente a poche decine di quanti di luce. Crescendo tra 50 a circa 80 quanti la probabilità aumenta e tra 100 e 200 quanti è presente una zona di incertezza, in cui il soggetto talvolta risponde “yes” e talvolta “no”. Sopra i 200 quanti, corrispondente ad un'intensità notevole, si ha quasi certezza di una risposta affermativa costante.

Il criterio di probabilità di risposta affermativa scelto da **Pirenne** è del 60%, nel grafico del Professore è rappresentato al 50% (solitamente usato nei manuali, il valore è infatti arbitrario). **L'informazione completa è data dall'intera funzione sigmoide** e, mentre la zona nera è data dal fit dei dati, la zona grigia rappresenta la zona di incertezza del fit (variazione data dal campionamento).



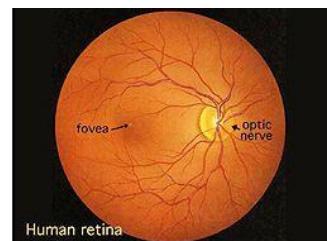
## Esperimento di Pirenne

L'esperimento di Pirenne prevede un compito yes-no, in cui a seguito della presentazione di uno stimolo, il soggetto deve dire se lo ha percepito o meno. Con tale esperimento Pirenne intende indagare il *limite assoluto*, il più basso possibile (limite inferiore di sensibilità dei bastoncelli), della sensibilità visiva. La sensibilità cambia a seconda dell'ambiente (a seconda che sia molto o poco illuminato) – nella retina sono presenti **coni**, fotorecettori specializzati per la visione diurna, e **bastoncelli**, specializzati per la visione notturna per cui la massima sensibilità possibile è di notte, in condizione con poca luce. Prima di poter iniziare l'esperimento è necessario porre il soggetto nelle condizioni ottimali, facendolo restare per almeno 30 minuti nella stanza al buio. Il nostro sistema visivo si adatta a condizioni di illuminazione diverse (adattamento neurale all'intensità luminosa), per cui se il soggetto è al buio, i coni si spengono e i bastoncelli gradualmente aumentano la propria sensibilità. L' *adattamento al buio* è un fenomeno lento (circa 20/30 min.), che termina al raggiungimento della massima sensibilità.

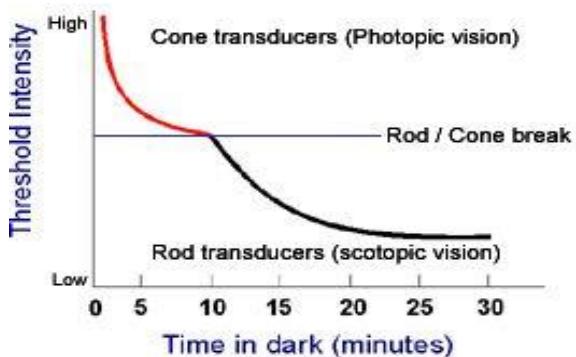


Il soggetto fissa un punto diverso (“FP”) dal disco (“D”), che è l’oggetto di valutazione. Il centro della visione, ossia quello utilizzato quando si fissa qualcosa, è infatti la fovea (il punto di massima visione/acuità diurna), in cui però non sono presenti bastoncelli.

Al buio la massima risoluzione possibile non si ha fissando un punto dato, ma spostando lo sguardo leggermente di lato, data la superiore concentrazione di bastoncelli in tale area, e quindi migliore risoluzione.



## Rappresentazione del fenomeno dell'adattamento al buio:



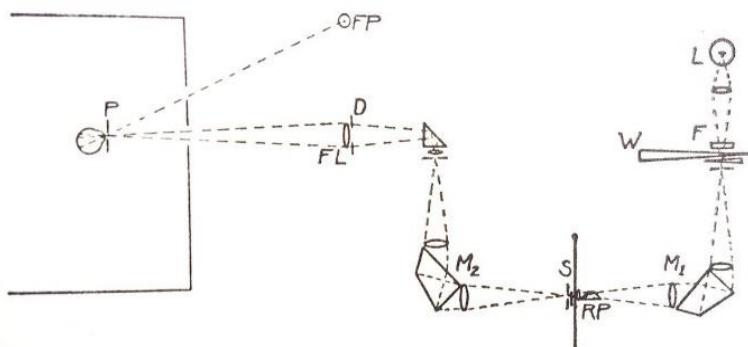
La rappresentazione è costituita da un *plot della sensibilità* in funzione del tempo passato al buio. Per la precisione, è osservabile la misura di una soglia. Ad una diminuzione della soglia corrisponde un aumento della sensibilità. La curva tende a scendere e ad appiattirsi fino ad arrivare, dopo circa 10 minuti, ad un “gomito”, per poi scendere di nuovo molto rapidamente.

Il “gomito” della curva, così osservabile, corrisponde al punto in cui i coni si spengono ed iniziano ad entrare in gioco i bastoncelli. Si verifica di fatto l’adattamento dei due fotorecettori. Nella zona prima del “gomito” è possibile osservare che la sensibilità tende ad appiattirsi (zona mesotica) poiché rappresenta una zona in cui i bastoncelli iniziano a funzionare.

\*(visione notturna: scotopica; visione diurna: fotopica)

## Descrizione dell'esperimento:

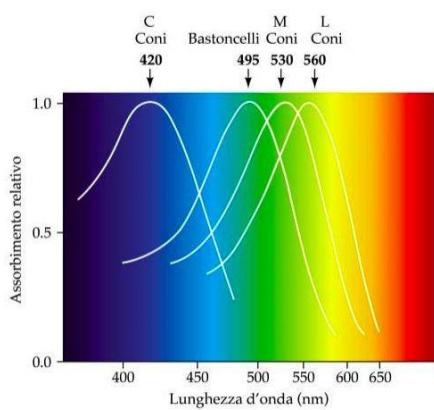
Partendo da sinistra: la sagoma rettangolare rappresenta la stanza in cui è presente il partecipante (rappresentato da un occhio e identificato con la lettera “P”). Il partecipante ha un occhio bendato (esperimento fatto monocularmente) e ha due buchi davanti a sé. Uno è collegato al punto di fissazione (“FP”).



Attraverso il secondo buco è possibile vedere un disco (“D”) illuminato (“FL”, Foot-Lambert: misura di intensità luminosa, attualmente in disuso).

L’apparato ottico restante, che serve a controllare come viene illuminato il disco, ha origine da una lampada (“L”), da cui esce una luce fissa di cui è possibile regolare l’intensità, in modo tale da ottenere una luminanza molto debole. A tale scopo è presente un filtro (“F”), che toglie luce e un prisma (“W”).

Il prisma, grazie al *fenomeno della rifrazione*, devia i raggi luminosi il cui grado di deviazione dipende dalla lunghezza d'onda del raggio.

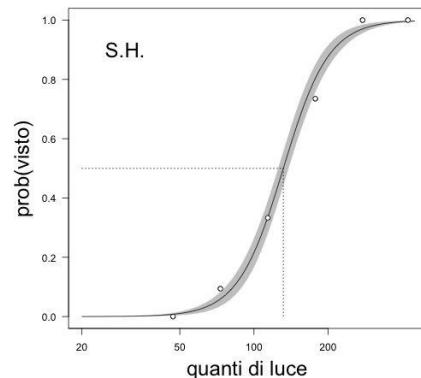


Sfruttando tale fenomeno, è possibile costruire un “*filtro mono-cromatico*”, ossia un filtro che seleziona solo la luce all’interno di una gamma di lunghezze d’onda molto strette (fa passare solo la luce a circa 500 nm). La luce che arriva al disco (“D”) ha quindi lunghezza d’onda precisa e intensità controllabile. La lunghezza d’onda utilizzata nell’esperimento (500nm) corrisponde al punto di massima capacità di assorbimento da parte del foto-pigmento dei bastoncelli.

Il grafico rappresenta una funzione psicometrica che descrive i dati di una sessione sperimentale di Pirenne sul soggetto S.H. (uno dei 3 soggetti sperimentali: Pirenne e i suoi due collaboratori).

La funzione, una sigmoide, è data dal fit dei dati di Pirenne rappresentati dai sei punti, corrispondenti all’intensità luminosa di sei dischi misurata con il metodo degli stimoli costanti.

\**Esistono due soglie possibili: una comportamentale/reale (100-200 quanti) e una del fotorecettore/bastoncello (1-10 quanti), e nel concetto di efficienza vengono presi in considerazione entrambi.*



**La conclusione** alla quale è giunto Pirenne è che la soglia assoluta dei bastoncelli è posta tra i 100 e 200 quanti di luce (*quanto*: pacchetto discreto di energia). Tale soglia è tuttavia misurata su dati comportamentali\* (si misura dalla risposta del soggetto nel compito yes-no) e conduce al successivo quesito: **quale è la quantità effettiva di quanti di luce che arrivano ai bastoncelli?**

Una quota viene infatti persa nel viaggio dello stimolo dal disco all’occhio, nell’attraversamento della cornea, nell’attraversamento del liquido che riempie l’occhio e solo una porzione riesce ad arrivare alla retina. Anche nella retina la luce deve attraversare altri strati di cellule: alla fine del percorso, Pirenne ha stimato che il numero di quanti di luce che riesce effettivamente ad arrivare su 1 bastoncello si aggira tra 1 e 10 quanti. Ciò indica che in condizioni ideali la sensibilità visiva è il massimo possibile (1 quanto di luce).

Sottostante a questo ragionamento vi è una distinzione che si è soliti attuare in psicofisica, tra **sensibilità** ed **efficienza**: **Sensibilità (sensitivity)**: indica la prestazione di un osservatore in un compito di detezione o discriminazione. **Efficienza (efficiency)**: dati i vincoli fisici e biologici (la luce non può essere meno di 1 quanto, la retina è fatta in un certo modo...), la prestazione reale di un osservatore viene espressa in proporzione alla prestazione ideale, un modello teorico di come un compito potrebbe essere fatto in maniera ottimale in condizioni ideali.

---

## 2. Metodo dei limiti

Intensity	Trial Number									
	10	9	8	7	6	5	4	3	2	1
10	Y			Y		Y		Y		Y
9	Y			Y		Y		Y		Y
8	Y			Y		Y		Y		Y
7	Y			Y		Y		Y		Y
6	Y			Y		N		Y		Y
5	Y			N	Y		Y	Y		Y
4	N	Y		N		N	Y	Y	Y	N
3		N		N		N	N	N	N	
2		N		N		N			N	
1		N		N		N			N	
Crossover	4.5	3.5	5.5	4.5	6.5	4.5	3.5	3.5	3.5	4.5

Threshold = mean crossover = 4.5

Dato un numero scelto di valori di “I” (intensità), si presentano ripetutamente al soggetto delle serie ascensioni e discensioni, in maniera ordinata (una serie ascendente è data ad es. partendo da un valore molto basso di “I” e salendo verso la soglia). **Si procede chiedendo al soggetto se ha visto o meno lo stimolo e si misura dove avviene un'inversione di giudizio (passaggio da una serie di “yes” ad un “no” e viceversa).** Il giudizio si inverte nella zona di incertezza (crossover) e se si procede per un certo numero di volte con le serie è possibile osservare che queste inversioni tenderanno a collocarsi sempre più frequentemente nella zona di incertezza. Quindi, di fatto, con una media di questi punti di inversione e la relativa dispersione, si ha un’idea *dell’ampiezza della zona di incertezza*.

## 3. Metodo dell’aggiustamento

**Nel metodo dell’aggiustamento il soggetto controlla direttamente l’intensità dello stimolo.** Attraverso un apposito strumento (es. una manopola per regolare il volume di uno stimolo acustico), si richiede in ogni prova al soggetto di aggiustare l’intensità “I” fino ad arrivare al punto in cui lo stimolo è appena percepibile. Tale punto rappresenta il punto di massima incertezza (stima della soglia).

La procedura ha il vantaggio di essere molto veloce e pratica, ma ha lo svantaggio della variabilità di ciò che può essere considerato “*appena percepibile*” da persona a persona. Per ovviare al problema si evita l’utilizzo del compito yes-no, privilegiando compiti di discriminazione. Nei *test di confronto* il soggetto è tenuto a fare una scelta discriminativa tra due stimoli (es. qual è il più pesante tra due stimoli), anche quando questi appaiono simili tra loro. La funzione psicometrica che ne deriva è identica a quelle viste in precedenza, con la differenza che non si parla di probabilità di aver discriminato lo stimolo o meno, ma di probabilità di dire che il test è maggiore dello standard.

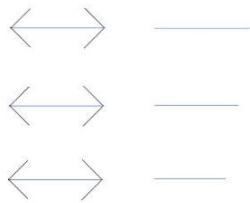
In questa situazione è possibile ragionare sulla distinzione tra **PES (punto di uguaglianza soggettivo)** e **PEO (punto di uguaglianza oggettivo)**. Il PEO è il punto in cui il test è oggettivamente uguale allo standard, mentre il PES è il punto in cui, facendo l’esperimento di discriminazione, si riscontra che per il soggetto il test è uguale allo standard, secondo una valutazione soggettiva.

Ad esempio, prima di fare un giudizio discriminativo, si fa fare uno sforzo fisico di sollevamento di un manubrio ad un soggetto, stancandone il bicipite, e successivamente si fa fare l’esercizio discriminativo: potrebbe succedere che, avendo affaticato il muscolo, il manubrio sembri più pesante, e dunque il PEO non sarà uguale al PES: questo si sposta e la sigmoide shifta rispetto l’asse x.

### Illusione di Muller-Lyer:

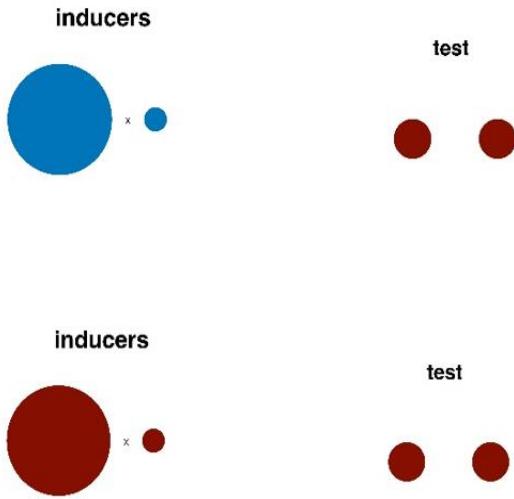
Nell’**illusione di Muller-Lyer** risulta evidente come la direzione delle frecce influenza il modo in cui

#### PES e PEO



viene elaborata la lunghezza di un segmento. In immagine è possibile osservare 3 segmenti uguali a sinistra e 3 segmenti diversi a destra, confrontabili a coppie orizzontali. Nonostante l’esito delle risposte sembri tendere ad una considerazione della coppia centrale come costituita da segmenti uguali (PES), è la prima coppia in alto ad essere costituita di fatto da segmenti della stessa lunghezza (PEO).

## Esperimento Overall

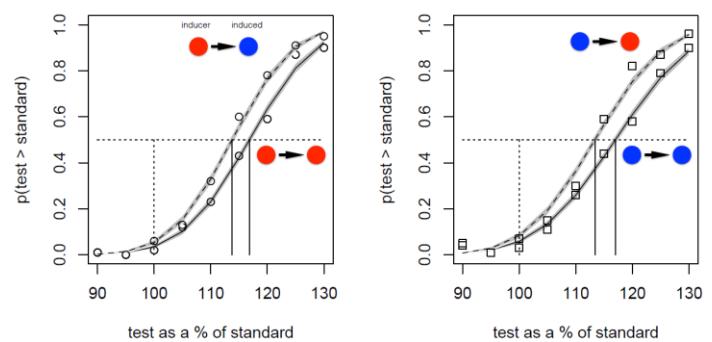


Un ulteriore esempio di distinzione tra PES e PEO è dato dall'illusione tra dischi nell'esperimento Overall:

Nel compito viene richiesto al soggetto di attuare una discriminazione di grandezza tra due dischi della stessa dimensione (Test) dopo aver osservato, ponendo il punto di fissazione dello sguardo in "x", l'immagine presentante due cerchi di diverse dimensioni (Inducers). Il punto di uguaglianza soggettivo (PES) emerge quale diverso rispetto al punto di uguaglianza oggettivo (PEO): nonostante i due dischi del test siano identici, vengono soggettivamente percepiti come diversi, uno più grande dell'altro. Ai soggetti viene richiesto dunque un compito discriminativo e il metodo utilizzato è il metodo degli stimoli costanti.

In immagine sono rappresentate le funzioni psicometriche, sull'asse delle x è presente "I", che in questo caso rappresenta la grandezza del test espresso in percentuale. 100% rappresenta il PEO, punto di uguaglianza oggettivo. È possibile notare che le curve sono spostate rispetto al PEO, verso destra. Il punto di uguaglianza soggettivo (PES) è diverso dal PEO, e tale diversità è dovuta ad un **effetto di contesto**: il contesto grande rimpicciolisce, il contesto piccolo ingrandisce.

### experiment 1: overall



La differenza tra i due grafici è data dal fatto che i colori sono stati incrociati (contesto rosso – test blu, oppure, contesto rosso – test rosso) al fine di attuare anche una manipolazione di congruenza del colore tra contesto e test.

L'effetto maggiore constatabile è quello del "**contrasto di grandezza**", ma è possibile in aggiunta notare una dipendenza anche dal colore: se il contesto, ossia il colore, dei dischi Inducers è dello stesso colore dei dischi di test, lo shift/contrasto è maggiore; viceversa lo shift è minore. **È possibile dunque evidenziare una dipendenza non solo dalla dimensione ma anche dal colore dello stimolo.**

## Lezione 4 – Psicoacustica e audiogrammi – Metodi di Stevens

Concludiamo il discorso sui metodi di Fechner, i **tre metodi di psicofisica indiretta**: abbiamo visto come funziona il *metodo degli stimoli costanti* e come funzionano gli altri due metodi che sono più o meno la stessa cosa con qualche differenza nella maniera in cui si fanno le diverse osservazioni, ed abbiamo anche visto la differenza tra **PES** e **PEO**.

Questo tipo di misurazioni si utilizzano anche da un punto vista pratico, dimostrando che questi principi non valgono solo per la visione, ma anche per altri aspetti e altre applicazioni della psicofisica.

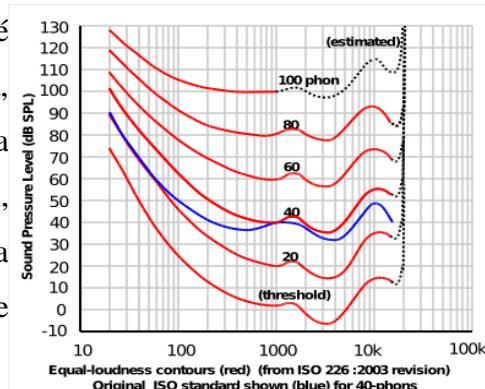
L'esempio più classico dell'applicazione della legge di W-F è nel campo della **psicoacustica**, ovvero la funzione psicofisica per la *loudness*, dove se noi fissiamo la frequenza di un singolo tono e poi ne modifichiamo l'ampiezza dell'onda (*pressure level*), ovvero il livello di pressione che viene esercitato sulle molecole dell'aria dall'onda otteniamo una **misura di intensità** del suono. Se noi variamo questo e andiamo a misurare le soglie, troviamo una curva che rappresenta la legge di W-F.

In realtà, questo è interessante fino a un certo punto (la funzione psicofisica per la *loudness* di un singolo tono), poiché siamo esposti a suoni che hanno un'ampia gamma di frequenze; noi sentiamo suoni a frequenza molto bassa e suoni a frequenza molto alta e invecchiando prediamo le frequenze altissime.

Invece di studiare una funzione psicofisica semplice come quella della *loudness* si possono costruire delle funzioni psicofisiche più complesse, in cui la variabile fisica non è più una, ma più di una, ad esempio in questo caso vediamo un **audiogramma**:

l'audiogramma rappresenta una funzione psicofisica in cui la soglia (*sensibilità*) viene plottata in funzione di due cose; una è quella sull'asse delle X (la frequenza del suono).

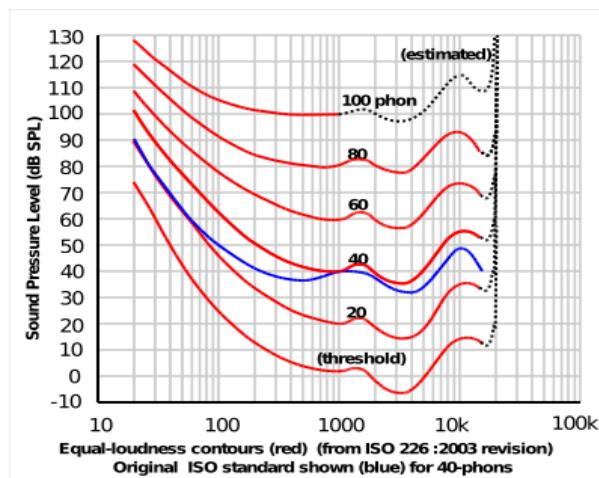
La scala della frequenza del suono è una scala un po' strana, perché vediamo che va da 10 a 100 Hz, poi da 100 a 1000 Hz e così via, vediamo che la distanza da 10 a 100 è uguale alla distanza da 100 a 1000, questa tipologia di grafico rappresenta gli assi logaritmici, perché uguale distanza sull'asse corrisponde in termini numerici a uguali rapporti, cioè c'è lo stesso rapporto tra 10 e 100 e fra 100 e 1000 e così via.



Quindi una delle due variabili considerate è la frequenza, non più il tono a 4 kHz, ma la sensibilità in funzione delle diverse frequenze e le diverse curve corrispondono a diversi livelli di intensità del suono, quindi se guardiamo la prima curva (quella in basso) vediamo che la soglia varia da decisamente molto alta per frequenze molto basse a sinistra, e poi scende aumentando la frequenza e, più o meno tra 1000 e 10k (circa a 4k Hz) nel range di quel singolo tono, la soglia è più bassa possibile e poi torna a risalire (i punti rappresentano un'estrappolazione, una stima).

Il punto di massima sensibilità, ovvero la soglia più bassa possibile, si va a collocare proprio intorno ai 4000/5000 Hz, il range di frequenza in cui noi parliamo, quindi il nostro udito è utilizzato per sentire in questo range di frequenza.

Le altre curve, quelle di sopra, sono invece delle misurazioni di sensibilità sopra soglia, quindi non è la capacità di distinguere la presenza di un tono dalla sua assenza, ma la capacità di distinguere fra due toni, e vediamo che più o meno la soglia si va a collocare più in alto, perché essendo suoni più intensi la sensibilità è meno buona.



Questo è un esempio di audiogramma, da notare è che in basso in rosso ci sono dei contorni di uguale loudness e indica che ognuno dei contorni rossi rappresenta una combinazione di decibel e frequenza che producono in un ascoltatore la sensazione di uguale loudness, come dire che affinché due suoni uno a frequenza bassa e una a frequenza che rientra nel range di massima sensibilità intorno a 4000Hz, perché ci appaiano ugualmente intensi, bisognerà che quella bassa frequenza sia fisicamente più intensa, perché la nostra sensibilità non è ugualmente buona. Ad esempio, vuol dire che se noi abbiamo un quartetto jazz con la cantante il tecnico del suono deve tirare su il volume del contrabbasso o il basso, perché altrimenti sembrerà meno intenso.

Questo rappresenta un fenomeno interessante, **stimoli che sono dal punto di vista fisico diversi possono apparire uguali rispetto a una qualche dimensione percepita (in questo caso la loudness)**. Questa maniera di mappare la sensibilità acustica è molto completa, perché ci dice per suono di diversa intensità e diversa frequenza il **comportamento del nostro apparato uditivo** e quindi ha anche molte applicazioni dal punto di vista clinico.

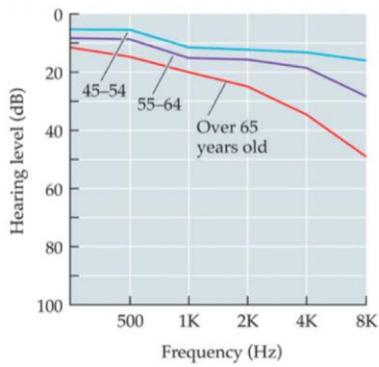
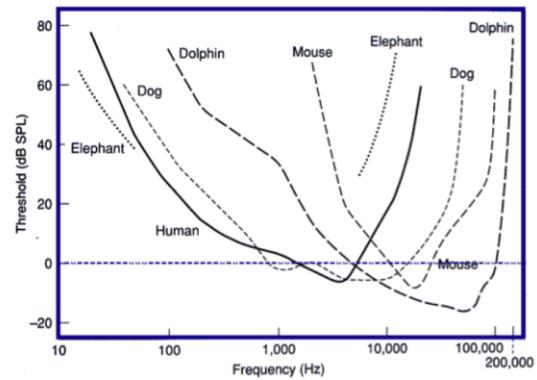
Un'applicazione, diversa da quella clinica, ma teoricamente interessante è che confrontando delle curve come queste possiamo farci un'idea della **struttura del mondo uditivo di organismi diversi**. Anche se non possiamo andare a chiedere a degli organismi non umani che cosa sentono, possiamo fare degli esperimenti di discriminazione con opportune tecniche e possiamo stimare delle *curve di udibilità* (audiogrammi) per un elefante oppure per un cane, topo o delfino e se confrontiamo queste curve vediamo per esempio che il delfino ha un mondo uditivo che è fatto di frequenze molto più alte di quelle che abbiamo noi, è spostato verso destra in maniera molto marcata rispetto all'uomo.

Questo rappresenta un'altra tipica rappresentazione della psicofisica, ci consente di avere un'idea della struttura del mondo percepito non solo di umani diversi, ma anche facendo confronti tra le specie. Un altro tipico problema è provare a capire come è fatto il mondo percettivo di un organismo umano che ancora non è in grado di parlare (neonato) e si possono fare delle misure di questo tipo, utilizzando delle tecniche che non sono molto diverse da quello che si utilizzerebbero con un cane e quindi per esempio scoprire che non è vero, come si pensava fino agli anni 60, che i bambini di poche settimane sono ciechi, ma sono immaturi dal punto di vista visivo, non si è ancora completato lo sviluppo di certe strutture neurali nell'occhio, però si possono misurare delle curve di visibilità e quindi hanno un certa zona in cui sono sensibili e possono rispondere a stimoli visivi.

Un altro esempio in cui abbiamo un'applicazione clinica è il fatto che le curve di udibilità si possono **perdita dell'udito** utilizzare per valutare deficit di tipo uditivo e **un fenomeno che assolutamente evidente è il fatto che con l'invecchiamento c'è un decadimento della sensibilità per le frequenze più alte**.

Questo lo vediamo dalle curve che confrontano tre decadi, questo è un esempio di audiogramma che vedremmo nel momento in cui andremmo in una clinica. La maniera in cui viene ricavato è mettere il soggetto all'interno di un cubo schermato ai suoni e con delle cuffie si fa un breve esperimento di psicofisica; si misurano delle soglie per stimare la sensibilità in funzione della frequenza del suono.

## curve di udibilità



## Metodi adattivi, Stevens e Teoria della detezione del segnale

- Le applicazioni moderne dei metodi di Fechner vengono definiti **metodi adattivi**. Sono i metodi di Fechner per misurare le soglie resi più efficienti in maniera da riuscire a stimare delle soglie con un'economia di tempo;
- **La psicofisica di Stevens**, un pò diversa da quella di Fechner, non solo perché arriva un secolo dopo, ma anche da un punto di vista metodologico;
- La **teoria della detezione del segnale** che è un approccio diverso per la misura della sensibilità che viene usato tantissimo nella ricerca.

I **metodi adattivi** sono sostanzialmente delle variazioni di quello che viene chiamato “*stimoli costanti*”, ma resi più efficienti. Ci sono diversi metodi per stimare delle soglie, però ci limiteremo a spiegare il metodo della “scalinata” (*staircase*) che rappresenta la “madre” dei metodi adattivi che sono sostanzialmente tutti i metodi di questo tipo alla cui base c’è un’idea fondamentale:

andiamo a cercare una soglia, abbiamo un qualche metodo e il nostro computer è in grado di controllare i parametri di uno stimolo che ci interessa, di cui vogliamo misurare una soglia e, un po’ come il metodo dei limiti si parte soprasoglia, quindi si dovrà scendere per andare verso la soglia. **A differenza del metodo dei limiti gli stimoli da presentare non sono predefiniti o fissati in anticipo, ma vengono di volta in volta ricalcolati adattivamente ovvero in funzione delle risposte del soggetto**, questo consente anche di modificare leggermente la procedura. Di fatto la configurazione si modifica da sola automaticamente anche in funzione delle differenze individuali.

L’idea è che se le nostre risposte indicano che siamo in un range che per noi è sicuramente soprasoglia (questo si vede se diamo sempre la stessa risposta), allora vuol dire che quel range non è interessante e devo spostarmi da quella zona e quindi un metodo come la staircase in una situazione del genere cambia lo stimolo di molto e abbassa, per esempio, l’intensità.

Se invece, vediamo che le risposte iniziano ad essere variabili, ovvero ci sono inversioni, vuol dire che mi trovo nella zona interessante, nella zona di incertezza e quindi l’algoritmo che gestisce questi tipo di metodo è costruito per farci rimanere in questa zona di incertezza dove raccogliamo dei dati utili.

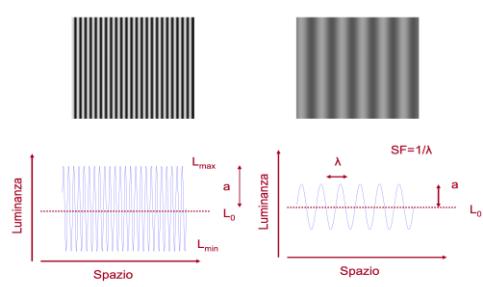
**Pensando alla funzione psicometrica (che si ricava con il metodo degli stimoli costanti, funzione a forma di “S”), è un po’ come dire che la zona interessante di quella funzione è la parte centrale della “S”, la parte in cui il soggetto è incerto perché la soglia è lì dentro. Le code invece non hanno alcun interesse perché il soggetto dice o sempre sì o sempre no.**

Quindi l’idea di un metodo adattivo in fondo non è altro che questo, inutile stare in quelle parti estreme in cui il soggetto dirà sempre “sì” o sempre “no”. Cerchiamo di andare nella zona interessante, nella zona di incertezza, e lì facciamo più misure possibili.

Un’altra cosa tipica del modo della staircase è che noi possiamo immaginare come il metodo dei limiti si parte soprasoglia e si scende sotto soglia e una volta che si è arrivati alla zona di incertezza si rimane lì e si fanno tante misure.

Il problema di questo, un po’ come nel *teorema dei limiti*, è che se noi cambiamo lo stimolo sempre in un’unica direzione (cioè scendendo), nel partecipante si forma un’**aspettativa** e questo può generare degli errori che non sono percettivi, ma legati all’aspettativa (so che sto scendendo quindi mi aspetto di non vedere più lo stimolo ad esempio). **Per evitare questo, si utilizzano almeno due staircase che vengono interlacciate, ovvero gli stimoli vengono randomizzati grazie al computer**, quindi non sappiamo se lo stimolo che stiamo vedendo in quel momento a cui dobbiamo rispondere appartiene ad una o all’altra staircase, quindi questo scoraggia gli errori di aspettativa.

Un esempio di dati che possiamo raccogliere facendo questo tipo di esperimento (e che usiamo per introdurre qualcosa che ci servirà a breve quando parleremo di visione di basso livello, ovvero il **grating – graticcio o reticolo, ma la traduzione letterale è staccionata**) e rappresenta uno stimolo che viene utilizzato ogni qual volta si fanno delle misure molto accurate per stimoli visivi di un certo tipo, ovvero delle **discontinuità orientate o contorni**.



Il graticcio è uno stimolo ben definito nel dominio dello spazio da una funzione matematica molto semplice che è la funzione *seno*, cioè una variazione ciclica della luminosità che arriva da quello stimolo (luminanza), e in funzione dello spazio questo varia ciclicamente da alta a bassa (da bianco a nero). Dal punto di vista fisico c’è una luce intensa e una luce debole e il passaggio è graduale che non è detto che sia visto come tale, in quanto cambia in funzione della distanza.

**Uno stimolo come questo è definito perfettamente e ha solo due parametri** che sono la sua **frequenza spaziale** e il suo **contrasto**. La **frequenza spaziale** fa riferimento al numero di barre che ci sono, quante alternanza di nero e bianco ci sono nell'unità di spazio, mentre il **contrasto**, dal punto di vista visivo, è la differenza tra il massimo e il minimo di quell'onda, cioè tra la luminanza massima e quella minima.

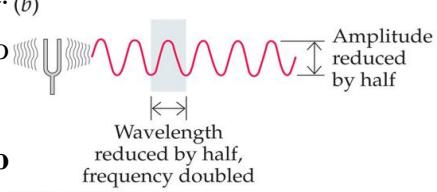
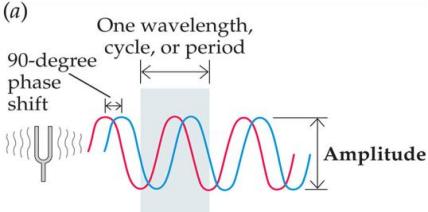
I due grating hanno diversa frequenza spaziale, perché in quello di sinistra nell'unità di spazio vi sono più alternanza di nero e bianco, c'è più livello di dettaglio (**la frequenza spaziale è la misura del livello di dettaglio che è presente in uno stimolo**). Inoltre, c'è anche differenza di contrasto perché quello di sinistra è più contrastato e quello di destra meno contrastato.

**“SF” indica la frequenza spaziale che rappresenta il reciproco della lunghezza d'onda**, a cui generalmente non si fa riferimento perché quando parliamo di grating prediligiamo il contrasto da un punto di vista visivo. Tale analogia l'abbiamo con il suono perché quest'idea di caratterizzare così uno stimolo visivo è venuta negli anni 60 proprio facendo un'analogia con la maniera in cui si descrive il suono.

Un suono è uno stimolo completamente diverso da quello visivo, nel senso che la frequenza è rispetto all'unità di tempo, il numero di onde che si verificano in una certa finestra di tempo.

Dal punto di vista matematico è sempre un'onda, se prendiamo un diapason, gli diamo un colpo e questo mette in vibrazione molecole nell'aria producendo un'onda che si propaga con una certa frequenza nel tempo e con una certa ampiezza del suono (l'equivalente del contrasto nel visivo), cioè la differenza tra massimo e minimo di quell'onda. (b)

Quindi un suono intenso ha un'ampiezza grande, così come uno stimolo molto contrastato sembra più saliente di uno meno contrasto.

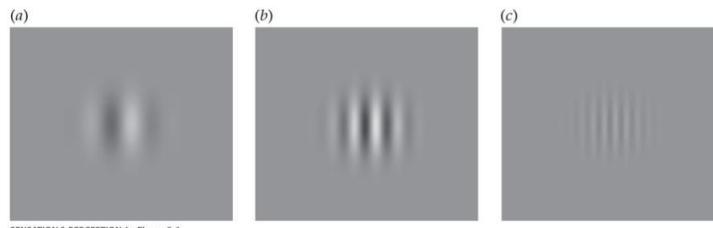


**Ciò che definisce i grating sono la frequenza spaziale e il contrasto**

e quello che si nota subito se proviamo a variare frequenza spaziale e contrasto è che il **contrasto apparente**, ovvero quello percepito di differenza tra nero e bianco **non è una funzione ovvia del contrasto fisico** e questo è ciò che abbiamo notato prima negli audiogrammi, l'ampiezza di un onda da un punto di vista fisico non mappa in maniera automatica sulla loudness (sull'ampiezza percepita), ma dipende anche dalla frequenza del suono quanto ci sembra intenso.

SENSATION & PERCEPTION 4e, Figure 1.16  
© 2011 Sinauer Associates, Inc.

Stessa cosa succede nella visione, infatti analizzando questo esempio:

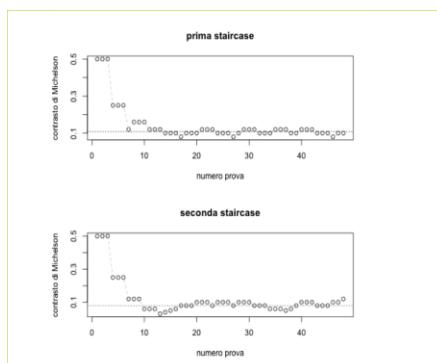


SENSATION & PERCEPTION 4e, Figure 3.6  
© 2015 Sinauer Associates, Inc.

abbiamo tre grating che sono con diverse frequenze spaziali (quello di sx bassa frequenza spaziale), quindi ci sono frequenza spaziale bassa, media e più alta.

Di questi tre grating il contrasto apparente più alto (più bianco e più nero) è senza dubbio *l'esempio b*, però in realtà dal punto di vista fisico il contrasto è uguale in tutti e tre gli esempi. **Il motivo per cui ci appare diverso dipende dalla sensibilità al contrasto non è costante ma dipende dalla frequenza spaziale, noi siamo ottimizzati per rilevare i contrasti, ovvero i contorni, in un range intermedio e nel caso della visione corrisponde statisticamente al range di livello di dettaglio che è più rilevante nella nostra nicchia ecologica per percepire gli oggetti.** Le cose che hanno un livello di dettaglio molto fine non sono rilevanti dal punto di vista comportamentale, mentre lo sono quelle cose con cui possiamo interagire dal punto di vista fisico e corporeo.

Per misurare ciò, vedremo che per ricavare la funzione di sensibilità al contrasto, dobbiamo misurare delle soglie e la maniera è utilizzare due *staircase*, cioè un esperimento di questo tipo: prendo un grating a un certo livello di frequenza spaziale e vado a modificare il contrasto cercando la soglia per il contrasto attraverso il metodo dell'aggiustamento, per esempio, con due bottoni in cui il bottone si preme riducendo il contrasto in cui è appena percepibile.



Oppure posso fare due *staircase* interlacciate, in cui se ne prendiamo una sull'asse delle Y c'è il **contrasto detto di Michelson** perché viene preso in percentuale, ovvero **differenza tra massimo e minimo diviso la somma del max e del min**, quindi diventa una **percentuale di contrasto**. Nel momento in cui si fa l'esperimento "yes/no" mostrando un grating, partiamo molto sopra soglia e vediamo che la procedura è adattiva perché ci sono tre risposte sempre a quel livello di contrasto e alla terza volta il programma capisce che siamo lontani dalla zona interessante e quindi scende bruscamente, riducendo il contrasto. Si ripete la stessa cosa, per tre volte e quindi altro salto, un po' meno grande rispetto al primo perché si presume che siamo arrivati alla soglia. Il pallino singolo in basso rappresenta che a quel livello di contrasto il soggetto ha invertito la risposta e allora il programma

alla terza volta il programma capisce che siamo lontani dalla zona interessante e quindi scende bruscamente, riducendo il contrasto. Si ripete la stessa cosa, per tre volte e quindi altro salto, un po' meno grande rispetto al primo perché si presume che siamo arrivati alla soglia. Il pallino singolo in basso rappresenta che a quel livello di contrasto il soggetto ha invertito la risposta e allora il programma

adattivamente ha di nuovo aumentato il contrasto e vediamo che a questo punto i cambiamenti di contrasto sono diventati molto piccoli, e con un piccolo aumento per tre volte il soggetto l'ha riconosciuto, poi siamo scesi e così via. L'idea è molto semplice, stiamo facendo tante misurazioni sempre intorno ad una zona di incertezza e ad ogni singola prova il programma controlla se il nuovo stimolo apparterrà alla prima o alla seconda *staircase*, quindi non è detto che le tre prove all'inizio vengano fatte consecutivamente e in questo modo si evitano gli errori di aspettativa dando una misura più pulita della sensibilità.

Le due *staircase* vengono plottate separatamente e contengono una moltitudine di dati; posso prendere, come nel metodo dei limiti, i punti di inversione facendo la media e quella è la soglia in base alla prima *staircase* e la stessa cosa la posso fare con la seconda *staircase* e non è detto che vengano identiche, ma presentano delle similitudini. Tuttavia, posso dire che la soglia di contrasto per questo particolare grating si colloca attorno al 10%.

## Stanley Smith Stevens

Stevens oltre alla statistica, si è occupato molto anche di psicofisica, ed è molto noto in tale ambito perché ha proposto che la forma generale della funzione psicofisica non è quella di W-F, ma è quella descritta da questa equazione:

$$S = cI^a$$

Chiamata “*legge di Stevens*”, essa ha delle analogie, ma anche delle differenze con la legge di W-F e consiste nel fatto che secondo Stevens la forma generale della funzione psicofisica non è logaritmica, ma è una **funzione potenza, è una funzione in cui la sensazione è proporzionale all'intensità dello stimolo fisico elevata a un esponente** (<sup>a</sup>).

Fechner ricavò la funzione logaritmica utilizzando il metodo analitico, prendendo la legge di Weber e attraverso la matematica dimostrò che se è vera tale legge allora la funzione psicofisica deve essere una legge logaritmica e sostanzialmente sono stati altri che hanno dimostrato che i dati si adeguano in molti casi a quel tipo di previsione.

Stevens ebbe un approccio molto diverso, inventò una serie di metodi per stimare S in funzione di I e poi ha mostrato che utilizzando questi metodi il miglior *fit* ai dati è la funzione potenza; i metodi di Stevens vengono chiamati metodi di **psicofisica diretta**, perché Stevens non passa dalla soglia, non ricava una funzione tipo audiogramma: nel caso di **Stevens la sensibilità viene misurata chiedendola ai soggetti, attraverso una procedura in cui si danno delle risposte che indicano la struttura dell'esperienza percettiva**.

L'idea fondamentale era quella che Stevens chiamava “**stima di magnitudine**”, che si riferisce a un continuum di intensità.

In un esperimento in cui dobbiamo fare una stima di magnitudine prendiamo un peso e diamo la stima di quel peso e, assumendo di usare una scala in maniera sensata, quando prendiamo un altro peso che sembra il doppio più pesante, daremo una risposta doppia rispetto alla prima. Pertanto, risulta possibile chiedere apertamente e ci permette di fare una stima numerica dell'intensità della propria sensazione.

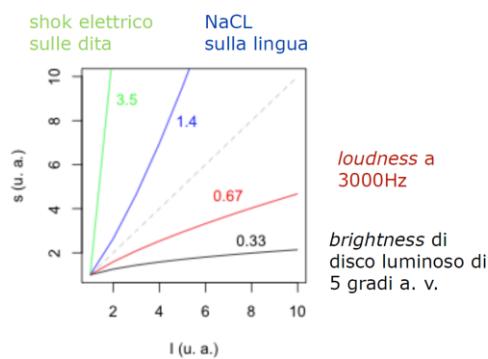
È consigliabile standardizzare i numeri per avere sempre la stessa unità di misura, così come l'ancorare i giudizi, e nel caso del peso, dare un peso di riferimento e successivamente una serie di stimoli. Questa idea è molto utilizzata soprattutto nella scala del dolore dei medici (es: quanto è forte il tuo mal di testa).

Una variante della stima di magnitudine è il **confronto cross modale**, invece di dire un numero che dovrebbe essere, per noi, proporzionale alla intensità della sensazione, quello che Stevens faceva era dare il controllo su un altro tipo di canale sensoriale, per esempio attraverso un potenziometro in cui la forza della presa corrisponde al dolore che stai provando, il loro utilizzo è molto immediato rispetto alle persone ed è molto naturale.

Esempio “**Momu**” e “**Keki**” di **Wolfgang Köhler** per sottolineare la naturale corrispondenza percettiva che c'è tra stimoli che vengono processati da canali sensoriali diversi.



## Funzione psicofisica di Stevens



La formula della funzione psicofisica di Stevens viene mostrata nel grafico in diversi domini:

invece di mettere x e y sugli assi delle quantità ben definite, c'è scritto u.a. ovvero unita arbitrarie. Si è cambiata la scala di diversi tipi di continuità fisiche, in quanto ci sono cinque funzioni psicofisiche e corrisponderebbero a cinque diversi tipi di stimoli che andrebbero misurati con la loro unita di misura.

Questo ci permette di plottare sullo stesso grafico cinque curve (quella tratteggiata è una retta), vicino ad ogni curva c'è un numero che nella formula corrisponde ad “a” ovvero l'esponente a cui dobbiamo elevare l'intensità per disegnare la curva (formula  $S=cI^a$ ). La “c” rappresenta la costante di proporzionalità e serve ad aggiustare il numero tenendo conto dell'unità di misura che noi abbiamo tenuto uguale a 1.

Abbiamo ripreso da Stevens quattro valori diversi dell'esponente e notiamo che due di questi esponenti sono più piccoli di 1, quando questo avviene la curva non è molto diversa dalla funzione di W-F anche se matematicamente non è logaritmica, più o meno racconta la stessa storia (**la sensibilità cresce rapidamente all'inizio e poi sempre di meno fino ad appiattirsi**) e ciò vale nel dominio dello stimolo acustico e visivo. **Gli altri due sono esponenti maggiori di 1** e la funzione potenza si comporta in maniera diversa dalla legge di W-F, ovvero **è accelerata positivamente** (all'inizio la sensibilità non cresce rapidamente, ma quasi subito la curva si impenna), **incrementi più piccoli producono effetti più grandi e questo si osserva con stimoli di tipo avversivo sui nostri sistemi sensoriali** (es: sale sulla lingua o shock sulle dita). **Quindi Stevens riteneva che questa fosse una proprietà più generale dei sistemi sensoriali, come vista e udito che sono compressivi (la curva si appiattisce e assomiglia alla legge di W-F), mentre altri sono espansivi e quindi invece di produrre una sensibilità che si estende in un range molto ampio di intensità e riducendo progressivamente la sensibilità succede il contrario, perché non appena si supera un certo livello di intensità non è più adattivo essere sensibili, ma conviene "staccare la spina", come sputare quello che si ha in bocca.** La curva tratteggiata rappresenta un'eccezione, perché rappresenta la funzione potenza di Stevens per la percezione di un'estensione (quanto grande ci sembra un qualcosa) e questo lo si può rilevare per via tattile o visiva e l'esponente è 1 e quindi di fatto si osserva una retta.

## Introduzione alla teoria di detezione del segnale



Lo introduciamo osservando dei puntini ovvero il *cinematogramma di punti casuali*, ciò che succede è che vengono scelte posizioni a caso dove mettere un punto che poi scompare dopo pochi secondi e si muove coerentemente a destra o a sinistra o in maniera casuale, questo stimolo contiene una componente che chiamiamo **segnale** e un'altra componente che chiamiamo **rumore**. Supponiamo di fare un esperimento come questo con il metodo di Pirenne e ci si chiede se abbiamo visto un movimento coerente verso destra e alla fine c'è qualcuno che dice di averlo visto per il 95% delle volte, questo ci sembra suggerire che la sensibilità per il movimento coerente è molto buona perché riesce a rilevare questo movimento nonostante la presenza del rumore; **posso concludere che la persona è molto sensibile al movimento?** La risposta potrebbe essere affermativa però in generale no perché questa prestazione potrebbe non dire nulla sulla sensibilità a questo movimento. Potrebbe essere semplicemente la propensione di questa persona a rispondere sempre di sì anche quando ha un vago sospetto del movimento.

## **Lezione 5 – Teoria della detezione del segnale – Inizio percezione visiva**

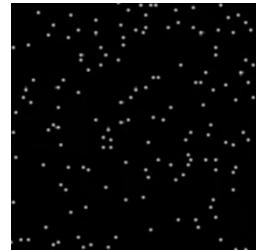
### **Teoria della detezione del segnale**

Terminiamo il discorso sulla psicofisica parlando della teoria della detenzione del segnale e parliamo non più della “*psicofisica classica o fechneriana*”, in cui la **sensibilità** si misura indirettamente misurando soglie, oppure abbiamo visto la psicofisica di **Stevens** che ha un carattere più psicologico, ed è detta “*psicofisica diretta*” perché analizza in modo diretto le esperienze dei soggetti e si chiede ai soggetti di mettere alla prova le proprie capacità discriminative apportando opportuni accorgimenti. Esiste poi un terzo approccio che è preferibile in diverse situazioni, dove la sensibilità viene misurata con **compiti di detezione**, ovvero compiti in cui viene chiesto di classificare degli stimoli decidendo in una scelta forzata se è presente un **segna**le. Un segnale può essere qualsiasi cosa che è necessario rilevare, sia un compito psicofisico, oppure esperimenti di psicologia cognitiva, dove si deve cercare un target in un “campo di distrattori”.

Questa è la **teoria della detezione del segnale** che nasce negli anni ‘60 ed è un insieme di tecniche che hanno larga applicazione anche in contesti di tipo tecnologico; es. tecnici di radiologia, i tecnici di radiologia devono trovare i segni di una patologia da una immagine radiologica. Il problema risiede sempre nell’efficienza della tecnica nel rilevare le patologie, quindi utilizzare la teoria della detezione del segnale per rilevare patologie.

***Di cosa si tratta quando parliamo di teoria della detezione del segnale?***

Rimandiamo all’esempio della immagine dei puntini: Compito yes/no, dove si richiede di riconoscere la presenza di un **segna**le, in questo caso si tratta di un movimento coerente di punti, e bisogna dire se c’è oppure no. In questo specifico caso la nostra prestazione è in realtà un cattivo indicatore della nostra sensibilità perché contaminata dal nostro criterio di risposta ovvero a come si decide di rispondere. Paradossalmente un soggetto potrebbe dire che vede sempre il segnale anche quando non è vero (si parla anche di **desiderabilità sociale** nei confronti dello sperimentatore).



*Altro esempio: non so se il segnale c’è o meno, quindi si fa un compito di classificazione: lo stimolo c’è o non c’è? C’è più segnale o rumore?* Se il segnale è più forte del rumore è molto più semplice discriminarlo e viceversa.

## Slide: "Un altro ipotetico esperimento di percezione del movimento è:"

In questa slide si discute sulla probabilità condizionata, indicata dai simboli %( $\rightarrow/\rightarrow$ ).

Prendiamo un soggetto che deve rispondere ad un compito di detezione del segnale simile a quello dei punti visto in precedenza e gli chiediamo di rispondere se vede un movimento verso destra. Il risultato della sua performance è il seguente:

1. Segnale presente verso destra: %( $\rightarrow/\rightarrow$ ) = 95%.

Il soggetto risponde che nel 95% delle prove ha visto un segnale/movimento verso destra;

2. Segnale non presente: %( $\rightarrow/\leftarrow$ ) = 90%.

Il soggetto risponde che nel 90% delle prove ha visto un segnale/movimento verso destra, anche se questo non era presente.

**Basandosi soltanto sulla prima osservazione potrebbe sembrare che il soggetto abbia un'elevata sensibilità alla rilevazione del segnale in questo compito, ma aggiungendo la seconda osservazione capiamo che in realtà è il soggetto che risponde anche quando il segnale non c'è.**

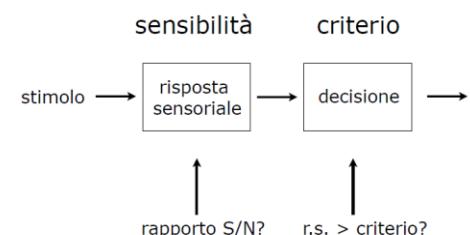
## Modello del compito di detezione

Dato uno stimolo la decisione che il soggetto prende nel dire se il segnale c'è o meno dipende da **due processi**:

1. Uno di questi processi ha ad a che fare con la **sensibilità** vera e propria, cioè con la risposta sensoriale, l'input dei processi percettivi influenzati dalla sensibilità e anche dal rapporto segnale/rumore;
2. Sulla base della risposta sensoriale vi è un **processo post percettivo di natura decisionale**: sulla base dell'evidenza sensoriale il soggetto decide se rispondere sì o no. Lo possiamo immaginare come un processo che confronta il criterio interno con la forza della risposta sensoriale e lo shift della risposta sì/no è identificato dal **punto di criterio o criterio di risposta**: se l'intensità della risposta sensoriale supera il criterio rispondo sì, altrimenti rispondo no.

Ciò diviene un problema di modellizzazione, un problema statistico: sulla base dei dati che ho, in un compito di classificazione, *come è possibile separare queste due componenti?*

*Come posso stimare la componente di sensibilità indipendentemente dal criterio di risposta del soggetto?*



Se otteniamo informazioni non solo sui veri positivi che il soggetto fa nella sperimentazione, ma anche sui falsi, cioè sugli errori che il soggetto commette, se utilizzo queste due informazioni in maniera opportuna **posso separare questi due aspetti della prestazione**.

Il primo parametro, la sensibilità, in psicofisica ci interessa molto di più del criterio nel contesto dello studio della percezione/psicofisica.

In psicofisica si utilizzano molti metodi, ma una distinzione fondamentale è quella tra compito tipo **Pirenne**, di tipo si/no (un compito per certi versi **introspettivo**) dove non esiste una maniera oggettiva di calcolare risposte corrette, e **compito di classificazione**, dove distinguiamo prove dove il segnale è sovrapposto al rumore oppure prove in cui c'è solo rumore:

- **yes-no: come il compito di Pirenne (non esiste maniera di calcolare %C);**
- **classificazione:  $S+N$  o solo  $N$ ? (include catch trials dove c'è solo rumore per cogliere in fallo il partecipante, quindi ci sono due maniere di rispondere correttamente).**

**Legenda:**

- **S: signal (segnale)**
- **N: noise (rumore)**

In un **compito di detenzione del segnale** a seconda del tipo di stimolo ci sono due maniere di rispondere bene (*hit e correct rejection*) e due maniere di rispondere male (*miss e false alarm*).

**Legenda tabella:**

- **HIT:** positivo corretto;
- **MISS:** risposta sbagliata;
- **CORRECT REJECTION:** corretta negazione della presenza dello stimolo;
- **FALSE ALARM:** dire che si è percepito lo stimolo anche quando non c'è.

		stimolo	
		segnale	rumore
risposta	si	hit	false alarm
	no	miss	correct rejection

Esempio: *Trial SEGNALE+RUMORE ( $S+N$ ):* possiamo rispondere positivamente bene (FARE UN **HIT – Positivo corretto**), oppure possiamo rispondere che non abbiamo visto lo stimolo (FARE UN **MISS**).

In un contesto come questo di segnale-rumore, fare un hit o miss sono le uniche due cose che possono succedere, quindi la somma di tutti gli hit e miss sarà la somma delle prove fatte e saranno **uguali a 1**, cioè l'evento certo perché non ci sono altre alternative.

Esempio di trial di classificazione: ci sono 100 prove di tipo S+N dove la performance è 70 hit e 30 miss; questa somma (70+30) deve fare 100 e pertanto anche la proporzione sarà 70 su 100 e 30 su 100 che mi da 1. La frequenza assoluta di hit e di miss viene divisa per il totale delle prove e otteniamo una frequenza relativa, ovvero una **stima di una probabilità**. Tutto ciò accade quando ci troviamo nel riquadro ROSSO dell'immagine, ovvero quello a sinistra, in un compito di classificazione possiamo trovarci o a sinistra dell'immagine (riquadro rosso) o a destra (riquadro blu).

Altro Esempio: ci troviamo in una situazione *catch trials*, dove abbiamo solo rumore (cioè parte blu dello schema in alto). La risposta corretta è quella dove si dice di no, quindi si fa un rifiuto corretto: non c'è stimolo e non lo vedo (**CORRECT REJECTION**). Oppure si può rispondere che lo stimolo c'è quando invece non c'era (**FALSE ALARM**). Anche in tutti questi casi la somma delle frequenze relative deve dare **1**.

Da questo discende che la tabella sopra è **ridondante** perché non occorre calcolare tutte e quattro queste probabilità, ne bastano due e si possono scegliere una qualsiasi coppia di indicatori per fare la stima, ad esempio HIT E MISS.

		stimolo	
		segna	rumore
risposta	si	H	FA
	no	1-H	1-FA

**Per convenzione in psicofisica/teoria della detezione del segnale si parte dalla proporzione di HIT (H) e di FALSI ALLARMI (FA) per fare i calcoli:**

Esempio: esperimento con 100 prove S+N e 100 prove N; potrei avere 70 su 100 HIT e 25 su 100 FALSI ALARMI. Queste due proporzioni sono le stime empiriche della probabilità per quel soggetto in quelle determinate condizioni di fare un positivo corretto o un falso positivo.

Questo stesso identico schema si usa anche in medicina quando si fa un test clinico, che è un test di detezione del segnale, dove c'è sempre una misura e dopo bisogna decidere se c'è la malattia c'è o meno. C'è anche similarità con la **statistica inferenziale** perché una tabella molto simile si ha quando per esempio uno statistico decide di rifiutare o meno l'ipotesi nulla. Lo schema visto in tabella in alto deriva dal lavoro di statistici che hanno sviluppato l'approccio dell'inferenza statistica.

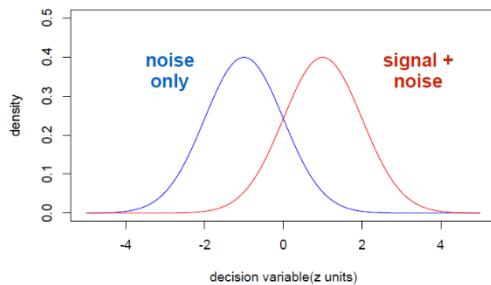
### Noi lavoreremo con delle proporzioni:

1. **H**, ovvero proporzione di HIT;
2. **FA**, ovvero proporzione di FALSE ALARM.

Non calcoleremo mai  $1-H$  e  $1 -FA$  cioè MISS e CORRECT REJECTION, perché sarebbero uno il reciproco dell'altro.

**Come si ricava una misura di sensibilità da queste grandezze?**

**modello gaussiano con uguali varianze**



Ci sono molte maniere di farlo, noi vediamo il caso più semplice ed utilizzato, ovvero quello basato sul **modello gaussiano con uguali varianze**.

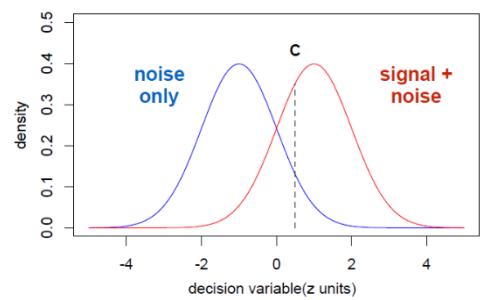
Per interpretare queste proporzioni, cioè H ed FA, si deve fare riferimento ad un modello statistico di come il soggetto prende la decisione, questo modello teorico è basato su due assunzioni:

1. La distribuzione delle risposte interna dell'organismo ad uno stimolo (S+N o solo S), quando si hanno tante prove, è un campione estratto da delle stesse curve normali, ovvero gaussiane, che hanno tutte la stessa forma e la stessa varianza, cioè la stessa dispersione.

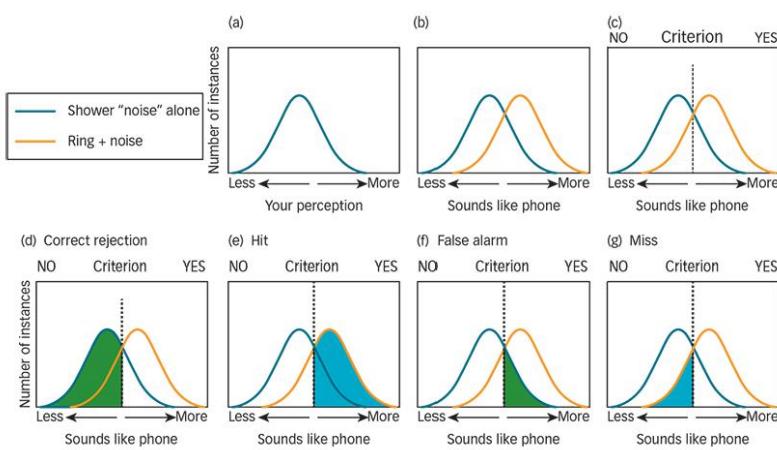
Rifacciamoci al grafico per capire meglio: sono due distribuzioni di probabilità teoriche espresse in punti z; immaginiamo una variabile interna all'organismo che è espressa in unità standard cioè punti z, e ci sono due distribuzioni: una, quella in blu, è la risposta dell'organismo nelle prove in cui c'è solo rumore, mentre l'altra è la risposta quando c'è segnale più rumore, in rosso. La distribuzione delle curve dipende da quanto è diversa la risposta dell'organismo in presenza del segnale; le curve si sovrappongono e quindi si può stimare quanto è forte la risposta dell'organismo da questa sovrapposizione. Dove si collocheranno queste due distribuzioni nel piano cartesiano sull'asse delle X **dipende da quanto è diversa la risposta dell'organismo in presenza del segnale**, seguendo questo filo logico possiamo pensare il grado di sovrapposizione tra le distribuzioni come una maniera di stimare quanto è forte dell'organismo, quanto il segnale produce un cambiamento nel sistema sensoriale.

2. Risposta dell'organismo quando c'è sia il segnale che il rumore.

Ragioniamo su cosa si intende con “conoscere una stima della proporzione di HIT e di FA”: **quando dobbiamo produrli cosa facciamo?** Il valore della variabile decisionale dipende dalla **presenza del segnale** (cioè dalla sensibilità), e dal **criterio di risposta**: al di sopra di una determinata forza dell'evidenza sensoriale rispondo di sì, in figura è la linea tratteggiata indicata con C, e il valore soglia è il corrispondente valore sull'asse delle X, solitamente questo livello C interseca entrambe le gaussiane.



**Fissato un c**, cioè un livello della variabile che sta sull'asse delle x, esso divide entrambe le distribuzioni in due parti, il totale dei casi sarà la somma di ciò che sta da una parte e dall'altra del segmento verticale.



Esempio classico della doccia: faccio la doccia e mi porto il telefono in bagno che suona. La doccia produce un rumore di fondo e a questo si sovrappone il telefono che suona: il compito di detezione in questo caso corrisponde alla domanda: **il telefono sta suonando oppure no?**

La situazione è quella in cui la distribuzione corrisponde a solo rumore (a) mentre un'altra corrisponde al suono del telefono (b), e poi c'è il criterio di risposta (c). Fissato il criterio (c) su queste distribuzioni ci sono quindi **4 aree** nel grafico che corrispondono alle proporzioni di **Correct rejection (d)**, **Hit (e)**, **False alarm (f)**. e **Miss (g)**.

Correct rejection (d): una **correct rejection** viene rappresentata dalle prove della distribuzione blu e siccome l'intensità della risposta era minore del criterio noi abbiamo risposto di no correttamente. **L'area verde sotto la distribuzione normale corrisponde alla probabilità teorica di una correct rejection.**

Hit (e): una **hit** sono nella distribuzione arancione dove la variabile interna supera il criterio e quindi rispondo di sì e **la probabilità teorica, sotto la distribuzione normale, di una hit è l'area segnata in azzurro**.

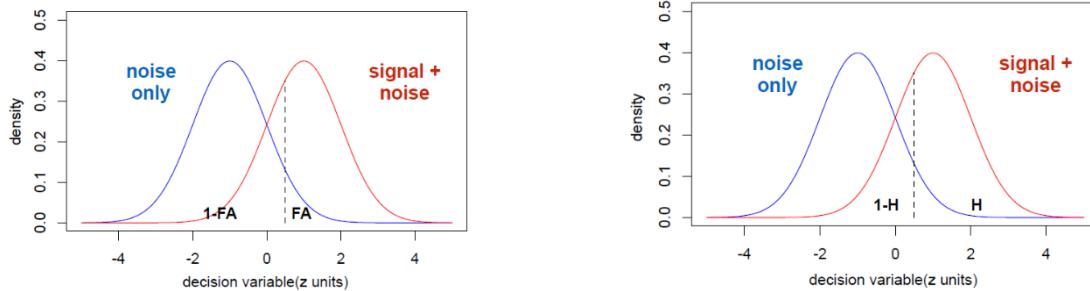
Negli errori la risposta sensoriale superava il criterio, e quindi ci troviamo nella parte a destra delle distribuzioni, e sono quelle prove in cui a causa della distribuzione probabilistica della risposta sensoriale quest'ultima supera il criterio e, ad esempio nel false allarm quindi nella distribuzione "solo noise", abbiamo detto di sì, ma in modo incorretto.

Le due zone verdi sommate ci danno la totalità delle prove N (blu), dividendo le risposte tra risposte corrette e sbagliate e la stessa cosa per hit e miss, le due zone blu ci danno la totalità della prova S (arancione).

Tutto ciò che abbiamo appena detto è un **modello teorico**: se la distribuzione della risposta interna è gaussiana e se HIT e FALSE ALLARM sono due gaussiane con uguale varianza allora posso pensare in questa maniera alle probabilità di commettere hit, miss ecc...

Quando invece si fa l'esperimento, invece si parte dal dato reale e avendo il dato reale si stima empiricamente la probabilità reale di fare un hit o un falso allarme perché misuro la frequenza di hit e false alarm dividendo per il totale. Per interpretarli si ritorna al modello teorico e si assume che la stima empirica dovrebbe corrispondere all'area rappresentata nel secondo grafico da sinistra, quell'area la stimo pari alla proporzione di hit.

Esempio: se faccio 70 HIT su 100, la mia area corrisponderà al 70% della guassiana e di conseguenza i MISS saranno il 30%. Se la proporzione di falsi allarmi è 25%, vuol dire che l'area in verde alla destra del criterio nella distribuzione blu sarà  $\frac{1}{4}$  del totale.

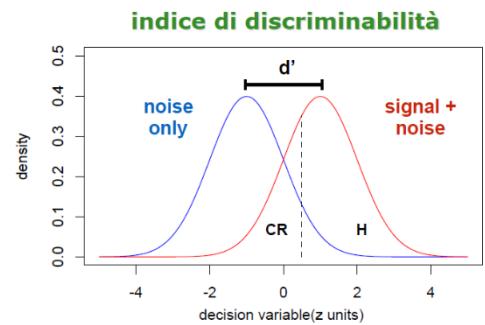


Distribuzione solo rumore (blu)	Distribuzione segnale + rumore (rosso)
alla destra del criterio vi è la probabilità dei FALSE ALLARM (FA), mentre a sinistra del criterio sotto la curva blu vi è la probabilità delle CORRECT REJECTION (CR = 1-FA).	Alla destra del criterio vi è la probabilità degli HIT (H), mentre a sinistra del criterio sotto la curva rossa vi è la probabilità dei MISS (M = 1-H).

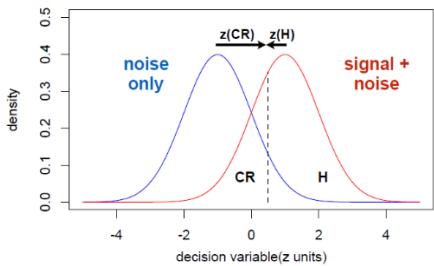
Siamo dunque arrivati a una **concettualizzazione facile da trattare dal punto di vista statistico**: data una certa area sotto la distribuzione normale bisogna trovare il valore sull'asse delle X che corrisponde a quell'area, cioè il quantile della distribuzione normale che corrisponde al punto che divide la distribuzione in quella maniera. Dunque il problema si risolve nel trovare dei valori sull'asse X perché la sensibilità è indicata da quanto bene il sistema risponde al segnale, che è equivalente a dire "*Qual è l'effetto del segnale su quelle distribuzioni?*" cioè di quanto il segnale sposta verso destra la distribuzione rossa rispetto alla blu. **Se un soggetto è particolarmente sensibile a quel determinato segnale allora la distribuzione rossa si sposterà tanto, si differenzierà molto dalla distribuzione blu, mentre se non si è molto sensibili/non lo si è per niente a quel segnale allora le due distribuzioni saranno sempre di più sovrapponibili.**

## Indice di discriminabilità

**La capacità di discriminare, quindi la sensibilità, è misurabile come lo spostamento della distribuzione rossa rispetto quella blu verso destra.** Dato che lo spostamento avviene in orizzontale, per misurare questa distanza devo trovare il punto sull'asse X che corrisponde al picco della distribuzione rossa e al picco della distribuzione blu, indipendentemente dal criterio. **La distanza tra i due picchi delle due gaussiane è la sensibilità;** l'indice per indicare la discriminabilità/sensibilità nella teoria della detezione del segnale è  **$d'$**  ( **$d$  prime**). **Il  $d'$  è una misura di sensibilità.**



**modello gaussiano con uguali varianze**



Per capire a livello statistico questa distanza è necessario studiare le due distribuzioni separatamente, prima quella del rumore (N, in blu) e poi quella del segnale (S, in rosso). Se prendiamo in considerazione soltanto la distribuzione blu allora il suo picco corrisponde alla distanza che c'è tra il quantile che corrisponde a C (il segmento tratteggiato verticale) e al quantile che corrisponde al picco. Quella distanza in unità standard di punti Z corrisponde al numero di rifiuti corretti (CR)  $\rightarrow z(CR)$ .

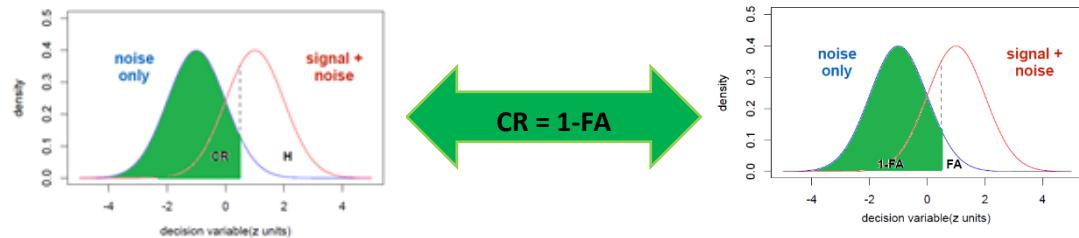
Infatti i punti z in una distribuzione normale in corrispondenza della media (del picco della distribuzione) corrispondono a 0, quindi dal picco andando verso destra saranno di valore positivo di modulo pari al massimo a 0,5 (che corrisponde alla metà della distribuzione) (lo stesso vale andando verso sinistra, dove si trovano valori negativi e di modulo al massimo pari a 0,5) e questa distanza può essere trovata stimando il punto z.

Se ci troviamo nella distribuzione del segnale, ovvero quella rossa, il discorso è il medesimo, infatti la distanza che intercorre tra C e il picco della distribuzione sarà  $z(H)$  e con questa semplice trasformazione trovare il quantile a partire dalla proporzione (che determino empiricamente dai dati) è abbastanza facile.

Nell'esperimento stimo la proporzione e poi trovo il quantile, quindi a partire dall'area trovo il punto sull'asse delle X, poi sommo quelle due distanze e ottengo  $d'$ .

$$d' = z(H) + z(CR)$$

Questo calcolo si può fare in molti modi e visto che solitamente si utilizzano hit e falsi allarmi e non correct rejections si attua una trasformazione:



La proporzione dei rifiuti corretti (CR) è uguale a  $1 - \text{la proporzione Falsi allarmi (1- FA)}$ :

$$d' = z(H) + z(1 - FA)$$

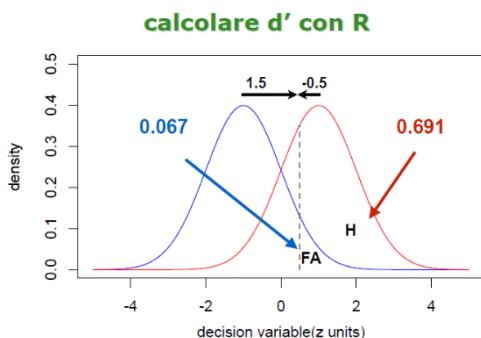
Inoltre, dato che la distribuzione normale è **simmetrica**, quindi i punti Z rispetto allo 0 crescono in maniera positiva a destra e in maniera negativa a sinistra in maniera simmetrica: questo significa che la distanza di  $z(2) = z(-2)$ . Quindi  $z(1-FA)$  è uguale a  $-z(FA)$ :

$$d' = z(H) - z(FA)$$

### Calcolo della formula

$$\begin{aligned} CR &= 1 - FA & z(1-FA) &= -z(FA) \\ d' = z(H) + z(CR) &\longrightarrow d' = z(H) + z(1 - FA) &\longrightarrow d' &= z(H) - z(FA) \end{aligned}$$

$d'$  è la differenza tra il quantile corrispondente alla proporzione di HIT (H) e il quantile corrispondente alla proporzione di FALSE ALLARMS (FA), quindi **per trovare  $d'$  bisogna calcolare il quantile a partire da un'area in una distribuzione normale**. Naturalmente si può trovare  $d'$  anche utilizzando la formula dell'inizio, è solo per convenzione che si utilizzano HIT e FA.



HIT: corrisponde all'area a destra del criterio ma sotto la curva rossa e corrisponde a 0.691 (è il 70% della distribuzione).

Mettiamo alla prova questo ragionamento con un esempio numerico sfruttando R.

Esempio: Abbiamo già calcolato la proporzione di HIT e di FA. FA: corrisponde all'area a destra del criterio sotto la curva blu, ed è pari a 0.067.

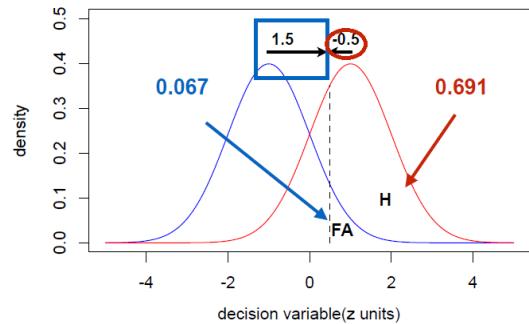
A partire dalle due proporzioni (aree) è necessario ricavarne il quantile corrispondente:

Il quantile corrispondente ad un'area è pari alla distanza dal picco, nel nostro caso:  $FA = z(FA) =$  La distanza dal picco della distribuzione blu al criterio C è pari a 1.5 (riquadro in blu).  $H = z(H) =$  La distanza dal picco della distribuzione rossa al criterio è pari a -0.5 (ellisse in rosso).

Per farlo in R, cioè calcolare d':

```
fa <- qnorm(0.067)
h <- qnorm(0.691)
dp <- h-fa
dp
[1] 1.9972
```

**Stima di d'**



Questo è il calcolo che ho fatto io, perché secondo me il prof ha sbagliato la slide, in quanto è inutile partire dalla distanza, che è quello che dobbiamo trovare. Comunque, per completezza metto anche i calcoli del prof:

```
> fa <- 1 - pnorm(1.5)
> h <- 1 - pnorm(-0.5)
> dp <- qnorm(h) - qnorm(fa)
> dp
[1] 2
```

I punti z, i quantili della distribuzione normale standard, sono numeri che vanno approssimativamente da -3 a 3 orientativamente. Dunque d' è una misura di sensibilità espressa come una differenza tra due punti z e proprio per questo è difficile che si ottengano valori alti, già un d' pari a 2 è grande. Infatti se troviamo un d' pari a 4, vuol dire che ci troviamo agli estremi di due curve che sono molto distanti tra loro, per avere degli intervalli indicativi di interpretazione:

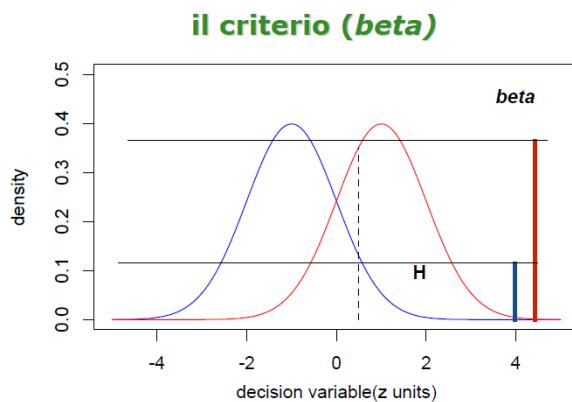
1.  $d' < 1 \rightarrow$  Scarsa sensibilità;
2.  $2 < d' < 3 \rightarrow$  Buona sensibilità;
3.  $d' > 4 \rightarrow$  Errore nei calcoli.

## Calcolo di d' #2

```
dprime <- function(h,fa) {
  zh <- qnorm(h)
  zfa <- qnorm(fa)
  zh - zfa
}
```

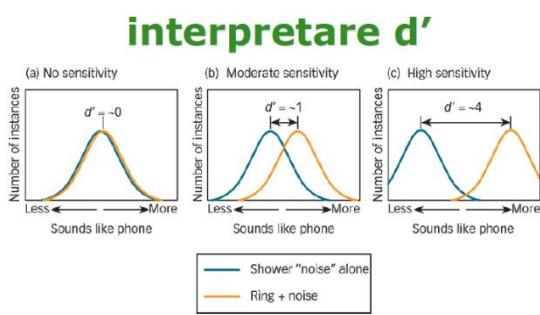
Inserendo dei valori al posto di h ed fa ci viene restituito il valore di d'.

## Calcolo del criterio



Qualche volta si è interessati al calcolo del criterio e ci sono molte maniere di farlo, quello che si utilizza più spesso non fa più riferimento alla distanza delle due curve dal criterio, ma ci interessa capire dove il soggetto ha collocato il criterio rispetto a queste due ipotetiche distribuzioni. Questo si può caratterizzare ad esempio prendendo il rapporto fra due altezze, perché a seconda di dove nelle due distribuzioni l'altezza della curva, che tecnicamente rappresenta la densità della distribuzione normale in quel punto, l'altezza della curva cambierà. Se le due distribuzioni sono sovrapponibili la collocazione del segmento verticale porterà alla medesima altezza del segmento, mentre se si sposta la curva blu verso sinistra allora l'altezza dell'intersezione tra il segmento verticale e le due curve sarà differente. In immagine è possibile vedere che la curva blu corrisponde al segmento verticale blu, che è nettamente più corto dell'altezza ricavabile dall'intersezione tra il criterio e la curva rossa. Il segmento verticale rosso rappresenta la densità della distribuzione rossa al criterio ( $X=\text{Criterio}$ )<sub>curva rossa</sub>, mentre il segmento verticale blu rappresenta la densità della distribuzione blu al criterio ( $X=\text{Criterio}$ )<sub>curva blu</sub> e possiamo notare come quello blu sia molto più breve di quello rosso, ciò significa che il rapporto tra le due altezze diventerà sempre più grande e viceversa. Solitamente comunque in psicofisica si è più interessati a  $d'$  che al criterio  $C$ .

## Interpretare $d'$



Interpretazione di  $d'$ : più è grande  $d'$  più è grande la distanza tra le due distribuzioni; se sono sovrapposte  $d'$  sarà = 0 e questo vuol dire assenza di sensibilità. È il caso dell'immagine a sinistra.

Se sono non sovrapposte ma c'è un'ampia zona di sovrapposizione allora la sensibilità c'è, ma non è grande (es.  $d' = 1$ ). Ed è il caso dell'immagine centrale.

Più aumenta la distanza, più  $d'$  sarà grande e ad esempio un  $d'$  pari a 4 vuol dire una buona sensibilità, difficilmente riscontrabile in un esperimento.

Finita la parte di psicofisica.

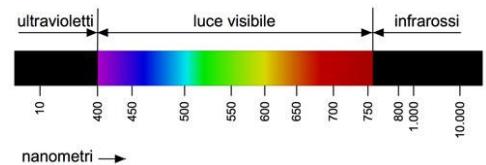
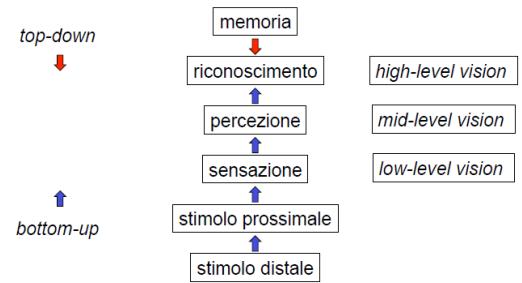
## Seconda parte del corso: Le tappe iniziali della visione

In psicologia si distinguono processi dall'alto verso il basso e dal basso verso l'alto, ovvero processi *top down* e *bottom up*. Questo è molto rilevante quando si ragiona sulla percezione, in quanto è fatta di cose che avvengono a livelli diversi di elaborazione. Non dobbiamo necessariamente pensare a questi processi come *conseguenti nel tempo*, in quanto non è così; sono aspetti che hanno a che fare con elaborazioni periferiche e più vicine al mondo esterno del sistema nervoso, oppure ad elementi più complessi ed integrativi.

Spesso, quando si ragiona nel contesto della **percezione**, si fa una distinzione tra **tre funzioni** della percezione che non sono necessariamente in sequenza temporale e non sono perfettamente distinguibili talvolta tra loro. Però da un punto di vista didattico ha senso dividerle e affrontarle separatamente, tali funzioni per la percezione della visione sono:

1. **High level vision;**
2. **Mid level vision;**
3. **Low level vision.**

Questa suddivisione in tre funzioni rappresenta un'ulteriore scomposizione della dicotomia top-down / bottom-up. La visione di basso livello, che è quella che affronteremo per prima, ha ad a che fare con l'elaborazione iniziale dell'informazione ottica che arriva ai recettori. Nell'analisi della visione di basso livello si fa per primo riferimento a **cosa sia lo stimolo**, cioè quali sono le caratteristiche dello stimolo per la visione ovvero la **luce, ma la luce in che senso?** La luce è una piccola parte dello spettro elettromagnetico quella che può essere captata dal sistema recettoriale dell'occhio, ed è anche detta luce bianca/visibile. Poi bisogna chiedersi come questa informazione viene codificata, non solo nell'occhio, ma in tutti i processi neurali che portano all'elaborazione dello stimolo luminoso. È importante ricordare che noi non vediamo con l'occhio, bensì con il cervello, anzi, vediamo con il sistema visivo che è diviso in molte parti diverse che interagiscono tra di loro.



Quindi in termini prettamente **fisiologici**, parliamo della **via retino-genicolato-striata**, cioè la via che va dalla retina (codifica/trasduzione dell'informazione visiva) al genicolato (struttura intermedia) e che poi proietta su uno strato di V1, cioè l'area visiva primaria. ➔ Questo è ciò che un fisiologo definirebbe visione di basso livello.

La visione è quindi un processo cognitivo fatto di riconoscimenti dell'oggetto e che serve a guidare i movimenti, e questi possono avere sia funzione cognitiva (muovendomi scopro qualcosa dell'ambiente che prima non potevo vedere) che performativa (serve a far accadere qualcosa fisicamente). Tutto ciò avviene dopo la visione di basso livello, entrando nel contesto della visione ad alto livello, dove la visione si interfaccia ad altri processi cognitivi, come per esempio la memoria o l'attenzione. “Riconoscere”, che per **David Marr** (fisiologo della visione) vuol dire “vedere”, vuol dire mettere in relazione l'informazione in entrata con qualcosa che è già presente nella nostra memoria (per es. il riconoscimento di un volto).

Per visione di **alto livello** si intende anche la modulazione dei processi percettivi attraverso l'attenzione; per esempio secondo alcuni studiosi l'esperienza cosciente che abbiamo attraverso la percezione del mondo e la nostra consapevolezza del mondo stesso, richiedono la locazione di risorse attentive. In tal modo si afferma che senza attenzione non c'è coscienza, e senza di essa non potrebbe avvenire alcun riconoscimento.

Pertanto **High level vision** e **Low level vision** sono abbastanza chiari e definiti come concetti. Molto meno ovvio è cosa si intenda con **Mid level vision** perché molto più complesso; è un concetto che si è fatto strada negli ultimi anni del '900, in quanto ci si è resi conto che tale livello di elaborazione è indispensabile, che non è ancora influenzato da processi cognitivi superiori, ed è comunque più complesso e di natura più integrativa e globale rispetto alla semplice codifica dell'informazione che è il prodotto della visione di basso livello. Per gli psicologi della percezione quasi sempre si parte dal livello intermedio, perché esso racchiude i problemi più interessanti.

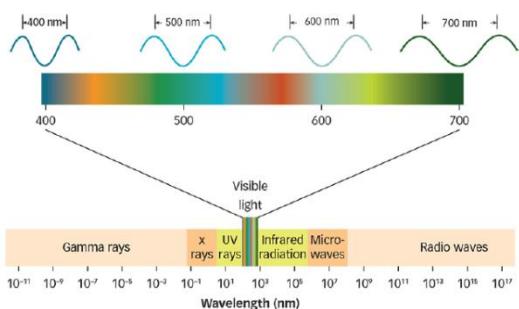
## Livello intermedio o Mid- Level Vision:

Ci rifacciamo alla proprietà del sistema nervoso della **convergenza**: muovendosi verso la corteccia il cablaggio dei neuroni converge verso un unico neurone che è in grado di integrare l'informazione, e i campi recettivi delle cellule diventano sempre più ampi, meno legati allo schema di riferimento retinico e più legati allo schema di riferimento esterno. **Nella visione di basso livello è utile ragionare sui campi recettivi della singola cellula, mentre nella visione di livello intermedio si comincia a ragionare su campi recettivi molto più complicati che risultano dalla convergenza di singole cellule, ed hanno la capacità di codificare proprietà più globali.** ➔ Non è molto semplice comprendere a cosa ci si riferisce quando si parla di livello intermedio, ma un modo per differenziare visione di basso livello e visione di livello intermedio è quello di far riferimento al singolo neurone e porsi la domanda: **“Cosa è in grado di codificare questa singola cellula con l'informazione contenuta nel suo campo recettivo?”**

Esempio: leggi di organizzazione della Gestalt; sono un esempio di visione di livello intermedio, cioè un processo integrativo e organizzativo dove vengono messe insieme più cose.

Pertanto la visione di livello intermedio è fondamentale perché responsabile dei **processi di unificazione/segregazione** cioè processi che di fatto creano unità di rappresentazione che hanno un significato biologico.

Cominciamo quindi concretamente a parlare delle **tappe iniziali della visione**:



La visione non inizia nell'occhio ma prima dell'occhio; questa è stata una conquista della metà del '900 dove si è iniziato a capire che per comprendere come funzionano la percezione, e naturalmente la visione, è utile ragionare su come è fatto lo **stimolo**, che per la visione è l'energia elettromagnetica, che è presente in natura in una gamma di

lunghezze d'onda enormi (per es. raggi  $\gamma$ , infrarossi, micro onde, onde radio ecc...) e che noi usiamo in molti modi. In fisica possiamo descrivere la luce in modi diversi a seconda del fenomeno: a volte è utile descriverla come **pacchetti discreti di energia** e quindi come *fotoni* o *quanti di luce* (meccanica quantistica), mentre per altri fenomeni è più utile descrivere la luce come un fenomeno di natura **oscillatoria**, cioè come un'onda, pensata come un raggio di luce che è orientato nello spazio.

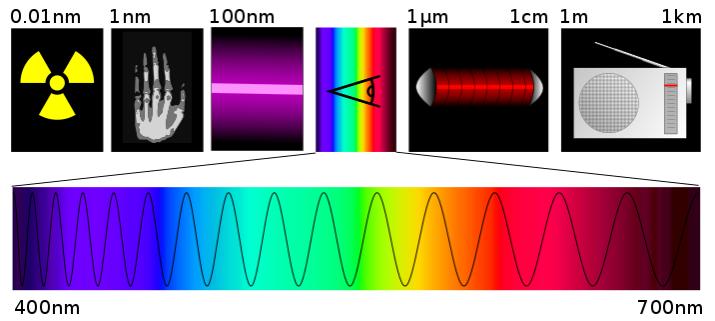
Per uno studioso della percezione è quest'ultima chiave interpretativa che è più utile per spiegare i fenomeni percettivi ed è anche la maniera in cui l'uomo ha sempre pensato la luce. Le **teorie classiche** su come funzionava la visione erano tutte basate sull'*idea del raggio di luce*, anche se gli studiosi del tempo non avevano la minima idea di cosa fosse in realtà la luce e pensavano che essa uscisse dall'occhio e “andasse a toccare l'oggetto”, questa è la *teoria estromissionsitica* della visione che anche **Euclide** sposava. Vi era anche una teoria alternativa di **Democrito** secondo la quale dagli oggetti si stavano delle immagini che viaggiavano nello spazio e l'occhio catturava queste immagini.

È stato solo nel '600 che gli scienziati si sono resi conto che **l'occhio funziona come una specie camera oscura**, ovvero un sistema formato da una lente che mette a fuoco i raggi di luce sul retro della lente facendo esperimenti su occhi di animali morti.

Dunque per i fisici la luce ha questa doppia possibile natura, ma per gli studiosi della percezione **la luce è un insieme di raggi**, che devono stare su un gamma piccolissima dello spettro elettromagnetico per essere utili alla visione.

**Immagine:** è lo spettro elettromagnetico, il range di variazione della lunghezza d'onda di tutto lo spettro elettromagnetico va da  $10^{-11}$ nm (raggi  $\gamma$ ) a  $10^{17}$ nm (onde radio).

Ma noi non siamo equipaggiati per rilevare visivamente l'energia elettromagnetica in tutta questa gamma, la luce bianca/visibile è quella piccola porzione di spettro che va da onde con lunghezza d'onda pari a 390 (il blu/viola) a 700nm (il rosso).



Nello spettro visibile la parte più lunga dello spettro elettromagnetico è il rosso, la parte intermedia è giallo-verde mentre la parte più corta è il blu. Questa è una maniera facilitata di descrivere lo spettro elettromagnetico, ma l'idea che la luce sia colorata è completamente sbagliata: **la luce non è colorata**, è soltanto energia elettromagnetica. Il colore è una proprietà della nostra percezione e vediamo gli oggetti che interagiscono con la luce grazie alla luce stessa: parlare di “luce colorata” è sbagliato, quello che dovremmo dire è che la luce ha una certa gamma di lunghezze d'onda che sono concentrate in una piccola parte dello spettro.

Due motivi a sostegno del fatto che la luce non sia colorata:

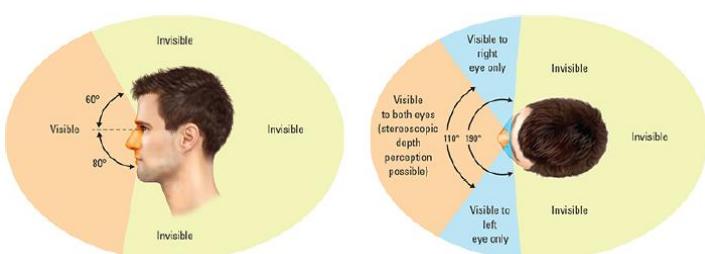
1. Lo stesso **Newton** nell’”*Ottica*” (1704) afferma che i raggi di luce di per sé non sono colorati, i colori sono qualcosa di diverso dalla luce;
2. È facile dimostrare che il colore percepito è indipendente dalla composizione spettrale; è possibile, infatti, trovare esperimenti e dimostrazioni in cui si fa arrivare all’occhio la stessa composizione spettrale di luce, ma si vedono colori diversi, ed è anche possibile l’opposto ovvero dove l’energia elettromagnetica è diversa si vede lo stesso colore.

**Quindi il colore è altro e attiene al dominio della percezione e non a quello della fisica.**

Domanda studente: la stessa lunghezza d’onda è percepita in modo diverso da diverse specie animali? (api, gatti...)

Bruno: un ratto, un falco, un daltonico... non sono in grado di fare certe discriminazioni ma sono in grado di farne altre ma queste sono tutte speculazioni perché non sappiamo bene come questi soggetti percepiscono il mondo.

Posto che ciò di cui la luce è composta è energia elettromagnetica dobbiamo ora interrogarci su come il nostro organismo è in grado di ottenere informazioni da questa energia e un concetto molto utile in questo ambito è quello di **campo visivo**.



**Definizione:** è la zona di spazio da cui posso ricevere informazione ottica.

Il campo visivo ha estensione in verticale sull’ordine di 140 gradi e in orizzontale di circa 180 gradi. Noi vediamo con due occhi quindi abbiamo due campi visivi mono-oculari ampiamente sovrapposti; quindi c’è un campo visivo binoculare che viene “visto” e che manda informazioni ad entrambi gli occhi, mentre in periferia ci sono zone monoculari, viste da un occhio solo. Importante: la luce viaggia ad una velocità molto grande rispetto ad altre velocità che sono rilevanti al nostro comportamento (*300 mila km al secondo* è la velocità della luce). È velocissima, quindi rispetto alla scala di grandezze rilevanti per la nostra nicchia ecologica la luce è quasi istantanea. In generale in un qualsiasi ambiente, in tutti i punti dello spazio arrivano un numero infinito di raggi di luce contemporaneamente che provengono da tutte le direzioni possibili e questo è fondamentale del perché la visione è così adattiva e ricca di informazioni per noi.

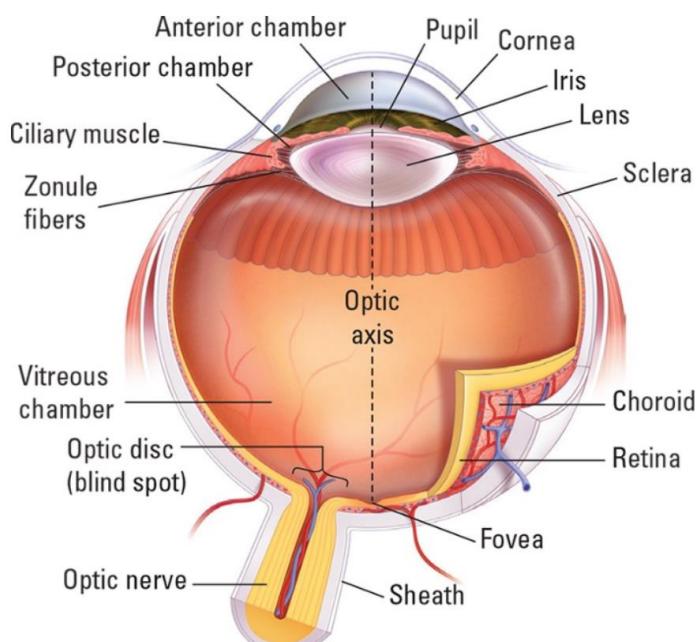
Paradosso filosofico dell'albero che cade: se nella foresta non c'è nessuno, **un albero che cade farà rumore?** Se non c'è qualcuno che può sentire questo rumore, è davvero esistito? Per certi versi questo è ciò che accade nella visione, perché l'informazione ottica esiste anche se l'occhio non c'è in quel momento ad assistere alla scena.

L'occhio non misura tutta l'energia elettromagnetica, ma è in grado di elaborare solo quella parte di informazione compatibile con la sua sensibilità spettrale e questo dipende dalle **opsine** contenute nei fotorecettori che sono sensibili a determinate lunghezze d'onda e non ad altre. Lo stimolo per la visione è molto ricco, ma per cercare una definizione soddisfacente per descrivere lo stimolo della visione bisogna utilizzare un **approccio sistematico**, cioè non bisogna ragionare su una singola proprietà ma bensì sull'interazione; in questo caso c'è il livello fotochimico di un organo e di come l'occhio è fatto, e il livello fisico, cioè ciò che è la luce.

Dalla retina quindi parte la visione, lo stimolo di partenza è la luce e tutto ciò ha senso solo in una visione sistematica.

Il campo visivo ha la sua particolare forma che dipende dalla struttura dell'occhio stesso, l'occhio è un sistema la cui funzione principale è quella di **mettere a fuoco la luce su una superficie** che è dotata di fotorecettori.

L'occhio ha due lenti:

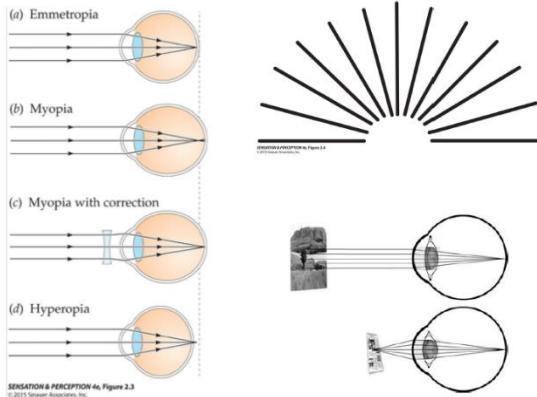
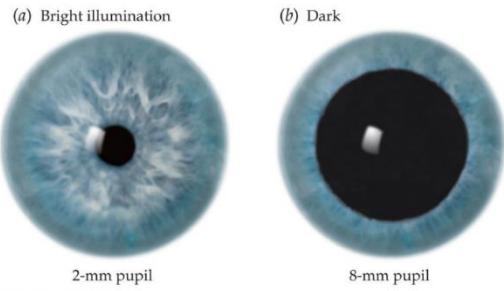


1. **Cornea:** ha un potere di rifrazione (capacità piegare i raggi) molto grande;
2. **Cristallino** struttura molto importante perché permette l'**accomodazione**.

La luce passa attraverso la pupilla e poi attraverso il cristallino, dopo aver attraversato il liquido che mantiene in forma l'occhio, arriva alla parte posteriore dell'occhio (la **retina**) e ci sono anche dei vasi capillari che si occupano del nutrimento della retina.

La **retina** è una struttura sottile fatta di più strati da cui esce un fascio di fibre, cioè il **nervo ottico**, che porta appunto l'informazione verso il **genicolato**. L'occhio è una struttura tridimensionale simmetrica; l'asse ottico è quella direzione che passa attraverso la pupilla e arriva ortogonale alla superficie della retina, è importante l'asse ottico perché nella sua intersezione con la retina troviamo un piccolo avvallamento che è la **fovea**.

La **pupilla** si dilata e si restringe a seconda della quantità di luce, ma questa caratteristica non è molto importante ai fini di ciò che noi stiamo studiando.

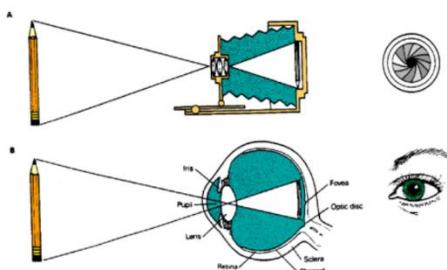


Un altro fattore non molto importante negli studi sulla percezione, ma importante ai fini dell'indagine di patologie, è che l'occhio può non essere sempre perfetto, generando per esempio miopia o ipermetropia. Queste patologie sono dovute al fatto che l'occhio è o troppo lungo o troppo corto e la luce tende ad essere messa a fuoco o davanti o dietro la superficie retinica e per correggere questo difetto strutturale bisogna correggere il sistema ottico aggiungendo un'altra lente, ovvero gli occhiali.

Tutto ciò ha a che fare con una funzione di natura proiettiva, ottica... ed è poco interessante perché ha poco ad a che fare con la funzione di natura percettiva.

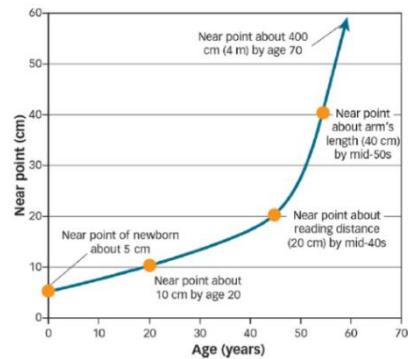
L'astigmatismo invece è dovuto al fatto che la cornea non è perfettamente simmetrica e quindi alcune direzioni del campo visivo risultano meno a fuoco di altre e può essere studiato attraverso la "somministrazione" di un'immagine a ventaglio (in alto a destra).

Il cristallino è una **lente elastica**, quindi nell'occhio c'è un meccanismo che permette di sfruttare questa elasticità per cambiare la forma della lente in funzione della necessità di messa a fuoco. Quando avviciniamo un oggetto la forma del cristallino rilassato non è più adeguato per mettere a fuoco l'oggetto vicino, quindi avviene il processo di **accomodazione** che modifica la forma del cristallino e porta a fuoco l'immagine sulla retina. Questo processo è un processo molto efficiente quando si è molto giovani, perché come è anche possibile vedere in immagine, quando si hanno meno di 10 anni si riescono a mettere a fuoco oggetti presentati a 5cm dall'occhio, ma questa capacità decresce con il passare del tempo perché il cristallino diviene sempre più rigido e l'accomodazione meno buona. L'accomodazione è un fenomeno più interessante rispetto alla miopia o all'ipermetropia perché ha ad a che fare con la **distanza di un oggetto**, cioè data una certa età più un oggetto è distante più l'oggetto tenderà ad essere a fuoco o fuorifuoco e quindi qualche studioso ha pensato che questo potrebbe essere uno di quei meccanismi coinvolti nella capacità di valutare le distanze, grazie al fatto che per mettere a fuoco dobbiamo utilizzare la forza muscolare, ovvero dobbiamo usare dei muscoli per cambiare la forma del cristallino e che il cervello possa avere accesso a questa informazione (tensione muscolare dei muscoli ciliari) per "inferire" sulla distanza di un oggetto.



Si tende a pensare all'occhio come a una macchina fotografica, ed effettivamente ci sono delle analogie, in quanto per *mettere a fuoco* entrambi sembrano essere varianti di una **camera oscura**: un raggio di luce entra in un piccolo foro e se si mette una parete alla distanza giusta si riesce a mettere a fuoco l'immagine che viene da fuori e che, tra l'altro, viene ribaltata come nell'occhio.

Il problema è che tutto questo ci fa pensare che la visione inizi con la formazione di un'immagine sulla retina (**immagine retinica**), ma in realtà sulla retina non c'è una immagine, questa è solo una semplificazione: ciò che in realtà succede è che i raggi vengono messi a fuoco sulla retina, ma poi il lavoro che compie la retina è molto diverso da quello che succede quando si impressiona una pellicola con la luce. Tra gli studiosi della visione c'è anche **Keplero**, che fu uno dei primi a vedere sull'occhio di un animale morto un'immagine, e così pensò che la visione fosse il prodotto di più immagini sulla superficie posteriore dell'occhio, ma in realtà ora sappiamo che non è così, quello è solo ciò che lui ha visto.



Infatti se si inizia a pensare alla visione con il concetto di **immagine retinica** allora sembra come se la visione fosse un trasferimento di immagini che vengono passata ad un decisore/interprete nella testa che guarda l'immagine e decide cosa c'è in questa immagine; **ma non è così che funziona la percezione.**

Per capire meglio come funziona la percezione è opportuno spiegare e comprendere bene gli step dell'elaborazione retinica, ragionando ad esempio sul come la retina sia fatta.

## 1. Inversione destra-sinistra e sopra-sotto

La retina inverte destra e sinistra, sopra e sotto e allora se la visione partisse dall'**immagine retinica**, **come è possibile che vediamo il mondo nel verso giusto e non rovesciato?** Questa domanda in realtà ha senso solo se pensiamo alla visione come “un qualcuno che guarda un'immagine impressa sulla retina e decide cosa sta guardando” e per questo è un problema senza senso. Per comprendere come realmente funziona la percezione dobbiamo liberarci dell'idea che sulla retina ci sia una immagine e invece pensare al sistema retino-genicolato-striato come un sistema in grado di campionare ed elaborare la luce.

Da questo punto di vista sono molto interessanti sono degli esperimenti dove si utilizzano **prismi** che alterano la normale mappatura della luce verso la retina. **Cosa succede se mettiamo degli occhiali che girano l'immagine retinica? Cosa si vedrà?** Questi esperimenti hanno dimostrato che noi siamo in grado di adattarci in fretta a qualsiasi tipo di cambiamento in questo tipo di mappatura, inizialmente il mondo è un po' strano (← cit. / io direi: *muoversi in questo nuovo “spazio percepito” è difficile e guidiamo i nostri movimenti in maniera sbagliata/difficoltosa*) ma molto rapidamente siamo in grado di ricalibrare sia la maniera in cui vediamo i movimenti che la maniera in cui vediamo il mondo.

Il primo a fare questi esperimenti è stato uno psicologo americano di nome **Stratton** negli anni '20, questi esperimenti sui prismi sono ampiamente trattati nel libro Bruno-Pavani “*Perception – A Multisensory Perspective*”. Il più famoso ricercatore a studiare questo fenomeno è stato **Ivo Koler** uno psicologo austriaco che ha dedicato molti anni agli studi con i prismi. Visione di un **video** in cui viene illustrato un esperimento sui prismi, dove c'è il soggetto che indossa degli occhiali con dei prismi che invertono la destra con la sinistra e si fa vedere come all'inizio non riesca ad interagire funzionalmente con il mondo, ma dopo due settimane si è completamente adattato. *Nome video: LEAVING IN A REVERSE WORDL*. È un video che riguarda esperimenti sull'adattamento a lungo termine e che mostrano come il sistema visivo sia in grado di adattarsi ad una certa mappatura dell'informazione visiva sulla retina e poi sulla corteccia. All'inizio del video il soggetto che mette gli occhiali/prismi vede invertito destra e sinistra, ma piuttosto rapidamente si vede **l'adattamento** del soggetto. Mano in mano il soggetto smette di percepire il mondo in modo strano, si adatta e tutto gli sembra normale.

Nella parte finale dell'esperimento viene misurato l'**aftereffect**, ovvero l'**effetto postumo**, che è la misura più chiara della bontà dell'adattamento: il soggetto toglie gli occhiali e lo si sottopone ad un qualsiasi test di guida di un movimento e più l'adattamento è stato completo e più la direzione del movimento sarà sbagliata. Gli aftereffect solitamente durano poco e rapidamente si torna alla normalità, il corpo si ricalibra rapidamente a ciò che era stato abituato per tutta la vita.

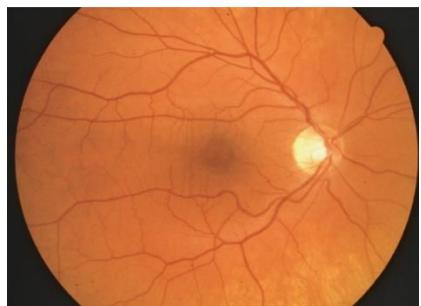
Esiste un uomo che è nato con la testa invertita, quindi percepisce il mondo sopra sotto, ma per lui non è un problema perché è nato così e si è creato una rappresentazione del mondo coerente con ciò che ha sempre visto.

Tutti questi esempi sono stati fatti per sottolineare come **il concetto di immagine retinica non è utile per capire come funziona la percezione** e che effettivamente il problema dell'inversione dell'immagine è veramente un *falso problema* e questo è ancor più vero se si va a vedere la struttura anatomica della retina.



## Struttura della retina dal punto di vista anatomico

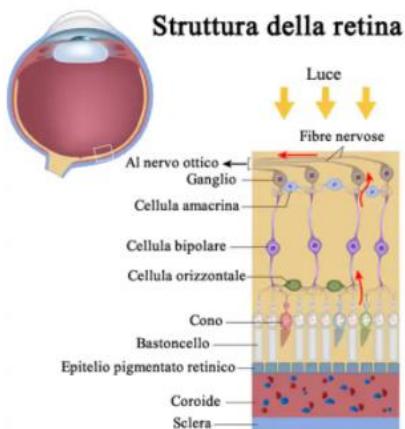
Dal punto di vista anatomico la retina è molto diversa da ciò che si potrebbe pensare, infatti una volta compreso che è costituita da fotorecettori la si potrebbe immaginare come un mosaico di fotorecettori messi uno affianco all'altro, ma in realtà non è così.. La retina ha una struttura **molto disomogenea**.



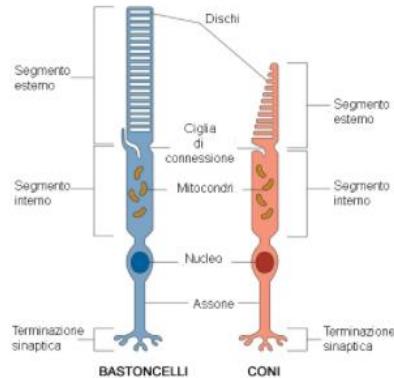
In immagine vi è una fotografia di una retina dove si può vedere molto bene la presenza di un sistema di vene, arterie e capillari molto ramificato e posto sopra lo strato dei fotorecettori, ha un colore arancione e c'è un punto dove il colore è più intenso che rappresenta la **fovea** ed infine un cerchio più chiaro che rappresenta il **disco ottico** o **macchia cieca** perché è il punto dove si raggruppano gli assoni delle cellule gangliari e formano il **nervo ottico**, la via dal quale l'informazione retinica viene convogliata verso il **nucleo genicolato laterale** .

Il motivo per il quale la fovea ha un colore più scuro è dovuto alla struttura stessa della retina, che è fatta a strati:

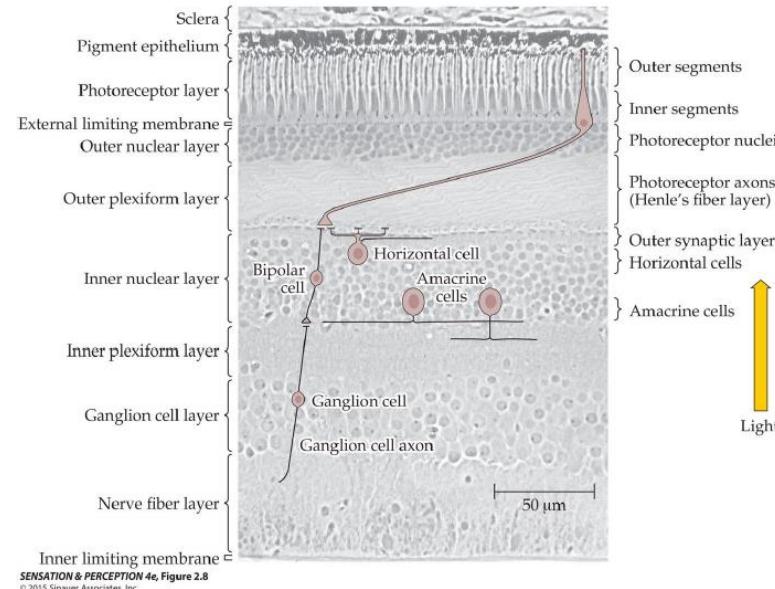
- Nella parte più “profonda” della retina vi è lo **strato di fotocettori**, e considerando la struttura dei fotorecettori, nella parte “ancor più profonda” della retina vi sono i segmenti esterni dei fotorecettori che sono quelli che contengono l’opsina, che è un fotopigmento sensibile alla luce;



In immagine abbiamo una rappresentazione schematica degli strati di cellule della retina. Andando dal basso verso l’alto c’è la membrana interna all’occhio, poi c’è uno strato di fibre nervose, poi vari strati di cellule gangliari e infine abbiamo strati interni dove sono presenti diversi tipi di cellule intermedie dette *bipolari*, *amacrine* e *orizzontali*. Poi solamente a questo punto troviamo i nuclei dei fotorecettori di coni e bastoncelli e poi i segmenti esterni che rappresentano l’epitelio pigmentato che riceve la luce e da cui inizia il processo di fototrasduzione, ovvero la conversione dell’energia elettromagnetica contenuta nella luce in segnale neurale.



- Strati di cellule che filtrano in parte la luce che arriva ai fotorecettori stessi. Nella fovea, che è un piccolo avvallamento, queste cellule sono spostate di lato e permettono quindi alla luce di raggiungerla direttamente e questo giustifica il colore più scuro.



La freccia di destra in figura ci fa vedere che la parte interna dell'occhio è in basso mentre la parte esterna è in alto; quindi la luce prima di arrivare ai segmenti esterni deve attraversare diverse strutture che in parte la riflettono e in parte la assorbono e quindi ci si potrebbe chiedere perché, se l'occhio riflette un'immagine, lo fa su una superficie che riceve un'immagine così deteriorata.

I fotorecettori sono tanti e in generale ci sono molti più bastoncelli che coni. Tuttavia ciò che più è importante sottolineare è che la disposizione dei fotorecettori sulla superficie retinica è ampiamente **disomogenea**.

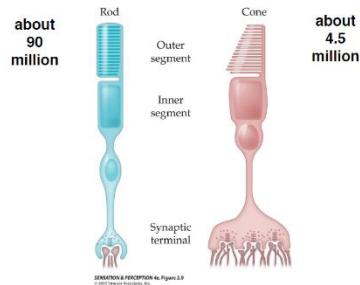
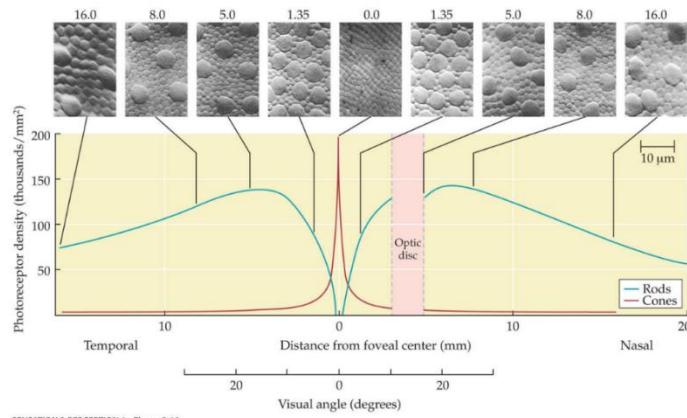


Figure 2.10 Photoreceptor density across the retina



SENSATION & PERCEPTION 4e, Figure 2.10  
© 2015 Sinauer Associates, Inc.

Questo di solito viene rappresentato come nell'immagine: sull'asse delle ordinate (Y) troviamo la densità dei fotorecettori (espressa in migliaia su  $\text{mm}^2$ ), mentre sull'asse delle ascisse (X) troviamo la distanza dalla fovea → **Il grafico rappresenta quindi la densità dei fotorecettori in funzione della posizione sulla retina.** Le due curve rappresentate nel grafico riguardano rispettivamente coni (la curva in rosso) e bastoncelli (la curva in blu).

La distanza sulla retina è misurata in direzione orizzontale, trascurando quella verticale, la retina ha un'estensione di circa  $30\text{mm}^2$ , tuttavia a volte è preferibile misurare la distanza sulla retina in **gradi di angolo visivo**, questo perché la retina è curva sia perché il campo visivo è curvo. Quindi la misura in angoli individua una direzione, questo non è un sistema di riferimento cartesiano (X e Y), ma è un sistema di coordinate polari: direzione (ovvero l'angolo) e distanza (ovvero un vettore) e tramite queste due misure si individuano delle coordinate nello spazio.

Se prendo come 0 la fovea, per la quale passa l'**asse ottico** e che quindi può essere considerato il centro di questo sistema di riferimento spaziale è possibile muoversi sia nella direzione nasale che in direzione temporale in gradi di angolo visivo: ciò che è possibile vedere in immagine è che i coni (in rosso) sono presenti in maniera rilevante solo nella parte centrale della retina che coincide con la fovea e nella sua zona immediatamente circostante; muovendosi di un paio di gradi di angolo visivo e andando sempre più in periferia verso la retina temporale e quella nasale i coni sono molti di meno e variano sensibilmente nella dimensione (vedere immagine dove le immagini in alto rappresentano il **mosaico retinico** a diversa distanza dalla fovea) e notiamo inoltre che in termini di densità i bastoncelli aumentano moltissimo. In fovea ci sono soltanto coni di piccole dimensioni, non ci sono bastoncelli e non ci sono neanche “coni corti”, sensibili al blu; infatti ci sono solo coni rossi e verdi. Spostandosi in periferia cambiano le foto e vediamo dei dischi più grandi e dei dischi più piccoli. I dischi grandi sono coni che occupano più spazio e sono molto più sparsi e quindi meno densi ed infatti la curva rossa decresce drasticamente. Mentre i dischi più piccoli visibili nelle immagini a destra e sinistra dello 0 sono i bastoncelli, anche questi tra l’altro tendono ad aumentare di dimensione spostandosi verso la periferia della retina, ma questa variazione non è così drastica come quella dei coni. Anche la densità dei bastoncelli decresce in funzione dell’**eccentricità retinica** (distanza dal centro della retina), ma in modo meno vistoso rispetto ai coni.

Se analizziamo quanto detto finora innanzitutto si parla di due immagini quella “vista dai coni” e quella “vista dai bastoncelli”, ma comunque si tratta di immagini molto strane, che sono ad alta risoluzione solo in 2° di angolo visivo, in tutto il restante campo visivo l’immagine è a bassissima risoluzione. ➔ Ulteriore prova dell’erroneità del concetto di immagine retinica.

#### *Esempio del prof a lezione*

Per comprendere a quanto corrispondono 2° gradi di angolo visivo: stendete il braccio, alzate il pollice, chiudete un occhio e guardate il vostro pollice, l’unghia del vostro pollice corrisponde a 2° di angolo visivo.

**METAFORA:** La visione è paragonabile all’andatura di un cieco che “scopre il mondo” con il bastone e integra questi tocchi successivi per crearsi esperienza del mondo, il bastone sarebbe la fovea che esplora lo spazio e poi queste piccole immagini di 2° vengono integrate insieme per creare il percepito visivo. Possiamo infine notare come a livello del disco ottico le curve si interrompono perché non ci sono fotorecettori, quindi oltre ad essere a fuoco solamente in 2° dell’angolo visivo questa “immagine retinica” avrebbe anche un enorme buco in mezzo.

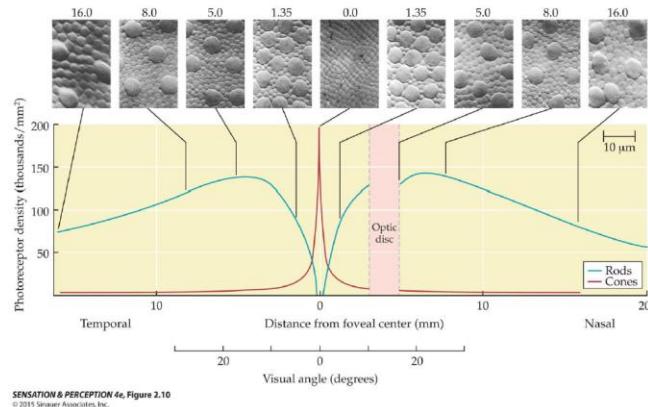
## Lezione 6 – Early e Mid vision – Bande di Mach – Griglia di Hermann

Nell'ultima lezione abbiamo iniziato a riflettere su cosa fa la retina e come funziona la visione, la funzione fondamentale della retina è quella di **trasdurre l'energia elettromagnetica in un segnale neurale**.

Per molto tempo si è pensato che l'occhio e la retina funzionassero come una macchina fotografica, ovvero che fossero in grado di “stampare” delle immagini del mondo esterno e inviarle ad un livello successivo, più centrale e sofisticato, che si occupa di elaborarle. Questa teoria vede la visone come un processo a 2 stadi, percezione ed elaborazione, un processo che “*cattura*” le informazioni del mondo esterno sotto forma di *immagini* e un secondo processo di “*ragionamento*”.

Questa idea è molto comune anche in filosofia in cui vengono distinti dati di senso “grezzi” e le percezioni, più cognitive. Tuttavia analizzando cosa fa la retina dal punto di vista neuroanatomico e dei segnali neurali che producono le cellule retiniche ci si rende conto che questa idea non ha senso.

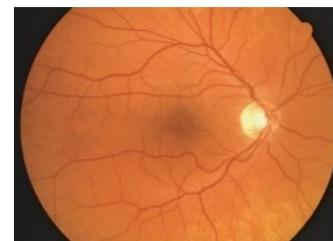
Figure 2.10 Photoreceptor density across the retina

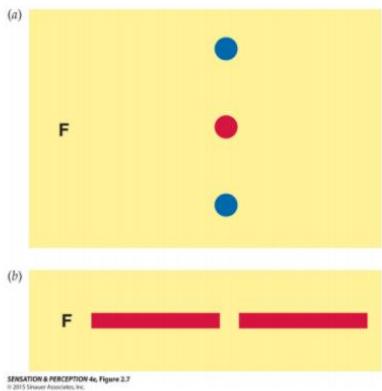


porzione del *campo visivo* più grande. Ciò dovrebbe implicare una minore risoluzione dell'immagine a colori (tipica della visione diurna) nel complesso, eccezion fatta per la parte messa a fuoco dalla fovea. Tuttavia nella percezione cosciente, quotidiana, non abbiamo questo fenomeno e il mondo sembra perfettamente dettagliato.

Problema a cui abbiamo accennato l'ultima volta è quello della **macchia cieca o blind spot**: la macchia cieca è perfettamente visibile con un *oftalmoscopio* (strumento oculistico per osservare il fondo dell'occhio) e risulta come un disco bianco al lato della fovea (nella retina nasale) e più grande di questa.

In particolare possiamo notare come la **superficie dei fotorecettori nella retina sia notevolmente disomogenea**. Abbiamo recettori diversi (**coni** e **bastoncelli**) e distribuiti in modo differente: i coni si addensano principalmente nella parte centrale, la **fovea**, e sono molto meno nel resto della retina. I coni sono anche notevolmente più grandi, quindi integrano informazioni ottiche (sotto forma di energia elettromagnetica) da una





In questa macchia non è possibile l'acquisizione di info visive in quanto non ci sono fotorecettori. Naturalmente la parte di campo visivo che cade all'interno della macchia cieca di un occhio non ricade all'interno della macchia cieca dell'altro occhio.

(Osservando con un solo occhio la slide nel punto di fissazione “F” e ponendo la slide alla giusta distanza arriverete a un punto in cui il pallino accanto ad F scompare, perché rientra nella macchia cieca).

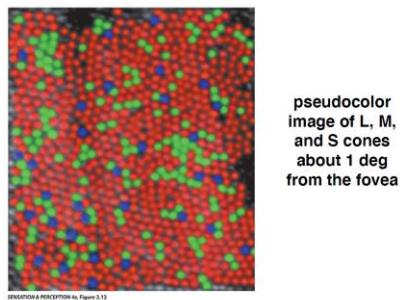
Ne risulta che **il campo visivo monoculare di un occhio include, oltre alla parte periferica, anche la zona corrispondente alla macchia cieca dell'altro occhio.**

Nella quotidianità non ci accorgiamo della presenza di tale macchia in quanto la nostra esperienza del mondo è il risultato di una **integrazione di informazioni binoculari**, che avvengono già nella corteccia visiva primaria, anche se non si sa bene come ciò avvenga. Interessante è che questo **“riempimento” della zona cieca avvenga in maniera “intelligente”**.

Ossia il cervello (che normalmente ha informazioni dall'altro occhio), anche in assenza di info provenienti dall'altro occhio, fa una ipotesi ragionevole sulla base delle info presenti nel contesto (osservando F noi non vediamo un buco nero al posto del pallino ma vediamo il colore dello sfondo che “invade” la zona dove c’è il pallino). Ancora più interessante è se osserviamo la figura in basso nella slide (dove c’è la riga rossa), possiamo notare che la zona cieca viene riempita non con lo sfondo ma avviene un completamento della riga che non risulta più spezzata ma un tutt’uno. Ne risulta che questo riempimento è già un processo di visione intermedia in quanto è già un processo di tipo integrativo che colma quel gap integrando l’informazione locale con quella trans locale (più globale). Da tali evidenze si è passati all’analisi del principio col quale il cervello scelga di colmare l’informazione mancante con giallo (colore sfondo) o col rosso (colore linea). Si ritiene ciò sia dovuto alla **distinzione figura-sfondo** per cui si dà la precedenza alla riga (una cosa, uno stimolo saliente) piuttosto che allo sfondo (una non-cosa). Ciò è stato dimostrato anche ruotando la riga in verticale, quindi, a prescindere dall’orientamento del segmento, il gap viene riempito completando la figura.

Ne risulta che già nella early vision inizia una elaborazione dell'immagine, o comunque ciò avviene nel passaggio tra early-vision e visione intermedia. Quindi, già a livello della retina parte una forma molto primitiva di interpretazione (*“Al prof piace pensare che l'occhio già dalle prime fasi compia un lavoro attivo di campionamento delle info visive presenti nella luce a livello di energia, tale campionamento avviene seguendo i limiti strutturali dell'occhio e ciò comporta un processamento già alla base nell’‘impacchettamento’ delle informazioni.”*)

Ne consegue che la nostra interpretazione avviene con meno informazioni rispetto a quelle presenti nell'ambiente. Noi abbiamo la sensazione di percepire tutto in modo chiaro e nitido ma ciò non è vero, sono presenti una serie di stimoli che noi non notiamo e non elaboriamo a livello cosciente, sia che siano **sotto-soglia** (ovviamente ignorati) ma anche stimoli **sopra-soglia**.



In questa immagine possiamo vedere la *distribuzione dei coni* (la seconda immagine non corrisponde alla fovea perché in questa non ci sono i coni C, quelli blu), che hanno la sensibilità spostata nella parte più corta dello spettro visivo).

Abbiamo anche accennato alla differenze che ci sono tra **visione diurna e visione notturna (fotopica e scotopica)** le cui differenze sono elencate nella tabella.

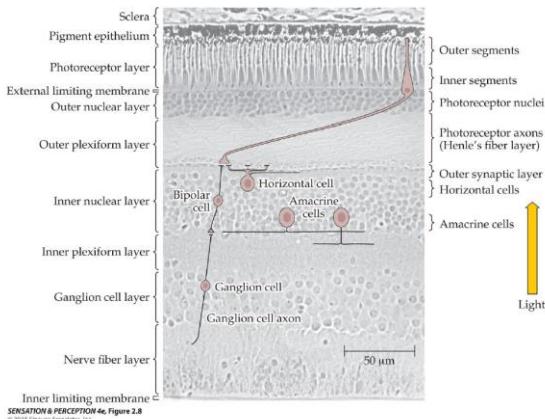
I bastoncelli sono di più, hanno una sensibilità maggiore (ossia una soglia d'attivazione più bassa in quanto richiede meno quantità di luce per attivarsi) e hanno una maggiore risoluzione temporale (utile nella percezione del movimento).

Ultimo punto riguardante la struttura della retina riguarda la trasmissione dell'informazione dai fotorecettori in poi. È importante ricordare che i fotorecettori NON producono potenziali d'azione, bensì producono dei **potenziali graduati**: ossia dei potenziali proporzionali all'intensità della luce. Il potenziale d'azione esce dalle **cellule gangliari** collocate sull'ultimo strato della retina.

**TABLE 2.2**  
Properties of human photopic and scotopic vision

Property	Photopic system	Scotopic system
Photoreceptors	4–5 million cones	90 million rods
Location in retina	Throughout retina, with highest concentration close to fovea	Outside of fovea
Acuity (detail)	High	Low
Sensitivity	Low	High

SENSATION & PERCEPTION 4e, Table 2.2  
© 2015 Sinauer Associates, Inc.



Non ci interessa troppo com'è strutturata la rete neurale e i vari strati della retina in quanto fonte tutt'ora di dibattiti, ci limitiamo a distinguere vari strati composti da **3 tipologie di cellule intermedie**: quelle **bipolari**, quelle **orizzontali** e quelle **amacrine** che sono collegate "in modo molto complicato" alle **cellule gangliari** che possono essere di diverso tipo (**Magno, Parvo e Konio**) che corrispondono a diverse tipologie di input retinico.

Adesso analizziamo la natura neurale del segnale che viene emesso dalle cellule gangliari e quali implicazioni ha nella comprensione dei **processi visivi di basso livello**.

*(Il prof sostiene l'idea dell'osservazione naturalistica, per studiare la percezione bisogna imparare a guardare il mondo con l'occhio dello scienziato, che non si interessa di capire come è fatto il mondo e come possiamo interagirci, bensì ci interessiamo al processo attraverso il quale noi percepiamo, una sorta di metapercezione. Ciò è utile per notare tutta una serie di cose che possono essere utili a formulare delle ipotesi da testare → atteggiamento fenomenologico)*

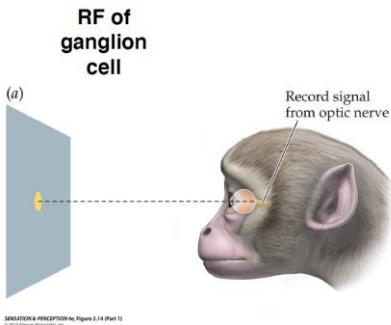


Per analizzare la natura del segnale emesso dalla retina possiamo usare un approccio fenomenologico: un fenomeno che possiamo notare quotidianamente è quello delle **bande di Mach**, un aspetto irrilevante nella nostra interazione col mondo, ma interessante per capire il processo visivo di basso livello e il segnale che esce dalla retina.

*(Mach → unità di misura della velocità di propagazione del suono sulla terra, notevolmente più lento della velocità della luce) Mach propone un ragionamento di come funziona la visione e osserva queste "bande":*

Osservando il bordo inferiore dell'ombra orizzontale potete notare in corrispondenza del bordo sopra una striscia più scura e sotto una più chiara, queste strisce sono le bande di mach. (**Concorso:** chi fotografa delle bande di mach può vincere un punto in più all'esame).



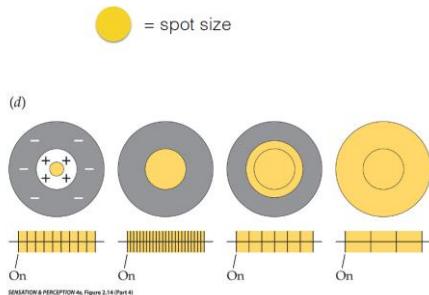


**Perché vediamo queste “bande”?** Già dagli anni '50 è stato possibile studiare il **campo recettivo** dei neuroni del sistema visivo e il segnale che viene fuori dalle cellule gangliari. Il campo recettivo di un neurone del sistema visivo è la porzione di spazio del campo visivo in cui deve cadere lo stimolo per attivare il neurone, ha a che fare con una certa porzione di spazio e ha a che fare con quello che attiva il neurone. I coni, più grandi, sono capaci di ricevere energia elettromagnetica da una porzione più grande del campo visivo. In fovea si ha una maggior concentrazione di coni che hanno un campo recettivo più piccolo e ciò è associato alla maggiore capacità di rispondere ai potenziali di dettaglio. Tramite tecniche di registrazione con microelettrodi è possibile registrare i campi recettivi delle cellule gangliari.

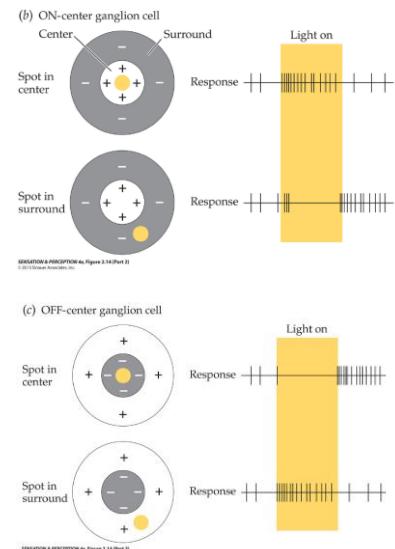
Questo campo recettivo è organizzato in maniera concentrica, disco-anello, opponente. Ci sono 2 tipologie di cellule gangliari: **centro-on** e **centro-off**, le prime se colpite nella parte centrale scaricano più potenziali d'azione e quindi una risposta di tipo eccitatorio, mentre nella parte periferica hanno una risposta di tipo inibitorio, le seconde il contrario.

Queste cellule sono state individuate con esperimenti particolari in cui vengono proposti degli stimoli al soggetto e si “ascolta” (grazie a degli amplificatori particolari) il potenziale d'azione emesso dalla cellula. Questa trasformazione del segnale neurale in segnale acustico ci permette di percepire direttamente, tramite i nostri sensi, il lavoro della cellula.

Spostando lo stimolo nei diversi punti del campo visivo si può valutare il campo recettivo dei neuroni e il tipo di risposta delle varie cellule.



Altro tipo di risultato lo ottengo ponendo il soggetto davanti a uno **stimolo luminoso che man mano si ingrandisce**: se lo stimolo colpisce la parte eccitatoria centrale del campo recettivo si ha un aumento graduale della risposta del neurone, se però lo stimolo si ingrandisce fino a colpire anche la parte inibitoria, paradossalmente più luce non vuol dire più eccitazione e si ha una riduzione della risposta del neurone. A questo punto ci troviamo davanti a una cellula che è molto eccitata, ma anche molto inibita e queste due situazioni si annullano a vicenda.



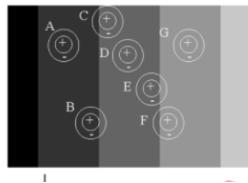
Ciò significa che per una cellula gangliare lo stimolo ideale è uno stimolo disomogeneo, con una parte illuminata e una meno, con una grande differenza di intensità stimolante tra zona ON e zona OFF, in quanto, come detto, un eccesso di luce che colpisce la zona inibitoria causa una riduzione della risposta. **La cellula gangliare misura l'intensità della relazione tra intensità luminose rispetto allo spazio.**

**Misurano il contrasto. Quindi misura delle differenze.**

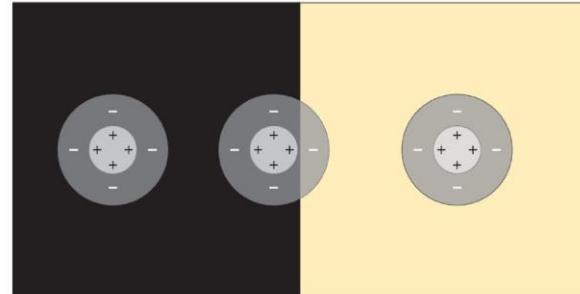


In questa immagine possiamo notare le bande di Mach.

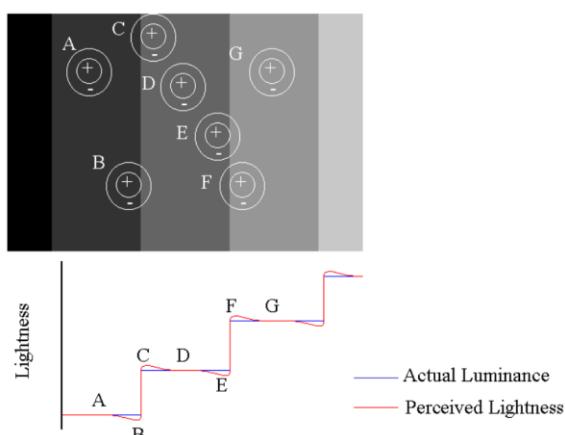
Si nota abbastanza chiaramente un grigio sfumato che a un certo punto presenta una banda più scura accompagnata da una più chiara. In questa immagine possiamo vedere la distribuzione della **luminosità apparente** (“*lightness*” → quanto bianco mi sembra un punto di una immagine). Noi vediamo un andamento “a scalino” con vari livelli di grigio omogenei al loro interno ma nettamente separati gli uni dagli altri. (nero che diventa più scuro, poi più chiaro)



**Quegli avvallamenti in prossimità del bordo sono le bande di Mach, sono la conseguenza della struttura del nostro sistema percettivo.**



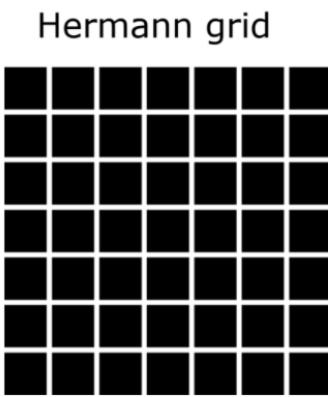
SENSATION & PERCEPTION 4e, Figure 2.15  
© 2013 John Wiley & Sons, Inc.



Se proviamo a sovrapporre a quella immagine dei campi recettivi (contrassegnati da varie lettere). Per esempio “A” riceve una intensità luminosa uguale sia nella parte eccitatoria che in quella inibitoria, mentre “B” ha un campo recettivo che sta a cavallo del confine tra 2 intensità di grigi (più scuro e uno più chiaro). B non ha un equilibrio tra eccitazione e inibizione ma sarà più inibito rispetto ad “A” perché parte dello stimolo più chiaro ricade in parte della zona inibitoria. “D” ha un equilibrio eccitazione-inibizione, mentre “C” è meno inibito in quanto parte del campo inibitorio sta nella zona più scura rispetto a D.

“D” ha un equilibrio eccitazione-inibizione, mentre “C” è meno inibito in quanto parte del campo inibitorio sta nella zona più scura rispetto a D.

A, D e G hanno una risposta uguale, anche se colpite da diverse intensità luminose hanno un equilibrio eccitazione-inibizione uguale: ciò dimostra che l'intensità della luce risulta irrilevante per queste cellule: ***quindi come da uno stesso segnale iniziale percepiamo dei colori diversi?*** Le cellule gangliari percepiscono i contrasti e permettono *a livello trans locale* (nella percezione intermedia) di percepire le forme.



Dando un'occhiata alla **griglia di Herman** possiamo notare tanti quadratini neri alternati da righe e colonne. Nei vari incroci si notano dei quadratini più chiari che però appaiono solo se fissi i quadrati scuri, se cerco di mettere a fuoco sul quadratino chiaro questo sparisce: chiaro esempio di illusione cioè ci accorgiamo che il sistema visivo ci sta ingannando. Abbiamo anche un esempio di griglia di Herman “scintillante” in cui una sovrapposizione di zone grigie su campi bianchi non sempre ma solo in zone periferiche producendo una specie di scintillio.

## Lezione 7 – Spiegazione bande di Mach e griglia di Hermann – Effetto Craik-O’Brien

### Le bande di Mach

*Si possono fotografare le bande di Mach?*

Mach bands



Non è semplice individuare (e ancor di più fotografare) le bande di Mach e tale incertezza può dipendere sicuramente dalla distanza da cui la si osserva. Questo suggerisce che c'è un intervento dei campi recettivi presenti sulla retina che rilevano quella *discontinuità* e quelle variazioni di *intensità di luce*. Va, inoltre, tenuto presente che le bande di Mach sono nel sistema visivo e non lì fuori; per cui la riproduzione fotografica delle bande di Mach ha dei limiti proprio perché non riuscirà mai a riprodurre esattamente le intensità che sono presenti nel mondo.

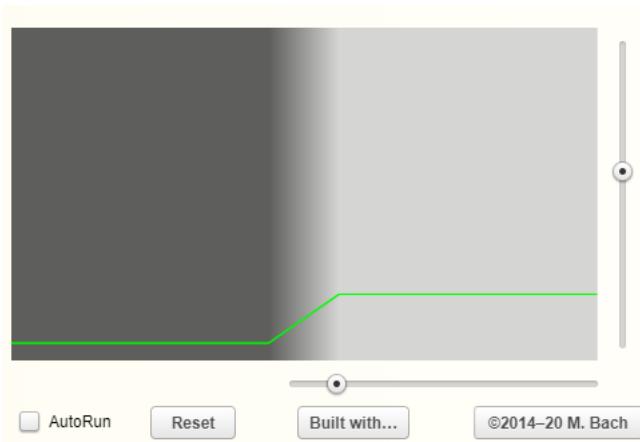
Le bande di Mach dipendono da due dimensioni:

- **Contrasto;**
- **Scala spaziale.**

Nel sito “*Optical Illusions & Visual Phenomena*” del ricercatore Michael Bach, dell’Università di Friburgo, si trovano dimostrazioni di percezione (illusioni e simili) in spiegazioni interattive ([→<https://michaelbach.de/ot/>](https://michaelbach.de/ot/)). Attraverso un approccio psico-fisico ha messo a confronto lo stimolo fisico con il risultato percettivo -ossia, con quello che si vede.

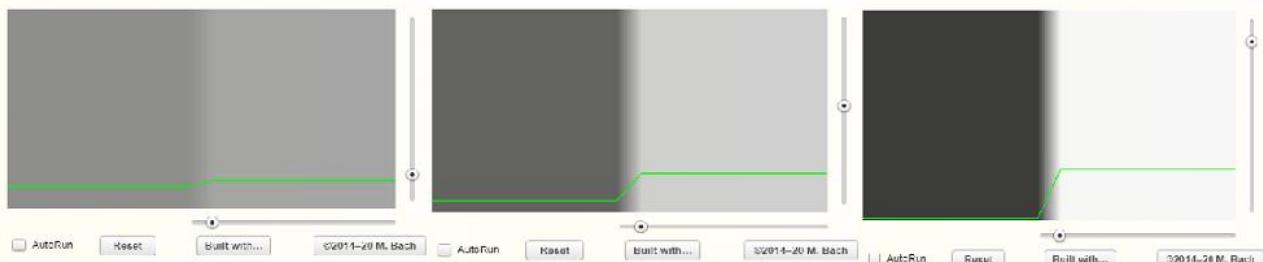
Lo stimolo fisico, in questo caso, è la “spezzata grande” (marcata dalla linea in verde) che esprime una rappresentazione del *profilo di luminanza* di quella immagine, ossia come l’intensità della luce riflessa (in questo caso dallo schermo) varia in funzione dello spazio.

Qui la variazione è solo in orizzontale (varia solo la scala spaziale) e, quindi, si osserva:



- una zona a bassa luminanza (zona a sx scura);
- una rampa che riflette una crescita graduale di luminanza;
- il susseguirsi di una zona in cui la luminanza è ancora più alta (zona a dx chiara).

**Cosa succede aumentando/ diminuendo il contrasto?**



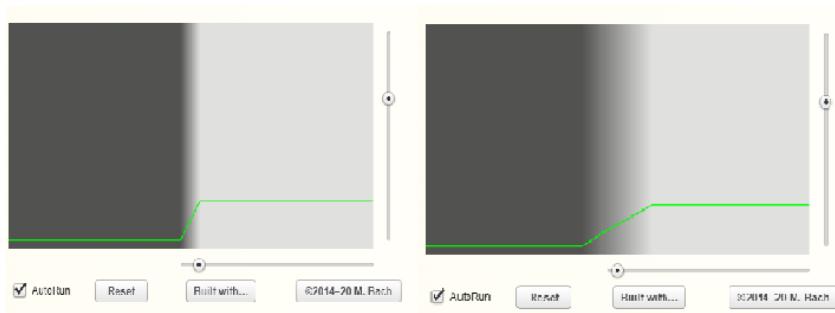
Le bande di Mach si notano nel punto in cui c’è un cambiamento e, quindi, nel punto in cui la luminanza ha una discontinuità (dove inizia e finisce la rampa). Aumentando il contrasto, naturalmente, il nero diventa più nero ed il bianco diventa più bianco; di conseguenza la transizione fra nero e bianco diventa sempre più profonda e più marcata quanto più aumenta il contrasto: questo è il caso in cui le bande diventano più evidenti.

Poco contrasto → bande di Mach poco visibili;

Molto contrasto → bande di Mach decisamente più visibili.

### Cosa succede aumentando/diminuendo la scala spaziale?

C'è solo una rampa, un cambiamento molto graduale senza la brusca discontinuità. Questo cambiamento lo percepiamo come un *gradiente di luce* (ad esempio, come un'ombra sfumata o una superficie reclinata rispetto alla fonte di illuminazione che fa percepire l'ombreggiatura). Quindi pensando di regolare questo fattore, come per il contrasto, avremo alcune discontinuità graduali e altre meno graduali. Ciò succede in una zona piatta con intensità diversa.



In prossimità di queste discontinuità i campi recettivi delle cellule gangliari rilevano uno sbilanciamento di eccitazione/inibizione in una certa direzione o nell'altra.

Nell'immagine in alto a sinistra si vede uno scalino puro, e le bande sono maggiormente visibili. Naturalmente questo esperimento richiederebbe di controllare diverse variabili, tra cui la distanza del partecipante dallo schermo, perché questa distanza determina quanto è grande sulla retina questo particolare stimolo.

### Effetto ottico della griglia di Hermann

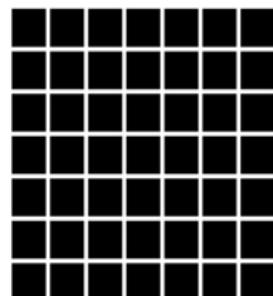
La capacità del sistema visivo di basso livello dipende da:

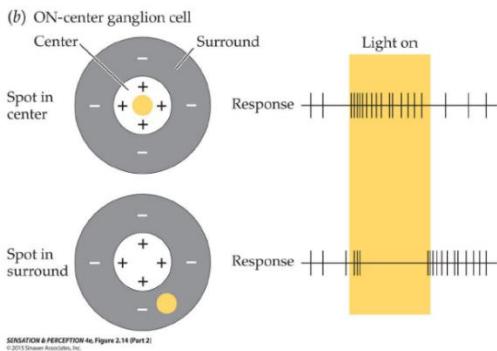
- **Contrasto**, ossia quanto è grande la differenza tra il massimo ed il minimo di intensità (fra la zona nera e la zona bianca)
- **Fattore di scala spaziale**, lo stacco tra le due zone di luminanza.

Prendiamo in considerazione una cellula retinica centro-on per spiegare l'effetto causato dalla griglia di Hermann:

- sulla linea bianca piena la scarica di centro-on è quasi ottimizzata perché il centro è ben illuminato ed il surround -che non deve essere stimolato per permettere alla cellula di funzionare correttamente- resta spento;
- nel punto di incrocio delle linee bianche, invece, il surround viene illuminato quindi la luce, ricoprendo anche la periferia della cellula centro-on, crea un'inibizione di quest'ultima provocando un effetto luminoso minore (rispetto alla stessa cellula colpita al centro, come nel caso precedente).

Hermann grid





più luce ricade nella periferia, più sarà presente un'inibizione della risposta della cellula.

Infatti, nella griglia di Hermann, quando il campo recettivo capta lo stimolo:

- consiste nella linea che fa da *corridoio* (la linea bianca “-” ) avrà un po' meno inibizione rispetto a quando è collocato sull'intersezione, proprio perché quando è collocato sulla linea bianca, l'anello inibitorio del campo recettivo, per lo più, becca il nero e quindi poca luce e poca inibizione (infatti, la cellula risponderà abbastanza bene);
- quando, invece, lo stimolo consiste nell'*intersezione* (dove le linee della griglia si incrociano “+”) la periferia inibitoria del campo recettivo prende del bianco sia a destra che sinistra e sia sopra che sotto (perché essendo una croce ci sarà in proporzione molta più inibizione).

### ***Il colore percepito dipende dall'attivazione delle cellule gangliari?***

L'assunzione di base per spiegare questo fenomeno è problematica; ci sono delle dimostrazioni contrarie come, ad esempio, le *griglie dalle forme irregolari* (es: al posto dei quadrati ci sono rombi) per cui, in certe condizioni, l'effetto può sparire completamente anche se continuano ad esserci corridoi, intersezioni e croci; quindi, non dovrebbe fare nessuna differenza.

La griglia di Herman non è interessante solo per il fatto che si vedono delle ombre scure nelle intersezioni, ma il punto cruciale risiede nel fatto che le ombre “tendano a sfuggire” e cioè che se provi ad inseguirle per fissarle, non le vedi.

Negli anni '70 e '80 la comunità scientifica che si occupava di questo genere di ricerca ha cominciato a studiare questi meccanismi molto basilari che servono a codificare le discontinuità nella luce e, quindi, risulta semplice pensare che un effetto come quello della banda di Mach sia, in qualche modo, coinvolto in quei meccanismi del nostro sistema visivo che ci consentono di rilevare i bordi.

L'effetto della **banda di Mach** sembra la conseguenza di un meccanismo cellulare per cui in presenza di una discontinuità questo cerca di esaltarla e di renderla più visibile, migliorandone il contrasto; la differenza di luminanza in prossimità del bordo viene accentuata.

Le cellule gangliari sono **centro-periferia/disco-anello**. Immaginiamo che queste cellule centro-on (e periferia-off) abbiano il campo recettivo scalato sulle dimensioni della griglia; il diametro del disco sarà pari alla distanza fra i quadrati neri della griglia, quindi la linea luminosa ricade nel centro del campo recettivo. La periferia-off ha un effetto inibitorio e questa inibizione, di nuovo, dipende dall'intensità della luce:

**NB:** La retina non fotografa lo stimolo ma produce già una prima elaborazione dell'informazione, ossia cercare di:

- rendere più evidenti certe discontinuità
- estrarre dove ci sono dei cambiamenti interessanti.

A rendere più complessa questa elaborazione ci sono elementi come il “fattore di scala”, per cui certi cambiamenti risultano più interessanti se avvengono ad una certa scala, in proporzione alle altre scale di cambiamenti che ci sono nell'immagine.

Questi meccanismi sono, quindi, “locali” e cioè avvengono localmente in una porzione della nostra retina da cui poi viene propagato il segnale. La percezione non dipende solo da queste informazioni locali, chiaramente ci sono degli aspetti più globali che entrano in gioco quando si presentano illusioni del genere. Il fatto che l'intersezione caschi in **fovea** (dov'è presente una ricca concentrazione di coni) e che quindi quando si prova a fissare lo stimolo<sup>3</sup> questo tende a scomparire, significa che questo stimolo va a finire nella zona di massima acuità visiva della retina, ricoprendo tutto il campo recettivo della cellula e creando delle “macchie cieche” che si distribuiscono al centro delle intersezioni del resto della griglia e che sfuggono quando si cerca di fissarle.

Il campo recettivo delle **cellule gangliari** presenti in fovea è più piccolo perché riceve informazioni da fotorecettori che sono distribuiti in maniera più densa e che, quindi, sono piccoli anche loro. La **risoluzione** (o Scala) del *mosaico di fotorecettori* presenti in fovea è migliore e, quindi, anche la **risoluzione** delle **cellule gangliari** è migliore.

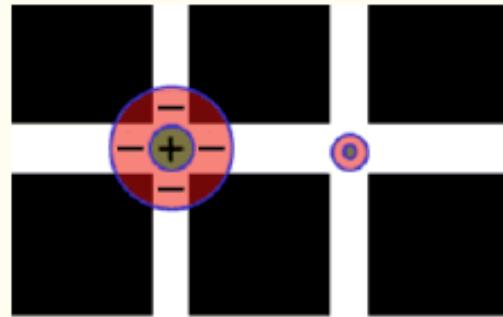
**NB:** la cellula gangliare confronta tanti output di una certa porzione di fotorecettori che vanno a finire nella sua zona-on (centro) con quelli che sono nella zona-off (periferia): misura il contrasto rispetto allo spazio, operando così un **confronto spaziale**.

---

<sup>3</sup> quella zona di ombra/ pallini scuri che si presenta quando si guardano le “+” e che quando si cerca di fissarle tendono a scomparire.

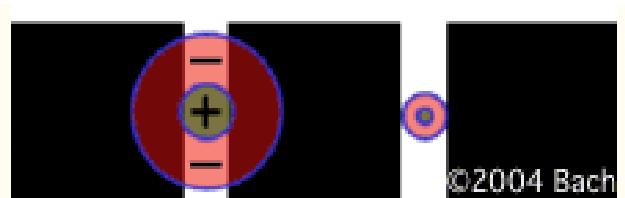
Quindi, se:

Lo **stimolo cade in fovea** → sui campi recettivi delle cellule gangliari centro-on avviene qualcosa di molto simile a quando fisso *l'intersezione* della griglia di Hermann: trovando la giusta distanza, mi potrei trovare in una situazione ideale in cui proprio l'intersezione mi provoca sia il massimo di *eccitazione* e sia il massimo di *inibizione* (perché lo stimolo va a beccare sopra, sotto, a destra, a sinistra del campo recettivo) e, quindi, oltre a la zona eccitatoria colpisce anche la zona inibitoria, motivo per cui vedo le macchie nere sparse.



Lo **stimolo cade in porzioni periferiche della retina**

→ sui campi recettivi delle cellule gangliari centro-on avviene qualcosa di molto simile a quando fisso, invece, un corridoio della griglia di Hermann: in periferia i recettori per la visione diurna (coni) sono molto più sparsi e più grandi rispetto a come si presentano in fovea e, di conseguenza, anche i campi recettivi delle cellule gangliari che li elaborano, in questa porzione di retina, risultano molto più grandi. In periferia, quindi, le cose sono diverse: essendo il campo recettivo delle cellule molto più grande, potrebbe succedere che la stessa zona eccitatoria non risulti abbastanza eccitata (perché il campo recettivo prende, per esempio, un po' degli spazi neri della griglia) per cui la cellula non si attiva come dovrebbe.



Dunque, cambiando la scala del campo recettivo, si va a modulare la quantità di eccitazione e di inibizione; il cambiamento di scala dipende da quanto è grande questa griglia sulla nostra retina e, quindi, dalla distanza da cui osserviamo la griglia.

**NB:** Lo stesso discorso è osservabile nelle bande di Mach: se si cambia l'ampiezza della rampa variando la sua scala spaziale, ci si trova una situazione ottimale in cui queste bande diventano molto evidenti e poi tutta una serie di situazioni intermedie in cui si vedono e non si vedono.

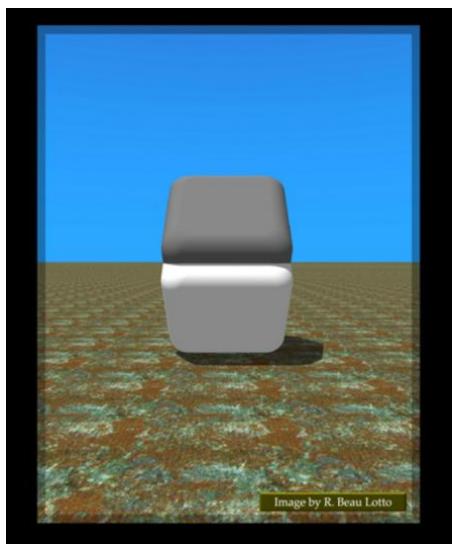
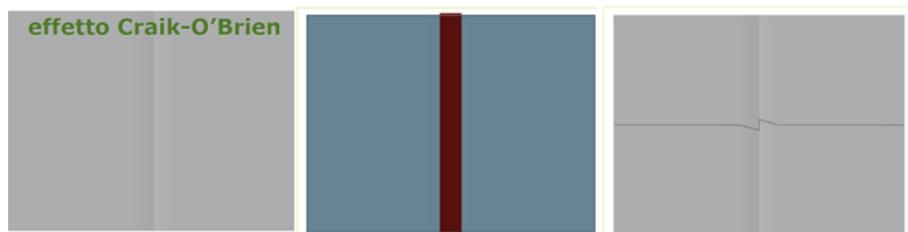
**Il sistema visivo lavora su discontinuità, ossia su punti in cui cambia qualcosa e quella è l'informazione che viene elaborata**, che viene codificata e che poi viene usata per assegnare un colore all'intera superficie. La macchia cieca è un processo di completamento, cioè vengono riempiti dei buchi utilizzando le informazioni del contesto; il cervello *riempie* la superficie sulla base delle informazioni e questa è una cosa veramente importante perché ci costringe proprio a ridefinire la maniera in cui pensiamo allo stimolo per la visione: lo stimolo, per la percezione in genere, è sempre di natura relazionale e anche nel caso della visione perché, anche a livelli base, rileva una discontinuità o un cambiamento e non rappresenta mai una quantità assoluta.

Va bene considerare una quantità assoluta di luminanza che determina il colore ma ciò che importa è come il sistema visivo codifica le discontinuità (e quindi il *contrasto*) e non solo.

### Effetto Craik-O'Brien

Il sistema visivo codifica le discontinuità in termini di “segnali di contrasto” e, quindi, di **differenza di intensità luminosa** ma anche lungo la **dimensione cromatica** (almeno per la visione diurna quando usiamo tutti e tre i tipi di coni).

Nell'effetto Craik-O'Brien osserviamo delle variazioni di luminanza rispetto alla *scala spaziale* (variazione in un'unica dimensione).



Questo è una situazione ben più complessa perché anche questa è un'immagine bidimensionale però, dal punto di vista delle variazioni e delle discontinuità, non sono più unidimensionali ma sono bidimensionali e quindi si possono individuare delle forme, e suggerire ad altri sistemi -che stanno a metà tra l'elaborazione di basso e di alto livello- del vostro sistema visivo una particolare disposizione spaziale di oggetti tridimensionali nell'ambiente.

Infatti, è impossibile non vedere una superficie di supporto, un cielo, un orizzonte, un oggetto tridimensionale ed anche l'illuminazione che proviene da una precisa direzione (dall'alto e leggermente da sinistra che poi è quello che succede sempre nella realtà).

Su questa superficie di supporto illuminata dal sole c'è un oggetto o forse due, uno accanto all'altro, che appaiono tridimensionali; quindi, è una situazione ben più complessa di quella dell'effetto Craik-O'Brien ma ispirata a quest'ultimo perché le due superfici che appaiono bianche e nere in realtà sono ad una stessa luminanza. Si potrebbe dire che in realtà hanno lo stesso colore però questo non avrebbe molto senso, perché stiamo parlando di una situazione complessa in cui la luminanza dipende anche da quanta luce arriva su quelle superfici. L'intensità della luce è la stessa sia per l'oggetto posto sopra che per quello posto sotto.

Ci sono due maniere in cui noi potremmo convincerci di questo:

- la prima sarebbe prendendo un fotometro o un *bidimensiometro* (una specie di video camera che voi puntate contro la superficie e vi dà una lettura della luminanza in quel punto, è un misuratore dell'intensità luminosa effettiva); potremmo puntarlo sul centro del confetto di sotto e di sopra e con nostra sorpresa scopriremmo che viene fuori lo stesso numero.
- tramite lo *schermo di riduzione*, utile e lowcost: basta un cartoncino con un foro da cui poterci guardare, si elimina il contesto e si lascia disponibile solo l'info locale - è ciò che succede approssimativamente nel campo recettivo di una cellula gangliare quando elabora la sua piccola parte dello stimolo.

Per quanto riguarda l'immagine sopra potrei avere uno schermo di riduzione con due buchi in corrispondenza del centro delle due superfici ed a quel punto posso verificare se veramente tutto il contesto, tutto quello che suggerisce la tridimensionalità eccetera, **corrisponda alla vera luminanza oppure suggerisca altro**.

A questo punto abbiamo ridotto abbastanza e si vede chiarissimamente che sono due grigi uguali, cioè sembra quasi impossibile pensare che quello di sotto sia bianco (come appariva all'inizio).



Naturalmente è molto istruttivo ed interessante studiare questi fenomeni di natura locale (come quelli che producono le bande di Mach) però bisogna ricordarsi che questi fenomeni rappresentano i mattoncini che fanno da base a dei processi più complicati (come quello di *riempimento* visto nell'effetto Craik-O'Brien) che hanno a che fare con la rappresentazione di superfici (es. oggetti) e, cioè, di cose che cominciano ad avere una forma ed un colore grazie ad un processo: **l'interpretazione globale di una scena**; questo processo chiama in causa non solo delle cose che io posso studiare in un reticolo (come nella griglia di Hermann o per la scala con cui avviene un cambiamento nel profilo di luminanza e di frequenza spaziale) ma anche delle cose più complicate: le *interpretazioni globali della struttura tridimensionale della scena*

e, quindi, l'interazione di un oggetto tridimensionale con l'illuminazione (in questo processo entra in gioco molto delle conoscenze acquisite tramite l'esperienza). Quello che succede è una cosa di grandissima sofisticazione perché il punto di partenza, naturalmente, è il fatto che vengono codificate le discontinuità orientate.

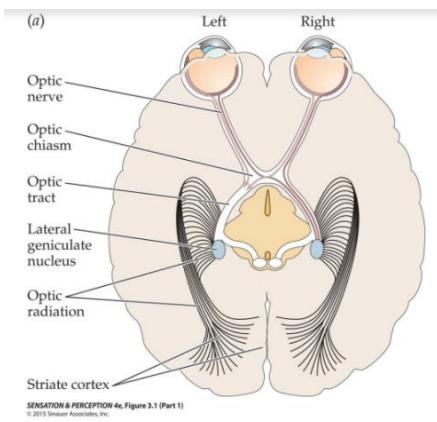
In ultima analisi, osservando ancora l'immagine sopra, si nota una discontinuità allineata con l'orizzonte. Quella è una discontinuità in cui, come nell'effetto Craik-O'Brien, c'è un salto netto da una zona più scura ad una zona più chiara e poi andando verso l'alto il grigio scuro della linea diventa più chiaro fino a stabilizzarsi sul grigio medio e sotto, invece, il grigio scuro della linea tende a schiarirsi stabilizzandosi sullo stesso grigio medio. Il segnale locale suggerisce al sistema visivo che la parte di sotto va riempita con una cosa più chiara e la parte di sopra va riempita con una più scura, motivo per cui osserviamo il confetto più scuro sopra ed il confetto più chiaro sotto. Quella zona grigia lì in mezzo non può più essere il colore della superficie ma sarà un'ombreggiatura dovuta dall'interazione delle superfici con l'illuminazione.

Quindi, questo aspetto “interpretativo” (di alto livello) fa sì che il confetto di sotto appaia ancora più bianco e ciò accade perché, pur essendoci una luminanza bassa, questa non viene attribuita al fatto che la superficie è scura di suo ma viene attribuita al fatto che è una superficie bianca però poco illuminata (grazie alla presenza di quella semplice linea orizzontale grigia che interagisce con la luce proveniente dall'alto, ci risulta che il confetto di sopra faccia ombra sul confetto di sotto).

**Noi non notiamo tutte le ombre, ma sono proprio loro ad avere un effetto molto potente su come si percepisce la tridimensionalità nella realtà.**

## Lezione 8

Uno psicologo della percezione cerca di analizzare i fenomeni e riflettere su di essi dal punto di vista teorico. **Cosa succede al prodotto dell'elaborazione della retina quando viene inviato alle stazioni successive? Qual è il significato della proiezione della retina dal genicolato a V1?**



In quest'organizzazione gli assoni delle cellule gangliari raggiungono una particolare struttura, **chiasma ottico**, e la prima parte della produzione si chiama **nervo ottico**. Dopo il chiasma, cambia nome e diviene **tratto ottico** e raggiunge il **genicolato**.

Dal genicolato si diparte un'enorme fascio, che è la **radiazione ottica**; questa è un po' come un'autostrada di informazioni che arriva a V1.

V1 viene chiamata corteccia striata perché si vede proprio una striscia bianca (chiamata *striscia di Gennari*, per lo studente che la scoprì).

Il solco, in basso che divide i due emisferi, si chiama *scissura calcarina* (il nome derivava da qualcuno che ci vide il piede di una gallina).

V1 in realtà gira intorno a questa scissura.

La proiezione **retino-genicolata-striata** è quindi molto ben visualizzabile osservando il cervello. Nell'immagine sono state asportate alcune parti (es. il cervelletto) in modo da metterla particolarmente in luce.

La radiazione ottica è quasi come un'ala di un uccello, è molto molto preminente.

Nell'immagine è quasi come si stesse guardando con una risonanza magnetica: in grigio è il cervello, sovrapposto in giallo (tecnica della trattigrafia) sono messe in evidenza le fibre della radiazione, in rosso c'è il genicolato (da cui si diparte la radiazione).

Quando si parla del genicolato e della natura delle sue proiezioni, è importante comprendere che arrivano delle informazioni che sono lateralizzate, non in base all'occhio di provenienza ma in base all'emicampo di provenienza; ciò è dovuto al fatto che al livello del chiasma il campo visivo è diviso in campo visivo di sinistra e di destra.

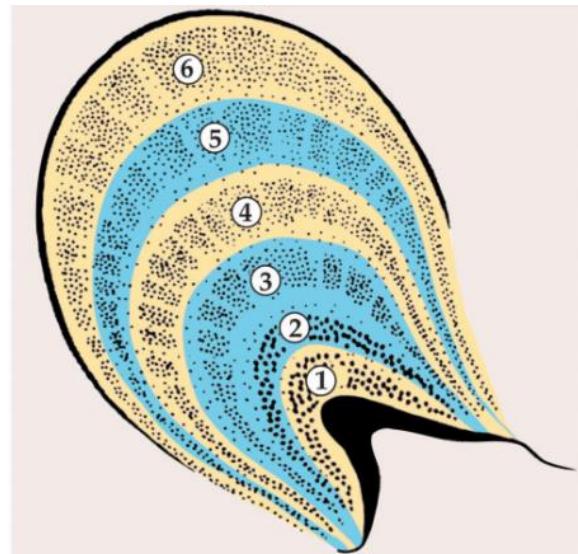
Per la natura della proiezione della luce sui due occhi, *l'informazione che arriva sull'emicampo di sinistra va a finire sulla retina nasale di sx e sulla retina temporale di dx.*

Dall'informazione da queste due emiretine, la retina nasale di sinistra, nel nervo ottico arriva al chiasma e decussa, finendo nel genicolato di destra, mentre la retina temporale di destra rimane omolaterale. La conseguenza di questo è che nel genicolato di destra c'è una rappresentazione dell'emicampo visivo di sinistra, ovvero quello controlaterale, e viceversa per l'emicampo visivo di destra.

**N.B.** Questa cosa si trascina anche nella corteccia visiva primaria, in quanto le informazioni che hanno a che fare con l'occhio di provenienza sono preservate al livello del genicolato ed anche al livello dei primi strati della corteccia visiva primaria. (genicolato = perché è a forma di ginocchio)

Nel **genicolato laterale** si trovano sei strati evidentissimi, che mantengono segregata l'informazione dei due occhi, perché in esso c'è una proiezione dell'emicampo e non più dell'occhio.

Negli strati si alternano occhio destro/sinistro e si alternano anche fibre provenienti dalla retina (che hanno a che fare con le cellule magno/parvo cellulari) e che sono i primi quattro strati (partendo dall'alto) di **parvo cellulari**, gli ultimi due **magno cellulari** (quindi provengono dalla periferia della retina).



Il genicolato di destra mappa l'emicampo di sinistra, ma tenendo separati i due occhi.

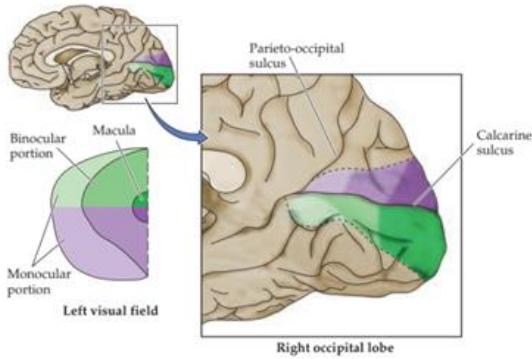
La radiazione che esce dal genicolato preserva ciò: la radiazione di dx, che contiene informazioni di sx, va a finire su V1 di dx; la radiazione ottica si può suddividere in una parte superiore e inferiore.

All'interno di ogni genicolato c'è un'ulteriore distinzione fra la parte superiore e inferiore del campo visivo. Il *campo visivo inferiore* proietta sulla *parte superiore della retina*, **questo poi viene preservato nella mappatura al genicolato, e poi fino a V1**, perché il campo visivo inferiore viene rappresentato dalla parte superiore della radiazione ottica e finisce nella parte superiore di V1 (quella che si trova al di sopra della scissura calcarina). Viceversa per il campo visivo superiore.

Questo a livello globale, è ciò che si intende per **mappa retinotopica**: non vuol dire che in V1 c'è una fotografia della retina e che sulla retina vi è una fotografia del mondo esterno, ma vuol dire che vengono preservate delle *relazioni spaziali*. In V1 è ancora presente l'informazione sull'occhio di provenienza, sul quadrante, e a livello di dettaglio è possibile definire una posizione specifica sulla retina e comprendere dove andrà a mappare in V1.

Le relazioni spaziali vengono preservate, ciò non vuol dire che siano preservate in maniera metrica, in quanto quello che viene preservato è la *topologia*, ovvero l'ordine e le proprietà globali (es. essere a dx rispetto a qualcos'altro). La natura della mappa retinotopica in corteccia è distorta, rispetto a quella originale della retina. Questo è ciò che viene chiamato **magnificazione corticale**.

### visual field maps onto V1



A sinistra è rappresentato il **campo visivo**, dove c'è una *porzione monoculare* (superiore e inferiore), la *porzione binoculare*; poi c'è la rappresentazione della zona centrale intorno alla fovea, la **macula**, (zona in cui vi è più fotopigmento).

La rappresentazione di ciò in corteccia è molto distorta (la parte più scura è quella che corrisponde alla rappresentazione della macula; la parte intermedia

corrisponde alla porzione monoculare del campo visivo e poi quella binoculare). In proporzione, rispetto allo spazio occupato sul campo visivo, sulla corteccia la parte centrale è sovrarappresentata, questo obbedisce al fatto che sulla retina vi sono molti più fotorecettori impacchettati nella parte centrale (e determina che vi debbano essere più unità di elaborazione a livello successivo). Dal punto di vista funzionale vuol dire che ciò che si vede alla fine è quello che la parte centrale della retina acquisisce.

Il cervello dedica moltissime risorse ad elaborare ciò che viene dalla parte centrale della retina ed in proporzione molte di meno per il resto del campo visivo.

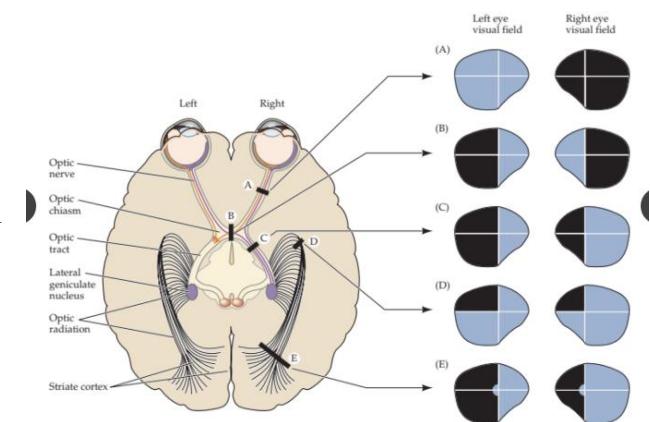
Il risultato finale è che sembra tutto uguale.

## **Cosa succede quando una persona subisce un danno neurologico ai diversi livelli della via retinogenicolato-striata?**

Esistono diversi tipi di cecità, non basta dire “non vede”, in quanto la cecità è differente in base al problema da cui nasce. Anche i ciechi possono essere deafferentati ad esempio.

Esistono molte forme di cecità, es. cecità corticali (parziali o totali) a causa di un danno neurologico di questa via. È interessante chiedersi cosa può succedere.

Nell'immagine sono visibili le possibili zone in cui si verifica una lesione.



La cosa più semplice è quella che succede in A o E, perché: **in A vi è il taglio del nervo ottico**, l'effetto è l'equivalente di perdere un occhio, in quanto viene interrotta la comunicazione tra l'occhio e la stazione successiva (viene interrotta prima del chiasma). Questa non è una condizione particolarmente grave in quanto l'occhio sinistro prende molto del campo visivo destro, si perde praticamente soltanto la parte monoculare dell'occhio destro, e tutti i processi binoculari (ma non sono così fondamentali per la sopravvivenza, anche per la percezione dello spazio non è particolarmente essenziale, si riesce ugualmente a fare quasi tutto); **in E viene interrotto integralmente una delle due radiazioni ottiche**, es. emisfero di dx, questo comporta che non si abbia più nessuna rappresentazione a livello corticale di tutto l'emicampo di sinistra, questo è più difficile da immaginare su cosa voglia dire (**emianopsia**), ci si può abituare a questa condizione, ma si ha comunque un campo visivo molto più ridotto.

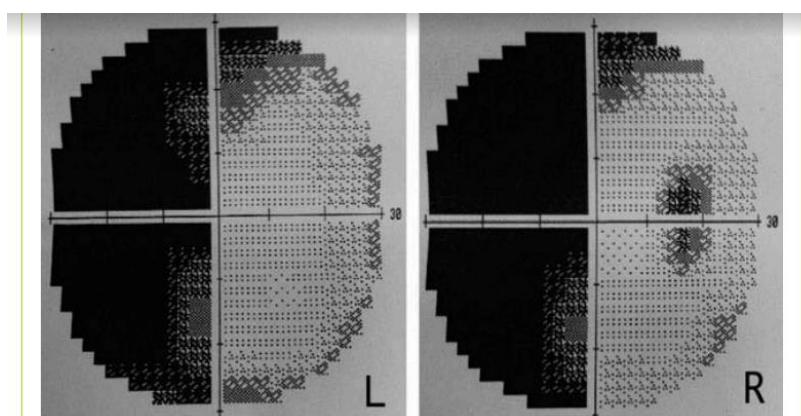
[Un altro tema affascinante (che verrà approfondito nelle lezioni di cognizione) è il **neglect** che non riesce a prestare attenzione a ciò che succede nell'emicampo di sx. Molto spesso il paziente neglect ha un danno nell'emisfero di dx, non è per niente semplice distinguere tra emianopsia e neglect. L'unica maniera di saperlo con certezza assoluta è quella di vedere se passa qualcosa attraverso la radiazione ottica oppure no (c'è molta letteratura che si occupa del problema tra diagnosi differenziale e neglect).]

Tornando all'immagine, si può notare che vi sono una serie di situazioni intermedie di danni visivi: ad esempio **C è un danno al livello del tratto ottico** (post chiasmatico) che è molto diverso da A, in quanto **si comporta come una lesione nella radiazione**, ovvero fa perdere un emicampo. La differenza tra C ed E è che, in E la parte centrale del campo visivo resta preservata, mentre in C no. È molto controverso se in V1 la zona della macula sia rappresentata in entrambi gli emisferi o no, secondo alcuni sì, ciò avrebbe a che fare con come viene costruita la mappa retinotopica, ovvero se nell'emisfero di destra che sarebbe tutto nel caso di sinistra ma la macula sia di destra che di sinistra, con l'emianopsia quindi la macula dovrebbe restare un po' preservata.

**In D viene immaginata una lesione alla radiazione**, ma non a tutta come in E, ma **solo alla parte superiore**, ciò conduce una condizione che si chiama **quadrantanopsia** ovvero **l'assenza di visione in un quadrante dell'emicampo visivo**.

**N.B.** il campo visivo può essere diviso in quattro quadranti, che sono divisi da un sistema di riferimento retinotopico che passa per la fovea e quindi vi è un meridiano orizzontale e verticale che passano per la fovea definendo i quattro quadranti.

**In F vi è una lesione al livello del chiasma**, in cui vanno perse le informazioni che decussano, mentre vengono preservate le informazioni che rimangono nello stesso emisfero e la conseguenza di ciò è che **si perde il campo visivo periferico di entrambi gli occhi**, ovvero risulta una visione come se si vedesse attraverso un tubo, ovvero vi è una riduzione del campo visivo e si vede solo la parte centrale.



**Figure 1.** Visual field testing results, documenting M.B.'s almost complete loss of the left hemifield (L = left eye, R = right eye). Dark areas on the uppermost part of the right hemifields are artifacts due to blepharoptosis. The dark area in the center of the right hemifield of the right eye corresponds to the blind spot. The visual field assessment was performed using Humphrey's automated perimetry and the SITA threshold testing program.

Questo è un esame fatto da un oftalmologo su un paziente studiato in quanto aveva perso la capacità di leggere, non riconosceva le parole leggendo. Ciò di solito è associato ad un danno nell'emisfero di sinistra, **il paziente invece aveva un danno nell'emisfero destro**, quindi **aveva un'emianopsia a sinistra**. Per comprendere la gravità di ciò

l'oftalmologo doveva fare un test del campo visivo, controllando la fissazione dei punti luminosi e misurando la sensibilità.

Nell'immagine più è chiaro e migliore è la sensibilità, più è scuro e peggiore è la sensibilità.

Si vede chiaramente che *i due quadranti di sinistra, sia nell'occhio destro che sinistro sono tutti neri*, vi è una piccola isola di migliore sensibilità avvicinandosi al meridiano verticale.

L'*emicampo preservato* ha comunque *due zone nere, quelle sono il disco ottico* (ovvero la macchia cieca) di destra.

Inoltre nel quadrante superiore, sempre vicino al meridiano verticale, vi è un'altra zona (approssimativamente triangolare) in cui è nero, ovvero non c'è sensibilità, ed è semplicemente l'ombra della palpebra (tipica delle persone anziane).

## Sensibilità al contrasto

Questo è un problema molto studiato a partire dagli anni '60, periodo in cui sono state sviluppate alcune tecniche che convergono verso lo stesso tipo di risposta: una delle principali funzioni che avviene a livello della corteccia visiva primaria -prima che vi sia la fusione binoculare- è rappresentare discontinuità orientate, cioè le zone del campo visivo dove ci sono dei cambiamenti (dove la luce ha discontinuità) però includendo anche una rappresentazione di come questa discontinuità è orientata nel campo visivo. Ciò vuol dire che se si ha un segmento, si ha inevitabilmente un *bordo* (verticale, orizzontale, diagonale) e, non solo, le discontinuità orientate sono a diversi *livelli di dettaglio*:

- se sono a un livello di dettaglio **fine** significa che tali cambiamenti sono molto localizzati (es. se si ha un margine molto netto in cui c'è un passaggio molto brusco da una zona più chiara ad una più scura);
- oppure **un margine sfumato** (come quello di un'ombra) in cui il passaggio dal chiaro allo scuro è molto più graduale e quindi avviene su scala più larga.

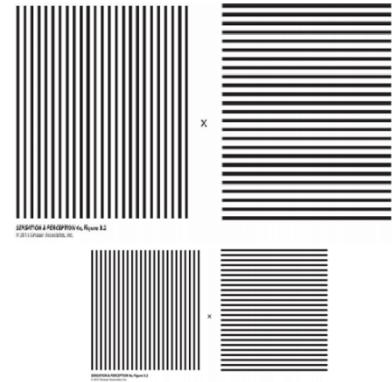
Questa idea generale sostiene che il sistema visivo di basso livello produce una mappa in cui ci sono le discontinuità aumentate a diversi livelli di dettaglio. Ciò emerge sia da tecniche di tipo *psicofisico* che da tecniche di tipo *fisiologico*.

## Psicofisica

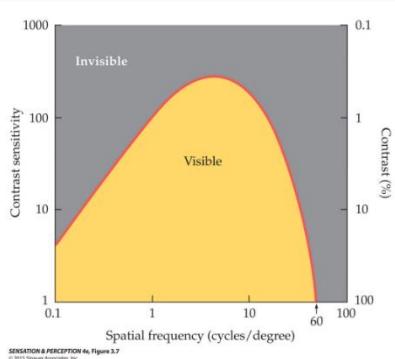
Bisogna comprendere esattamente cosa si intende in psicofisica per “*sensibilità al contrasto*”.

**La sensibilità al contrasto** è la capacità di rilevare una discontinuità (differenza) tra una zona più chiara ed una più scura. Questa è fondamentale perché ha a che fare anche con l’acuità (*capacità di cogliere i dettagli fini*).

**Esempio:** se ci viene mostrata questa figura e ci si allontana, la figura si rimpicciolisce sulla retina e, ad un certo punto, è talmente piccola che non si riesce più a distinguere tra le barre verticali e orizzontali. Ciò che accade è che la differenza tra il nero e il bianco (il contrasto) diventa sotto-soglia, ovvero è più piccola della soglia differenziale e quindi non si riesce a discriminare tra verticale e orizzontale. Si può facilmente immaginare una procedura tipo *Fechner*, in cui si riesce a misurare la soglia differenziale. Questo ha a che fare con la capacità di risolvere i dettagli fini, ci sono dei limiti superiori alla capacità di risolvere i dettagli fini.



La visione umana, però, non riguarda solo la capacità di cogliere i dettagli fini ma anche la misura della *sensibilità al contrasto nella zona delle frequenze spaziali più alte*. Un quadro completo della sensibilità al contrasto dovrebbe, infatti, misurare la soglia usando molti grating diversi (a bassa, intermedia o alta frequenza) e mappando la sensibilità in funzione della frequenza spaziale. Ciò produce la funzione psicofisica per la **sensibilità al contrasto**.



Nel grafico vi è una funzione che descrive come la sensibilità al contrasto varia in funzione della frequenza spaziale.

**Frequenza spaziale** vuol dire livello di dettaglio:

- Frequenze spaziali basse, sono situazioni in cui la sensibilità al contrasto è stata misurata con un grating a bassa frequenza, ovvero in cui vi sono delle barre molto sfumate e molto ampie che occupano una parte molto rilevante del campo visivo;
- Frequenze spaziali alte, vengono misurate con barre molto piccole e definite che occupano una parte contenuta del campo visivo.

**NB**→ La sensibilità al contrasto è l'inverso della soglia: con i grating si misura la soglia assoluta, ciò consiste in un compito in cui si deve distinguere il grating da un campo omogeneo.

La relazione inversa tra *sensibilità* e *soglia* è resa esplicita nel grafico:

- sulla sinistra vi è la sensibilità al contrasto;
- sulla destra vi è un asse diverso in cui viene rappresentato il contrasto della soglia.

La *sensibilità* cresce e raggiunge il suo massimo più o meno intorno a cinque/sei cicli del grating per grado di angolo visivo (nella parte più alta della curva nell'asse di sinistra) che, però, guardando l'asse di destra corrisponderà alla *soglia* più piccola -essendo la sensibilità al contrasto l'inverso della soglia; quindi, andando verso l'alto decresce anziché crescere. Il picco (spike) avviene, quindi, dove c'è il massimo della sensibilità e il minimo di soglia. Ciò significa che in quella zona si è in grado di distinguere un grating anche poco contrastato da un campo grigio omogeneo (soglia differenziale molto bassa).

**NB**→ La frequenza spaziale in realtà è il numero di cicli del grating per grado di angoli visivi (in quanto il campo visivo si estende per 180°). La frequenza spaziale si misura in gradi e riguarda l'ampiezza dello spazio che quello stimolo occupa nel campo visivo. È una misura di densità che indica quanti cicli del grating sono presenti in un grado di angolo visivo.

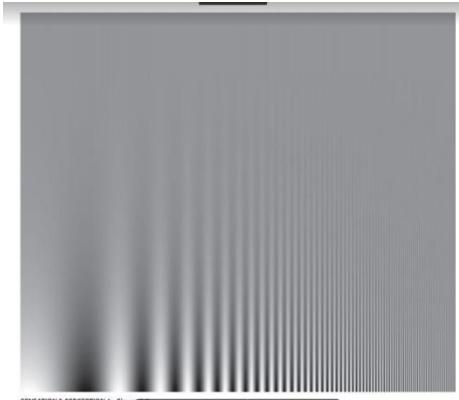
**NB2**→ per “ciclo” si intende “quante volte il grating passa dal nero al bianco”.

**Esempio:** se passa una sola volta dal nero al bianco vuol dire che è un grating in cui non c'è molta azione; se è 1/10 vuol dire che non passa neanche completamente dal nero al bianco ma è come un gradiente sfumato che, lentamente, diventa o più chiaro o più scuro. Invece, 60 cicli in un grado di angolo visivo indicano un grating molto fitto, in cui si alternano 60 volte parti chiare con parti scure.

Questa è una funzione psicofisica per cui in ogni grating viene presa la soglia assoluta e poi viene plottata in funzione della sequenza spaziale: il risultato assomiglia a quel tipo di funzione psicofisica vista con gli audiogrammi sonori. Gli audiogrammi sonori sono dei plot della soglia (sensibilità) in funzione della frequenza del suono, che non è spaziale ma temporale. Visto che vi sono anche ragioni storiche in queste cose, gli audiogrammi sonori sono sempre plottati avendo in mente la soglia e non la sensibilità, e quindi dando una funzione a forma di U.

Per quanto riguarda la *frequenza spaziale*, invece, si ragiona in termini di sensibilità e non di soglia e quindi il punto più alto è quello in cui la soglia è più bassa e la sensibilità più alta, a forma di U rovesciato.

Al di là di questo, in realtà, queste due tipologie di funzioni psicofisiche si somigliano moltissimo perché sono entrambe delle mappature della sensibilità in funzione della frequenza che nel caso del suono è temporale, nel caso della vista è spaziale.



Questa figura è stata creata disegnando dei reticolati in cui, invece di avere una frequenza spaziale costante rispetto allo spazio, si parte da una frequenza spaziale bassa che aumenta gradualmente andando verso destra. Va pensata come un'immagine bidimensionale in cui la frequenza spaziale, da sinistra verso destra, aumenta in maniera costante. Questa *variazione della frequenza spaziale* è la stessa a qualsiasi altezza ci si ponga, perché ciò che viene cambiato andando dal basso verso l'alto è il *contrasto*.

Quindi:

- Andando da sinistra verso destra ci sono dei reticolati la cui frequenza spaziale aumenta;
- Andando dal basso verso alto i reticolati sono a contrasto sempre più basso.

Guardando la parte alta della figura è difficile accettare che lì ci sia ancora un certo livello di contrasto ed una variazione della frequenza spaziale, in quanto appare omogeneo; ciò avviene perché il *contrasto fisico* in quel punto è *al di sotto della soglia* e, quindi, non riusciamo a vederlo: indipendentemente dalla frequenza spaziale, si è al di fuori della sensibilità visiva al contrasto.

Spostando l'attenzione verso la parte bassa della figura succede che, ad un certo punto, vi è distinzione sempre più netta; in questo caso il risultato osservabile dipende dalla frequenza spaziale:

- a *sinistra* si deve arrivare quasi a metà dell'asse "y" per cominciare a vedere che c'è una differenza tra nero e bianco;
- al *centro* ciò avviene in un punto dell'asse "y" molto più alto;
- a *destra* avviene in un punto più in basso dell'asse "y".

La forma esatta di ciò dipende da quanto si è distanti dallo schermo, in quanto anche questo incide sulla frequenza spaziale.

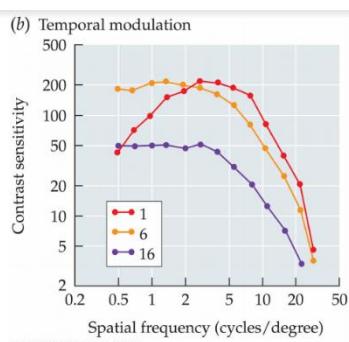
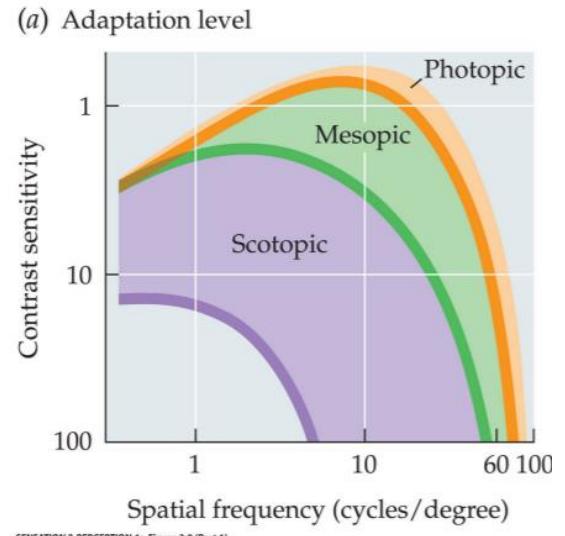
Grossomodo, la zona in cui c'è la transizione da grigio-reticolo ha più o meno la forma della funzione della sensibilità al contrasto, perché è proprio ciò che viene rappresentato: il punto in cui vi è la soglia dal vedere grigio a vedere qualcosa di contrastato. Per tale motivo è utile la rappresentazione della capacità di rilevare discontinuità orientate, perché la percezione del contrasto ha a che fare con la capacità di rilevare discontinuità e questo lo si può fare non solo in funzione della frequenza spaziale, ma anche in funzione *dell'orientazione del grating*.

Si possono studiare le CFS (contrast sensitivity function) a diversi livelli di orientazione avendo, quindi, un quadro completo della capacità di rilevare discontinuità orientate a diverso livello di dettaglio. Ciò consente, poi, di fare dei confronti andando a vedere come il quadro generale della sensibilità varia in funzione di altri fattori.

**Esempio:** come cambia la CSF a seconda del livello di adattamento della retina: visione notturna/scotopica, diurna/fotopica o intermedia/mesopica.

La visione fotopica, in generale, ha un migliore livello di dettaglio (rileva dettagli più fini) rispetto a quella scotopica, perché si estende nelle frequenze spaziali più alte. Anche la sensibilità al contrasto è migliore in condizioni fotopiche rispetto a quelle scotopiche.

La sensibilità al contrasto dipende anche dalla frequenza temporale; infatti, nel caso in cui si faccia accendere e spegnere il grating la forma della CSF cambierebbe anche in funzione della frequenza temporale.



In funzione dell'età queste modalità cambiano: invecchiando non solo si diventa presbi, ma si osserva anche un cambiamento nella forma della CSF. È un cambiamento non drammatico ma consiste comunque in una perdita di sensibilità al contrasto nella capacità di rilevare queste cose, a tutti i livelli e non soltanto per i dettagli più fini. Confrontando la *forma* della dimensione di sensibilità al contrasto ci si può fare un'idea dell'esperienza visiva che hanno i diversi organismi; per esempio, l'uomo -che può descriverla a parole- rispetto ad altri organismi.

Questa è una rappresentazione della *visione di sensibilità al contrasto di un adulto*: la forma non è esattamente identica alla precedente perché questa rappresenta dati reali su un adulto e su un bambino a diverse età.

Quando si nasce la retina non è ancora del tutto sviluppata. Non è strano, quindi, che un bambino dopo un mese abbia una capacità di percepire il contrasto che è quasi esclusivamente concentrata nella parte più bassa delle frequenze spaziali, ossia è capace di rilevare il contrasto solo ad un livello di dettaglio molto grossolano.

Questo certe volte viene anche visualizzato immaginando che il bambino non veda in modo ottimale ma soltanto in modo sfumato (come se vedesse l'immagine senza tutti i dettagli fini). Si può immaginare un proiettore in cui l'immagine viene proiettata molto sfocata; non è, però, solo una sfocatura ma è una perdita di contrasto apparente (il bianco non è più tanto bianco e il nero non è più tanto nero). Con lo sviluppo si vede chiaramente che la curva sale e, quindi, la sensibilità migliora e gradualmente si estende anche nella parte dei dettagli più fini a livello delle frequenze spaziali più alte.

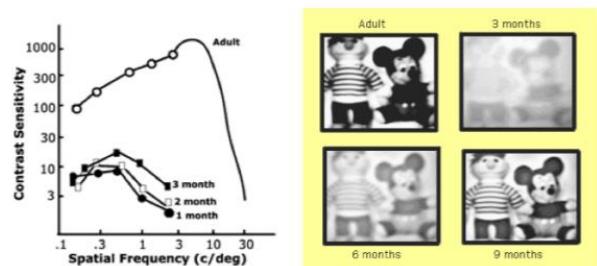
Fino agli anni '60 si credeva che i neonati fossero ciechi. Questi dati dimostrano, appunto, che non è così in quanto possono rilevare una sensibilità al contrasto ma è poco sviluppata rispetto a quella dell'adulto.

Un altro fattore presente, sia nel bambino che nell'adulto, è un limite nella capacità di rilevare la rappresentazione delle forme. La discontinuità orientata rappresenta i segmenti, le forme e come questi sono orientati nello spazio. Questo è uno dei vincoli fondamentali alla capacità dell'uomo di segmentare il mondo in figure dotate di forma. L'uomo è molto sensibile a fare questo tipo di cose, in un range intermedio di frequenze spaziali né altissimo né bassissimo.

Le cose cambiano se ci si confronta con altre specie animali per cui si può misurare la sensibilità al contrasto nonostante la mancanza di linguaggio. Ad esempio, per i rapaci la sensibilità al contrasto si estende di più nella zona delle frequenze spaziali alte (dettagli molto fini); per altri animali, invece, si estende per le frequenze spaziali più basse.

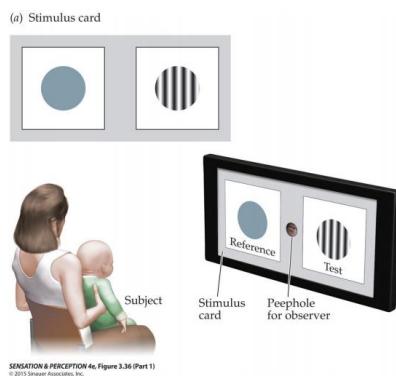
La capacità, delle diverse specie, di costruire delle unità percettive dotate di un significato biologico -che garantisce l'interazione nel loro ecosistema- è un po' diversa da quella dell'uomo.

## CSF e sviluppo



## *Come si può misurare la funzione di sensibilità al contrasto in un bambino o in un ratto?*

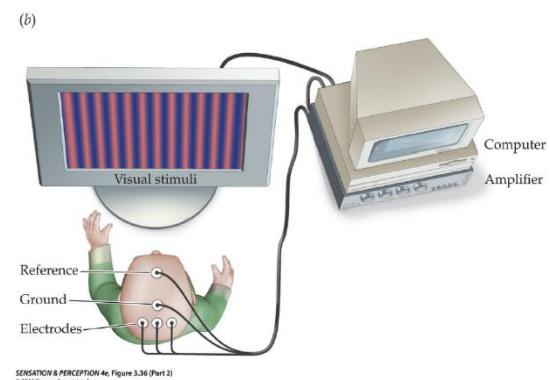
Nel caso dei bambini si può fare un **test di preferenza nella fissazione**: è basato sulla semplice idea (vera anche negli animali) che *una cosa interessante (dove c'è un cambiamento), produce un comportamento di fissazione*. Una cosa meno interessante (dove non c'è cambiamento) non viene fissata dal bambino. In questo modo si può misurare la capacità di discriminare tra un campo omogeneo e un grating: se il bambino discrimina ed è interessato, dovrebbe guardare di più dove c'è il cambiamento; se non vede nessuna differenza non mostrerà interesse e non ci sarà fissazione.



Durante questi test viene presentato sullo schermo uno stimolo di riferimento (campo grigio omogeneo) e un grating nel quale varia il contrasto. Ad ogni presentazione si misura quanto tempo il bambino guarda a destra piuttosto che a sinistra, ossia quali dei due stimoli presentati preferisce. Può essere fatto con un'assistente o con una videocamera che regista gli occhi del bambino<sup>4</sup>. Attraverso questo principio si ha una statistica di quanto tempo il bambino ha speso guardando a destra o a sinistra per ogni livello di contrasto; se ha guardato più verso il grating significa che il bambino è in grado di discriminare i due stimoli visivi.

Con un animale adulto non umano si possono usare tecniche in cui gli viene insegnato a discriminare tra stimoli attraverso il rinforzo con un premio; ad esempio, se l'uccello becca il grating riceve un premio - perché significa che non va a caso, bensì sta mettendo in atto una discriminazione.

Possono essere utilizzate anche delle tecniche elettroencefalografiche, in cui si presentando un grating che si muove nel campo visivo del bambino si osserva la differenza di potenziale nell'EEG. Questo viene chiamato SWEEP VEP.



<sup>4</sup> Eye tracking

## **Lezione 9**

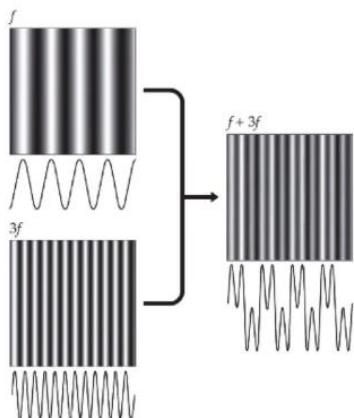
### **Codifica delle discontinuità orientate**

Il **sistema visivo di basso livello** codifica le *discontinuità orientate cioè quello che*, a livello percettivo, diventa la base per percepire i contorni degli oggetti.

Cosa centra la sensibilità al contrasto con la percezione delle discontinuità orientate? Centra perché la funzione di sensibilità al contrasto descrive la capacità di rilevare un contrasto, ossia una discontinuità nella distribuzione della luce. Ad esempio, prendendo delle semplici barre variamo:

- l'orientazione nello spazio o il contrasto;
- il livello di scala in cui questo reticolo presenta questa discontinuità, ossia la **frequenza spaziale**.

### **somme di reticolli**



La *sensibilità al contrasto* è interessante sia dal punto di vista clinico che dal punto di vista comparativo (es: confrontare la capacità di percepire le discontinuità orientate in diversi tipi di organismi).

Ma c'è un motivo ancora più profondo per cui si studia la psicofisica.

A sinistra ci sono due reticolli:

- quello in alto ha una frequenza spaziale più bassa, perché nella stessa unità di spazio c'è un numero minore di **cicli** di alternanza bianco/nero, quindi il livello di dettaglio è minore

- quello in basso ha una frequenza spaziale più alta, perché presenta un numero superiore di cicli e ciò conferisce al grating un livello di dettaglio più alto rispetto al reticolo in alto.

La differenza di frequenza spaziale si vede anche dall'**onda** che descrive il *profilo di luminanza*: l'onda di sopra ha una frequenza più blanda nello spazio rispetto a quella di sotto che ha una frequenza più fitta. Inoltre, anche il simbolo “f” (che sta per “frequenza”) è indicativo: il graticolo sotto è denominato “3f” perché ha una frequenza 3 volte maggiore rispetto al graticolo di sopra.

Questi graticoli sono una semplice funzione matematica “seno”, la quale descrive il loro profilo di luminanza: y=sen(x) è una funzione che descrive l’onda perfettamente perché la luminanza varia solo in una direzione, ortogonalmente rispetto alla direzione del reticolo.

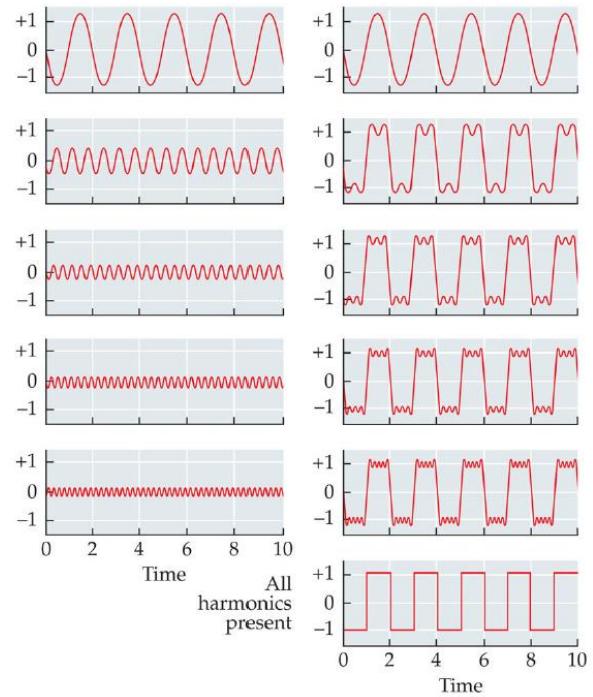
Si può pensare alle due funzioni come a dei vettori in cui per ogni pixel dell’immagine c’è un numero che corrisponde all’intensità dello schermo in quel punto. Questo significa che noi possiamo *sommare* questi due vettori e, ciò che ne viene fuori è un’altra funzione più complicata (quella a destra) e, quindi, un altro reticolo che rappresenta la somma dei due reticolari ( $f+3f$ ).

I reticolari possono essere usati per costruire immagini più complesse, semplicemente prendendone due e sommandoli.<sup>5</sup> Questa viene chiamata la “**Sintesi di Fourier**” un fisico che ha studiato matematicamente cosa succede quando si prendono dei reticolari e si sommano, ottenendo un reticolo più complesso.

Poi se prendiamo la seconda funzione ( $3f$ ) più complicata e la sommiamo la terza armonica [ $3f + (f+3f)$ ], succede che la forma d’onda cambia ancora e così via fino ad ottenere un’onda quadra.

Questa regola vale anche per i suoni. Quando noi produciamo una nota, questa ha una *frequenza di oscillazione fondamentale* e poi ha delle frequenze di oscillazione che sono più elevate (con un’ampiezza più piccola) e sono, rispetto alla fondamentale, in certi rapporti interi. L’ottava è la prima armonica di un suono che corrisponde esattamente al doppio della frequenza fondamentale. Queste frequenze che si differenziano dalla fondamentale sono quelle che poi vanno a determinare un certo tipo di timbro (di voce, ad esempio).

La stessa cosa vale per un reticolo. Se prendiamo un reticolo che ha una singola frequenza spaziale e lo sommiamo ad altri reticolari otteniamo reticolari più complessi -così come la nostra voce è costituita da un’onda più complessa che è la somma di più sinusoidi.



*SENSATION & PERCEPTION 4e, Figure 3.10*

© 2015 Sinauer Associates, Inc.

<sup>5</sup> con “sommare” si intende prendere ogni pixel e sommare le sue intensità con le intensità dei pixel corrispondenti all’altra immagine.

Ciò che è più interessante è che Fourier ha trovato dei metodi matematici per fare anche l'opposto della sintesi: è possibile prendere un segnale complesso, come per esempio un suono che corrisponde ad un'onda quadra o ad una sua approssimazione (ultima onda del grafico in basso a destra della prima pagina), e sottoporlo ad un algoritmo che si chiama trasformata di Fourier il quale scomponete il segnale

*J. Physiol.* (1968), **197**, pp. 551-566  
With 7 text-figures  
Printed in Great Britain

551

**APPLICATION OF FOURIER ANALYSIS TO THE VISIBILITY OF GRATINGS**

By F. W. CAMPBELL AND J. G. ROBSON  
*From the Physiological Laboratory, University of Cambridge*

(Received 10 November 1967)

**SUMMARY**

1. The contrast thresholds of a variety of grating patterns have been measured over a wide range of spatial frequencies.
2. Contrast thresholds for the detection of gratings whose luminance profiles are sine, square, rectangular or saw-tooth waves can be simply related using Fourier theory.
3. Over a wide range of spatial frequencies the contrast threshold of a grating is determined only by the amplitude of the fundamental Fourier component of its wave form.
4. Gratings of complex wave form cannot be distinguished from sine-wave gratings until their contrast has been raised to a level at which the higher harmonic components reach their independent threshold.
5. These findings can be explained by the existence within the nervous system of linearly operating independent mechanisms selectively sensitive to limited ranges of spatial frequencies.

complesso e fa in modo che questo venga espresso come somma di tante funzioni semplici. Il risultato ottenuto dopo la trasformata di Fourier è uno *spettro*, cioè una serie di componenti armoniche che sono implicite in un segnale più complesso.

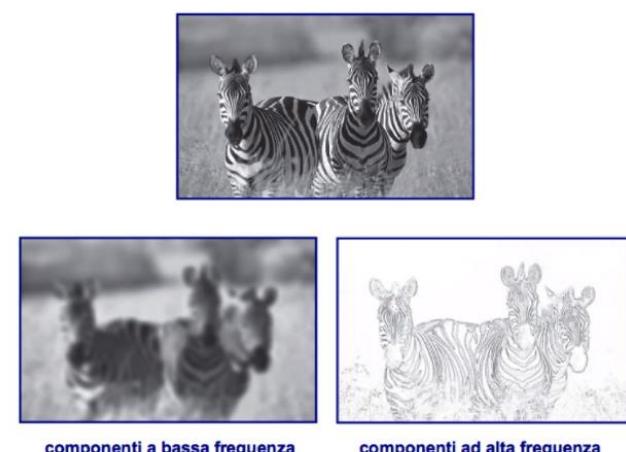
Ora si può fare un passo successivo per capire cosa c'entra tutto questo nella *psicofisica della visione*.

A sinistra viene riportato il lavoro di Campbell e Robson, dell'Università di Cambridge, pubblicato nel 1968 sul "Journal of Physiology": "Applicazione dell'analisi di Fourier alla visibilità dei gratings". I due studiosi, usando metodi psicofisici e misurando funzioni di sensibilità al contrasto, hanno testato se è vero che ogni segnale complesso può essere espresso come somma di tanti segnali più semplici (funzioni seno); in particolare hanno testato se l'analisi di Fourier vale per ogni tipo di segnale complesso, incluse tutte le immagini che vediamo.

Ad esempio, nell'immagine a destra, quella in alto può essere pensata come la somma di tante immagini più semplici in bianco e nero (che corrispondono alle sue componenti armoniche).

Ma per le immagini (come le foto) non funziona semplicemente così. La differenza tra un grating e questa immagine consiste nel fatto che il grating è una funzione unidimensionale, cioè l'intensità varia lungo un'unica dimensione -che è quella ortogonale alle barre.

## immagini naturali



componenti a bassa frequenza

componenti ad alta frequenza

Nell'immagine con le zebre, invece, la variazione dell'intensità luminosa non è in un'unica dimensione, bensì in due; bisogna, quindi, trovare il modo di generalizzare l'idea di Fourier da 1 dimensione a 2 dimensioni.

Quindi, l'analogia tra una somma di gratings e la somma delle componenti (ad alta e bassa frequenza) delle immagini in 2D è concreta; si può proprio pensare che in queste immagini ci siano:

1. tante discontinuità nell'intensità luminosa;
2. che queste discontinuità sono orientate (nel senso che alcune sono verticali, altre orizzontali e diagonali);
3. che sono a diversi livelli di dettaglio;
4. e che alcune sono più localizzate mentre altre sono di distribuite in una zona.

L'immagine (come quella delle zebre) è fatta da tante discontinuità, quindi è possibile scomporla in una serie di immagini che catturano singolarmente ogni aspetto di queste discontinuità in 2D -che è quello che corrisponde alla scala del dettaglio (più o meno dettaglio), ossia al continuum che riguarda la frequenza spaziale. Le componenti a bassa frequenza delle tre zebre sono:

-le *componenti a bassa frequenza* definiscono le **zone sfuocate**, ossia quei punti in cui il cambiamento di luminanza è a livello di dettaglio più ampio e globale;

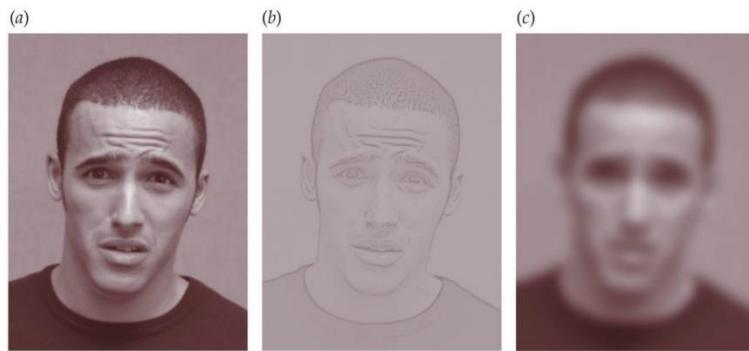
-le *componenti ad alta frequenza* definiscono i **dettagli fini**, ossia quei punti in cui ci sono dei cambiamenti abbastanza bruschi nell'intensità di luce.

Partendo da questa idea Campbell e Robson hanno studiato le immagini in 2D semplificandole in un'unica dimensione misurando la sensibilità ai graticoli.

**NB:** alla base deve essere chiara l'idea che qualsiasi immagine complessa può essere pensata come una somma di tante immagini più semplici in cui le discontinuità vengono rappresentate a diversi livelli di dettaglio (come una serie di trasparenti messi uno sopra l'altro in cui sono presenti sempre più o meno dettagli).

Questo ovviamente vale anche per un volto, non solo per le zebre.

Es: Nell'immagine "a" possiamo vedere un volto che esprime sorpresa/disgusto: l'immagine "b" è la foto originaria fatta passare attraverso un filtro chiamato "Passa alto" che lascia passare solo le frequenze spaziali più alte, perdendo però quelle frequenze più basse che invece sono filtrate nell'immagine "c" con il filtro Passa basso.

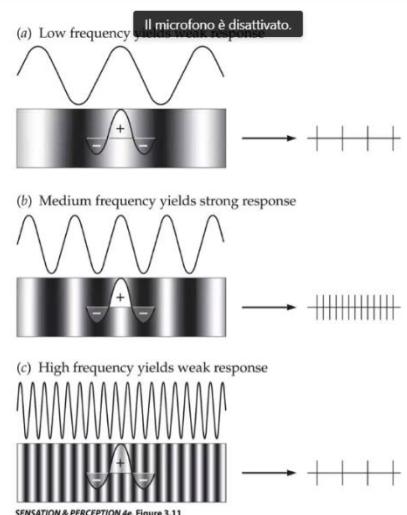


SENSATION & PERCEPTION 4e, Figure 3.34  
© 2015 Sinauer Associates, Inc.

Questo discorso torna molto bene con quello che riguarda i **campi recettivi**, le **cellule gangliari** ed il **Nucleo Genicolato Laterale**. Questi campi recettivi sono fatti da un centro che è una specie di disco e una periferia che è una specie di anello (es: cellule centro-on).

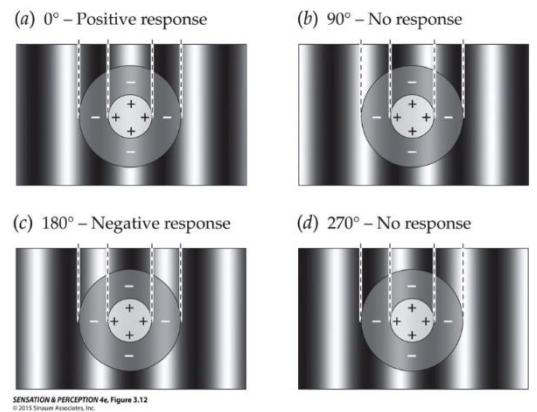
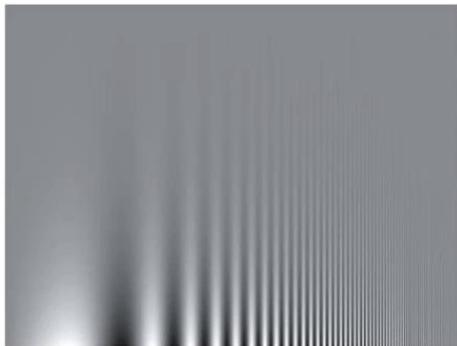
Questo può essere rappresentato rispetto allo spazio, per esempio, se il centro è ON da una porzione di reticolo su cui va a cadere il campo recettivo della cellula gangliare, il che vuol dire che al centro ci sarà una risposta eccitatoria e in periferia una inibitoria. Quindi può essere pensato come un qualcosa che a seconda di dove va a finire sul reticolo (e anche a seconda di quanto è ampio questo campo recettivo) produce una risposta più o meno ottimale : quando il reticolo corrisponde proprio esattamente dal punto di vista spaziale al rapporto tra l'area occupata dal disco e quella occupata dall'anello, lì ci sarà la risposta ottimale di quella cellula e questo produce l'equivalente di un filtro che codifica l'informazione a quella frequenza lì e non ad altre.

Nella seconda immagine a destra vediamo che c'è la risposta ottimale perché il bianco va a finire tutto nella zona eccitatoria e il nero in quella inibitoria, quindi ho il massimo di eccitazione e il minimo di inibizione. Con frequenze più alte o più basse (prima e terza immagine a destra) invece non succede questo, nei campi eccitatori e inibitori ci sono zone eccitatorie e inibitorie mescolate e quindi la risposta non è ottimizzata.



SENSATION & PERCEPTION 4e, Figure 3.11  
© 2015 Sinauer Associates, Inc.

Nell'immagine successiva abbiamo la stessa cosa ma, cambiando non la frequenza spaziale del reticolo, ma la sua posizione nello spazio che matematicamente si esprime come **la fase della funzione seno**. La fase si esprime con un angolo perché dice se cominciamo a calcolare l'intensità dal seno di 0, di 90°, di 180° o 270°. È l'equivalente quindi solo di shiftare il punto di partenza: se partiamo da un punto in cui la sinusode ha un valore piccolo, intermedio o alto. Questo produce lo stesso effetto, se ho una cellula gangliare Centro ON e la nostra sinusode è shiftata in modo che il bianco va a cadere sulla zona eccitatoria e il nero sulla inibitoria , lì avrò una risposta positiva ma , se shiftiamo verso destra come in C, SF adaptation demo: test display



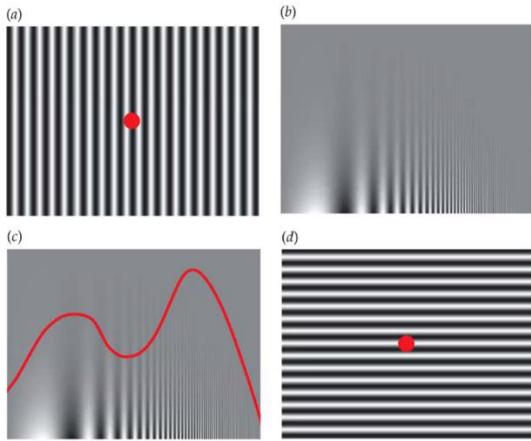
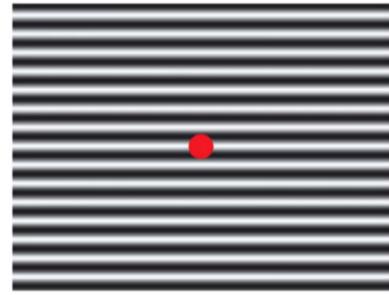
il campo recettivo rimane lì e il centro on va sul nero quindi non c'è eccitazione e il bianco va nella zona inibitoria, quindi avrò il massimo di inibizione. Questo dimostra che le cellule gangliari non codificano solo il contrasto ma sono anche dei **filtri** per le frequenze spaziali.

→ Esperimento in classe per testare l'idea di Campbell e Robson che il sistema visivo di basso livello può essere descritto dal punto di vista funzionale come una serie di filtri che sono selettivamente sintonizzati sulle discontinuità orientate ad un certo livello di dettaglio

(c'è un filtro che becca le discontinuità sulle basse frequenze, uno sulle intermedie e uno sulle alte): a sinistra abbiamo un **paradigma di adattamento**. Adattamento vuol dire che noi misuriamo la sensibilità per qualcosa e poi creiamo delle condizioni per cui il sistema visivo si adatta, cioè riduce progressivamente la propria sensibilità (es: ho stimolazione continua come un odore e dopo un po' non lo sento più). Fenomeno per cui anche senza perdere completamente la percezione, la presentazione ripetuta di un qualcosa produce selettivamente una riduzione di sensibilità per quel qualcosa.

L'immagine in basso alla pagina precedente rappresenta la funzione di sensibilità al contrasto, in questa figura si vede la soglia, cioè il confine tra il grigio e il grating e questo confine ha la caratteristica forma ad U rovesciato asimmetrico.

A sinistra abbiamo uno stimolo di adattamento, nel senso che è un reticolo e per adattarmi a questo reticolo dobbiamo fissare il pallino rosso e contare fino a 60. Poi subito dopo si mostra l'immagine di prima, cosa è cambiato? Una prima ipotesi è che si produce un'immagine consecutiva, cioè si produce un'immagine che è dovuta al fatto che sulla retina c'è stato un processo di adattamento globale e quindi una perdita di sensibilità in funzione di quello che ho visto prima, quindi è come se proiettassi un'immagine su qualcos'altro.



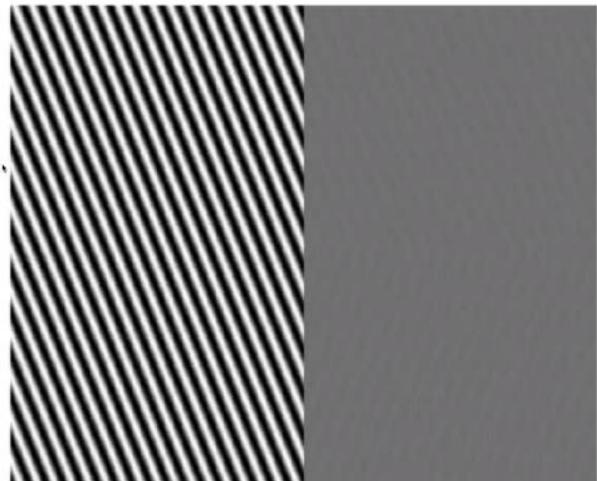
Se invece ora ci adattiamo allo stimolo *a* e poi vediamo l'immagine *b*, può succedere che dopo l'adattamento ci dovrebbe essere una perdita di sensibilità del sistema, però questa perdita dovrebbe essere selettiva per la frequenza spaziale del reticolo in questione. Si può dimostrare che è selettiva per la frequenza spaziale del reticolo in questione e questa è un'evidenza del fatto che c'è un meccanismo specifico responsabile di questo range di frequenze spaziali, quindi noi *quel* range abbiamo adattato nello specifico e non gli altri sensibili ad altre frequenze spaziali.

Rifacendo l'esercizio dunque si dovrebbe vedere ciò che è illustrato nell'immagine C a sinistra, cioè la curva che delimita la zona di passaggio dal reticolo al grigio che però non è una U rovesciata ma una U rovesciata con un inarcamento. Il buco dell'inarcamento è situato in corrispondenza della frequenza del reticolo a cui ci siamo adattati. Questo si verifica solo quando c'è adattamento ad A e poi mostrato C, ma non se mi adatto a D e mostro B, questo perché questo ipotetico filtro non solo è selettivo per la frequenza spaziale (quindi abbiamo stancato quella frequenza e in quel punto è saltato fuori l'adattamento) ma anche selettivo per l'**orientazione**. Quindi non basta stancare la frequenza spaziale, è necessario che sia la frequenza spaziale di un reticolo orientato come l'immagine test, siccome in A ho "stancato" sia la giusta frequenza sia il giusto orientazione allora è venuto fuori questo avallamento, questa riduzione selettiva, mentre in B la frequenza spaziale era quella giusta ma non era la giusta orientazione.

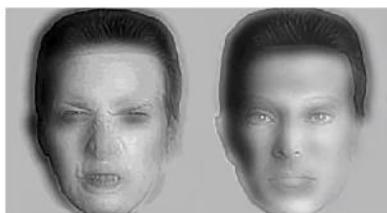
Se in questa immagine fissiamo il primo reticolo nella metà di sinistra e dopo guardiamo la seconda metà a basso contrasto (che contiene due reticolli orientati in senso opposto), riusciamo a vedere solo il reticolo in basso a destra, come se quello in alto sparisse.

Questa è un'altra conseguenza dell'adattamento.

Adattando prima il sistema visivo a un reticolo con la stessa frequenza ma con orientazione solo un po' diversa succede che quando mostro lo stimolo neutro, è come se dall'insieme delle attivazioni di questi filtri togliessi una componente che è quella degli orientamenti equivalenti a quella che ho adattato prima, quindi l'insieme di queste attivazioni si sbilancia dall'altra parte (effetto postumo di spostamento dall'altra parte).



## Dr. Angry & Mr. Smile



Esperimento con i volti: questi due volti identificano un volto arrabbiato (a destra) e uno più rilassato (a sinistra). Se sfochiamo l'immagine stringendo le palpebre (o se sei miope), i due volti si scambiano di posto. Perché succede?

Ognuna di queste due immagini è fatta di due volti: si prendono entrambe e si filtrano con un filtro passa-basso e un filtro passa-alto dove ho 2 immagini di Dr.Angry, una con le alte frequenze ed una con le basse (idem per Mr. Smile). Poi prendiamo queste 4 immagini e le ricombiniamo incrociandole, quindi prendo Mr.Smile a bassa frequenza e lo sommo a Mr.Angry ad alta frequenza e viceversa. Il risultato è che questi volti hanno delle ombre strane date dal processo di somma delle immagini. Quando sfochiamo è come se applicassi un filtro passa-basso e levo le componenti ad alta frequenza (i dettagli fini). Mentre se guardiamo da vicino vediamo le immagini ad alta frequenza e Dr Angry e Mr Smile si invertiranno.

## La gioconda



La Gioconda ha un **sorriso enigmatico**, il carattere enigmatico non ha a che fare con la percezione di basso livello .I filtri di cui abbiamo parlato, conditi con il ciclo percezione-azione ci danno una spiegazione ad un'ipotesi interessante che cerca di codificare il fenomeno di questo carattere enigmatico della Gioconda: dobbiamo pensare il quadro come un'immagine complessa, quindi si può pensare come una somma di tante immagini con diverse frequenze spaziali (alte e basse) e le vediamo rappresentate

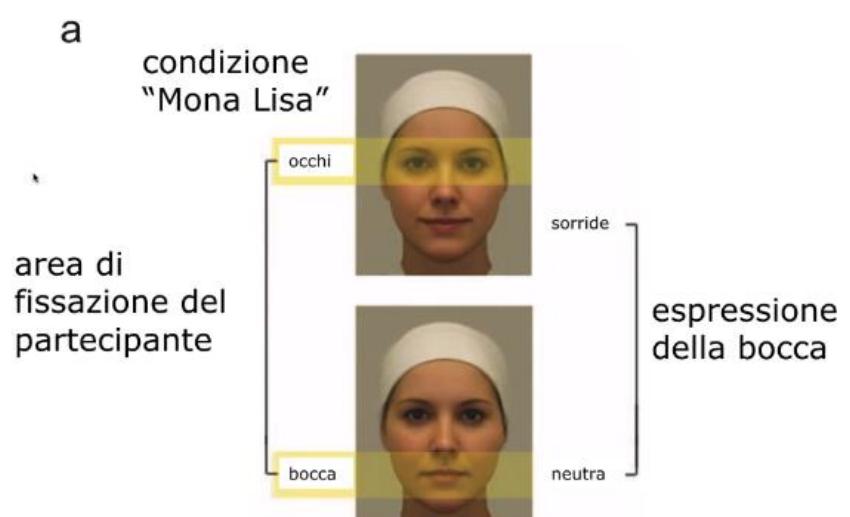
nell'immagine a sinistra dove l'immagine a destra in alto è stata passata in un filtro passa-basso mentre quella sotto con un filtro passa-alto.

Negli anni '70 la neuroscienziata Susan Beckmoore si è accorta che se faccio questa operazione di filtraggio, a causa di come sono disposte le discontinuità nel dipinto, succede che la zona che corrisponde alla bocca, che ha un sorriso asimmetrico, è più accentuato nell'immagine a bassa frequenza (risulta più incurvato) rispetto a quella ad alta frequenza. Questo è un primo elemento.

È come se dicesse che la Gioconda, secondo un'analisi dell'ottica di Fourier, sorridesse di più nelle immagini a basse frequenze piuttosto che nelle immagini ad alta frequenza. Se è vero che il sorriso è più nelle basse frequenze che nelle alte, allora questo potrebbe essere parte del motivo per cui il sorriso risulta enigmatico, enigmatico vuol dire che ti pone una domanda cui non sai rispondere, è un sorriso elusivo. Questa neuroscienziata, per spiegare questa elusività, ha proposto la seguente spiegazione, e cioè che quando guardiamo un quadro noi non teniamo gli occhi fermi (guardare vuol dire fare una scansione dell'immagine perché la visione è un processo attivo di saltare da una fissazione all'altra attraverso i movimenti saccadici) ma mettiamo in un atto una serie di saccadi saltando da certe zone ad altre. Quando si guarda un'immagine i movimenti oculi non sono casuali ma tendono ad andare verso certe zone che sono più salienti di altre, nei volti in particolare si tenderà a soffermarsi su occhi e bocca e in misura minore su punta del naso, mento e attaccatura capelli. Il motivo di ciò è dato dalla salienza di queste parti per la comunicazione interpersonale e anche per motivi di riconoscimento perché la relazione spaziale tra occhi e bocca dice molto sulla geometria del volto. In questo quadro le saccadi inizialmente sono su bocca-

occhi per poi passare alle mani, poi a lungo raggio il tronco e in seguito lo sfondo. Cosa succede quando guardo il volto e faccio questo pattern di fissazioni? Quando fisso la bocca uso la **fovea**, quindi la mia via reticolo-genicolato-striata riceva un segnale ad alta definizione , quindi un segnale ad alta frequenza, il che vuol dire che quando fisso la bocca l'informazione contiene anche l'equivalente neurale del volto filtrato per le alte frequenze, quindi riceve un segnale in cui il volto sorride di meno o non sorride affatto .Ma non appena vado sugli occhi, gli occhi sono ad alta frequenza ma la bocca rimane in una zona in cui l'acuità della retina è meno buona e i campi recettivi son più grandi, (è come filtrare le frequenze) ,ciò comporta che la bocca va in periferia della retina e quindi vanno a scomparire le alte frequenze, il che ci porta a percepire che la bocca sembra sorridere di più (discorso simile alla griglia di Hermann). Quando fisso gli occhi, la bocca sembra sorridere di più perché da questa parte del corpo ricevo più basse frequenze e come sappiamo dalla teoria, il sorriso è più accentuato nelle basse frequenze (quando lo fisso infatti sembra sorridere meno, perché va nelle alte frequenze). Questa caratteristica determina l'elusività, perché quando cerco di catturarlo non c'è più. Ma questa è una pura e semplice ipotesi che non ha la pretesa di spiegare del tutto il fascino della Gioconda.

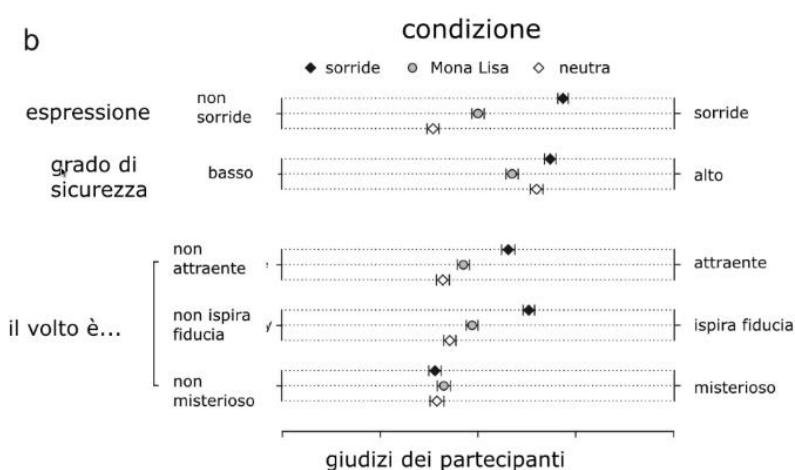
Un gruppo tedesco ha provato a vedere se si riesce a portare in laboratorio l'idea di Susan Beckmoore (?) e di testarla. Loro hanno usato volti di persone sconosciute perché la Gioconda è troppo conosciuta quindi si sarebbero creati Bias legati alle conoscenze pregresse. Questi volti sono stati creati semplificando il più possibile le informazioni a disposizione, per esempio togliendo i capelli, componente molto saliente di un volto. Si lasciano però le sue caratteristiche: occhi, bocca e naso.



Il compito: guardare queste foto e mentre le guardo un sistema monitora i movimenti oculari che classifica quando guardo gli occhi o la bocca e in funzione di questo modifica l'immagine stessa mentre la guardiamo.

In particolare c'è la condizione "Mona Lisa" in cui la natura di questa modifica è la seguente: quando il soggetto fissa gli occhi l'immagine è con la bocca appena sorridente, quando la fissazione è sulla bocca il sorriso non c'è più ma c'è un'espressione neutra. Il sistema riproduce il gioco di switch dell'espressione simile alla Gioconda perché quando fisso gli occhi, la bocca va nelle basse frequenze (e sembra sorridere) e viceversa. Il soggetto non si accorge che le immagini hanno questo cambiamento perché innanzitutto c'è questo gioco di basse e alte frequenze ma soprattutto perchè, quando noi moviamo gli occhi esiste il fenomeno della **soppressione saccadica**: durante il breve tempo in cui muoviamo gli occhi da una saccade all'altra, il sistema visivo interrompe l'output retinico. Ciò viene sfruttato per operare un cambiamento nello stimolo molto poco saliente e il soggetto non se ne accorge affatto.

Qui sotto troviamo i dati: c'erano 3 condizioni, una sperimentale e 2 di controllo:



- MonaLisa: sperimentale (scambio di emozioni a seconda di dove si fissa)
  - Volto sempre neutro : di controllo
  - Volto sempre sorridente: di controllo

Veniva dato un questionario dove si chiedeva ai soggetti di esprimere il grado di accordo con delle affermazioni (misura discreta da 1 a 7 o misura continua con una linea che va da “completamente disaccordo” a “completamente d'accordo”). Le barrette che delimitano gli indicatori rappresentano gli intervalli di fiducia , calcolati su un campione abbastanza ampio.

Il questionario prevedeva 5 domande:

- la prima riguarda l'espressione: “*questo volto sorride, quanto sei d'accordo?*”

La media della condizione neutra è spostata verso il no, la media della condizione sorride sempre è spostata verso l'essere completamente d'accordo, e guarda caso la condizione MonaLisa è precisamente a metà strada, le persone sono incerte. Questo essere in dubbio può essere interpretato appunto come qualcosa di elusivo dell'esperienza. Questi ultimi hanno anche un grado di sicurezza più basso, anche se non di tanto.

- Le altre domande cercano di sondare aspetti più legati all'estetica: “*il volto è attraente?*”

Nella condizione MonaLisa abbiamo sempre risposte a metà strada, la condizione di sorriso è più propenso a dire che è attraente mentre quello neutro che non è attraente, questo perché le persone preferiscono i volti sorridenti.

- “*Ispira fiducia?*”: i risultati sono identici.

Infine : - “*Quanto ti sembra misterioso questo volto?*” in questo caso la media delle risposte del gruppo MonaLisa è la più alta delle 3, quindi il risultato è nella giusta direzione però l'effetto è talmente minuscolo che non si può essere sicuri che questa differenza non sia data da un errore casuale (gli intervalli sono ampiamente sovrapposti) o da grandi differenze individuali.

## **Lezione 10**

### **Selettività per l'orientazione**

Fin qui ci siamo occupati del problema della **sensibilità al contrasto**, problema che è stato analizzato con tecniche psicofisiche (gratings). È stato visto come questi tipi di metodi psicofisici (metodologie di misurazione di soglie e studi sull'effetto dell'adattamento) tendono a suggerire che esistano meccanismi dedicati per la codifica di discontinuità orientate a diversi livelli di dettaglio, ossia **contorni** (“contorni” è un termine che si può usare per la prova d'esame, per essere più sintetici). Questi meccanismi possono essere pensati come delle specie di filtri neurali, cioè meccanismi che filtrano un certo livello di *frequenza spaziale* (un certo livello di dettaglio) e una certa *orientazione di contorno nello spazio*.

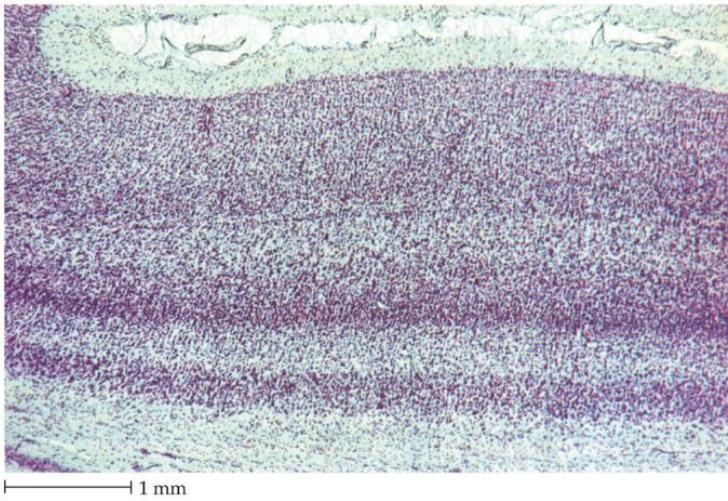
Questa “storia” rimane parallela ad un altro lavoro riportato da altri fisiologi, che hanno messo in luce come nella parte finale della via retino-genicolato-striata (visione di basso livello, V1), una caratteristica molto saliente di moltissime cellule presenti è la **selettività per l'orientazione**.

Da notare che il termine usato è **ORIENTAZIONE** (non orientamento) → in inglese la parola che si usa è “*orientation*”: si misura in gradi, es. ci dice se un contorno è  $0^\circ$ ,  $90^\circ$  ecc.

Oltre a questo, fin qui abbiamo visto che si possono avere delle manifestazioni psicofisiche sul fatto che ci siano delle unità che sono selettive per l'orientazione. La scorsa settimana è stato analizzato un esperimento di adattamento selettivo a un grating (reticolo), con una certa frequenza spaziale ed una certa orientazione. **L'adattamento avviene tramite un'osservazione prolungata** → l'adattamento produce una riduzione della sensibilità per il contrasto, quindi il reticolo, la cui orientazione è stata adattata, avendo ridotto la sensibilità per quella orientazione, diventa indistinguibile da un campo grigio. Importante è tenere in considerazione il fatto che è selettivo (avviene solo per un reticolo e non per l'altro).

Questi meccanismi non sono sensibili solo alla *frequenza spaziale*, ma anche all'*orientazione*: funzionano da filtri per una particolare orientazione, o meglio per un certo range di orientazioni, definito da una particolare funzione detta **curva di sintonizzazione**.

I protagonisti della lezione di oggi sono: David **Hubel** e Torsten **Wiesel**, premio Nobel per le ricerche sul sistema visivo di basso livello. Sono stati i ricercatori che hanno scoperto l'esistenza di queste particolari unità neurali, in V1, selettive per l'orientazione.

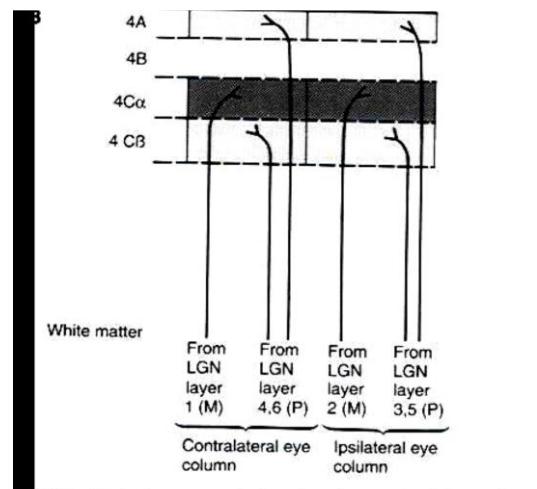


L'immagine a fianco rappresenta una sezione di V1: chiamata *corteccia striata* per la presenza della striscia bianca in alto, striscia di fibre nervose mielinizzate (assoni), molto visibile, che distingue, dal punto di vista citoarchitettonico, V1 dalle altre cortecce circostanti (V2 e V3, che sono invece extrastriate). Come il genicolato, anche V1 è diviso in 6 strati diversi: il numero 4 è quello su cui

arrivano le fibre che provengono dal genicolato stesso, ovvero le fibre della **radiazione ottica**. A sua volta il 4° strato è suddiviso in 3 sottostrati (A, B, C). Da un punto di vista fisiologico e psicologico, la cosa più rilevante è che in questi tre sottostrati l'input proveniente dalla radiazione ottica, segregato in base all'occhio di provenienza, viene ricombinato. Quindi si passa da una *rappresentazione monoculare*, dove i due occhi sono tenuti distinti, ad un livello successivo, dove troviamo una *rappresentazione binoculare*, dove l'input arriva da entrambi gli occhi.

Questo ha implicazioni da molti punti di vista: noi vediamo con due occhi distinti, ma la nostra esperienza del mondo è comunque unitaria, unica, quindi il cervello effettua una  **fusione binoculare**. In questo processo di fusione sono contenute delle informazioni importanti che riguardano lo spazio e la tridimensionalità del mondo (discorso particolarmente complesso, legato alla percezione dello spazio).

A fianco abbiamo uno dei tanti modi in cui si può vedere e mappare come l'input dal genicolato raggiunge i diversi sottostrati del 4° strato di V1: la cosa importante da notare è che all'inizio sono ancora segregati, in base all'occhio (contro e ipsi laterale), segregazione che continua sulla base della tipologia di cellule gangliari di provenienza (magno, parvo).



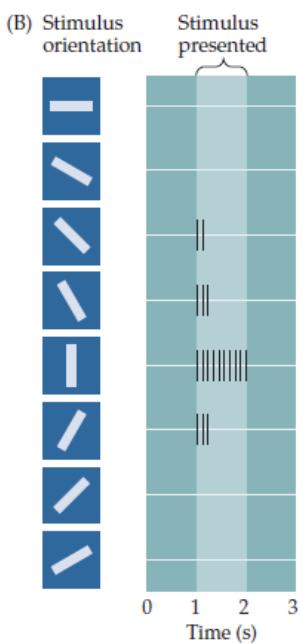
## Come Hubel e Wiesel hanno scoperto la selettività per l'orientazione?

Studiando l'area visiva primaria del cervello del gatto. Gli animali che loro studiavano erano anestetizzati, non erano gatti coscienti. Si riconosce che il meccanismo di cui ci si stava occupando è ancora precedente all'esperienza cosciente, è semplicemente il comportamento di questi particolari filtri per l'orientazione.

Ciò che è stato trovato è che, per molte tipologie di cellule in V1, lo stimolo in grado di produrre un eccitamento non era, a differenza di quanto si osserva nel genicolato e nelle cellule della retina, uno stimolo ad organizzazione concentrica (disco-anello), ma uno stimolo **orientato nello spazio → barra o contorno** (una discontinuità orientata).

Spesso le scoperte scientifiche si associano ad un mito sulla loro nascita (digressione su questo aspetto).

La cosa importante da capire è che il fenomeno della selettività per l'orientazione, che si osserva in queste cellule, non sta a significare che le cellule rispondono solo a barre che hanno un certo numero di gradi di inclinazione, ma significa che c'è una curva di sintonizzazione.

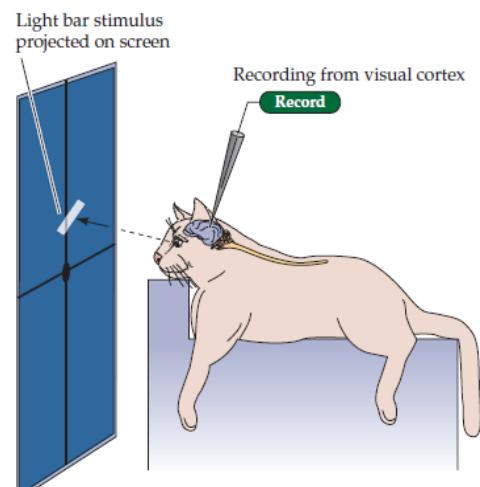


Nella figura a fianco a seconda dell'orientazione della barra presentata sullo schermo, nella posizione retinotopica appropriata, troviamo una modulazione della frequenza di scarica.

Per esempio, questa cellula, risponde in maniera ottimale ad un contorno verticale, ma risponde anche ad un contorno leggermente inclinato rispetto alla verticale, mentre è completamente silente nei confronti di un contorno orizzontale.

[Visione di due video → cellula del genicolato di cui viene mappato il capo visivo eccitatorio. Nel secondo video si vede l'attivazione di una cellula semplice.]

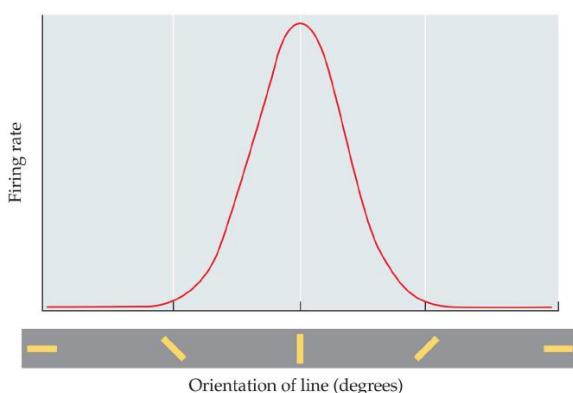
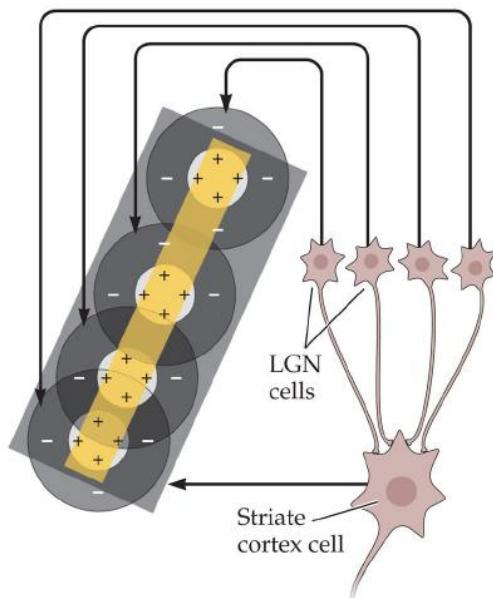
(A) Experimental setup



**CELLULE SEMPLICI** → cellule dove si deve trovare una particolare orientazione e posizione nel campo visivo. Quando si mette una luce lì la cellula spara, quindi quella è la zona eccitatoria, non è un disco ma una specie di barra; se si mette la luce ai lati di questa si ha una riduzione della frequenza di scarica (zona inibitoria). **Quando si cambia orientazione la cellula non si attiva** → l'interpretazione funzionale a cui si pensa è che queste unità neurali si comportano come dei rilevatori di linee e contorni: base per costruire la rappresentazione di una forma, di una superficie e poi di un'oggetto dell'ambiente.

Da un punto di vista psicofisico, rimane però più corretto pensare a queste cellule come dei *filtri neurali*, selettivi per l'orientazione e per la frequenza spaziale e che fanno parte di una rete neurale che si occupa di percepire le forme.

È abbastanza naturale pensare che questo fenomeno della selettività all'orientazione avvenga grazie alla convergenza di una serie di cellule nel genicolato, che corrispondono a particolari locazioni retiniche (posizioni nel campo visivo), le quali vengono codificate da dei campi recettivi centro-periferia (concentrici), e che sono allineate lungo una certa orientazione. Queste, convergendo sulla cellula semplice, producono un'unità neurale che è selettiva per quella specifica orientazione e non altre.



L'immagine raffigura la **curva di sintonizzazione**: studiando e misurando la frequenza di scarica in funzione dell'orientazione di diverse barre, si può arrivare ad un risultato come questo, in cui si definisce il range di orientazione a cui la cellula risponde. Vi sarà un'orientazione dello stimolo a cui la cellula risponde in modo ottimale, e altre che determineranno un'eccitazione meno marcata.

Oltre a questo, va aggiunto che Hubel e Wiesel hanno analizzato e descritto anche altre tipologie di cellule, sempre selettive all'orientazione ma con campi recettivi più complessi, come le:

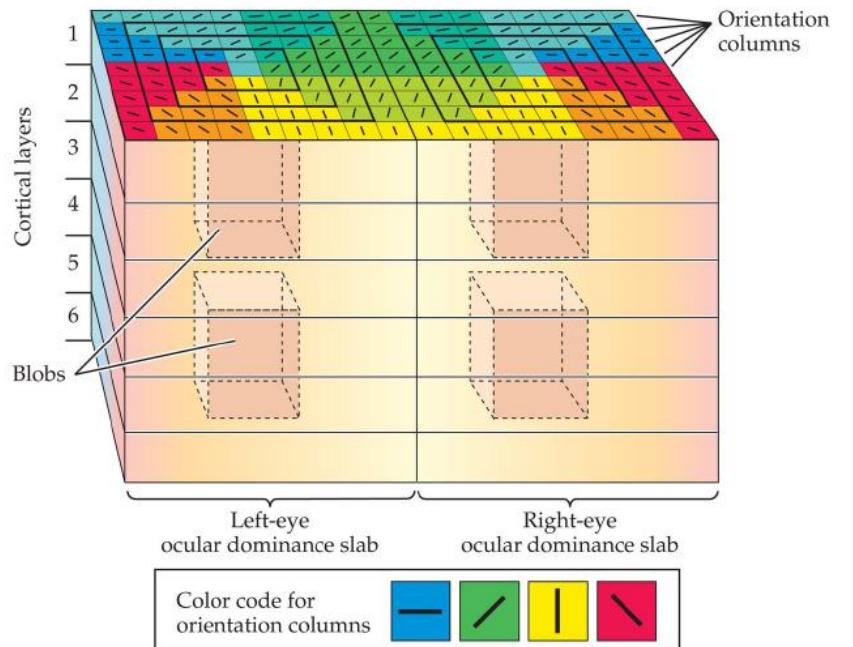
- **Cellule complesse;**
- **Cellule ipercomplesse;**
- **Cellule end-stopping** → cellule che posseggono un campo recettivo sensibile alla parte finale dello stimolo, dove preferiscono linee di una certa lunghezza

Sono tutti meccanismi che hanno a che fare con questo fenomeno generale della selettività all'orientazione (su cui focalizzare bene l'attenzione).

Infine, i due autori si sono anche dedicati a cercare di capire com'è fatta **l'architettura di V1**, ovvero come sono organizzate sulla superficie della corteccia striata, queste cellule: ciò che è stato riportato è che questa organizzazione è incredibilmente regolare.

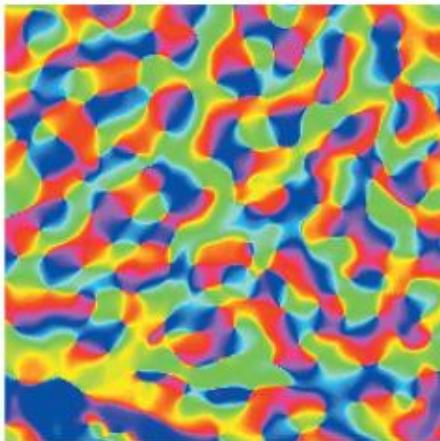
La tecnica utilizzata consisteva nel penetrare con un microelettrodo, in direzione ortogonale o parallela, la corteccia visiva. Con questa tecnica si sono accorti della presenza di un'alternanza regolare di attivazioni, che erano funzioni dell'occhio di provenienza.

Qui si vede che uno accanto all'altro vi sono le **COLONNE DI DOMINANZA OCULARE**: qui si trovano unità funzionali sensibili all'orientazione per input provenienti dall'occhio sx o dx (c'è ancora la segregazione in base all'occhio di provenienza).



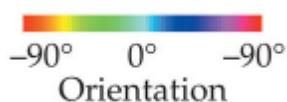
Oltre a queste sono presenti anche le **COLONNE DI ORIENTAZIONE** (orientation columns), dove cambiando l'orientazione dello stimolo assistiamo ad eccitazioni differenti e selettive delle cellule presenti.

Questi due elementi, insieme, producono una vera e propria **struttura modulare**, organizzata in colonne per l'orientazione accanto ad altre colonne di orientazione, con un'organizzazione sovraordinata in base all'occhio.



In seguito, hanno scoperto che in mezzo a queste colonne, vi sono delle altre zone, chiamate **BLOB**, dove vi troviamo unità funzionali non sensibili all'orientazione ma capaci di **codificare il contrasto di colore**.

Un'altra cosa curiosa è che questa struttura non è propriamente reale: è stata inferita tramite un'analisi accurata, ma tecniche più moderne hanno mostrato che questa organizzazione in V1, più che a delle colonne assomiglia a delle girandole.



Nell'immagine a sinistra, facendo un color coding in base all'orientazione preferita delle cellule, si nota come ogni colore corrisponde ad una specie di zona sensibile alla stessa orientazione. Al centro di ogni girandola vi è un blob, aspetto che suggerisce un meccanismo che mette insieme il contrasto di colore con l'informazione di orientazione.

Connettiamoci ora ad un discorso sul versante psicologico: *Che cosa vuol dire percepire linee e contorni? Fino a che punto questo discorso fatto sulla fisiologia della visione può spiegare la percezione di linee e contorni?*

Per rispondere a queste domande facciamo riferimento a **Gaetano Kanizsa** (1913-1993), uno degli psicologi italiani più influenti del secolo scorso e probabilmente uno dei più importanti scienziati della visione, di formazione Gestaltista. Essendo di origine ebrea, la sua carriera universitaria è stata inizialmente ostacolata durante il periodo fascista. Subito dopo la guerra è diventato professore ordinario di psicologia generale, fondando l'Istituto di Psicologia a Trieste.



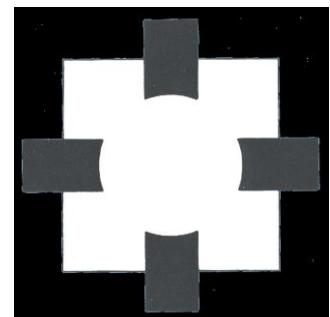
Kanizsa è famoso per il suo **triangolo**, scoperta di per sé molto antica: nel 1954, al congresso nazionale degli psicologi italiani, Kanizsa si presenta con una relazione su linee e contorni, in cui presenta il famoso triangolo. L'anno successivo pubblica un articolo intitolato “*Margini quasi-percettivi in campi con simulazione omogenea*”. L'articolo inizia con un'analisi di natura fenomenologica, ovvero un'analisi su cosa sono, nella nostra esperienza percettiva cosciente, linee e contorni. L'autore faceva notare che esiste una specie di *continuum* nella fenomenologia di quelle che si riconoscono come linee e contorni, che corrisponde al *continuum* di ciò che Kanizsa stesso considerava il pensare e il vedere, due aspetti diversi della nostra cognizione, della nostra vita mentale.

Kanizsa argomentava in modo principalmente visivo, mostrava le cose che analizzava: il suo stile di ricerca era l'analisi di situazione (disegni) che illustrassero il concetto, evidenza empirica non basata su un'analisi statistica ma su ciò che concretamente tutti potevano vedere.

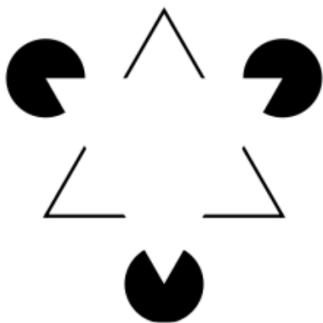
Il primo concetto che si trova nell'articolo è quello di **LINEA IMMAGINARIA**: si tratta di linee non esistenti ma solo “pensate” dal soggetto che osserva gli stimoli ambientali (per esempio, se si pensa a due punti a distanza ravvicinata si può immaginare una linea che li collega senza che questa sia realmente presente).

Questa tipologia è distinta dall'autore dalla **LINEA VIRTUALE**: fenomeno per cui, in certi casi, vi è una tendenza spontanea a strutturare il campo di elementi lungo una certa direzione, provocando la formazione di linee virtuali che esulano da quello che realmente e concretamente è presente (per esempio, punti vicini tra loro possono essere organizzati in una linea orizzontale), questa tipologia di linea diventa saliente nell'esperienza percettiva del soggetto.

Esiste anche la **LINEA “REALE”**, che viene differenziata dalle altre due in quanto, in questo caso, ciò che noi percepiamo corrisponde ad una effettiva discontinuità orientata nella stimolazione. È bene utilizzare il termine “reale” in modo relativo, in quanto anche le linee virtuali possono essere riconosciute come reali dal soggetto, che ha esperienza di queste a livello percettivo. Queste linee reali possono diventare margini reali quando sono presenti in veste di confine che delimita una superficie (esempio: un cerchio e un triangolo, hanno margini e contorni che sono reali perché corrispondono ad una discontinuità orientata, a un certo livello di dettaglio, nel campo visivo).



Secondo Kanizsa esisteva però anche un'altra tipologia di margine: nel 1954-55, ha coniato il termine **MARGINE QUASI-PERCETTIVO** → tipologia di margine proposto per la prima volta da un autore tedesco, la cui particolarità è che da un punto di vista fisico ci si accorge del fatto che non è realmente presente una discontinuità (cerchio), anche se globalmente sembra essere percettivamente presente.



Dall'analisi di queste figure Kanizsa stesso propone il suo famoso **triangolo**: riconosciuto come una sorta di interfaccia tra la visione di basso e di alto livello.

Kanizsa faceva notare come il triangolo sembri avere un margine, sembra esistere una discontinuità come margine del triangolo. Egli spiega questo fenomeno tramite principi psicologici di origine gestaltista: sarebbero, in questo caso, in gioco vari meccanismi percettivi-fisiologici, uno tra i quali è la tendenza a percepire uno stimolo in modo semplice e più economico possibile, che rifletterebbe una preferenza per le figure più semplici e regolari → meccanismo che tende a spingere il sistema visivo verso un'interpretazione che consente di codificare la comparsa del triangolo, invece che dover analizzare stimoli più complessi come i cerchi neri incompleti. Lo stesso discorso può essere fatto anche per il triangolo nero, caratterizzato da interruzioni dovute sempre alla presenza del triangolo bianco, che sembra stargli davanti.

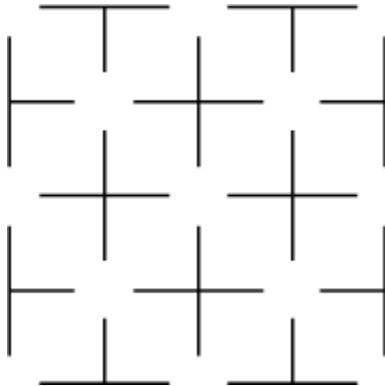
Tutte questi meccanismi secondo Kanizsa sono manifestazioni del processo di formazione delle superfici: processo di integrazione dei contorni, che vengono messe insieme, creando oggetti che si segregano da altri oggetti e dallo sfondo (*segregazione figura-sfondo*). Egli faceva notare come questi principi sono responsabili, quindi, della formazione di unità percettive, di forme, della loro segregazione rispetto ad altre o allo sfondo e di come questo produca anche una specie di proto-stratificazione nello spazio: il *triangolo quasi-percettivo* viene percepito come “davanti” agli elementi neri. **Si parla quindi di diversi piani di distanza e di un ordinamento specifico nello spazio.**

Questi particolari stimoli, nel corso del tempo, hanno subito delle modificazioni teoriche a livello di terminologia: Kanizsa negli anni '50 li ha denominati “**margini quasi-percettivi**”, termine modificato quasi subito in quanto non particolarmente identificante questi particolari oggetti.

A partire dagli anni '70 vennero chiamati “**margini anomali**”: si vede un margine a livello percettivo, ma nello stimolo localmente non esiste una reale discontinuità orientata, si vede dunque un contorno (una linea) anche all'interno di un campo visivo a stimolazione omogenea (→ lì le unità di Hubel e Wiesel non registrano nulla). Successivamente, un famosissimo ricercatore e teorico della percezione, Richard Gregory, negli anni '80, pubblicò su *Nature* un articolo intitolato “**Contorni cognitivi**” in cui presentava il triangolo di Kanizsa. Per molto poco tempo è stato utilizzato questo termine che è poi stato sostituito da “**contorni illusori**” (ad oggi ancora utilizzato).

Di seguito qualche altro esempio di figura illusoria:

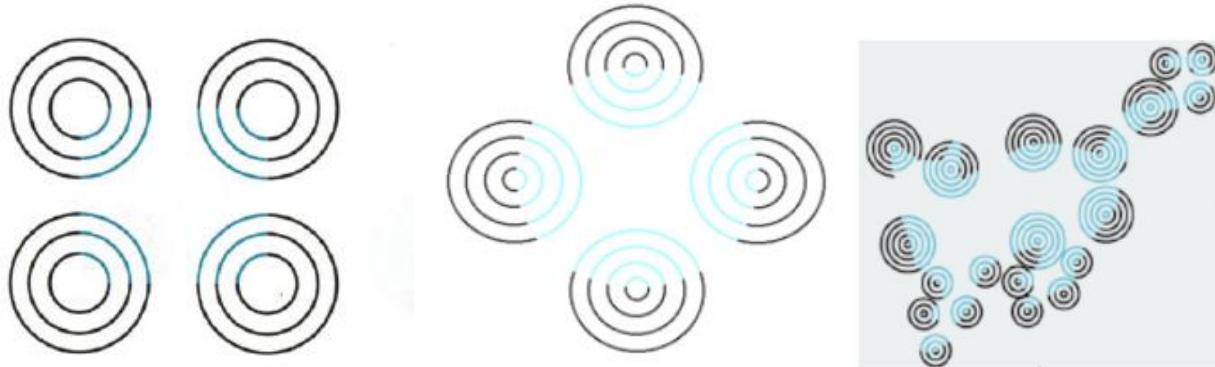
In questa immagine si percepisce la zona centrale delle linee avente forma di cerchio, che risulta essere



anche più luminoso, vi è quindi anche un aspetto legato alla percezione del contrasto (della luminosità).

Possono essere percepite anche delle linee bianche, sovrapposte a quelle nere.

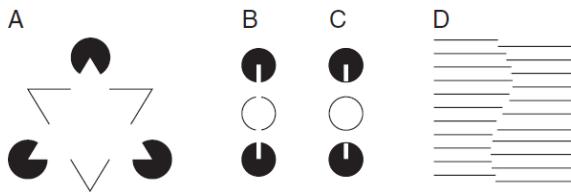
Una vera e propria spiegazione di questo fenomeno in letteratura non si trova ancora.



Queste altre tipologie di illusioni fanno sì che l'occhio umano percepisca il bianco dello sfondo come unificato al colore interno delle immagini (azzurro chiaro), unificandolo nell'immagine del quadrato (nella prima), del cerchio (nella seconda) e del gatto (nella terza), quando invece non c'è una reale differenza di colore sullo sfondo.

Questi fenomeni (chiamati anche “*effetto neon*”) sembrano dare supporto all'idea che denominare questi stimoli come quasi-percettivi sia effettivamente riduzionistico.

Quando Kanizsa mostrava il suo triangolo e chiedeva ai fisiologi se si potesse spiegare con la selettività all'orientazione come funzionasse il fenomeno stesso, in realtà non aveva del tutto ragione: negli anni '80, un fisiologo tedesco è riuscito a dimostrare che il ragionamento teorico sulla selettività all'orientazione poteva effettivamente spiegare, almeno in parte, il fenomeno del triangolo. Si tratta di una tipologia di selettività all'orientazione che è diversa da quella studiata da Hubel e Wiesel. Il lavoro di ricerca è stato poi pubblicato su *Science*.



**FIGURE 5.14** Sensitivity to illusory contours in V2. **A**, **B**, and **D** illustrate illusions in which many report seeing a white triangle (**A**), a vertical white line (**B**), or a curving boundary between two surfaces (**D**). **C** illustrates how just a small modification to the image can remove the illusion. Source: Von der Heydt, Peterhans, and Baumgartner, 1984. Reproduced with permission of AAAS.

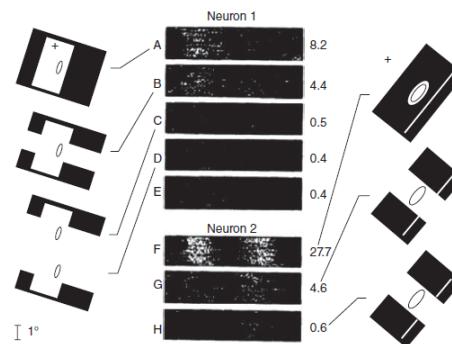
Sulle immagini che sono state proposte vale il fenomeno del triangolo di Kanizsa.

- A) Triangolo di Kanizsa (non utilizzato in quanto oggetto complesso);
- B) e C) sono stimoli in cui sulla barra bianca centrale vige lo stesso fenomeno che si trova sul triangolo bianco in A.

L'idea del ricercatore era quella di andare ad analizzare se potessero esistere delle cellule in V2 (nella scimmia), che rispondono all'**orientazione della barra illusoria** (**B**), ma non rispondono alla situazione tipo **C**, in cui non si vede una barra bianca illusoria.

Successivamente, sono stati utilizzati dei contorni generati dall'allineamento dei terminatori di linee (**D**): si vede in questo caso un contorno illusorio curvilineo.

La figura mostra un ingrandimento di un'immagine presente nell'articolo, dove troviamo il neurone 1 e 2: si osserva quindi la modulazione di frequenza di scarica di due singoli neuroni della corteccia V2, area extra striata, della scimmia. La colonna al centro (puntini bianchi) rappresenta la **frequenza di scarica**: più sono densi, più il neurone scarica. Ogni pannello nero si riferisce a diverse tipologie di stimolo, a sx per il neurone 1 e a dx per il neurone 2. I numerini presenti si riferiscono ad una statistica che indica la frequenza di scarica (più sono alti più la frequenza è alta).



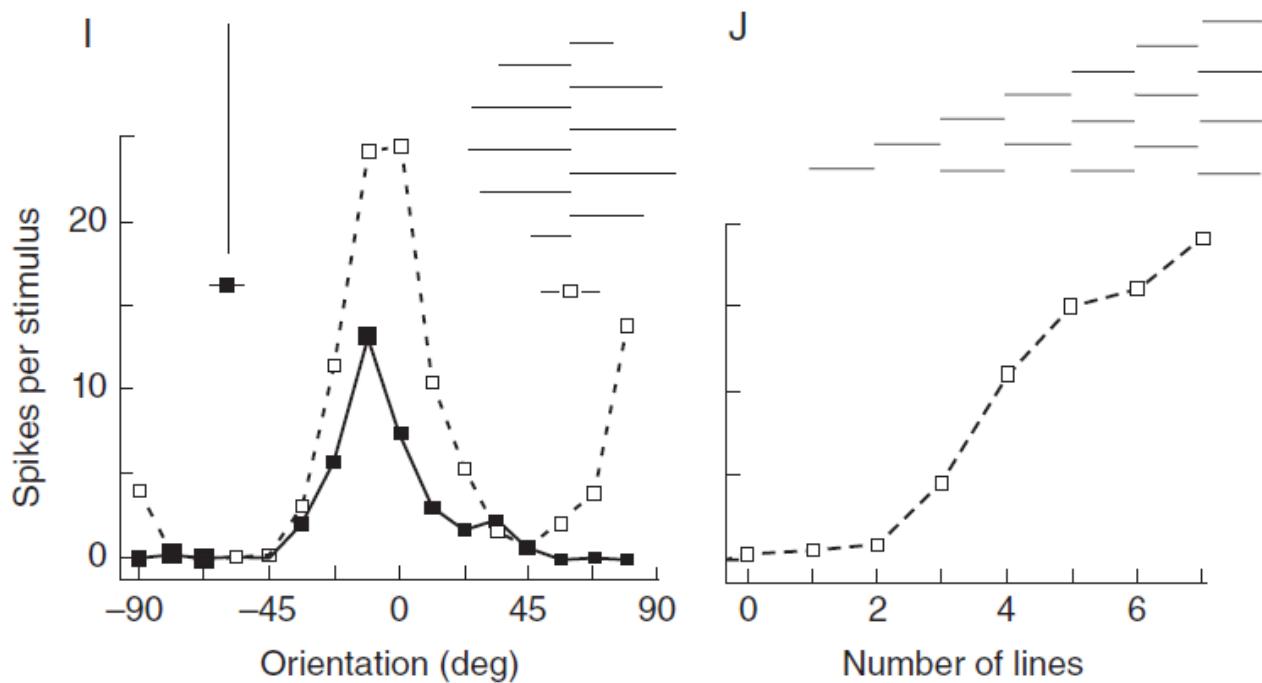
**FIGURE 5.15** The responses of two V2 neurons reported by von der Heydt, Peterhans, and Baumgartner (1984; center column), with the stimuli producing these responses shown to the left (for neuron 1) and to the right (neuron 2). For both neurons, the receptive fields are represented by ellipses. Neuron 1 responds most strongly to the right-side edge of an oriented bar (raster plot A), but also to the analogous illusory edge (raster plot B), a response that is abolished when either the top or bottom half of the stimulus is removed (raster plots C and D). Raster plot E shows spontaneous activity. Neuron 2 responds most strongly to a narrow oriented bar (raster plot F), less strongly to the illusory bar (raster plot G), and effectively not at all to a minor modification that abolishes the illusory percept (raster plot H). Source: Von der Heydt, Peterhans, and Baumgartner, 1984. Reproduced with permission of AAAS.

Gli stimoli che venivano presentati sono:

- A) Stimolo in cui è presente un margine reale: rettangolo bianco e uno nero. L'ellisse corrisponde a il *campo recettivo retinotopico* della cellula in V2, da cui regista. Il rettangolo bianco viene fatto scivolare nel campo visivo della scimmia fino a scivolare sull'ellisse. In questo punto la cellula risponde a quella particolare orientazione (risposta abbastanza vigorosa);
- B) Stimolo in cui non vi è più una vera discontinuità orientata, il campo visivo è omogeneo, però anche in questo caso la cellula risponde, anche se un po' meno vigorosamente;
- C) Composto solo dalla parte superiore dello stimolo B. La cellula risponde molto meno vigorosamente rispetto a B;
- D) Composto dalla parte inferiore dello stimolo in B. La cellula risponde molto meno vigorosamente rispetto a B;
- E) Presentazione di un campo nero. Cellula non risponde.

**Questa cellula sembra rispondere ad un margine reale, ma risponde anche al margine illusorio: non vi è una vera differenza tra i due tipi di stimoli.**

- F) Stimolo in cui è presente un margine reale bianco, che produce una forte frequenza di scarica;
- G) Stimolo in cui si presenta un margine illusorio (margine di Kanizsa), la cellula continua a scaricare in modo simile alla situazione F;
- H) Situazione di controllo, in cui viene leggermente modificata la figura in maniera da annullare la possibilità di interpretare la barra come davanti a due rettangoli neri, ecco che non abbiamo risposte ma solo, presumibilmente, un po' di frequenza di scarica spontanea;



**FIGURE 5.16** Data from neuron 3 of von der Heydt, Peterhans, and Baumgartner (1984). Panel I shows that the tuning of this V2 neuron is to a (real) black bar oriented nearly vertically (solid black lines on plot). The dashed lines show the cell's tuning to the illusory line created by the stimulus to the right is nearly identical (although the response gain is higher for the illusory than the “real” stimulus). Panel J shows the sensitivity of this neuron to the strength of the illusory percept, its responses increasing as a function of the number of lines added to the gratings in the stimulus. Source: Von der Heydt, Peterhans, and Baumgartner, 1984. Reproduced with permission of AAAS.

Facendo riferimento ad un terzo neurone, sono stati presi stimoli composti da barre verticali “reali” che determinano una frequenza di scarica massima da parte della cellula (curva di sintonizzazione per l’orientazione → curva nera).

Successivamente ha proposto uno stimolo in cui è presente una barra verticale illusoria (*barra di Kanizsa*), realizzata tramite uno stimolo circolare composto da barre orizzontali interrotte al centro. In questo caso la curva di sintonizzazione mostra che questa cellula risponde in modo più marcato rispetto alla condizione precedente (curva tratteggiata).

Questo fenomeno diventa sempre più saliente ed evidente quante più barre orizzontali si inseriscono nello stimolo: nel grafico J si vede la frequenza di scarica media della cellula in rapporto al numero di segmenti inseriti nello stimolo → **più è il numero di linee più è vigorosa la frequenza di scarica.**

## Lezione 11

### Recap lezione precedente

L'ultima cosa di cui abbiamo parlato sono questa categoria di fenomeni chiamati figure illusorie tra cui il **triangolo di Kanizsa**, come principale prototipo. Con questa illusione, **è possibile immaginare una maniera in cui il fenomeno della selettività per l'orientazione di un contorno, può essere applicato anche alle figure illusorie immaginando un meccanismo neurale come quello descritto**. Tali neuroni permettono di considerare un *margine “illusorio”* come un *margine “reale”*, non distinguendoli. La cellula è selettiva per un margine indipendentemente dal fatto che ci sia o non ci sia nel suo campo recettivo una discontinuità orientata, cioè una differenza di densità luminosa. Evidentemente questo tipo di cellula ha un campo recettivo più complicato e non riceve solo informazioni da quel punto a livello locale, ma tiene conto anche del contesto. Questo è anche un esempio di un processo di elaborazione dello stimolo che non verrebbe più classificata come *early vision*, in quanto avviene un processo maggiore di integrazione delle informazioni. Qualcuno infatti in tale ambito preferisce parlare di **campo percettivo** invece di **campo recettivo** in quanto non sono più i recettori a contare effettivamente, ma sono le interazioni reciproche che permettono l'integrazione dell'informazione.

### Ecologia della visione

Abbiamo esaminato il fenomeno per la selettività per l'orientazione e per la percezione delle figure illusorie da un lato come farebbe uno psicologo, guardando le illusioni e riflettendoci sopra, ma dall'altro come lo farebbe un fisiologo, vedendo come risponderebbe un'unità neurale davanti a tali immagini. Per mettere insieme queste due cose e per capire il senso di questo processo è necessario analizzarli anche da un terzo punto di vista, che è quello dell'**ecologia della visione**.

In uno dei primi incontri abbiamo parlato dell'idea proposta da **David Marr** il quale sosteneva come per spiegare la percezione avremmo dovuto mettere insieme 3 livelli di spiegazione diversi:

1. *Livello dell'implementazione* e cioè come l'informazione viene implementata in un qualche hardware che può essere il cervello o un computer;
2. *Livello dell'algoritmo* in cui si cerca di capire il processo indipendentemente dall'hardware (taglio più psicologico);
3. *Livello della teoria computazionale* che cerca di capire qual è la funzione adattiva di un determinato processo biologico come può essere la visione.

Per capire il significato dal punto di vista adattivo di un processo, per Marr è necessario studiare una serie di cose che hanno a che fare con i vincoli fisici come la gravità, ma anche i **vincoli biologici** e cioè **i limiti che un organismo possiede**. Applicando l'idea di Marr si vede come l'algoritmo di Kanizsa, e cioè quel processo che egli descriveva quando parlava del triangolo, il processo che tende a completare le cose per trovare una soluzione in termini più semplici e completi, non era altro che un *modo molto più economico* rispetto alla descrizione puramente geometrica di uno stimolo dove avremmo avuto molte interruzioni. Secondo Kanizsa l'algoritmo è ispirato a questi principi e infatti di base questo è conseguenza di una sequenza di calcoli neurali che cercano di arrivare ad una soluzione generale dal punto di vista percettivo. Una parte di questa implementazione potrebbe essere dovuto alle **cellule semplici** di Hubel e Wiesel che codificano sulla base dell'orientamento spaziale.

Dal punto di vista dell'ecologia della visione, si potrebbero fare diverse domande tra cui: ***per quale motivo è adattivo per un organismo avere un sistema visivo che inventa figure che non ci sono?***



Per cercare di capire, possiamo partire da una foto scattata da un artista cinese, Liu Bolin, per una campagna di moda di Moncler. **Cosa vediamo in questa foto?** C'è un personaggio nella foto che sembra quasi trasparente, ma non del tutto invisibile. Qui se facessimo un'analisi localmente in varie zone dell'immagine andando a vedere cosa c'è nello stimolo, scopriremmo che in alcune aree ci sono dei contorni reali e in altri punti ci sono invece dei contorni illusori (vediamo il contorno ma non c'è discontinuità). Questo succede per esempio se osserviamo le gambe dell'artista. L'artista gioca con un fenomeno chiamato **mimetismo** che è un po' il problema computazionale di come rendersi invisibili e non in senso fantascientifico. Per riuscire a rendersi invisibili bisogna prendere le sembianze di ciò che abbiamo alle nostre spalle e infatti Liu Bolin si dipinge il corpo con quelli che sono gli oggetti presenti alle sue spalle.

***Ma perché questo è interessante rispetto al problema iniziale?***

Proviamo ad osservare questa immagine. Se proviamo ad osservarla sicuramente ci saranno persone che riusciranno a scorgere qualcosa mentre altri no. In realtà all'interno di essa si può osservare una mucca. È molto interessante questo stimolo da tanti punti di vista. Intanto si faccia caso al fatto che questo è uno stimolo leggermente degradato, come una fotografia venuta male. Quando noi vediamo la mucca, stiamo cercando di trovare una maniera di



piazzare l'immagine dove ci sono i contorni che in parte saranno definibili come virtuali (parzialmente connessi ma non si vede realmente il contorno), ma anche quelli illusori.

Dal punto di vista visivo e non solo, noi siamo delle “**macchine per classificare**”, abbiamo una potente capacità di trovare delle regolarità e percepire delle strutture anche quando sono molti deboli o piene di rumore. Quindi, ad esempio, guardiamo le nuvole e vediamo delle immagini; guardiamo le stelle e vediamo le costellazioni. Molto spesso questo è il motivo per cui le persone sono convinte di aver visto gli ufo o altre immagini. Queste cose vengono dette **paraidolie** che ci permettono di osservare cose che in realtà non sono presenti sfruttando la nostra estrema sensibilità. Anche per l'immagine della mucca siamo vicini a tali condizioni in quanto il nostro cervello proietta delle immagini che in realtà non sono presenti. Tale fenomeno però non è solo visivo ma succede anche nell'udito. Per esempio, se dovessimo prendere un disco e porlo sul giradischi e facessimo girare nel senso opposto il disco, noi inizieremmo a sentire dei rumori totalmente disorganizzati. Su tale fenomeno molte persone hanno costruito veri e propri miti come per esempio quello riferito ai Beatles e al fatto che si potesse sentire la voce del diavolo.

Facendo un'analisi rigorosa di come è strutturata la luce presente nell'ambiente che arriva al nostro occhio, ci si accorge che l'informazione sui contorni localmente, cioè prendendo delle zone piccole della nostra retina che costituiscono il campo recettivo delle cellule specifiche per i contorni, spesso non è così ben definita la situazione, anzi. In un certo senso nell'ecologia della nostra visione, in maniera meno estrema della foto della mucca, queste cose possono succedere e naturalmente sono quelle che diventano importanti dal punto di vista evolutivo se siamo una preda o un predatore.

Per fare un esempio pratico possiamo pensare alle **rane**. Alcune di esse cercano di nascondersi dai predatori mimetizzandosi con le superfici. D'altra parte, però i predatori cercano di usare delle strategie per riuscire a individuare rapidamente la loro preda, contrastando tale mossa attuata dalla rana. Anche la sogliola usa la stessa strategia per nascondersi sul fondo del mare. Osservando un video dove è presente la sogliola, notiamo come **nel momento in cui si mette in movimento, si crea un contorno ben visibile cosa che scompare nel momento in cui si riferma**. Tale contorno è illusorio come quello di Kanisza in quanto è un prodotto del nostro cervello.

Tutto quello che abbiamo detto fino ad ora è un po' una risposta alla domanda iniziale relativa alla **capacità adattativa** di tale comportamento per noi esseri viventi. Dire che lì c'è qualcosa non è uno sbaglio perché di base lì non c'è un contorno diverso tra la superficie e l'animale. Noi riusciamo a superare tale limite nell'ecologia del sistema visivo in maniera efficace: è proprio tale capacità ad essere adattiva.

L'illusione non deve essere percepita come qualcosa di estremamente negativo in quanto può essere anche qualcosa da intendere come positivo.

Capiamo quindi come un primo modo per superare tale limite è sicuramente **il movimento** e un altro modo per farlo è quello di **confrontare più punti di vista tra loro**. Per alcuni psicologi evoluzionistici è proprio questo il motivo che si nasconde dietro l'evoluzione di una suddivisione binoculare dove i campi visivi sono ampiamente sovrapposti, cioè la zona binoculare del campo visivo è molto grande. Questo aiuta ad esplorare l'ambiente con due punti di vista diversi aiutando ad individuare queste forme nascoste. Ci sono però altri animali che hanno due occhi (campo visivo binoculare) con una zona di sovrapposizione nulla in quanto i loro occhi puntano verso l'esterno, sono laterali e li aiutano ad avere un campo visivo molto più ampio, vedendo anche dietro. Solitamente essi sono prede e questa conformazione del sistema visivo gli serve per tenere sotto controllo l'ambiente.

Se vediamo un altro video riguardo il cane dalmata notiamo la stessa identica cosa: **il movimento permette di individuare rapidamente l'animale che da fermo si confonde con lo sfondo**. Questo non è assolutamente un'illusione anzi, è molto evidente nel momento in cui si muove. Infine, diventa saliente anche la struttura tridimensionale, la quale emerge.

## Canny detector

La questione del mimetismo è diventata sempre più importante quando ci si è iniziati a porre il problema riguardo l'intelligenza artificiale intorno agli anni 80 e di come si potesse creare un *sistema di visione artificiale*. Lo stesso David Marr è considerato un pioniere dell'approccio computazionale e della visione artificiale. Ci sono stati tantissimi studiosi del settore che si sono posti il problema relativo all'individuazione dei contorni degli oggetti una volta effettuata una fotografia. All'inizio sembrava un problema abbastanza semplice individuarli in quanto bisognava osservare le discontinuità o il cambio di luminanza. Nel momento in cui si è iniziato a lavorare per davvero su questo ci si è accorti che in realtà non è così banale. Si può fare un algoritmo che cerca di individuare queste discontinuità però esso avrà una tendenza a trovare discontinuità dove in realtà non ci sono e viceversa.

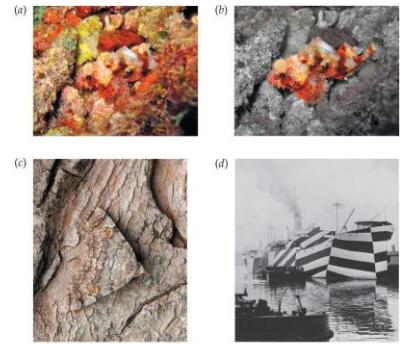


Una delle versioni più recenti è il **CANNY DETECTOR** che è molto più sofisticato di quello che era presente negli anni 80, ma allo stesso tempo notiamo dalla figura come, se andiamo a confrontare la foto con l'analisi effettuata dal detector, ci sono dei punti dove non vengono trovati i contorni e altri dove invece vengono trovati contorni quando in realtà nella foto non ci sono. In generale da questo problema qui si è capito che un'analisi locale dello stimolo funziona abbastanza male se si vuole capire dove sono i contorni degli oggetti.

### ***Ma perché è così importante avere un meccanismo che individua le figure illusorie?***

Perché in un certo senso tutte le figure che noi vediamo sono illusorie. Non nel senso che ci sbagliamo a vederle, ma nel senso che sono il **risultato di un'attività integrativa** che mette insieme uno stimolo che può essere localmente non del tutto esauriente. Questo torna molto con il fenomeno del *mimetismo*, ma molto spesso si sfrutta la proprietà contraria e cioè quella del rendere più grande la figura per incutere timore nei confronti del predatore/aggressore.

Per esempio, durante le due guerre mondiali si sfruttava tale capacità illusoria sulle navi dipingendole in modo tale da creare delle forme e dei contorni che portavano il puntatore della nave nemica a puntare nella direzione sbagliata. In natura non è una cosa inesistente, basti pensare al manto delle zebre il quale assomiglia molto a questo tipo di mimetismo o dazzle. Questo principio di creare forme che non ci sono lo troviamo anche in altre forme di mimetismo come per esempio nella farfalla.



SENSATION & PERCEPTION 4e, Figure 4.21  
© 2015 Sinauer Associates, Inc.

Possiamo concludere sottolineando come questa tendenza a percepire contorni illusori non dovrebbe essere pensata come illusione, ma come la **manifestazione di una tendenza molto adattiva di costruire contorni su una base di informazioni che non sempre sono del tutto esaustive**.

Dopo aver fatto la trasduzione dal livello basso al livello intermedio, si è iniziata ad avere un'idea su quali siano i processi che danno luogo ai processi visivi. Solitamente questi processi sono etichettati come **organizzazione percettiva**. Gli studiosi della percezione hanno cercato di dare una definizione di quello che è la **middle vision**, ovvero uno stadio dell'elaborazione visiva la cui definizione è approssimativa: è uno stadio che avviene dopo che sono state estratte le caratteristiche di base, basic features il cui significato è “contorni”.

Molte features sono caratterizzate dal colore. L'estrazione delle loro caratteristiche è un lavoro che appartiene alla **visione di basso livello**, o anche definite **caratteristiche locali**, ovvero sono cose che possono essere rilevate con un'analisi locale dell'informazione codificata dalla retina. Tutto questo avviene prima che ci siano processi più complessi che permettono il riconoscimento e la comprensione di come è fatta una scena, partendo dalla struttura tridimensionale di quest'ultima. Ci sono processi di riconoscimento molto più complessi che chiamano in gioco la memoria, l'attenzione e altri processi di ordine superiore.

La visione intermedia rappresenta un **livello percettivo presemantico**, in cui la percezione è ancora indipendente dall'elaborazione percettiva di alto livello. In questo modo è più facile produrre delle illusioni. Ad esempio, la percezione dei *piani inclinati impossibili*, ovvero percepire uno stimolo orientato che non rispetta le leggi della fisica e quelle relative alla forza di gravità.

### **Cosa avviene a livello di middle vision?**

A questo livello sono presenti dei **processi integrativi**, di elaborazione globale a lungo raggio responsabili dell'elaborazione di unità percettive (si potrebbe parlare anche di oggetti, ma secondo il prof parlare di oggetti equivale a parlare di qualcosa di più complesso). Per **unità percettiva** invece si intende qualcosa fatto di elementi dello stimolo, messi insieme, suddivisi e segregati. **Questa attività di selezione, unificazione, differenziazione che risulta essere molto elementare, è una forma di categorizzazione, in cui la visione suddivide il mondo in entità, che devono avere un significato biologico, ovvero deve essere adattivo percepirlle.**

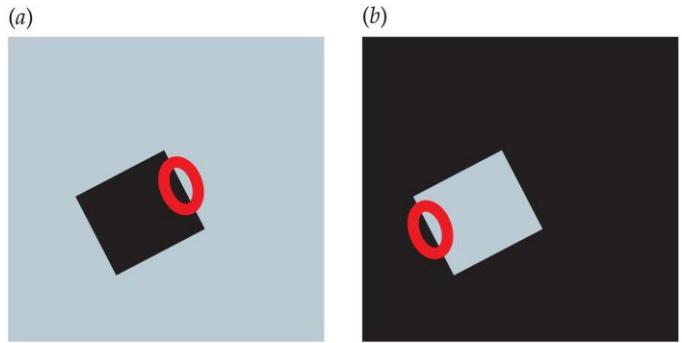
Il riconoscimento di uno stimolo è una forma di **categorizzazione**. Si parla invece di **protocategorizzazione**, quando si parla di *figura e sfondo* e di suddividere e segregare uno stimolo. Ad esempio, quando si parla una lingua conosciuta si tendono a percepire le pause. Gli stacchi presenti tra una parola e l'altra sono il frutto di un lavoro di segregazione e ricategorizzazione delle parole all'interno di una frase. La categorizzazione e la ricategorizzazione viene effettuata attraverso il match degli elementi presenti in memoria. La differenza nel percepire i suoni appartenenti ad una lingua conosciuta e i suoni presenti in lingue sconosciute è la possibilità di segregare e categorizzare un suono.

La percezione può essere segregata in tre categorie:

1. **Border-ownership** ovvero l'appartenenza di un contorno, (quello che potremmo definire “proprietario”);
2. **Segmentazione figura sfondo;**
3. **Formazione di unità.**

#### 1. Analizziamo il **border-ownership**

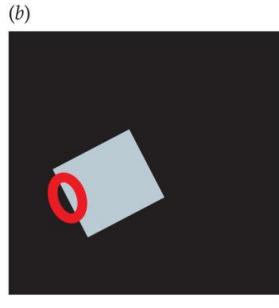
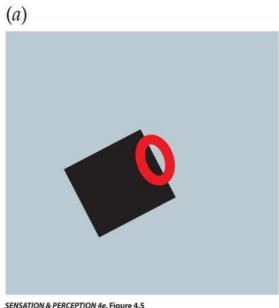
Pensiamo alle cellule semplici di Hubel e Visel, caratterizzate per la specificità di orientazione. Il campo percettivo delle cellule è orientato, possiamo percepirllo come un'ellisse, il quale è costruito per caratterizzare una selettività orientata.



SENSATION & PERCEPTION 4e, Figure 4.5  
© 2015 Sinauer Associates, Inc.

In A e B è possibile visualizzare la stessa figura (quadrato ruotato su sfondo omogeneo). In A il quadrato è nero e lo sfondo è grigio, mentre in B è presente il contrario. Se si provasse a mettere la zona di una di queste due figure nel campo recettivo di una di queste cellule semplici, precisamente dove è presente l'ellisse rosso, esse **sarebbero ugualmente eccitate**, in quanto, non in grado di distinguere la situazione A e B, quindi a livello di V1 abbiamo una medesima risposta. In A il proprietario del bordo è il lato nero, i margini degli oggetti hanno una funzione unidirezionale. In A il margine è di una figura nera in B invece è il grigio e non il nero, si potrebbe affermare che solo le figure hanno uno sfondo, gli sfondi percettivamente non hanno nulla. Dal punto di vista percettivo ciò che è saliente nella nostra esperienza è la **forma delle figure**. Il vuoto fra le figure è difficile da identificare. Una delle competenze più importanti per imparare a fare il disegno di uno stimolo veritiero, è quello di imparare dei trucchi per comprendere i vuoti tra le figure. Ciò nonostante a livello di V1 questo aspetto non risulta essere ancora possibile, in quanto ci troviamo ancora nella early vision.

In V2 invece, troviamo alcuni neuroni che sono in grado di fare una distinzione molto più particolareggiata, che permette di creare una risposta selettiva per A rispetto a B.

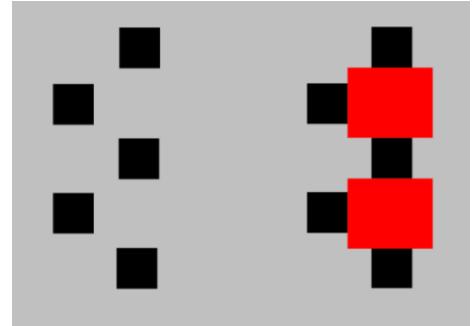


SENSATION & PERCEPTION 4e, Figure 4.5  
© 2015 Sinauer Associates, Inc.

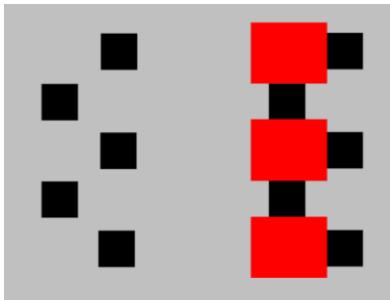
Queste cellule, in V2, hanno una selettività non solo per l'orientazione, ma anche per la proprietà del bordo. A questo livello quindi, ci troviamo di fronte qualcosa di più complesso per l'elaborazione. Ad esempio alcune cellule di V2 possono rispondere maggiormente ad (a) rispetto che (b) e viceversa.

Per comprendere il **border ownership** utilizziamo una dimostrazione, utile per cercare di semplificare cosa si intende effettivamente per border ownership.

Ci sono una serie di elementi che dal punto di vista geometrico rimangono sempre gli stessi: 5 elementi rappresentati da 5 quadrati neri su sfondo grigio.



Aggiungendo degli stimoli non aumenta il numero di oggetti da vedere ma al contrario, si riduce. Questo aspetto può essere spiegato dal fatto che, aggiungendo oggetti, si attivano dei processi relativi alla decisione di chi è proprietario del bordo e quindi anche di unificazione e segregazione tra figura e sfondo.

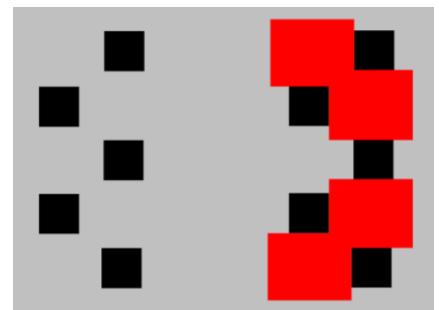


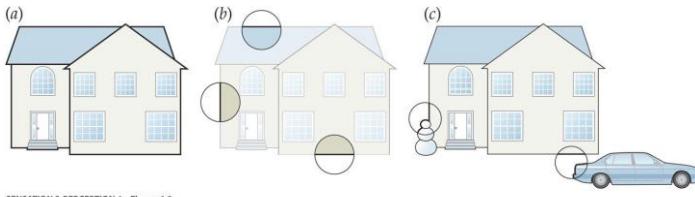
Nell'immagine si vede che aggiungendo due quadrati rossi, succede che uno dei bordi o due dei bordi dei quadrati neri non rappresentano più i bordi del quadrato. (Per esempio, uno dei quadrati neri a sinistra è proprietario di tutti e 4 i bordi, a destra invece, è proprietario solo di 3 dei segmenti il quarto invece non può essere associato ad un bordo nero).

In questo esempio, osserviamo che cambia la classificazione di chi è proprietario del contorno. Tutti i quadrati neri si unificano diventando una figura unica, in cui un unico oggetto nero è sovrapposto da figure rosse.

Cambiando la posizione dei quadrati rossi e aggiungendone uno, posso notare nuovamente gli stessi 5 elementi neri.

In questa immagine notiamo due figure rosse con forma irregolare e si nota inoltre che il numero di contorni tra il nero e il rosso aumenta sempre di più.





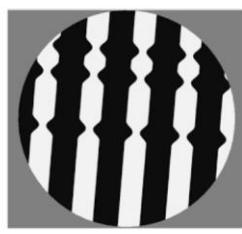
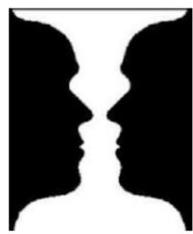
SENSATION & PERCEPTION 4e, Figure 4.9  
© 2015 Sinauer Associates, Inc.

Partendo dall'analisi di basso livello dei contorni ci chiediamo come queste informazioni sono integrate a lungo raggio. Questo è un processo presemantico, importante per comprendere come rappresentiamo il mondo.

Gli oggetti sono il risultato del processo che abbiamo nella nostra mente in cui vengono prodotte delle unità salienti che hanno una funzione adattiva dal punto di vista biologico.

## Il problema della segregazione figura sfondo

Nell'immagine si vedono tante macchie nere, alcuni di questi elementi diventano unità (il resto sfondo). Percepire il cane indica che la figura è data dall'integrazione di contorni reali ed altri virtuali. La caratteristica importante è la presenza di fattori che tendono a spingere il sistema visivo verso la segregazione figura sfondo. Inoltre, una volta che si è compreso quali sono i fattori che ci permettono di percepire uno figura, si può provare a vedere il medesimo stimolo in diverse maniere. L'ambiguità è presente quando lo stimolo è impoverito.



In queste immagini è possibile osservare le classiche situazioni che vengono chiamate **situazioni di distabilità o ambigue**, definite così perché presentano due soluzioni (o si vedono i profili neri o il calice bianco). Nello stimolo ci sono due soluzioni che hanno due evidenze. Due percezioni nelle figure sfondo. È impossibile vedere entrambe le figure nello stesso tempo: infatti o si notano i calici o i profili. La maniera spontanea di percepire le due figure avviene per mezzo

della segregazione figura sfondo (sfondo bianco due profili e viceversa). La stessa cosa succede con l'ascolto della musica polifonica, nella fuga a 4 voci il brano apparirà sempre diverso.

## Lezione 12

### Le leggi di Wertheimer

*Quali sono le regole che il sistema visivo tende a seguire quando deve elaborare lo stimolo a livello globale unendo diversi elementi?*

Questo è un tema tradizionale per la psicologia della percezione e centrale per gli psicologi della Gestalt (anni '20). La **Gestalt** è stato un approccio allo studio dei processi cognitivi molto influente per tutta la psicologia del '900; nacque in Germania come reazione al panorama psicologico di quel tempo: il *comportamentismo* nel mondo anglosassone (approccio focalizzato sull'apprendimento e scarsamente interessato alla percezione e caratterizzato da una ricerca sulle relazioni esterne tra stimoli e risposte e un divieto assoluto di ricercare leggi interne di processi o rappresentazioni non direttamente osservabili) e lo *strutturalismo* di Wundt sia in Germania che negli Stati Uniti (un approccio molto elementaristico: per studiare la percezione solevano individuare un **catalogo di sensazioni elementari**, gli elementi fondamentali da cui inizia il processo percettivo, per poi studiare i principi che combinavano questi elementi portando alla percezione). Questi approcci erano totalmente opposti alla Gestalt, nata dalla tradizione filosofica di **Husserl**, caratterizzata da un approccio anti-elementaristico e dalla ricerca di meccanismi interni alla mente per capire il funzionamento della percezione. Tra gli anni 20 e gli anni 30 in Germania la psicologia della Gestalt ha prodotto innumerevoli ricerche ma poi, a causa del nazismo, è stata spenta in quanto i maggiori esponenti (**Wertheimer**, **Koffka**, **Kohler**) persero i loro incarichi accademici poiché ebrei. Dovuti emigrare negli Stati Uniti, non riuscirono a prendere posizioni di spicco e la psicologia della Gestalt rimase sullo sfondo del panorama psicologico mondiale. Alcune eccezioni ci furono in Italia (università del nord-est: Padova, Trieste, Bologna) e in Giappone. Successivamente la *rivoluzione cognitivista* degli anni '60 ha ibridato la psicologia con la ricerca sull'intelligenza artificiale e, tornando a parlare di modelli di processi interni, vi è stata una riscoperta della Gestalt (tutt'ora presente nel mondo: "neogestaltisti").

Già negli anni '20 i gestaltisti avevano affrontato il problema della visione intermedia: *quali sono i principi a cui il sistema visivo ubbidisce per rappresentare in modo globale le informazioni provenienti dalla visione di basso livello?* Le **leggi dell'organizzazione** descrivono il funzionamento di 3 processi: chi possiede il margine, come viene fatta la segmentazione figura-sfondo, come vengono raggruppati gli elementi. Questo è ciò di cui si occupa **Max Wertheimer** nel suo famoso articolo degli anni 20'.

## **Ma qual è la tesi basilare della Gestalt (che non è solo una teoria della percezione)?**

*"The basic thesis of gestalt theory might be formulated thus:  
there are contexts in which what is happening in the whole cannot be deduced from the characteristics of the separate pieces, but conversely; what happens to a part of the whole is, in clear cut cases, determined by the laws of the inner structure of its whole."*

Max Wertheimer, Gestalt theory. (translation of lecture at the Kant Society, Berlin, 1924)

Wertheimer in una conferenza la definisce così: "Vi sono dei contesti in cui ciò che accade ad un tutto non può essere dedotto dalle caratteristiche degli elementi. Al contrario, ciò che accade a una parte del tutto è, nei casi più chiari, determinata dalle leggi della struttura interna del suo tutto."

Secondo Wertheimer esistono delle entità psicologiche ("the whole" o, in tedesco, "Gestalten" ovvero "unità strutturata") che stanno ad un livello ontologico diverso rispetto agli elementi che le compongono: dipendono da **interazioni** tra gli elementi, ma non sono riducibili agli elementi stessi; una sorta di **"proprietà emergente"**. Il classico esempio di **Von Ehrenfels** (che ha coniato il termine "gestalt") per spiegare questo tipo di entità psicologica è quello che riguarda le **melodie musicali**: una melodia non è la semplice somma delle note di cui è composta; dimostrazione di ciò è che si può avere la stessa melodia facendo una trasposizione della tonalità in cui è scritta con note completamente diverse: ciò che conta non sono le note, ma le interazioni tra esse (intervalli). Questo è un concetto rivoluzionario, per l'epoca e anche attuale: **esistono nel mondo fisico e biologico dei concetti che non sono riducibili agli elementi che compongono il tutto, anzi, sono le proprietà del tutto che incidono sugli elementi.**

Esempio: il tempo atmosferico è una proprietà globale che dipende dall'interazione di un'enorme quantità di elementi locali (temperatura, pressione, vento, altitudine ecc).

Per i gestaltisti i percetti sono proprietà emergenti dell'interazione fra più elementi: essi concepivano il cervello come un'implementazione neurale di una serie di campi di forze elettriche (campi magnetici di organizzazione che producono delle strutture). Sebbene questo concetto sia sbagliato, **l'idea che la percezione cosciente emerge come un salto qualitativo improvviso attraverso l'interazione di tanti elementi (rete neuronale) è abbastanza moderna.**

Nel lavoro del 1923 (Ricerche sulla teoria della gestalt) Wertheimer descrive **7 leggi dell'organizzazione percettiva** e poi menziona una **tendenza**, secondo lui globale, alla "**pregnanza**" (un concetto un po' vago che si riferisce alla preferenza globale del nostro sistema visivo di creare delle forme visive pregnanti ovvero semplici, complete).

Premessa: Sebbene si dica che furono coniati dagli psicologi cognitivistici negli anni 50'-60', i termini **top-down/bottom-up** per descrivere direzioni di elaborazione erano stati in primis utilizzati dai gestaltisti: **nelle prime 5 leggi si fa riferimento a fattori bottom-up**, che riguardano le proprietà dello stimolo e la cui organizzazione è guidata dall'esterno – dunque indipendente dall'esperienza del soggetto – e riflettono tendenze ad organizzare l'informazione visiva in maniera pre-semantica ed automatica, indipendente dall'attenzione; **le ultime due leggi sono top-down** e si riferiscono al set – dunque i gestaltisti riconoscevano che l'organizzazione superiore potesse essere influenzata dall'esperienza del soggetto.

Per dimostrare le leggi dell'organizzazione percettiva Wertheimer utilizzava una modalità molto semplice: disegnando su carta delle matrici di punti e manipolandone le proprietà strutturali, osservava quali strutture globali emergevano, un approccio di autosperimentazione.

Modificare la relazione di distanza tra punti ha delle conseguenze: ad esempio nell'immagine sottostante i punti tendono ad appartenersi in coppie non arbitrarie; ovvero è necessario uno sforzo attentivo per accoppiare b-c ed essendo per sua natura focalizzato, è difficile per la maggior parte delle persone vedere l'intera sequenza simultaneamente in questo modo (a/bc/de). Tuttavia è molto semplice vedere l'intera sequenza come un insieme di gruppetti ab/bc/de.

"Normalmente, questa fila viene percepita come una sequenza di gruppetti di punti ab/cd/ef..., e non come a/bc/de...  
Per la maggior parte delle persone è impossibile vedere la seconda soluzione lungo tutta la fila simultaneamente."

• • • • • • •

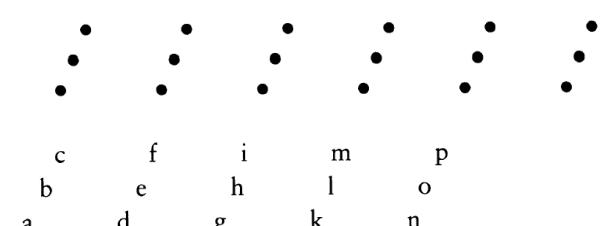
"Intendo proprio vedere, non semplicemente capire. [Qui] si impone una serie di gruppetti obliqui, inclinati da sinistra in basso a destra in alto, con la struttura ab/cd/ef... La struttura alternativa a/bc/de [...] è molto più difficile da realizzare."

Anche in questo altro esempio, sebbene ci siano tantissime strutture legittime, si tende a vedere la struttura ab/cd/ef.

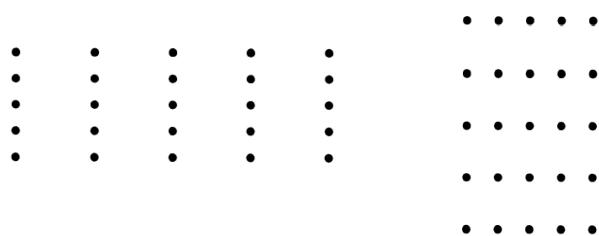
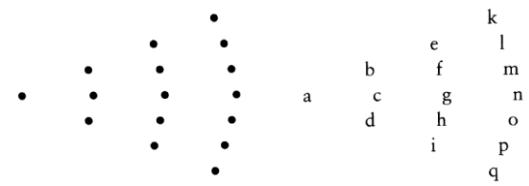
• • • • • • •

*"Inoltre [a/bc/de] è quasi impossibile da percepire con la stessa evidenza simultaneamente, per tutta la lunghezza della costellazione. Se poi sforzandoci riusciamo a ottenere tale risultato, esso si rivela assai più labile del primo, probabilmente per la difficoltà di mantenere stabile la fissazione e l'attenzione."*

*"ceg/fhk/iln [...] è praticamente impossibile da vedere simultaneamente su tutto l'insieme."*



*“...non si vede (o si coglie soltanto con difficoltà) un’altra segmentazione in via di principio adeguata, e cioè kebadiq/lfchp/mgo/n.”*



*“...[a sinistra] si vedono tipicamente delle colonne verticali e [a destra] delle righe orizzontali.”*

*“Se la costellazione contiene pochi punti, allora è più facile vedere la configurazione alternativa; ma il risultato non è più così chiaro, e tutto diventa instabile. Per esempio [...] a/bc/de/f diventa più accessibile.”*



Questi esempi sono per dimostrare il **primo principio della percezione** (“**Il principio della prossimità/vicinanza**”): “A parità di altre condizioni le unificazioni si attuano in base alla distanza minore”.

La formazione di una gestalt avviene lungo quelle direttive che comportano una minore distanza tra gli elementi.

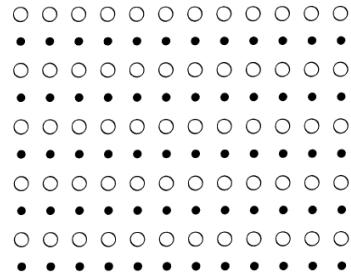
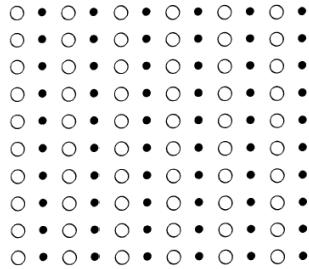
N.B. In tutte le leggi della Gestalt viene premesso “a parità di altre condizioni” poiché nell’elaborazione dello stimolo non viene considerato solo un fattore, ma tutti, ed è necessario vengano equiparati.

Altro esempio:

*“Sia data una fila di punti equidistanti, distinti a due a due per colore, su un campo omogeneo. [...] Domina la configurazione che raggruppa gli elementi uguali.”*



*“Sia data [...] , ancor meglio, una matrice bidimensionale. [A sinistra] si vedono le verticali; [a destra] le orizzontali.”*



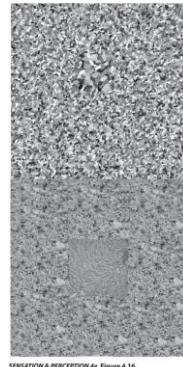
**Il secondo principio è la somiglianza:** *“A parità di altre condizioni, tendono a unificarsi gli elementi simili.”*

Wertheimer si chiese anche:

“ Che cosa accade quando [...] sono presenti due di tali fattori?”



“I due fattori possono cooperare [...] oppure agire uno contro l’altro.”



SENSATION & PERCEPTION 4e, Figure 4.16  
© 2015 Sinauer Associates, Inc.

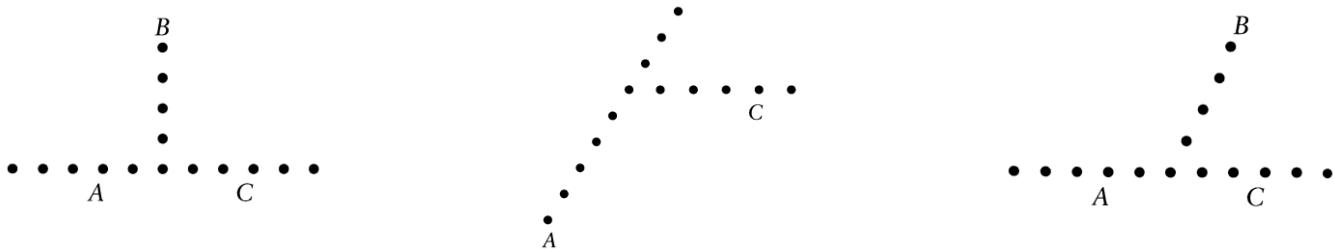
In queste situazioni ambigue si può dire che il risultato percettivo sia descritto da un *modello probabilistico*: non in tutte le condizioni tutti quanti vedono lo stesso.

**Il terzo principio è il destino comune**, ovvero la somiglianza nel modo in cui si muovono gli elementi: quelli che si muovono nella stessa direzione vengono unificati.

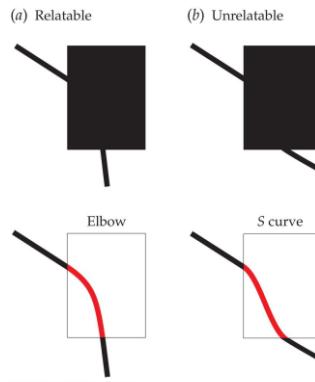
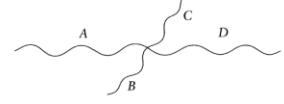
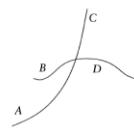
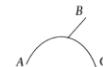
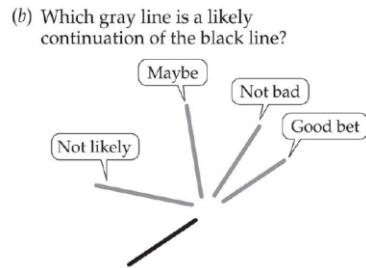
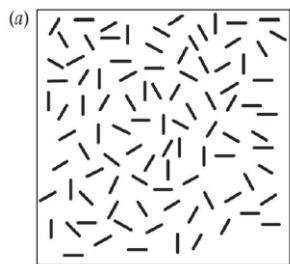


*“Cambiamento anti-strutturale, in cui vengono accomunati dal medesimo destino elementi appartenenti a gruppi statici differenti.”*

**Il quarto principio è la continuità di direzione**, ovvero la somiglianza nella curvatura locale. E' il motivo per il quale tendiamo a vedere le sequenze di punti dell'esempio sottostante non come l'incontro di 3 segmenti, ma come incontro di 2 segmenti in un punto.



**Continuità di direzione (con curve)**

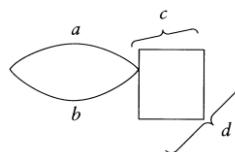


SENSATION & PERCEPTION 4e, Figure 4.31

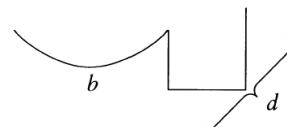
Questo principio spiega perché tendiamo a percepire segmenti parzialmente nascosti alla vista poiché occlusi da altre superfici opache.

Qui ad esempio, nell'immagine a sinistra, si tende a percepire il segmento nero dietro all'occludente più grande come se descrivesse una curva dolce poiché è la soluzione che massimizza la continuità di direzione.

**L'ultimo principio di organizzazione bottom-up è la chiusura:** a parità di altre condizioni, si tende a formare una gestalt che produca una forma chiusa piuttosto che una forma aperta (ad esempio, nel triangolo di Kanizsa, la figura a forma di "pacman" trova la sua chiusura nella formazione di un triangolo illusorio).

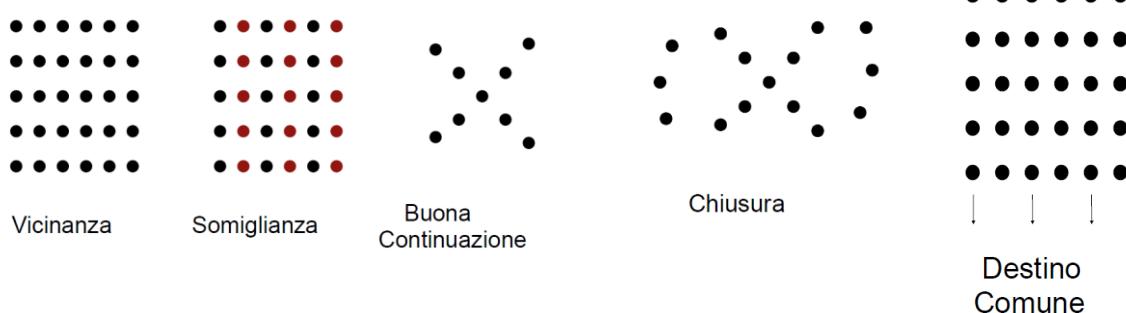


**ab/cd e non ac/bd**



Riassumendo, le leggi bottom-up (guidate dal basso) di organizzazione percettiva sono: vicinanza, somiglianza, buona continuazione, chiusura e destino comune (che pare essere il più “influente”).

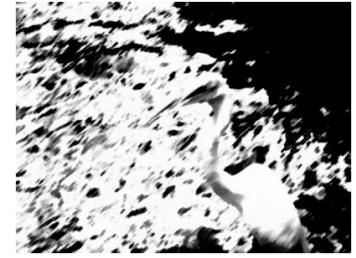
### Leggi di Wertheimer e leggi di organizzazione



Le forme di organizzazione top-down (guidate dall’alto) invece sono il **set oggettivo** e il **set soggettivo**.

Il **set oggettivo** è quel principio per il quale il sistema visivo ha una certa inerzia: raggiunta una certa organizzazione, tende a preservarla nel tempo.

Immagine con airone: rimuovendo informazioni cromatiche e sul contesto, lo stimolo viene ridotto ad una configurazione di punti bianchi e neri che è molto più difficile da organizzare. Una volta vista l’immagine, è molto più immediato ritrovarla anche nella configurazione di punti: l’organizzazione acquisita tende a mantenersi. Sebbene sia una forma di organizzazione top-down, non ha a che fare con processi cognitivi di alto livello come la memoria a lungo termine o le conoscenze dichiarative.



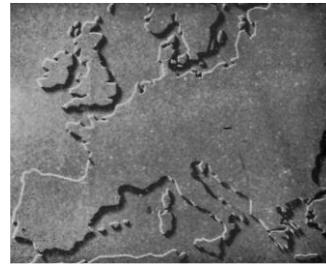
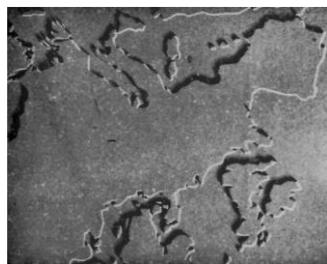
Al contrario, il **set di impostazione soggettiva** dipende dai processi cognitivi superiori: nell’immagine della mucca precedentemente mostrata, l’acquisizione dell’informazione linguistica tramite suggerimento verbale (“È una mucca”) attiva la rappresentazione interna dell’aspetto di una mucca e dunque le conoscenze dichiarative contribuiscono all’organizzazione.





Immaginando di guardare “the spinning dancer” (<https://www.youtube.com/watch?v=2RSsoTJA6cA>) dall’alto, **in che senso sembra ruotare? La maggior parte delle persone risponderà in senso orario. Ma perché vista l’ambiguità dello stimolo?** Per una “innata” preferenza. Wertheimer spiegava questa “preferenza” tramite i fattori top-down.

Immagine di carte geografiche in rilievo dell’Europa: se nella prima sono facilmente intuibili i confini europei che si creano mettendo in rilievo i territori; nella seconda è molto più difficile in quanto insolitamente sono le acque ad essere messe in rilievo. Ciò accade poiché spontaneamente interpretiamo le ombre in un certo modo: nel nostro mondo è molto più probabile che la luce venga dall’alto piuttosto che dal basso (ci siamo evoluti in un ambiente illuminato dal sole, dunque dall’alto). Di conseguenza tendiamo ad interpretare la direzione delle ombre e del rilievo alla luce di questa conoscenza pregressa. Questo è un esempio di influenza di un fattore top-down, regolato dalle nostre conoscenze sul mondo.



Sebbene possa sembrare che la percezione dipenda dalle informazioni dello stimolo (prime 5 leggi di Wertheimer), in realtà dipende anche dalle conoscenze che abbiamo sul mondo (es. Spinning dancer). I Gestaltisti spiegano questi fenomeni tramite le leggi bottom-up e top-down, invece, nel suo manuale di ottica fisiologica, **Von Helmholtz** spiega le percezioni così:

*“Le attività psichiche che ci portano a inferire che di fronte a noi, in un certo posto, c’è un certo oggetto con certe caratteristiche, generalmente non sono attività coscienti, ma inconsce. Per quel che riguarda il loro risultato, sono equivalenti a una conclusione, nel senso che l’azione osservata sui nostri sensi ci consente di formarci un’idea della possibile causa di questa azione; nonostante, di fatto, accada invariabilmente che soltanto le stimolazioni nervose siano percepite direttamente (cioè le azioni e mai gli oggetti esterni).”*

*Handbuch der physiologischen Optik  
III volume  
“Le percezioni della Visione”*

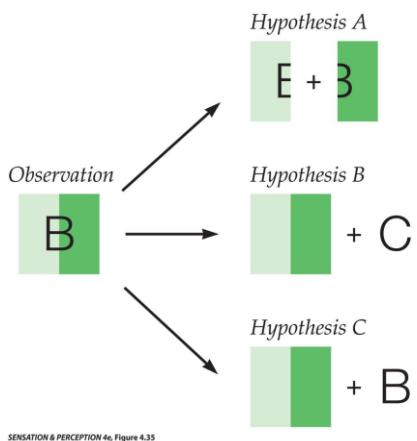
#### §26. Considerazioni generali sulla percezione

- Illusioni sensoriali 1-6
- Difficoltà nell’osservazione delle sensazioni soggettive 6-10
- Influenza dell’esperienza 10-18
- Accordo tra apparenza e realtà 18-24
- Conclusioni induttive 24-29
- Influenza dei movimenti 29
- Legge della causazione 29-35
- Nota storica 35-37

Nella percezione vi sono delle premesse e nelle premesse vi è un equivalente neurale di un'inferenza inconscia e la conclusione è il percepito.

Noi non siamo in contatto diretto con il mondo esterno, cioè l'ambiente, ma dobbiamo utilizzare l'informazione che ci arriva (in questo caso attraverso la luce) per ricostruire quello che c'è nel mondo esterno.

Secondo Von Helmholtz dunque la percezione è una sorta di **ragionamento deduttivo le cui conclusioni sono la nostra migliore ipotesi probabilistica su quale è stata la causa di quella particolare stimolazione**; nel gergo contemporaneo si parla di “*problema dell'ottica inversa*”: vi è una causa dell'informazione ottica (la luce interagisce con gli oggetti, si riflette e incontrando il nostro occhio passa la stimolazione); **ma qual è la causa della stimolazione?** E' un processo simile a quello dello scienziato: si fanno delle ipotesi ed in base a dei criteri si sceglie la migliore.



Come Von Helmholtz, anche i gestaltisti prevederebbero che il risultato fosse l'ipotesi B (in base ai 7 principi). **Qual è la relazione tra questi due approcci della percezione dunque?**

Von Helmholtz continua dicendo:

“Poichè le percezioni degli oggetti esterni hanno la natura di idee [ovvero sono parte del nostro vissuto mentale, della nostra esperienza cosciente. Non sono cose fisiche], e poiché le idee sono invariabilmente attività della nostra energia psichica [ossia sono il risultato della nostra attività mentale], le percezioni non possono che essere il risultato dell'energia psichica. Quindi, a rigore, la teoria delle percezioni appartiene al dominio della psicologia [N.B. Von Helmholtz era un medico e un fisico].. Tuttavia anche qui si apre un ampio campo di ricerca sia per la fisica che per la fisiologia, in cui si devono determinare, con metodo per quanto possibile scientifico, le specifiche proprietà dello stimolo fisico e della stimolazione fisiologica responsabili della formazione di una particolare idea relativa alla natura degli oggetti percepiti”.

*“La regola generale che determina le idee della visione, quelle che si formano ogniqualvolta gli occhi subiscono un’impressione, con o senza l’aiuto di strumenti ottici, è la seguente: vengono sempre immaginati [=percepiti] come presenti nel campo della visione degli oggetti con le proprietà che essi avrebbero dovuto possedere allo scopo di produrre le stesse impressioni sul meccanismo nervoso che si sarebbero prodotte se gli occhi fossero stati usati nelle normali condizioni ordinarie.”*

Questo significa che noi vediamo in funzione dell’informazione che viene data dallo stimolo, non in funzione di ciò che c’è veramente nella realtà. All’epoca questa era una conquista concettuale molto importante poiché da un punto di vista del realismo, di cosa è il reale e di cosa sono le illusioni.

Ed aggiunge:

*“Accade così che, quando le modalità di stimolazione degli organi di senso sono inusuali, si formino idee incorrecte sugli oggetti: idee che quindi sono state descritte come illusioni dei sensi. Naturalmente, in questi casi non c’è nulla di sbagliato nell’attività degli organi di senso e nei corrispondenti meccanismi nervosi che producono l’illusione. Tutti devono funzionare in base alle leggi che governano le loro attività una volta per tutte.”*

Von Helmholtz intende che esistono dei processi percettivi che sono automatici, adattivi (ovvero ci permettono di formarci una rappresentazione del mondo adeguata per interagire con il mondo) che guidano il comportamento in modo da permettere la riproduzione e continuare il processo di evoluzione della specie. Sebbene tali processi siano normalmente adeguati, in alcuni casi vi possono essere delle discrepanze: **le illusioni**.

Le illusioni sono interessanti perché sono delle **condizioni limite del funzionamento del nostro sistema percettivo e come tali ci danno delle preziose informazioni sulle sue caratteristiche, sulle sue modalità di funzionamento**. Questa è l’idea alternativa a quella dei gestaltisti (secondo la quale la percezione è il risultato dell’interazione di forze autonome-descritte nei principi della gestalt) poiché vede la percezione come un’implementazione neurale in possesso di inferenza inconscia.

In realtà l’approccio dei gestaltisti e quello di Von Helmholtz ai principi dell’organizzazione percettiva non sono poi tanto diversi: sono entrambi basati sull’idea che a questo livello la percezione sia una costruzione del cervello e che il risultato finale (il percepito?) dipende da come trattate determinate premesse (sebbene per i gestaltisti siano le interazioni tra gli elementi del campo visivo e per Helholtz siano dei vincoli ad un ragionamento probabilistico).

N.B. Molti modelli contemporanei della percezione sono una sorta di compromesso tra queste due cose: cercano di capire come fare un modello formale che preveda la probabilità di un certo risultato percettivo, in funzione dell'equilibrio di queste due cose in un contesto di tipo probabilistico, si parla di “modelli bayesiani” della percezione.



**Bayes** fu un teorico della probabilità del 1700 che mostrò come calcolare una probabilità – detta probabilità a posteriori – a partire da delle conoscenze pregresse e sulla base di nuovi dati. La cosiddetta regola di Bayes sul calcolo della probabilità dice che se si hanno delle conoscenze a priori – ad es. Il 60% degli studenti di psicologia sono donne – e sulla base di nuovi dati – prendo a caso uno studente e scopro che è di psicologia – posso ridurre il livello di incertezza a posteriori su un'altra probabilità – ad es. la probabilità che quel particolare studente sia una donna. Dunque è una forma di ragionamento probabilistico in cui vengono combinate informazioni già conosciute e nuovamente acquisite e sulla base di quest'ultime aggiornate le credenze sul mondo. Questo teorema applicato alla percezione significa che il percepito è il livello di conoscenza sul mondo attuale che è risultato della combinazione di conoscenze già apprese (es. la luce proviene dall'alto) e nuove informazioni acquisite al momento (es. una certa configurazione di ombre nello stimolo).

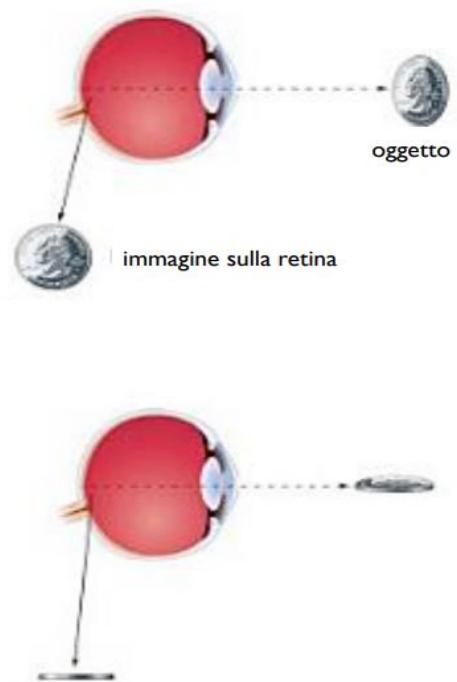
## Lezione 13

### Il punto di vista generico

Tutto quello che abbiamo detto fino ad ora sui principi di organizzazione, le leggi della Gestalt, il problema di interpretazione probabilistica delle leggi della Gestalt, ecc. si può anche pensare da un altro punto di vista, più elegante, più semplice, che si riassume in un'unica idea: **il vero principio generale che determina il fatto che noi andiamo ad unificare in base alle leggi della Gestalt o a un principio probabilistico (scegliere la soluzione che sembra più probabile) è che il nostro sistema visivo (ma non solo il nostro) è costruito in modo da lavorare in base ad un'assunzione molto generale che non riguarda la natura del mondo, ma la natura della maniera con cui interagiamo col mondo, cioè il tipo di posizione che possiamo assumere nel mondo quando guardiamo gli oggetti.** In altre parole: *qual è la statistica (la distribuzione di probabilità) dei punti di vista che noi tendiamo ad occupare quando campioniamo la luce con gli occhi.*

L'idea è che il cervello tende ad assumere di trovarsi in un **punto di vista generico**, non speciale, che non ha niente di particolare (questo punto di vista ha quindi una **alta probabilità di essere occupato**) e tende a rifiutare l'idea che si trovi in un punto di vista inusuale (speciale, improbabile). Per capire meglio il concetto di punto di vista generico (anche detto *non accidentale*) si provi a mettere una moneta davanti a un occhio e si ruoti fino a che il lato non diventa parallelo alla linea di mira. Sulla retina all'inizio si proietta un cerchio, ruotandola si proiettano ellissi sempre più schiacciate: tutte queste proiezioni contengono margini curvi. Statisticamente in tutti questi punti di vista c'è qualcosa che non varia: la presenza di margini curvi. Solamente quando si arriva alla situazione estrema (la singolarità), cambia qualcosa: i margini curvi scompaiono. Questa è la distinzione tra i punti di vista normali/generici e un punto di vista accidentale/inusuale/non generico (con probabilità che tende a zero e che si realizza soltanto usando un occhio solo, tenendo la testa perfettamente ferma e trovando una particolare inclinazione della moneta).

Possiamo anche aggiungere che la curvatura è praticamente sempre presente sulla retina (la probabilità tende a 1).





Questo principio si applica a qualsiasi cosa e ad altre proprietà degli stimoli. Questo oggetto proietta sulla retina una certa figurazione di contorni, cambiando il punto di vista si ottengono distorsioni (i cerchi diventano ellissi, ecc.), però rimangono disponibili le proprietà di connessione dei contorni (proprietà non accidentali).

Il cervello ricava l'informazione sulla base della proiezione sulla retina e per passare dall'informazione retinica al riconoscimento della bicicletta assume un punto di vista generico.



Quindi se nello stimolo ci sono margini connessi il ragionamento è il seguente: “*dato che sono in un punto di vista generico, non accidentale, se i margini sono connessi nello stimolo allora saranno connessi anche nel mondo*”. Possiamo quindi prevedere che se genero uno stimolo in cui i margini sono connessi, noi tenderemo a vederlo come un oggetto tridimensionale in cui i margini sono connessi, ma in realtà può essere un’illusione come in questo caso:

Questo è un principio di natura probabilistica ma la probabilità non è basata su premesse o conoscenze che ho imparato a scuola, **è invece basata sui punti di vista che io assumo quando mi trovo nel mondo a interagire con gli oggetti. Non è una regolarità della fisica ma appartenente alla nostra ecologia (della maniera in cui ci muoviamo e campioniamo la luce quando vediamo).**



Possiamo quindi ora capire come funziona *l’illusione della sedia riducente* precedentemente vista: troviamo il punto di vista dove tutti i contorni vanno a registro (si incontrano nei punti giusti), in maniera da generare l’immagine bidimensionale che è la stessa che ci sarebbe se ci fosse una sedia vera.

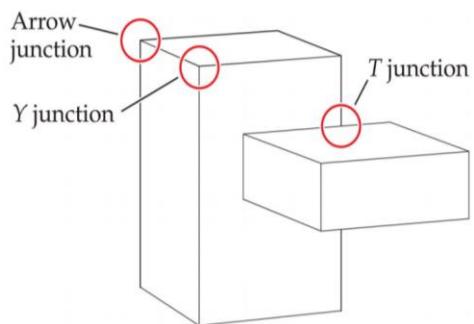
Uno psicologo gestaltista direbbe: “*Questo è il manifestarsi dei principi di organizzazione per prossimità, buona continuazione, chiusura e somiglianza. I segmenti vengono organizzati insieme e diventano un oggetto*”.

Von Helmholtz direbbe: “*Noi vediamo la sedia perché è la soluzione più probabile*”.

Queste due affermazioni si possono mettere insieme dicendo: “*Noi siamo costruiti in maniera tale da avere introiettato questa regolarità statistica dei punti di vista che tendiamo ad occupare. Interpretiamo questo secondo il principio che se sono connessi nell’immagine sono connessi anche nel mondo, perché altrimenti vorrebbe dire che occupo un punto di vista “strano” che ha una probabilità di avvenire talmente piccola da non essere economica la sua assunzione*”.

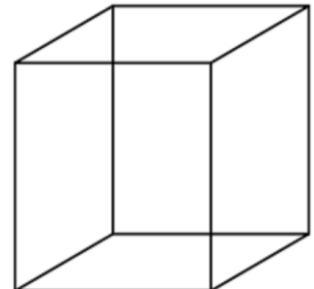
**Questa è la base di tantissime illusioni: creare un punto di vista che è accidentale** (es: lo spioncino della sedia di Ames serve a creare un punto di vista accidentale).

Un programma di visione artificiale, quando deve cercare di ricostruire la struttura tridimensionale di un oggetto a partire dai contorni, va a vedere subito come sono fatte le giunzioni tra i contorni.



Per gli oggetti artificiali questi contorni hanno delle caratteristiche (a freccia, a Y, a T) che corrispondono a particolari relazioni tra le superfici: a meno di non essere in un punto di vista strano, una giunzione a T corrisponde a una superficie che sta davanti e una che sta dietro, le Y e le giunzioni a freccia corrispondono allo spigolo di un oggetto che punta verso di noi o nell’altra direzione.

Questo è il **cubo di Necker** (una figura *bistabile*). Si può vedere la tridimensionalità in due maniere diverse: i due spigoli in basso sono appartenenti alla faccia in primo piano, oppure la faccia in primo piano è quella coi due spigoli più in alto. Questa bistabilità del cubo di Necker in realtà non è così forte come si descrive nei libri, infatti, di solito quando si mostra l’immagine per la prima volta ad osservatori *naïve*, questi dicono di vederlo con la faccia di sinistra in avanti. Nonostante dal punto di vista proiettivo siano equivalenti, una soluzione è più facile dell’altra, questo presumibilmente perché normalmente vediamo oggetti dall’alto verso il basso, appoggiati su qualcosa (prima soluzione), piuttosto che dal basso verso l’alto, sospesi (seconda soluzione, meno probabile).



Stesso motivo per cui si tende a vedere la ballerina (vista nella lezione precedente) ruotare in senso orario; la preferenza è per un punto di vista più probabile rispetto ad uno un po' meno probabile.

Nell'ambito della *psicologia dell'arte* si individuano certe tipologie di arte figurativa basati su questo



gioco sul punto di vista dell'osservatore. Uno di questi giochi è l'**anamorfosi** in cui si creano immagini che sembrano completamente estranee al resto del quadro, ma se lo si osserva da un punto di vista particolare. In questo esempio, se si osserva il quadro da un punto di vista molto laterale, l'ellissi diventa quasi circolare e sulla retina viene proiettato un teschio (il cervello si comporta come se non fosse in un punto di vista non generico e il teschio compare).

Ci sono molti artisti di strada che giocano con questo fenomeno.



PISCINA 3D di Julian Beever (foto presa dal web)



PISCINA 3D vista dal "punto sbagliato" - di Julian Beever (foto web)

In Italia ci sono moltissimi esempi nell'architettura e nelle decorazioni delle chiese:

**Andrea da Pozzo (1685). Gloria di S. Ignazio, affresco sul soffitto della chiesa di S. Ignazio a Roma**

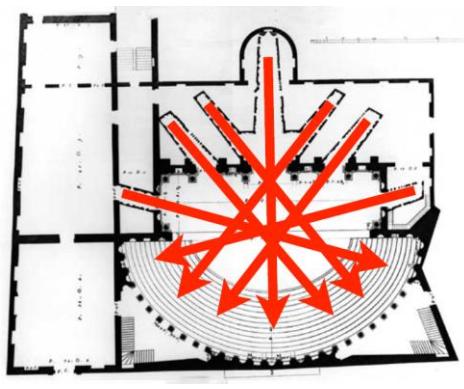
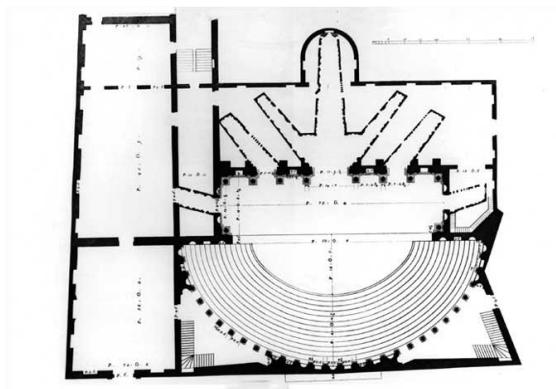
Questo gioco è chiamato “*sfondato*”, per avere l'impressione che il tetto sia aperto bisogna camminare all'interno della chiesa fino ad un punto segnato sul pavimento (funziona perché il cervello non rifiuta l'idea di essere in un punto di vista generico).



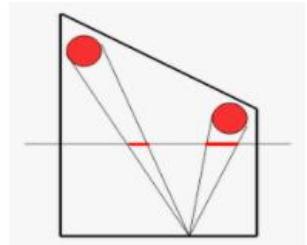
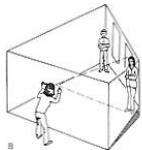


**Vincenzo Scamozzi (1584). Scene fisse per il palcoscenico del Teatro Olimpico a Vicenza.**

L'effetto funziona perché in platea le persone sono disposte a semicerchio e occupano più o meno dei punti di vista speciali. Guardando la piantina si vedono come sono i corridoi dove sono costruite in prospettiva inversa le case che danno l'impressione che la strada continui per molte decine di metri. A seconda di dove ci si trova in platea si vede principalmente un corridoio dal punto di vista giusto per produrre l'illusione (se una persona percorre il corridoio si rimpicciolisce o si ingrandisce).



Nella **stanza di Ames** c'è uno spioncino che è un *punto di vista non generico* dal quale sembra di vedere una stanza normale, ma a seconda di dove si posizionano le persone dentro alla stanza sembrano nani o giganti e se si spostano nella stanza si rimpiccioliscono o ingrandiscono (la stanza è distorta: ha una pianta trapezoidale, il pavimento e il soffitto sono inclinati, le pareti sono trapezi. **Il cervello assume di essere in un punto di vista generico e i trapezi vengono interpretati come rettangoli; la persona che in realtà è più vicina, viene percepita alla stessa distanza di quella più lontana e sembra un gigante).**







STABILIZED IMAGES typically fade as in the illustrations on this and the following two pages. The parts of a profile drawing that stay visible are invariably specific features or groups of features, such as the front of the face or the top of the head.



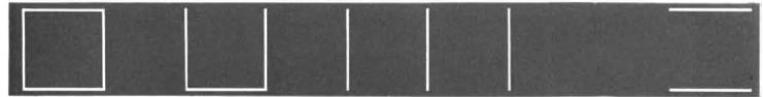
MEANINGLESS CURLICUES first come and go in random sequence. But after a while small groups of curlicues organized in recognizable patterns start to behave as units. This suggests that they have themselves become meaningful perceptual elements.



MONOGRAM formed of the letters *H* and *B* also seems to illustrate the importance of elements that are meaningful because of past experience. When the monogram breaks up it is the recognizable letters and numbers within it that come successively into view.



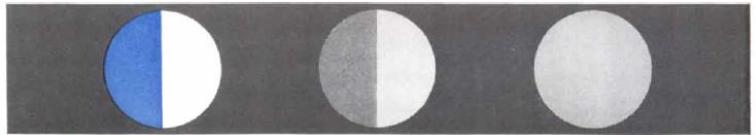
WORDS containing other words behave in much the same manner as the monogram. Here, for example, the subject sees new words made up of letters and parts of letters in the original. He is far less likely to report seeing meaningless groups of letters such as *EER*.



LINES act independently in stabilized vision, with breakage in the fading figure always at an intersection of lines. Adjacent or parallel lines may operate as units. This independent action of lines tends to support the cell-assembly theory of perception.



PLANES operate as units in three-dimensional figures. In this Necker cube (which gives an illusion of reversing in stabilized vision as well as in normal vision) a line may act alone. But usually lines defining a plane operate together, leaving parallel planes.



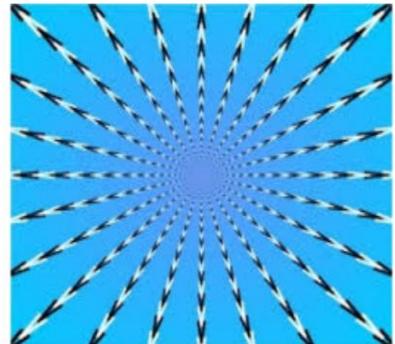
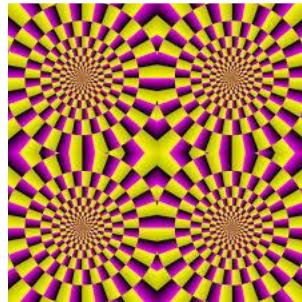
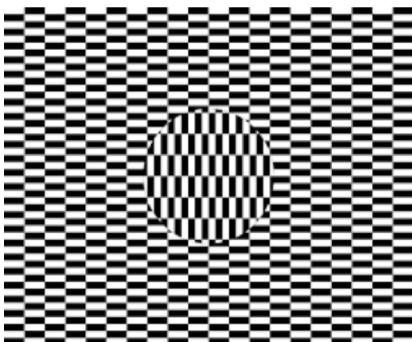
SENSE OF COLOR is lost with particular speed. A two-color field like this fades almost immediately when stabilized, to leave two values of gray; then the brightness difference disappears. The stabilized technique promises to be useful for studying color vision.

Se si stabilizza uno stimolo semplice di un contorno che da una parte ha un colore (*discontinuità orientata*), succede che prima non si vede più il colore poi il contorno svanisce e uno dei due colori invade l'altro.

**Un'immagine stabilizzata è un'immagine che svanisce**, infatti l'uomo non è equipaggiato per vedere immagini stabili, è come se per vedere il mondo stabile si dovesse elaborare stimoli continuamente mutevoli (in movimento sulla retina).

Per percepire un movimento non è neanche necessario che ci sia un movimento nel mondo (c'è una mancanza di correlazione): posso avere situazioni dove non c'è nulla che si muove ma vedo il movimento, oppure anche il contrario, una cosa che si muove ma la vedo stabile.

Ci sono molte illusioni a riguardo:

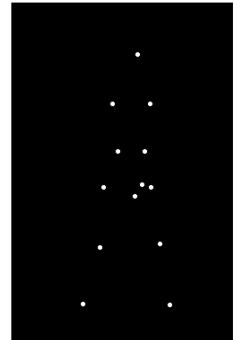


Non possiamo dire che vediamo il movimento perché c'è qualcosa in movimento, ma evidentemente il movimento è il risultato di una computazione neurale.

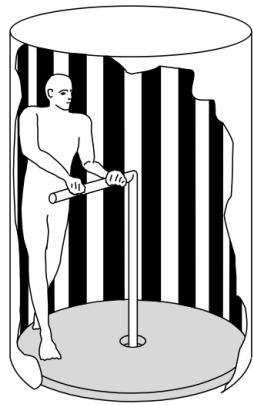
## Funzioni della percezione del movimento

La codifica del movimento da parte del sistema percettivo (prima basso livello poi livelli successivi) è importante non solo per dare inizio al processo di percezione (se non c'è movimento le immagini scompaiono), ma anche da altri punti di vista:

- **La legge del destino comune:** il movimento comune di elementi rappresenta il fattore più potente in assoluto per produrre **unificazione e segregazione** (fare emergere un'unità percettiva coesa);
- Il movimento contiene informazioni fondamentali per la percezione della **tridimensionalità** (della struttura tridimensionale degli oggetti e delle distanze relative nel mondo). Quindi è un discorso che riguarda la percezione della forma e dello spazio, in immagine **M. Duchamp**, Rotorelief;
- Noi tendiamo spontaneamente, in certe condizioni, ad interpretare i movimenti anche come uno **stimolo sociale** non solo traiettorie e dunque a donare **qualità espressive** al movimento. Se si prendono forme senza senso e si fanno muovere, le persone che devono descrivere questi eventi tendono ad attribuire intenzioni ed emozioni come se fossero organismi guidati da scopi ed emozioni. Su questi ragionamenti viene ricondotta la comprensione delle azioni altrui, la *cognizione incarnata*;
- **Il movimento biologico.** Stimoli in cui si cerca di catturare gli aspetti cinematici del movimento di un organismo togliendo tutti gli aspetti legati ai contorni e alle forme ma preservando i movimenti relativi. Queste informazioni sono utili non solo per capire le traiettorie, ma anche le conseguenze di interazioni tra forze (la dinamica di un movimento, non solo la cinematica), in immagine dei punti apparentemente casuali che una volta in movimento fanno chiaramente riconoscere una persona che cammina [http://www.lifesci.sussex.ac.uk/home/George\\_Mather/Motion/](http://www.lifesci.sussex.ac.uk/home/George_Mather/Motion/);
- **Multisensorialità:** le informazioni sul movimento sono molto ricche di dati che aiutano a ricostruire lo spazio e questo è importante per il controllo della postura e della locomozione (forse più importante dell'informazione fornita dal sistema vestibolare). In apparati sperimentali si dà uno stimolo in movimento che il cervello interpreta come un cambiamento di postura rispetto all'ambiente e si genera un conflitto tra l'informazione visiva per la postura e quella vestibolare e quasi sempre prevale l'informazione visiva. Vediamo due esempi:



## Postura e locomozione

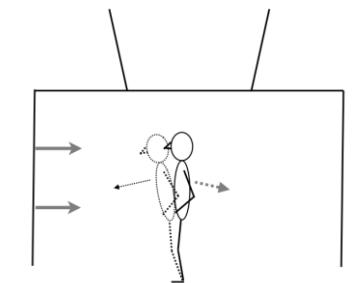


Vediamo il *tamburo ottico-cinetico* ovvero una stanza cilindrica con pareti mobili.

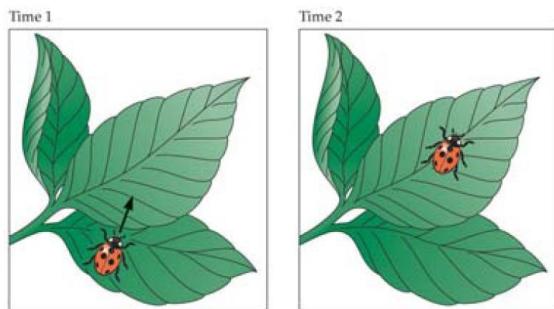
Quando le pareti si muovono, sulla retina viene percepito un lento movimento, tipicamente non si vede la stanza ruotare ma si percepisce di essere in movimento nella stanza ferma.

Poi la *stanza mobile*, ovvero una stanza quadrata con le pareti leggermente sospese rispetto al pavimento. Il compito del soggetto è di stare fermo in mezzo alla stanza; muovendo leggermente la stanza (parete che si avvicina), il

partecipante ha l'impressione di sbilanciarsi in avanti e si attiva un riflesso di controllo della postura che tira il corpo indietro, ma siccome il corpo non era sbilanciato si perde l'equilibrio.

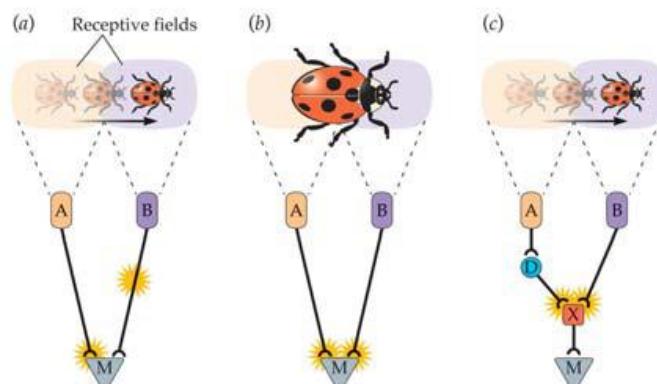


## Codifica del movimento: informazione locale

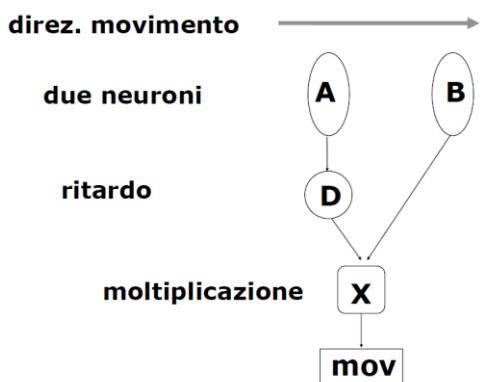


A livello retinico viene registrato quello che c'è e da un punto di vista fisico un movimento **è un cambio di posizione rispetto al tempo**.

I recettori retinici non possono registrare il movimento perché l'animale muovendosi produce un effetto prima su una certa zona della retina e poi su un'altra zona avendo quindi un ritardo tra queste due situazioni, ed un ipotetico neurone che dovrebbe codificare questo movimento verrebbe stimolato in tempi diversi (potrebbero essere due oggetti separati o lo stesso oggetto in un momento successivo).

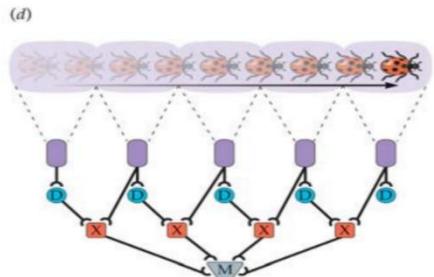


Inoltre un **oggetto grande** potrebbe occupare entrambi i campi recettivi e produrre paradossalmente una situazione in cui i recettori di movimento vengono stimolati simultaneamente.



Quindi per avere un'unità neurale che sia in grado di codificare le caratteristiche di un movimento che viene registrato dalla retina, occorre una *circuiteria neurale* più complessa, in cui l'attivazione di due zone della retina arriva a un recettore di movimento (M) passando prima per l'unità X che ha lo scopo di combinare insieme gli input retinici ma uno dei due deve passare prima per un'ulteriore stazione che ha il compito di produrre un ritardo (D=delay) del segnale neurale.

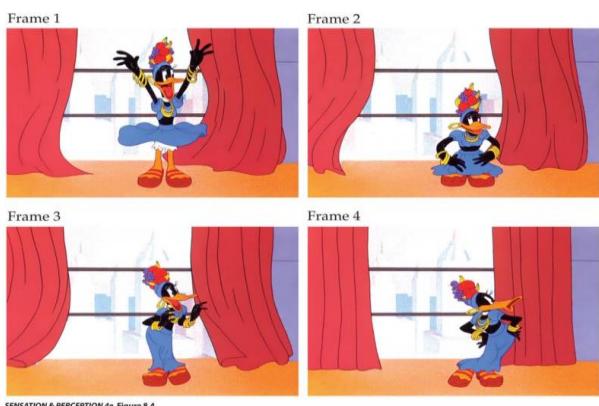
Questa struttura di base viene chiamata “**recettore di Reichardt**”. Le unità M sono in grado di codificare la direzione e la velocità dello stimolo in movimento (codifica del movimento retinico). Se il movimento è da sinistra verso destra, il segnale di A viene ritardato da D e raggiunge X allo stesso tempo del segnale di B facendo raggiungere la soglia di X per sparare (*integrazione*). Se il movimento è da destra verso sinistra i segnali arriveranno su X separati nel tempo non permettendogli di raggiungere la soglia per sparare: questo recettore ha quindi una **preferenza (selettività) per movimenti in una certa direzione**. La selettività è anche per una certa velocità, perché lo stimolo ottimale è quello in cui il ritardo imposto da D compensa esattamente il tempo che lo stimolo ci ha messo per spostarsi da A a B (le due informazioni arrivano esattamente insieme generando il massimo di segnale per M). Modulando l'entità del ritardo si produce una selettività per la velocità: se il ritardo è piccolo sarà più selettivo per un movimento veloce e viceversa.



Questo principio si può applicare ad una stringa di fotorecettori retinici ognuno col suo ritardo; l'insieme di questi campi recettivi definisce il campo recettivo di questo recettore di Reichardt, selettivo per un movimento verso destra con una velocità che dipende da D.

Per il principio *dell'adattamento* (già visto per la frequenza spaziale e orientazione), dovrebbe essere possibile adattare questi rilevatori “affaticandoli” e producendo una perdita di sensibilità per la direzione di un movimento a una certa velocità e quindi osservare un *aftereffect* (effetto postumo). Quando smetto di guardare uno stimolo in movimento e guardo un oggetto stazionario, avendo ridotto la sensibilità per una specifica direzione, lo stimolo fermo dovrebbe sembrare muoversi nella direzione opposta (*effetto cascata*). Questa impressione di movimento è del tutto svincolata dalla percezione della forma.

Per la struttura del recettore di Reichardt dovrebbe essere possibile produrre l'impressione di un movimento convincente mostrando una sequenza di immagini statiche, perché non è importante che ci sia un movimento reale da A a B, ma che ci sia l'attivazione di A che viene ritardata seguita dall'attivazione di B; quando il tempo che è passato fra l'attivazione di A e quella di B matcha il ritardo, il rilevatore di Reichardt viene stimolato e si vede il movimento (movimento stroboscopico, quello che succede al cinema e nei cartoni animati).



Questa non è un'illusione ma semplicemente per il rilevatore di Reichardt non è distinguibile da un movimento, produce esattamente lo stesso effetto.

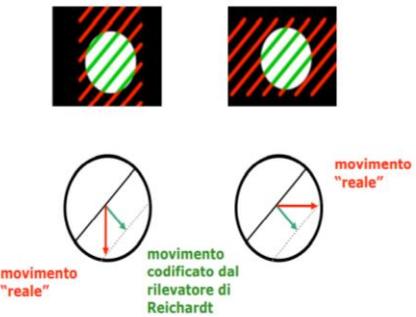
Ricapitoliamo le caratteristiche del rilevatore di Reichardt:

1. **Selettivo per la direzione;**
2. **Selettivo per la velocità;**
3. **Prevede aftereffect di movimento;**
4. **Prevede il movimento stroboscopico.**

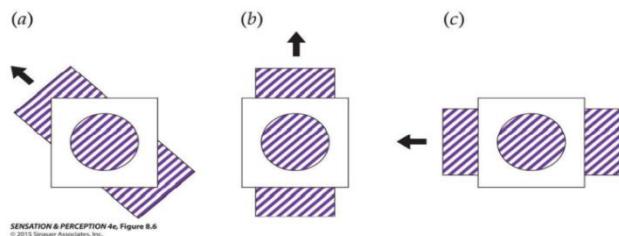
## Il problema dell'apertura

Queste unità neurali hanno un campo recettivo che è limitato nello spazio (campo recettivo *localizzato*, che può dare solo informazioni locali). Il problema è di natura computazionale (non neurale), è un limite intrinseco di una qualsiasi architettura (artificiale o biologica) che deve cercare di **ricostruire un movimento a partire da informazioni che sono localizzate nello spazio** (acquisisce informazioni solamente da una parte molto piccola localizzata nello spazio di un contorno che si muove e quello che sta al di fuori di questa zona localizzata semplicemente non c'è). Questo problema vale anche per la situazione opposta all'apertura: un contorno che si estende infinitamente nel campo visivo e in qualsiasi punto lo guardi l'informazione è sempre la stessa.

All'interno di un ipotetico campo recettivo di un recettore di Reichardt (la zona bianca) i segmenti cambiano nello spazio-tempo in maniera esattamente uguale (l'informazione è la stessa) nonostante i movimenti reali siano uno verso il basso e l'altro verso destra. La stessa osservazione varrebbe anche se i reticolari fossero infinitamente lunghi, perché non avremmo niente che si muove verso il basso o verso destra ma tutto si muove esattamente nella stessa maniera.



**Il recettore di Reichardt è capace di codificare solo il movimento di un contorno nella sua componente ortogonale (perpendicolare) all'orientazione del contorno.** Non è capace di distinguere tra una serie di movimenti reali in direzioni diverse ma che hanno uguale componente ortogonale al contorno.



Il recettore di Reichardt rileva solo un segnale che è la somma di due (o più di due) unità, una delle quali è stata ritardata. L'informazione locale per il movimento è un'informazione in cui il recettore codifica la direzione e la velocità di un contorno nella direzione ortogonale alla sua orientazione e quindi è ambiguo.

Noi vediamo i movimenti globali degli oggetti accuratamente risolvendo il problema dell'apertura legato all'informazione locale, mettendo insieme più informazioni locali in stadi di elaborazione successivi (come abbiamo visto per le forme passando dalla visione di basso livello a quella di livello intermedio).

## Lezione 14

Nella scorsa lezione si è introdotto il tema della *percezione del movimento*, più in particolare il problema dell'apertura. **Il punto fondamentale è capire come il sistema visivo intermedio risolva il problema dell'apertura, a partire da quello che viene codificato dai singoli recettori di Richard.** La struttura logica di questa argomentazione non è molto dissimile dalla percezione dei contorni e delle forme, ovvero **ci sono dei limiti intrinseci nell'informazione contenuta nei singoli segnali locali**, però il limite viene superato grazie ad un'operazione di **natura integrativa**: mettendo insieme, organizzando, combinando, questi segnali locali. Questa attività integrativa è tutt'altro che banale. Il professore ed un suo collega hanno scritto un breve articolo di rassegna sull'argomento (come i processi di organizzazione risolvano il problema dell'apertura).

**Perceptual Organization and the Aperture Problem** [a](#)  
Nicola Bruno and Marco Bertamini  
The Oxford Handbook of Perceptual Organisation (Forthcoming)  
Edited by John Wagemans  
  
Online Publication Date: Apr 2015  
Subject: Psychology, Neuropsychology  
DOI: 10.1093/oxfordhb/9780199668880.013.046

### [\[–\] Abstract and Keywords](#)

We discuss how the visual system can achieve coherent object motion, overcoming the aperture problem and the edge classification problem. The solution cannot be reduced to relatively simple vector operations, such as averaging or intersecting constraints in velocity space, but is instead a complex form of perceptual organization, which takes into account the spatial structure of the stimulus.

Keywords: motion perception, aperture problem, perceptual organization

### **Cos'è il problema dell'apertura?**



Bisogna immaginare uno schermo nero con al centro un buco circolare, attraverso il quale è possibile vedere un reticolo che si muove.

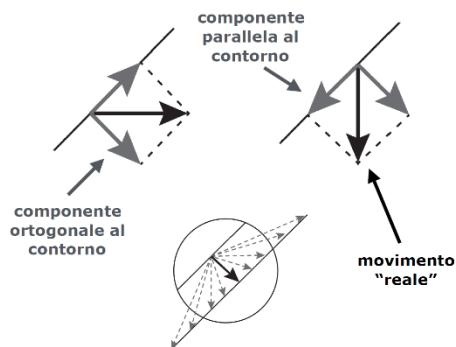


Per rilevare il movimento fisico del reticolo, nell'animazione vengono anche mostrati i pezzi del reticolo che sarebbero nascosti dietro allo schermo nero (il rosso è dietro), il verde lo vedo.

Si può intanto notare che una cosa è il **movimento reale** del reticolo, (i vettori hanno due posizioni ortogonali a 90 gradi l'una rispetto all'altra), nel movimento reale però se ci guardiamo con attenzione analitica (attenzione solo a come si muovono soltanto i segmenti verdi), ci rendiamo conto che ciò che succede all'interno del buco bianco è **esattamente uguale nelle due animazioni**, ovvero che all'interno di quel buco si può notare un movimento, in diagonale, da destra verso sinistra - dall'alto verso il basso.

Ora, questo movimento, è quello che verrebbe visto da una persona in queste condizioni. **Che cosa riescono a codificare di quel movimento i recettori di R?** Questa componente del vero movimento è l'unica che i recettori di R possono codificare in queste condizioni, in quanto il recettore di R non è in grado di ricostruire, di registrare il movimento reale di un contorno che passa attraverso il suo campo recettivo.

Analogia tra il campo recettivo del recettore (zona in cui il recettore è sensibile) e quello che si vedrebbe attraverso il buco. Questa analogia è giusta fino ad un certo punto, perché quando si guarda il buco attraverso lo schermo il risultato di quello che noi vediamo non è soltanto l'output del recettore di R, **ma è anche il risultato di passaggi successivi**, però questo illustra bene il fatto che noi non saremmo in grado di distinguere tra questi due reticolari (i due movimenti dei reticolari), se li guardassimo attraverso un buco. Uno dei motivi di ciò è che l'*informazione sarebbe solo quella locale* che verrebbe rilevata a basso livello, ovvero il **movimento ortogonale alla direzione del reticolo**. Più esattamente possiamo dire che comunque noi vediamo il movimento, qualsiasi sia il sistema, se noi abbiamo un contorno che è privo di ulteriori discontinuità (che sia un campo recettivo limitato o illimitato), **questo contorno è sempre uguale**, cioè un segmento senza ulteriore caratteristica che possa essere rivelata attraverso un qualche meccanismo visivo, il contorno è sempre omogeneo, un segmento sempre uguale. Allora il movimento di questo contorno si scomponerebbe sempre per la sua cinematica in due componenti.



La freccia nera indica la direzione reale del movimento, che ha due componenti:

- una **ortogonale** al contorno;
- l'altra che è **parallela** al contorno, e l'insieme delle due componenti fa sì che il contorno si sposti in orizzontale verso destra.

Si vede che, il problema fondamentale del “problema dell’apertura” è che il nostro sistema visivo è sensibile all’orientazione ed è sensibile al cambiamento di posizione rispetto al tempo di quel contorno, però la componente parallela al contorno, il nostro sistema visivo non ha modo di rilevarla, poiché quel movimento è nella direzione del contorno stesso, quindi non c’è nessun cambiamento nello spazio-tempo in quella direzione, il cambiamento avviene solo nella direzione ortogonale al contorno, questo è il vero motivo per cui vale il problema dell’apertura: **è un problema di natura proiettiva, che ha a che fare con quali sono i cambiamenti spazio-temporali che si danno in uno stimolo quando c’è un contorno omogeneo che si muove.**

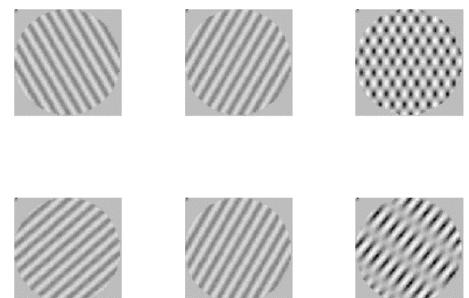
Questo è il motivo per cui il nostro sistema visivo in queste condizioni non sarà in grado di distinguere tra il movimento orizzontale e quello verticale, **perché la componente parallela l’ha persa** e rimane solo la componente ortogonale al contorno stesso.

Il problema dell'apertura è quella cosa per cui i *recettori locali di movimento sono in grado di rilevare solo le componenti ortogonali del movimento e non il movimento reale*, quindi c'è una forma di **ambiguità** (figura in basso al centro) tra tutta una serie possibile di movimenti reali che qui sono indicati in grigio con le tratteggiature. Tutti hanno la componente ortogonale, le componenti parallele vanno perse come se non ci fossero, sono tutti quindi movimenti **equivalenti** dal punto di vista della ricezione dei recettori di R.

### **Come viene superata questa fondamentale ambiguità?**

Le componenti locali di movimento di ogni singolo contorno possono venir combinate con un modello molto semplice: la combinazione lineare, il che significa **somma**, o **media**.

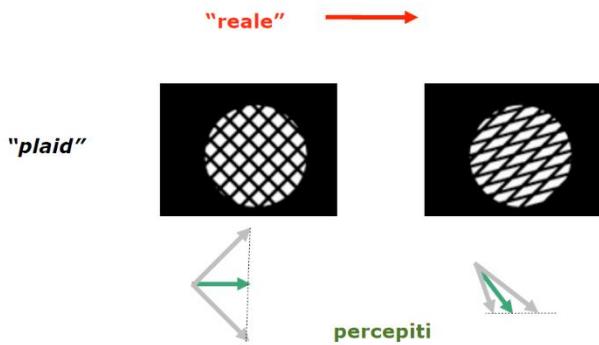
Ciò è stato studiato molto a livello psicofisico, utilizzando un tipo di stimolo in movimento che non è un singolo grating che si muove, bensì più grating. Si fa la somma dei reticolati, dove c'è una singola componente con una certa frequenza spaziale, un certo contrasto, orientazione e possono essere sommati, **generando così un segnale più complesso**. Questo può essere fatto anche con i reticolati che si muovono: prendiamo ogni singolo frame dell'animazione e sommiamo, ottenendo un segnale particolare che in psicofisica viene chiamato **Plaid** (richiamo alle "fantasie" dei plaid). **Un Plaid è la somma di due reticolati**. Quello che si può vedere chiaro è che il plaid si muove come un'unità, (a volte può succedere che due reticolati non si comportino così) in questo caso si vede un plaid unitario e esso ha una direzione di movimento che non è quella di nessuna dei due reticolati.



**Domanda:** *Un plaid è la somma di due reticolati che devono per forza essere in movimento?*

**Risposta:** Solitamente si usa questo termine per i reticolati in movimento, sì. Noi li chiamiamo reticolati per il teorema di Fourier, quello che noi diciamo sulla somma dei reticolati vale per la somma di qualsiasi componente. Potrei mostrarvi anche due figure che sovrapposte e organizzate nel giusto modo danno lo stesso risultato (movimento globale).

**In che direzione si muovono questi plaid che sono nell'ultima colonna a destra? Quello in alto?** Dal basso verso l'alto, in verticale. **E quello nella riga di sotto?** Da sinistra verso destra leggermente in diagonale verso l'alto.



Questi esempi ci mostrano che, se noi abbiamo due reticolari, localmente sui contorni di questi reticolari vengono registrati dei movimenti che sono nella direzione ortogonale all'orientazione del reticolo stesso. Se osserviamo il plaid di sinistra si deve immaginare che in tanti punti locali di quei contorni c'è un campo recettivo piccolo e in quel punto il movimento registrato è soltanto ortogonale all'orientazione del contorno, quindi dal punto di vista locale in quel plaid ci sono principalmente **due tipi di segnali locali**: uno che va verso l'altro e verso destra, e un altro che va verso il basso e verso destra, **perché quelle sono le direzioni ortogonali ai contorni che compongono il plaid**.

Come si vede dal disegno, nel plaid di sinistra se si fa la media dei due vettori si crea un vettore con una velocità molto maggiore, la freccia verde. **La media è proprio la direzione in cui si vede muovere il plaid quando viene messo in movimento.**

**Domanda:** *Quando l'immagine è statica, perché assumiamo che un vettore vada per esempio dall'alto verso destra e non verso sinistra in basso, visto che sarebbe la stessa componente ortogonale?*

**Risposta:** Perché bisogna pensare a quel plaid come composto da due grating, uno dei due grating ha i segmenti orientati in diagonale verso le “10” e le “2” dell’orologio, a 45 gradi, uno ha una componente ortogonale in una direzione e l’altra in un’altra. Quando è statico non ha nessuna componente di movimento in nessuna direzione, è solo quando si muove che il recettore di R rileva il movimento ortogonale alla direzione dl contorno.

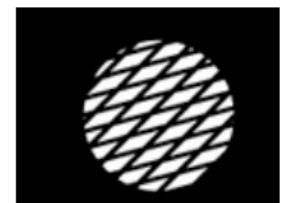
Una possibile soluzione del problema dell’apertura è dire: se io due contorni li metto insieme posso fare la media di queste componenti ortogonali, il quale rappresenta un modello di integrazione semplice, ovvero basta immaginare che ci sono dei meccanismi neurali che sono molto compatibili con proprietà di convergenza delle cellule del sistema visivo superiore, e questi sistemi fanno l’equivalente, ovvero la **media dei due vettori**. Il movimento reale di questo plaid è verso destra, il movimento percepito è anche verso destra quindi vediamo correttamente il movimento, superando a primo occhio il problema dell’apertura.

**Domanda:** *L'orientazione è sempre la stessa?*

**Risposta:** Si, la velocità è diversa, perché la regola del parallelogramma viene usata per sommare i vettori. Se si immagina di costruire un parallelogramma aggiungendo ai due vettori grigi (riferimento alle slide) due altri segmenti paralleli a quelli che sono disegnati, a quel punto la freccia verde arriva fino al vertice opposto del parallelogramma e sarebbe lunga il doppio, **quindi la lunghezza del vettore rappresenta la velocità, mentre il verso della freccia rappresenta la direzione.**

Come avviene questo è stato oggetto di dibattito, perché qualcuno sostiene che non è necessario fare la media dei vettori perché ci sono le intersezioni di quei contorni in quei punti, quindi il problema dell'apertura non vale, perché c'è una **discontinuità**, il recettore in quel punto di fatto registra un movimento in orizzontale. Molte ricerche hanno provato a vedere se cambiando le orientazioni, o il contrasto, il movimento rimane coerente o meno con le previsioni che avremmo facendo la media vettoriale, ed in molti casi la risposta è sì.

Ed esempio è questo è un altro plaid che è ottenuto sommando due reticolati che hanno delle orientazioni diverse, anche in questo plaid il movimento reale è verso destra, però se il sistema fa la media dei vettori in questo caso non dovrebbe essere in grado di recuperare il vero movimento del plaid, dovrebbe vedere il movimento illusorio o sbagliato perché se lui fa la media dei vettori le componenti ortogonali sono quelle grigie, e la media dei vettori è una freccia verde orientata in diagonale verso il basso e verso destra.



**In che direzione vedete muovere questo plaid?**

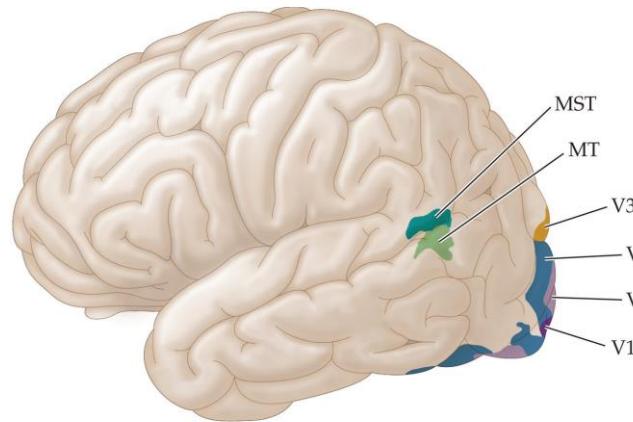
Non in orizzontale sicuramente, si muove verso la freccia verde... Ci sono varie alternative al fare la media dei vettori, ma ciò che hanno in comune questi modelli è che sono tutti basati su di una operazione lineare, ovvero una sorta di somma pesata in qualche maniera (somma vettoriale), un tipo di operazione che non ha bisogno di fare dei ragionamenti “intelligenti” su come vengono organizzati questi contorni, abbiamo due segnali e li sommiamo.

**Domanda:** Questo modello è stato dimostrato a livello fisiologico oltre che matematico?

**Risposta:** È stato dimostrato che questo tipo di processo integrativo è implementato da tutte le cellule sensibili al movimento che non si comportano come recettori di R, ma come delle cellule che rispondono al movimento globale del plaid, fanno l'equivalente della somma vettoriale e queste si trovano nell'**area MT**.

Se la media dei vettori corrisponde a quello che vedi non significa che quello che vedi corrisponde al movimento reale. La teoria è credibile proprio perché in certi casi il sistema può sbagliare. Le illusioni sono utili se c'è una teoria che le prevede.

**Dove sono questi rilevatori di movimento locale?**



Una risposta completa ancora non si è trovata. Tre indizi importanti:

1. La proiezione dorsale da V1 fino alla corteccia parietale posteriore riceve informazioni dalla via magnocellulare ed è noto che se si hanno delle **lesioni alle vie magnocellulari** del sistema visivo queste producono un deficit nel percepire il movimento di oggetti grandi e che si muovono rapidamente (movimento globale);
2. Cosa più interessante, i pazienti che hanno delle lesioni focali alla area MT spesso presentano l'**akinetopsia**, deficit nel percepire il movimento globale degli oggetti, è un sintomo bizzarro in cui gli oggetti apparentemente cambiano posizione senza continuità spazio-temporale, vedono come con la luce stroboscopica, immagini statiche, ma che in realtà sono in movimento. È molto invalidante (es: attraversare la strada);
3. Si è studiata la selettività al movimento dei neuroni nell'area MT, e si vede che ci sono tantissime unità neurali in quella zona che sono selettive per una particolare direzione di movimento che vedremo quale.

**Come possiamo pensare alla risposta direzionale di queste cellule, che sono sensibili al movimento in MT?**

Ci sono due ipotesi:

1. Potrebbe essere che queste cellule rispondono come se fossero dei recettori di R, cioè rispondono mostrando un singolo grating alla componente ortogonale di movimento. Rispondono bene indipendentemente dalla direzione del grating, a patto che abbiano la stessa componente ortogonale;
2. Nel caso di un plaid, esso sarà fatto di due componenti ortogonali, la distribuzione delle risposte possono essere alternativamente nella direzione ortogonale di un plaid o dell'altro. Oppure la direzione preferita potrebbe essere la media vettoriale.

Con queste due ipotesi, in laboratorio si è andato a studiare **a cosa corrispondevano la selettività per la direzione del movimento.**

Ciò si può fare in V1 dove ci sono le cellule complesse che rispondono al movimento, la loro selettività al movimento è ortogonale all'orientazione del grating. In questo caso si comportano come dei recettori di R. Anche mostrando un plaid, a seconda della cellula presa in considerazione, troviamo che rispondono ad una delle due direzioni ortogonali, ovvero come dei recettori di R.

**In V1 non ci sono cellule sensibili al movimento, la cui sensibilità alla direzione corrisponde al movimento globale del plaid. In V1 non si risolve il problema dell'apertura.**

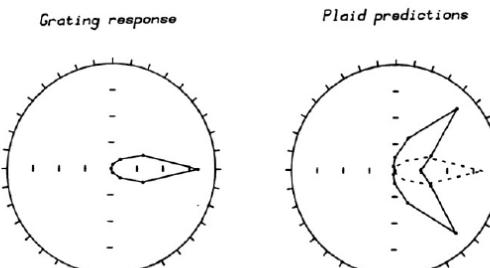


FIG. 9. Hypothetical data illustrating component and pattern directional selectivity. See text for details.

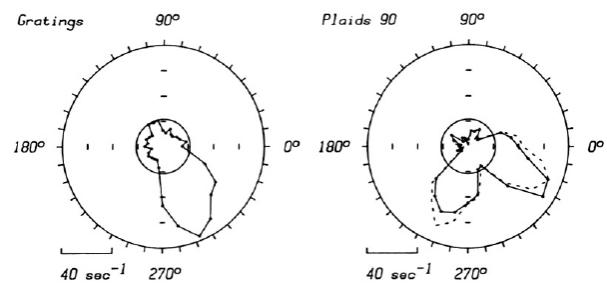
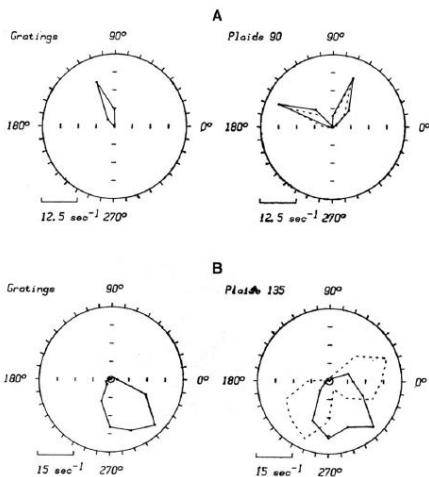


FIG. 10. Directional selectivity of a special complex cell recorded in area 17 of a cat. The spatial frequency was 1.2 c/deg, and the drift rate was 4 Hz. On the left is shown the neuron's tuning for the direction of motion of single gratings, and on the right is shown the neuron's response to moving 90 deg plaids. The dashed curve on the right shows the expected response of a component direction selective neuron. The inner circles in each plot show the neuron's maintained discharge level. For this cell the component correlation was 0.976, and the pattern correlation was -0.076 ( $n = 32$ ).

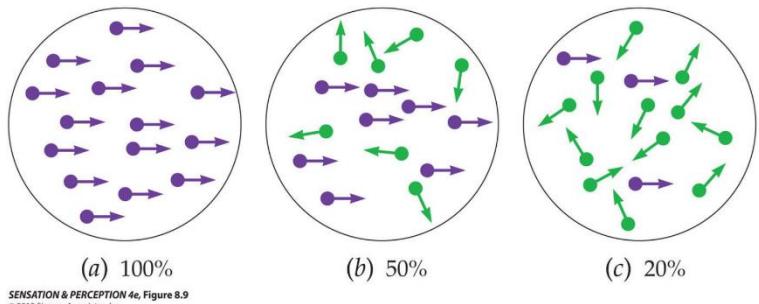


Se facciamo la stessa cosa in MT, nel **40%** delle cellule la risposta direzionale è la stessa di V1 (A), un **quarto** circa delle cellule si comporta in maniera diversa (B): se mostriamo un grating, risponderanno alla direzione ortogonale, se mostriamo un plaid non risponderanno più ad una delle due direzioni ortogonali, ma nella direzione che corrisponde alla media dei due vettori che compongono il plaid, quindi è un dato molto solido e incontrovertibile.

Ciò si può confermare in altre maniere, **Newsome** ha pubblicato un altro classico lavoro di lesione nella scimmia (1985). Se lesioniamo MT nella scimmia l'effetto non è macroscopico, **ma si può vedere se produce una riduzione nella sensibilità per il movimento globale**. Il piano di ricerca è semplice, misuro le soglie, e confronto scimmie lesionate in MT con scimmie con lesioni in aree diverse o in nessuna area.

La maniera di misurazione delle soglie è basata su di un **cinematogramma di punti causali**, cioè un cartone animato che contiene dei punti a caso da un frame all'altro, random.

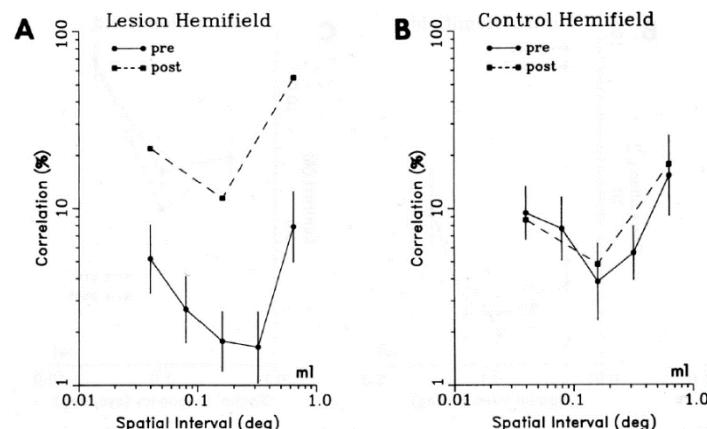
Questi punti possono avere una **componente di movimento globale**, cioè si possono muovere tutti quanti in una direzione, con una certa velocità. Per misurare una soglia si può variare il rapporto tra segnale e rumore, cioè invece di avere il 100% dei punti che si muovono con una componente comune di movimento, solo una certa percentuale si fanno muovere in maniera coerente.



SENSATION & PERCEPTION 4e, Figure 8.9  
© 2015 Sinauer Associates, Inc.

**COERENZA o CORRELAZIONE:** percentuale che esprime il rapporto tra segnale e rumore. La coerenza quindi verrà confrontata tra scimmie lesionate e non.

## Analisi della lesione in MT



Si veda il grafico qui affianco, le curve rappresentate somigliano alla curva della sensibilità al contrasto, cioè la variabile dipendente sull'asse y che è una misura di sensibilità che in questo caso è la soglia (l'inverso). Sull'asse delle x invece c'è la distanza che percorre il punto da un frame all'altro, quando si muove in maniera coerente, quindi è una curva che dimostra come la

**sensibilità al movimento abbia un certo andamento di sintonizzazione rispetto alla velocità.** Sull'asse delle x e dell'y i dati sono plottati su una scala logaritmica, questi dati sono misurati su un range abbastanza ampio di velocità e sensibilità.

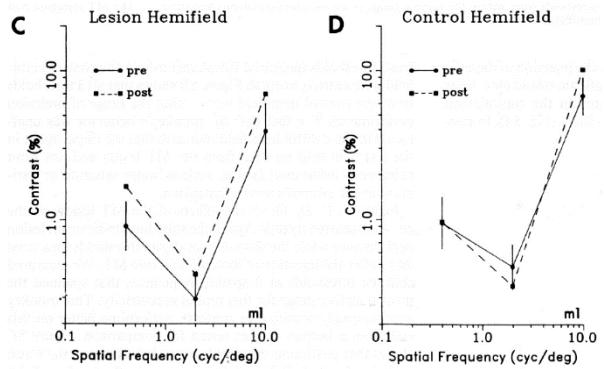
Sul grafico di sinistra vediamo come cambia la sensibilità della scimmia prima e dopo dell'operazione: **la soglia si alza di molto, la forma della curva rimane simile, ma la sensibilità è peggiorata.**

Nel caso dei dati preoperatori il dato medio viene presentato con una specie di segmento verticale, **cosa è?** L'errore standard o forse un intervallo di fiducia intorno ad una media, che vuol dire più o meno due errori standard, questo segmento ci dà anche un'informazione perché le misure della soglia sono state fatte in più sessioni preoperatorie, quindi si può stimare attraverso l'intervallo di fiducia quanto è precisa questa stima, in condizioni ottimale la soglia si colloca più o meno al 2% e questa stima potrebbe essere sbagliata più o meno di un ulteriore 1%.

**La lesione (soltanto ad un emisfero) nell'area MT ha prodotto una riduzione della sensibilità per il movimento globale, in accordo con l'ipotesi,** vediamo inoltre la completa sovrapposizione dei due grafici nei soggetti di controllo.

### Cosa succede alla sensibilità al contrasto?

Quello che si vede che in entrambi gli emicampi visivi le due curve sono ampiamente sovrapposte, quindi si può essere ragionevolmente sicuri che **la lesione ha largamente risparmiato la capacità di percepire le discontinuità orientate**. Se si è pignoli si può dire confrontando a e b, che una piccola differenza l'ha avuta anche la sensibilità al contrasto, quindi la lesione neurologica in MT ha prodotto anche deficit alla sensibilità al contrasto, tuttavia un'altra cosa che si può dire è che questo effetto è molto piccolo, basta guardare l'ordine di grandezza (1%).



Entrambi questi grafici fanno riferimento ad uno studio del 1988 di **Newsome & Paré**, pubblicato su J Neurosc, il primo grafico rimanda alla *sensibilità al movimento* a seguito di una lesione focale ad MT, mentre il secondo alla *sensibilità al contrasto* a seguito della medesima lesione.

Nelle prossime lezioni si vedrà come il problema dell'apertura dipenda da dei meccanismi più globali, più intelligenti, come questa dimostrazione di Michael Bach: <https://michaelbach.de/ot/mot-motionBinding/>

## Lezione 15

### Altre soluzioni al problema dell'apertura

Fino alla scorsa settimana si è parlato della percezione del movimento.

#### **Cos'è il problema dell'apertura?**

Le cellule al livello più basso (recettori di Reichardt) dell'organizzazione percettiva visiva non riescono esattamente a cogliere il movimento nella sua globalità, ma solo un movimento ortogonale all'orientazione del contorno che si muove (cioè, non possono misurare direttamente il movimento reale di un contorno ma sono sensibili solo a quella componente del movimento ortogonale all'orientazione del contorno stesso).

Il motivo per cui accade questo è che qualsiasi movimento reale dal punto di vista della cinematica (descrizione delle traiettorie dei movimenti) lo si può scomporre in 2 componenti: uno ortogonale e uno parallelo al contorno tramite la regola del parallelogramma.

La componente parallela al contorno il recettore non lo può misurare perché sul contorno non ci sono discontinuità e quindi perde una parte di quel movimento in una certa direzione.

Abbiamo anche parlato dei modi con cui si può risolvere il problema dell'apertura, soprattutto di quale processo è utile a ciò: la strategia per mettere insieme più movimenti locali è fare la media delle componenti vettoriali locali. Questa cosa è stata ampiamente indagata anche dal punto di vista fisiologico: sappiamo che MT è implicata in questo tipo di processo. Tuttavia, le cose sono più complicate di così, infatti la media degli output dei recettori di Reichardt non predice ciò che si percepisce. Quindi ora analizzeremo delle situazioni in cui la percezione del movimento globale non è semplicemente una media di vettori, ma è il risultato di processi più complessi.

### Effetto barber pole

È un cilindro in cui è disegnata una spirale e quando esso ruota il recettore di Reichardt dovrebbe comportarsi come se vedesse le diagonali attraverso un'apertura (quindi, localmente sono segnali di movimenti che hanno una direzione ortogonale all'orientazione del contorno).



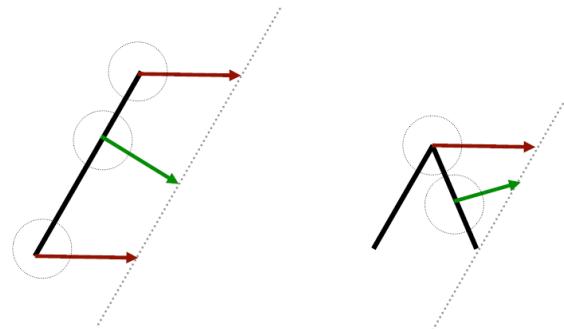
Quindi per le strisce diagonali dovrebbero esserci segnali diagonali a 45° (di fatto, se li si guardasse attraverso un buco, a seconda se il barberspole ruota in un senso o nell'altro, vedremmo i contorni scivolare in diagonale verso dx in basso o verso sx in alto).

Dato che non lo osserviamo così, noi vediamo solo un movimento verso l'alto o verso il basso in verticale. Se considerassimo i recettori di Reichardt messi lungo questo contorno, considerando le caratteristiche globali di una scena complessa ci rendiamo conto che i contorni (discontinuità orientate) hanno la struttura più complessa (non sono semplicemente un segmento che si estende all'infinito senza nessuna caratteristica riconoscibile).

Ci sono dei punti speciali chiamati “**terminatori**” dei contorni, cioè dove si interrompono (possono essere punti di incrocio in un plaid, un segmento con una lunghezza finita e che quindi si interrompe, uno spigolo). **Questi punti sono speciali perché dal punto di vista dei segnali locali di movimento per essi non vale il problema dell'apertura** (perché qui ci sono delle discontinuità identificabili sul contorno).

Quindi, i segnali locali di movimento sono almeno di 2 tipi diversi:

- Quelli che si trovano *lungo il contorno* (in verde) (soggetti al problema dell'apertura ed il rilevatore di Reichardt che codificherà solo la componente ortogonale);
- Quelli dove ci sono *terminatori di un contorno*, in cui il rilevatore di Reichardt viene stimolato da un cambiamento di posizione rispetto al tempo di un elemento che è identificabile da una discontinuità e quindi questo segnale è diverso da quello ortogonale alla orientazione del contorno. Questi contengono l'informazione veridica (cioè, l'informazione che corrisponde al movimento reale del contorno).



Quindi, è questo che accade nell'effetto barberpole: lungo il contorno diagonale i rilevatori di Reichardt non potranno che registrare la componente ortogonale all'orientazione del movimento, ma nel punto dove c'è il margine nel cilindro, il segmento si interrompe (c'è una discontinuità nel segmento stesso) e quindi percepiamo la componente di movimento in verticale. Quindi, il sistema deve fare la media vettoriale di tante componenti vettoriali e soprattutto deve mettere insieme i segnali oggetti al problema dell'apertura con quelli veridici e deve decidere quanto peso dare a questi 2 segnali.

**Nell'effetto barberpole diamo più peso ai segnali dei terminatori e poco agli altri, ecco perché vediamo il movimento in verticale e non ortogonale all'orientazione dei contorni.**

### *Come fa il sistema a decidere quali siano più o meno credibili?*

Non basta dire che il sistema sia costruito in maniera tale che presta attenzione ai segnali dei terminatori (esempio: se riduciamo progressivamente un segmento che si muove e che occupa uno schermo grande, non avendo molte informazioni sui movimenti relativi di questo segmento rispetto al contesto, se esso è molto lungo il movimento tende ad essere nella direzione ortogonale alla sua orientazione. Mentre se si riduce la lunghezza di questo segmento, succede che ci sono sempre meno segnali lungo il segmento e rimangono sempre 2 segnali terminatori, quindi cambia il numero di segnali di un tipo rispetto all'altro, si riduce il peso dei segnali ortogonali e dovrebbero diventare più preponderanti nel carico del movimento globale i segnali dei terminatori. Quindi cambiando la lunghezza del segmento, lo si vede in una direzione diversa a quella ortogonale all'orientazione del contorno.)

L'effetto barberpole è un “*effetto etichetta*” fuorviante che si applica a qualsiasi cornice di qualsiasi forma e orientazione all'interno di cui si fa muovere un contorno.

Ad esempio in questa animazione i gratings si muovono in diagonale



da sinistra a destra, ma se focalizziamo l'attenzione solo sul *rettangolo verticale* possiamo vedere come il grating sembra muoversi in verticale, anche se nei lati corti ci sono terminatori che vanno in orizzontale; il *rettangolo orizzontale* si muove in orizzontale, ma ci sono sempre i terminatori nei lati corti che vanno in verticale; quindi se è vero che il movimento dei terminatori si trascina nel calcolo del movimento globale tutto il resto, dovrebbe esser possibile far cambiare direzione allo stesso movimento fisico semplicemente mettendolo in una cornice che cambia di orientazione, infatti lo si vede nella *S rovesciata* dove in alto c'è un movimento orizzontale, gomito, va in verticale, gomito, torna orizzontale; il *cerchio* corrisponde alla configurazione per studiare il problema dell'apertura, i tanti movimenti dei terminatori e nel cerchio sono in tutte le direzioni (partono in alto e scivolano lungo il contorno circolare) quindi localmente la direzione del movimento sarebbe tangente alla circonferenza e quindi il punto cambia continuamente direzione.

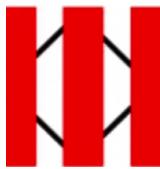
Quindi il motivo per cui quando guardiamo un contorno che si muove attraverso un buco lo vediamo muoversi lungo la direzione ortogonale all'orientazione del contorno, **non è che vediamo quello che ci dice il recettore di Reichardt, ma in quelle condizioni mettiamo insieme i terminatori con i segnali ortogonali** (i terminatori sono in tutte le direzioni in una gamma di  $180^\circ$  definite dal verso del movimento e quindi si cancellano l'uno con l'altro, gli unici coerenti sono i segnali ortogonali).

Quindi il problema dell'apertura è un qualcosa che definiamo facendo uno sforzo di astrazione, chiedendoci cosa riesce a registrare un recettore di Reichardt (modello di un detettore locale di movimento) che non vede nulla in sé per sé, **è il nostro sistema visivo che produce una percezione come risultato di un processo integrativo**. Questa è quasi una conclusione paradossale, infatti il motivo per cui vediamo il movimento in diagonale non è per via del problema dell'apertura, in realtà, a causa di esso solo indirettamente, per quei segmenti ci sono delle componenti in direzione ortogonale al contorno e poi ce ne sono lungo i terminatori. Quindi il risultato del dato percettivo dipende sempre da questa attività integrativa.

I terminatori (le discontinuità che si muovono lungo il contorno della cornice in cui c'è un segmento) non spiegano del tutto perché le superfici che abbiamo nel mondo sono in 3 dimensioni: sono figure davanti a sfondi, presuppongono una **segmentazione figura-sfondo**, un'**organizzazione dei contorni nello spazio** e che il sistema quando c'è un segmento **decida chi sia il proprietario dei contorni**, chi è la figura e chi è lo sfondo.

Questo vale anche quando i segmenti sono in movimento e anche quando ci sono terminatori di questi segmenti in movimento (esempio: nell'effetto barberpole per quei terminatori che scorrono lungo la cornice il cervello implicitamente "decide" che quei terminatori appartengono al segmento e che quest'ultimo si muove davanti ad uno sfondo e quindi li mette insieme come/con componenti ortogonali). Questo ci consente di fare una previsione: se manipoliamo la struttura globale dell'organizzazione in 3 dimensioni, l'organizzazione figura-sfondo, la border ownership di questi contorni dovremmo poter, senza cambiare nient'altro dei movimenti locali, cambiare la percezione della direzione del movimento stesso dei contorni.

Vediamo un altro esempio, in questo caso abbiamo 4 segmenti che si muovono accoppiati in verticale: lungo i contorni ci sono punti in cui vale il problema dell'apertura (ci sono tanti piccoli vettori in diagonale che specificano un movimento nella direzione ortogonale all'orientazione di quei contorni), non vediamo quella direzione perché ci sono i terminatori che si muovono in verticale e i recettori locali registrano questo movimento verticale e quindi questi catturano il movimento globale e perciò si vedono i movimenti in verticale.



Tuttavia aggiungendo 3 superfici rosse che stimolano il nostro **processo di organizzazione percettiva** così che quei 4 segmenti diventino un'unica figura che si completa dietro di esse. Il movimento singolo locale di quei 4 pezzi visibili è fisicamente uguale a prima, eppure l'impressione è di una losanga che si muove orizzontalmente dietro le bande.

A quanto pare la maniera in cui il nostro sistema visivo tratta questi segnali locali è molto sofisticata, non è semplicemente un'operazione lineare, ma è un'operazione capace di fare qualcosa di molto più difficile dal punto di vista della modellizzazione: **è capace di dire che in certi casi i segnali che vengono dai terminatori sono un'informazione molto utile e quindi la utilizza per decidere come si sta muovendo l'oggetto a sinistra, ma se invece il sistema ha altre informazioni che gli indicano che i terminatori non sono i punti in cui i segmenti davvero s'interrompono, sono solo la parte visibile di un segmento che continua dietro la banda** (quindi in quei punti il proprietario del bordo non è il segmento ma la banda rossa, perché il nero in quei punti non ha terminatori).



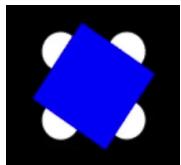
Questa è una forma di classificazione e di interpretazione dell'informazione molto sofisticata e altamente simbolica. Siamo passati dal dare tutto il peso per l'interpretazione dell'informazione ai terminatori locali o essere completamente ignorati se vengono a causa di un processo di organizzazione classificati come non terminatori nel **processo di ownership**, quindi rimangono solo i segnali ortogonali (ognuno dei 4 segmenti ha un movimento locale in diagonale, alcuni a sinistra e altri a destra) ed il sistema non può che metterli insieme con una media vettoriale in orizzontale ed ecco che vediamo questo effetto.

### **Effetto breathing square**

Si tratta di un quadrato che ruota, ma sembra *contrarsi ed espandersi*.

#### **Come mai?**

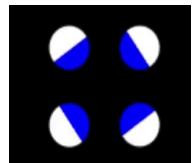
Questa componente di movimento che vediamo è in un certo senso “illusoria”. Questa particolare illusione dà questo effetto quando i vertici degli angoli scompaiono dietro gli occludenti (cioè, quando si toglie l'informazione sui terminatori). Ecco perché in quei momenti dove la rotazione diventa meno visibile e diventano più salienti le componenti ortogonali all'orientazione dei contorni (puntando verso il centro o allontanandovisi a seconda della fase della rotazione).



Tolti gli *otturatori*, si vede bene che il quadrato ruota, perché dove ci sono i punti dove si incontrano i segmenti ci sono gli effetti dei terminatori che influenzano la visione di questo movimento. Tutto questo discorso sulla segmentazione figura-sfondo se fosse vero dovrebbe essere possibile riprodurlo anche in situazioni minimali. Dovrebbe succedere anche se semplicemente si prende un segmento, lo si fa muovere attraverso un'apertura circolare, il segmento se non ha nessun elemento individuabile, darà il solito effetto che sembra muoversi lungo la direzione ortogonale. Se però si mette un elemento individuabile, esso dovrebbe catturare il movimento reale del segmento e quindi a quel punto sembrerà cambiare direzione.

Il punto bianco è una figura su uno sfondo nero (il segmento) che è a sua volta una figura su uno sfondo bianco (il cerchio), etc., allora potremmo fare un'altra previsione: il puntino bianco non cattura il movimento del segmento nero, ma si crea un'organizzazione gerarchica in piano diversi (il puntino si muove rispetto al suo sfondo nero in una certa direzione e lo sfondo nero si muove in un'altra direzione rispetto al suo schema di riferimento circolare).

**Se si guarda il puntino bianco sembra muoversi in diagonale, invece il suo movimento reale è in orizzontale.** Quindi, l'idea è che la percezione globale di movimenti sia il risultato di un processo di organizzazione gerarchica di movimenti distinti (ogni movimento è un movimento rispetto ad un altro).



## Interazioni multisensoriali di alto livello: Stream – Bounce effect

Spesso la percezione del movimento non è solo visiva, ma ci sono anche informazioni che provengono da altri canali sensoriali (Esempio: ogni volta che l'oggetto in movimento è in contatto con noi, quando l'oggetto in movimento è una sorgente sonora quindi c'è un processo di percezione uditiva del movimento).

Fino ad ora abbiamo parlato di diversi modi con cui poter risolvere il problema dell'apertura, dal *movimento locale* (registrato dai recettori di Reichardt) al *movimento globale* dell'oggetto:

- Una parte di questo processo sembra essere spiegato dal fatto che il sistema fa una media vettoriale dei vettori locali (processo ancora di early vision anche se è già di tipo integrativo);
- Il movimento dei terminatori può interagire con le componenti ortogonalni e quindi fare una media vettoriale non basta;
- Come viene trattato il movimento dei terminatori dipende anche da come si classificano questi, quindi da un processo che ha a che fare con la border ownership e con l'interpretazione globale della scena (cosa sta davanti e cosa dietro, cosa si muove rispetto a qualcos'altro, etc.);
- In certi casi questo processo di interpretazione globale non è solo visivo, ma può tener conto anche di informazioni provenienti da altri canali sensoriali, quindi anche attraverso interazioni multisensoriali il sistema sembra in certi momenti risolvere il problema dell'apertura.

### **Perché dobbiamo andare a considerare la percezione del movimento oltre la modalità visiva, uditiva e aptica?**

Perché l'area MT+ (sotto-area di MT) è attivata da stimolazione elettro-tattile della lingua sia in individui normodotati sia in individui ciechi congeniti. L'**area MT**, quindi, non sembra essere coinvolta solo nella percezione visiva del movimento, ma sembra essere anche attività da un movimento meccanico su un distretto cutaneo (hanno scelto la lingua perché è una delle zone più sensibili dal punto di vista tattile).



#### Research report

Beyond visual, aural and haptic movement perception: hMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals

Isabelle Matteau<sup>a,\*</sup>, Ron Kupfers<sup>b,c,\*</sup>, Emiliano Ricciardi<sup>d</sup>, Pietro Pietrini<sup>d,e</sup>, Maurice Ptito<sup>a</sup>

<sup>a</sup> Ecole d'Ingénierie de l'Université de Montréal, Montréal QC, Canada H3C 2B2

<sup>b</sup> Institute of Neuroscience & Pharmacology, Novum Institute, Copenhagen University, 2200 Copenhagen, Denmark

<sup>c</sup> PET Unit, Rigshospitalet, Copenhagen University, 2100 Copenhagen, Denmark

<sup>d</sup> Laboratory of Clinical Biochemistry and Molecular Biology, University of Pisa, Pisa, Italy

<sup>e</sup> Department of Laboratory Medicine and Molecular Diagnostic, Pisa University Hospital, Pisa, Italy

#### ARTICLE INFO

Article history:  
Received 11 December 2009  
Received in revised form 8 April 2010  
Accepted 2 May 2010  
Available online 11 May 2010

Keywords:  
Cross-modal plasticity  
Sensory substitution  
Supramodality  
Vision  
fMRI  
Tongue display unit  
Dorsal stream  
Motion

#### ABSTRACT

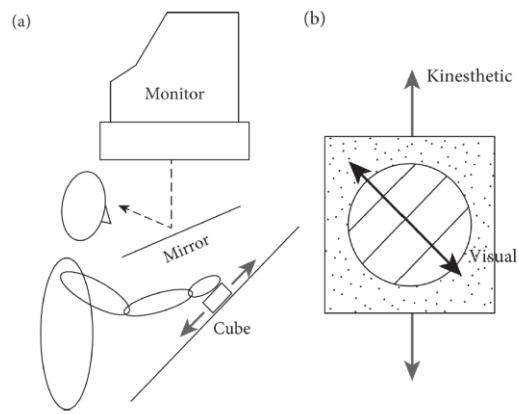
The motion-sensitive middle temporal cortex (hMT+ complex) responds also to non-visual motion stimulation conveyed through the tactile and auditory modalities, both in sighted and in congenitally blind individuals. This indicates that hMT+ is truly responsive to motion-related information regardless of visual experience and the sensory modality through which such information is carried to the brain. Here we determined whether the hMT+ complex responds to motion perception *per se*; that is, motion not perceived through the visual, haptic or auditory modalities. Using functional magnetic resonance imaging (fMRI), we tested brain responses in two groups of subjects: sighted and congenitally blind individuals who had been trained to use the tongue display unit (TDU), a sensory substitution device which converts visual information into electrotactile pulses delivered to the tongue, to resolve a tactile motion discrimination task. Stimuli consisted of either static dots, dots moving coherently or dots moving in random directions. Both groups learned the task at the same rate and activated the hMT+ complex during tactile motion discrimination, although at different anatomical locations. Furthermore, the congenitally blind subjects showed additional activations within the dorsal extrastriate cortical pathway.

These findings extend our knowledge about the sensory and functional organization of hMT+ complex by showing that this cortical area processes motion-related information *per se*, that is, motion stimuli that are not visual in nature and that are administered to body structures that, in humans, are not primarily devoted to movement perception or spatial location, such as the tongue. In line with previous studies, the differential activations between sighted and congenitally blind individuals indicate that lack of vision leads to functional rearrangements of these supramodal cortical areas.

© 2010 Elsevier Inc. All rights reserved.

Vediamo alcune dimostrazioni psicofisiche sul fatto che il problema dell'apertura sembra essere influenzato anche da **informazioni cross-modali**.

### Come si può dimostrare questo?



Hu and Knill (2010) Curr Biol 20(10): R436–7

In questo esperimento vengono combinati informazioni di natura tattile e muscolare (info che vengono dal movimento che si fa col braccio, cinestesia) vengono combinati col l'informazione visiva sul movimento di qualcosa.

Il setting fu creato per far sì che si potesse decorrelare (controllare indipendentemente) il movimento visto di un oggetto (un grating che si muove attraverso un'apertura) combinandolo con un movimento della mano (utilizzando il

monitor di un computer, uno schermo che presenta un'immagine, uno specchio e un manipulandum). Quindi, il soggetto tiene in mano un cubo che si può muovere lungo una rotaia (quindi la direzione del movimento è obbligatoria) in verticale. La mano e il cubo non si vedono perché davanti ad essi c'è uno specchio a 45° e in esso si vede il riflesso di quello che viene proiettato sul monitor sopra allo specchio. Questo apparato sperimentale consente di presentare lo stimolo visivo che, dal punto di vista dell'ottica proiettiva dello stimolo stesso, è collocato dietro lo specchio (ad una distanza che è pari alla distanza dell'occhio dallo specchio).

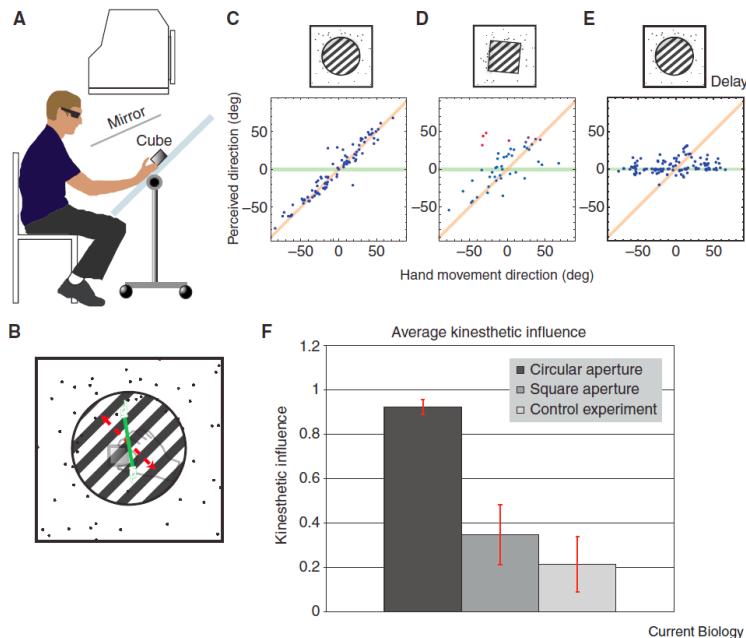
Con questo tipo di apparato si crea una situazione in cui vediamo il contorno muoversi sullo schermo dietro lo specchio proprio dove c'è la nostra mano sul manipulandum, quindi dal punto di vista percettivo è come se fossero nella stessa posizione nello spazio l'oggetto che muovo e la nostra mano che non vediamo ma che sentiamo e l'oggetto che vediamo nello schermo.

Quindi si crea in quel punto virtuale dello spazio una coerenza spaziale di 2 informazioni di movimento: quella visiva dello stimolo che si muove attraverso un'apertura circolare creando un segnale visivo orientato nella direzione ortogonale all'orientazione del contorno; questo viene combinato con un segnale di natura cinestesica (relativo ai segnali somato-sensoriali) che arrivano alla corteccia somato-sensoriale sul movimento del nostro arto.

Di questi 2 segnali può essere manipolato il **grado di concordanza** fra l'*informazione visiva e cinestesica* (Esempio: se il movimento della mano è in verticale e il contorno è così com'è, ci sarà un conflitto, visivamente dovrebbe muoversi in diagonale, ma dal punto di vista aptico dovrebbe muoversi in verticale).

## Quindi, quale tra i 2 segnali ha prevalenza?

Se li si vuole mettere in accordo, cioè, per far sì che il segnale cinestesico sia nella stessa direzione del segnale visivo, nell'apparato o si ruota lo stimolo visivo in modo da rendere i 2 movimenti molto diversi.



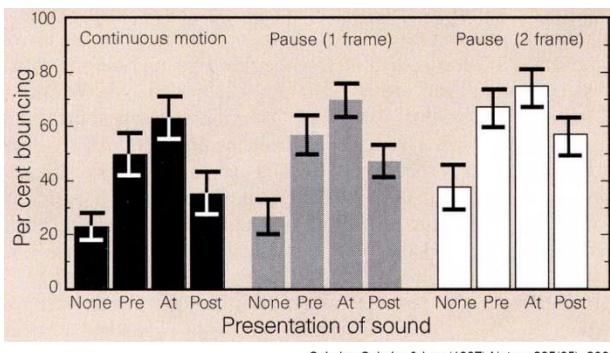
- A.** L'apparato sperimentale;
- B.** Dimostrazione del fatto che il soggetto non vede né la propria mano né il manipulandum però lo si sente dietro lo specchio. Il movimento del manipulandum è verticale, mentre il movimento visivo del grating è nella direzione ortogonale all'orientazione dei contorni;
- C.** Il primo grafico si riferisce alla situazione descritta in B ed è un plot della direzione percepita di movimento espressa in gradi (in che direzione si muove

apparentemente), in funzione del movimento della mano (quindi ruotando la rotaia). C'è un'associazione molto forte fra queste 2 variabili (correlazione molto vicina ad 1): **il movimento della mano determina il movimento in cui vediamo muovere il grating**. Quindi il sistema in queste condizioni di movimento volontario e intenzionale usa questa informazione per risolvere il problema dell'apertura, ma la usa se ha di nuovo delle informazioni contestuali coerenti con l'idea che queste due cose si appartengano (cioè, che il soggetto stia proprio muovendo ciò che vede);

- D.** Qui lo stimolo è uguale a quello su cui è basato il grafico in B e C. Tuttavia, di diverso hanno sfasato nel tempo le 2 info a diversi canali sensoriali, c'è un *delay*. Quindi il cervello non riesce più a mettere insieme il segnale visivo con quello cinestesico e infatti il grafico è completamente diverso: indipendentemente dal movimento della mano (asse delle X), la direzione percepita del movimento è sempre su quello che viene classificato come 0 (la componente ortogonale). Quindi, **a seconda del parametro del grado di concordanza spazio-temporale, il sistema tende a mettere insieme** (risolve il problema dell'apertura e vediamo l'oggetto muoversi nella direzione in cui lo muoviamo) **o a separare i segnali perché troppo sfasati** (non percepiamo il movimento come causato dal movimento della nostra mano e quindi l'unica informazione che ci rimane è la componente ortogonale all'orientazione del contorno che vedremo indipendentemente dal movimento della nostra mano).

## Stream-bounce effect

Le interazioni multisensoriali possono venire a tanti livelli diversi, anche a livelli più alti di interpretazione della struttura di una scena dal punto di vista spaziale tridimensionale, ma anche dal punto di vista dell'interpretazione delle forze che entrano in gioco in un certo evento che vediamo (relazione di causa-effetto). Questo è stato studiato usando l'**effetto stream-bounce** (scivolare-rimbalzare). Se fissiamo la croce di fissazione dovrebbe accadere che in certi momenti le palline rimbalzano e in altri che scivolano l'una sull'altra e per via della continuità di direzione del movimento ci si aspetterebbe una prevalenza dello stream.



Ma se all'improvviso s'introduce un segnale di tipo acustico nel momento in cui i pallini si toccano, succede che avviene una modifica di come il soggetto può vedere questo *movimento bistabile*.

Quest'effetto è stato studiato in modo dettagliato e si è visto che a seconda di dove si presenta il suono, si può produrre uno spostamento molto visibile nella statistica

delle risposte bounce (cioè, la probabilità che una persona dica che li vede rimbalzare) o fare lo stesso col reciproco per le risposte stream. Le *barre nere* si riferiscono ad un movimento continuo come quello di prima, se non si presenta nessun suono, la probabilità che le persone dicano di vederle rimbalzare è molto bassa (barra none), c'è un bias nel vederle nello stream. Se però si **introduce un suono** le barre cominciano ad alzarsi (cioè, la probabilità di vedere il rimbalzo aumenta), soprattutto se c'è coincidenza spazio-temporale (cioè, se il suono accade nel momento in cui i 2 pallini si toccano visivamente); se il suono arriva un po' prima funziona meno; se arriva ancor più in ritardo la probabilità di bounce diminuisce sempre più.

Questo si definisce "**finestra di integrazione spazio-temporale**": quando bisogna mettere insieme un segnale visivo con uno acustico, di solito, questa integrazione non avviene solo quando sono perfettamente coincidenti nel tempo, ma all'interno di una finestra e può anche essere modificata la sua grandezza (questo è il motivo per cui quando vediamo un film con ritardi nel doppiaggio, succede che dopo un po' modifichiamo la nostra finestra di integrazione spazio-temporale audio-visiva e si smette di accorgersi di questo problema, come anche nelle videoconferenze, se il ritardo è troppo grande non è proprio possibile combinare le informazioni).

Questo fenomeno usando l'effetto stream-bounce funziona bene anche se oltre all'effetto continuo si pone anche una pausa (barre grigie e bianche): mettendo una piccola pausa nel momento in cui compare il suono, dal punto di vista visivo viene rinforzata l'impressione del rimbalzo (perché le cose quando rimbalzano per un attimo si fermano), questo dipende anche dalla lunghezza della pausa.

L'**effetto stream-bounce** potrebbe derivare da una forma di *suggerimento* (sentiamo un rumore, si dà quella descrizione, ma noi per davvero non vediamo un effettivo cambiamento nella direzione del movimento).

**Sekuler e Lau** hanno pubblicato un articolo su Nature in cui affrontano questo problema in cui descrivono un esperimento di controllo cruciale per la sua comprensione.



Conclusioni sul problema dell'apertura e sue soluzioni:

1. **Media vettoriale;**
2. **Terminatori;**
3. **Proprietà del confine e interpretazione della scena;**
4. **Interazioni multisensoriali.**

The effect of sound on visual motion could represent some generalized attention effect, evocable by any salient transient. We examined this possibility by testing three conditions with 15 new observers. In one condition, a sound (440 Hz, 100 ms, 80 dB) came on only at the point of coincidence. In another condition the same sound, equally intense, was on for the entire visual display, but was turned off for 100 ms at the moment of coincidence. In the final condition no sound was presented. As before, sound onset significantly increased bouncing reports ( $P < 0.01$ ). However, sound offset produced results indistinguishable from those with no sound at all. This suggests that the sound's impact on visual motion is not the product of heightened attention at the moment of coincidence, but may require an acoustic event that signals a collision between moving objects.

## Lezione 16

### **Percepire un mondo stabile: le costanze percettive**

**Costanza di posizione:** quando si muovono gli occhi e il corpo, tutti i movimenti fanno sì che l'occhio occupi dei punti di vista diversi nel tempo (relativamente poco diversi quando facciamo solo dei micro-movimenti oculari durante le fissazioni oppure molto diversi quando ci spostiamo velocemente, come ad esempio guidando l'auto in autostrada). Questo fa sì che sulla retina **la proiezione del mondo esterno di tutti gli oggetti, quindi la struttura spazio-temporale della luce che arriva all'occhio, si modifichi continuamente**. Sulla retina non ci sono infatti delle "foto", semmai ci sono dei "filmati". I movimenti che determinano la non stabilizzazione del "filmato" vengono registrati nel sistema ma ciò che vediamo risulta stabilizzato, infatti la visione risulta perfettamente stabile, evidenziando quanto sia perfetta la costanza di posizione.

**Costanza di forma:** quando un soggetto si muove nella stanza, chi lo osserva non lo vede deformarsi. Durante il movimento il soggetto cambia i punti di vista nella spettatore e nella retina la proiezione del

suo corpo subisce una serie di trasformazioni proiettive che non determinano però la percezione di deformazione di quanto si osserva.

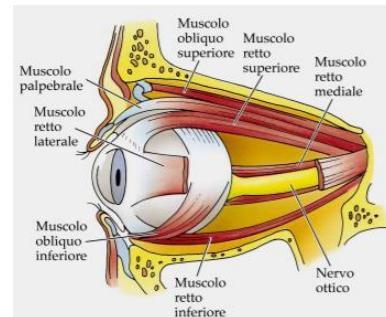
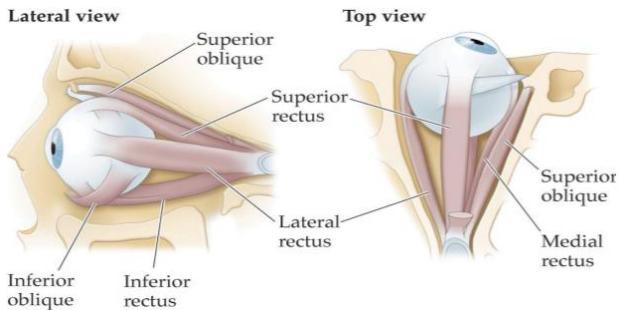
**Costanza di grandezza:** nel caso in cui un soggetto si allontana o avvicina a degli spettatori, ci sono dei cambiamenti evidenti della grandezza della proiezione del sul corpo sulla retina dello spettatore. Viene comunque mantenuta una percezione di costanza di grandezza.

**Costanza del colore:** il colore degli oggetti rimane lo stesso, salvo piccole situazioni particolari in cui sono notabili variazioni. Ad esempio, la percezione di un foglio di carta bianco non cambia con condizioni di illuminazione diverse (lampada, luce solare ecc.).

### Costanza di posizione\*:

L'argomento della costanza di posizione consente di occuparsi di un *tema multisensoriale*: riflettere sulle implicazioni del movimento per la percezione e viceversa, dato che si usano i dati percettivi per guidare i movimenti.

## Processi visuomotori riferiti ai movimenti oculari

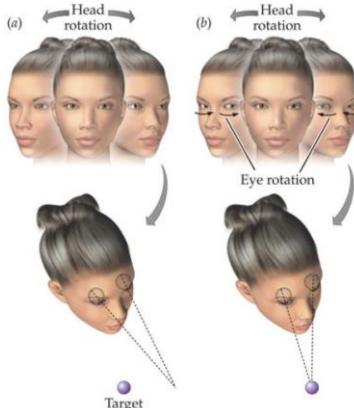


In immagine sono rappresentati i *muscoli oculomotori*. Possono far muovere l'occhio in tre modi diversi: orizzontale, verticale e ciclovergenza.

Questi muscoli sono in grado di implementare diverse tipologie di movimento oculare:

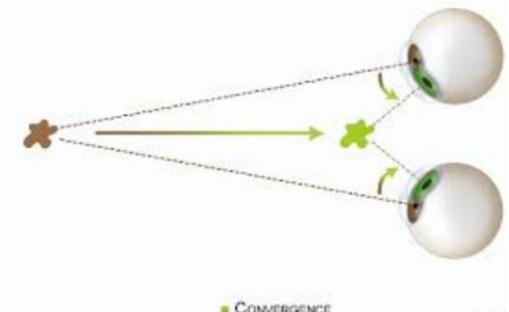
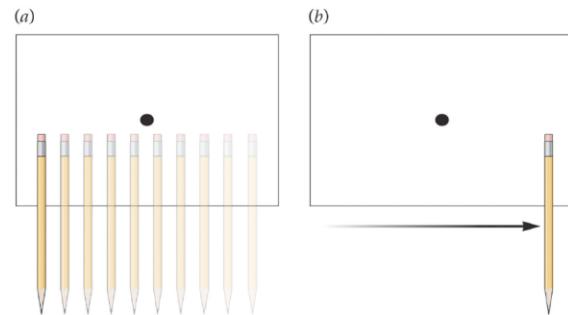
- Il riflesso vestibolo-oculare;
- L'inseguimento lento;
- La convergenza;
- I movimenti saccadici.

I primi tre divergono dai movimenti saccadici sotto molti punti di vista, di seguito esemplificati.



Il **riflesso vestibolo-oculare**, anche detto “*Doll’s Reflex*”, consente di mantenere la fissazione su un oggetto quando si ruota la testa: gli occhi in questo caso controruotano in maniera automatica al fine di mantenere la fissazione.

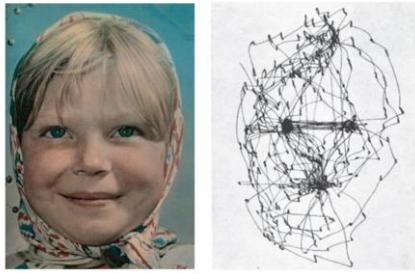
Anche il **movimento di inseguimento lento**, in inglese *Smooth Pursuit*, serve a mantenere la fissazione. Immaginando di fissare un punto dritto davanti a sè e muovendo un oggetto nel campo visivo. Ad esempio un movimento da sinistra verso destra come in figura, l’oggetto scivola sulla propria retina e si mantiene la fissazione in fovea sul punto. Tipicamente però succede che quando un oggetto passa davanti e il soggetto cerca di metterlo in fovea, lo fissa e poi cerca di mantenerlo in fovea, quindi l’occhio ruota, si alza ecc, cercando di seguire il movimento dell’oggetto. Questo movimento di inseguimento è definito lento perché può essere eseguito solo per oggetti che si muovono a velocità non eccessiva. Il movimento non è perfetto perché il soggetto non riesce a mantenere sempre l’oggetto perfettamente in fovea. Ci sono infatti dei continui aggiustamenti che fanno sì che l’oggetto inseguito sia più o meno fisso sulla retina. È comunque da ricordare che durante il movimento di inseguimento es. della matita, tutto ciò che è dietro o davanti la matita stessa si sposta.



Nel movimento di **convergenza**, che ha l’obiettivo di mantenere la fissazione, l’angolo formato dai due occhi si modifica in modo da mantenere la fissazione su un oggetto a seconda che sia vicino o lontano. Più l’oggetto è vicino più gli occhi si incrociano e viceversa.

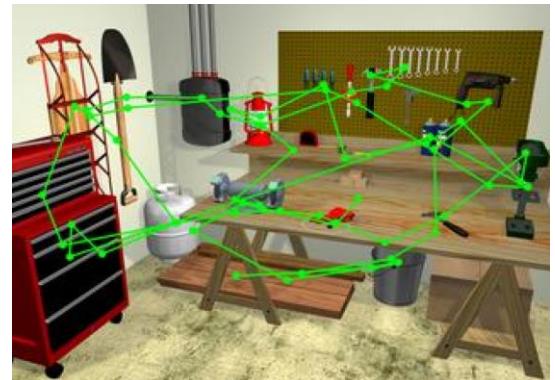
Questi tre movimenti hanno in comune l'esser guidati da **feedback sensoriali** (afferenti). Esistono dei circuiti che online, durante il movimento, tengono sotto monitoraggio uno *stato desiderabile* (mantenere la fissazione, ossia mantenere in fovea l'oggetto) e quindi calcolano un segnale di errore (quanto è discostato dallo stato desiderabile) che se presente porta a far correggere online il movimento stesso. Tale meccanismo, simile a quello di un termostato, tende dunque a ridurre la discrepanza tra lo stato attuale e quello desiderabile, ed è di tipo afferente. Ricordiamo dunque che esistono due principali **flussi di informazione**:

1. **Afferente**, dai muscoli (*proprioceuzione e cinestesia*) verso la corteccia o dalla retina (*visione*) (visione) verso la corteccia;
2. **Efferente**, nell'altra direzione, quindi dal cervello alla periferia (*movimento volontario*).



I **movimenti saccadici** sono velocissimi, simili a salti improvvisi che l'occhio compie per passare da una fissazione ad un'altra. Ne esistono di vari tipi (raggio più o meno breve, deliberati o riflettivi, guidati a memoria ecc) e hanno un ruolo fondamentale nell'attività esplorativa che dà inizio all'esplorazione e permette di acquisire informazioni sul mondo esterno guardandolo. Sono un elemento fondamentale del ciclo percezione-azione.

In immagine è possibile vedere un'illustrazione dello "Scanpath": il percorso di scannerizzazione che mostra le fissazioni (pallini) e la traiettoria dell'occhio da una fissazione all'altra. Questa dinamica è svolta continuamente quando si guarda una scena e **la caratteristica fondamentale che distingue i movimenti saccadici da altri movimenti oculari**, oltre la velocità, è **che non sono guidati sulla base del feedback afferente, ma sono pre-programmati**. Questo è un aspetto fondamentale perché caratterizza due tipologie distinte di trasformazioni sensomotorie.



Le **trasformazioni senso-motorie** sono dei processi psicologici che producono output motori (movimenti) a partire da input sensoriali (il soggetto per muoversi deve avere delle informazioni sul mondo esterno). Dunque l'output di questa tipologia speciale di trasformazioni sono dei movimenti che costituiscono delle **azioni**, ovvero degli insiemi di movimenti coordinati voltati ad uno scopo.

Per guidare dei movimenti è necessario coordinare input di **natura visiva** con input di **natura somatosensoriale**, relativi alla **proprioceuzione** (percezione di com'è il proprio corpo nello spazio) e alla **cinestesia** (percezione di come ci si sta muovendo nello spazio).

Quando si deve eseguire una trasformazione sensomotoria è necessario risolvere un problema assolutamente non banale.

## problemi inversi



Nella percezione visiva il cervello deve risolvere un **problema inverso**. A partire da un oggetto tridimensionale, calcolando la proiezione in due dimensioni si ha un problema di *geometria proiettiva*, che è un **problema diretto** che ha un'unica soluzione possibile. Sul cervello il problema è inverso perché sulla retina arriva l'informazione bidimensionale e partendo dalla proiezione si cerca di inferire/indovinare/ricostruire l'oggetto tridimensionale

che l'ha creata. La soluzione non è univoca: qualsiasi ellisse in 2D può essere la proiezione di diversi oggetti in 3D, come nell'esempio della stanza di Ames o nell'esperimento delle tre sedie.

La trasformazione sensomotoria è simile perchè anch'essa è costituita da un problema inverso. Se si definiscono una *serie di traiettorie e forze* che si vogliono applicare a livello muscolare, è possibile calcolare il movimento che ne risulterà. Nel caso della trasformazione sensomotoria si deve fare l'opposto: **si parte dallo stato finale desiderato** (Esempio: la mano che deve afferrare la tazza) **e si deve risolvere il problema inverso**, ovvero **definire quali sono le traiettorie e le forze necessarie per arrivare a quello stato finale**.

Questo problema è risolvibile in due maniere diverse: utilizzando l'informazione afferente o efferente, infatti esistono **due principali strategie per la guida percettiva di un'azione**:

1. A **circuito chiuso**, basata sul controllo on-line, che utilizza il feedback sensoriale e che risulta essere poco costosa, ma lenta;
2. A **circuito aperto**, la quale necessita di pianificazione, produce dei movimenti balistici e risulta essere veloce, ma molto costosa.

Si può utilizzare il feedback sensoriale durante il movimento o risolvere il problema inverso, analiticamente, prima di iniziare a muoversi e poi lanciare il movimento. Ciò corrisponde alle due fondamentali strategie per guidare il movimento: “**closed-loop**” (a circuito chiuso) o “**open-loop**” (a circuito aperto). Il termine loop indica appunto la presenza di un circuito ciclico che determina una **ripetizione**. Le due strategie corrispondono all'utilizzo dei due tipi di informazioni (afferente o efferente): il controllo a circuito chiuso indica che durante il movimento (online) viene costantemente controllato un segnale di errore. Il movimento viene monitorato durante la sua esecuzione e quindi continuamente viene acquisito il feedback sensoriale su dov'è la mano rispetto la tazza, in un dato momento e durante il movimento, che quindi viene corretto mediante questa strategia.

Vantaggio: richiede poco investimento cognitivo (anche un organismo semplicissimo, unicellulare, può muoversi utilizzando il gradiente di una sostanza nutritiva – random walk).

Svantaggio: è lento perché richiede di completare dei cicli di acquisizione dell'informazione afferente e dunque il riaggiornamento della trasformazione sensomotoria e quindi degli impulsi dati ai muscoli. La lentezza dipende dal tipo di feedback utilizzato: il feedback visivo richiede il tempo che impiega l'informazione dalla retina ad arrivare sulla corteccia in V1 (circa un 150-200 millisecondi). Sembra ci siano alcune aree in cui l'informazione arriva un po' prima come ad esempio MT, in cui si vede un'attivazione prima di V1. Il feedback somatosensoriale sembra sia un po' più veloce (dell'ordine di 60/80 millisecondi dalla presentazione dello stimolo). I tempi dunque non sono affatto nulli e questo rende impossibile svolgere azioni molto veloci.

Data la possibilità di eseguire movimenti molto rapidamente, questa strategia non può essere l'unica possibile e in merito è nato un dibattito il cui quesito è se l'informazione afferente (il feedback sensoriale) fosse necessario per poter guidare i movimenti. Si era provato a rispondere a questa domanda in maniera diretta: prendendo una scimmia e deafferentandola delle fibre nervose che lungo la colonna vertebrale portano l'informazione sensoriale al cervello. È stato così possibile vedere che all'inizio la scimmia deafferentata sembrava mostrasse una **paresi**, non riusciva cioè a guidare i movimenti. In realtà non era così: nei primi esperimenti veniva de-afferentato un braccio alla scimmia, che sembrava paralizzato perché non più usato dalla scimmia. Provando a deafferentare entrambe le braccia è stato possibile invece vedere che la scimmia iniziava ad usare le braccia in un modo *non equivalente* a prima, ma **riusciva comunque a guidare i movimenti anche in assenza di feedback sensoriale**. Ne è un ulteriore esempio quando si dorme ponendo il braccio in alto, riducendo in maniera drastica l'afflusso di sangue al braccio, producendo una sorta di deafferentazione momentanea, con fastidiosa sensazione annessa, che però non esclude la possibilità di movimento.

È possibile muovere gli arti anche in assenza di feedback sensoriale: **strategia a circuito aperto**. Questa strategia presuppone che si acquisisca il feedback sensoriale, ad esempio di natura visiva, e sulla base di questa informazione iniziale si formuli una specie di **programma motorio**: rappresentazione interna di natura predittiva, proiettata verso il futuro, di come si intende muovere l'arto. Una volta che il programma motorio è stato completato viene eseguito senza ulteriori modifiche. Un **programma motorio è una rappresentazione astratta dell'azione da svolgere creata prima di iniziare a muoversi contenente informazioni sulle caratteristiche invarianti del movimento, ovvero ordine, struttura temporale (tempi relativi) e forze relative**.

La differenza fondamentale tra la strategia a circuito chiuso e aperto è che la prima, simile ad un *termostato*, gradualmente permette di spostare un oggetto monitorando continuamente le posizioni durante l'estrinsecarsi del movimento. Nel caso del circuito aperto invece il movimento programmato viene eseguito senza possibilità di intervento una volta avviato (si dice infatti siano di natura *balistica*).

**I movimenti lenti degli occhi sono principalmente guidati a circuito chiuso, mentre i movimenti saccadici sono principalmente guidati a circuito aperto, anche se non completamente.** (Ancora oggi si discute infatti quale sia l'eventuale contributo di feedback sensoriali durante i mov. saccadici)

Attenzione alla distinzione: Circuito chiuso - occhi aperti / Circuito aperto - occhi chiusi.

In un circuito aperto non si possono eseguire azioni molto veloci. Il problema di eseguire un movimento a circuito chiuso invece è l'assunzione di un programma interno del movimento da eseguire per il soggetto, inosservabile dall'esterno (si può solo inferire), da eseguire per raggiungere il risultato. Il programma motorio prevede che venga specificato l'ordine con cui il soggetto eserciterà determinate attivazioni muscolari, con che struttura temporale (quanto verrà attivato un determinato muscolo o un'altro nella sequenza), con quale forza/intensità avverrà l'attivazione, la selezione dell'azione e la scelta dell'arto da impiegare.

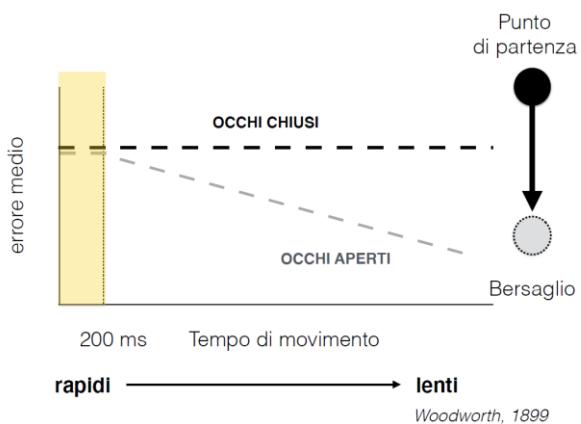
**È risaputo che quando si esegue un movimento vengono utilizzate entrambe le strategie al fine di realizzare il movimento stesso nel miglior modo possibile.** Dato che la programmazione molto precisa è costosa dal punto di vista cognitivo, si è soliti eseguire la prima parte del movimento molto rapidamente, guidata a circuito aperto, sulla base di un programma motorio non troppo preciso dal punto di vista dei parametri cinematici-spaziali. Nella parte finale del movimento invece, addetta all'interazione con l'oggetto cui è diretto il movimento, viene introdotta una fase a circuito chiuso, più lenta e in cui si sfruttano i benefici di un sistema che consente di affinare la parte finale del movimento senza averla dovuta programmare in maniera precisa. Ciò è visibile tipicamente nei movimenti di afferramento.

Alla fine dell'800 **Woodworth** si occupò nella sua tesi di dottorato di *guida dei movimenti volontari*. Con metodi rudimentali ha potuto osservare come viene guidato lo spostamento di una mano in un compito semplice e ha potuto osservare che esistono due modi di guidare la mano: a circuito chiuso e aperto.

Esperimento: viene utilizzato un apparato somigliante ai vecchi sismografi, avente un rullo su cui scorre un foglio di carta e un ago collegato a sensori che rilevano le vibrazioni. L'ago sul rullo dà un tracciato in funzione del tempo (la rotazione del rullo ha una durata nota) delle vibrazioni registrate dai sensori. Woodworth dà uno stilo/ matita in mano al soggetto e chiede di spostare l'oggetto da sinistra a destra, ripetutamente mentre il rullo gira, potendo così ricostruire la distanza percorsa in funzione del tempo. Il compito è di immaginare che il punto di partenza del movimento precedente è il bersaglio del movimento successivo. In questo modo aveva di fatto un bersaglio registrato sul rullo, la traiettoria, la velocità del movimento e soprattutto poteva calcolare un margine di errore.

Adesso l'esperimento è effettuabile attraverso un *touch-monitor*. Il programma in questo caso prevede che venga posto un punto di partenza e un bersaglio. Si chiede al soggetto di porre l'indice sul punto di partenza e di spostarsi sul bersaglio. Si possono fare infinite variazioni, ad esempio facendo scomparire il bersaglio quando il dito viene staccato, togliendo così il feedback sulla posizione, oppure chiedendo di strisciare il dito, saltando ecc.

Un compito del genere si presta ad essere fatto anche con i *movimenti saccadici*: se si ha un sistema di registrazione dei movimenti oculari, si chiede al soggetto di mettere l'occhio sul punto di partenza e si presenta un bersaglio. Il compito è quello di fare una saccade sul bersaglio e può essere eseguito anche da una scimmia.



Woodworth faceva fare questo tipo di compito in due condizioni diverse: **occhi chiusi** e **occhi aperti**. Viene registrato e plottato l' errore medio di un gruppo di soggetti in funzione del tempo di movimento dal punto di partenza a quello di arrivo. Questo grafico è pensabile come un grafico in cui nell'asse delle X si ha la dimensione che va dai movimenti rapidi a quelli lenti. **Woodworth** ha osservato che nella condizione ad occhi chiusi l'errore medio era indipendente dal tempo di movimento. La linea risultante è piatta. Se lo stesso compito viene svolto ad occhi aperti succede che nei movimenti ultrarapidi non c'è un'influenza del tempo di movimento. Più è lento il movimento, più è basso l'errore.

È possibile eseguire dei **movimenti in maniera balistica** (senza utilizzare il feedback sensoriale come nel caso degli *occhi chiusi* – ci si deve basare sull'informazione sensoriale prima di iniziare il movimento). In questo caso anche l'errore si basa sulla valutazione eseguita prima di iniziare il movimento. Se gli occhi sono aperti, più il movimento è lento, più tempo si ha per eseguire dei cicli di aggiustamento del movimento e di conseguenza, più ci si muove lentamente più il movimento è accurato. Woodworth chiamava questa distinzione il “*movimento basato sull'impulso iniziale*” (corrispondente al programma motorio) e il “*controllo corrente*”.

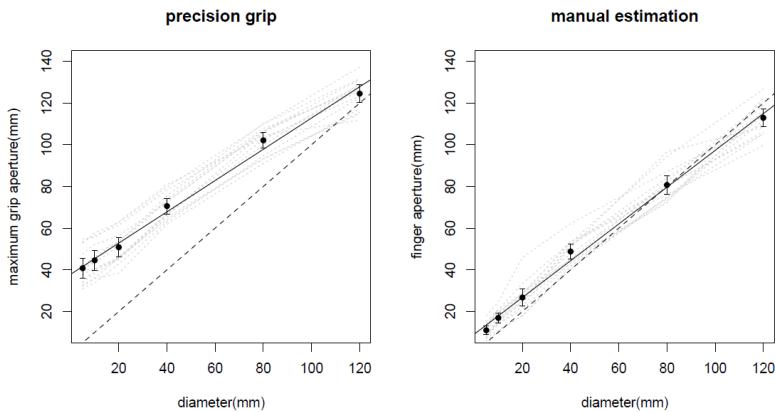
Quanto detto si nota bene anche nell'analisi della cinematica del *movimento di raggiungimento e preensione* e tuttora è presente un dibattito su come davvero sia guidato il movimento di raggiungimento e preensione. La classica caratterizzazione è quella che si deve a , secondo cui il soggetto utilizza una **strategia mista**: si esegue una prima parte del movimento a circuito aperto con lo scopo di portare la mano rapidamente verso l'oggetto da afferrare. In questa fase è possibile osservare che il movimento accelera e una volta raggiunto un picco di velocità rallenta vicino l'oggetto. Viene co-articolato, ossia eseguito nello stesso tempo, un movimento di apertura delle dita, che varia a seconda dell'oggetto da afferrare (es. in base alla grandezza, pesantezza, posizione dell'oggetto nello spazio ecc).

In immagine è possibile osservare il movimento di raggiungimento e preensione attraverso il “pincer grip” (presa a pinza), che è una forma di presa di precisione. C’è una fase molto veloce in cui la mano si apre e arriva in prossimità dell’oggetto. L’apertura della mano è proporzionale alla grandezza dell’oggetto, ma non riflette esattamente la grandezza dell’oggetto. L’apertura supera infatti quella necessaria all’afferramento dell’oggetto (questo costituisce un “margine di sicurezza”). In questo modo non occorre calcolare in maniera precisa quant’è grande l’oggetto in questione, ma è sufficiente un calcolo approssimativo. Questa componente è a circuito aperto. Arrivati al picco di velocità, si arriva al punto in cui le dita raggiungono la massima apertura: la mano si muove velocemente e si apre molto. Segue una fase di decelerazione, in cui la mano si muove in maniera più lenta e in cui il feedback sensoriale acquista maggiore importanza. Questa fase somiglia al “controllo corrente” di Woodworth. In questa fase il soggetto posiziona le dita in maniera precisa su punti selezionati per il contatto con l’oggetto stesso.

Il tutto ha una dipendenza con la velocità con cui si esegue il movimento: maggiore è la velocità del gesto di afferramento, minore è la possibilità di utilizzo del feedback sensoriale per guidare il movimento.



Figura 1.5. Raggiungimento e preensione nella presa di precisione. La mano si solleva durante lo spostamento e le dita si aprono, raggiungendo il punto di massima apertura poco dopo la prima metà del movimento. Sia l’ampiezza della massima apertura sia la durata delle fasi precedente e successiva alla massima apertura dipendono dalle caratteristiche dell’oggetto.



L’immagine si riferisce a dati reali riguardanti la massima apertura delle dita durante due tipi di risposta motoria: a sinistra, la prensione con le dita (**precision grip**), a destra la stima manuale (**manual estimation**). Nel compito di stima manuale non viene richiesto di afferrare l’oggetto ma di indicare con le dita quanto sembra grande. Tale compito viene chiamato in psicofisica “**confronto cross-modale**”, in cui viene chiesto di uguagliare in un canale sensoriale (l’apertura delle dita) la magnitudine di un’altra stimolo (quanto è grande ciò che viene visto).

L’apertura delle dita è plottata in funzione del diametro dell’oggetto (per l’esperimento in questione sono stati utilizzati dei dischetti). Come visibile arriva fino a 12 cm.

Sono entrambi funzioni lineari: **l’apertura delle dita è una funzione lineare della grandezza dell’oggetto**. Come detto da **Stevens**, l’esponente nella funzione psicofisica di un’estensione di una lunghezza non è una funzione potenza, ma diventa una retta. I dati confermano che effettivamente il riconoscimento cross-modale avviene con accuratezza (se l’oggetto è di 20 mm, viene riconosciuto come di 20mm). Le linee tratteggiate sottili grigie nel grafico rappresentano i dati dei singoli soggetti, dunque la variabilità individuale e si nota come sono addensati intorno alla retta di regressione.

La linea solida tratteggiata nera (benchmark) rappresenta il punto di riferimento reale ed è una retta avente intercetta 0 e pendenza 1: equazione  $y=x$ , le due variabili sono perfettamente correlate. Questa retta tratteggiata è pensabile come un’ipotesi nulla, un modello teorico rappresentante un agente ideale perfettamente accurato nel riportare la grandezza dell’oggetto visto. Consente di confrontare i dati reali con l’osservatore ideale.

Osservando la differenza con il compito motorio (manual estimation), che riflette un programma motorio interno, si evidenzia effettivamente una diversità. I dati sono in una retta ma la pendenza e l’intercetta della retta sono diverse da 1 e 0. La pendenza risultante è -1, diversa dall’agente ideale. L’intercetta non è 0 ed è come se il sistema aggiungesse sempre una costante che portasse ad aprire le mani sempre più del necessario.

## **Il comparatore**

Ricordiamo il concetto di **programma motorio**, questo è una rappresentazione astratta creata prima di iniziare a muoversi che contiene informazioni sulle caratteristiche invarianti del movimento, ovvero ordine, struttura temporale (tempi relativi) e forze relative.

È possibile acquisire informazioni preziose sui processi cognitivi studiando pazienti che hanno dei problemi neurologici, prima di vedere uno studio sulla **deafferentazione somatosensoriale** è doveroso specificare le differenze tra *neuropsicologia clinica* e *neuropsicologia cognitiva*.

**Neuropsicologia clinica:** disciplina che si occupa della riabilitazione del paziente che ha un deficit di natura cognitiva.

**Neuropsicologia cognitiva:** studia i pazienti neurologici allo scopo di imparare qualcosa sui processi cognitivi.

Ci sono pazienti che hanno perso le afferenze somatosensoriali, totalmente o in parte, a seguito di malattie al sistema nervoso periferico. Non sono presenti segnali tattili (compresi quelli relativi a cinestesia, propriocezione). In particolare a livello scientifico-mediatico sono conosciuti due pazienti deafferentati, poiché molto studiati. Uno dei due (**I.W.**), protagonista di un libro scritto da un neurologo, ha sofferto la deafferentazione da giovane ed è stato molto aiutato dal neurologo a recuperare/sostituire le informazioni mancanti, riuscendo ad essere uno dei pochi pazienti che sono riusciti a camminare (la difficoltà nella deambulazione è dovuta alla mancanza di informazioni somatosensoriali). I.W. riusciva a camminare e a stare in piedi utilizzando la visione per il controllo della postura.

L'altro paziente (G.L) invece, una donna molto studiata da gruppi francesi, è in sedia a rotelle e non può camminare.

In immagine è possibile osservare una mappatura delle **deafferentazione somatosensoriale** (ossia la sensibilità nei vari distretti corporei in funzione del tipo di stimolo tattile dato). La stima della sensibilità viene eseguita mediante la misurazione delle soglie:

- **Nero:** totale assenza di sensibilità, non si riesce a misurare una soglia;
- **Grigio con righe diagonali:** deficit severo, soglia molto più alta del normale;
- **Grigio con punti:** deficit lieve, soglia poco più alta del normale;
- **Bianco:** normale.

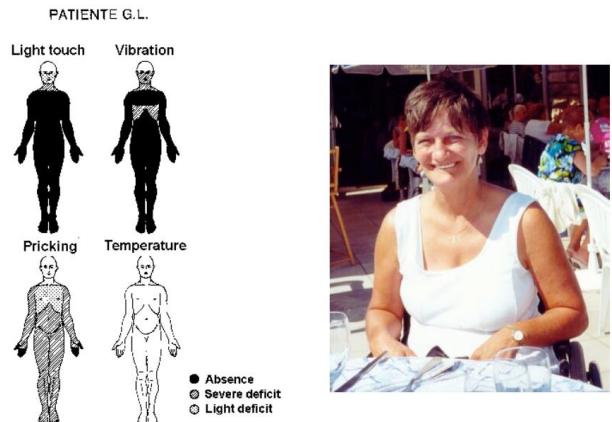
Innanzitutto notiamo come la **sensibilità per la temperatura** è normale, infatti i recettori del caldo e del freddo sono nella pelle, ma il problema della paziente G.L. non è cutaneo, ma si trova a livello delle radici dorsali delle fibre che vanno verso la corteccia, si colloca più o meno a livello della colonna vertebrale. Le vie per la temperatura sono diverse rispetto a quelle degli altri segnali, quindi vengono risparmiate: il soggetto sente caldo o freddo, ma non percepisce la pressione.

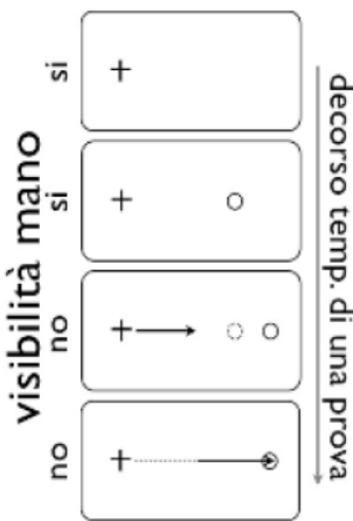
Nel caso del “**light touch**” la sensibilità è praticamente nulla nel corpo, G.L. ha una severa deafferentazione nel collo e lieve in volto (sulla zona dei capelli non è stato probabilmente misurato). Un quadro simile è presente per lo **stimolo vibratorio**. Migliore il quadro, anche se denota un deficit generale piuttosto severo, per il **pricking** (puntura con ago, di natura dolorosa molto lieve).

Questi pazienti, come intuibile, hanno gravi difficoltà nelle attività della vita quotidiana poiché non hanno alcun feedback somatosensoriale a sostegno della consapevolezza e controllo delle proprie azioni (è sempre necessario l'utilizzo della visione).

L'esperimento, di seguito esposto e svolto su G.L., rappresenta una delle dimostrazioni più convincenti dell'esistenza di **modelli interni di movimento di natura predittiva** di come il movimento si svolgerà dopo che è partito.

## deafferentazione somatosensoriale





In immagine è possibile osservare un esperimento svolto con G.L., attraverso l'utilizzo del **paradigma double-step** (paradigma molto utilizzato per lo studio sia dei movimenti manuali, sia dei movimenti saccadici). È una variante del compito di Woodworth, in cui è presente un bersaglio da raggiungere con la mano, cui si aggiunge una seconda fase in cui, dopo che il soggetto è partito, il bersaglio si può spostare (la distanza si può allungare o accorciare). Nel secondo step il movimento deve essere aggiustato per compensare la variazione.

A sinistra lo schema illustra com'è fatta una prova dell'esperimento: i quadri simboleggiano lo schermo di un computer in cui, in funzione del tempo (decorso temporale della singola prova) compare inizialmente un punto di fissazione in cui porre il dito quando si inizia (punto di partenza) e poi compare un bersaglio. Il compito del soggetto è di spostare il dito indice in modo tale da puntarlo sul bersaglio. Nel double-step può capitare che il bersaglio si sposti e, nel caso dell'esempio, si allontana. **Il compito del soggetto, mentre si sposta, è di aggiornare i parametri del movimento in maniera da arrivare al bersaglio.** Da notare che durante questa prova viene controllata la visibilità della mano: all'inizio quando il bersaglio compare la mano è visibile, non appena il movimento inizia la mano non è più visibile.

La situazione è mista perché non vedendo più la mano non è possibile fare un circuito chiuso dal punto di vista visivo. È un circuito chiuso da un punto di vista multisensoriale perché il bersaglio è visibile e si ha l'informazione somatosensoriale su dov'è la mano durante il movimento. Per i *soggetti non deafferentati* si ha questa forma cross-modale di circuito chiuso per aggiornare i parametri del movimento.

G.L., essendo invece de-afferentata, quando si toglie la visione della mano non ha più idea di dove sia la mano. Non può eseguire il circuito multisensoriale, non può confrontare la somatosensazione della posizione della mano con quella vista del bersaglio. Dovrebbe avere dunque problemi di esecuzione del compito. In realtà non ha problemi e la sua prestazione risulta quasi indistinguibile da quella dei soggetti normali.

Nel grafico, a destra sono visibili due tipologie di dati: il picco di velocità e la posizione finale. Osservando la posizione finale (cerchietti) è possibile valutare la prestazione. Nei grafici sono visibili vari livelli della posizione del bersaglio rispetto al punto di partenza e ci sono varie possibilità.

Le tre distanze (vicino – 12; medio – 18; lontano - 24) del bersaglio rispetto al punto di partenza si riferiscono a tre tipi di prove di controllo (single step) cui seguono due prove double step, in cui il bersaglio intermedio si sposta tornando indietro (18 - 12) oppure andando in avanti (18 - 24). Sull'asse delle y c'è la posizione della mano misurata in gradi di angolo visivo. Le tre linee orizzontali nel grafico sono riferite al soggetto ideale. Ad esempio, guardando le prime 3 prove single step, è possibile vedere che i soggetti di controllo sono conformi alla previsione. Osservando il grafico riferito a G.L. non si notano risultati particolarmente discostanti. La differenza risulta nella maggiore variabilità dei dati, tipico del paziente che è meno capace di ripetere lo stesso movimento.

Nelle prove single step, G.L. vede inizialmente la propria mano e successivamente esegue il movimento senza vedere la propria mano. Può aver effettuato una sorta di calcolo a parte dalla primaria osservazione della propria mano e del bersaglio, poi, dato che durante il movimento non ha più la sua afferenza somatosensoriale, la sua risposta è equiparabile. Non si riscontra un bias o errore sistematico.

Nelle prove double-step, in cui vi è uno spostamento da 18 a 12 oppure da 18 a 24, il soggetto dovrebbe arrivare a 12 o 24, e non a 18, essendo il target iniziale. I controlli rispettano l'aspettativa e arrivano alla posizione finale (cerchietti) di 24, compensando perfettamente il double-step. Nel caso 18-12 è presente un leggerissimo “overshoot”: leggera tendenza a non andare abbastanza indietro, compatibilmente all'errore casuale. I soggetti di controllo riescono ad eseguire questa compensazione. I dati di G.L. mostrano che anche lei è spaccata su 24 nella prova 18-24 ed è solo leggermente non abbastanza indietro nella prova 18-12.

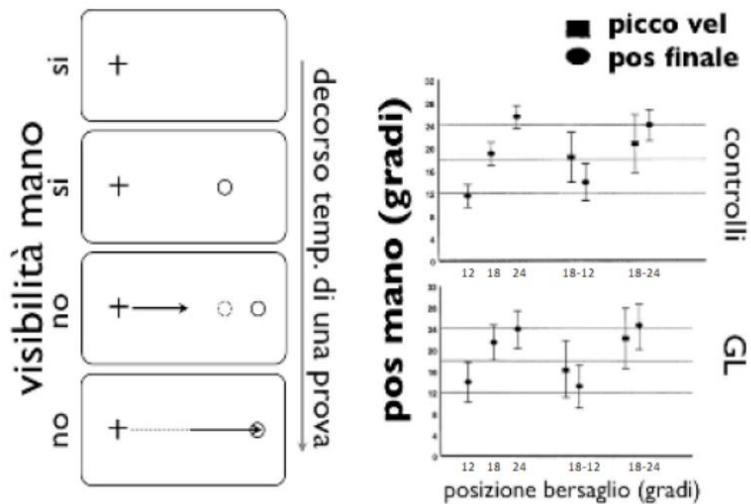
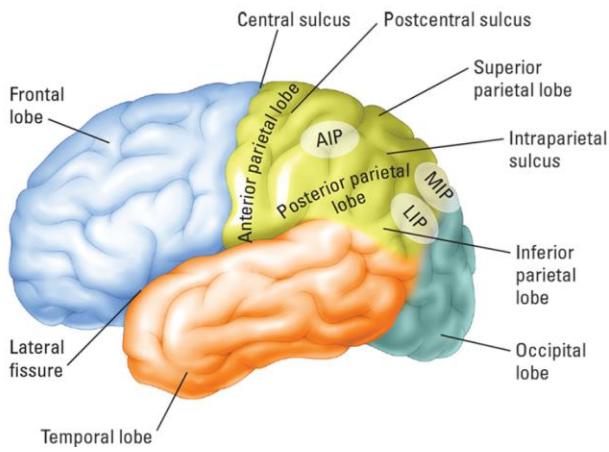


Figura 1.10. A sinistra, le fasi di una prova nel paradigma double-step. Posto il dito sulla posizione iniziale (croce); compare un bersaglio da raggiungere (disco); durante lo spostamento della mano, il bersaglio si sposta; il partecipante deve modificare il movimento già iniziato per raggiungere la nuova posizione. La mano è visibile prima di iniziare il movimento ma diventa invisibile durante il movimento. A destra, i risultati medi relativi a un gruppo di partecipanti senza patologie e alla paziente GL. Le prove in cui il bersaglio è a 12, 18 e 24 gradi sono prove normali. Quelle indicate con 18-12 e 18-24 si riferiscono a situazioni double-step in cui il bersaglio rispettivamente si sposta all'indietro, in avanti come nell'esempio a sinistra (adattato da Bard et al., 1999).



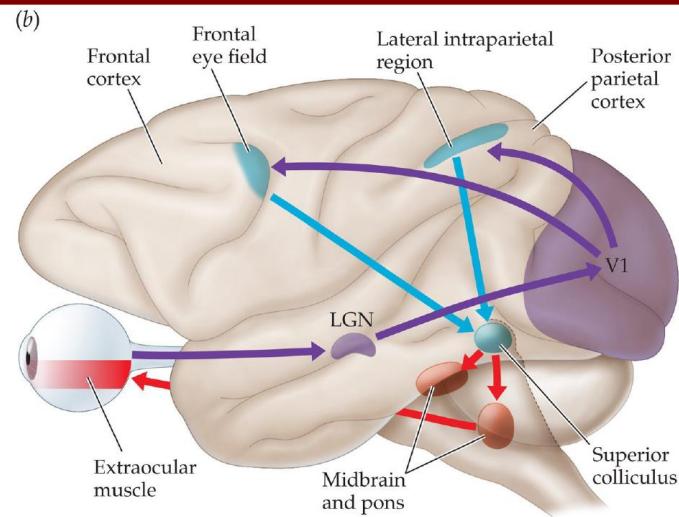
**L'unica spiegazione possibile di come G.L. riesca ad eseguire la compensazione è che, quando esegue i movimenti, non utilizza il feedback di natura somatosensoriale, ma utilizza una rappresentazione interna del movimento di natura predittiva.** Il suo programma motorio specifica come il movimento si svolgerà una volta iniziato. In questo modo **lei può aggiornare il programma motorio di dov'è la sua mano in un certo momento e modificarlo, in maniera predittiva, perché ha previsto dove sarà prima di iniziare a muoversi, riuscendo così a compensare.**

Anche al di fuori dell'ambito dei movimenti saccadici esistono delle prove convincenti dell'esistenza delle rappresentazioni interne del movimento che hanno natura predittiva (come se sapessimo come si svolgerà il movimento prima di eseguirlo e fossimo in grado di fare delle previsioni su quelle che saranno le conseguenze sensoriali di quel movimento).

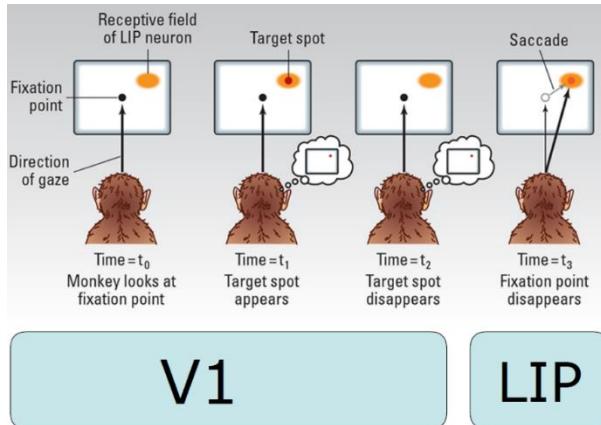
## Come avviene la stabilizzazione del mondo

L'immagine mostra alcune aree presenti nella parte posteriore della corteccia parietale, in particolare l'**area LIP** (intraparietale laterale). Queste aree sembra siano parte di un circuito neurale coinvolto nel **processo di stabilizzazione del mondo** che include l'input visivo, quindi dalla retina al genicolato, alla corteccia visiva primaria (V1). Da V1 il segnale va sia nella corteccia parientale posteriore sia verso i campi frontali degli occhi. Da notare che sia **LIP** che **FEF** inviano un segnale all'indietro verso i collicoli superiori, che a loro volta mandano un segnale di retroazione efferente a livello del mesencefalo verso l'occhio.

Figure 8.16 Muscles of the eye (Part 2)



SENSATION & PERCEPTION 4e, Figure 8.16 (Part 2)  
© 2015 Sinauer Associates, Inc.



Sulla modalità di stabilizzazione del mondo è presente un dibattito. Secondo un'ipotesi, sostenuta da **William James**, *la stabilizzazione del mondo avviene mediante un segnale afferente dei muscoli*. Secondo tale concezione, il soggetto muovendo i muscoli oculomotori genera un segnale afferente disponibile in corteccia che dà informazioni su come si sono mossi gli occhi e permette così una compensazione. Il dibattito si accende con altri neuroscienziati, tra cui **Hermann von Helmholtz**, che sosteneva invece che ciò non era possibile, sulla base di due argomentazioni:

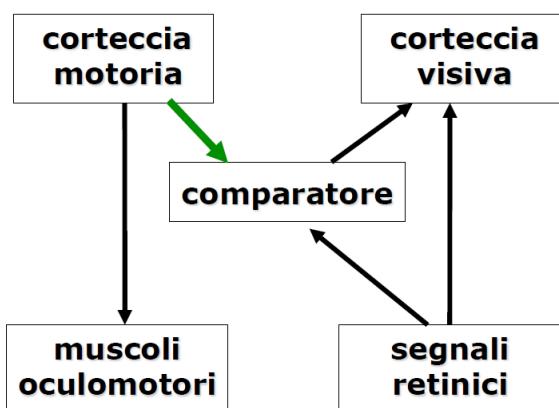
- 1) Il segnale afferente dei muscoli è troppo lento rispetto i movimenti saccadici;
- 2) Osservazioni in cui veniva preso un soggetto e gli veniva temporaneamente paralizzato, uno dei muscoli oculomotori di un occhio, mediante iniezione di curaro. **Helmholtz aveva notato che, nel caso di un soggetto con occhio paralizzato, quando gli viene chiesto di eseguire un movimento saccadico nella direzione impossibilitata, il mondo sembra spostarsi nella direzione opposta.** Normalmente quando l'occhio si sposta sulla retina avviene una proiezione dell'immagine retinica, ma la percezione del mondo rimane stabile. **In questo caso il soggetto vuole fare una saccade, ma questa non avviene, quindi il mondo sulla retina non si sposta ma nella percezione si sposta nella direzione opposta a quella del movimento desiderato.** Questo dimostra che l'afferenza somatosensoriale non è cruciale, ma è presente un **segnale efferente, ossia il fatto di aver deciso di muovere l'occhio**. Nella corteccia e quindi nel circuito, è stata segnalata un'intenzione ed è stato formulato un programma motorio predittivo che poi viene inviato agli occhi e fa eseguire la saccade. La teoria di James deve essere dunque rifiutata.

La conclusione di Helmholtz è che deve esserci una “**sensazione di innervazione**” o, come è stata chiamata poi negli anni 50 da Von Holst, “**copia efferente**”. La copia efferente è una rappresentazione interna del movimento su come si svolgerà una volta iniziato ed è dunque una **rappresentazione di tipo predittivo**.

Prima ancora di eseguire il movimento saccadico, il cervello sa che tipo di trasformazione dovrà aspettarsi dall'input retinico in conseguenza di questo movimento saccadico perché la copia efferente di questo movimento (segnaletico motorio inviato alle cortecce motorie) viene inviata altrove, probabilmente al **collicolo superiore**. Esiste un meccanismo, detto “**il comparatore**”, che ha il ruolo di confrontare la copia efferente del movimento (la modifica attesa dell'informazione retinica) con l'informazione retinica corrente. Quando il soggetto muove l'occhio e sulla sua retina “tutto il mondo” scivola e il cervello è consapevole che tale scivolamento è dovuto al movimento oculare perché lo confronta con la copia efferente, ossia lo scivolamento atteso.

Questo meccanismo permette anche di distinguere tra ciò che rimane stabile e ciò che si sposta. Ciò spiega il motivo per cui il mondo sembra muoversi nella direzione opposta al movimento che il soggetto intende eseguire: se quel che fa il comparatore è “movimento atteso” meno “movimento osservato”, se l'esito è 0 il mondo è fermo, se il movimento atteso è 10 e il movimento osservato è 0 perché l'occhio è paralizzato, allora sottraendo avviene un movimento nella direzione opposta, a -10. Da un punto di vista computazionale il modello può essere rappresentato come segue:

## copia efferente



La corteccia motoria invia un segnale ai muscoli oculomotori, per muovere l'occhio. Allo stesso tempo la retina manda un segnale alla corteccia visiva e al comparatore.

Tuttora il dibattito è aperto sull'eventuale esistenza di *altri meccanismi di stabilizzazione del mondo*. Alcuni ipotizzano un contributo da parte dei segnali afferenti, altri l'esistenza di altre informazioni visive sulla stabilità del mondo utilizzabili, quindi segnali retinici, altri sostengono che in realtà non ci sia nulla da stabilizzare perché il cervello ha una maniera efficiente di calcolare le posizioni degli oggetti ad ogni fissazione e ad ogni fissazione viene ricalcolato.

Quanto visto è un esempio di come il cervello tende a **sopprimere i segnali sensoriali autoprodotti** e questo ha una **funzione adattiva**. Nel caso dei movimenti oculari, i movimenti sulla retina prodotti a causa dei movimenti del proprio occhio non sono infatti degni di interesse. Nel caso si passi il dito su una superficie c'è un pattern di attivazioni sulla pelle non percepito perché si percepisce l'oggetto esterno. Tale

principio per cui **il soggetto lavora su rappresentazioni predittive delle conseguenze sensoriali dei propri movimenti e che una di queste funzioni di queste rappresentazioni predittive è l'annullamento dei segnali sensoriali autoprodotti**, è molto adattivo.

Nello specifico caso della visione e dei movimenti saccadici questo processo si chiama **soppressione saccadica** ovvero una **riduzione della sensibilità durante la saccade** che serve ad **eliminare i motion streak**.

Questo potrebbe essere una delle spiegazioni del motivo per cui non è possibile farsi il solletico da soli. Dal punto di vista delle neuroscienze esistono due tipi di solletico: *knismesi* (sensazione borderline tra piacevole e dolorosa, che si ha quando viene accarezzata leggermente la pelle) e *gargalesi* (che provoca la risata e che non può essere autoprodotto).



i-PERCEPTION

*i-Perception* (2010) volume 1, pages 69–72

a Pion publication 

[dx.doi.org/10.1068/i0387](https://doi.org/10.1068/i0387)

ISSN 2041-6695

[perceptionweb.com/i-perception](http://perceptionweb.com/i-perception)

## How the brain makes the world appear stable

Bruce Bridgeman

Department of Psychology, University of California Santa Cruz, Santa Cruz, CA 95064, USA;  
e-mail: [bruceb@ucsc.edu](mailto:bruceb@ucsc.edu)

Received 7 May 2010, in revised form 9 September 2010; published online 5 November 2010

**Abstract.** Space constancy, the appearance of a stable visual world despite shifts of all visual input with each eye movement, has been explained historically with a compensatory signal (efference copy or corollary discharge) that subtracts the eye movement signal from the retinal image shift accompanying each eye movement. Quantitative measures have shown the signal to be too small and too slow to mediate space constancy unaided. Newer theories discard the compensation idea, instead calibrating vision to each saccadic target.

**Keywords:** vision, efference copy, corollary discharge, space constancy, spatial orientation

Un ulteriore studio su come il cervello rende il mondo stabile:

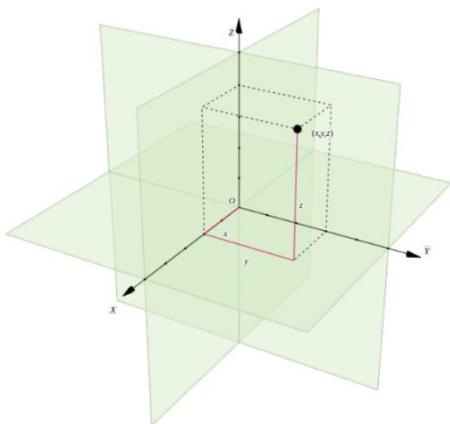
## Lezione 17

### Percezione dello spazio

*La percezione è intrinsecamente spaziale.*

Circa la percezione dello spazio per la **visione**, noi tutti, abbiamo un'esperienza percettiva spaziale cosciente e consapevole, cioè coscientemente riusciamo a percepire un oggetto in una “maniera spaziale”, cioè **collocato nello spazio**. Questo è sicuramente vero per la visione, ma anche per l'**uditore**, infatti, anche attraverso l'uditore **percepiamo i suoni collocati nello spazio**, cioè percepiamo i suoni come provenienti da delle posizioni nello spazio esterno. Anche per la **somatosensazione** vale lo stesso: quando si tocca qualcosa, si acquisiscono informazioni non solo su quell'oggetto, **ma anche dove è collocato nello spazio rispetto al mio corpo**. Per i **sensi chimici** questa capacità è meno spiccata nella nostra specie, ad esempio l'**olfatto**, abbiamo mantenuto una primitiva capacità di individuare nell'ambiente il luogo da dove proviene la fonte di un odore, ma l'olfatto non rappresenta la primaria modalità per la collocazione spaziale. Infine, anche quando si mette del cibo in bocca per assaporarlo, c'è una sorta di **esplorazione tattile orale**, fatta tramite la bocca, che consente di apprezzare le caratteristiche del cibo: forma e resistenza. Tutte queste capacità “somatosensoriali” della bocca influiscono sulla capacità di discriminare la qualità del cibo. ➔ Dunque, **la percezione è intrinsecamente spaziale**.

*Il primo errore circa la comprensione dello spazio: la GEOMETRIA EUCLIDEA.*



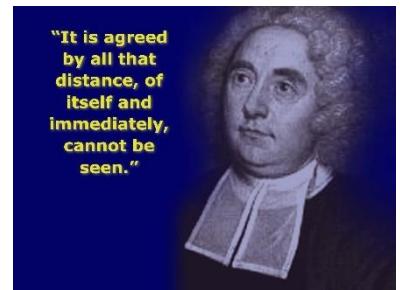
Classicamente lo strumento per lo studio della percezione dello spazio è stata la **geometria euclidea**. Lo studio dello spazio è avvenuto grazie al potentissimo strumento della geometria, che si è dimostrato utilissimo per formalizzare problemi di natura spaziale. Esistono tanti diversi tipi di geometria, ma tradizionalmente è stata utilizzata la **geometria euclidea**, cioè quella geometria basata su due idee:

1. L'oggetto viene identificato attraverso uno schema di riferimento formato da **i tre assi ortogonali, senza dunque possedere un'area**;
2. Di pensare **lo spazio come una specie di contenitore, di volume** che è riempito di **possibili posizioni** che sono i punti nello spazio tridimensionale euclideo.

Questo tipo di geometria **dà un'idea dello spazio altamente astratta**, perché ogni punto **non ha un'area**: corrisponde solamente a una tripletta di coordinate, detto **volume**. Come vedremo invece, dal punto di vista della percezione, lo spazio è qualcosa di completamente diverso. Non dobbiamo pensare lo spazio della percezione come una specie di contenitore astratto, vuoto in cui ci sono delle posizioni così definite. Dunque, l'utilizzo della geometria euclidea nello studio della percezione, dà qualche problema.

## George Berkeley

Quando si introduce lo studio dello spazio con questi presupposti “geometrici” si fa riferimento alle teorie di uno dei primi autori di un trattato sulla visione come processo cognitivo: **George Berkeley**, uno degli esponenti della filosofia empirista inglese del ‘700, autore di un saggio: “*La nuova teoria della visione*”.



**Berkeley pensa lo spazio percettivo come a un singolo punto, nello spazio vuoto.** (Noi cercheremo di confutare questa teoria, non è giusto pensarla in questi termini).

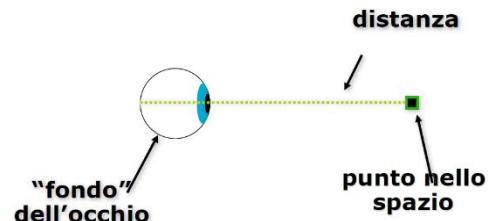
La teoria di Berkeley circa il funzionamento della percezione dello spazio è sintetizzata da questa sua famosa affermazione:

“*It is agreed by all that distance, of itself and immediately, cannot be seen*”, “*Tutti quanti sono d'accordo che la distanza non sia immediata, quindi non può essere vista*”.

La visione risulta essere dunque “*mediata*”, e la diretta conclusione di ciò è che è **impossibile vedere le distanze**. Questa affermazione paradossale è una logica conseguenza della geometria euclidea: proprio perché un oggetto nello spazio è definito solo dalle sue coordinate, cioè è definito da una tripletta di coordinate, senza possedere una vera e propria area, **non è un vero e proprio oggetto, ma semplicemente un punto nello spazio esterno**.

La distanza, dice Berkeley, non è altro che una retta che si estende dal punto nello spazio fino a un corrispondente osservatore, sul fondo dell'occhio sulla retina.

Da questo presupposto si conclude che la distanza **non può essere giudicata oggettivamente**. Non possiamo avere informazioni oggettive sulla distanza, perché se si spostasse un punto nello spazio, da una posizione vicina a una più lontana, cioè se aumentassi o diminuissi la distanza, mantenendolo lungo la stessa direttrice, lungo la stessa retta, **il punto sul fondo dell'occhio su cui cade questo punto rimane sempre lo stesso**.





Quindi la visione non può essere immediata, diretta, ma solo mediata da un giudizio, **da un processo di valutazione cognitiva**. Pertanto, quello che Berkeley sostiene è che la distanza non è un qualcosa che viene visto immediatamente, direttamente sulla base di informazioni contenute nella luce che arriva all'occhio. Secondo Berkeley la distanza è un qualcosa che noi dobbiamo imparare a giudicare, a stimare con un **processo di ragionamento**, come una specie di *ragionamento probabilistico*. Questa idea ha influenzato molti altri come ad esempio Helmholtz. **Non è giusto però pensare allo spazio percettivo utilizzando il modello della geometria euclidea, lo spazio non è coerente.**

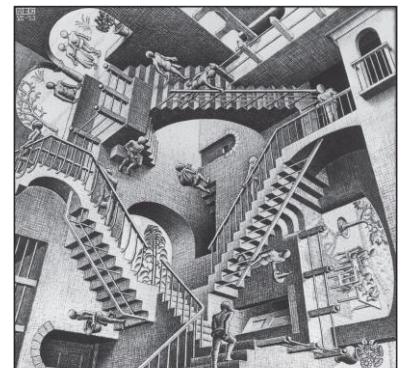
Un'altra idea sicuramente sbagliata circa la nostra intuizione dello spazio, è che lo spazio tridimensionale che noi percepiamo sia coerente dal punto di vista geometrico. Noi siamo sensibili a proprietà degli stimoli che ci dicono qualcosa sulla struttura tridimensionale dello spazio, o alle varie posizioni relative degli stimoli tra loro nello spazio. Ma non abbiamo affatto bisogno di far sì che questo sia inserito in un modello coerente, non contraddittorio, complessivo, di una scena o un ambiente tridimensionale.

Ci sono due tipologie di dati che confermano che **non abbiamo un'unica rappresentazione dello spazio**:

- Il primo dato è offerto dal lavoro di **Escher**, un artista grafico olandese: **le scale impossibili**.

Escher ha lavorato tantissimo sull'idea di presentare degli **spazi impossibili** dal punto di vista della coerenza globale, ma connessi tra di loro in un'unica rappresentazione.

Se guardiamo questa immagine *a livello locale* la scena è del tutto coerente, tuttavia se si cerca di seguire come si sviluppa la *scala globalmente* ci si rende conto che una cosa del genere non potrebbe esistere nello spazio tridimensionale. Questo sembra suggerire che non abbiamo nella testa un'immagine globale geometricamente coerente dello spazio, **semai abbiamo molte rappresentazioni dello spazio che sono funzionali a diverse modalità di interazione con l'ambiente**. ➔ Questi dati sono molto coerenti con l'idea che non abbiamo un'unica rappresentazione dello spazio.



SENSATION & PERCEPTION 4e, Figure 6.4

2. Altri dati molto convincenti di questa teoria vengono dallo studio dei pazienti con il **neglect** o sindrome di **eminegligenza spaziale unilaterale**, che colpisce un emicampo visivo. Di solito si verifica per l'emicampo di sinistra, perché il neglect è la conseguenza di un danno all'emisfero di destra, in alcune aree specifiche. **L'emeinegligenza spaziale unilaterale non è un problema di percezione, ma di cognizione**, in quanto non è una **emianopsia** (forma di cecità corticale che implica di non riuscire a vedere nulla nell'emicampo contralaterale alla lesione), infatti **il paziente è capace di vedere l'emicampo di sinistra, solo che non riesce ad elaborarlo**.

Ad esempio, l'**emianopsia sinistra** è data da una lesione alla V1 dell'emisfero destro, questa lesione si manifesterà col fatto che il soggetto non riesce a misurare delle soglie di detezione o delle soglie di sensibilità al contrasto. Il paziente non riesce a misurare nulla nell'emicampo di sinistra, è come se lì non ci fosse nulla, non ci fosse più sensibilità. Invece nell'**emeinegligenza spaziale unilaterale** potenzialmente c'è la possibilità di notare gli stimoli presentati nell'emicampo affetto, **ma vengono trascurati** ed è per questo che viene definito *neglect*, ovvero "negligenza". Il neglect è un tipo di problematica legata alla capacità di **prestare attenzione a quello che accade nell'emicampo affetto dal neglect**, che di solito è quello di sinistra.

Il neglect sembrerebbe essere legato ad un problema che riguarda **il meccanismo che lega percezione, attenzione e coscienza**. Ossia al fatto che quando si processa dal punto di vista percettivo uno stimolo, questo viene in qualche maniera ulteriormente manipolato, processato dall'attenzione che produce come risultato che lo stimolo possa entrare a far parte dell'esperienza cosciente. Inoltre, è stata dimostrata una **dissociazione** tra neglect per lo *spazio vicino* al corpo e neglect per lo *spazio lontano* dal corpo. **Questo suggerisce che queste due parti del nostro ambiente, vicino e lontano dal corpo, sono rappresentate in maniera separata**: infatti si può avere un danno neurologico che danneggia una di queste rappresentazioni, ma lascia libera l'altra.

Questa conseguenza sembra essere strettamente legata anche a degli aspetti di natura multisensoriali differenti e all'interazione tra percezione e azione. Infatti, nello *spazio molto vicino al corpo* (distanza data dal braccio) ciò che conta è l'**interazione tra la vista e la somatosensazione**. Nello *spazio più lontano* invece l'esplorazione visiva presuppone l'utilizzo dei movimenti oculari e altri specifici meccanismi di percezione dello spazio e ovviamente la somatosensazione è poco importante. Contano di più la locomozione e l'udito, quest'ultimo è un ottimo supplemento di informazione di oggetti che sono distanti dal proprio corpo.

## Halligan e Marshall (1991)

Questi ricercatori inglesi, si sono imbattuti in un paziente con neglect abilissimo a giocare a frecce e si sono chiesti come fosse possibile che riuscisse a giocare a frecce presentando un neglect. Se appunto c'era un problema nella locazione dell'attenzione nella parte di sinistra del bersaglio, in linea teorica non sarebbe stato in grado di mirare correttamente.

Questo paziente aveva un *neglect classico*: alla richiesta di descrivere gli oggetti in una stanza, lui riportava tipicamente solo gli oggetti nella parte destra della stanza e non quelli presenti nella parte sinistra. Un altro test classico è chiedere ad un paziente di disegnare un orologio o di copiare un orologio: il paziente di solito tende a disegnare un orologio tutto deformato, spostato verso destra, magari disegna anche un orologio regolare, ma mette tutti i numeri da 1 a 12 tutti nella parte destra.

Halligan e Marshall perciò hanno confrontato gli errori commessi dal paziente sia nello spazio peripersonale, che in quello lontano. Per fare questo hanno utilizzato il **test di bisezione di linee** che consiste nel segnare a metà un segmento orizzontale, indicare con un tratto di penna il centro di questo segmento. Pazienti con neglect **tendono a sbagliare nell'indicare il centro del segmento**: si spostano molto più a destra del centro obiettivo. Immaginate di avere un segmento di 20 cm. Si dovrebbe rispondere tracciando un tratto di penna a metà cioè a 10 cm, mentre un paziente con neglect traccia il segmento ad esempio a 15 o 16 cm. Quello che accade presumibilmente è che a causa del neglect la parte di sinistra del segmento non entra a pieno titolo nella coscienza, per cui nell'esperienza è come se non ci fosse.

In questa immagine sono riportati i dati riguardanti le due condizioni del compito di bisezione:

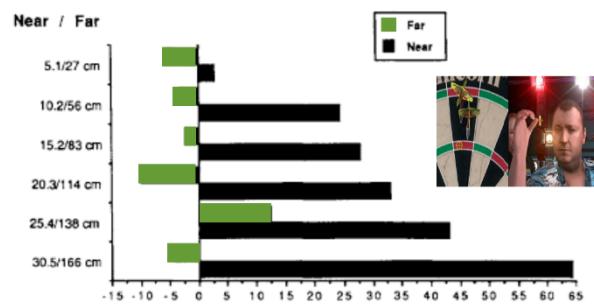


FIG. 1.17 A comparison of neglect for near and far space on line bisection.

- **Far, lontano (verde)** → venivano utilizzati dei tabelloni bianchi, con disegnato sopra dei segmenti molto più grandi, collocati più distanti e il paziente utilizza un puntatore laser;
- **Near, vicino (nero)** → classico compito: foglio di carta con un segmento in cui si indica il centro con la matita.

Innanzitutto questo grafico è un diagramma a barre, e queste rappresentano un numero, tale numero l'**errore di bisezione**. I numeri positivi sono rappresentati verso destra lungo l'asse, il negativo verso la sinistra dell'asse, positivo vuol dire neglect, mentre negativo rappresenta l'errore nella direzione opposta.

## fonti di informazione sulle distanze



L'altra cosa importante da capire è che sull'asse delle Y c'è inusualmente la variabile indipendente, mentre l'errore (variabile dipendente) è sull'asse delle X. Quindi la variabile indipendente è sull'asse verticale e non su quella orizzontale (Y e non X). La variabile indipendente è indicata da una coppia di segmenti (vicino e lontano) dati da due numeri separati da una “/” . Questi numeri rappresentano le lunghezze in centimetri dei due segmenti accoppiati, (es.: 5.1 cm e 27 cm). Il primo numero si riferisce al segmento vicino (che era di 5.1 cm), il secondo al corrispondente segmento lontano (27 cm), questo perché nella condizione lontano avevano utilizzato dei segmenti più lunghi, in maniera da compensare esattamente la riduzione ottica, il rimpicciolimento sulla retina, dovuta alla distanza.

Dunque il segmento vicino era fisicamente più piccolo di quello lontano, ma sulla retina del partecipante la dimensione era uguale: **l'informazione sulla lunghezza retinica era la stessa, perché la lunghezza del segmento sulla retina era sempre la stessa!**

[Sarebbe interessante provare a calcolare il rapporto delle due grandezze es. ( 5.1/27) oppure ( 10.2/56) approssimativamente è circa 1 su 5. Il rapporto è costante, perché il la distanza del segmento sulla retina è fissa.] **E' importante capire che questi segmenti sono tutti equivalenti sulla retina, ma in realtà diversi dal punto di vista della lunghezza fisica** (perché sono collocati vicino o lontani dal soggetto).

Sul grafico si vede chiaramente che c'è una differenza abissale tra il nero (near) e il verde (far). L'errore dello spazio vicino (nero), è chiaramente sintomatico del neglect, ossia sempre nella direzione di destra. Inoltre **l'errore diventa sempre più grande, quanto è più grande il segmento da bisecare, quindi c'è proporzionalità diretta tra errore e lunghezza del segmento.**

La diretta conseguenza di questo è che se il segmento è molto piccolo possono essere commessi errori di alcuni centimetri, mentre se il segmento è grande l'errore può essere molto più evidente.

Per esempio nel segmento di 30,5 cm, il partecipante invece di mettersi sul centro esatto, cioè 15. 25, si metteva sul centro più altri 5 o 6 cm, quindi su 21.75 centimetri (tanto vicino alla fine del segmento). I numeri sull'asse delle X sono millimetri e non centimetri.

Nello spazio lontano (verde) invece l'errore tende ad essere leggermente nell'altra direzione, quella di sinistra, eccetto che in un caso che è spostato un po' verso destra (25,4/138 cm).

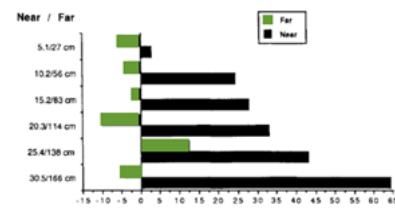


FIG. 1.17 A comparison of neglect for near and far space on fine bisection.

**Quindi il paziente manifesta una chiara dissociazione fra neglect nello spazio vicino e neglect nello spazio lontano.** Questo risultato è proprio coerente con l'idea che **non c'è un'unica interpretazione dello spazio.** Torneremo sul tema della presenza di diversi meccanismi per la percezione, quando, finita la percezione dello spazio visivo, vedremo la percezione dello spazio acustico, comunque dal punto di vista visivo non c'è dubbio che esistono molti meccanismi diversi che ci consentono di rilevare informazioni sulle distanze. Quello che ci chiediamo ora è quali sono i *meccanismi neurali* che sembrano essere responsabili dell'elaborazione di questo tipo di informazione.

## **Classificazione delle fonti di informazioni nella percezione dello spazio**

Esistono diverse modalità in cui possiamo classificare le fonti di informazioni sulle distanze:

### **1. Distinzione tradizionale duale**

Una maniera molto semplice è distinguere tra informazioni che richiedono l'uso di due occhi, ossia **binoculari** e informazioni che richiedono l'uso di un occhio solo, **monoculari** (informazioni **monoculari vs. binoculari**). Siamo in grado di vedere la distanza e di percepire lo spazio anche con un occhio solo, non è necessario utilizzarne due, anche se la fenomenologia dello spazio percepito binocularmente, soprattutto quello vicino, possiede delle caratteristiche particolari come la **stereopsi** (percezione della tridimensionalità).

### **2. Distinzione tripartita**

Un modo in cui è possibile classificare le diverse fonti di percezione visiva dello spazio è attraverso una *classificazione tripartita*:

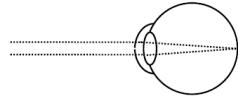
1. Fonti di informazioni legati ad aspetti **oculomotori** (accomodazione e convergenza);
2. Fonti di informazioni legati alla **parallasse** (monoculare e binoculare);
3. Fonti di informazioni di natura **pittorica** (ce n'è una lunga lista accumunata dal fatto che questo tipo di informazioni non richiedono né il movimento, né l'utilizzo dei due occhi. ES: volume tridimensionale, prospettiva..).

Analizzando la classificazione tripartita delle informazioni per la percezione dello spazio scopriremo i meccanismi che sottendono a queste tipologie di fonti e che ci permetteranno di rispondere alla domanda di Berkley, ovvero **come avviene la rilevazione delle informazioni sulla distanza**.

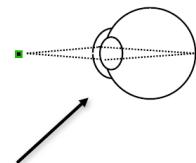
L'**accomodazione** è quel meccanismo per cui il *cristallino*, cioè la lente dell'occhio, può cambiare di forma. Cambiando di forma, cambia anche il suo **potere di rifrazione**, quindi è in grado di piegare di più o di meno i raggi luminosi all'ingresso.

Questo è uno dei meccanismi attraverso i quali l'evoluzione ha permesso la risoluzione del “*mettere a fuoco le cose*”, cioè la visione ad alta risoluzione, su oggetti più lontani o più vicini.

**oggetto  
lontano**



**oggetto  
vicino**



**accomodazione**

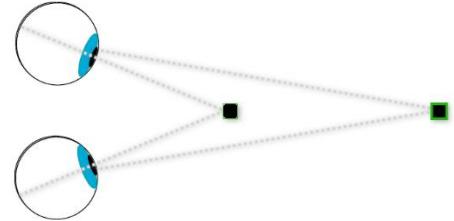
### **Dal punto di vista della percezione delle distanze cosa c'entra questa messa a fuoco?**

L'accomodazione è una forma di lavoro muscolare: il *muscolo ciliato* si rilassa e di conseguenza il cristallino che ha una certa elasticità, cambia la sua forma, la sua curvatura, aumentando in questa maniera il suo potere di rifrazione. Il motivo per cui si diventa *presbiti* infatti è perché invecchiando tutti i tessuti diventano un po' più rigidi, tra cui anche il cristallino, che perdendo di flessibilità non riesce più a cambiare forma in maniera efficiente, perdendo dunque la capacità di “accomodare”. La messa a fuoco è fondamentalmente un lavoro muscolare, un qualcosa che dipende da un'attivazione dei muscoli ciliati. Tale informazione risulta essere di tipo afferente (muscolare), cioè un'informazione che il cervello potrebbe avere sullo base dello stato di accomodazione muscolare dell'occhio e quindi implicitamente sulla **distanza a cui è l'oggetto**.

Ad esempio se si aumenta l'accomodazione con un lavoro muscolare dell'occhio, il cervello riceve implicitamente un'informazione sulla distanza dell'oggetto, ossia che si sta avvicinando. Questo meccanismo può, in certi casi, **contribuire all'impressione di tridimensionalità**, o interagire soprattutto con altri meccanismi. Lo vedremo quando parleremo dello *stereoscopio*, uno strumento che permette di produrre delle disparità binoculari, ossia una parallasse binoculare, cioè una forma di visione delle distanze utilizzando due occhi, in cui l'accomodazione, può contribuire a favorire oppure ostacolare la visione binoculare. In generale però diciamo che **questo meccanismo di percezione delle distanze non è molto preciso** e non è neanche straordinariamente utile al di là di oggetti che sono molto vicini. L'accomodazione per la percezione della distanza è importante più o meno fino alla distanza del proprio braccio o poco più.

Quando si perde la capacità di accomodare, bisogna allontanare l'oggetto sfuocato, infatti gli occhiali da presbite si usano per leggere, non per guardare le cose distanti (dove il processo di accomodazione non è importante).

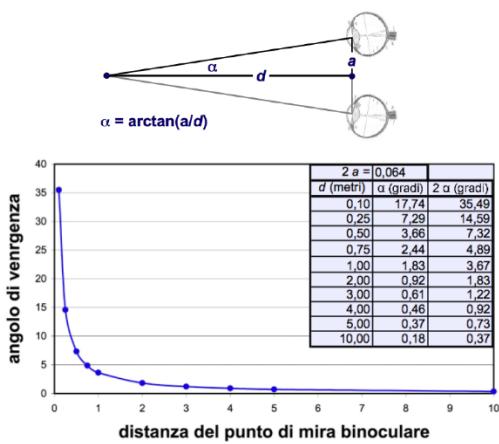
**La convergenza** è un meccanismo binoculare: quando un oggetto si avvicina, gli occhi, per mantenerlo sulle fovee, devono aumentare il grado dell'angolo formato dalle loro linee di mira. Viceversa, se l'oggetto si allontana questo angolo deve diminuire.



Anche la convergenza così come l'accomodazione è un **lavoro muscolare**: c'è un informazione afferente da parte del *muscolo oculomotore* di quanto gli occhi convergono o meno, che appunto dà l'input al cervello sulla **distanza** dell'oggetto (vicino o lontano).

Questo corrisponde all'idea di Berkeley quando diceva di “imparare a *giudicare* le distanze”, aveva in mente proprio questo concetto ossia di giudicare attraverso informazioni che provengono dai muscoli: ad esempio tocchiamo gli oggetti, e toccandoli ci facciamo un'idea di quanto siano distanti da noi (così avviene anche per quanto riguarda gli stimoli visivi). Si potrebbe parlare di una forma di apprendimento associativo, in cui il tatto insegnava alla vista a valutare le distanze. La stessa cosa accade per la convergenza: **impariamo ad associare determinati stimoli visivi per riuscire a stimare la distanza**.

Tuttavia, anche per la convergenza, come per l'accomodazione, si può fare un ragionamento più preciso sulla relazione tra lavoro muscolare e percezione della distanza: l'informazione che potenzialmente si può ricavare è di scarsissima utilità al di là di distanze ravvicinate al corpo.



In questa immagine è raffigurata la **trigonometria della convergenza**: se si fissa un punto nello spazio, i segmenti che si estendono da questo punto alle due fovee degli occhi formeranno un angolo, che è appunto **l'angolo di convergenza** ( $\alpha$  per ogni occhio).

Questo angolo è dato dalla formula in cui compare la funzione arcotangente, ossia il rapporto fra la distanza interoculare ( $a$ ) e la distanza tra il punto intermedio tra i due occhi e il punto stesso nello spazio esterno ( $d$ ).

Alfa è l'angolo di convergenza, ed è l'arcotangente di  $a/d$ :  $\alpha = \arctan(\frac{a}{d})$ .

Questa è una semplice relazione trigonometrica, che consente di calcolare in funzione della distanza fra gli occhi e l'oggetto esterno fissato binocularmente, quanto è grande l'angolo di divergenza e quindi quanta potenziale informazione sulla distanza c'è nel lavoro muscolare di convergenza degli occhi.

Come si vede dalla funzione rappresentata dalla curva blu quando l'oggetto è molto *vicino* al corpo (su un *range* che corrisponde più o meno alla distanza dal braccio cioè 50-60 cm), l'angolo di divergenza cambia dimensione notevolmente, (da 35 a circa 5 gradi). Quindi in una distanza di 10-20 cm c'è una grande possibilità di distinzione dell'angolo di convergenza. Tuttavia **allontanandosi** non cambia quasi più niente, ci sono delle differenze molto piccole che tendono ad appiattirsi sempre di più.

**Possiamo concludere che i due indici oculomotori (accomodazione e convergenza), hanno un ruolo nella percezione delle distanze, ma solo in un *range* di distanza molto limitato: molto vicino al corpo. Inoltre le fonti di informazioni oculomotorie sono legate a segnali afferenti a livello muscolare dei muscoli che permettono il movimento dell'occhio.**

### **Parallasse di movimento e parallasse binoculare**

Sono invece importantissime e fondamentali nella **percezione della tridimensionalità** e della **distanza** le fonti di informazioni basate sulla **parallasse**. Sono importanti per il fatto che manipolando queste fonti di informazioni è possibile produrre **stereopsi**, ossia la **percezione della tridimensionalità** come un qualcosa di realisticamente solido e volumetrico (avere la sensazione che l'oggetto che si sta vedendo si possa afferrare, toccare, in quanto ha una sua materialità).

Infatti, il termine stereopsi, dal greco: *opsein* → “vedere” e *stereo*, un aggettivo che non vuol dire contrariamente a quanto si pensi “doppio”, ma “solido”. È vero che uno dei meccanismi per ottenere il fenomeno della stereopsi implica di guardare con due occhi (così anche per la *stereofonia* → ascoltare i suoni con due orecchie). Però *stereo* in greco non vuol dire doppio, bensì **solido, dotato di volume**. Quindi stereopsi vuol dire “*visione solida*” ossia di oggetti che sembrano dotati di un volume, realisticamente tridimensionali. Anche *stereofonia* vuol dire percezione del suono collocato in un ambiente tridimensionale, come se venisse da dei punti dello spazio ben definiti.

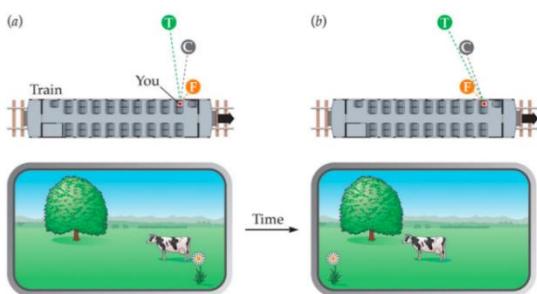
La stereopsi si può avere anche guardando con un occhio solo, se nello stimolo è presente una **parallasse di movimento**. Si può avere l'impressione di vedere un oggetto solido, tridimensionale anche guardando con un occhio solo, purché ci sia movimento. Ci sono anche dei *casi paradossali* di stereopsi che si possono ottenere senza movimento e con un solo occhio, attraverso gli **indicatori pittorici**.

### **Che cos'è la parallasse?**

È un termine che viene dall'astronomia. **La parallasse è lo spostamento di qualcosa che viene registrato per via ottica, di solito in funzione del passaggio del tempo.** Ad esempio, guardando il cielo stellato con un telescopio la parallasse è quel qualcosa per cui le stelle cambiano di posizione relativa fra un momento della notte e un altro. Sulla base di questo facendo della trigonometria si può calcolare la distanza fra le stelle o la distanza delle stelle dall'osservatore.

La parallasse di movimento nella visione segue lo stesso principio: immaginate un osservatore monoculare, il quale spostandosi nell'ambiente, occupa nel tempo successivamente due punti di vista diversi. Ciò che accade è che le proiezioni degli oggetti e dell'ambiente sulla retina dell'osservatore occuperanno **posizioni retinotopiche diverse**. Questo cambiamento di posizione rispetto al tempo è una parallasse di movimento, dunque se gli oggetti sono statici e il soggetto è in movimento, il cambio di posizione degli oggetti è solo funzione della posizione nello spazio dell'osservatore.

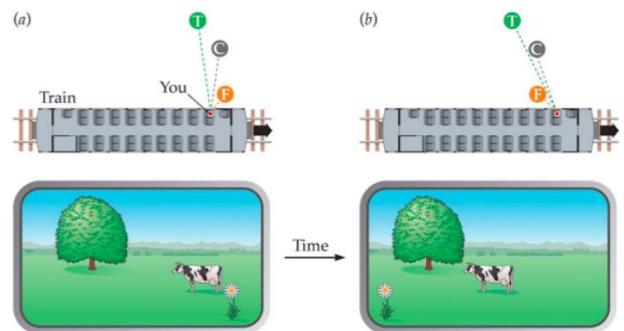
È possibile attraverso la trigonometria o anche capendo le regole di funzionamento della parallasse, ricostruire informazioni sulle **posizioni relative**: durante un movimento vi sono relazioni fisse tra la distanza di oggetti fra di loro (posizioni relative). Queste regole fisse permettono sia la stereopsi che la percezione delle distanze di un oggetto.



In immagine vediamo la rappresentazione del principio che sta dietro la parallasse di movimento.

In questo caso, invece di pensare ad un osservatore in movimento, i cui oggetti a loro volta si muovono sulla sua retina, pensiamo una situazione più semplice. Immaginiamo un *osservatore su un treno in movimento* da sinistra verso destra, che guarda fuori dal finestrino. Potremmo applicare a ciò che si vede dal finestrino lo stesso principio della retina: è come uno schermo sul quale si proiettano gli oggetti che sono sul mondo esterno. Gli oggetti nel mondo esterno sono tre: una **mucca** (*cow-c*), un **albero** (*tree-t*) e un **fiore** (*flower-f*).

L'obiettivo mentre il treno si sposta è di fissare la mucca. L'osservatore è fermo sul treno, quindi la mucca si sposterà mano a mano rispetto al finestrino. Visto che il treno va verso destra, la mucca si sposta verso sinistra. Se invece del finestrino considerassimo la proiezione sulla retina si creerebbe esattamente la stessa situazione: per la parallasse di movimento mantenendo comunque sempre un punto di fissazione (la mucca), non cambia molto se è l'osservatore a muoversi o se questo rimane fisso e a spostarsi sono gli oggetti (proiezioni diverse sulla retina). Sia che l'osservatore mantenga l'occhio sempre nella stessa direzione, quindi è la fissazione che si sposta, oppure che lo stesso mantenga la fissazione sulla mucca, quindi spostandosi in avanti è l'occhio che ruota (che mi sposti io o che si spostino gli oggetti fissati) non è una questione cruciale rispetto al problema della parallasse di movimento.



La parallasse di movimento è un **cambio di posizioni relative tra gli oggetti osservati non rispetto al finestrino, cioè non rispetto all'osservatore**. Rispetto al finestrino tutti gli oggetti (l'albero, il fiore, la mucca) si spostano nella direzione opposta al movimento del treno, ciò che cambia durante il movimento sono le **posizioni relative tra gli oggetti** stessi che si spostano in modo diverso rispetto all'**oggetto di fissazione** (la mucca). Quindi la parallasse di movimento è un cambio di posizioni relative tra gli oggetti osservati rispetto al **punto di fissazione**, è un movimento relativo in **coordinate retinotopiche**.

Questi spostamenti delle posizioni relative tra gli oggetti rispetto al punto di fissazione obbediscono a due principi molto semplici e chiari:

- 1) La prima regola:** se gli *oggetti* sono *vicini* rispetto al punto di fissazione (c) si spostano nel campo visivo dell'osservatore nella *direzione opposta* al movimento.

Ad esempio → il treno va da sinistra verso destra, quindi il fiore (f) che prima era a destra della mucca (c) ora è a sinistra della stessa, lo spostamento sulla retina è nella direzione opposta al movimento dell'osservatore. Mentre per l'albero (t) è il contrario, in quanto si muove nella stessa direzione della mucca (c). **Gli oggetti più lontani dal punto di fissazione si spostano nella medesima direzione del movimento dell'osservatore.**

**NB:** l'oggetto si sposta **rispetto al punto di fissazione, non rispetto al movimento all'osservatore**.

**Ricapitolando:** se l'oggetto è lontano (albero) dal punto fissato (mucca), il suo spostamento è nella stessa direzione del movimento; se invece l'oggetto è vicino (fiore) al punto fissato, il suo spostamento è nella direzione opposta a quella del movimento.

- 2) **La seconda regola:** gli *oggetti più vicini* (fiore) all'oggetto fissato (mucca) si spostano più lontani da questo, e gli *oggetti più lontani* (albero) dal punto di fissazione si spostano più vicini all'oggetto fissato.

Queste regole della parallasse di movimento danno un **ordinamento degli oggetti in profondità e dunque anche delle distanze fra gli oggetti**.

La parallasse di movimento non dà informazioni riguardo la **distanza egocentrica, cioè fra la fissazione e l'oggetto fissato, ma dà informazioni su quanto dista l'oggetto fissato, in termini relativi, da un altro oggetto**:

- Sia per la **relazione ordinale**: l'oggetto più vicino alla fissazione si sposta in direzione contraria rispetto al movimento dell'osservatore e l'oggetto più lontano nella stessa direzione del movimento.
- Sia per la **relazione metrica**: l'oggetto più lontano si sposta di meno dell'oggetto più vicino, che si sposta di più dal punto di fissazione.

**L'entità dello spostamento è maggiore, tanto maggiore è la distanza dal punto di fissazione.**

L'albero è un po' più vicino alla mucca di quanto non lo sia il fiore. Essendo un po' più vicino si sposta un po' di meno, ossia disegna un percorso da sinistra verso destra che è meno grande del percorso fatto dal fiore nella direzione opposta. Tutto questo dipende dalla vicinanza o lontananza iniziale dalla mucca. Quindi il percorso del fiore nella direzione opposta nella direzione del movimento è maggiore, talmente maggiore che passa da destra verso sinistra, mentre l'albero rimane sempre a sinistra.

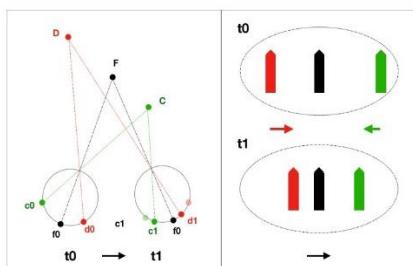
Tutto ciò naturalmente avviene per questo percorso e lasso di tempo... Se invece ci si continua a spostare succederà che il fiore uscirà completamente dal campo visivo, cioè dal finestrino, mentre l'albero passerà dall'altra parte del campo visivo, cioè a destra della mucca. Poi a un certo punto non potremmo neanche più fissare la mucca in quanto anch'essa sarà uscita dal campo visivo.

Quindi la parallasse di movimento è un insieme di movimenti relativi che avvengono sulla retina rispetto a un determinato **oggetto di fissazione**. Ovviamente ciò che viene fissato e proiettato sulla retina non si muove, perché rimane fisso in **fovea**. È tutto il resto a muoversi e tutti questi movimenti relativi hanno un

verso e una velocità, cioè sono dei *vettori orientati* con una determinata lunghezza che ne rappresenta la velocità. Dunque grazie alle regole della parallasse di movimento il verso e la velocità del vettore spostamento dicono la posizione relativa e la distanza relativa degli oggetti fra di loro.

Il problema è che io non vedo solo le distanze relative nel mondo esterno, ma vedo anche l'esperienza percettiva della **distanza egocentrica** ossia quanto dista da me quel determinato oggetto di fissazione (ad esempio la mucca). Questo tipo di informazione non viene data dalla parallasse di movimento.

### Esempio



Immaginiamo di avere tre colori, rosso nero e verde disposti davanti a voi su un tavolo in piedi.

Il pennarello nero (**F**) al centro fa da **punto di fissazione**, l'**oggetto (C)** è più **vicino** e l'**oggetto (D)** è più **lontano** dall'oggetto fissato.

La fissazione avviene con un occhio solo.

Sulla slide sono disegnati due occhi non perché stiamo guardando con due occhi, ma perché immaginiamo di spostare l'occhio, da una posizione a un'altra posizione in due tempi diversi (**t0** e **t1**). Lo spostamento dell'occhio avviene da sinistra verso destra, in cui la fissazione viene mantenuta su f. Ciò significa che il punto f rimane in fovea, sempre nella stessa posizione, sulla retina.

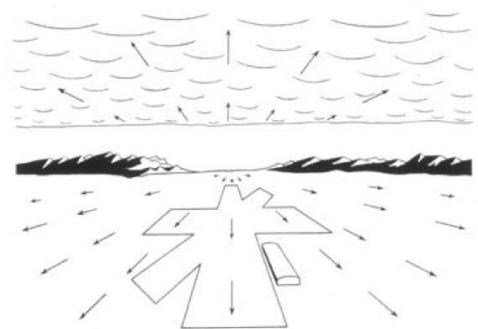
È possibile notare che rispetto ad **F**, il punto **C** cambia di posizione nel tempo, da **t0** a **t1**, si sposta da **c0** a **c1** così come anche il punto **D** cambia di posizione, si sposta da **d0** a **d1**.

**Questi cambi di posizione sono in direzioni opposte, e hanno anche un'estensione sulla retina che è proporzionale alla distanza che **C** e **D** hanno con **F**.**

Quello che dobbiamo notare è che il pennarello rosso si avvicina da una parte mentre il pennarello verde si avvicina dall'altra. Le direzioni sono diverse! → Questa slide fa vedere la differenza tra quello che osservate e quello che di fatto succede sulla retina, che è tutto invertito rispetto a ciò che vedete.

Questo principio generale della parallasse di movimento è stato chiamato in tante maniere nella letteratura psicologica, per esempio spesso viene chiamato **flusso ottico**.

Il flusso ottico è quel campo di velocità, cioè di spostamenti orientati nello spazio, che avviene sulla retina quando si è in movimento verso un determinato punto e si fissa l'oggetto verso il quale ci si sta dirigendo. Fissando e tenendo lo sguardo nella direzione della locomozione, si forma sulla retina un **pattern caratteristico di velocità** che la riempie.



È un **pattern di natura espansiva**: tutto si espande a partire dal punto che si sta fissando che viene chiamato **il fuoco dell'espansione**, quest'ultimo rispetto al resto del campo visivo invece rimane fermo, non si espande. Più ci si allontana rispetto alla fissazione, più il campo di velocità è grande, quindi: in periferia le cose si espandono molto rapidamente, andando verso il centro, verso il punto di fissazione, si espandono molto meno. Questo tipo di situazione è tipica della *locomozione in linea retta*, quando si fissa nella direzione del proprio spostamento, quando si riduce la distanza rispetto a una superficie verso la quale si sta andando. Ad esempio si va verso una parete e si crea un fuoco di espansione che è funzione del cambiamento della distanza relativa tra il soggetto che si avvicina e la parete stessa.

Nella slide viene riportato l'esempio del pilota che deve fare atterrare un aereo. Il pilota deve mantenere allineato l'aereo sulla pista di atterraggio e continuamente osservare la direzione del movimento, ma ha anche bisogno di un giudizio di distanza egocentrica per calcolare quando far appoggiare le ruote del carrello. **Questo tipo di compito è possibile svolgerlo proprio grazie alla percezione di questo tipo di pattern espansivo, chiamato flusso ottico. Il punto di fissazione**, quindi dove non c'è espansione, è **proprio ciò che si utilizza per mantenere l'aereo allineato rispetto alla pista di atterraggio**. Dunque siamo sensibili al *tasso di espansione*, cioè quanto velocemente si espande l'ambiente attorno a noi: **più siamo vicini all'oggetto, più aumenta il tasso di espansione**.

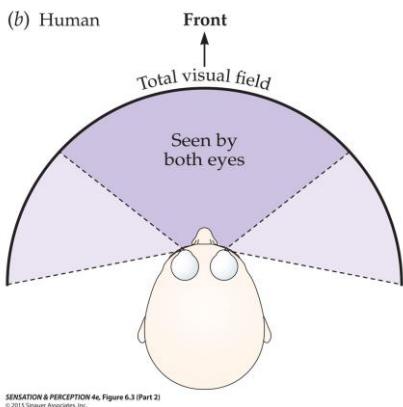
C'è una letteratura molto interessante in relazione a questo tipo di informazioni della parallasse di movimento: per esempio alla nostra capacità di fermarci prima di sbattere contro un ostacolo, alla capacità di guidare un veicolo...ma non solo nell'uomo...un articolo di psicofisica riusciva a mostrare che quando i gabbiani si tuffano dall'alto in acqua molto rapidamente per prendere i pesci hanno il problema di valutare la distanza prima dell'impatto, perché durante il tuffo devono chiudere le ali che altrimenti distruggerebbero visto che entrano in acqua a grandi velocità. **È stato dimostrato che i gabbiani utilizzano il fuoco di natura espansiva nella parallasse di movimento per guidare questo movimento di chiudere le ali.**

**Quando sono lontano il fuoco si espande poco rispetto ai margini, alle periferie, ma quando sono vicino si espande molto di più rispetto ai margini.**

Vedremo un'altra forma di parallasse → la **parallasse binoculare**: invece di vedere un osservatore monoculare, che cambia la propria posizione nel tempo, esamineremo la *situazione binoculare*. La parallasse binoculare ha una geometria quasi uguale, in cui l'osservatore non si muove, ma ha due punti di osservazione simultaneamente disponibili, ossia due occhi e quindi il campo di posizione non è rispetto al tempo ma è rispetto allo spazio, dato che osservo con due occhi simultaneamente.

## Lezione 18

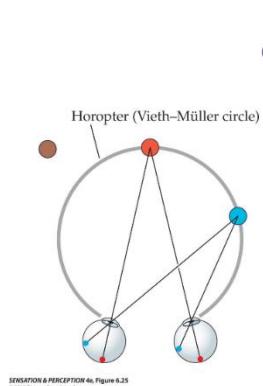
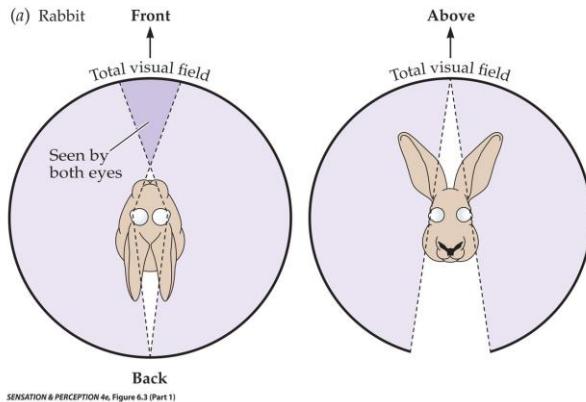
### Stereopsi binoculare



Il campo visivo umano si può suddividere in due porzioni: il **campo visivo binoculare** (visto da entrambi gli occhi) e **due campi monoculari** (questi ultimi più periferici). Questo tipo di organizzazione del campo visivo non lo si riscontra in tutti gli organismi: la *lepre*, ad esempio, presenta gli occhi in una posizione più lateralizzata e possiede quindi un campo visivo molto più ampio del nostro che le consente di vedere in modo efficiente sia lateralmente al corpo, sia posteriormente (quindi una visione pressoché a 360 gradi);

la porzione binoculare, tuttavia, risulta fortemente limitata in questa specie rispetto alla nostra. Le *scimmie* invece possiedono un'organizzazione simile alla nostra. La visione binoculare comporta diversi vantaggi: innanzitutto, nel campo visivo binoculare, vi sono dei punti dello spazio che godono di una particolare proprietà, ossia, **cadono su punti corrispondenti delle due retine**.

Osservando l'esempio fornito: il pallino rosso corrisponde al punto di fissazione, quindi cade, per definizione, sulle due fovee (in punti quindi corrispondenti nelle due retine); tuttavia, anche il pallino azzurro **cade su punti corrispondenti delle due retine**, in quanto, considerando le coordinate retinotopiche, presenta la medesima distanza dalla fovea in entrambi gli occhi.

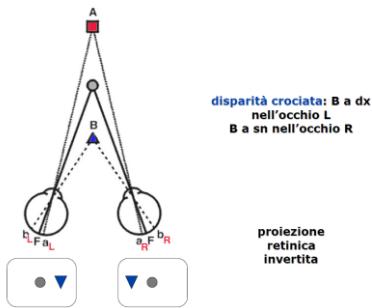


In generale, dato un punto di fissazione più o meno vicino, attraverso lo stesso punto passa sempre un luogo di punti dalla forma approssimativamente circolare, definito "**cerchio di Vieth-Müller**" o **oroptere**, il quale gode della proprietà per cui **tutti i punti che lo costituiscono cadono in punti corrispondenti delle retine** (quindi  $disparità\ retinica = 0$ ). All'infuori dell'oroptere, pertanto, gli oggetti del campo visivo non godono di questa proprietà e, quindi, le loro proiezioni non cadono su punti corrispondenti delle retine (vedi il punto marroncino o viola della slide).

Tale proprietà è definita **disparità retinica** e ne godono i punti che non cadono sull'oroptere, perciò si verifica quando le due proiezioni retiniche non cadono su punti corrispondenti delle due retine. La disparità retinica non è altro che la conseguenza di una parallasse binoculare, quindi di una variazione di posizione non rispetto al *tempo* come nel caso della parallasse di movimento, bensì rispetto allo *spazio*, cioè tra la posizione occupata dall'occhio sinistro e quella occupata dall'occhio destro.

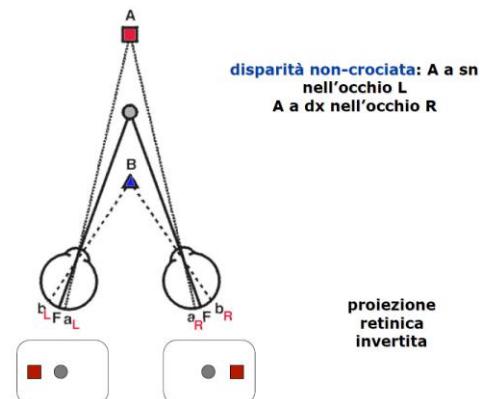
La disparità retinica può essere di due tipi:

1. **Disparità crociata** (*entro l'oroptere*): dato un punto di fissazione, un oggetto interposto tra il soggetto e il punto di fissazione cade nell'emiretina temporale di un occhio e nella nasale di un altro; la proiezione retinica risulta quindi invertita, per cui l'oggetto apparirà a destra rispetto alla fovea dell'occhio sinistro e a sinistra rispetto alla fovea dell'occhio destro.



In altre parole: dato un punto di fissazione (il pallino grigio dell'immagine), il quale cade sull'oroptere e non ha disparità in quanto cade sulla fovea (**F**) in entrambi gli occhi, rispetto al punto di fissazione un oggetto più vicino al punto di vista godrà di una proprietà definita **disparità crociata**. Ciò vuol dire che, facendo riferimento all'immagine la proiezione dell'oggetto **B** (triangolo blu) cade nella retine temporali dei due occhi, però la proiezione retinica è invertita e quindi l'oggetto **B** apparirà a destra della fovea dell'occhio sinistro e a sinistra della fovea dell'occhio destro. Questo tipo di disparità viene chiamata crociata perché se si volesse annullarla bisognerebbe aumentare la convergenza, ovvero **incrociare gli occhi ulteriormente**.

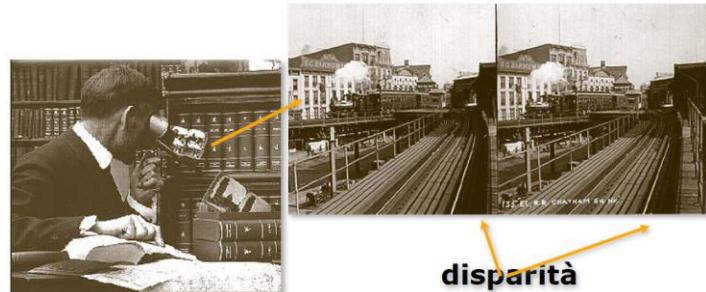
2. **Disparità non crociata** (*oltre l'oroptere*): concerne gli oggetti oltre il punto di fissazione. L'oggetto cade a sinistra rispetto alla fovea nell'occhio sinistro e a destra nell'occhio destro. Per annullare tale disparità occorrerebbe **incrementare la divergenza**, quindi rendere più paralleli gli occhi.



Quanto descritto è analogo a quello che succede nella parallasse del movimento, **per cui gli oggetti più vicini rispetto al punto di fissazione si muovono nella direzione opposta al movimento dell'osservatore, mentre quelli più lontani seguono la medesima direzione**.

La regola generale è la seguente: **gli oggetti più vicini alla fissazione hanno una disparità crociata, mentre gli oggetti più lontani rispetto alla fissazione hanno una disparità non-crociata.**

Le disparità retiniche, conseguenza della parallasse binoculare, **concorrono alla percezione della tridimensionalità**; a tal proposito, è stato fondamentale il contributo di **Charles Weathstone**, il quale, a fine '800, ha inventato lo **stereoscopio**. Tale strumento ha consentito di dimostrare che **non serve necessariamente un'immagine 3D per percepire la tridimensionalità, ma si può ricorrere anche stereoscopio**



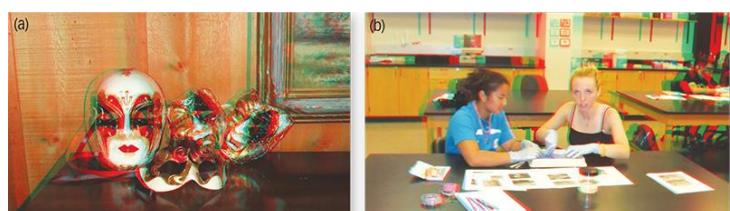
**disparità**

**ad immagini 2D se queste vengono presentate separatamente nei due occhi e se sono in grado di indurre una disparità retinica.** Venivano utilizzate fotografie stereoscopiche (**stereogrammi**), prese da due punti di vista leggermente diversi, che venivano sviluppate e montate nello stereoscopio, il quale serve a presentare separatamente le immagini ai due occhi. Lo stereoscopio 800esco presentava tuttavia un



*problema di accomodazione* (l'accomodazione è quel fenomeno attraverso il quale si cerca di mettere a fuoco delle immagini\fotografie che si trovano a pochi centimetri dagli occhi), gli strumenti moderni di fatto presentano una piccola lente che permette di sopperire a tale problematica.

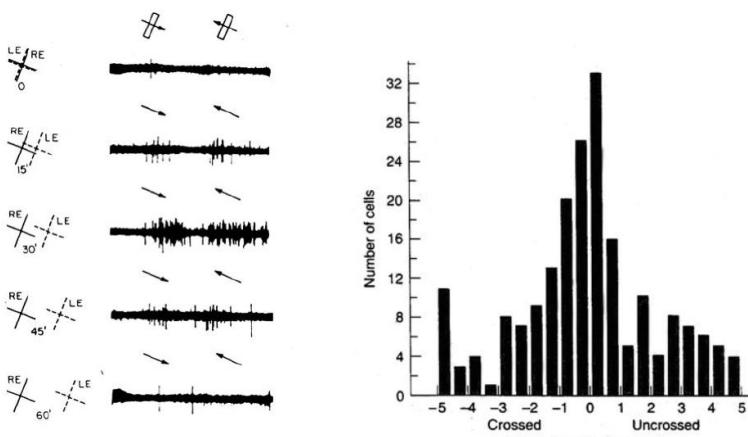
Un'altra tecnica per creare l'illusione di profondità consiste nell'utilizzo degli **stereogrammi anaglifici**, i quali risultano identici, ma sono costituiti da colori



leggermente diversi: in questo caso le due immagini sono sovrapposte in un unico riquadro e, se osservate mediante due filtri di colore complementare l'uno rispetto all'altro, forniscono un'illusione di tridimensionalità.

## Basi neurali della stereopsi

La disparità binoculare può essere informativa relativamente alla posizione di un oggetto nello spazio, posizione relativa all'occhio e al punto di fissazione. La **quantità di disparità retinica, inoltre, permette di conoscere quanto sono lontani gli oggetti (in termini relativi) rispetto al punto di fissazione**; ciò suggerisce che nel sistema visivo potrebbero esistere dei *neuroni selettivi per la disparità*, per cui alcuni risponderanno per una disparità crociata, altri saranno selettivi per la non crociata. Tali neuroni presenterebbero inoltre una **curva di sintonizzazione per la disparità**, ossia un *range di disparità ottimali* che provocano la massima frequenza di scarica della suddetta unità (analogamente a quanto accade per la sintonizzazione all'orientazione o alla velocità del movimento). Queste unità si trovano in **V1** e rappresentano il **punto di integrazione binoculare** (elaborazione di alto livello).



sua attivazione è ottimale quando fra l'occhio sinistro e l'occhio destro c'è una disparità di circa 30 minuti di grado. **Dunque con una disparità di circa 30 minuti di grado abbiamo il massimo dell'aumento della frequenza di scarica.** Per 0 minuti di grado o 60, l'attivazione risulta invece pressoché nulla. In V1 vi sono quindi numerose unità neurali sensibili alla disparità, alcune specializzate per la crociata, altre per la non crociata (la maggior parte per le disparità relativamente piccole).

In questi anni si è largamente discusso sull'importanza del periodo critico nello sviluppo della capacità di utilizzare la parallasse binoculare per vedere le distanze: **alla nascita, il neonato non è in grado di utilizzarla per percepire le distanze** (ciò non vuol dire che il neonato veda il "mondo piatto", ma è più sensibile alla parallasse di movimento); **tuttavia la parallasse binoculare matura velocemente nell'arco del primo anno di vita, già intorno ai sei/sette mesi.** La **stereoacuità** è un indicatore di tale capacità ed indica la disparità minima (la soglia) al di sopra della quale il soggetto percepisce la profondità.

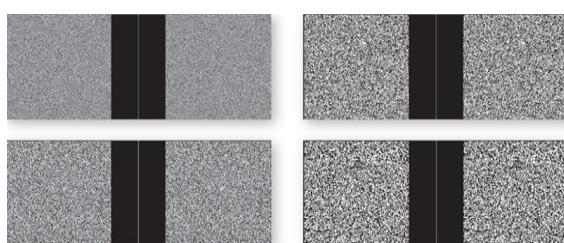
In immagine la colonna di sinistra rappresenta degli stimoli con dei contorni orientati a 90° l'uno rispetto all'altro (una sorta di croce), che stimoleranno in V1 delle unità sensibili selettivamente all'orientazione; in un altro strato di V1 si può trovare inoltre un'unità che risponde esclusivamente quando questo contorno viene presentato a entrambi gli occhi e la

Il primo anno viene ritenuto cruciale per sviluppare una corretta visione binoculare delle distanze e, di fatti, bambini che presentano alla nascita strabismo marcato e che sviluppano la condizione di “*occhio pigro*” (per cui il cervello sopprime l’immagine di uno dei due occhi e il cervello utilizza solo uno dei due occhi, come se vedesse monocularmente), non sviluppano tale capacità, proprio perché non vengono stimolate le unità binoculari sopra citate; occorre quindi intervenire tempestivamente per ripristinare una corretta visione e favorire lo sviluppo della parallasse binoculare. Tuttavia, ad oggi tale periodo critico rigido è stato messo in discussione, poiché alcuni soggetti hanno recuperato da adulti la capacità stereoptica; un caso celebre è quello di **Susan Berry**, il cui strabismo è stato corretto solo a 48 anni: nonostante l’età avanzata, dopo tale correzione e a seguito di un training per implementare la coordinazione tra i due occhi, l’artista ha recuperato improvvisamente la percezione della profondità. **Evidenze come questa sembrano quindi suggerire che sia possibile sviluppare la stereopsi anche oltre il periodo critico.**

Ci possiamo porre due domande a proposito della parallasse binoculare:

**1. La parallasse binoculare è condizione sufficiente per vedere il mondo in 3D, o devono esservi ulteriori condizioni?**

Per rispondere a tale domanda occorre creare una situazione in cui l’unica informazione sulla presenza di un oggetto e sulla sua struttura tridimensionale sia data dalla parallasse binoculare, per cui non vi sono informazioni a livello monoculare su forme, colori, tridimensionalità ecc... A tal proposito si utilizzano gli **stereogrammi a punti casuali**.



Osservando le immagini, monoculturalmente non si percepisce nulla; in realtà, tali stereogrammi nascondono un segnale di disparità retinica che viene creato prendendo alcuni dots in un’immagine e spostandoli - preservando le loro relazioni spaziali - nella copia corrispondente.

In questo modo, nell’immagine monoculare si percepiscono solo i puntini disposti a caso, mentre nel confronto tra le due immagini si crea una sorta di *correlazione spaziale*, per cui in una vi è una zona in cui un sottoinsieme dei punti è il medesimo, solo spostato di una certa disparità.

**N.B. c'è un errore in quanto i numeri nel quadrato arancione dovrebbero essere gli stessi.**

1	2	1	2	1	2	2	1	2	1
1	2	1	1	2	1	2	2	2	2
2	1	1	2	1	2	1	1	1	2
1	2	2	2	2	1	2	1	2	1
1	2	1	1	2	1	2	2	2	1
2	1	2	2	2	2	1	2	1	2
1	2	1	1	1	2	1	2	1	1
1	2	2	2	1	2	2	1	2	2
1	2	1	2	1	2	2	1	2	2
2	1	2	1	1	1	1	2	1	2

1	2	1	2	1	2	2	1	2	1
1	2	1	1	2	1	2	2	2	2
2	1	1	2	1	2	1	1	1	2
1	2	2	2	1	2	1	2	1	2
1	2	1	1	2	1	2	2	2	1
2	1	2	2	2	1	2	1	2	1
1	2	1	1	1	2	1	2	1	1
2	1	2	2	2	2	1	2	2	1
1	2	1	1	1	2	1	1	2	1
1	2	1	2	1	2	2	2	1	2

FIGURE 7.20 How Random-Dot Stereograms Are Made.

To make a random-dot stereogram, you make a random grid of black and white dots. Each number in the figure represents either a black (1) or white (2) dot. You then copy the image to make two such images. But in the second image, you shift a central section of the first image to the right or left. Thus, the same pattern is represented in each image, but part of it is shifted. When we look at this through a stereograph, the shifted part will appear either in front of or behind of the rest of the dots, depending on which direction it was shifted.

Inoltre, percepisce i 4 puntini immediatamente a sinistra del margine verticale (2,1,2,1 quelli in blu), mentre il destro percepisce i 4 puntini immediatamente a destra del margine verticale (2,1,1,2). Tale tipo di struttura stimola le unità sensibili alla disparità anche in assenza di qualsiasi altra informazione.

Questo particolare tipo di stimolo (**stereogramma a punti casuali**) è stato inventato negli anni 60 dall'ingegnere ungherese **Béla Julesz** e fondamentalmente consiste in una tecnica psicofisica che consente di studiare la visione binoculare della distanza e, quindi, come il cervello risponde alle disparità retiniche.



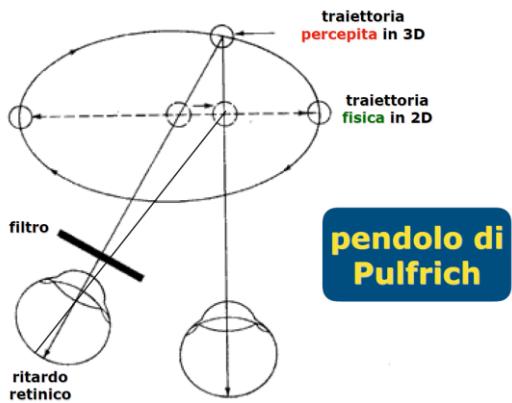
Béla Julesz (1928-2003)

Queste due immagini presentano una struttura binoculare (ricavabile solo attraverso il confronto tra le immagini monoculari): la posizione sulla retina del quadrato arancione è diversa nell'occhio sinistro rispetto al destro; l'occhio sinistro,

- l'oggetto non esiste nelle viste monoculari
- emerge solo grazie alla fusione di elementi disparati

In immagine vediamo una coppia di stereogrammi a punti casuali, sul cui fondo vi è un'immagine stereoscopica, che non esiste nel campo monoculare e che non è visibile fintanto che il cervello non fonde l'input dei due occhi; la sua percezione non può essere quindi supportata da cellule semplici - che non sono in grado di riconoscerne i contorni - ma da **cellule di second'ordine**, le quali estraggono **l'informazione dalla disparità retinica**. Pertanto, come sosteneva Julesz, **tramite gli stereogrammi a punti casuali è possibile bypassare l'apparato visivo periferico e proiettare direttamente le informazioni alla "retina della mente"**, dove le proiezioni di occhio destro e sinistro vengono **combinate nella corteccia visiva**. Egli definiva il campo binoculare (la "retina della mente") "*l'occhio ciclopico*", proprio perché in grado di convertire le immagini dei due occhi e la disparità in un'unica percezione.

Altro esempio per dimostrare che la disparità è condizione sufficiente **pendolo di Pulfrich**:



La tecnica prevede che venga fatto oscillare un pendolo su un piano (che quindi descrive una traiettoria in due dimensioni) e che venga posto un filtro su uno dei due occhi che *cambia lo stato di adattamento della retina*, ossia **riduce la quantità di luce che entra nella retina e quindi un occhio è meno illuminato dell'altro** (più adattato al buio). L'occhio meno illuminato inizia quindi a lavorare più lentamente a causa di questo diverso adattamento: mettere un filtro è come porre un ritardo ad una delle due retine, poiché, riducendo l'intensità luminosa, *la retina cambia la propria finestra di integrazione temporale* (l'occhio lavora più lentamente, in modo da raccogliere più fotoni nel tempo prima di produrre l'output). L'occhio coperto da un filtro, producendo l'output in ritardo segnala al cervello una posizione dell'oggetto un po' precedente rispetto all'occhio senza filtro. **Tale condizione è equiparabile alla produzione di una disparità retinica**, perché uno dei due occhi segnalerà una posizione diversa del pendolo rispetto all'altro occhio. In questo modo il pendolo sembra di fatti oscillare in 3D, descrivendo una traiettoria ellittica, la cui direzione dipende da quale occhio ha il filtro.

**Quindi, per concludere, la parallasse binoculare è condizione sufficiente per la percezione della stereopsi e la disparità retinica è molto utile da un punto di vista adattivo, soprattutto per le specie predatorie.**

**2. La parallasse binoculare è condizione necessaria per la stereopsi?** NO, poiché esistono altri meccanismi che consentono la percezione 3D.

**Rogers e Graham** nel loro lavoro "Similarities between motion parallax and stereopsis in human depth perception" hanno utilizzato tecniche basate su *random dots* per investigare la **sensibilità del sistema visivo umano a modulazioni sinusoidali di profondità**, specificate da informazioni contenute nella parallasse di movimento; le soglie riferite alla percezione della profondità risultano essere massimamente piccole quando la frequenza spaziale delle corrugazioni in profondità è tra 0.2 e 0.5 cicli x grado di angoli visivo. Tali dati sono stati confrontati con le soglie equivalenti che si ritrovano nella percezione della profondità specificata da disparità binoculari, utilizzando un apparato e una procedura psicofisica simile.

La tecnica prevede che venga fatto oscillare un pendolo su un piano (che quindi descrive una traiettoria in due dimensioni) e che venga posto un filtro su uno dei due occhi che *cambia lo stato di adattamento della retina*, ossia **riduce la quantità di luce che entra nella retina e quindi un occhio è meno illuminato dell'altro** (più adattato al buio). L'occhio meno illuminato inizia quindi a lavorare più lentamente a causa di questo diverso adattamento: mettere un filtro è come porre un ritardo ad una delle due retine, poiché, riducendo l'intensità luminosa, *la retina cambia la propria finestra di integrazione temporale* (l'occhio lavora più lentamente, in modo da raccogliere più fotoni nel tempo prima di produrre l'output). L'occhio coperto da un filtro, producendo l'output in ritardo segnala al cervello una posizione dell'oggetto un po' precedente rispetto all'occhio senza filtro. **Tale condizione è equiparabile alla produzione di una disparità retinica**, perché uno dei due occhi segnalerà una posizione diversa del pendolo rispetto all'altro occhio. In questo modo il pendolo sembra di fatti oscillare in 3D, descrivendo una traiettoria ellittica, la cui direzione dipende da quale occhio ha il filtro.

Vision Research Vol. 22, pp. 261 to 270, 1982  
Printed in Great Britain

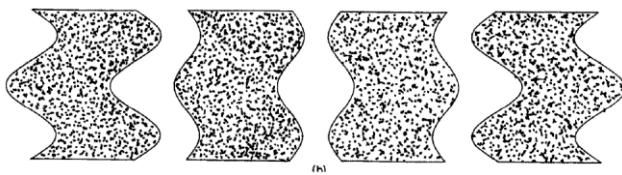
0042-8989 \$2.00/0.00 0  
Pergamon Press Ltd

#### SIMILARITIES BETWEEN MOTION PARALLAX AND STEREOPSIS IN HUMAN DEPTH PERCEPTION\*

BRIAN ROGERS and MAUREEN GRAHAM  
Psychological Laboratory, University of St Andrews, St Andrews, Scotland

(Received 15 May 1981)

**Abstract**—Random dot techniques were used to investigate the human visual system's sensitivity to sinusoidal depth modulations specified by motion parallax information. Thresholds for perceiving depth were found to be smallest when the spatial frequency of the depth corrugations was between 0.2 and 0.5 c/deg visual angle. These data were compared with the equivalent thresholds for perceiving depth corrugations specified by binocular disparity using similar apparatus and psychophysical procedures. The similarity between the sensitivity functions is suggestive of a closer relationship between the two systems than has previously been thought.



In breve: gli autori hanno creato **stereogrammi con corrugazioni/modulazioni sinusoidali di profondità** e poi l'equivalente con la parallasse di movimento (la geometria è molto simile, ma nella parallasse di movimento l'occhio singolo assume posizioni diverse nel tempo, mentre in quella binoculare gli occhi, simultaneamente, assumono due posizioni): confrontando, tramite lo stesso apparato, tali condizioni, hanno verificato che la funzione di sensibilità (cioè la variazione rispetto alla frequenza spaziale) di tali meccanismi risulta molto simile (praticamente equivalenti).

**Ciò suggerisce che tali meccanismi siano molto affini tra loro e che un osservatore monoculare in grado di utilizzare la parallasse di movimento dovrebbe avere quindi un'esperienza di stereopsi simile a quella di un osservatore binoculare** (tranne per oggetti molto vicini al corpo, per cui è migliore la percezione derivante da disparità retinica dal momento che sono richieste discriminazioni molto fini difficili da fare con un occhio solo).

Esempio: è estremamente difficile infilare il filo nella cruna dell'ago utilizzando solo un occhio rispetto all'utilizzo dei due occhi.

Si può concludere, quindi, che a meno di un fattore di scala (perché le scale sono diverse ed è dovuto al fatto che il movimento non può essere su un range più grande della disparità) le soglie per la parallasse di movimento e le soglie per la disparità retinica sono praticamente equivalenti rispetto alla funzione spaziale.

**Quindi possiamo affermare che la parallasse binoculare NON è condizione necessaria alla stereopsi e che esistono altri meccanismi relativi alla parallasse di movimento che sono in un certo senso equivalenti.**

**Filippo Brunelleschi** già in epoca rinascimentale aveva prodotto una percezione visiva che aveva permesso di comprendere come anche solo con indici pittorici monoculari (quindi visione monoculare statica) si può dedurre la stereopsi.

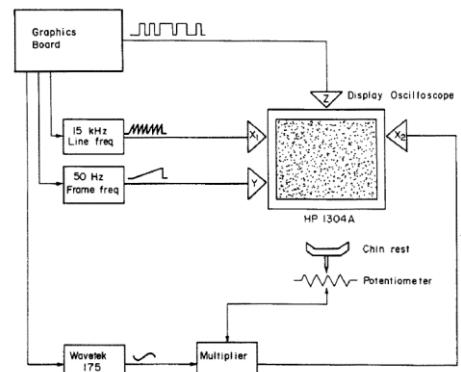
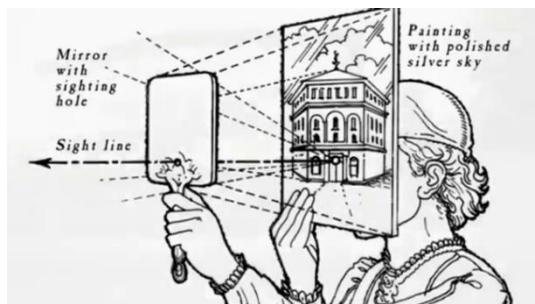


Fig. 1. A static random dot pattern was displayed on the oscilloscope screen using a raster technique. The Matrox ALT 256 graphics board provided the Z (intensity) signal as well as the sync pulses for the line (X1) and frame (Y) ramp generators. The entire pattern could be "distorted" by an additional waveform fed to the X2 input, whose amplitude was controlled by the position of the subject's head on the chin rest.



## Lezione 19

### Percezione Visiva dello Spazio

**Domanda:** Ruolo della parallasse binoculare

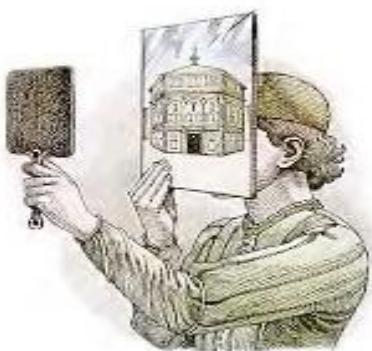
Posto che la parallasse binoculare sembra essere una *condizione sufficiente* perché si abbia esperienza di stereopsi visiva, abbiamo provato a chiederci se sia una condizione necessaria per la stereopsi.

**Risposta:** no, non è necessaria. Si può avere quello che sembra, dal punto di vista psicofisico, un risultato equivalente anche per un osservatore monoculare, se si presentano informazioni che dipendono dalla parallasse di movimento. Questo vale in condizioni particolari anche per il tipo di informazione che un osservatore può ricevere grazie alla struttura della luce quando tiene la testa ferma e guarda con un occhio solo, cioè utilizzando quelli che vengono chiamati gli **indici monoculari pittorici di profondità**.

Per introdurre il discorso, parliamo della rappresentazione sottostante.

**Filippo Brunelleschi** fu uno dei *massimi teorici della prospettiva*.

Il suo biografo racconta che per procurarsi i clienti e dimostrare la sua padronanza di questo strumento nell'arte figurativa, Brunelleschi ha costruito una dimostrazione particolare: si era posizionato in corrispondenza dell'ingresso della Chiesa della Piazza di Santa Maria Novella a Firenze, guardando verso il battistero e aveva dipinto un quadro in prospettiva.



A questo quadro aveva praticato un forellino (creando un *punto di vista speciale*) in corrispondenza del **punto di fuga** della prospettiva, cioè il **punto in cui si dovrebbe collocare un singolo occhio per ricevere esattamente il tipo di struttura ottica che avrebbe un osservatore guardando il vero battistero**. Brunelleschi dopo aver creato questo apparato chiedeva ai possibili clienti di osservare il quadro mettendo la testa sul retro di esso e tenendo uno specchio davanti.

Le persone vedevano il riflesso del quadro nello specchio, quindi vedevano un'**immagine virtuale** collocata dietro lo specchio. L'immagine corrispondeva esattamente al tipo di visione che l'osservatore, il quale teneva la testa ferma e guardava con un occhio solo, avrebbe avuto se fosse stato veramente nella piazza a guardare il battistero da quel particolare punto di vista. I clienti rimanevano veramente

meravigliati perché avevano l'impressione di vedere una scena tridimensionale e realistica; avevano una specie di esperienza di stereopsi, pur in assenza di qualsiasi informazione legata alla parallasse.

A partire da questo è interessante chiedersi quale sia la base di questo tipo, un po' paradossale, di stereopsi monoculare sulla base degli **indici pittorici**.

Quando guardiamo un quadro in prospettiva, una fotografia o un quadro iperrealistico, possiamo apprezzare la sofisticata tecnica di chi ha dipinto il quadro, ma non abbiamo mai l'esperienza di stereopsi; vediamo che si tratta di un quadro con strutture ottiche dipinte su superfici bidimensionali.

### **Indici pittorici di profondità e il loro rapporto con la prospettiva**



Scene reali fotografate: scene bidimensionali che se viste con due occhi e muovendo la testa riportano informazioni di parallasse di movimento e parallasse binoculare che non sono quelle giuste per la scena tridimensionale. Ci sono però altri tipi di informazioni che si correlano con la distanza rappresentata:

- La **grandezza relativa** è qualcosa che il sistema visivo utilizza per informazioni sulle distanze: le cose più piccole sul piano pittorico tendono ad essere viste come più lontane (Ad esempio nell'immagine il vaso di fiori occupa molto più spazio nella scena rispetto alla casa);
- le **conoscenze** che possiamo avere sugli oggetti (Ad esempio le case sono grandi e i vasi di fiori più piccoli): anche queste potrebbero entrare in gioco. Quindi, non solo **grandezze relative, ma anche conoscenze sulle grandezze che statisticamente tendono ad avere gli oggetti**. Sono cose che hanno una valenza più probabilistica che assoluta (potrebbe essere il vaso di fiori sia in realtà enorme o che la casa sia un modellino);
- la **posizione sul piano pittorico: le cose più alte sul piano pittorico** (gli alberi, in questo caso) **tendono ad essere più lontane delle cose più basse nel piano pittorico**. Questo dipende dal fatto che solitamente noi siamo in piedi su una superficie di supporto come è in questo caso il terreno. Queste superfici di supporto sono di solito più o meno statisticamente orizzontali quindi le cose più vicine tendono a proiettare sulla nostra retina come se fossero più in basso (fino ad arrivare all'orizzonte).

Nell'immagine sono presenti meno oggetti familiari (che hanno una grandezza che sappiamo statisticamente stare in una certa gamma). Orizzonte, cielo, mare, sabbia, pietra: **c'è una stratificazione in piani di profondità che dipende dalla posizione sul piano pittorico ed è rispetto all'orizzonte**: il mare viene percepito come più distante rispetto alla pietra bianca.

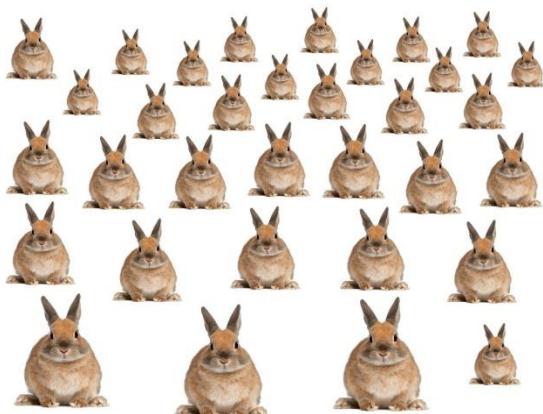


In questa foto invece possiamo osservare un panorama in una valle di montagna: chiaramente c'è una relazione tra quanto gli oggetti sono grandi e dove sono posizionati sul piano dell'immagine. Esempio: il è sentiero vicino, la strada più lontana, poi ci sono sia case vicine che case lontane. Se si fa attenzione si possono vedere le montagne sullo sfondo: le montagne che sono più vicine sono più in basso rispetto a quelle lontane e le montagne più sono lontane e meno sono contrastate, tendono a diventare più azzurrine e sempre meno contrastate rispetto al cielo. Questi indizi pittorici vengono sfruttati dai pittori quando cercano di fare un quadro che rappresenta una scena tridimensionale.

### **Regole della prospettiva: gli indici pittorici di profondità**

- Le persone vicine sono più grandi e più basse sul piano pittorico;
- I ciottoli sulla strada sono più grandi e più bassi quando sono vicini e diventano più piccoli allontanandosi.

Questi indici di profondità vanno ad interagire con la grandezza con cui percepisco questi oggetti:



SENSATION & PERCEPTION 4e, Figure 6.11  
© 2015 Sinauer Associates, Inc.

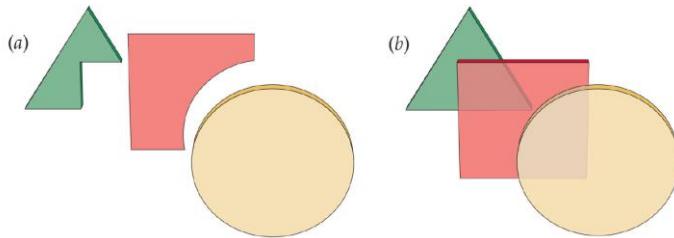


G. Caillebotte (1877) Rue de Paris, temps de pluie, oil on canvas [Art Institute of Chicago]

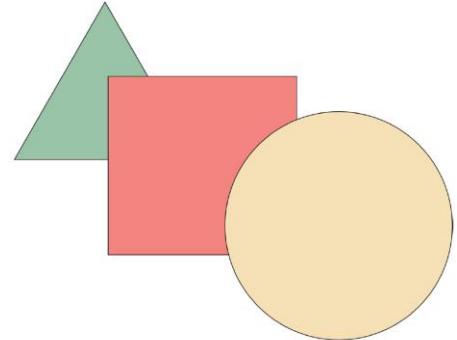
Esempio → Il leprotto in basso sembra piccolo, non perché la proiezione sulla retina sia piccola di per sé, ma a causa della sua posizione sul piano pittorico che lo fa sembrare più vicino. **Il calcolo complessivo di grandezza e distanza lo fa sembrare più piccolo rispetto agli altri sulla stessa fila.**

**Gli indici pittorici di profondità**, vengono appunto detti “pittorici” perchè sono quelli che un pittore ha a disposizione se vuole provare a dare l’impressione di profondità in un quadro e sono:

1. **Posizione sul piano pittorico;**
2. **Grandezza relativa;**
3. **Interposizione:** certi oggetti sembrano nascondere alla vista altri oggetti. Qui entra in gioco in maniera critica come il sistema analizza l’intersezione tra i contorni. Quando i contorni formano tra di loro una struttura che si può ricondurre ad un qualche tipo di T (giunzione a T) di solito viene interpretato come una conseguenza di un’interposizione in profondità.



SENSATION & PERCEPTION 4e, Figure 6.5  
© 2015 Sinauer Associates, Inc.



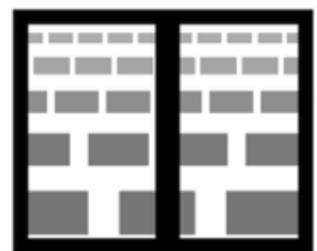
SENSATION & PERCEPTION 4e, Figure 6.5  
© 2015 Sinauer Associates, Inc.

**Fenomeno di completamento amodale:** la superficie occlusa sembra continuare dietro alla forma che occlude. **Amodale** perchè questa esperienza di completamento non si accompagna alla percezione reale

**dell’oggetto** (della parte occlusa); la modalità visiva non è presente per quella parte lì, ma si trova una sorta di “*presenza percettiva*”;

#### 4. **Contrasto relativo/ Prospettiva Aerea:**

a sinistra è presente una struttura di quello che noi potremmo vedere guardando attraverso una finestra vedendo dei rettangoli. Alcuni sono più in basso e altri più in alto e diventano sempre più piccoli e **sempre meno contrastati andando verso l’alto**. Tutte queste cose sono indici pittorici di profondità, in particolare **il fatto che diventino meno contrastati rappresenta appunto il contrasto relativo o prospettiva aerea** (*ha a che fare con la prospettiva rispetto al cielo - utilizzata anche da Leonardo da Vinci per dipingere la Gioconda*.



Indici pittorici che utilizzano le **ombre** e le **ombreggiature**:

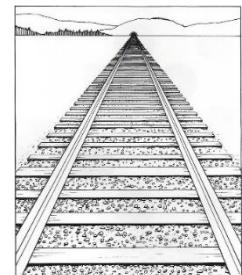
## 5. Chiaro - Scuro

Sulla superficie c'è un gradiente di colore grigio coerente con una certa struttura tridimensionale e con un'illuminazione che deriva dall'alto. Se questi oggetti sono delle sfere allora saranno più illuminati in alto e si faranno ombra da soli nella parte più in basso. **Questo genera un processo di inferenza della struttura tridimensionale sulla base di questa ombreggiatura e anche sulla base di assunzioni che riguardano la direzione dell'illuminazione.**

Questo vale per le ombre che ci sono sulla superficie dell'oggetto (ombre proprie) e vale anche per le ombre che l'oggetto proietta su altre superfici. Questo dà informazioni sulle **relazioni di contatto** tra un oggetto e una superficie su cui è posato: ad esempio nell'immagine una sfera sembra posata su un piano mentre l'altra sembra sollevata.

Tutte queste informazioni sui diversi indici pittorici di profondità possono essere combinate in un *quadro in prospettiva*:

Esempio – Rotaie di un treno: queste su un piano pittorico non sono parallele, ma vengono dipinte convergenti, le distanze delle coppie di punti diventano man mano minori → gradiente di grandezza relativa. Vanno anche verso l'alto sul piano pittorico: i diversi pezzi delle rotaie devono diventare sempre più piccoli e sempre più alti → indice pittorico del fatto che le rotaie si allontanano da noi e procedono in profondità.



L'idea è che **dovrebbe essere possibile in certe condizioni, anche utilizzando solo questi indici, produrre una impressione di stereopsi**. Un po' *paradossale perché di solito la stereopsi si accompagna ad una attività esplorativa e utilizza sempre le informazioni legate alla parallasse binoculare e di movimento*. Nel caso di Brunelleschi viene impedita la normale attività esplorativa ma in realtà produce un'impressione più realistica.

Questo porta a interrogarci su una questione: **Conflitto di informazioni tra Parallasse e Indici Pittorici**  
Interroghiamoci su cosa succede quando provo a prendere gli indici pittorici mettendoli in contrasto con le informazioni legate alla parallasse → Questo non succede nel mondo reale perché **tutti i processi di informazione visiva e percezione dello spazio sono tutti a registro e noi non abbiamo mai condizioni di conflitto nella visione naturale**. Non è difficile però creare condizioni in cui questo conflitto succede ed analizzando questo possiamo capire molte cose sulla percezione dello spazio e sulla nostra psicologia visiva.

Esempio → categoria di affreschi dell'arte barocca in cui il pittore ha cercato di riprodurre una scena in cui la parete non ci fosse e si vedesse invece qualcosa di tridimensionale.

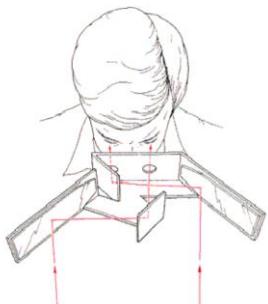


*Soffitto della Chiesa di Sant'Ignazio a Roma, pittore barocco Andrea Pozzo*

In questa chiesa c'è un'analogia del “giochino” di Brunelleschi: un punto al centro della chiesa, se uno si mette sul disco si trova nel punto di fuga della prospettiva dipinta da Pozzo. L'impressione, guardando verso l'alto, è che sia quasi stereoscopica la visione anche se stando molti metri sulla nostra testa non c'è parallasse binoculare perché gli occhi sono troppo vicini e la scena è troppo lontana (è come se uno guardasse con un occhio solo o con due occhi con lo stesso punto di vista). **Tenendo quindi la testa ferma l'unica cosa che è presente a livello visivo sono gli indici pittorici che, di nuovo, sono sufficienti per produrre un'esperienza che è quasi stereoscopica**. Sembra che gli elementi architettonici reali della chiesa continuino, in questo spazio illusionistico, verso l'alto e poi si aprano verso il cielo. Continuando a guardare in alto si può provare a spostarsi: così facendo succede che la struttura ottica proiettata sulla nostra retina si sposta in conseguenza al movimento; se questa fosse una reale struttura tridimensionale si dovrebbe produrre un'informazione legata alla parallasse di movimento che corrisponde a quella struttura tridimensionale. Questo però non accade perché non c'è una vera struttura tridimensionale, ma è presente un affresco bidimensionale → la parallasse di movimento non è appropriata a quella illusionistica struttura tridimensionale ma è appropriata alla piattezza dell'affresco.

Nella chiesa si crea un conflitto tra la parallasse di movimento e gli indici pittorici. *L'illusione della struttura tridimensionale non va persa, però mentre mi muovo questa sembra deformarsi.* Come se il cervello tentasse di trovare un compromesso/spiegazione per il fatto che gli indici pittorici specificano una certa struttura tridimensionale mentre la parallasse di movimento un'altra. Una maniera di risolvere questo problema è di interpretare il tutto come: “la struttura tridimensionale si deforma” e, cambiando io di posizione, **la deformazione compensa la mancanza di parallasse di movimento.** Questo è sintomo del fatto che *anche nella visione naturale queste cose vengono messe insieme nella stima globale della tridimensionalità della scena.*

## Test con Pseudoscopio



Pseudoscopio → divisore dove si guarda, attraverso due buchi, il mondo circostante. Non vediamo direttamente, ma dopo un gioco di riflessioni con degli specchi che servono a scambiare i punti di vista dei due occhi: all'occhio destro arriva quello che normalmente arriverebbe all'occhio sinistro e viceversa.

**Quando guardate il mondo attraverso lo pseudoscopio, quindi, gli indici pittorici rimangono invariati (poiché sono riferiti a viste monoculari), ma la parallasse binoculare è invertita rispetto alla struttura tridimensionale del mondo che vediamo.**

Esempio: se un oggetto A sta davanti ad un oggetto B, guardando attraverso lo pseudoscopio specifica al contrario. È interessante da fare perché non è che improvvisamente il rilievo del mondo si inverte, non succede che tutte le cose concave sembrano convesse e viceversa, **il mondo sembra abbastanza normale, a meno che non ci siano condizioni particolari.** In quel caso si può produrre una situazione in cui le cose si invertono in profondità e quindi si avranno illusioni (come nel caso di Fra Pozzo e delle deformazioni illusorie).

**La base della percezione dello spazio, almeno nel sistema visivo dell'uomo, è la visione/parallasse binoculare, ma evidentemente non è proprio così perché guardando attraverso lo pseudoscopio gli indici pittorici prevalgono sulla parallasse binoculare e noi vediamo il mondo più o meno normalmente.** C'è la sensazione che qualcosa di strano ai propri occhi viene fatto, ma questo dipende da tante cose tra le quali il fatto che guardare attraverso uno pseudoscopio è una situazione molto artificiale. Non ci dovremmo sorprendere che in certe condizioni sia possibile avere una impressione molto realistica di tridimensionalità anche solo con gli indici pittorici, che sono importanti e in certi casi prevalgono sulla parallasse binoculare.

Altro gioco con la prospettiva che hanno fatto in molti pittori tra i quali **Patrick Hughes**, pittore inglese: quadri fatti con un principio di **prospettiva inversa**, ovvero sono quadri non dipinti su una tela normale bidimensionale, ma su specie di bassorilievi, su una tela che è una specie di superficie corrugata.



Esempio : tela che invece di essere piatta è composta da specie di diedri i quali fanno un gioco di pareti inclinate con spigoli che vengono verso di noi o che si allontanano da noi. Gioco simbolico sulla pittura: il rettangolo centrale che sembra essere, nell'immagine, lo sfondo è in realtà più vicino a noi, mentre le pareti di destra e sinistra si avvicinano a noi.

Se si osserva il quadro da lontano sembra un quadro normale con certe relazioni spaziali, mentre se ci si avvicina ad una determinata distanza e ci si muove accade che il quadro sembra animarsi, *sembra cominciare a "muoversi"*. Se poi ci si avvicina ulteriormente il movimento scompare e diventa saliente il fatto che si stia guardando un bassorilievo e non una tela bidimensionale.

Link: <https://michaelbach.de/ot/sze-reverspective/index.html>

Cosa succede dal punto di vista dei processi della percezione dello spazio: **quando siamo lontani dal quadro lo guardiamo con due occhi (che sono abbastanza vicini) quindi la differenza fra le due immagini retiniche è quasi nulla cioè non c'è parallasse binoculare**. Anche se siamo lontani e il quadro non è troppo grande il nostro movimento non produce grandi cambiamenti sulle nostre retine, non c'è praticamente parallasse di movimento. **Quindi cosa c'è?** Ci sono solo gli indici pittorici e noi vediamo il quadro in funzione di quelli, come lo stiamo vedendo adesso nella slide.

**Avvicinandoci al quadro cosa succede però? Cominciano a diventare più salienti e informativi gli indici legati alla parallasse e quindi entra in gioco un conflitto fra questi e gli indici pittorici.** Questo conflitto il sistema visivo cerca di risolverlo: *c'è una zona in cui il peso relativo di queste fonti di informazione è più o meno equivalente e quindi il cervello non sceglie una sola fonte ma le utilizza tutte*, cercando di metterle insieme.

**Se mi avvicino ancora di più** la parallasse binoculare diventa cruciale, dominante, poiché io vedo con due occhi una cosa molto vicina: **le disparità retiniche sono molto grandi, molto sovra soglia e quindi vedo il quadro corrugato e si annulla l'importanza degli indici pittorici.**

**Quando guardiamo un quadro lo facciamo esattamente utilizzando quei processi di osservazione dello spazio che usiamo per guardare il mondo reale e cerchiamo di dare una soluzione ragionevole di quello che stiamo vedendo.** Questo ha implicazioni per come noi possiamo pensare a certi fenomeni che sono stati studiati dalla psicologia dell'arte.

## **Trompe-l'œil**

Esiste un genere pittorico chiamato Trompe-l'œil (ingannare l'occhio, in francese), che rappresenta la pittura illusionista per eccellenza.

Il genere trompe-l'œil presuppone una **grande maestria tecnica nell'utilizzare indici pittorici** e può dare l'impressione, se lo si guarda dal punto di vista giusto, tenendo la testa ferma guardando con un occhio solo, che la scena rappresentata sia una scena tridimensionale. L'inganno dell'occhio però non avviene quasi mai in forma completa. **Solitamente quando si guarda un quadro trompe-l'œil è simultaneamente consapevole del fatto che la scena possa sembrare reale ma in realtà non lo sia.** Simultaneamente consapevole di vedere una rappresentazione di una scena tridimensionale, molto sofisticata da un certo punto di vista, ma allo stesso tempo un quadro piatto; questo perchè guardiamo il quadro con due occhi, muovendoci.

Gli indici pittorici, quindi, sarebbero quelli adeguati a produrre stereopsi (es: Brunelleschi con lo specchio) però, solitamente, al museo, non guardiamo un quadro in questo modo e quindi tutti gli indici di parallasse specificano piattezza.

Alcuni teorici del trompe-l'œil → **Omar Calabresi:** aveva una teoria secondo la quale questo genere di pittura non sarebbe un “inganno dell'occhio” → l'osservatore non scambierebbe quella scena per una vera scena tridimensionale, ma metterebbe in atto un “**gioco cognitivo**” → *gioco in cui il pittore ti mostra quanto è bravo ma, allo stesso tempo, lo sa che tu sai che sta cercando di ingannarti.*

Questo lo diciamo perché è interessante da contrastare con quello che succede invece in quelle situazioni che noi chiamiamo **Realtà Virtuali**.

Esempio: **Realtà Virtuale Immersiva** (VR): quella in cui si verifica il fenomeno chiamato “*Presenza/Presence*”.

Nella **realtà virtuale immersiva** il soggetto ha esperienza della “*Presence*”, intendendo la **stereopsi**. Gli oggetti nella realtà virtuale sono appunto virtuali, ma l’impressione è che siano volumi nello spazio del tutto realistici, con i quali possiamo interagire. Abbiamo un visore stereoscopico, con cui si hanno due viste separate per i due occhi, allo stesso tempo abbiamo anche qualcosa in grado di tracciare la posizione della mano, testa e corpo. **Il computer quindi può ricreare in tempo reale stimolazioni ottiche (quello che viene presentato sui due schermi) tenendo conto di come noi ci siamo spostati, simulando la parallasse di movimento e la parallasse binoculare, oltre a tutti gli indici pittorici.** La simulazione è perfettamente fedele alla realtà (a parte qualche leggero ritardo nell’aggiornare l’immagine) e gli indici pittorici e di parallasse sono quasi tutti perfettamente a registro. Questa è la conseguenza del principio per il quale **il cervello risponde in base alla stimolazione che gli diamo: se diamo una stimolazione equivalente a quella del mondo reale il cervello risponde permettendoci di vedere un mondo che sembra reale.**

Quando invece guardiamo un quadro accade una cosa molto diversa rispetto alla realtà virtuale → *la situazione di Brunelleschi è una specie di realtà virtuale dei poveri* (cit.) - realtà virtuale minimale perché elimina il conflitto tra indici di parallasse e indici prospettici, costringendoci a guardare il quadro attraverso il buco, mettendolo poi in uno spazio virtuale attraverso lo specchio. Questo però non è ciò che succede tipicamente quando guardiamo il quadro. Dal punto di vista della psicologia dell’arte, di una persona che guarda un quadro e vive un’esperienza estetica, **quello che succede è che questi indici sono sempre in conflitto: la parallasse binoculare di movimento specifica piattezza mentre gli indici pittorici specificano tridimensionalità.** Nonostante ciò noi non percepiamo, guardando un quadro, alcuna deformazione.

Esempio: Quando si guarda un ritratto, muovendoci davanti, sembra che questo ci segua, però, di norma, vediamo quella rappresentazione tridimensionale e la interpretiamo correttamente anche se ci muoviamo e quindi anche se non occupiamo il punto di vista che sarebbe “quello giusto” (punto di fuga della prospettiva che è stata dipinta). Questo fenomeno è noto nella psicologia dell’arte prospettica, chiamato da **Michael Kubovy** (ha scritto il libro “*Psicologia della prospettiva e dell’arte del rinascimento*”) ***Robustness della prospettiva.***

**Robustness della prospettiva:** fenomeno per cui la prospettiva non si deforma, ma sembra funzionare nella stessa maniera anche quando guardiamo il quadro da un punto di vista generico. La *prospettiva* come tecnica di rappresentazione dello spazio in due dimensioni è cognitivamente robusta.

**Questo suggerisce che quando noi guardiamo un quadro attiviamo un processo che non è solo percettivo, ma anche cognitivo. C'è qualcosa che compensa il fatto di non essere nel punto di fuga esatto e questo significa che noi riusciamo a trasportarci virtualmente nel punto di vista rispetto al quale quella prospettiva sarebbe giusta per vedere la tridimensionalità.**

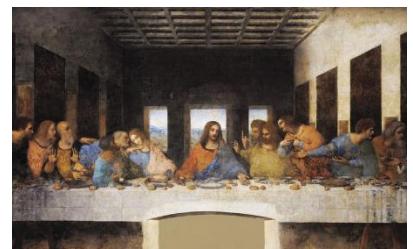
Teoria basata su questa idea: secondo Kubovy il motivo per cui le persone sono così affascinate dai quadri in prospettiva è che questi consentono all'artista di fare un gioco legato al punto di vista reale e al punto di vista immaginato dell'osservatore.

*“Quando percepiamo un quadro che è stato dipinto in prospettiva rispetto ad un particolare punto di vista, (centro prospettico, punto di fuga) il nostro sistema percettivo inferisce la posizione di questo centro di proiezione. Allora abbiamo la sensazione, cognitivamente, che stiamo guardando quella scena da quel particolare punto di vista.”*

**Kubovy**

Guardo il quadro e so che non sono in quel punto esatto, ma inferisco dov'è. Un po' come se l'artista avesse attivato dentro di noi dei processi percettivo-cognitivi/ metapercettivi in cui ci trasporta mentalmente dal punto in cui siamo davvero, quando guardiamo la scena, ad un punto diverso.

Questo si presta a dare interpretazioni di certe scelte che gli artisti hanno fatto per costruire quadri in prospettiva. Kubovy fa un esempio riferendosi al quadro “*Ultima Cena*” di Leonardo: affresco dipinto sulla parete di un refettorio, molto in alto, in modo che quando si entra nel refettorio si è su un punto di vista reale più basso rispetto a quello virtuale mostratoci nell'affresco, che corrisponde al punto di fuga della prospettiva.

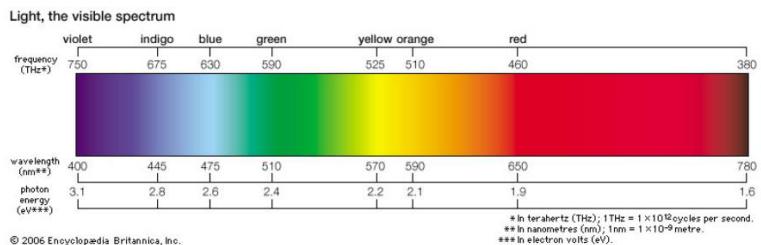


L'effetto di questo è che lo spettatore guarda il quadro e cognitivamente ha un senso di elevazione, quindi l'esperienza estetica progettata da Leonardo è un **gioco metapercettivo sul significato spirituale che lui voleva dare all'affresco**.

## 20 Lezione

### Luce, Riflettanza e Colore

Questo è lo **spettro visibile**. Il fatto che l'energia elettromagnetica a cui noi possiamo rispondere (cioè lo spettro visibile) sia compreso fra 400/500 e 650/700 nm, non va confuso con il colore.



Spesso si dice che la luce con lunghezza d'onda corta è blu e quella lunga è rossa e quindi implicitamente si afferma che “*la luce è colorata*”, in realtà la luce **non è colorata**. Il colore è un **costrutto mentale**, è un **prodotto del processo di percezione**, non è una proprietà fisica.

Questa non è un'idea nuova, infatti già **Newton** nell'*Optica* aveva scritto questo:

“*..the rays to speak properly are not coloured.*”

[**Isaac Newton – 1730**]

Traduzione: “se parliamo con proprietà di linguaggio, i raggi luminosi non sono colorati.”

“*There is no red in a 700 nm light, just as there is no pain in the hooves of a kicking horse.*”

[**Steven Shevell - 2003**]

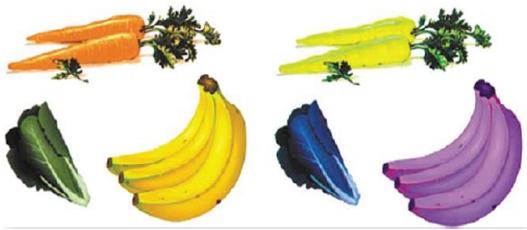


FIGURE 6.3 Familiar Colors

Objects with their familiar characteristic colors, and the same objects but with odd colors.  
? Objects get their characteristic colors because they reflect light at particular wavelengths. Ripe bananas absorb most light but reflect light at about 550 nm, giving them their yellow color.

La stessa idea la ritroviamo in **Steven Shevell**, esperto americano del colore, che pochi anni fa nel suo articolo ha scritto che: “*Non esiste il rosso in una luce di 700 nm, esattamente come non esiste il dolore negli zoccoli di un cavallo che ci dà un calcio.*”

Questo per problematizzare dall'inizio un'idea che spesso nel linguaggio comune troviamo accettata in modo acritico, ciò viene spesso riportata nei libri purtroppo. Per esempio, in un libro di testo viene riportato come le banane assorbo la maggior parte della luce, ma riflettono la luce intorno a circa 500 nm e questo gli conferisce il colore giallo. **È vero?**

Nell'*Optica* Newton dice qualcosa di simile:

“*Corporum naturalium colores e genere radiorum derivantur quos maxime reflectunt.*”

[**I. Newton 1670-72 Optica, Parte II, lectio 9**]

Traduzione: “*I colori dei corpi naturali derivano dal tipo di raggi che riflettono in massimo grado.*”

Quindi, il colore che noi percepiamo corrisponderebbe a una proprietà fisica degli oggetti, cioè **la loro tendenza a riflettere raggi in una certa gamma dello spettro visibile e ad assorbirne altre parti**. Il colore è il corrispettivo percettivo di una proprietà fisica, il modo in cui il cervello codifica una proprietà fisica delle superfici, ovvero la loro **riflettanza**.

**Cos'è la riflettanza?** È quella proprietà fisica definita da questa equazione semplificata:

$$R(\lambda) = \frac{I_r(\lambda)}{I_i(\lambda)}$$

**R** (riflettanza) è una funzione, cioè una proporzione senza unità di misura, che assume valori diversi in funzione di un'altra variabile: la *lunghezza d'onda* ( $\lambda$ ).

Quindi R è una particolare funzione: **funzione di lambda**. Ed è uguale al rapporto fra due altre funzioni: la **funzione che descrive l'intensità (I) della luce riflessa (r)** e la **funzione dell'intensità (I) della luce incidente (i)**.

Ad esempio, una sorgente di luce come il sole emette una fonte di illuminazione che è uno spettro, (una funzione di lambda) il quale arriva ad una superficie e in parte viene assorbito e in parte viene riflesso. Il rapporto fra ciò che viene riflesso e ciò che viene assorbito è R (riflettanza spettrale).

Questo vuol dire che: se tutta o quasi tutta la luce che arriva su una superficie viene riflessa questo rapporto tenderà ad essere molto vicino a 1 e quindi R sarà tipo: 0.95, 0.91 ecc. che equivale a dire che il 95% della luce che arriva sulla superficie viene riflessa e molto poca viene assorbita.

Se, di contro, la luce che arriva viene assorbita allora il rapporto sarà approssimativamente tendente a 0 e quindi R sarà: 0.05, 0.10 ecc.

Se mettiamo le superficie chiare sotto il sole si scaldano di meno di quelle scure, perché quella superficie scura (e i materiali che la compongono) tendono ad assorbire tutta la luce e quindi la trasformano in calore scaldandosi molto di più, viceversa per quelle chiare.

**Ma cosa succede davvero in natura? Com'è fatta la riflettanza di un corpo naturale?**

Un laureando del prof ha svolto una tesi di laurea sulla riflettanza prendendo una grande varietà di oggetti naturali (foglie, pezzi di corteccia, fagioli, pigne, petali di fiori ecc.) e ha usato uno *spettrofotometro* con una camera chiusa.



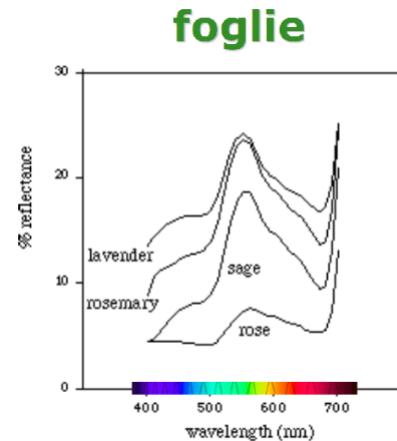
Per utilizzare lo spettrofotometro si inseriscono i vari campioni di oggetti al suo interno, dove c'è una fonte di illuminazione di composizione spettrale/spettro noto e ci sono dei sensori di energia elettromagnetica. La camera interna è tutta dipinta con una vernice che assorbe la luce, per cui i sensori registrano solo la luce che viene riflessa dal campione. Essendo lo spettro dell'illuminante noto, viene misurata la luce riflessa attraverso il rapporto fra luce riflessa e luce assorbita.

Il risultato è un'approssimazione della **funzione di riflettanza spettrale**. Approssimazione perché R di lambda è una funzione continua, definita per tutti i valori possibili di quella gamma, però essendo lo spettrofotometro uno strumento fisico non si possono misurare infiniti valori, quindi di fatto viene restituito un vettore di numeri associati a un vettore di lunghezze d'onda.



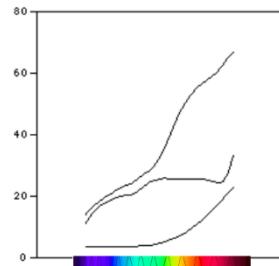
Ad esempio se immaginiamo di essere su R, nel dataframe c'è una colonna con la lunghezza d'onda campionata ogni 5 nm e poi un'altra colonna che esprime il rapporto fra l'intensità della luce nota nell'illuminante a quella lunghezza d'onda e l'intensità della luce misurata dal sensore a quella lunghezza d'onda.

Usando lo spettrofotometro si possono ricavare delle funzioni di riflettanza, ad esempio nell'immagine è possibile osservare le funzioni di riflettanza di diversi tipi di *foglie verdi* e infatti possiamo vedere come c'è una certa somiglianza nella forma complessiva della funzione di tutte le foglie. Infatti, la foglia si comporta come una specie di filtro che blocca proporzionalmente di più le lunghezze d'onda molto corte (400-500 nm) e lascia passare le lunghezze d'onda al centro dello spettro e tende di nuovo a bloccarle verso la parte finale dello spettro, e infine c'è il picco che quasi sconfinava nell'infrarosso (che va al di fuori dello spettro visibile). Si tratta di colori verdi con sfumature diverse, quindi sembrerebbe che ciò che noi percepiamo come verde è questa particolare forma che ha la funzione di riflettanza, mentre la diversa sfumatura dipende dal fatto che questa particolare forma si può collocare più in alto o in basso (verdi più chiari e verdi più scuri).



Le foglie secche invece hanno tipicamente una riflettanza molto diversa, non si comportano come filtro che lascia passare soprattutto la parte centrale dello spettro, ma come una funzione che tende a far passare sempre di più la luce, a seconda di quanto è grande la lunghezza d'onda. Questo sembra suggerire anche come le foglie secche abbiano un colore diverso, giallastro, rosso e non verde.

La riflettanza in questi grafici va intesa come percentuale e non come proporzione (rapporto fra luce riflessa e luce incidente x 100).



**secche**

**Studente:** “Si può parlare di riflettanza sullo spettro completo? Senza selezionare la lambda massima, ma prendendo tutto lo spettro?”

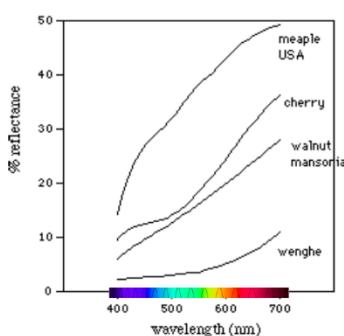
**Prof:** “Certo, questa è la funzione intera. Quando diciamo  $R$  di  $\lambda$  si intende quella particolare  $f$  di  $\lambda$  che descrive quella particolare forma.”

**Studente:** “Quindi la luce incidente viene sempre selezionata, percorre tutto lo spettro però non viene mai somministrata completamente tipo la luce bianca?”

**Prof:** “Dal punto di vista tecnico non sono sicura di come sia fatto il circuito di questi sensori, concettualmente questa funzione  $R$  di  $\lambda$  viene stimata basandosi su una interpolazione di più o meno una quarantina di punti, misurati separatamente.”

Quando diciamo  $R$  di  $\lambda$  intendiamo una di queste curve, le quali descrivono dove nello spettro visibile la superficie tende a riflettere e dove ad assorbire.

## legno



Le foglie secche hanno un andamento crescente e quasi lineare, e questi andamenti sono molto simili a oggetti come il legno.

*Maple* = acero americano;

*Cherry* = ciliegio;

*Walnut* = noce;

*Wenghe* = legno africano.

Queste curve hanno un andamento simile, nel senso che non mostrano una preferenza per una certa zona dello spettro, ma mostrano un incremento della riflettanza più o meno costante (monotonico e continuo) in funzione di  $\lambda$ . Questo suggerisce che c'è una certa corrispondenza fra la forma di quelle curve e il colore che noi percepiamo, ma se prendiamo il *colore dei fiori*, le cose cambiano di molto.

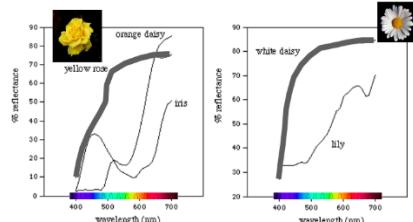
Ad esempio, la *margherita arancione* ha un certo andamento che è diverso da quello di una *margherita bianca*. Quest'ultima non riflette nella parte più bassa dello spettro, mentre riflette in modo quasi uniforme da 500 nm in su, mentre la margherita arancione a circa 500 nm la sua capacità di riflettere aumenta, rimane costante per un po' e poi sale di nuovo e riflette nella parte medio-lunga dello spettro.

Quindi anche questo sembrerebbe obbedire al principio: diversa forma della curva = diverso colore percepito. Anche questo esempio parrebbe confermare che a funzioni di riflettanza simile corrispondano colori simili.

Se tuttavia prendiamo i grafici delle funzioni di riflettanza corrispondenti al petalo della *rosa gialla* e alla

### fiori

*margherita* possiamo vedere che sono molto simili (anche se non uguali) però i due fiori sono differenti e questo crea effettivamente un problema.



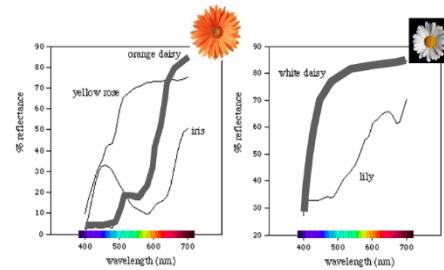
**In conclusione si può affermare che ci sono delle grossolane corrispondenze tra colore e il pattern caratteristico che mostrano queste funzioni, ma non è così semplice affermare quello che diceva Newton ovvero che il colore corrisponde alla riflettanza degli oggetti.**

**Studente:** “*Dal punto di vista delle proprietà fisiche della materia da cosa dipende questo grado di riflettanza degli oggetti?*”

**Prof:** “*Dipende da come il materiale di cui è composto l'oggetto interagisce con i fotoni. Questo trova spiegazione a livello di fisica della materia*”.

Quando un raggio di luce indice su una superficie possono succedere 3 cose: può essere **assorbito**, **riflesso** o **attraversare** la superficie. Attraversandola poi può succedere che la sua direzione viene alterata e questo è il **principio della lente**, una lente fa piegare un raggio di luce, fenomeno noto come **rifrazione** della luce.

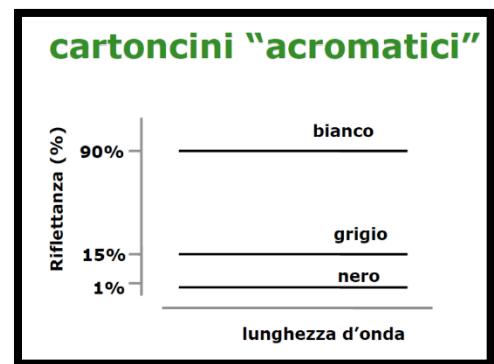
### fiori



Per quanto riguarda la riflessione questa può avvenire in due modi diversi: alcuni materiali riflettono come *specchi*, in particolare il raggio di luce viene riflesso con un angolo che è uguale all'angolo di incidenza, quindi i raggi riflessi mantengono le loro relazioni spaziali ed è per questo ci vediamo riflessi in uno specchio. Mentre altre superfici tendono a riflettere la luce in modo *diffuso*, cioè quando la luce colpisce la superficie c'è un'uguale probabilità che venga riflessa in qualsiasi direzione. Quindi questo tipo di superfici non preservano la struttura, le relazioni spaziali tra i raggi, per cui non ci vediamo riflessi in superfici di questo tipo. Questa è la ragione per la quale alcune superficie sono opache ed altre si comportano come specchi, questo dipende dalle caratteristiche fisiche del materiale di cui è fatta la superficie: più una superficie è *granulosa*, più tende a riflettere in maniera diffusa, mentre una superficie liscia tende a riflettere in modo speculare (tipo il metallo).

**Quindi ci sono vari fenomeni fisici al di là dell'assorbimento o della riflessione di cui stiamo parlando che contribuiscono in condizioni naturali a come noi percepiamo il colore, ma l'aspetto più importante è comunque sapere se la superficie tende più ad assorbire o riflettere.**

**Ipotesi psicofisica derivante da Newton:** c'è una *corrispondenza esatta* fra la riflettanza, che è una caratteristica fisica degli oggetti, e il colore percepito. In questo senso il colore sarebbe il modo in cui il cervello codifica la riflettanza degli oggetti e lo fa ad esempio per sapere di che oggetto si tratta. Questo potrebbe essere anche una forma di **ottica inversa**, cioè all'occhio arriva una certa luce e sulla base di questa luce il cervello cerca di inferire il materiale di cui è fatto un oggetto. ***Ma è questo il colore?***



## Colori acromatici: colori senza colore

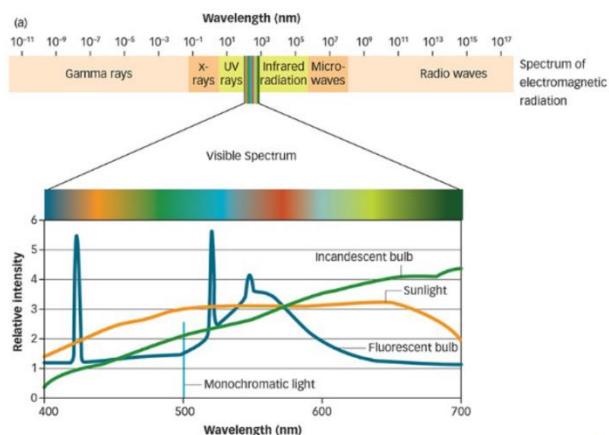
Esistono anche dei **colori** che chiamiamo **acromatici**, cioè colori senza colore. Una delle dimensioni dei colori percepiti è quella che va dal bianco al nero passando dalla scala dei grigi. Ad esempio, prendendo dei cartoncini graduati per tutte le tonalità di grigio e ciò che viene fuori misurando la riflettanza con lo spettrofotometro è una funzione piatta. Quindi, un cartoncino bianco avrà una riflettanza alta, per cui tenderà a riflettere il 90% della luce incidente senza avere una selettività per la lunghezza d'onda quindi in egual misura sia per le lunghezze d'onda più corte, medie e lunghe. Un cartoncino grigio avrà una riflettanza intorno al 15% e uno nero intorno allo 1%.

A livello percettivo il grigio si trova a metà strada tra il grigio e il bianco, tuttavia sulla riflettanza nel grafico è possibile vedere come in realtà questo si trovi molto più vicino al nero che al bianco e questo è sempre dovuto alla **relazione logaritmica tra la sensazione e la proprietà fisica**.

Il grafico rappresenta ciò che è possibile ottenere con i cartoncini acromatici, tuttavia esistono materiali che hanno anche una riflettanza maggiore, ad esempio il *gesso* ha il 95% di riflettanza, mentre è impossibile andare sotto l'1%. Questo per dire che è non esiste in natura un oggetto che riflette o assorbe il 100% della luce e che rimane un concetto teorico.

Visto che abbiamo parlato di **riflettanze**, cioè il rapporto tra luce incidente e luce riflessa, proviamo a vedere com'è fatta la **luce incidente** e anch'essa è una funzione di lambda.

In immagine è rappresentato lo spettro visibile in relazione all'**intensità relativa**, cioè quanta energia è presente in funzione della lunghezza d'onda in diversi tipi di potenziali luci incidenti.



La curva arancione (etichettata come “**sunlight**”) è approssimativamente la luce del sole a mezzogiorno, che si comporta quasi come una luce acromatica in quanto ha uno spettro quasi piatto. Dobbiamo comunque tener presente che la *luce del sole* durante il giorno può cambiare molto ad esempio in funzione dell'ora o del tempo atmosferico.

Quello che però è più interessante notare sono le differenze dal punto di vista della composizione spettrale tra la luce del sole e la *luce artificiale*. Nell'immagine si possono vedere due esempi: la luce di una **lampadina fluorescente** e di una **lampadina incandescente** e si può notare come gli andamenti di queste due siano molto diversi da quelli della luce solare. Ad esempio la luce fluorescente è composta solo da 3 gamme molto ristrette di lunghezze d'onda con un picco vicino a 400 nm, uno appena oltre 500 nm e un ultimo picco intorno a 550 nm e ha pochissima energia nella parte destra dello spettro. In questa figura è anche mostrata la **luce monocromatica**, non dobbiamo pensare alla luce monocromatica come corrispondente ad un solo colore, ma dobbiamo ragionare in termini fisici: in questo caso questa particolare luce monocromatica corrisponde ad una gamma molto ristretta dello spettro visibile, più o meno 500 e 501 nm, è molto difficile ottenere nella realtà una luce così ben definita.

## Psicofisica del colore acromatico

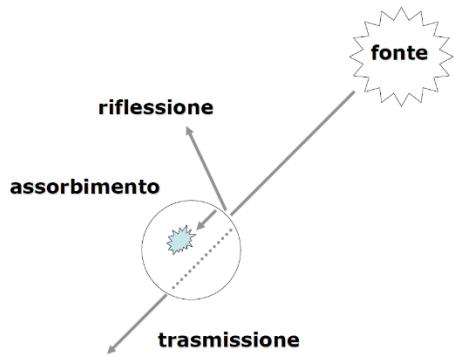
Vediamo come la riflettanza e l'illuminazione trovano un'applicazione in psicofisica, ovvero vediamo nel dettaglio la relazione fra riflettanza e colore percepito. Per fare questo è necessario fare una semplificazione, infatti se facciamo riferimento ad oggetti naturali la riflettanza è sempre il rapporto fra due funzioni e questo è molto complicato.



Esiste un settore della percezione che si occupa del colore facendone una ulteriore semplificazione, ovvero studia il mondo nelle diverse sfumature di grigio: la **percezione del colore acromatico**. Questa drastica semplificazione permette di determinare delle leggi e delle regole che poi possono essere utilizzate anche nella percezione del colore, immaginiamo dunque di studiare un sistema percettivo che non veda il colore, ma che veda un mondo fatto solamente di sfumature di grigio, un po' come fossimo soggetti affetti da *acromatopsia*.

Il grande vantaggio di questa semplificazione è che si può parlare di riflettanza non come funzione, ma come un singolo numero. Tornando all'esempio dei cartoncini acromatici questi hanno uno **spettro di riflettanza piatto**, per cui la riflettanza è costante in funzione di lambda, e quindi basta un numero per descriverla e non più una funzione. D'ora in avanti quando si parlerà di riflettanza si farà riferimento ad un singolo numero, compreso tra 0 e 1, dato dal rapporto tra luce riflessa e luce incidente (intesi a loro volta come singoli numeri).

Immaginiamo di avere una superficie che si comporta come il cartoncino acromatico e quindi possiamo trascurare lambda. Possiamo pensare alla **fonte** di illuminazione, come una certa quantità  $I$  che raggiunge la superficie di un oggetto e possono succedere tre cose:



1. **Riflessa**;
2. **Assorbita** e trasformata in calore;
3. **Trasmessa** attraverso la superficie.

La quantità di cui stiamo parlando è la **luminanza**, cioè la luce che potenzialmente può arrivare all'occhio. La luminanza è il prodotto tra **riflettanza ( $R$ )** (predisposizione della superficie a riflettere più o meno la luce) e **illuminazione ( $I$ )** (quantità di luce che incide sulla superficie):  $L = R \times I$ . Gli psicofisici chiamano questa formula **l'equazione fotometrica fondamentale**:  $L = R \times I$ .

Lo stimolo che arriva all'occhio non è la riflettanza in senso proprio, le varie funzioni di riflettanza di cui abbiamo parlato precedentemente non sono quelle che arrivano all'occhio. **Infatti la luce che arriva all'occhio è la riflettanza moltiplicata per l'illuminazione** e a seconda del tipo di illuminazione, la luce che arriva all'occhio cambia.

L'equazione fotometrica fondamentale è una “*Micky mouse version*” della visione, cioè una versione semplificata della vera fisica dell'interazione fra gli oggetti e la luce. Infatti affinché la formula fotometrica fondamentale sia valida, è necessario che siano presenti determinate assunzioni:

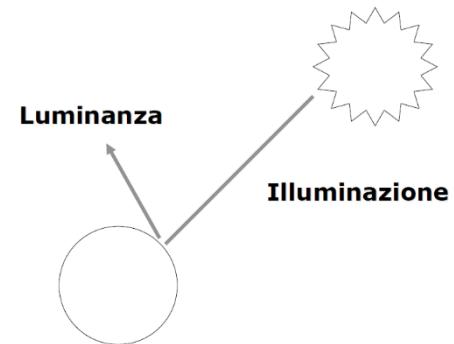
- Le **superfici** devono avere uno **spettro di riflettanza piatto** (tipo il cartoncino, quindi non oggetti o superfici naturali);
- Ci deve essere un'**unica fonte di illuminazione** sull'oggetto (non ci devono essere interazioni secondarie). Nella realtà questo non si verifica mai, perché dobbiamo pensare che ogni oggetto riflette parte della luce che arriva dal sole, diventando a loro volta una fonte di illuminazione. In un ambiente naturale c'è un gioco molto complesso di riflessioni tra tutti gli oggetti presenti che modificano l'illuminazione che arriva in un punto di un particolare oggetto;



- Si considera solo la **riflettanza diffusa**, ovvero solo quel tipo di riflettanza delle **superficie opache**, dove c'è solo la componente diffusa e non si tengono in considerazione le riflessioni speculari. Quindi non si tiene conto di specchi, vetri, superfici metalliche molto lisce ecc.

Per gli oggetti naturali questa formula non va bene in quanto queste assunzioni non vengono mai rispettate.

Ad esempio, questa palla da biliardo non va bene perché è illuminata da due fonti di illuminazione, la quantità di luce che arriva sulla superficie dipende dalla struttura tridimensionale dell'oggetto quindi ci sono zone in cui questa pallina è più illuminata e zone in cui è meno illuminata (differenza tra ombre portate e ombre proprie). Essendo poi una superficie liscia si comporta un po' come uno specchio, per cui riflette la fonte di luce e infatti sulla superficie in alto a destra è possibile vedere una zona più chiara che corrisponde alla fonte di illuminazione.



Queste proprietà che hanno a che fare con le ombre e con la specularità della superficie, le stiamo volontariamente tralasciando ora, ma sono importanti perché contengono informazioni sulla struttura tridimensionale sulla scena: la distribuzione delle ombre sulla superficie è correlata con la forma della superficie e con la geometria della direzione dell'illuminazione. Per esempio queste 4 palline da biliardo sono illuminate da una lampada che sta in posizione diversa rispetto all'oggetto.

Fatte queste semplificazioni abbiamo ben definito lo spazio degli stimoli in cui ci collochiamo per proseguire la spiegazione: *ci occuperemo di cartoncini acromatici con un unico illuminante*, per cui vale l'equazione fotometrica fondamentale, che definisce la relazione fra la luminanza, la riflettanza e l'illuminazione.

**La luminanza è lo stimolo effettivo che può arrivare all'occhio e che dipende dal prodotto fra riflettanza e illuminazione.**

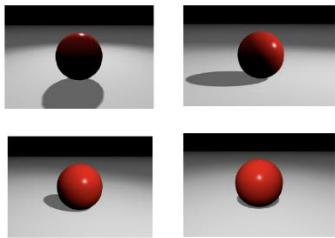
Questa definizione però ha un problema: dato che il cervello deve ricostruire il colore a partire dalla luce arriva all'occhio, riflettanza e illuminazione sono confuse, dato un valore di lambda ci sono infiniti valori di R e di I il cui prodotto può darmi quel valore di lambda (ad esempio, per fare 90 ci sono infiniti numeri che moltiplicati danno quel risultato).

Dal punto di vista percettivo R non può essere misurata direttamente, in quanto non c'è uno spettrofotometro nell'occhio che può misurarla direttamente.

Quindi questa equazione fondamentale, per quanto semplificata, è un buon modo di capire il problema e può essere usata per fare della psicofisica.

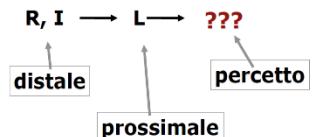
Ad esempio, possiamo chiederci come fa il cervello a darci un percetto di colore dato che a *livello prossimale* (sulla retina) c'è una distribuzione di luminanze, le quali però sono il prodotto di due *valori distali* (nel mondo esterno), inestricabilmente mescolate nella luminanza.

Quei punti di domanda (**catena psicofisica**) ci fanno capire che il cervello si rappresenta sia le proprietà dei materiali e sia quelle dell'illuminazione, perché noi quando guardiamo il mondo colorato abbiamo una



certa impressione del colore, delle superfici, cioè attribuiamo proprietà agli oggetti. E allo stesso tempo abbiamo una certa impressione soggettiva di come gli oggetti siano illuminati, non li misuriamo lo spettro dell'illuminazione, ma abbiamo un "senso" del fatto che siano più o meno illuminati e di quali sono più o meno le caratteristiche di illuminazione.

### catena psicofisica



In psicofisica del colore acromatico queste due quantità vengono chiamate, nonostante un po' di confusione dal punto di vista terminologico:

- **Brightness**: intensità percepita **della luce** riflessa da una superficie. È una proprietà dell'illuminazione, non della superficie stessa, si tratta dell'intensità della luce. In italiano viene chiamata "chiarezza" o "luminosità", ma è meglio utilizzare la terminologia inglese.
- **Lightness**: quella proprietà che attribuiamo alla **superficie**, cioè la "*riflettanza percepita*" o "il colore della superficie nella scala dei grigi".

La differenza fondamentale tra i due termini è che la *Brightness* è una proprietà che percepiamo attribuita **alla luce** che viene riflessa, mentre *Lightness* è una proprietà che si attribuisce alla **superficie**.

#### *Esempio di giudizio su Lightness e Brightness*

Se prendiamo una superficie scura, tipo una lavagna, la guardate e vi sembra di un certo colore (nera), questo è un giudizio basato sulla Lightness che si percepisce.

Se accendiamo una luce molto forte, la lavagna riflette molta più luce di prima perché c'è molta più illuminazione e quindi la lavagna risulta molto più illuminata, ma è comunque nera. Questo è un giudizio basato sulla Brightness.

Un'altra differenza importante è che la **riflettanza nel mondo fisico è costante**, mentre l'**illuminazione è variabile**. Per essere ancora più precisi, la riflettanza è costante nella misura in cui la composizione materiale fisico-chimico di un materiale è costante. Se prendiamo una foglia ad esempio, adesso o fra una settimana la foglia si è seccata e ha cambiato la sua composizione fisico-chimica, c'è stato un processo di ossidazione, per cui l'interazione con l'illuminazione si modifica e quindi sua riflettanza è cambiata. Nella pratica della "*percezione del qui e ora*" possiamo considerarla come una proprietà costante del mondo, mentre l'illuminazione è continuamente variabile.

Dato che la luminanza è una quantità psicofisica che dipende dal rapporto tra una costante e una variabile, sarà anch'essa mutevole e quindi la luminanza cambia ogni qualvolta cambia l'illuminazione.

#### Ritornando all'esempio della lavagna:

La lavagna ha una certa riflettanza fissa, essendo nera è un materiale con bassa riflettanza (possiamo immaginare ad esempio il 5%), ma a seconda di quanto viene illuminata la luminanza cambierà perché rifletterà in termini assoluti tanta più luce tanto più è intensa l'illuminazione: sarà sempre il 5% ma chiaramente il 5% di 10 è molto meno del 5% di 1000 e quindi la luminanza cambierà in funzione dell'illuminazione.

Dal punto di vista percettivo la differenza importante fra Brightness e Lightness è che: **la prima si attribuisce alla luce varia al variare dell'illuminazione** (più è intensa la luce incidente e più sarà intensa la quantità percepita di luce riflessa dalla superficie), mentre invece quando si cambia l'illuminazione, la luminanza cambia, ma la Lightness, il colore percepito attribuito alla superficie, rimane più o meno costante.



Se prendiamo un foglio di carta sia che lo guardiamo sotto la luce del sole a mezzogiorno, sia che lo guardiamo dentro un ambiente poco illuminato, ci sembrerà sempre bianco, ci accorgiamo che ciò che cambia è la fonte di illuminazione, ma il colore rimane costante.

**Brightness si attribuisce alla luce e cambia con l'illuminazione nella nostra percezione, Lightness invece è una proprietà che attribuiamo alla superficie e tende a rimanere costante anche al variare dell'illuminazione.**

Questo è un altro esempio di **costanze percettive**, ovvero tendiamo a percepire colori approssimativamente costanti, anche quando ci sono cambiamenti drastici nell'illuminazione.

Nell'ambito del colore acromatico pone il problema di capire come questo sia possibile, perché l'informazione che arriva all'occhio, considerata a partire dall'equazione fotometrica fondamentale, **è un segnale che varia continuamente al cambiare dell'illuminazione**.

**Studente:** “Il fatto che percepiamo il colore in modo costante nonostante l'illuminazione cambi, deriva dal fatto che noi variamo la nostra sensibilità in base alla condizione di illuminazione mesopica, scotopica ecc.””

**Prof:** “In parte sì, contribuisce, però il problema rimane anche se io rimango in un certo ambito di stato di adattamento della retina. Se immagini di mettere dei fogli di carta presi tutti dalla stessa risma di carta e ti appare bianco, immagina poi di metterle in zone diverse della tua stanza. In queste zone diverse della stanza l'illuminazione sarà un po' diversa, però ti sembra sempre bianco il foglio.

Cambiare lo stato di adattamento ha più una funzione di consentirti di fare questo tipo di processo di percezione acromatica quando cambi l'ordine di grandezza dell'illuminazione di tanto, di adattarti ad un ambiente poco o tanto illuminato, ma non è la soluzione al problema di cui stiamo parlando.””

Un altro modo di rendersi conto di questo problema è facendo riferimento a un paio di dimostrazioni famose: questa è stata inventata da **Edward (Ted) Adelson**, uno psicologo della percezione che lavora all'MIT. Si è occupato del ruolo di MT nel fare l'integrazione dei vettori di movimento.

Studiando il colore acromatico, ha inventato questa dimostrazione: “**Checker and shadow**” (scacchiera e ombra) in grado di spiegare la **costanza di bianchezza**, ovvero che la Lightness di una superficie non corrisponde alla luminanza, ma alla proprietà del materiale, come una codifica della proprietà del materiale.



La dimostrazione ha ad fare con il colore percepito degli elementi bianchi della scacchiera marcati come A e B:

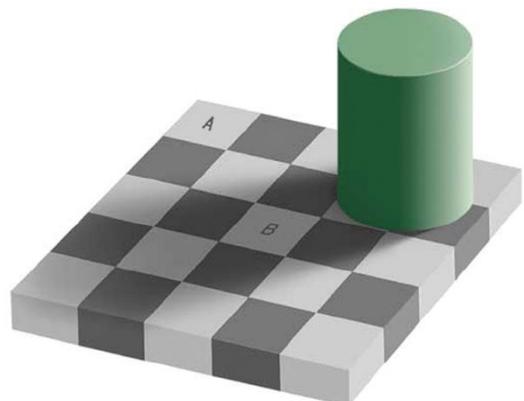
A e B appaiono, nell'immagine sopra, con la stessa lightness, cioè della medesima tonalità di grigio, però B appare in ombra, mentre A appare illuminato. Questo avviene a discapito del fatto che le luminanze di A e B sono molto diverse, per renderci conto di questo possiamo utilizzare un **fotometro** o uno **schermo di riduzione**: togliendo il contesto possiamo

vedere che B è molto più scuro di A. Però noi non vediamo questo, noi non vediamo la luminanza, noi vediamo ciò che il cervello fa per cercare di attribuire una proprietà alla superficie e contemporaneamente di separarla da quella che sembra essere la quantità di luce che la superficie riflette in funzione dell'illuminazione. **A e B sembrano dello stesso colore, ma hanno luminanze diverse.**

➔ Qui abbiamo visto un video che mostra il fenomeno opposto: due punti nello spazio hanno la stessa luminanza, ma hanno colore differente.

Quindi il colore acromatico percepito non è una funzione di L, non dipende dalla luminanza. Quindi dobbiamo chiederci **da cosa dipende? Qual è l'informazione che il sistema visivo utilizza per percepire il colore acromatico?** Questa domanda implicitamente riguarda anche la costanza bianchezza in quanto il colore acromatico tende a rimanere costante al variare di I.

## costanza di bianchezza



**Studente:** “Ma quindi la lightness esiste nel momento in cui si ha esperienza di quell’oggetto?”

**Prof:** “La lightness è il colore acromatico. Il colore, anche quello cromatico esiste nella tua testa. Il colore è una proprietà della percezione. È un aspetto della nostra vita mentale. È una delle dimensioni della nostra esperienze cosciente del mondo. Quindi, i raggi non sono colorati e la riflettanza non è colorata. Quello che possiamo dire è che il colore rappresenta un modo in cui il cervello etichetta le superfici.”

**Studente:** “Quindi se noi avessimo un oggetto che non abbiamo mai visto in un ambiente ambiguo e ad esempio percepiamo un grigio scuro non è una lightness in quel caso.”

**Prof:** “Noi tendiamo sempre quando vediamo una superficie ad attribuire una lightness e una brightness, ma hai ragioni, in certe condizioni questo può essere ambiguo.”

Studente: “Si può dire che il colore è un’interpretazione del nostro cervello della realtà?”

**Prof:** “Sì.”

**Studente:** “Se diciamo che si tratta di un’interpretazione, come si può dire che la reazione che abbiamo alla stessa realtà fisica porti ad uno stesso percepito tra soggetti diversi? Come posso dire che ciò che definisco rosso è lo stesso percepito che prova lei che lei chiama rosso?”

**Prof:** “Infatti non si può dire. È un classico dilemma-paradosso della percezione cromatica, o meglio, il paradosso delle altre menti. Dato che non abbiamo accesso diretto all’esperienza percettiva delle altre menti/soggetti, possiamo solo inferire questo sulla base di analogie e somiglianze strutturali.”

Quello che possiamo osservare direttamente in un compito psicofisico è la capacità di discriminare i colori, di dire se due colori sono diversi oppure uguali e per questo abbiamo delle tecniche psicofisiche molto affidabili che si basano sulle soglie. In questa maniera possiamo notare come la struttura percettiva dei colori di un individuo che ha anomalie nella percezione cromatica è diversa da quella di un organismo che non ha queste anomalie. Un daltonico sicuramente vede dei colori diversi da come li vedo io, però quale sia la qualità in modalità percettiva, la “rossità” di un rosso, come la definiscono i filosofi, per due individui differenti. Questo è ciò che i filosofi chiamano “*l’Hard Problem*”, il problema difficile della coscienza, il problema dei *qualia* e della qualità dell’esperienza e questo esula dal programma, è un problema di più alto dominio.

In questa sede ci limitiamo a descrivere a cercare di comprendere quale potrebbe essere la base di un aspetto della percezione cromatica, cioè il fatto che il colore rimane costante al variare dell’illuminazione e questo si può stabilire in soggetti diversi usando un compito psicofisico. Infatti tramite esperimenti di psicofisica si può di testare il **problema della costanza del colore acromatico**.

## Il principio del rapporto e il colore della Luna

Hans Wallach era un allievo di Wertheimer e Koffka, uno psicologo gestaltista ebreo e fuggito in USA, lavorò in un piccolo college vicino a Philadelphia vicino la Pennsylvania ed è stato un pioniere dello studio della percezione moderna. Ha eseguito lavori fondamentali sia sulla percezione del colore acromatico sia sulla percezione del movimento (il problema dell'apertura e l'effetto barber pole).

Il principio del rapporto spiega come mai la luminanza delle lettere non cambia, ma cambiando lo sfondo la loro bianchezza varia.

la luminanza di queste lettere non cambia

solo lo sfondo cambia

la luminanza di queste lettere non cambia

solo lo sfondo cambia

la luminanza di queste lettere non cambia

solo lo sfondo cambia

la luminanza di queste lettere non cambia

solo lo sfondo cambia

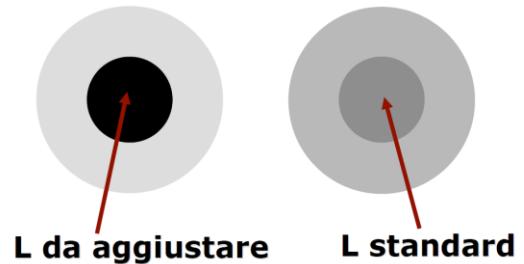
Questo suggerisce che la Lightness, ovvero il colore acromatico percepito, non è una funzione della luminanza, ma è qualcosa che dipende da una relazione, ovvero dipende dal contesto.

Wallach ha studiato questo fenomeno in modo sistematico, usando il **paradigma del confronto fra configurazioni disco-anello**.

L'esperimento consiste nel modificare attraverso dei filtri la luce proiettata da due diversi proiettori. Si generano degli stimoli semplici, come due dischi e due anelli e si può modificare la luminanza, cioè l'intensità della luce nel disco e nell'anello, in maniera indipendente.

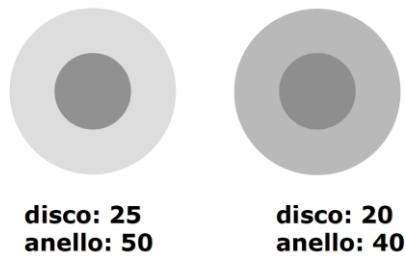
Si mette da una parte (nell'immagine è a destra) un disco con luminanza standard fissa e dall'altra parte si sceglie un anello e si aggiusta la luminanza, attraverso il metodo dell'aggiustamento di Fechner, del disco di confronto finché non sembrerà che il disco da aggiustare abbia la stessa bianchezza del disco standard. In pratica si fa un uguagliamento di bianchezza/Lightness. Se si fa questa cosa sistematicamente, con varie prove e varie luminanze standard, si trova che la lightness non dipende dalla luminanza, ma dal rapporto di luminanza fra il disco e l'anello. Cioè, quando la luminanza del disco e la luminanza dell'anello nella configurazione standard è uguale al rapporto fra la luminanza del disco e la luminanza dell'anello nella configurazione test i due dischi centrali sembrano avere la stessa Lightness.

## esperimento di Wallach



Ad esempio, a destra possiamo immaginare di avere un disco con luminanza di 20 (con unità di misura della luminanza = candele/m<sup>2</sup> ( $\frac{cd}{m^2}$ )) e l'anello, più chiaro,  $40 \frac{cd}{m^2}$ . Dunque il rapporto del disco standard è 1:2.

Il disco test apparirà uguale a quello standard quando nella configurazione test avremo lo stesso rapporto, ad esempio 25 e 50 o 50 e 100.



Questo è il principio di Wallach, per cui: la **Lightness non dipende dalla luminanza, ma dal rapporto fra luminanze adiacenti** e questo vale solo per le configurazioni semplici come il disco-anello, perché per le configurazioni più complicate la situazione è più complessa.

**Ciò equivale a dire che due grigi appaiono uguali se stanno in uguale rapporto con lo sfondo rispettivo.** Se il disco e l'anello, essendo adiacenti, ricevono la stessa luce incidente, allora il rapporto fra le luminanze sarà uguale al rapporto fra il prodotto della prima riflettanza per l'illuminazione diviso per il prodotto della seconda riflettanza per l'illuminazione. Dato che in questo secondo rapporto I è la stessa quantità perché è comune al disco e l'anello, allora il rapporto fra le luminanze L<sub>1</sub>/L<sub>2</sub> si riduce al rapporto fra le riflettanze:  $\frac{L_1}{L_2} = \frac{R_1*I}{R_2*I} = \frac{R_1*\#}{R_2*\#} = \frac{R_1}{R_2}$ .

Ossia, il principio di Wallach dice che se due superfici adiacenti sono ugualmente illuminate, allora il rapporto fra le loro luminanze corrisponde al rapporto fra le loro riflettanze, in un mondo acromatico. Dunque se due superfici sono illuminate allo stesso modo allora il rapporto fra le luminanze diventa indipendente dall'illuminazione. In questo modo abbiamo trovato qualcosa che rimane costante (il rapporto) anche al variare dell'illuminazione, ad esempio: anche se si cambia I, L<sub>1</sub> su L<sub>2</sub> rimane sempre uguale e corrisponde a R<sub>1</sub> su R<sub>2</sub>.

Questo è un principio in base al quale ci si aspetta che in tutte le condizioni in cui vale questa assunzione, cioè che quando l'illuminazione di due superfici è la stessa, allora il colore dovrebbe anche rimanere approssimativamente costante, perché il rapporto fra le due luminanze è rimasto costante. ➔ Questo è anche il motivo per cui la luna ci appare bianca.

La Luna è un oggetto che noi vediamo in condizioni assolutamente inusuali: di notte, in un ambiente vuoto e scuro dove non ci sono altre superfici. Questa è una particolare configurazione disco-contesto generale (contesto generale che ha sostituito l'anello) che contiene **solo un rapporto** fra due luminanze: la **luminanza della Luna e luminanza del contesto**. La Luna è illuminata dal sole, riflette molta luce e quindi ha un'alta luminanza, mentre il resto del cielo non ha superfici che possono riflettere la luce del

sole e per questo la “luce viene persa” (o per essere più precisi non c’è luce che torna indietro da quella zona del nostro campo visivo) e la conseguenza è che il contesto ha una luminanza molto bassa. In base al principio al principio di Wallach noi dovremmo vedere il colore della Luna in accordo col fatto che c’è una luminanza molto alta circondata da una molto bassa: un rapporto di luminanze molto ampio. È per questo motivo che noi vediamo la Luna bianca, ma la vediamo così solo perché è il prodotto di come noi attribuiamo il colore a quella superficie, non vuol dire che abbiamo correttamente recuperato una stima della riflettanza della Luna stessa.

Infatti se andassimo sulla Luna vedremo che è *grigia scura* e questo si può vedere anche dall’immagine dove è riportata una foto di un satellite della NASA. In questo caso la Terra è stata messa come sfondo, in una configurazione disco-anello rappresenterebbe l’anello, mentre la Luna il disco, e in questo caso la Luna è grigia.

**Sostanzialmente il colore viene attribuito in base al contesto**, il colore è qualcosa di *natura relazionale*, nel principio di Wallach il contesto è un anello monocromatico ed è molto semplice, ma in generale è così che decodifichiamo il colore. Nel caso della Luna non siamo in grado di recuperare il vero colore perché non siamo in grado di stimare correttamente la riflettanza della Luna, mentre nel nostro ecosistema vediamo neri gli oggetti con riflettanza bassa e bianchi quelli con riflettanza alta.

Riprendendo il discorso sui fiori: in certi casi il colore sembra andare insieme alla riflettanza, mentre in altri no, ed il motivo è dovuto al fatto che non siamo in grado di misurarla direttamente, ma proviamo a recuperarla, e quindi **attribuire un colore alle superfici**, in funzione di quello che è importante per noi nel nostro ecosistema. Il principio del rapporto funziona bene nel nostro ecosistema per determinare il “vero colore” delle superfici, mentre non funziona bene con la Luna perché ci appare bianca (sempre in funzione del principio del rapporto) anche se è grigio scuro, ma non è molto importante questo, perché a tutti gli esseri umani la Luna appare bianca e questo sottolinea ancora come **il colore è ciò che appare, non ciò che è da un punto di vista fisico**.

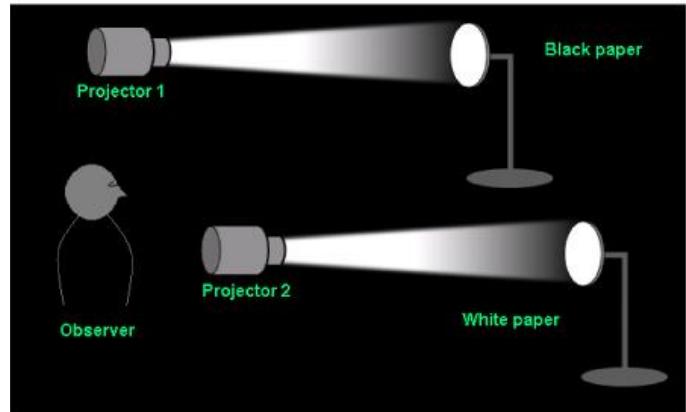


## Effetto Gelb

Da questo ragionamento discende una previsione contorta intuitiva, cioè il fatto che in laboratorio dovrebbe essere possibile produrre delle particolari illusioni sulla percezione del colore: ad esempio far apparire una superficie nera o grigio scuro come bianca. Tale fenomeno è stato riprodotto in laboratorio da **Gelb**, uno psicologo tedesco gestaltista, che ha la paternità sull'omonimo effetto.

*“L’effetto Gelb è un modo per portare in laboratorio l’effetto illusorio della Luna.”*

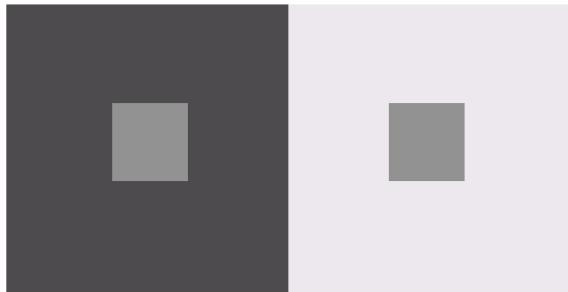
Per riprodurlo serve un cartoncino nero, lo si fissa su un supporto in grado di sospenderlo a mezz’aria, poi si spegne la luce e si accende un proiettore che illumina direttamente il cartoncino, senza altre superfici adiacenti al cartoncino: in questo modo esso apparirà bianco, come la Luna.



Nella stessa stanza in cui si è fatto questo si può anche aggiungere un’ulteriore condizione prendendo un altro proiettore che punta verso un cartoncino bianco e chiedere ad un utente di modificare la luminosità del secondo proiettore affinché il cartoncino bianco appaia uguale al cartoncino nero (che era stato precedentemente illuminato affinché apparisse bianco). Nonostante la riflettanza dei due cartoncini sia completamente diversa, infatti quello bianco ha un’alta riflettanza, mentre quello nero bassa, modificando l’illuminazione si possono ottenere due cartoncini che appaiono identici. Ovviamente l’illuminazione è differente, in quanto quella del proiettore che illumina il cartoncino nero deve essere molto maggiore di quella del proiettore che illumina il cartoncino bianco. *Si tratta di una violazione del principio del rapporto?* Infatti riducendo l’intensità della luce sul cartoncino bianco, le due luminanze sono differenti e quindi i cartoncini dovrebbero apparire diversi, però dobbiamo considerare che questi due dischi sono privi di un contesto, quindi non c’è un vero rapporto di luminanza, se non che c’è una zona molto illuminata circondata da una zona completamente buia.

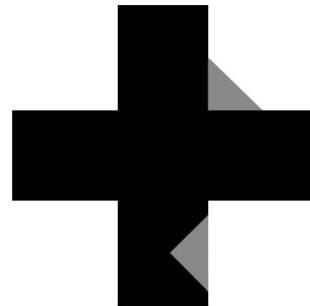
Dall’effetto Gelb si intuisce che bisogna capire come spostarci da un semplice ragionamento sui rapporti di luminanza a delle situazioni più complicate con dei contesti più articolati.

## contrasto simultaneo

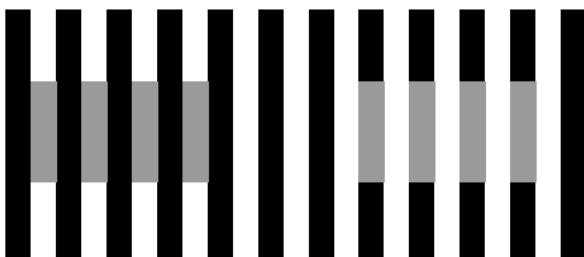


Questa è la classica illusione chiamata “**il contrasto simultaneo di lightness**” in cui i due quadratini grigi centrali hanno la stessa luminanza, però appaiono diversi a causa del fatto che stanno in rapporti diversi rispetto ai loro sfondi. Il quadrato di destra è più scuro dello sfondo e il quadrato di sinistra è più chiaro dello sfondo e i rapporti di luminanza sono opposti in quanto uno ha un incremento e l’altro un decremento. Pertanto i due quadrati appaiono diversi, ma non tanto diversi, perché in questa situazione non ci sono solo due rapporti disco-anello, ma c’è anche un terzo rapporto: quello fra i due sfondi. Quindi il sistema tiene conto di tutto quello che c’è nella figura, i due quadrati appaiono diversi, ma siamo ancora coerenti col classico principio del rapporto di Wallach.

Ma osservando una configurazione più complessa come la **croce di Koffka** in cui ci sono due triangoli grigi con stessa luminanza rivolti con il lato più lungo che confina con lo sfondo bianco, e quindi sono in rapporto con il bianco, e i due lati più corti che confinano col nero e che quindi sono in rapporto di luminanza con il nero.



Dal punto di vista dei rapporti di luminanza i due triangoli sono equivalenti in base al principio di Wallach, ma se si osservano bene si nota che uno dei



due è leggermente più bianco, quello in basso.

Koffka diceva che non sono solo **i singoli rapporti di luminanza a determinare la lightness, ma che essa dipende anche da dei processi di unificazione-segregazione, dalla geometria della scena, cioè da come il sistema unifica e crea le superfici rispetto agli sfondi.**

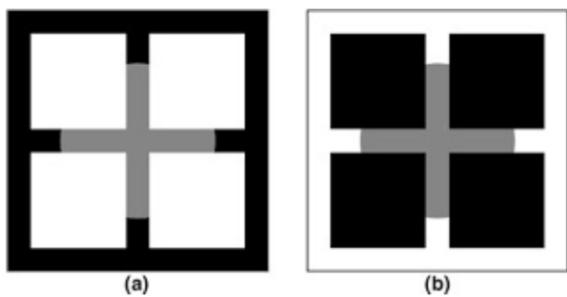
Ossia, il colore delle superfici è qualcosa che viene attribuito ad un processo di *middle vision*, non di *early vision*. Quest’ultimo infatti codifica i segnali locali (in questo caso i rapporti singoli di luminanza), ma poi tutto questo deve essere messo in un “tutto globale” che tenga conto dell’unificazione-segregazione. In questi termini, possiamo dire che il triangolo sopra sta sullo sfondo bianco, quindi viene visto soprattutto in rapporto al bianco e meno al nero e viceversa per il triangolo di sotto.

## Effetto White

Michael White era uno psicologo-psicoterapeuta australiano e dobbiamo ai suoi studi l'omonimo effetto: l' **effetto White**.

Nell'immagine osserviamo che il grigio della superficie di sinistra appare diverso da quello di destra nonostante i due grigi abbiano stessa luminanza. Questa illusione è molto più forte rispetto a quella della croce di Kofka, ma la spiegazione è la stessa: il grigio della superficie di sinistra pur avendo molti più contorni che confinano col nero, cioè molto più rapporto di luminanza col nero che col bianco, in realtà sembra essere dietro al nero, quindi viene visto prevalentemente in rapporto al bianco. Mentre quello di destra pur avendo molto più rapporto di luminanza col bianco, molto più contorno col nero, sembra essere davanti al nero e viene visto in rapporto al nero.

### variante di Todorovic



Questa a sinistra è una variante fatta da uno psicologo serbo di Belgrado, **Todorovic**, il quale ha trovato questa versione dell'effetto White che funziona ancora meglio. C'è un disco grigio chiaro a sinistra e un disco grigio scuro a destra. I dischi hanno la stessa luminanza, quello di sinistra è praticamente circondato solo da bianco e dovrebbe apparire più scuro, ma invece appare più chiaro e viceversa quello di destra, in cui il grigio confina quasi solo col nero per cui dovrebbe apparire più chiaro e invece appare più scuro. **Questo avviene perché il sistema che attribuisce il colore alle superfici non si basa solo sui rapporti di luminanza locali, ma tiene conto anche del contesto.** Quindi il disco di sinistra viene visto in rapporto allo sfondo nero anche se in realtà è uno sfondo nero coperto da i quadrati bianchi e viceversa a destra.

## Lezione 21

### La luna e il problema dell'ancoraggio

Il principio di Wallach ci dice che non bisogna pensare alla percezione del colore come il risultato di una certa quantità assoluta dello spettro di luce che arriva all'occhio, piuttosto bisogna pensare al colore acromatico di superficie su una scala di grigi, come un *meccanismo di natura relazionale* cioè un meccanismo che è sensibile ai **rapporti tra le luminanze di una figura e il suo sfondo**.

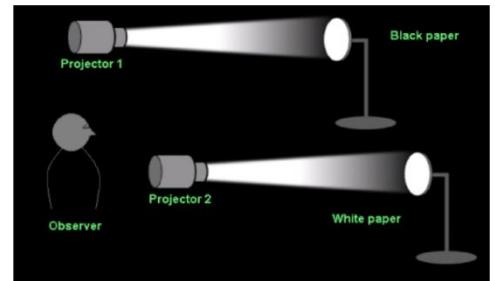
Si ritorna quindi alla **determinazione relazionale** e abbiamo visto come questa vale anche per i contorni, ma ancora di più vale per il colore.

Ovviamente non esiste solo questo come rapporto relazionale, ma ce ne sono altri, un esempio introduttivo utile a comprender il prossimo rapporto relazione è quello dell'**illusione del colore della luna**. Normalmente percepiamo la luna come bianca, ma è solo un'illusione perché in realtà la luna non è fatta di materiali bianchi, ma di materiali che sono di colore grigio scuro. Solo in condizioni particolari si può vedere che la luna è grigio scura, ma nel nostro mondo percettivo la luna ci appare sempre bianca.

### ***Ma perché ci appare sempre bianca?***

Il principio di Wallach ci dice che la luna ci appare bianca perché dipende dal rapporto fra un disco che è fortemente illuminato dal sole (che non vediamo perché è dall'altra parte della terra) e un campo che gli fa da sfondo che è il cielo buio, quindi è il **rapporto tra una luminanza bassa e una alta**, questo fa sì che la figura ci appare bianca e lo sfondo nero.

Si è parlato anche **dell'effetto Gelb** che ci permette di riprodurre l'illusione della luna. In una stanza buia, viene messo su un supporto rialzato un cartoncino di colore nero e viene illuminato in modo tale che la luce non arrivi in altre parti della stanza. Così osserviamo un equivalente della luna che è fortemente illuminata dal sole, ma circondata da un campo non illuminato e quindi uno sfondo molto scuro.



Lo stesso risultato si ottiene anche prendendo un cartoncino di colore bianco e illuminandolo di meno rispetto a quello nero. Quindi creiamo una situazione in cui c'è lo stesso rapporto fra le luminanze.

Non è importante di quale riflettanza è il materiale del cartoncino, ma quello che conta è il **rapporto fra le luminanze che creiamo tra figura e sfondo in funzione della riflettanza del cartoncino e dell'illuminazione che gli diamo**.

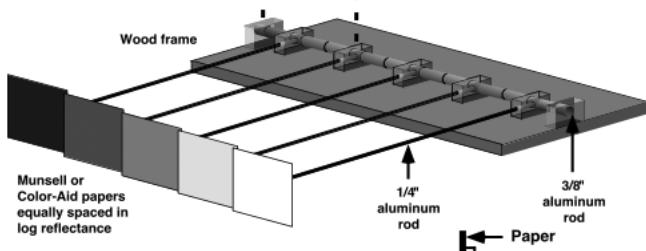
### **Il principio dell'ancoraggio di Wallach**

Wiki order of magnitude è una risorsa di Wikipedia che riassume tutta una serie di ordini di grandezza di quantità fisiche e tra i vari domini di questa risorsa c'è anche quello che riguarda la luminanza, ovvero l'intensità di luce che arriva all'occhio. Le stime del rapporto della luminanza in diverse condizioni sono:

- La luminanza della luna nel cielo di notte: rapporto enorme da 1 ad un 1000000;
- La luminanza del cielo al mattino con la luna: rapporto basso da 10 a 1000;
- La luminanza su sfondo bianco su un monitor è di 90 ad 1, mentre la luminanza su sfondo nero su un monitor è di 3 ad 1. Il rapporto di 30 ad 1 ( $90/3=30$ ) corrisponde al salto da bianco a nero. Ma la stessa cosa mi da bianco quando è un rapporto di 100 ad 1 o addirittura di 1000000 ad 1.

Non è così chiaro come il rapporto di luminanza determina il colore, perché ci sono rapporti molto diversi però poi il bianco è sempre bianco e il nero è sempre nero. **Come si spiega questo?**

Wallach si era accorto di questo problema e ci dice che un rapporto di per se' non è sufficiente per determinare un colore, perché questo rapporto deve essere *ancorato* mettendolo in relazione con una scala di grigi. E' sbagliato dire semplicemente che il rapporto da 1 a 30 è il passaggio da bianco a nero, perché in realtà potrebbe essere tante cose diverse. Così Wallach ha proposto il **principio dell'ancoraggio**, secondo cui se ho più rapporti presenti in una scena, quello che corrisponde al rapporto con la luminanza più alta di tutte, farà apparire la figura bianca. Quindi il bianco diventa la luminanza più alta in una scena mentre tutto il resto viene scalato in base ai rapporti che ci sono tra le luminanze. **Come si spiega questo?**



Pensiamo **all'effetto Gelb** e immaginiamo di avere 5 cartoncini con colori diversi, nero, grigio scuro, grigio intermedio, grigio chiaro e bianco, che vanno da una bassa riflettanza ad un'alta riflettanza e che riproducono tante lune di colore diverso.

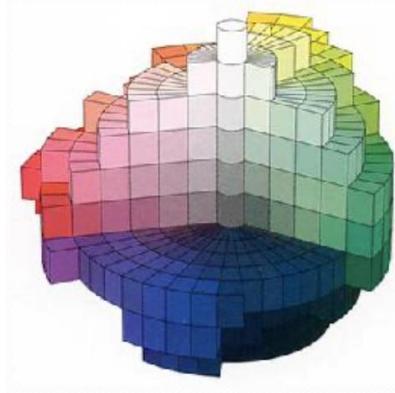
Montiamo i cartoncini su un supporto di alluminio con una base che permette ai cartoncini di essere in posizione rialzata e di ruotare verso l'alto così da non vederlo. Puntiamo la fonte di illuminazione verso il supporto, iniziamo a mettere il cartoncino nero puntato verso la luce e gli altri quattro invece ancora non si vedono. Fin qui abbiamo riprodotto l'effetto Gelb classico che riproduce l'illusione della luna perché il cartoncino nero diventa bianco.

Adesso però posizioniamo in direzione della luce il secondo cartoncino di colore grigio scuro e osserviamo che il cartoncino nero non appare più bianco ma grigio scuro, mentre quello grigio scuro adesso appare bianco. Questo è il principio dell'ancoraggio, **la luminanza più alta appare bianca e rispetto a questa, gli altri oggetti presenti nella scena vengono scalati**.

Ora posizioniamo anche il terzo cartoncino che è grigio intermedio verso la fonte luminosa e vediamo che il cartoncino grigio intermedio adesso appare bianco, mentre il cartoncino che prima appariva bianco che è grigio scuro, ora appare grigio intermedio e quello nero ora appare grigio scuro. Adesso posizioniamo il quarto cartoncino di colore grigio chiaro verso la fonte luminosa, questo appare bianco e tutti gli altri vengono scalati di conseguenza. Quindi il terzo cartoncino appare grigio chiaro, il secondo appare grigio medio e l'ultimo appare grigio scuro.

Infine posizioniamo il quinto cartoncino di colore bianco verso la fonte luminosa e vediamo che diventa bianco e tutti gli altri vengono scalati. A questo punto vedo i colori come sono davvero.

## Colore cromatico

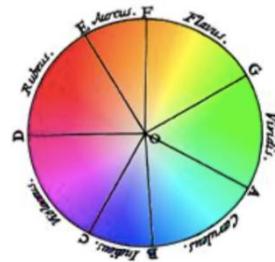


*Come possiamo caratterizzare dal punto di vista percettivo il colore? Quali sono le dimensioni e la struttura nella nostra esperienza cosciente dei colori?*

In figura vediamo una rappresentazione dei colori di tipo circolare, si tratta della versione moderna che in precedenza aveva dato **Newton**, il quale aveva fatto esperimenti di scomposizione della luce solare in diverse tonalità di colore ed aveva osservato una certa predisposizione dei colori distribuiti nello spettro.

Nella versione moderna l'idea è che ci sono **tre colori primari**: **rosso**, **blu** e **giallo** e possiamo pensare agli altri colori come il risultato di una miscela fra questi tre. Per esempio se prendiamo il blu e il giallo e li mescoliamo e se queste miscele sono paritarie in quantità, il risultato è una miscela di verde. Se ci mettiamo un po' più di giallo avremo una miscela giallo-verdastro, mentre se mettiamo un po' più di blu, avremo una miscela di blu-verdastro. Oppure se prendiamo il giallo e il rosso in egual quantità, creiamo l'arancione. Se invece aumentiamo la quantità di giallo avremo un giallo-rossastro, invece se aumentiamo la quantità di rosso avremo un rosso-giallastro.

Questo è un tentativo di organizzare la struttura percettiva dei colori che obbedisce al *criterio della scomposizione*. Quando guardiamo l'arancione vediamo un po' di rosso e un po' di giallo, mentre quando guardiamo il rosso, il giallo e il blu, che sono dei colori primari, non riusciamo a vedere miscele intermedie. Questa è l'idea che è stata poi sviluppata anche da altri dopo Newton come ad esempio **Goethe** che ha scritto un libro sulla teoria dei colori, in cui propone un circolo delle tinte dove il rosso è sostituito con il magenta, ma il principio di base rimane lo stesso.



Ancora oggi nei modelli contemporanei ritroviamo lo stesso principio e il colore viene visto esattamente come la nostra esperienza percettiva, ovvero in tre dimensioni:

1. La **tinta**: si muove lungo il circolo cromatico;
2. La **saturazione**: rappresenta quanto è pura quella particolare tinta rispetto all'assenza di tinta (il colore cromatico), quindi è la distanza sul cerchio di tinte da quell'asse centrale che rappresenta la dimensione del bianco al nero della visione cromatica;
- La **brightness**: rappresenta il livello di grigio che c'è dentro al colore, cioè quanto è vicino al bianco e quanto è vicino al nero.

In figura abbiamo uno spazio in 3D ove è rappresentato un doppio cono. È tridimensionale perché le dimensioni sono la tinta, la saturazione e la chiarezza, ma ha la forma di un doppio cono perché quanto più ci spostiamo verso il basso o verso l'alto in direzione della chiarezza, tanto più andiamo verso il nero e verso il bianco e meno avremo un colore saturato. Possiamo aggiungere del grigio ad un colore per saturarlo, ma se mettiamo del bianco non abbiamo più colore. Quindi la possibilità di avere un colore saturato si riduce sempre di più e di conseguenza la punta di questi due coni corrisponde a delle situazioni in cui la chiarezza è massima o minima e non c'è più ne differenza né di tinta né di saturazione.

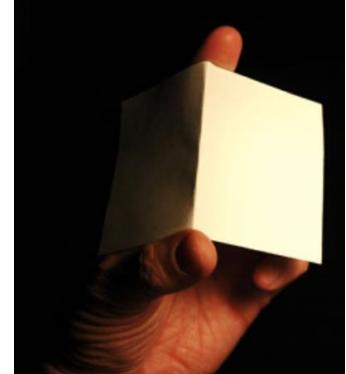
Questo è il **sistema Munsell** (valore, croma e tinta) che definisce ogni colore in riferimento ad un cartoncino appositamente tarato e questi cartoncini possiedono ognuno una coordinata lungo tre dimensioni che nel sistema Munsell sono tinta, croma e valore. **Il croma corrisponde alla saturazione, il valore corrisponde alla brightness e la dimensione lungo l'asse dal nero al bianco è quella cosa che varia muovendosi lungo il cerchio delle tinte.** Il sistema Munsell non è solo un sistema teorico, ma è concreto nel senso che è fatto di una serie di cartoncini colorati che possiamo riconoscere mettendoli a paragone con la parete della nostra stanza per esempio e trovare le coordinate che nel sistema Munsell corrispondono al colore simile alla parete. Questo fa sì che non sia esattamente un doppio cono, ma sia un **doppio cono irregolare** in cui mancano determinati pezzi e questo perché la struttura ideale nei colori naturali non si riscontra veramente. Nel sistema Munsell abbiamo più colori, come ad esempio il *marrone*, che non erano presenti nel modello di Newton.

Nel nostro computer sono presenti diversi modelli della percezione del colore che ci consentono di determinare che tipo di colore vogliamo utilizzare sul nostro schermo. Possiamo utilizzare il **sistema RGB**, sono tre dimensioni che corrispondono a rosso, verde e blu (quanto ci vogliamo mettere di ognuno dei tre colori). Si può scegliere anche il modello di Munsell che è **HSB**.

## Diedro di Mach

Alcuni psicologi della percezione nel secolo scorso si sono interrogati su quale fosse la maniera giusta di caratterizzare l'esperienza percettiva del colore e alcuni hanno dato una risposta diversa rispetto ai modelli visti fino adesso.

**Ernst Mach** (1838-1916) aveva osservato che il *diedro* corrisponde allo spigolo che si forma piegando a metà un cartoncino bianco, e se lo si espone ad una fonte luminosa si osserva che una metà risulta essere più illuminata dell'altra. Se poi il diedro viene anche posizionato verso di sé, si chiude un occhio tenendo la testa ferma e lo si fissa per qualche secondo, si nota che il diedro si appiattisce a formare due trapezi piatti. Questo succede perché chiudendo un occhio si è tolta la *parallasse binoculare*, non ci sono indici pittorici nel diedro se non che la forma tridimensionale delle due facce che proietta sulla nostra retina dei trapezi piatti. Si può anche riuscire ad invertire i trapezi e vedere che non punta più verso di se', ma punta verso la direzione opposta. Quindi quello che si nota quando il diedro si appiattisce e quando cambia di direzione, è che *i colori delle due facce del diedro cambiano*, non vediamo più una superficie bianca con una parte più illuminata e una meno illuminata, ma vediamo da una parte un bianco e dell'altra parte un grigio scuro. **Questo cambiamento si accompagna ad un cambio di qualità del colore in funzione della tridimensionalità apparente dell'oggetto e in funzione della sua relazione rispetto alla fonte di illuminazione.**



Questo tipo di osservazione del cambio drastico della qualità del colore, ha fatto molto pensare lo psicologo tedesco **David Katz**, che fece la sua tesi di dottorato sui *modi di apparenza del colore* e si chiese se i loro modi di apparenza sono riconducibili alle tre dimensioni di tinta, saturazione e chiarezza.

I diversi modi di apparenza dei colori per Katz sono 6:

1. **Colore di superficie:** corrisponde alla lightness, ha una certa proprietà specifica perché appare solido e compatto non si può attraversare con lo sguardo. Il colore superficiale si vede solo su una scena complessa con un certo grado di articolazione geometrica;
2. **Colore di riduzione o colore filmare:** lo apprezziamo se togliamo il contesto, quindi riducendo l'esperienza di colore a grado 0. Questo fenomeno lo possiamo osservare creando un buco di circa 1 cm di diametro e tenendo uno schermo davanti ad un occhio e fissando una superficie. Quello che appare attraverso il buco è che si vede ancora una figura colorata ma non ha più la compattezza di superficie, diventa piuttosto una specie di nebbia colorata, vediamo una minima esperienza di

colore (grado 0 del colore). Questa differenza è molto chiara, ad esempio se guardiamo una parete bianca se la osservo attraverso il colore di riduzione probabilmente diventa una specie di nebbiolina grigia e non più bianca. Questo richiama un po' la differenza tra lightness o brightness ma in questo caso stiamo parlando di colori non più di colori acromatici. Questa caratteristica è quella che si cerca di valutare quando si fanno gli esperimenti di colorimetria;

3. **Colore trasparente:** il colore può essere visibile su una superficie parzialmente trasparente;
4. **Colore di volume:** all'interno di una superficie trasparente viene messo del liquido o del gas che ha un suo colore e il colore ha delle caratteristiche percettive diverse da quelle del colore delle superfici (pensate al colore del mare);
5. **Colore di scintillio o colore riflesso:** certi colori sembrano avere una loro caratteristica legata all'illuminazione che vediamo scintillare su una superficie liscia;
6. **Glow o colore legato all'impressione:** il colore viene riferito ad una fonte luminosa cioè sembra appartenere non alla superficie opaca o metallica che produce degli effetti di luccichio ma può essere proprio attribuito ad un qualcosa che non riflette luce.

I primi due *modi di apparenza* di colore sono i più importanti.

#### APPROFONDIMENTO: GLI ESPERIMENTI DI COLORIMETRIA

Tutti i modelli colorimetrici sono basati su delle procedure psicofisiche in cui una persona deve fare un *compito di eguagliamento*. Si ha un disco, all'interno del quale una persona vede un colore di riduzione che ha una certa caratteristica spettrale. Ci sono dei controlli come delle manopole attraverso le quali il soggetto può manipolare la composizione spettrale di un altro disco che è quello di confronto. Il compito è appunto quello di eguagliare i due dischi. In un soggetto con un sistema visivo sano, è possibile uguagliare qualsiasi disco-test mescolando non più di tre colori primari → **principio della tricomatricità del sistema visivo umano**.



Questo è il motivo per cui è possibile dare una *sistematizzazione dell'esperienza dei colori* di riduzione utilizzando le tre dimensioni percepite, ma è anche il motivo per cui tutti i modelli colorimetrici hanno tre dimensioni, ed è anche il motivo per cui per riprodurre il colore su uno schermo ci servono tre diedri colorati che poi si mescolano.

## Illusioni percettive del colore

I modelli neuronali della percezione del colore non sono in grado ancora di spiegare certi fenomeni come le illusioni sulla percezione del colore.

Alcuni esempi si ritrovano in rete e tra questi quelli che hanno suscitato più scalpore sono stati il colore di una scarpa e di un vestito poiché vi è un disaccordo sulla percezione del colore.

Uno studio scientifico intitolato *the dress* ha reclutato 1400 partecipanti ai quali veniva chiesto di indicare di che colore percepivano il vestito in figura.

Le alternative di risposta erano:

1. Blu-bianco;
2. Bianco-oro;
3. Blu-nero;
4. Nero-oro.



Il 50% ha indicato correttamente il colore del vestito che è blu-nero quindi questo significa che 1 su 2 sbaglia. Un 30% ha indicato l'opzione bianco-oro, un 10% ha indicato blu-bianco ed infine un 10% ha indicato di vedere colori misti.

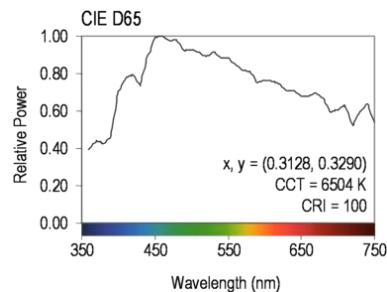
Queste situazioni si verificano perché la fotografia è *ambigua nell'illuminazione*, di conseguenza il nostro sistema visivo non riesce ad attribuire nella dimensione cromatica il colore corretto. A seconda del tipo di assunzione che il sistema elabora per percepire l'illuminazione (che è molto soggettivo) attribuisce un colore diverso.

## Lezione 22

### Il segnale cromatico

Il colore, da un punto di vista percettivo, è molto *complesso e multidimensionale*, di seguito vedremo **come** il cervello produce il percepito del colore. Il primo concetto da analizzare è chiamato dagli psicofisici: **segnale cromatico**.

**La riflettanza di una superficie è quella funzione che descrive come la superficie riflette la luce in funzione della lunghezza d'onda, cioè il rapporto tra luce riflessa e luce incidente in funzione di Lambda  $\lambda$ .** Gli oggetti naturali hanno delle forme tipiche della funzione di riflettanza: infatti **l'illuminazione solare ha una certa distribuzione di potenza (cioè di intensità) in funzione della lunghezza d'onda.** La luce solare è approssimativamente abbastanza piatta in funzione di Lambda: si dice “*luce bianca*” perché contiene luce relativamente intensa a più o meno tutte le lunghezze d'onda.



CIE standard illuminante D65 è stato definito dalla Commissione Internazionale per l'Illuminazione (Commission Internationale de l'Éclairage - **CIE**) che esiste da tantissimi anni e che ha il compito di fissare gli standard illuminotecnici. Questo **CIE D65**, in particolare, si riferisce alla sensibilità alla luce del sole a mezzogiorno.

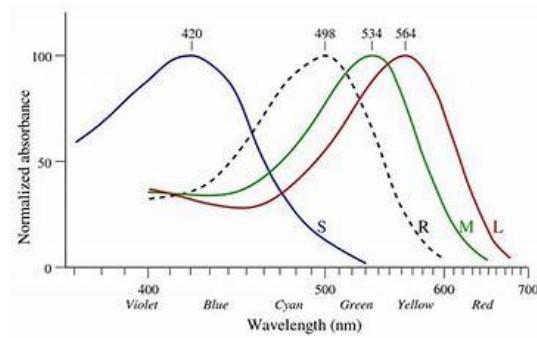
### **Qual è l'effetto di ciò sul sistema visivo?**

Nella *visione di basso livello* c'è la retina, su cui vi sono i fotorecettori (coni e bastoncelli); i fotorecettori in grado di vedere i colori sono i *coni* che garantiscono la visione dei colori in fotopica, diurna.

### **Come rispondono i coni alle diverse lunghezze d'onda?**

Le 3 tipologie di cono sono sensibili a una gamma relativamente ampia di luce a diverse lunghezze d'onda. Questo è rappresentato dal grafico in cui viene plottata la *sensibilità relativa* che dipende dalla capacità di assorbire i fotoni per i **coni S**, **coni M** e **coni L** in funzione della lunghezza d'onda.

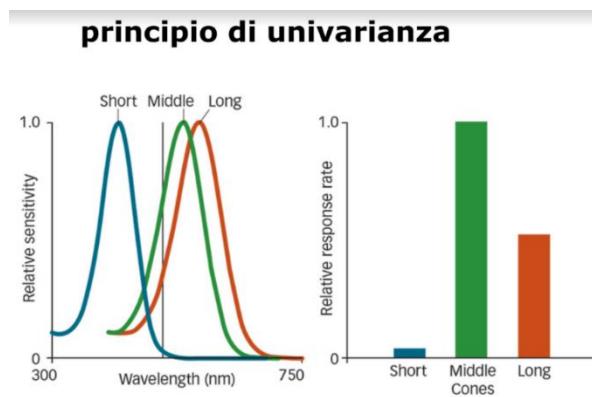
- I **coni S**, che sono etichettati con una curva blu, tendono ad essere sensibili di più alla luce nella zona delle **lunghezze d'onda più corte**, anche se la sensibilità si estende fino a quelle intermedie;
- I **coni M** e i **coni L** sono, invece, più sensibili a delle **lunghezze d'onda intermedie o lunghe** con una *lieve sovrapposizione* tra le 2 curve di sensibilità.



## Principio di univarianza

Vuole dire che un singolo cono (S, M o L) in realtà non è selettivo per la lunghezza d'onda; la **visione a colori** dipende da un sistema che distingue grazie alla codifica **congiunta** di tutti e tre i coni; perchè l'unica cosa che può fare il cono è produrre una risposta graduata più o meno intensa in funzione dell'*intensità* e della *lunghezza d'onda* dei fotoni della luce con cui interagisce.

Di fatto, il cono non sa qual è la lunghezza d'onda della luce che l'ha colpito: l'unica cosa che fa il cono, in sé e per sé, è produrre una risposta in termini neurali.



ES: Se immaginassimo di stimolare la retina con la luce monocromatica, si otterrebbero 3 risposte diverse nei 3 tipi di coni:

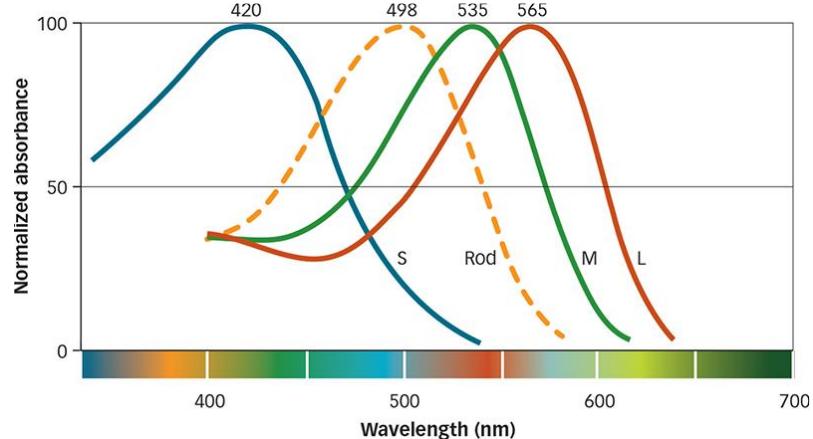
- la risposta sarà debole per i **coni S**;
- particolarmente intensa per i **coni M**;
- un po' meno intensa per i **coni L**.

Questo, semplicemente, perché quella luce monocromatica si colloca in un certo punto dello spettro e, in quel punto dello spettro, ci sono le sensibilità relative dei tre tipi di coni.

Si consideri la risposta di un cono teorico (linea tratteggiata), in funzione di Lambda, che abbia una certa curva che ne descrive la sensibilità spettrale:

- questo cono risponde bene al centro, intorno ai 500 nanometri, e poi la sua sensibilità decade andando verso le lunghezze d'onda più lunghe o più corte.  
 - Si immagini, ora, che questo cono venga *stimolato* con una luce di circa 550 nanometri: la sua risposta sarà *l'altezza della curva* in quel punto che leggiamo sull'asse delle Y ed avrà un certo valore.

- Però, ci rendiamo conto che esiste un'altra luce di lunghezza d'onda completamente diversa che produrrà esattamente la stessa risposta, proprio perchè questa è la forma della sua sensibilità spettrale.



Quindi, esistono tutta una serie di *coppie di lunghezze d'onda* diverse che producono un'onda, la quale a sua volta produce la stessa risposta recettoriale.

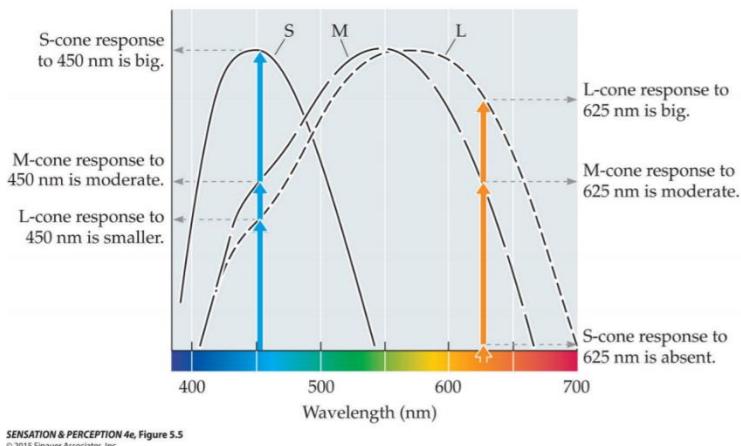
**NB:** Il cono *non* è in grado di distinguere: infatti può variare in una sola direzione (**principio di univarianza**) ossia produrre un potenziale più o meno forte e lo farà in funzione della sua sensibilità spettrale e anche dell'intensità della luce che lo colpisce.

Quindi si vede che per ognuna delle due luci monocromatiche, quello che sarà vero è che **ogni tipologia di cono produrrà un certo livello di risposta.**

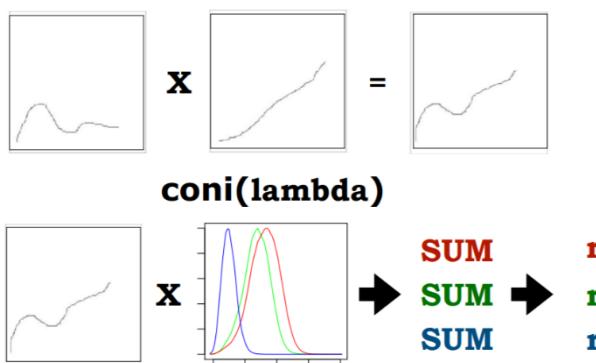
Per esempio, per la **luce azzurra** (450 nanometri), ci sarà una risposta debole da parte dei coni L, un po' più forte da parte dei coni M e decisamente più forte da parte dei coni S; mentre invece per la **luce arancione** (630 nanometri) non ci sarà praticamente risposta da parte dei coni S, un po' di più nei coni M e maggiormente da parte dei coni L.

**Quello che consente al sistema di distinguere fra queste due luci monocromatiche non è la risposta di un singolo cono, ma è la codifica congiunta da parte di tutti e tre i tipi di coni.**

**NB:** quello che è cruciale per la codifica dell'informazione, che poi viene utilizzata per produrre la percezione del colore, è il fatto che dai coni S, M e L escono 3 segnali che corrispondono a 3 numeri (le attivazioni di S, M e L).



**R( $\lambda$ ) I( $\lambda$ ) L( $\lambda$ )**



È come dire che lo *spettro della luce d'ingresso* viene ridotto/condensato in una misura che rileva quei tre numeri. Questa misura è ciò che gli psicofisici del colore chiamano **segnale cromatico**:

$$R \times I = L$$

L'equazione fotometrica fondamentale si è vista nel caso acromatico però nel caso di adesso stiamo parlando del colore in senso proprio; quindi, avremo:

- una **funzione di Riflettanza (R)** della superficie, in funzione di lambda (la lunghezza d'onda);
- una certa **Illuminazione (I)** che corrisponde alla distribuzione della potenza spettrale dell'illuminante, in funzione di lambda;
- Il prodotto di Riflettanza ed Illuminazione è l'equivalente della **Luminanza (L)**, nel caso acromatico, anch'essa in funzione di lambda.

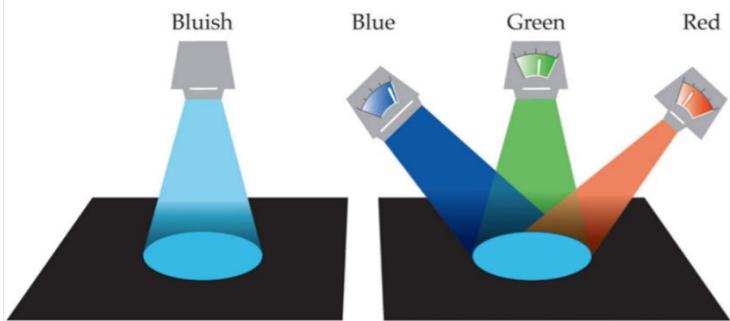
**NB:** È come *moltiplicare* ogni punto del vettore che descrive R di lambda per i punti corrispondenti del vettore che descrive I di lambda. Questi prodotti producono una nuova funzione di lambda che è la luce che viene *rimbalzata* dalla superficie che deve essere, poi, ulteriormente pesata con la sensibilità spettrale dei tre coni. All'interno di ogni classe di cono, il segnale che arriva all'occhio viene pesato in ogni punto e poi l'esito di questo prodotto in ogni punto della curva blu, verde e rosse mi danno tre numeri. Questi numeri sono l'integrale di questa funzione che risulta dal prodotto. Si possono pensare come a dei vettori di cui poi si sommano i valori fino ad ottenere un singolo numero.

**Questa tripletta di numeri dei coni S, M e L è quello che viene chiamato segnale cromatico:** è il punto di partenza, **il segnale da cui parte il processo che poi produce la percezione del colore.**

Questo è anche il motivo per cui il nostro sistema visivo è **un sistema visivo tricromatico**. È un'idea che è stata scoperta ben prima che noi potessimo misurare la sensibilità spettrale dei coni, di due ricercatori della visione **Thomas Young** e **Von Helmholtz** che, sulla base di osservazioni uditivo-psicofisiche, giungono alla conclusione che **il sistema visivo dal punto di vista della percezione del colore è tricromatico, cioè codifica il colore lungo tre dimensioni che dipendono dalle caratteristiche delle tre tipologie di coni.**

## Il principio di tricromaticità

In particolare, Thomas Young è anche responsabile della procedura di **uguagliamento cromatico** che è la base della colorimetria. Ha scoperto, utilizzando il confronto tra due luci, che data una luce standard (per esempio bluastra) è sempre possibile trovare una **miscela di al massimo tre luci monocromatiche**, tale per cui la miscela appare indistinguibile dallo standard.



Ad esempio, si può produrre una miscela fatta di tre luci monocromatiche che è in grado di produrre uno stimolo che è equivalente (viene visto identico) ad una qualsiasi luce standard. Questa è la base di tutti i metodi che utilizziamo oggi per riprodurre il colore sui nostri strumenti tecnologici: a partire dallo schermo del

computer, cellulare, scanner a colori, ecc. Tutti questi strumenti non emettono uno spettro di luce che è veramente equivalente a quello di un oggetto arbitrario nel mondo quando vogliono riprodurre una fotografia colorata, bensì sfruttano questo principio di Young e Helmholtz (**principio di tricromaticità**) e ne producono degli stimoli equivalenti al colore desiderato miscelando tre luci.

Esempio: lo schermo di un computer → Ogni pixel è fatto di tre diodi che sono tre luci monocromatiche (3 colori primari). In realtà non c'è nulla di primario, sono semplicemente tre luci monocromatiche scelte in posizioni diverse dello spettro in maniera tale che sia possibile emettere ciò che poi si vede; più luci si aggiungono alla miscela, meglio si riuscirà a riprodurre lo spettro ricercato.

Una cosa di cui c'è una prova molto evidente è che le persone sono poco consapevoli dell'importanza di come mettono lo schermo di un computer nell'ambiente, cioè come lo orientano rispetto le fonti di illuminazione; di come organizzano l'illuminazione nella stanza quando lavorano al PC. Una cosa che può essere estremamente stancante, non per l'occhio, ma per il nostro sistema visivo è il fatto di avere dei riflessi, oppure di avere delle luci sovrapposte al monitor che ne riducono il contrasto.

Questo principio di tricromaticità è, quindi, quello che viene sfruttato per la *riproduzione dei colori*. Esistono dei sistemi che consentono di specificare le caratteristiche di un colore, cioè di misurarlo.

**NB:** misurare un colore non significa misurare la riflettanza o l'illuminazione, ma vuol dire trovare un sistema e, quindi, un modello matematico che è in grado di identificare con i numeri in maniera univoca un colore.

Esempio: quando si va al negozio con un campioncino di pittura e si chiede di riprodurre un colore uguale, viene utilizzata una macchina che analizza il colore del campione e ne fa una pittura identica (o abbastanza simile).

Ai fini della *riproduzione* si creano gli **standard colorimetrici** e modelli matematici per la specifica del colore, ma è diverso dal modello di come è strutturato il colore **percepito**. Ci sono vari modelli che vengono continuamente aggiornati.

## Il modello CIE

### CIE

**spazio dei colori XYZ CIE 1931**

**derivato da esperimenti di  
uguagliamento svolti negli anni '20**

**modello per predire il metamerismo**

**XYZ = valori tristimolo = proporzione  
di tre "colori primari" teorici  
necessaria a uguagliare un "colore"  
test**

corrispondono ad una posizione in uno spazio tridimensionale teorico. Ogni *tripletta* identifica un colore e corrisponde alla proporzione di tre colori primari (che sono teorici non sono reali) le cui proporzioni sono necessarie a eguagliare un colore test.

Il modello originario è stato formulato nel 1931 dalla CIE ed è in continuo aggiornamento e raffinamento; è chiamato **"spazio dei colori XYZ"**. È stato derivato da esperimenti di uguagliamento per consentire di sviluppare un modello matematico che predice il *metamerismo cromatico*, cioè predice quali triplette di colori primari si devono combinare per produrre un dato colore.

I valori XYZ vengono chiamati **"valori tri-stimolo"**, e

## Dal segnale cromatico al colore percepito

Abbiamo definito il concetto di segnale cromatico, quindi analogamente alla maniera in cui abbiamo definito la luminanza (prodotto della riflettanza per l'illuminazione con la sensibilità spettrale dei coni). Questo pone un problema, **perchè l'illuminazione non è costante**, ma varia continuamente, quindi il segnale cromatico è qualcosa che non migliora direttamente una proprietà della superficie (noi vediamo un colore superficiale, lo attribuiamo alla superficie e soprattutto lo percepiamo come un qualcosa di relativamente stabile nello spazio e nel tempo → gli oggetti non cambiamo di colore in funzione dell'ora del giorno/clima/luce secondaria, **ma in funzione dell'ora del giorno cambia la forma della funzione di Lambda**). L'illuminazione cambia non solo per **l'intensità** (perchè andando verso il tramonto l'intensità dell'Illuminazione si abbassa), ma anche per la **composizione** (viene tagliata la parte più corta dello spettro in funzione della fase del giorno, per condizioni atmosferiche, presenza di luci secondarie, etc.). Prendendo in esempio un animale che vive nella foresta amazzonica, la luce del sole gli arriva all'occhio dopo essere stata ulteriormente riflessa da tutta una serie di superfici che sono principalmente verdi; quindi, tendono ad assorbire parte dell'illuminazione e rifletterne altre e questo cambia la composizione spettrale dell'illuminante.

Il problema è che il segnale cromatico è qualcosa che **varia in funzione** di tutti questi continui cambiamenti di I di lambda.

## La costanza del colore superficiale

La composizione/potenza spettrale relativa dell'illuminante CIE D65 su riportato, rappresenta il sole a mezzogiorno. Se cambiassimo l'illuminante e, invece di usare il sole a mezzogiorno usassimo una lampada al tungsteno (fonte di illuminazione artificiale, definita dallo standard CIE come *illuminante A*), si noterebbe che la scena prende una tinta giallastra, perché  $R \times I$  dà un risultato diverso, dato che  $I$  ha una forma diversa.

Nel mondo reale, quando l'illuminazione cambia (quando entriamo in una stanza dove c'è una lampada al tungsteno, dopo essere stati prima fuori sotto il sole), osservando le superfici non si notano cambi drastici nei colori.

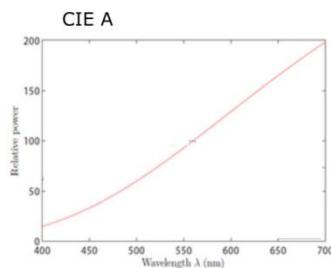
*“Il colore rimane relativamente costante al cambiare dell’illuminazione.”*

Questa è un'altra forma di quel principio generale che è la stabilità del mondo percepito; rientra tra le **costanze percettive** (costanza di posizione, della forma, di grandezza) e si chiama → “**Costanza del colore superficiale**”.



La costanza del colore non è perfetta, così come ogni costanza percettiva. Ci sono situazioni (illusione del vestito o della scarpa) in cui l'ambiguità sull'illuminazione con cui è stata creata quella fotografia può produrre delle **differenze individuali** di come una persona vede il colore nella stessa fotografia e alle stesse condizioni. Di solito, queste violazioni della costanza del colore sono associate al fatto che la superficie viene osservata in condizioni di illuminazione artificiale, particolarmente ristretta. Le luci che si utilizzano nelle strade quando c'è molta nebbia hanno una composizione spettrale particolarmente inusuale e ristretta in certe zone dello spettro.

## lampada al tungsteno



*Com'è possibile che il colore delle superfici rimane relativamente stabile nonostante il fatto che il segnale cromatico associato a queste superfici è continuamente variabile, a causa dei cambiamenti di illuminazione?*

Due ingredienti sono cruciali per produrre una percezione del colore approssimativamente costante al variare dell'illuminazione, a partire da segnali cromatici che cambiano:

1. **Informazioni di tipo relazionale (i rapporti)**: non sono i segnali cromatici di per sé, ma sono rapporti fra segnali cromatici che il sistema utilizza e questo trova una specie di analogia con il principio di Wallach;
2. **il contesto.**

Entrambe queste idee sono state chiarite per la prima volta da **Edwin Land**, considerato da Bruno l'ultimo grande inventore poliedrico della storia della tecnologia del '900. <sup>6</sup>Land è stato un vero e proprio pioniere della ricerca sulla fotografia a colori e gli va dato il merito di essersi reso conto che per riuscire a risolvere il problema di come fare le foto a colori, era **necessario capire come funziona la percezione cromatica nell'uomo**. Negli anni '60 Land ha fatto degli esperimenti considerati dei grandi classici della psicofisica della percezione del colore: come il sistema visivo umano produce dei colori approssimativamente costanti a fonte di cambiamenti continui dei segnali cromatici, causati da differenze nell'illuminazione.



Questi sono noti come **esperimenti con i Mondrian**. Spesso, in psicofisica, il progresso lo si fa quando si trova lo **stimolo giusto per fare degli esperimenti che catturano quello che interessa studiare**: quindi uno stimolo che non sia troppo complesso, ma abbastanza complesso da catturare il fenomeno di interesse. Land si era reso conto che per riuscire a creare un modello matematico della costanza del colore era necessario acquisire dei dati, utilizzando degli stimoli abbastanza complessi dal punto di vista cromatico, ossia stimoli che fossero ricchi di molti colori diversi e che, attraverso questa ricchezza e complessità, era possibile immaginare un algoritmo che avrebbe prodotto il primo esempio di soluzione al problema della

---

<sup>6</sup> Egli ha inventato la macchina fotografica istantanea, ha smesso di studiare, ha fondato una sua compagnia ed ha iniziato a realizzare macchine fotografiche (la polaroid) decidendo di inventare questo sistema che consentiva, per la prima volta, di fare la fotografia e nel giro di pochi secondi stamparla dalla macchina fotografica e, quindi, produrre la fotografia istantanea.

costanza del colore. Land aveva chiamato “**Mondrian**”<sup>7</sup> questi stimoli che aveva creato, in cui ci sono una grande varietà di cartoncini colorati.. Gli esperimenti con i Mondrian hanno lo scopo di mostrare che il colore che noi percepiamo è **indipendente da segnale cromatico locale** (quello che si misura in un punto della scena); mentre può essere predetto utilizzando un algoritmo che tenga conto della relazione tra **tutti e tre i segnali cromatici in tre canali indipendenti che corrispondono alle tre classi di cono**.

Questo in figura è una specie di *esperimento di uguagliamento*, perchè si possono confrontare le superfici di un Mondrian con la corrispondente superficie dell'altro Mondrian. Grazie al *fotometro*, si poteva misurare il segnale cromatico, la composizione spettrale della luce che veniva riflessa da una particolare superficie e poi, con un po' di calcoli, trovare il segnale cromatico corrispondente a quella superficie.

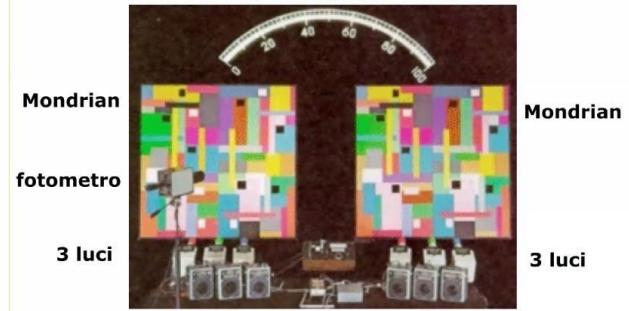
Per ognuno dei due Mondrian, ci sono degli illuminanti che sono fatti di tre luci, con un filtro rosso, verde e blu; combinando queste tre luci si può creare un illuminante, la cui composizione è anche nota.

Quindi, si può fare un **esperimento di costanza del colore** in cui:

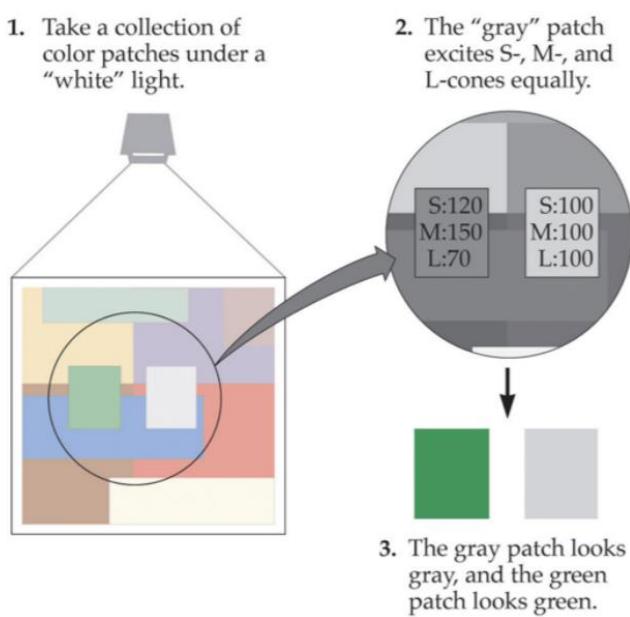
- una superficie a sinistra è illuminata in una certa maniera;
- la stessa a destra è illuminata con un illuminante che ha una composizione spettrale differente.

Si può andare a vedere se queste superfici appaiono diverse (o uguali) e fare esperimenti che vanno a sondare in quali condizioni si verifica la costanza del colore superficiale, oppure quando non si verifica.

## esperimenti con i Mondrian



<sup>7</sup> Il nome “Mondrian” è un chiaro riferimento al pittore olandese Pieter Mondrian, un astrattista che amava dipingere delle superfici di forme geometriche molto regolari (rettangoli, quadrati, etc.) con un colore omogeneo.



SENSATION AND PERCEPTION, Figure 5.18 (Part 1) © 2006 Sinauer Associates, Inc.

media e lunga dello spettro). Se prendiamo il cartoncino grigio, esso produce la stessa quantità di eccitazione nei coni S, M e L (altrimenti non sarebbe acromatico e, quindi, grigio). Si può immaginare che i coni S, M e L abbiano delle oscillazioni pari a 100 unità arbitrarie. Sotto la luce bianca, il cartoncino grigio appare grigio; il cartoncino verde non eccita tutti i coni in maniera uguale, ma stimola più i coni M e, proporzionalmente di meno, i coni L ed S (quello appare verde). Abbiamo una situazione in cui il segnale cromatico per il grigio è di un certo tipo, per il verde è di un altro tipo (il grigio appare grigio e il verde appare verde).

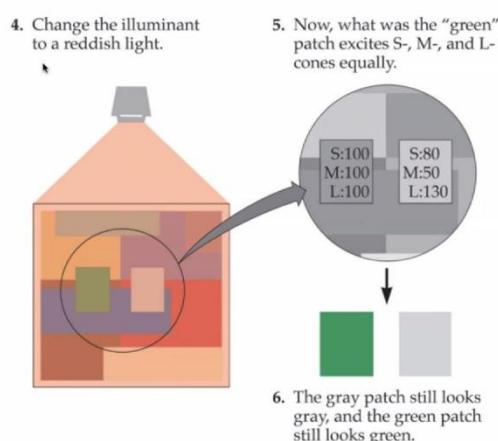
Si cambia l'illuminante aggiungendo potenza nella parte lunga dello spettro (si usa una **luce un po' rossastra**). Se aggiustata in maniera opportuna, possiamo riuscire a compensare il fatto che la superficie verde tende a riflettere nella parte centrale dello spettro, e trovare una situazione in cui il segnale cromatico che proviene da quella superficie che appariva verde (sinistra), adesso produce una uguale eccitazione da parte dei 3 coni (100, 100 e 100).

Cambiando l'illuminante ci si aspettava che:

- il segnale cromatico del cartoncino di sinistra fosse lo stesso che, invece, prima sotto la luce bianca era quello di destra.;

Land in questi esperimenti ha osservato che il segnale cromatico varia in funzione dell'illuminazione, però il colore non si può predire in base al segnale cromatico.

Basti immaginare che nel Mondrian ci sia una *“collection of colour patches”* (un insieme di cartoncini colorati) e di metterla sotto una **luce “bianca”** -luce che ha un certo tipo di composizione spettrale, più o meno equivalente la quantità di luce nei tre illuminanti (più o meno la stessa intensità della parte corta,



SENSATION AND PERCEPTION, Figure 5.18 (Part 2) © 2006 Sinauer

- che, invece, il cartoncino di destra adesso risulti diverso: data la luce rossastra, il segnale cromatico tende a sovra-rappresentare la parte lunga dello spettro e di meno l'altra, in quanto l'illuminante ha meno energia in quella zona.

Abbiamo creato una situazione in cui, dal punto di vista del segnale cromatico, adesso il cartoncino di sinistra dovrebbe apparire grigio, mentre quello di destra avrebbe un qualche tipo di cromaticità, ma questo non succede perché si verifica una **costanza del colore superficiale**, cioè Land osservò che il cartoncino di sinistra continuava ad apparire verde, mentre quello di destra continuava ad apparire grigio.

Secondo Land, questo può essere predetto con un modello matematico, che lui chiama il **modello Retinex**. Non vengono utilizzati i segnali cromatici in quanto tali, ma vengono utilizzati i “*segnali differenziali*”: confronto tra segnali cromatici centrali e quelli contestuali. Questa analisi viene fatta globalmente su tutta la scena.

## **Modello Retinex**

Il modello Retinex è il padre di tutti i modelli computazionali della costanza del colore.

Land era consapevole che questo processo non poteva essere un processo solo retinico che, bensì, coinvolge interazioni a livello cerebrale.

$$\text{retina+cortex} = \text{retinex}$$

L'idea fondamentale è che ci siano tre canali che acquisiscono un segnale, non importa che siano coni o sensori; questi canali si chiamano **S, M e L**, ossia i tre “**Retinex channels**”, in ognuno dei quali viene calcolata una **mappa di lightness** basata sui *rapporti di intensità* in quel canale (lightness più alta=bianco). Ognuno di questi Retinex vede un'immagine monocromatica, a toni di grigio. Il colore corrisponde alla tripletta per ogni superficie. **All'interno di quel canale, ci sarà una certa intensità che viene vista dai coni S nella posizione X e nella posizione X+1; quindi viene fatto il rapporto: viene calcolato il segnale differenziale fra questi due.** Basandosi su questi rapporti (*principio di Wallach*) viene costruita una mappa delle diverse bianchezze, le quali non corrispondono ad un colore percepito ma sono il segnale differenziale all'interno di quel canale di Retinex. È come se ci fossero tre immagini monocromatiche che corrispondono all'output di questi tre canali (S, M ed L), dopodiché queste tre immagini vengono messe insieme e il prodotto è il colore.

Un altro aspetto interessante **dell'algoritmo di Retinex è che incorpora anche l'idea che il rapporto più intenso** (ossia, la lightness più alta nella particolare “mappa di lightness” di quel Retinex) diventerà bianco e, cioè, viene utilizzato per scalare tutti gli altri.

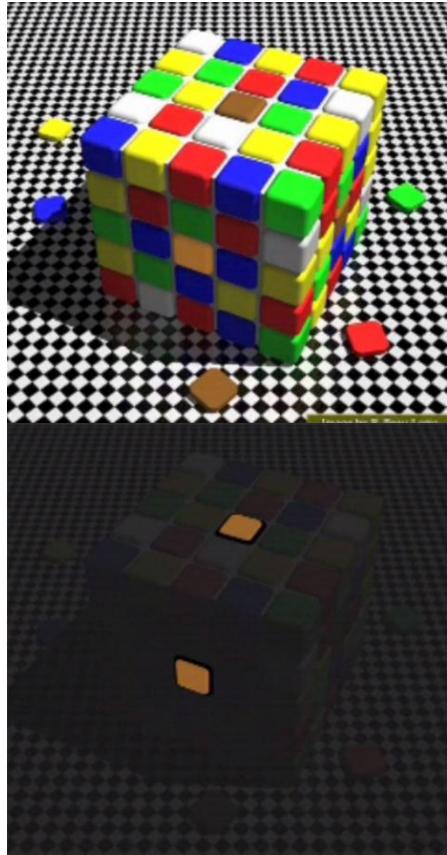
Land propone che ***questa particolare tripletta*** è quella che non varia al variare dell'illuminazione; con l'assunzione che l'illuminazione sul Retinex sia omogenea -e che tutto il Retinex sia illuminato alla stessa maniera, come negli esperimenti controllati con i Mondrian- molto del lavoro successivo ha avuto lo scopo di generalizzare queste idee a situazioni più realistiche; ciò vuol dire:

- da un lato, **avere una distribuzione spaziale dell'illuminazione che non sia costante in tutta la scena;**
- dall'altro, **permettere che la scena sia tridimensionale.**

Infatti, il modello Retinex originale funziona bene con i Mondrian in 2D e con un'unica illuminazione. **Nel mondo reale è più complesso**, però ci sono molti punti di contatto con quello che era originariamente il principio del rapporto nelle ricerche di Wallach, di qualche decennio prima.

**Di fronte ad una scena tridimensionale siamo in grado di tenere conto delle zone illuminate, con un lavoro di *integrazione* a lungo raggio fra le diverse relazioni fra superfici**, dal punto di vista cromatico.

#### Esempio - Cubo con dei Mondrian



Osservare la faccia che punta verso l'alto e la faccia che punta verso di noi e concentrarsi sul colore del quadratino centrale della faccia che punta verso l'alto; sembra marroncino mentre, invece, quello che sta nella faccia che punta verso di noi (che è in ombra) sembra di colore giallo/arancione. **Con il fotometro si misura il segnale cromatico di entrambi i quadratini e si scopre che è identico.**

Un metodo più semplice ed economico rispetto al fotometro potrebbe permetterci di sinceraci di quanto appena detto. Con lo **schermo di riduzione** si elimina il contesto (o meglio, si pone un contesto nero per i due quadratini); ancora una volta i segnali cromatici sono identici. Tuttavia, i colori appaiono diversi perché quei segnali cromatici vengono utilizzati in *relazione* a tutto il resto e tenendo conto del fatto che l'illuminazione è differente sulla superficie che punta verso l'alto e sulla superficie che punta verso di noi. Dato che sono illuminati diversamente, i segnali cromatici devono corrispondere a superfici diverse (questo è quello che si vede).

## Base neurale della percezione cromatica

Il sistema visivo umano/primate è in grado di fare una cosa abbastanza simile a quello che è descritto nel modello Retinex.

**Karl Gegenfurtner** ha scritto, nel 2003, una review su Nature nel settore di Neuroscienze “*Cortical mechanisms of colour vision*” in cui fa il punto su quello che noi sappiamo attualmente in merito ai **meccanismi corticali della visione cromatica**, con un particolare riferimento al “*problema della costanza del colore*”. Egli dice:

*“La percezione del colore è una componente centrale della visione nei primati. Il colore facilita la percezione degli oggetti e il riconoscimento degli oggetti e ha un ruolo importante nella nostra capacità di segmentare le scene e anche la memoria visiva. Inoltre, il colore dona all’esperienza visiva una componente estetica che è fondamentale per la percezione del mondo.”*

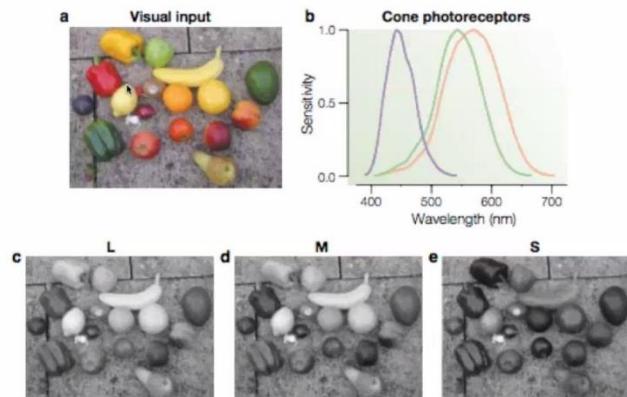
A dispetto di una lunga storia di ricerche e studi sulla percezione e visione del colore, ci sono ancora tante cose che dobbiamo imparare: **la base fisiologica della percezione cromatica**. Alcuni avanzamenti recenti nella nostra maniera di capire il funzionamento della retina, del **talamo** e del **genicolato**, ci hanno consentito di avere una visione nuova della componente estetica del colore. Questi studi stanno cominciando a indicarci **che il colore non è qualcosa che viene elaborato da solo, ma viene combinato con informazioni sia sulla luminanza** (il colore e l’intensità non sono due cose separate) **e anche con informazioni sulla forma degli oggetti** (tenere conto globalmente della struttura geometrica di uno stimolo). Quindi, il colore non è processato da solo, ma insieme a informazioni sulla luminanza e sulla forma visiva **da parte degli stessi circuiti neurali** -non solo circuiti per i colori ma sono anche circuiti in grado di dare informazioni sulla forma, sulla struttura tridimensionale, etc.- **allo scopo di ottenere una rappresentazione del mondo unitaria e robusta**<sup>8</sup>.

Gegenfurtner presenta, così, la struttura del sistema visivo per quanto riguarda la percezione cromatica: **l’input è dato dal prodotto fra delle proprietà di superfici (la loro riflettanza) e l’illuminazione.**

Esempio: ci sono delle superfici illuminate dalla luce solare; verso l’occhio rimbalza una certa luce (con una certa composizione spettrale ed intensità) la quale viene pesata con la sensibilità dei coni producendo il segnale cromatico per ogni punto di questa superficie: questo segnale cromatico può essere pensato come tre immagini acromatiche (come nel modello Retinex).

<sup>8</sup> “Robusta”, dal punto di vista computazionale, significa poco influenzato da casi particolari e, quindi, relativamente insensibile a cambiamenti, variazioni, rumore: costanza.

Queste tre immagini in toni di grigio rappresentano quella scena filtrata da una diversa sensibilità spettrale, in base al **principio di univarianza**; i singoli coni sono ciechi ai colori in quanto tali, vedono un'immagine in bianco e nero: è una maniera di rappresentarsi il tipo di informazione che viene codificata al livello dei tre segnali dei coni.



1. Questo è il primo stadio del segnale cromatico. Abbiamo, quindi, un **sistema tricromatico** (una tripletta di valori) **per ogni punto dell'immagine** e questi possono essere visualizzati come tre immagini monocromatiche. Ad esempio, la stessa superficie (l'avocado o il peperone giallo), nella *lightness map* ha un aspetto molto diverso a seconda che venga visto dai coni S, M o L e questo perchè questi tre coni hanno una sensibilità spettrale differente; la sua superficie ha dei valori differenti nella tripletta per le dimensioni S, M, L.
2. C'è poi uno stadio successivo, che è implementato da un circuito che coinvolge le *cellule gangliari* e le *cellule del genicolato* ed è responsabile di una **trasformazione di questa tripletta di valori** che spesso si dice sia un po' l'analogo della computazione Retinex, per via dell'utilizzo dei rapporti: vengono confrontate le triplettie di valori, **le quali non sono più all'interno di una classe di cono, ma fra classi di coni**.

In sostanza, lo “spazio tridimensionale S-M-L” e delle *relative attivazioni dei coni* viene trasformato in uno spazio in cui i segnali sono dei **segnali opposenti**. Il termine “opponenti” vuol dire che rappresentano ogni superficie come una posizione in degli assi teorici dimensionali, dove ci sono 2 colori primari che fungono da polo; c’è un canale opponente rosso-verde; uno blu-giallo e un colore opponente acromatico nero-bianco. Questo rende molto più diversi fra di loro i segnali cromatici, separandoli da quelli legati invece all’intensità (che è la luminanza).

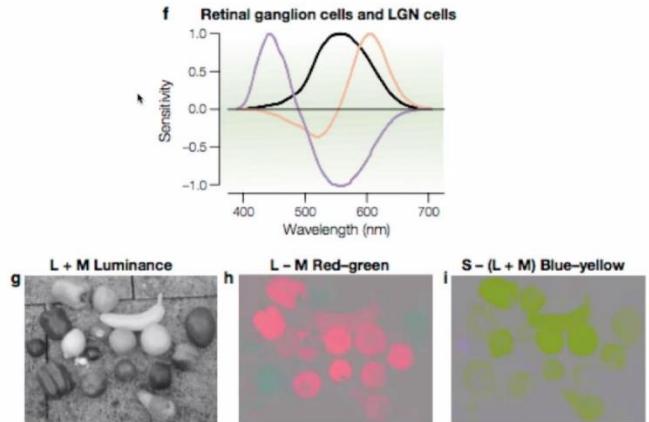
Se guardiamo i segnali S, M e L è evidente che le “*lightness maps*” che vengono prodotte dai coni L e M sono simili e correlati fra di loro; quelli dei coni S sono molto diversi perché la sensibilità spettrale è molto diversa.

**NB:** questa trasformazione in termini di **opponenza** ha l'effetto di **aumentare la risoluzione del segnale**, perchè rende maggiormente diverse fra di loro le due **dimensioni cromatiche e le separa dalla dimensione acromatica**. Un'altra funzione che dovrebbe avere come scopo questo primo livello della computazione, è **produrre un segnale che poi può essere utile per ottenere la costanza cromatica**.

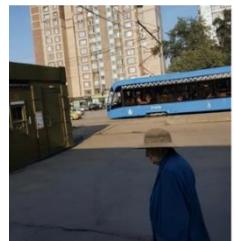
Non è ancora un segnale in grado di predire una dimostrazione -come quella del cubo con i Mondrian- perchè è un segnale che non sa nulla sulla struttura tridimensionale della scena.

Ci si pone il problema di dove vada a finire questo segnale in corteccia, quale sia il circuito corticale responsabile della produzione del colore in quanto tale e del colore “invariante” nonostante il variare dell’illuminazione. Su questo, nel 2003, Gegenfurtner dice:

“Noi sappiamo che i pazienti che hanno l'**acromatopsia** (pazienti neurologici che non vedono il colore) hanno **lesioni alla corteccia extra-striata**, in un’area che è omologa all’area V4 nella scimmia (si dice che il centro del colore sia V4). Tuttavia, i dati di neuroimmagine mostrano che ci sono risposte molto forti al colore anche nell’area striata V1 e V2 e, d’altro canto, è generalmente accettato che l’**area V4 sia anche un centro coinvolto nella percezione spaziale; è un centro che fornisce integrazione, cioè collegamento fra la visione, l’attenzione e la cognizione**. Pertanto, la nostra esperienza del colore probabilmente dipende dall’attività di molte aree corticali e, quindi, è necessario porre una maggiore enfasi alle computazioni, ossia sul come dal punto di vista algoritmico si possa ottenere la costanza del colore superficiale, piuttosto che focalizzarsi così tanto sulla localizzazione di regioni che sono importanti o meno per l’analisi del colore”.



Illusione: bande di Mach, due illuminazioni diverse, colore percepito non tanto in tinta ma nel senso del modo di apparire del colore, sembra trasparente sotto



## Lezione 23

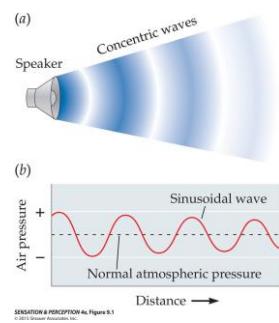
### Elementi di Multisensorialità 1: udito

In questa lezione verrà trattato il tema della **localizzazione degli oggetti nello spazio attraverso interazioni tra l'udito e la vista** ovvero dei processi audio-visivi in cui, così come anche nelle *interazioni visuo-tattili*, abbiamo degli aspetti legati alla localizzazione degli oggetti: nel caso delle audio-visive parliamo soprattutto di oggetti esterni quindi nello *spazio extra-personale*, mentre nel caso delle interazioni visuo-tattili è coinvolto lo *spazio peri-personale* in cui rientra anche il nostro corpo. Infatti siamo sempre consapevoli di dove sono localizzati gli arti e le diverse parti del nostro corpo nello spazio e questo ha un ruolo importante anche nella percezione del Sé. (Questo tema verrà trattato nella lezione successiva) mentre oggi si parlerà di udito quindi di come è fatto il sistema uditivo infatti questo è fondamentale per capire in che modo l'udito può darci informazioni sulla localizzazione degli oggetti e quindi poi chiederci cosa succede quando udito e vista possono collaborare.

Iniziamo col definire **cos'è un suono**.

Un suono è **il risultato percettivo della nostra capacità di processare energia meccanica ovvero informazioni che sono veicolate da onde sonore che, dal punto di vista fisico, sono delle oscillazioni| modificazioni cicliche di pressione che si propagano in mezzi elastici, ad esempio le molecole dell'aria**. Quindi un evento meccanico può mettere in vibrazione l'aria e questa vibrazione si propaga nello spazio. Naturalmente l'aria non è l'unico mezzo elastico infatti i suoni si sentono molto bene anche nell'acqua in cui vale lo stesso principio per cui le molecole di acqua vengono messe in vibrazione e propagano quest'onda. Inoltre, se pensiamo che Beethoven, pur essendo sordo, riusciva a comporre utilizzando un bastoncino di legno che posizionava contro il pianoforte e vicino all'orecchio e l'onda, propagandosi attraverso il legno che è un mezzo elastico, gli consentiva di farsi un'idea del suono che il pianoforte produceva.

Il suono ha una velocità di propagazione relativamente bassa rispetto a quella con cui si propaga la luce e da un punto di vista pratico i raggi luminosi sono istantanei, viaggiano talmente veloci che sulla scala del nostro ambiente è praticamente nullo il tempo che impiegano per arrivare da una superficie al nostro occhio. Nel caso del suono la cosa è un po' diversa perché viaggia velocemente, ma la sua velocità non è minimamente paragonabile a quella della luce.



La velocità della luce è circa  $300.000 \frac{m}{s}$ , mentre la velocità di propagazione del suono nell'aria ad una temperatura di  $20^{\circ}\text{C}$  è pari a  $344 \frac{m}{s}$ , la velocità del suono infatti dipende dal *mezzo* in cui il suono si propaga e anche da alcune *caratteristiche e proprietà del mezzo stesso*, infatti, ad esempio, la velocità del suono nell'aria a  $0^{\circ}\text{C}$  è pari a  $331,2 \frac{m}{s}$ .

In approssimazione lineare la velocità del suono varia in funzione della temperatura secondo la legge:

$$\alpha(T) = \sqrt{401,8 \cdot T} \frac{m}{s},$$

a livello fisico questa variazione dipende dalla **distribuzione molecolare** dei mezzi che incontra, infatti in un *solido*, in cui tutte le molecole sono collegate solidamente fra loro la velocità del suono è più alta rispetto alla velocità di propagazione del suono in un *gas* o in un *liquido* dove le molecole sono disposte incoerentemente. Proprio per questo motivo la velocità del suono è circa 5 volte più alta in acqua ( $1500 \frac{m}{s}$ ) rispetto che nell'aria, essendo l'acqua è meno comprimibile dell'aria la vibrazione viene trasmessa più rapidamente. Il nostro orecchio non è in grado di recepire il suono allo stesso sia in acqua che in aria, perché la conformazione del nostro padiglione auricolare è adattiva per la vita terrena e non subacquea: la vibrazione, viaggiando più velocemente, quando incontra il padiglione auricolare non viene trasmessa all'orecchio medio e interno nello stesso modo in cui viene fatto quando quest'ultima viaggia nell'aria e quindi il processo di trasduzione non può essere messo in atto.

**PER APPROFONDIMENTI:**

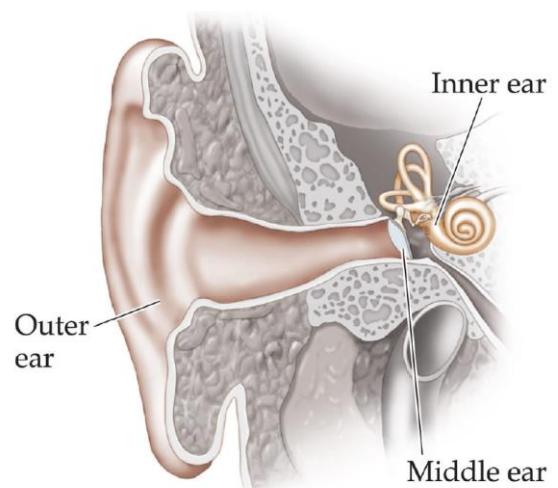
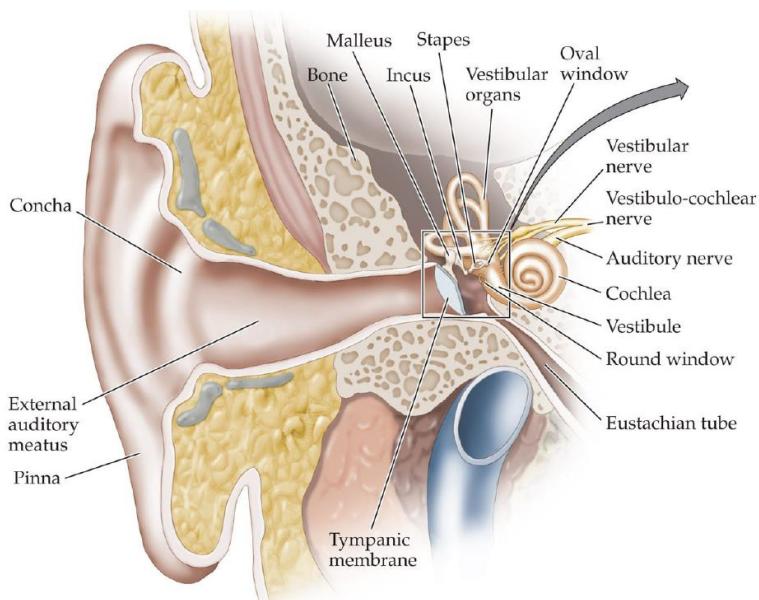
[https://it.wikipedia.org/wiki/Velocità\\_del\\_suono](https://it.wikipedia.org/wiki/Velocità_del_suono)



<https://www.chimicifisici.it/con-lorecchio-sottacqua-la-fisica-del-suono/#:~:text=il%20suono%20si%20propaga%20in,acqua%20%C3%A8%201500%20m%2Fs>

## Anatomia del sistema uditivo

L'organo che ci consente di effettuare la percezione di un suono è l'orecchio che può essere diviso in 3 parti: **esterna**, **media** ed **interna**. L'*orecchio esterno* è composto dalla pinna ovvero il **padiglione auricolare** che ha lo scopo di agire come un *microfono direzionale* in quanto aiuta a convogliare le onde vibratorie verso l'*orecchio medio*. In questa zona le vibrazioni raggiungono la membrana del **timpano** che a sua volta vibra in relazione alla vibrazione che c'è nell'aria/viene trasferisce dall'aria alla **membrana timpanica**.

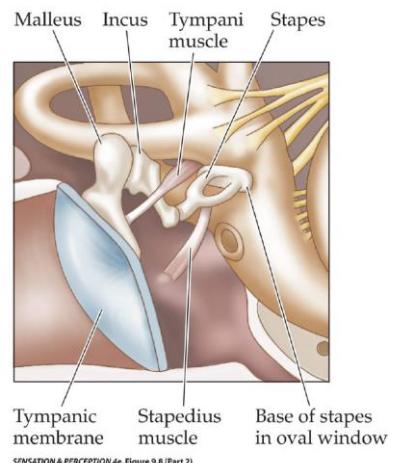


Quest'ultima, come un tamburo, vibra e trasmette a sua volta la vibrazione su un sistema di leve formato da tre ossicini ovvero **staffa**, **incudine** e **martello** chiamati in questo modo per la loro forma. A questo punto la vibrazione si propaga, sempre meccanicamente, grazie al sistema di leve fino ad un'altra membrana chiamata **finestra ovale** la quale a sua volta trasmette la vibrazione all'interno di una struttura chiamata **coclea**. Tale termine vuol dire “*chiocciola*” → la coclea

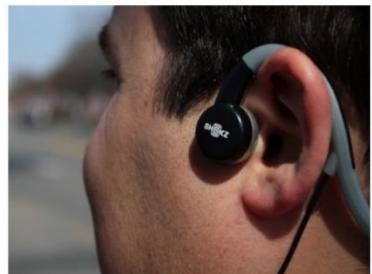
ha una forma di tubo arrotolato che appunto ricorda la forma di una chiocciola. All'interno della coclea a sua volta è contenuto un liquido che mette in vibrazione delle piccolissime protuberanze che si chiamano **stereociglia** e che vibrano conseguentemente alla vibrazione del liquido. Questa è una descrizione sintetica del **processo di audio-trasduzione** che descrive come l'energia meccanica contenuta nell'onda sonora viene trasformata in un segnale nervoso.

In immagine è rappresentato un ingrandimento del sistema degli ossicini all'interno della coclea e possiamo notare la *membrana timpanica*, il *martello*, la *staffa*, l'*incudine* e la *finestra ovale*. Inoltre al di sotto della finestra ovale abbiamo un'altra struttura rotonda, la **finestra rotonda** la cui funzione è quella di disperdere l'energia meccanica una volta che questa ha completato il suo percorso all'interno della coclea.

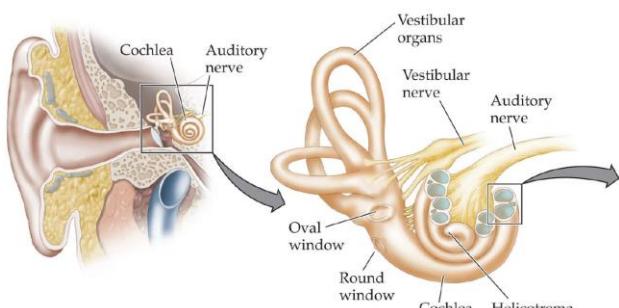
Notiamo che si tratta proprio di energia meccanica tanto è vero che è possibile udire anche in assenza di questo processo che coinvolge orecchio esterno, medio ed interno in quanto è possibile mettere la coclea direttamente in vibrazione attraverso energia meccanica che viene somministrata per via ossea (il sistema che usava Beethoven ponendosi il bastoncino di legno vicino l'orecchio) come avviene ad esempio per le *cuffie a conduzione ossea* che anziché produrre un segnale che mette in vibrazione l'aria dentro l'orecchio, mettono direttamente in vibrazione l'osso del cranio intorno all'orecchio consentendo così di sentire. Questa è una cosa molto interessante in quanto apre un canale sensoriale normalmente non disponibile una specie di "realtà aumentata" perché possiamo sentire simultaneamente quello che c'è nell'ambiente e, attraverso un canale separato, ascoltare ad esempio musica.



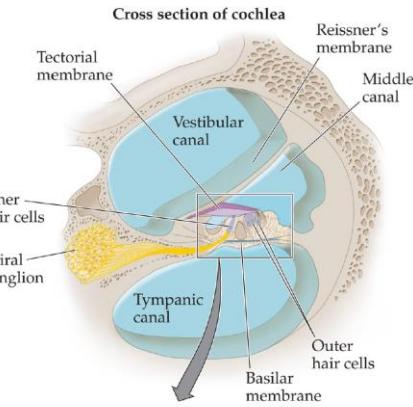
**bone conduction headphones**



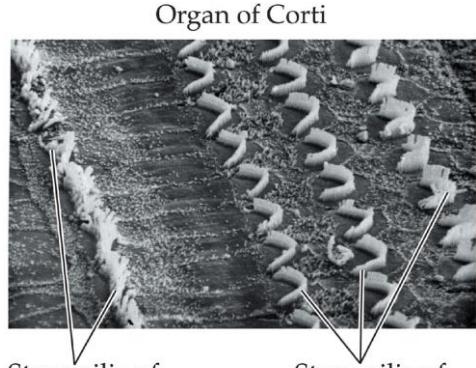
Inoltre come dicevamo, la trasduzione del segnale acustico avviene all'interno della coclea che poi forma il **nervo uditivo** e manda il segnale alle stazioni successive.



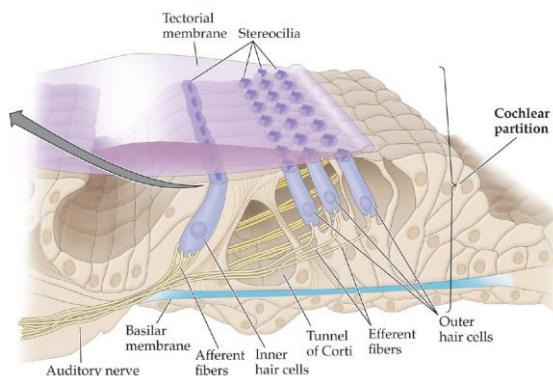
SENSATION & PERCEPTION 4e, Figure 9.9 (Part 1)  
© 2015 Sinauer Associates, Inc.



SENSATION & PERCEPTION 4e, Figure 9.9 (Part 2)  
© 2015 Sinauer Associates, Inc.



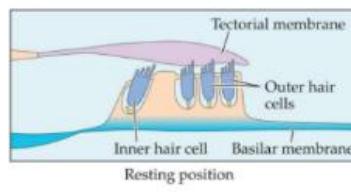
SENSATION & PERCEPTION 4e, Figure 9.9 (Part 4)  
© 2015 Sinauer Associates, Inc.



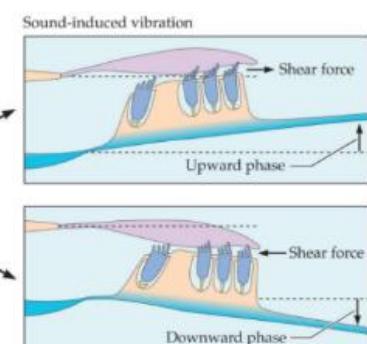
SENSATION & PERCEPTION 4e, Figure 9.9 (Part 3)  
© 2015 Sinauer Associates, Inc.

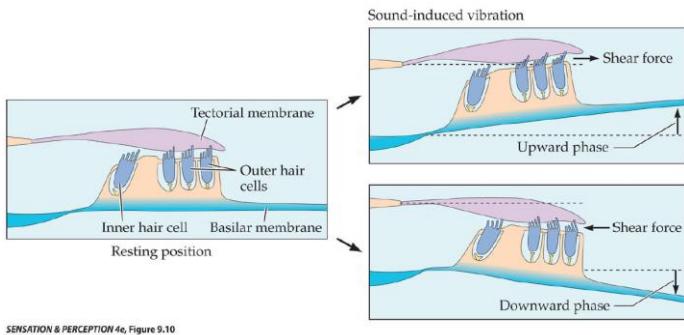
Questo meccanismo è consentito dalla presenza di canali nella coclea in cui il liquido può vibrare. In immagine vi è una rappresentazione in sezione della coclea e possiamo notare il **canale timpanico**, il canale **medio** ed il canale **vestibolare** ed al centro abbiamo la **membrana tectoria** che viene fatta vibrare e vibrando trasmette la vibrazione alle *stereociglia* dei recettori uditivi.

Si tratta dunque di un *sofisticato sistema di leve e membrane* che di fatto non soltanto trasmettono la vibrazione dall'area fino alle stereociglia, ma anche grazie al loro grado di maggiore o minore rigidità e all'azione delle leve producono una sorta di amplificazione del segnale che ha lo scopo di favorire la sua trasmissione fino al punto d'arrivo.



SENSATION & PERCEPTION 4e, Figure 9.10  
© 2015 Sinauer Associates, Inc.

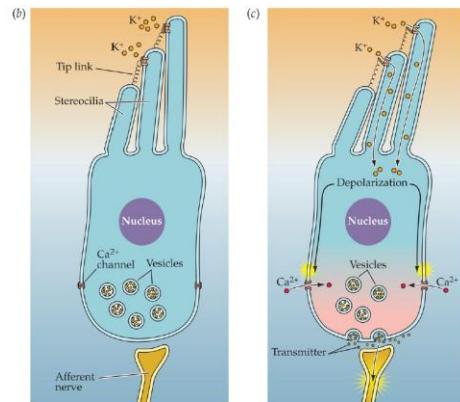




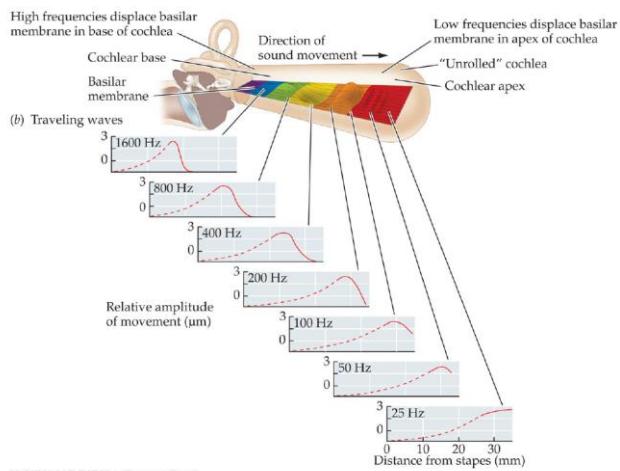
SENSATION & PERCEPTION 4e, Figure 9.10  
© 2015 Sinauer Associates, Inc.

Il movimento delle ciliege induce la variazione del potenziale graduato necessario ad iniziare il processo di traduzione che in questa sede non approfondiamo.

In immagine è possibile vedere come la vibrazione della membrana tectoria causi uno spostamento delle stereocilia a destra o a sinistra.



SENSATION & PERCEPTION 4e, Figure 9.11 (Part 2)  
© 2015 Sinauer Associates, Inc.



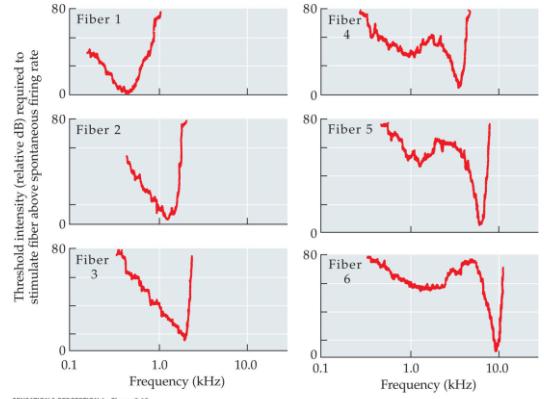
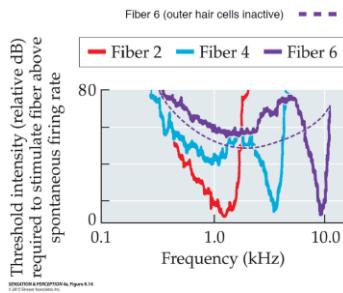
SENSATION & PERCEPTION 4e, Figure 9.12 (Part 2)  
© 2015 Sinauer Associates, Inc.

Alla fine) rispondono in maniera selettiva a bande di frequenza diverse (**codifica tonotopica**): alla **base** della coclea c'è una migliore selettività per le **frequenze più alte** mentre l'**apice** della coclea risponde meglio a **frequenze basse**. Tale selettività è particolarmente chiara con suono deboli in quanto un suono molto intenso contiene molta energia e tende a far vibrare tutta la coclea in modo identico, questo fa presupporre che ci siano anche altri meccanismi di percezione della frequenza.

Quindi andando dalla base all'apice della coclea, ovvero da porzioni più rigide a porzioni meno rigide, troviamo una selettività per frequenze via via più basse e queste sono delle vere e proprie *curve di sintonizzazione* alla frequenza in funzione della posizione sulla coclea ovvero la distanza dalla finestra ovale, a livello fisico l'apice della coclea è anche quella più larga ed è quella che vibra per le frequenze più basse.

**Codifica tonotopica:** parti diverse della coclea rispondono in maniera selettiva a bande di frequenza diverse. La base risponde meglio alle alte frequenze, mentre l'apice alle basse frequenze, ricordiamo inoltre la **seleattività è più chiara con suoni deboli**, questo perché un suono di intensità molto forte tende a far vibrare tutta la membrana della coclea.

Come notiamo dal grafico si tratta di una sintonizzazione piuttosto ampia, ad esempio per un suono di 400 Hz viene messa in vibrazione soprattutto una zona centrale della coclea però vibra comunque quasi tutta e quindi questo significa che basta aumentare un po' l'intensità del suono per far sì che la vibrazione riguardi l'intera coclea e questo meccanismo di sintonizzazione non sarà più così buono.



SENSATION & PERCEPTION 4e, Figure 9.13

Ad ogni modo tale meccanismo esiste ed è visibile anche da un punto di vista fisiologico: se registriamo da fibre che provengono da parti diverse della coclea troviamo che la soglia per produrre una risposta neurale si sposta lungo il continuum della frequenza in accordo con l'idea che ci sia una sintonizzazione per quella parte della coclea riguardo a frequenze più o meno alte.

Passiamo ora ad un secondo meccanismo.

La codifica tonotopica sicuramente esiste ed è stato il primo meccanismo ipotizzato per la codifica delle frequenze a livello sonoro. Già **Helmholtz** aveva immaginato che la coclea fosse come la tastiera di un pianoforte in cui in base al punto in cui si produce un evento percussivo, ossia viene messa in vibrazione una corda, sono chiamate in causa frequenze diverse. **Tuttavia**, come abbiamo visto, **questa codifica è abbastanza imprecisa perché funziona meglio con suoni non troppo intensi e non funziona con suoni a frequenza molto alta**. Difatti, se riguardiamo il grafico di sensibilità alla varie frequenze in base alla porzione della coclea notiamo che la massima frequenza è pari a 1600 Hz quando in realtà sappiamo che la sensibilità sonora si estende quasi fino a 10.000 Hz.

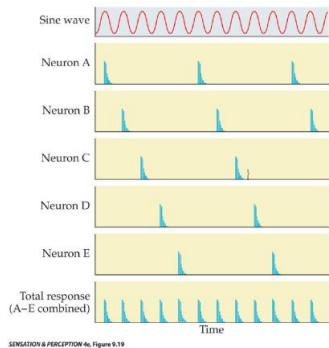
## Come si fa quindi a percepire questi suoni molto intensi?

E' stato chiarito che oltre al meccanismo di codifica tonotopica esiste anche un **meccanismo**, basato sul tempo più che sulla posizione, **di codifica temporale**

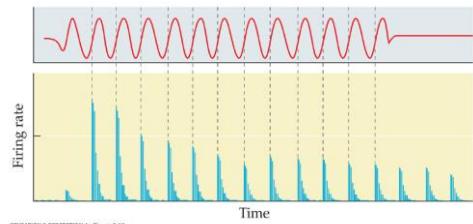
che si fonda sull'idea di fare una codifica non su una singola cellula che risponde quando c'è una certa energia in una certa frequenza, ma una codifica di popolazione ovvero basata sull'attività simultanea di più cellule. Ci si basa dunque sull'idea che parti diverse della coclea si sintonizzano su frequenze differenti grazie alla capacità di codificare un profilo temporale di un suono.

Questo avviene grazie a due processi chiamati:

**Aggancio di fase (*phase locking*):** un singolo neurone scarica sempre nello stesso punto del periodo dell'onda;



**CODIFICA TEMPORALE:** parti diverse della coclea si sintonizzano su determinate frequenze grazie a una codifica del profilo temporale del suono



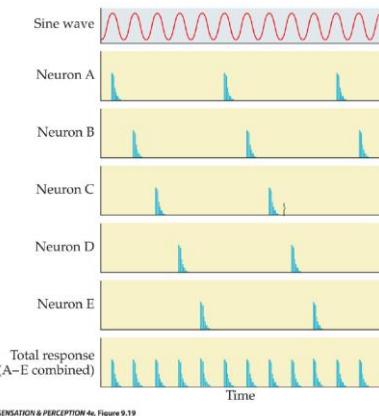
**Principio della scarica (*volley*):** una popolazione di neuroni effettua una codifica temporale "distribuendo" l'aggancio di fase.

1. Se abbiamo un **segnale ciclico, periodico**, questo oltre ad avere una frequenza ha anche una **fase** che indica il punto nel tempo in cui inizia questo segnale. Ad esempio, date 2 melodie se la seconda voce inizia un po' più sfasata nel tempo rispetto alla prima, questo sfasamento nel profilo melodico in questi cambiamenti di frequenza (la melodia) costituisce la differenza di fase. Dunque per **aggancio di fase** ci riferiamo all'idea che un neurone si aggancia alla fase dell'onda perché scarica sempre nello stesso punto dell'onda stessa. Ad esempio scarica quando l'onda ha il suo massimo.
2. Per quanto riguarda il concetto di **volley**, questo termine inglese ha molteplici significati. In questo caso ci si riferisce a quando vi è una sequenza di eventi coordinati che avvengono tutti insieme a cui segue un momento di pausa per poi riprendere con la sequenza. Possiamo trovare un esempio di volley se pensiamo a quando si coordinano i soldati a sparare. Quello che si faceva, dato che c'era bisogno di un po' di tempo per ricaricare il fucile, era disporre i soldati su delle file e poi si facevano dei volley successivi ovvero delle scariche successive di fuoco quindi la prima riga si

metteva in ginocchio e sparava poi la seconda riga in piedi e sparava mentre la prima ricaricava e così via. Questa è l'idea che queste scariche di gruppi di neuroni si distribuiscono tra di loro gli aggangi di fase e quindi l'insieme dei neuroni, dato che sono tutti agganciati alla stessa fase, ma non scaricano tutti nello stesso momento, producono una scarica che riesce a rappresentare la frequenza del suono in ingresso e riesce a farlo anche per suoni con frequenza molto più alta di quelli rappresentabili con la codifica tonotopica.

Dunque il **principio della scarica** rappresenta il fatto che una popolazione di neuroni effettua una codifica temporale "distribuendo" l'aggancio di fase.

Possiamo immaginare una cosa simile. Nel tempo l'onda, che ha un andamento ciclico, e ci sono dei neuroni che nel tempo scaricano agganciati sempre allo stesso punto dell'onda e questo crea un volley ovvero una codifica di popolazione perché ogni singolo neurone A, B, C, D, E scarica con una frequenza che è molto più bassa della frequenza dell'onda sinusoidale in ingresso. Quindi la frequenza di scarica del singolo neurone non rappresenta la frequenza vera dell'onda però combinando, con una **codifica di popolazione**, le frequenze di scarica di questi neuroni si va a ricostruire la frequenza del suono in ingresso.



#### APPROFONDIMENTO



In questa foto scaricata da internet possiamo notare innanzitutto delle Bande di Mach se guardiamo al centro dov'è il confine tra l'ombra e il marciapiede. La cosa più interessante è il fatto che l'allineamento tra il cappello e l'ombra produce un effetto di **unificazione-segregazione**: sembra che il cappello appartenga ad una struttura gialla che sta sopra (è presente anche un effetto di *somiglianza* e di *buona continuazione*) mentre il resto del corpo alla parte di

sotto; questo genera l'impressione di due diverse illuminazioni al punto che si fa fatica ad unificare il cappello con la persona che lo porta e quest'ultima, sembra quasi vista attraverso una superficie trasparente. L'immagine costituisce dunque un ottimo esempio di come questi processi legati alla **percezione della forma** e alla **struttura tridimensionale** di una scena interagiscano con il modo in cui viene percepito il colore, in questo caso non tanto il colore come tinta, nel senso che noi vediamo blu e giallo e probabilmente sono davvero questi i colori rappresentati, ma nel senso del modo di apparire del colore che in un caso sembra visto attraverso un mezzo trasparente ed in un caso invece no.

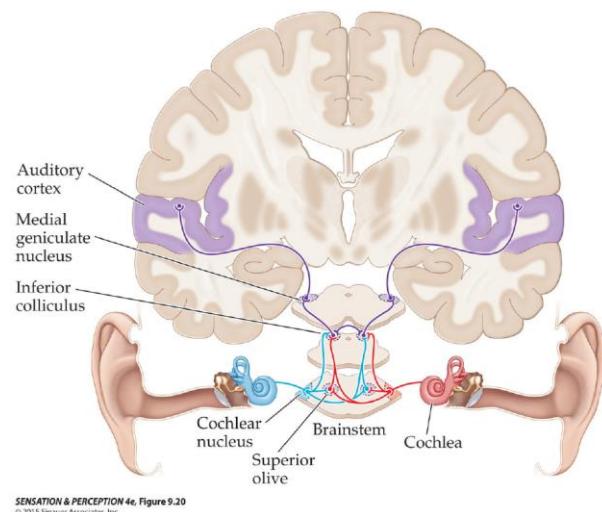
## La corteccia uditiva e la localizzazione delle sorgenti sonore

Adesso vediamo come possiamo parlare di **localizzazione uditiva e audio-visiva** da un punto di vista multisensoriale. Innanzitutto facciamo un breve cenno sulla localizzazione in base all'udito. Noi abbiamo parlato in precedenza di localizzazione visiva esaminando vari *meccanismi pittorici* o basati sulla *parallasse monoculare o binoculare* per la percezione dello spazio.

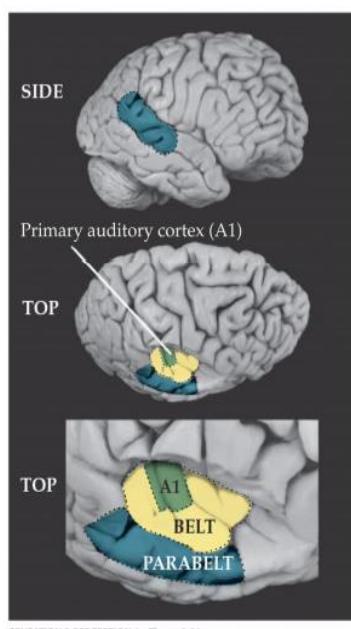
Nel caso dell'**udito** i meccanismi sono un po' diversi anche se anche qui troviamo dei meccanismi binaurali ovvero utilizzano il fatto che abbiamo due orecchie, ed altri monoaurali.

### Neurofisiologia dell'udito

Dopo la coclea, il **nervo acustico** invia il segnale al cervello e lo fa passando attraverso stazioni intermedie nel tronco encefalico ovvero il **nucleo olivare superiore** poi i **collicoli**, il **nucleo genicolato mediale** ed infine la **corteccia uditiva** situata nel giro superiore del lobo temporale.



Quest'ultima è suddivisibile in 3 parti:



SENSATION & PERCEPTION 4e, Figure 9.21  
© 2015 Sinauer Associates, Inc.

A1: primary auditory area; tonotopic

BELT: more complex sounds

PARABELT: more complex sounds and multisensory

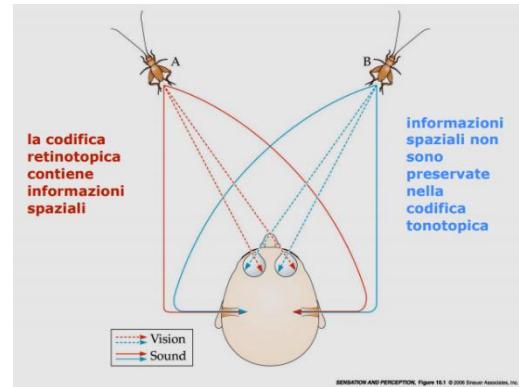
**A1 → area uditiva primaria** in cui vi sono cellule ancora legate alla codifica tonotopica. Abbiamo quindi un'analogia con la visione in quanto anche in V1 troviamo ancora la codifica retinotopica.

Abbiamo poi delle aree circostanti che vengono chiamate **cintura (BELT)** e **paracintura (PARABELT)** in cui troviamo aree sintonizzate a suoni più complessi quindi non singole frequenze e anche delle sensibilità dei campi recettivi che iniziano a diventare multisensoriali.

## Localizzazione del suono

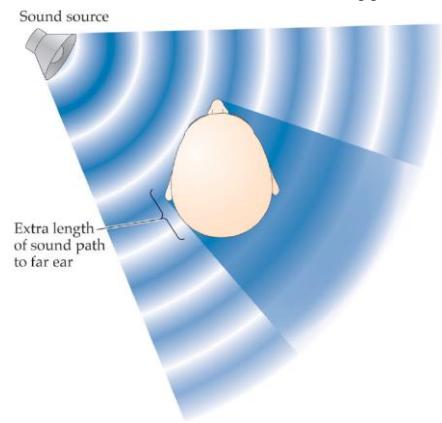
Noi siamo in grado di farci un'idea più o meno accurata della direzione da cui proviene un suono ovvero di localizzare una sorgente sonora. **Come riusciamo a farlo?**

Dal punto di vista visivo abbiamo una proiezione bidimensionale sulla retina di un mondo che è tridimensionale ma comunque la codifica retinotopica preserva degli aspetti della spazialità in questa proiezione: un punto sulla retina corrisponde ad una direzione nello spazio e anche se lungo questa direzione ci sono infiniti punti essi appartengono comunque a quella specifica direzione.



Di contro, dal punto di vista uditivo, abbiamo una codifica tonotopica in cui viene elaborata la frequenza di un suono e dunque non ci sono informazioni spaziali di per sé. Tuttavia il fatto di avere due orecchie ci consente di avere informazioni che sono correlabili allo spazio. Queste informazioni sono in prima battuta di natura binaurale e sfruttano il fatto che la velocità di propagazione del suono nell'aria sia lenta per cui ci sono delle differenze nel tempo di arrivo di un suono ad un orecchio rispetto all'altro in funzione di come è orientato il capo rispetto alla sorgente sonora stessa.

## ITD, interaural time difference



Ad esempio in immagine possiamo vedere una sorgente sonora posizionata a sinistra rispetto alla direzione del volto dell'ascoltatore ed il suono arriverà un po' prima all'orecchio sinistro rispetto al destro. Questa differenza di *tempo di arrivo (ITD, interaural time difference)* è piccola, ma apprezzabile dal punto di vista neurale grazie al fatto che il suono viaggia relativamente lento.

Dunque le differenze nel tempo di arrivo danno informazione sulla posizione di un luogo di punti chiamato **Azimuth**, ovvero il cerchio immaginario che possiamo tracciare intorno a noi sul piano orizzontale. Quindi ogni posizione su questo cerchio immaginario è descritta da un angolo e identifica una *direzione*, ovvero un segmento in cui ci sono tanti punti a distanze diverse, non una posizione specifica nello spazio tridimensionale.

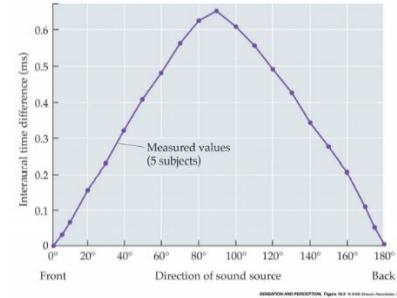
Se proviamo a pensare a questo cerchio intorno al nostro capo e al fatto che a seconda di come è posizionato il nostro capo rispetto alla sorgente sonora ci possono essere delle differenze nel tempo di arrivo del suono ad un orecchio rispetto ad un altro, ci possiamo chiedere:

- 1. In quale posizione deve essere una sorgente per produrre la massima differenza nel punto di arrivo (ITD)?** questa ci sarà quando la sorgente è direttamente davanti uno dei due orecchi quindi a destra o a sinistra, a  $90^\circ$  o  $270^\circ$  oppure ad ore 3 ed a ore 9 di un orologio.
- 2. In quale posizione deve essere una sorgente per produrre la minima o nulla differenza binaurale (ITD)?** La sorgente dovrà essere esattamente avanti o esattamente dietro quindi a  $0^\circ$  o  $180^\circ$
- 3. Cosa accade nelle posizioni intermedie?**

Ci sarà un segnale che sarà a metà strada e la differenza sarà proporzionale alla direzione. Possiamo quindi immaginare un grafico di questo tipo in cui viene plottata la **differenza interaurale**, (non è un calcolo geometrico, ma una *misura* vera e propria) in funzione della **direzione della sorgente**. Immaginiamo un segmento che si estende dal nostro naso a va verso l'infinito, in quella direzione  $0^\circ$  se perfettamente allineato alla simmetria del nostro viso il suono arriverà esattamente nello stesso momento alle due orecchie in quanto la distanza è la stessa. In questo caso la ITD, differenza di tempo interaurale sarà 0.

Se ci spostiamo verso destra quindi lungo un angolo da  $20^\circ$  a  $90^\circ$  la differenza aumenta fino a raggiungere il valore massimo a  $90^\circ$  e poi continuando a ruotare dietro il capo le differenze ricominciano a scendere tornando a 0 quando la direzione del suono è a  $180^\circ$  rispetto al capo.

Inoltre il grafico suggerisce che dal punto di vista della **differenza interaurale** noi non dovremmo essere in grado di distinguere tra una sorgente a  $20^\circ$  ed una a  $160^\circ$  perché queste dal punto di vista geometrico producono la stessa differenza interaurale, la differenza è la stessa solo che ci si sposta da una posizione anteriore ad una posteriore lungo un cerchio immaginario. Se proviamo a fare *questo esperimento* chiedendo ad un soggetto di chiudere gli occhi e poi gli presentiamo un suono derivante ad esempio da uno schiocco delle dita, prima **davanti** e poi **dietro**, e gli chiediamo dov'era localizzato il suono, vediamo che difficilmente sbaglia, riuscendo a localizzare il suono in modo corretto. Questa situazione comunque potrebbe suscitare più incertezza rispetto a quella in cui viene chiesto di localizzare un suono presentato a **sinistra** oppure a **destra**, sarebbe infatti una localizzazione molto saliente in quanto c'è una differenza molto grande nel tempo di arrivo interaurale ad un orecchio rispetto ad un altro se il suono è a destra o a sinistra rispetto alla localizzazione del volto.

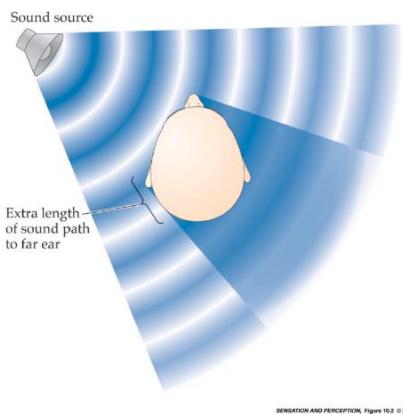


Questa idea è rappresentata anche in questa immagine in cui possiamo riconoscere l'**Azimuth** con i gradi e i tempi di arrivo interaurale.

Fra  $60^\circ$  ed il corrispondente punto a  $260^\circ$  (ovvero muovendosi nella direzione opposta di  $60^\circ$ ) la differenza è molto evidente: una a favore dell'orecchio destro e l'altra, nel senso contrario, a favore dell'orecchio sinistro.

Ma se ad esempio consideriamo  $60^\circ$  e  $120^\circ$ , che è il punto simmetrico dietro la testa, vediamo che **il tempo di arrivo interaurale** è sempre  $480\ \mu s$  (come nel caso precedente) e quindi non dovremmo essere in grado di distinguere, ma in realtà siamo in grado di farlo e infatti la localizzazione del suono non è dovuta esclusivamente al tempo di arrivo interaurale, ma anche ad altri processi di integrazione neurale.

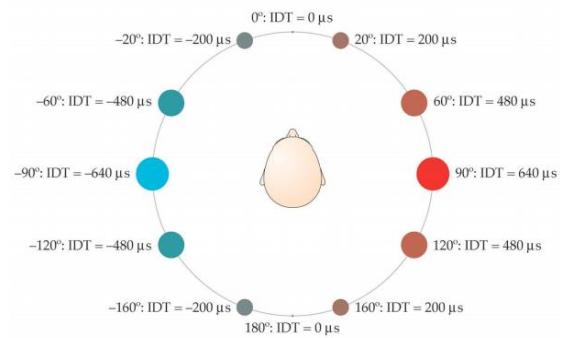
### **ILD, interaural level difference**



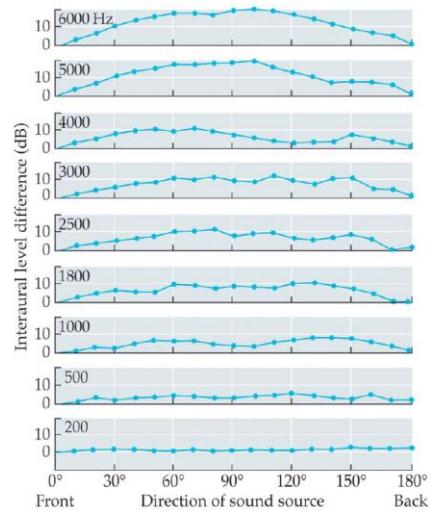
Ci sono anche altri meccanismi che consentono di localizzare la direzione da cui proviene una fonte sonora, uno di questi è quella che viene chiamata **ILD, interaural level difference**, laddove “level” si riferisce al **livello di intensità**, e rappresenta la differenza di ampiezza dell’onda. Questo meccanismo ha a che fare con il fatto che più strada deve fare il suono e più tende a *dissiparsi* ovvero *perdere di potenza* infatti, banalmente, più ci si allontana dalla fonte sonora e meno si sentono i suoni. Dunque anche questa

**differenza di intensità** può essere utilizzata per localizzare un suono: dato che deve fare più strada per arrivare ad un orecchio, tenderà ad essere leggermente meno intenso.

Possiamo osservare questo fenomeno in un grafico simile quello precedente in cui si vede che per suoni a diversa frequenza ci sono delle differenze nell’intensità del suono che sono abbastanza apprezzabili per suoni con frequenze relativamente alte e diventano sempre meno apprezzabili per suoni di frequenze più basse in quanto in un suono a bassa frequenza la lunghezza d’onda è maggiore e quindi questo dispendersi dell’energia su uno spazio molto breve diventa molto meno rilevante.



SENSATION AND PERCEPTION, Figure 15.4 © 2008 Sinauer Associates, Inc.



Ad ogni modo **entrambi i meccanismi danno informazione sulla posizione del suono rispetto all'Azimuth**. Nel caso della differenza di intensità (ILD) la precisione è meno alta in modo particolare per le basse frequenze e questo dipende dal rapporto fra la lunghezza d'onda e le dimensioni della testa (*il cono d'ombra sonora*).

- Anche le ILD danno informazione sulla posizione rispetto all'Azimuth;
- La precisione è tuttavia meno buona rispetto alle ITD e particolarmente scadente alle basse frequenze;
- Questo dipende dal rapporto fra la lunghezza d'onda e le dimensioni della testa (*cono d'ombra sonora*);
- L'asimmetria delle curve dipende da una terza fonte di informazione: la **forma del lobo auricolare**.

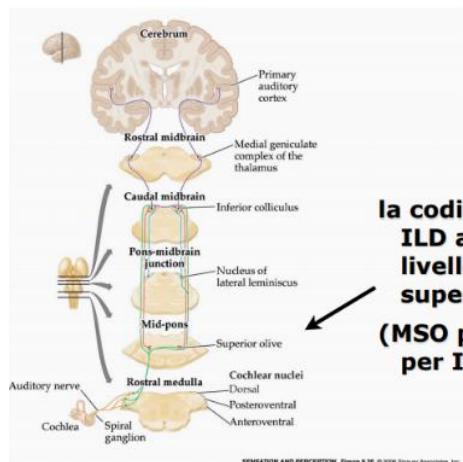


Inoltre, è importante anche notare che queste curve sono un po' asimmetriche e questo suggerisce che ci sia ancora una **terza fonte** di informazione per la localizzazione di natura monoaurale. Quest'ultima è legata alla **forma del lobo auricolare** che è estremamente variabile da persona a persona il che suggerisce che ci sono degli aspetti della percezione uditiva a carattere idiosincratico, leggermente diversi da persona a persona. Infatti *la forma della pinna* altera la composizione dell'onda meccanica in funzione del fatto che la pinna è orientata verso la parte anteriore e funziona un po' come un amplificatore direzionale: amplifica di più davanti rispetto alla parte posteriore. Tutto ciò

sicuramente è uno degli aspetti che ci aiutano a distinguere tra davanti o dietro rispetto all'Azimuth.

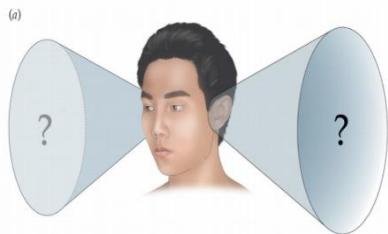
La codifica delle differenze interaurali avviene nel **nucleo olivare superiore**: nella *parte mediale dell'oliva superiore* sono codificate le differenze di tempo di arrivo (ITD) e in quella laterale le differenze di intensità (ILD).

Ad ogni modo tutti questi elementi ci danno un **angolo** e non una posizione nello spazio ovvero ci dicono da quale direzione proviene un suono e non quanto è distante da noi il suono stesso.



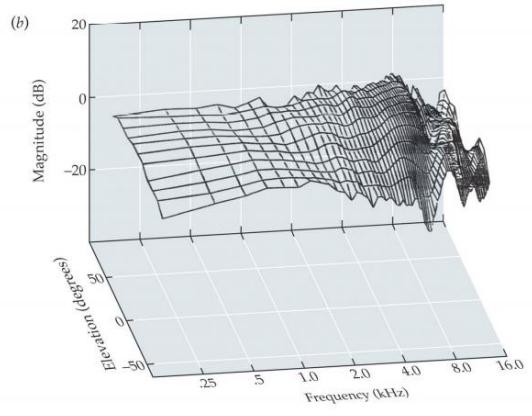
Infatti giudicare quanto è distante da noi un dato suono è molto più complesso anche perché è confuso con l'intensità originaria del suono: se un suono è più lontano è naturalmente più debole, ma è difficile giudicare se il suono diventa più debole perché pur essendo molto intenso è molto lontano, oppure perché è già debole di suo. Inoltre questi giudizi si riferiscono al **piano orizzontale**, sono localizzazioni della direzione rispetto all'Azimuth e non ci danno info rispetto all'elevazione di questa sorgente sonora. E' un po' come se estraessimo una componente di questa posizione su piano orizzontale, ma naturalmente un suono potrebbe essere a sinistra in alto o in basso rispetto al nostro naso ad esempio.

- ITD e ILD forniscono informazioni su una relazione spaziali: l'angolo formato dalle due sorgenti sull'Azimuth;
- Non forniscono informazioni sulla posizione assoluta dell'esorgenti, ma solo su quella relativa;
- Per ogni ITD e ILD esistono infinite coppie di posizioni distali corrispondenti (*cono di confusione*).



Detto ciò rimane il problema che vi sono *posizioni distali* che sono potenzialmente confondibili rispetto alla loro posizione davanti o dietro: per ogni ITD e ILD esistono infinite coppie di posizioni distali corrispondenti (il **cono di confusione**).

Abbiamo inoltre già detto che è presente una fonte di informazione che dipende dalla **forma del padiglione** auricolare che appunto sembra contribuire alla percezione dell'elevazione. Questo diventa un meccanismo davvero complesso perché chiama in causa non solo la codifica dei tempi e dell'intensità, ma anche la codifica della frequenza: in funzione dell'elevazione di un suono questo viene amplificato dalla pinna in modo selettivo sia rispetto alla sua ampiezza, ma in modo diverso rispetto alla frequenza a seconda dell'elevazione quindi la cosa diventa davvero complicata in quanto bisogna tener conto di 3 elementi. In immagine possiamo notare una superficie in uno spazio tridimensionale in cui però si vede che entro certi limiti, a seconda dell'elevazione del suono, ovvero della qualità di quello che arriva, varia sia in ampiezza che in frequenza quindi c'è una specie di filtraggio che contiene informazioni sull'elevazione.

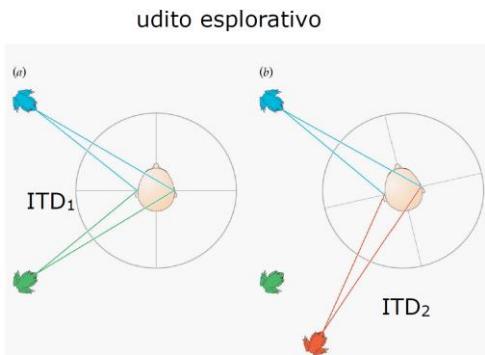


SENSATION AND PERCEPTION, Figure 10.9 (Part 2) © 2006 Sinauer Associates, Inc.

Al di là di questo che è un discorso che riguarda l'**informazione potenzialmente disponibile** ad un ascoltatore immobile monoauralmente e binauralmente, c'è anche qualcosa che assomiglia alla parallasse di movimento *perché noi quando ascoltiamo di solito ascoltiamo attivamente, tutta la percezione è un processo attivo* e ascoltare attivamente vuol dire "tendere l'orecchio" cioè letteralmente ruotare il capo o spostarsi nello spazio in modo da modificare la relazione geometrica che c'è tra la posizione della sorgente sonora e la posizione delle nostre orecchie. Anche questa è una preziosa fonte di informazioni dello spazio in quanto in questo processo esplorativo si rimuovono alcune delle ambiguità che abbiamo visto prima.

- Una terza fonte di informazione sulla posizione spaziale di una sorgente sonora, è fornita dalla forma del padiglione auricolare;
- L'intensità del suono viene modificata in modo selettivo in funzione della posizione rispetto all'azimuth e alla sua elevazione;
- La funzione che descrive questa modifica, tenendo conto anche della frequenza, si chiama *directional/head related transfer function (DTF)*.

### Esempio



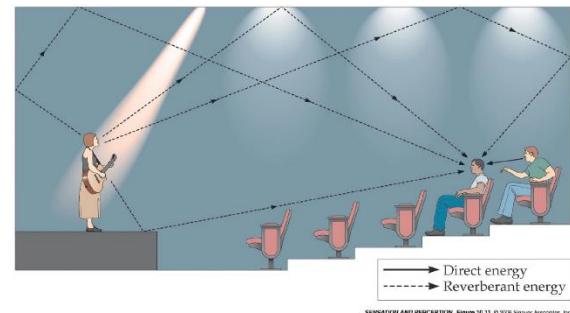
Supponiamo di avere lo sguardo che punta verso  $0^\circ$  quindi dritto davanti a noi e che alla nostra sinistra ci sia una fonte sonora (una rana che emette il verso "cra cra cra"). Questa fonte sonora noi la localizziamo a sinistra, ma la differenza interaurale nel tempo di arrivo è ambigua per il motivo che abbiamo appena visto: trascurando la pinna, il tempo di arrivo interaurale è esattamente lo stesso per due posizioni diverse, davanti e dietro, rispetto all'Azimuth. Questa ambiguità però è davvero tale solo se facciamo un'assunzione iperrestittiva e non realistica che noi ascoltiamo sempre tenendo la testa ferma. Tuttavia, percependo sempre in maniera attiva, quello che avviene è che nel momento in cui sentiamo una fonte sonora provenire da sinistra, basta ruotare leggermente la testa verso sinistra perché questo riveli al nostro sistema uditivo un'altra informazione legata al tempo di arrivo interaurale e cioè che adesso arriva un'altra ITD che di nuovo è coerente con 2 possibili posizioni nello spazio, diverse dalle precedenti, per quanto riguarda la rana verde ma non per quanto riguarda la rana blu. Le due posizioni nello spazio ci danno 2 altre possibili collocazioni, ma solo una di queste è coerente con una delle due precedenti alternative e quindi attraverso questo **meccanismo esplorativo** è anche possibile rimuovere quelle ambiguità di cui parlavamo prima ed è inoltre uno dei meccanismi più importanti per riuscire a localizzare la direzione di una fonte sonora rispetto all'Azimuth.

Rimane comunque un ultimo problema da chiarire riguarda la **posizione della sorgente rispetto alla testa** che è ancora ambigua rispetto alla **distanza assoluta**, cioè *sappiamo da che direzione proviene il suono, ma non sappiamo se la fonte sonora è vicina o lontana da noi.*

In generale non siamo molto bravi a valutare con precisione la distanza assoluta di una sorgente sonora, ma riusciamo comunque a farci una qualche idea anche se imprecisa grazie ad altri possibili meccanismi che sono legati innanzitutto all'**intensità assoluta (1)** → il suono si disperde perdendo di frequenza con una legge che è proporzionale all'inverso del quadrato della distanza (**Legge dell'Inverso del Quadrato**). **Dunque a parità di ampiezza del suono distale, più è distante e più tenderà ad essere debole quando arriva all'orecchio.**

Un altro aspetto che è informativo rispetto alla distanza è il **rapporto tra le alte e basse frequenze (2)** → più una sorgente sonora è lontana e più noi tenderemo a sentire solo le frequenze più basse perché queste hanno lunghezza d'onda più lunga e viaggiano di più perché si disperdoni più lentamente su una distanza maggiore. Questo è quello che possiamo osservare se, durante un concerto, passiamo vicino allo stadio, non sentiamo il suono della chitarra o dei cembali, ma quello del basso proprio perché le frequenze più basse viaggiano più lontano. Inoltre in un ambiente chiuso in cui ci sono degli echi siamo in grado di apprezzare il rapporto tra suoni più intensi che ci arrivano direttamente e suoni che possono essere classificati come **energia riverberante** ovvero echi che sono meno intensi e che si ripetono. Il **rapporto tra energia diretta ed energia riverberante (3)** → è il terzo elemento da considerare.

Immaginando di avere 2 sorgenti sonore in un teatro, una persona che ci parla dal sedile dietro di noi e la voce della cantante, quest'ultima ci arriva da una distanza molto maggiore, dopo che il suono è stato riflesso da varie superfici all'interno del teatro: c'è molta più energia riverberante nel suono della voce della cantante rispetto al suono della persona che ci parla direttamente da vicino. Anche questo fenomeno dà un'idea su se la sorgente sonora sia più nel nostro spazio peripersonale oppure no. Tutti questi meccanismi ci consentono quindi di localizzare la sorgente del suono in maniera più o meno precisa con vari vincoli che abbiamo visto.





A questo punto possiamo finalmente chiederci cosa succede quando dobbiamo localizzare un suono e oltre a sentire che il suono proviene da una certa direzione, vediamo la fonte del suono stesso. Il modello sperimentale che serve per studiare questo meccanismo è l'**effetto ventriloquo**, in cui vediamo la fonte del suono che sembra essere costituita dalla bocca della marionetta che si muove.

Naturalmente si tratta di un'**illusione** in quanto il suono non proviene dalla bocca della marionetta bensì dal ventriloquo che non parla con il ventre, ma utilizzando una particolare tecnica appresa che gli consente di produrre del linguaggio che è quasi normale senza muovere le labbra. L'effetto ventriloquo non si verifica soltanto negli spettacoli di ventriloquismo, ma anche guardando un film al cinema in quanto sembra che il suono provenga dalla bocca degli attori quando in realtà non è così: non sono presenti altoparlanti dietro lo schermo nella posizione in cui ci sono gli attori che parlano, bensì sono collocati nella sala ed in certi casi ci sono anche degli effetti stereofonici per cui sono prodotti ulteriori effetti di localizzazione del suono. **L'effetto ventriloquo fa sì che noi non sentiamo il suono provenire dagli altoparlanti ma dalla bocca degli attori.**

*“Tale effetto si verificava anche durante le prime lezioni di questo corso svolte al cinema in quanto il professore utilizzava il microfono per cui le persone che erano sedute negli ultimi posti della sala, quindi più lontane dalla fonte sonora, non sentivano il suono della voce direttamente ma, fisicamente, gli arrivava dagli altoparlanti provenienti dalla sala. Nonostante ciò avevano l'impressione che il suono uscisse dalla bocca del professore e non dagli altoparlanti.”*

Dunque, l'effetto ventriloquo è un buon modello per studiare la localizzazione sonora in quanto è una situazione in cui la direzione da cui proviene il suono è potenzialmente diversa dalla direzione da cui proviene l'evento visivo che genera il suono. Ad esempio, immaginiamo il ventriloquo che in certi momenti posiziona la marionetta vicino il volto per cui la direzione visiva e quella uditiva sono quasi coincidenti mentre, in altri casi, la marionetta sarà più lontana per cui lo scarto tra le 2 direzioni sarà maggiore. Il fatto che nell'effetto ventriloquo ci sia questa illusione, per cui il suono sembra provenire dalla bocca della marionetta che si muove in sincronia con le parole pronunciate dal ventriloquo, questo fatto tradizionalmente faceva pensare che la vista prevale sull'udito per cui non c'è una localizzazione audio-visiva, ma siamo semplicemente organismi visivi quindi se abbiamo info visive queste predominano su tutto. Tuttavia questa è solo una delle due possibili interpretazioni.

Ci sono infatti due categorie di possibili interpretazioni:

- **Dominanza visiva** → la visione “cattura” l’udito. Abbiamo questi meccanismi che sono molto più precisi di localizzazione di un oggetto visivo e quindi questo cattura anche la localizzazione uditiva. Potenzialmente avremmo queste fonti di informazioni binaurali e monoaurali che ci potrebbero aiutare a localizzare uditivamente la fonte del suono e che dovrebbero dirci che il suono proviene dal ventriloquo stesso e non dalla marionetta. Tuttavia l’informazione visiva sovrascrive queste informazioni e quindi ci sembra che il suono provenga dalla marionetta. Questa interpretazione appartiene a quella classe di spiegazioni che hanno una *visione unisensoriale* della visione, per cui per ogni tipo di compito abbiamo una modalità percettiva particolarmente appropriata.
- **Integrazione ottimale** → i due canali vengono integrati in base al **grado di affidabilità** (inverso della varianza).

Negli ultimi 10\15 anni c’è stato un grandissimo aumento di interesse per gli **aspetti multisensoriali** della percezione in quanto non c’è dubbio *che in condizioni naturali la percezione è un qualcosa che facciamo per via multisensoriale cioè usiamo tutta l’informazione che abbiamo*. Nonostante l’effetto ventriloquo sia più a sostegno della prima ipotesi, è stato fatto notare che esiste una possibile spiegazione che considera anche questo fenomeno come multisensoriale, ed è la teoria di quella che viene chiamata l'**integrazione multisensoriale ottimale**. Il termine “*ottimale*” fa riferimento ad un principio molto generale che sembra funzionare in molte situazioni e che deriva dalla statistica. Quest’ultima ci dice che se si deve stimare un parametro ignoto in una popolazione e si hanno a disposizione 2 campioni estratti da tale popolazione e con cui posso calcolare uno stimatore, la stima migliore non è scegliere uno dei 2 campioni (ad esempio quello più numeroso) ignorando l’informazione dell’altro bensì combinare l’informazione dei 2 campioni dando proporzionalmente più peso al campione più preciso ovvero quello che è meno variabile. **La precisione è dunque l’inverso della variabilità del campione ovvero della sua varianza**. Trasferendo quest’idea in un contesto multisensoriale, possiamo dire che abbiamo un campione visivo ed uno uditivo, li usiamo entrambi, ma li pesiamo in base al loro **grado di precisione**. Sapere quant’è preciso il segnale visivo e quello uditivo misurando il segnale neurale è molto difficile perché presuppone il fatto che *il cervello si rappresenti la precisione di questo segnale*, ed è tutt’ora aperto il dibattito, in certi ambienti di neuroscienziati, su quale potrebbe essere il codice neurale per rappresentare la precisione di un segnale sensoriale.

Però, dal punto di vista psicofisico, noi possiamo “raccogliere delle sensibilità” e vedere quanto queste variano rispetto a misurazioni ripetute. L’ inverso di questa variabilità è un buon indicatore della presunta precisione del segnale neurale. Inoltre precisione, nella teoria dei test, quella che viene chiamata “reliability” ovvero affidabilità → quanto ci posso credere.

### **Come potrebbe tutto ciò spiegare l’effetto ventriloquo?**

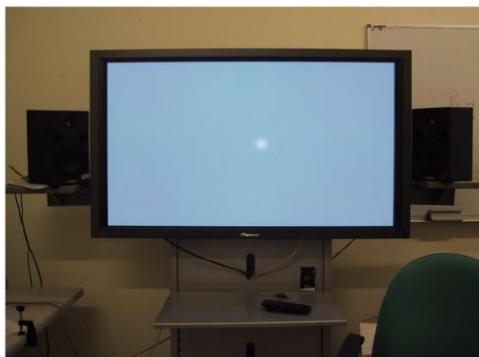
Potrebbe farlo in base al fatto che la localizzazione di un oggetto per *via visiva* dispone di meccanismi più precisi, almeno dal punto di vista teorico, rispetto a quelli uditivi. Quindi potrebbe essere che quella che sembra **dominanza visiva** è invece il risultato che in quelle particolari condizioni (marionetta con il ventriloquo) la localizzazione di una direzione di provenienza del suono per via *uditiva* è molto più sfumata, ovvero è un’informazione compatibile con molte più potenziali direzioni, rispetto a quanto non sia precisa l’equivalente localizzazione visiva. Dunque non è vero il fatto che viene usata solo la visione ma che, in queste condizioni, il **peso** assegnato all’informazione visiva è maggiore. Se tutto ciò è vero ne discende una previsione, cioè che dovrebbe essere possibile produrre un **effetto ventriloquo inverso**, in cui sembra che sia il suono a catturare la localizzazione visiva e ciò dovrebbe avvenire nelle condizioni in cui la localizzazione visiva diventa molto imprecisa. Immaginiamo una situazione in cui siamo di notte in mezzo alla nebbia e le forme sono molto sfumate, non si capisce bene dove siano collocate nello spazio, nel momento in cui emettono dei suoni, dovrebbe succedere che ci sembra di vederle più in accordo con la direzione di provenienza del suono di quanto non in accordo alla posizione vista.

A questo punto esaminiamo un classico esperimento fatto in Italia, a Pisa, da 2 ricercatori australiani: **David Burr** e **David Alais**. **David Burr** è un ricercatore di origine australiana, sposato con un’italiana, che insegnava all’Università di Firenze e **David Alais** collaborava con lui, mentre attualmente insegnava all’Università di Sydney.

Si sono chiesti: “**Come produrre un effetto ventriloquo inverso?**”. Per produrre un effetto ventriloquo inverso bisogna chiedere alle persone di localizzare una fonte sonora.

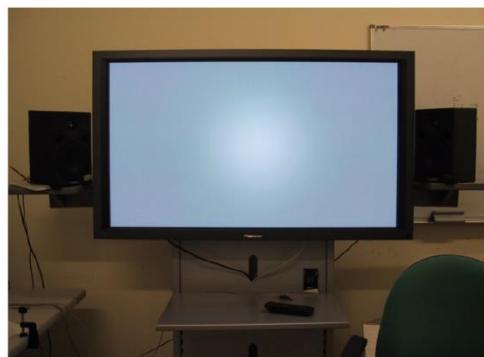
L’esperimento prevedeva l’utilizzo di uno schermo con delle casse posteriori quindi si poteva manipolare la direzione fisica da cui il suono proviene facendolo uscire da una delle casse. Sullo schermo era presentato un dischetto luminoso di cui veniva manipolata la precisione spaziale rendendolo più o meno sfocato: se il bordo del dischetto è molto netto (alta frequenza spaziale) la localizzazione è precisa, se il bordo è meno preciso la posizione vera di quel dischetto diventa un po’ più incerta quindi il segnale visivo è più impreciso. Questa variabile viene chiamata *blur*, ovvero il **grado di sfocatura** → maggiore è il grado di sfocatura più impreciso sarà il segnale visivo.

4 deg blur



Alais & Burr Current Biology, 2004

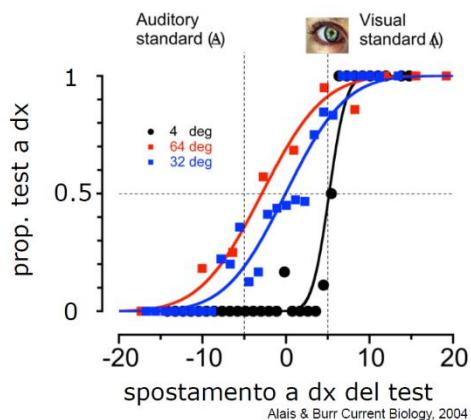
32 deg blur



Alais & Burr Current Biology, 2004

## Risultati

In questo esperimento chiedevano ai soggetti di localizzare la direzione della fonte sonora in funzione del grado di sfocatura che aveva il concomitante segnale visivo. E' un po' la situazione del ventriloquo ridotta all'osso in quanto abbiamo un segnale visivo e uno acustico che possono provengere da direzioni diverse.



Nel grafico che mostra i risultati dello studio sono presentate **3 funzioni psicometriche** (ovvero le classiche sigmoidi) che rappresentano l'andamento di una soglia. Sull'asse dell Y c'è la proporzione di risposte "il test è a destra" rispetto ad uno standard, quindi è una *capacità discriminativa* (che è la classica cosa che produce una sigmoide) in funzione di uno spostamento a destra del test. Quindi il valore pari a 0 vuol dire che sono perfettamente allineati, valori negativi "è a sinistra" ed infine positivi "è a destra". Le 3 sigmoidi si riferiscono a tre livelli di sfocatura: **poco**,

**intermedio, molto sfocato**. Sono inoltre indicati due punti di **uguaglianza oggettivi potenziali**, l'uno corrisponde ad uno standard visivo dove è il centro del disco sfocato, e l'altro allo standard uditivo, cioè dov'è collocata realmente la fonte sonora. In funzione di queste cose si può apprezzare che il **punto di uguaglianza soggettivo** (ovvero il punto in cui la sigmoide interseca una proporzione del 50%, il punto che identificheremmo con la **soglia** secondo il **criterio della massima incertezza**), è perfettamente coincidente con lo standard visivo quando il segnale è molto poco sfocato, quindi quando la visione è molto precisa cattura l'udito e la persona dice che sente provenire il suono dove vede l'oggetto → classico effetto del ventriloquo.

Tuttavia, **aumentando la sfocatura**, la sigmoide diventa più piatta e la *zona di incertezza aumenta*: allo stesso tempo però il punto di uguaglianza soggettivo si sposta indietro, ovvero verso la direzione dello standard uditivo, e naturalmente si sposta di più quanto più lo stimolo visivo è sfocato fino a quasi arrivare in corrispondenza del punto di uguaglianza oggettivo dello standard uditivo

→ **effetto ventriloquo inverso**, ed è un'evidenza sperimentale molto convincente a favore di questa idea dell'integrazione ottimale.

\*\*Ulteriori spiegazioni grafico:

Quando il test è molto spostamento rispetto allo standard, lo spostamento è talmente grande che non c'è incertezza: è sempre il 100% oppure dall'altra parte è lo 0%.

La cosa informativa è come si sposta la zona di incertezza nella parte centrale quando si può vedere l'effetto di questa pesatura.

## Lezione 24

### Multisensorialità, spazio e corpo

Nel momento in cui si passa da un livello basso ad uno alto di percezione, c'è un'operazione di trasformazione di coordinate, soprattutto dal punto di vista della **percezione spaziale**, ma non solo. Ad esempio, si è già visto per quanto riguarda la visione che essa inizia con una codifica di caratteristiche e *coordinate spaziali retinotopiche* e che poi queste diventano delle coordinate che non sono più ancorate sulla bidimensionalità della retina, ma diventano tridimensionali.

Il professore suggerisce che la percezione dello spazio non è quella dello spazio euclideo (un sistema di coordinate tridimensionali come quello che si studia in geometria), ma è più una rappresentazione dello spazio in termini di coordinate che si riferiscono all'ambiente (quindi la posizione degli oggetti rispetto ad altri oggetti).

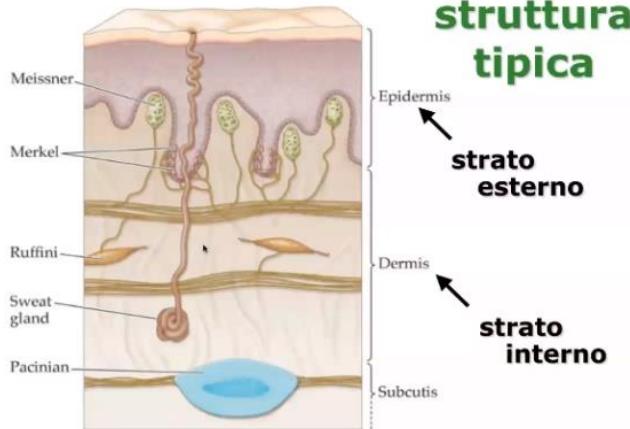
Questa distinzione si può anche etichettare come un passaggio da coordinate **egocentriche** (cioè riferite a qualche aspetto dell'organismo che percepisce, in questo caso la retina) a delle coordinate **allocentriche** (in cui i singoli oggetti vengono rappresentati rispetto ad altri

	Egocentriche	Allocentriche
<b>Visione</b>	retinotopiche	ambientali
<b>Udito</b>	tonotopiche	ambientali
<b>Tatto</b>	somatotopiche	ambientali
Trasformazioni di coordinate.		

oggetti che si vedono nell'ambiente). Lo stesso discorso si può fare per le *coordinate tonotopiche* della coclea che poi vengono utilizzate per la localizzazione uditiva degli oggetti; vedremo che questo vale anche per il tatto, nel passaggio da coordinate riferite al corpo, alla superficie del derma (somatotopiche) a coordinate ambientali, riferite all'ambiente.

Due cose su come funziona la **somatosensazione**:

- 1) Noi abbiamo *recettori tattili* che sono distribuiti sulla **pelle**: la pelle non è un organo come un occhio o un orecchio, ma una superficie con una certa struttura molto ampia (circa 2 m<sup>2</sup>) che corrisponde al 5-8% del peso totale del nostro corpo in base alla nostra struttura fisica;
- 2) I *recettori tattili* sono distribuiti nel derma in maniera molto diseguale, un po' come la retina che è molto disomogenea dal punto di vista della distribuzione dei fotorecettori, anche i **meccanorecettori sono distribuiti nella pelle in modo molto variabile in base al distretto somatico che analizziamo**.



### struttura tipica

Una struttura tipica è quella che vediamo sulla punta delle dita: abbiamo l'**epidermide** (lo strato esterno della pelle) e il **derma**; procedendo dall'esterno verso l'interno troviamo varie tipologie di meccanorecettori: quelli di **Meissner**, quelli di **Merkel**, quelli di **Ruffini**, quelli del **Pacini**. Sono classicamente distinti in base a due criteri:

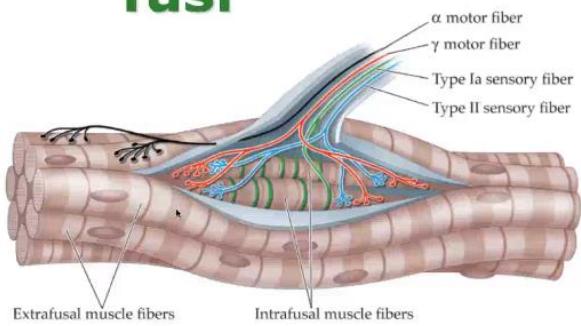
- 1) La **rapidità con la quale si adattano** (la capacità di produrre un segnale sostenuto nel tempo), quelli a *lento adattamento* continuano a sparare in presenza di uno stimolo continuo e quelli a *rapido adattamento* (RA) producono una risposta più transiente, segnalano l'inizio dell'evento meccanico e poi smettono di sparare;
- 2) La **grandezza del campo recettivo** (CA). Vengono distinti in tipo 1 e 2, quelli con campo recettivo più piccolo sono più vicini alla parte esterna della pelle (epidermide) e quelli con campo recettivo più grande sono più interni.

A volte si tende a sottolineare come queste caratteristiche dell'avere un campo recettivo più grande o più piccolo, un adattamento più lento o più veloce, si prestino di più ad essere utilizzate per acquisire informazioni in determinati tipi di compito tattile, come riconoscere un carattere braille o mantenere la stabilità di una presa o rilevare un tocco leggero. Il prof non è d'accordo con questi ragionamenti perché ritenere che a livello recettoriale ci siano dei meccanismi responsabili di compiti complessi come questi è sicuramente semplicistico, in quanto tutti questi tipi di compito sottendono sistemi molto più articolati.

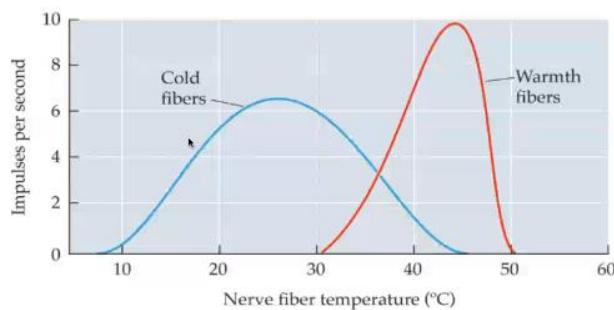
Oltre che nella pelle, noi abbiamo dei recettori somatosensoriali, estremamente importanti, anche nei *muscoli* e nelle *articolazioni*: sono i **fusi neuromuscolari** e sono dei recettori di stiramento.

Nel caso dei *muscoli* segnalano al cervello il fatto che un muscolo si stira, si allunga. Abbiamo poi dei **recettori di posizione** nelle *articolazioni* che sono in grado di segnalare se l'articolazione ha una certa apertura o meno.

## fusi



## termorecettori



per il freddo, in realtà hanno un'escursione abbastanza ampia, da dieci a quaranta gradi con un picco tra i 20 ed i 30 mentre quelle per il caldo sono più spostate verso temperature più alte. Ci sono poi anche i **recettori per il dolore o nocicettori**, ma la percezione del dolore è un tema ancora più complicato, con un'altissima componente cognitiva che non approfondiremo ulteriormente.

**La percezione tattile non è una misura passiva di forze meccaniche che interagiscono con la nostra pelle ma, come tutta la percezione, un processo attivo di esplorazione**, anzi, le ricerche sulla percezione delle forme attraverso la somatosensazione hanno contribuito in maniera fondamentale a chiarire quanto è importante concepire la percezione come un processo attivo di esplorazione.

Tutto ciò lo dobbiamo a **Gibson**, famoso per tanti motivi, è stato uno dei teorici più influenti della percezione. Nel 1962 ha pubblicato un lavoro che intitolato “*Osservazioni sul tatto attivo*” in cui egli dimostra che l'attività intenzionale è fondamentale nella percezione nel caso del tatto (poi lui perseguitò questa idea anche in altri campi, per esempio nella visione).

Negli anni '80 un filosofo della scienza, di nome **Harrè**, ha pubblicato un libro divulgativo “*20 esperimenti che hanno cambiato il mondo*”, nel quale parla dei 20 esperimenti che sono stati cruciali nelle diverse discipline per cambiare il punto di vista di una disciplina su un particolare problema, tra questi si trova un esperimento di psicologia che ora andremo a trattare, esperimento condotto da Gibson.

VOL. 69, NO. 6

NOVEMBER 1962

## PSYCHOLOGICAL REVIEW

OBSERVATIONS ON ACTIVE TOUCH<sup>1</sup>

JAMES J. GIBSON  
Cornell University



FIG. 1. The six forms to be identified by touch.

### action for perception: exploratory procedures



Gibson, nel lavoro, fa delle associazioni libere, un po' di **fenomenologia**, per farlo si costruisce una specie di scatola nella quale il soggetto può mettere la mano e fare dei *compiti di percezione tattile*, cioè toccare con la mano degli oggetti. Dall'altra parte la scatola è aperta e quindi lo sperimentatore (Gibson) può osservare e prendere nota. Gibson fa una serie di ipotesi sulla natura dell'esplorazione tattile e successivamente fa un esperimento più

controllato per cercare di confrontare la percezione tattile come **tatto passivo** (cioè quando sei con la mano e qualcosa passivamente ti preme sulla mano) o come **tatto attivo** (quando attivamente vai ad esplorare qualcosa muovendoti).

Per fare questo mise in piedi un esperimento semplice, utilizzando formine per biscotti: il “*cookie cutter experiment*” (formine biscotti) (storia confermata dalla vedova di Gibson, il quale aprì un cassetto di casa e si procurò le formine).

Gibson fa quindi un **esperimento di riconoscimento cross-modale**: prima di tutto chiede ai soggetti di associare un numero ad ognuna delle forme dei biscotti presentate (quindi la goccia 1, la mezzaluna 2, ecc.) poi li benda e in ordine casuale posiziona in mano ai soggetti una di queste sei formine chiedendo loro di riconoscerla tramite il numero associato (1,2,3,4,5,6). Questo viene fatto in **tre condizioni** diverse (Il Bruno, Pavani è più dettagliato per questo esperimento):

- 1) *tatto passivo*, condizione nella quale la mano è ferma e lo sperimentatore preme la formina contro il palmo della mano;
- 2) *tatto attivo*, la forma è sul tavolo ed il soggetto la esplora come vuole;
- 3) condizione di *tatto dinamico*.

Gibson si rende conto che i soggetti nella condizione di *tatto passivo* se la cavano benino, riescono a rispondere correttamente, un po' meglio del caso (una risposta casuale sarebbe un po' meno del 16/20%). Nella condizione di *tatto attivo* i soggetti non sbagliano mai, arrivano al 95/96% di risposte corrette. **Una differenza abissale, quindi, con il tatto passivo**, infatti in generale è molto più semplice ottenere informazioni molto accurate sulla forma di un oggetto se è possibile *esplorarlo in maniera attiva*.

Gibson si è chiesto quanto di questo risultato sia dovuto al fatto che la forma, nella condizione di tatto attivo, non sia statica, ma si *modifichi nel tempo* (perché esplorando con la mano, il pattern di informazioni somatosensoriali meccaniche sulla mano varia in funzione del tempo) e quanto invece sia dovuto al fatto che questa variazione è una variazione che è *contingente all'intenzione di muoversi* del partecipante. **Per questo ha introdotto la terza condizione, una specie di controllo.**

Nella condizione di *tatto dinamico* la forma veniva mossa sulla mano, ma in maniera passiva, cioè lo sperimentatore muoveva la forma sulla mano cercando di riprodurre il movimento con il quale si esplora con la mano. Questa condizione è stata chiamata *tatto dinamico*, perché c'è la componente del movimento, ma non è un movimento intenzionale.

In questa condizione la prestazione si colloca **ad un livello intermedio tra quella di tatto passivo e quella di tatto attivo**, quindi, la conclusione di Gibson era che **noi estraiamo molta più informazione da un pattern che si trasforma rispetto ad un pattern statico, e questa informazione è ottimale quando queste trasformazioni sono associate ad una esplorazione intenzionale**. (Lui diceva nel momento in cui la stimolazione è ottenuta e non imposta, cioè quando l'informazione arriva al sistema percettivo in funzione di un'esplorazione attiva).

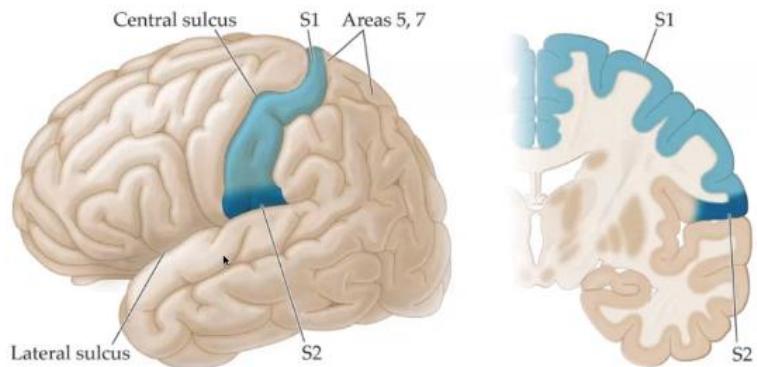
Rispetto a questo esperimento dobbiamo pensare che il risultato non è affatto ovvio: noi tendiamo a pensare che la percezione comincia con una specie di “fotografia del mondo esterno” (discorso fatto per

la visione), per quanto riguarda il tatto pensiamo che inizi con un pattern stabile di pressioni sulla nostra pelle, quindi, quando io ti premo contro la formina a forma di stella, si forma un pattern a forma di stella sulla mia mano e dico che è la formina a forma di stella. Ciò dovrebbe essere la **situazione ottimale** per percepire la stella. Al contrario, quando io esploro con la mano, sulla mia pelle c'è un'assoluta confusione, un susseguirsi caotico di pressioni che cambiano nel tempo, si modificano, prima in un punto, poi in un altro, quindi, *come è possibile che queste pressioni caotiche e variabili siano preferibili rispetto a qualcosa che in principio è già ben identificato e stabile?* Non a caso i ricercatori della visione hanno cominciato ad interrogarsi sul fatto che le *immagini stabilizzate* nella visione tendono a scomparire; è un po' lo stesso ragionamento per il tatto: il **tatto passivo** è come l'**immagine stabilizzata**, la “somatosensazione stabilizzata”, che però funziona male, funziona molto meglio il movimento, il cambiamento. L’idea sull’esplorazione è stata ripresa da molti ricercatori. C’è molta letteratura sulle diverse procedure esplorative della mano: la mano è, infatti, un organo estremamente complesso e noi utilizzandola tendiamo a mettere in atto una grande quantità di procedure esplorative.

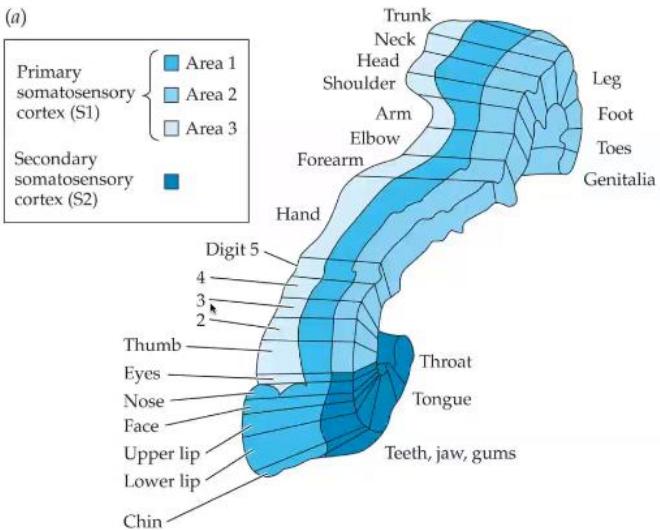
### **Cosa succede all’informazione somatosensoriale? Come il processo trasduzionale di energia meccanica viene processata?**

**S1 e S2 → Mappa somatotopica del corpo.**

Una mappa del corpo in cui a determinati distretti corporei corrispondono a diversi settori corticali in S1 e S2.



Detto questo, dobbiamo chiederci cosa succede all’informazione somatosensoriale, quando questa energia meccanica trascorsa arriva al cervello (saltando un po' di passaggi) nell’area somatosensoriale primaria che è la striscia di corteccia che si trova subito dopo il solco centrale, la parte più anteriore del lobo parietale.

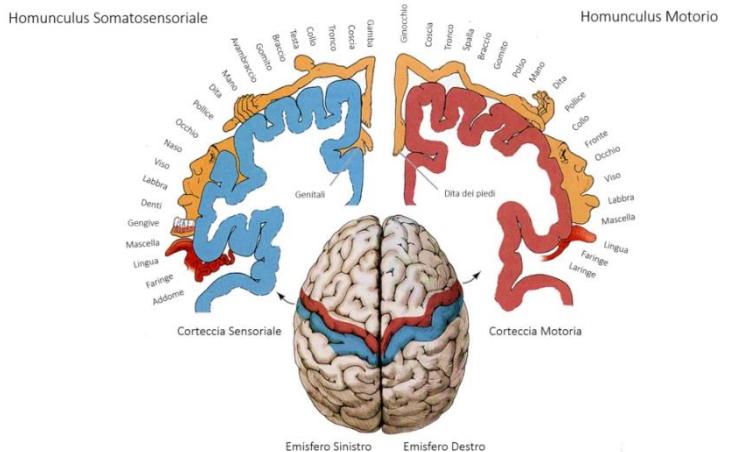


Essa è suddivisa di solito in due parti S1 e S2 che noi considereremo un'unica cosa. Sappiamo che nell'area S1 si trova una *mappa del corpo*, nel senso che a seconda del punto in cui ci collociamo in S1 e S2, troviamo unità neurali che ricevono input da distretti somatici diversi. In immagine è raffigurato un esempio della mappa somatotopica rappresentata e distribuita a livello corticale.

Partendo dal centro dell'emisfero, che stiamo considerando, in prossimità della *scissura interemisferica*, troviamo aree che ricevono input dei genitali, dalle dita dei piedi, dalle gambe poi salendo ancora troviamo il tronco, il collo, la testa, le spalle, il gomito, l'avambraccio, la mano, le 5 dita e poi ancora il volto, gli occhi, il naso, le labbra, il mento e poi le parti più interne della bocca, i denti, gengive, mandibola, la lingua e la gola. **Questa è una mappa somatotopica**, una mappa del corpo in cui **a certi distretti cutanei corrispondono specifiche locazioni in S1 e S2**, questa mappa ha una caratteristica (descritta per la prima volta da Penfield) cioè alcune aree corporee sono **sovrrappresentate**, cioè **la proporzione di corteccia dedicata a certe aree non è proporzionale all'area della pelle che corrisponde a quelle aree corporee**.

## Omunculus di Penfield

**Sovrapposizione corticale nelle aree di maggior interesse →** non c'è corrispondenza tra superficie corporea e superficie corticale, ma c'è una corrispondenza recettore-corteccia, questo principio si vede anche nella corteccia motoria, nella corteccia visiva ecc...



Il fenomeno è illustrato graficamente per quanto riguarda l'area somatosensoriale primaria e anche per quanto riguarda la corteccia motoria primaria, che contiene una mappa del corpo per certi versi simile, ma non uguale. La parte motoria è quella di destra dove c'è ancora di più una sovrarappresentazione delle parti, in particolar modo il volto e le mani, rispetto ad altre. Per esempio, dal punto di vista motorio, il tronco e le spalle non sono rappresentate mentre invece le gambe ed i piedi sono sottorappresentati.

Lasciando perdere l'aspetto motorio e concentrandoci sulla parte somatosensoriale (che è quella grigio scura) vediamo comunque questa sovra rappresentazione del volto, della bocca e delle mani, che da soli ricoprono quasi  $\frac{3}{4}$  dell'estensione di S1 ed S2. È una mappa del corpo **per certi versi distorta**, come se fosse una specie di omuncolo interno che è molto diverso da come noi percepiamo il nostro corpo.

## Soglia dei due punti



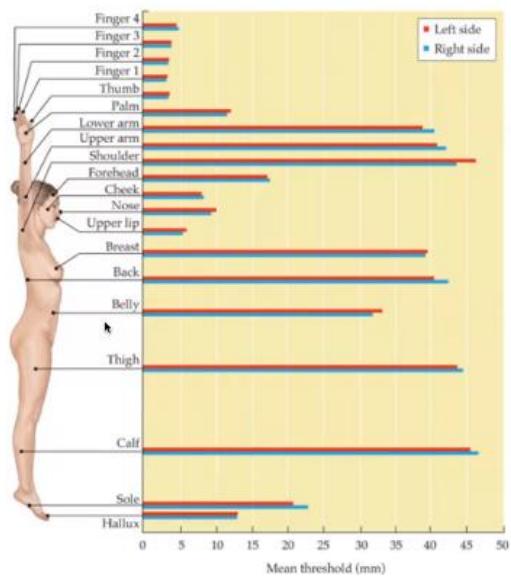
La **soglia dei due punti** è diversa in due diversi distretti corporei: è molto minore sul polpastrello, dove ci sono più meccanorecettori che permettono una maggiore *acuità tattile* (*maggior sensibilità*) rispetto a quelli distribuiti sulla superficie dell'avambraccio.

Quindi la soglia dei due punti è molto minore sul polpastrello rispetto all'avambraccio, utilizzando un piccolo oggetto riuscirò a percepire come due punti distinti due punti molto più vicini sul polpastrello rispetto all'avambraccio.

Per certi versi è possibile dimostrare che questa mappa, a livello corticale, provoca delle distorsioni a livello comportamentale. Questo si può dimostrare misurando la **soglia dei due punti che è quell'estensione sulla pelle tale per cui un soggetto è in grado di dire che una stimolazione duplice è realmente duplice** (distinguere tra una pressione con due punti e una con un punto solo).

## Densità recettoriale tattile nel corpo in funzione della “regola dei due punti”

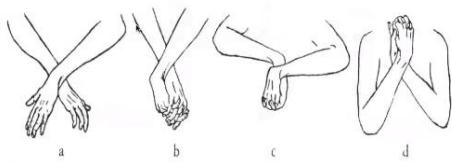
Se prendo un soggetto gli posso far notare che la stessa estensione sul polpastrello del pollice lui la sente come due punti, ma la sente come un punto solo sull'avambraccio; la soglia dei due punti è diversa in certi distretti corporei rispetto ad altri, precisamente qui vediamo una specie di funzione psicofisica: si vede la soglia (inverso della sensibilità) per tutta una serie di posizioni sul corpo e vediamo che in alcuni punti la soglia è particolarmente bassa (sulle dita, sul palmo della mano, sulla zona intorno al naso, alle labbra e le guance) mentre diventa più alta in altre zone (sulla schiena, sul torace, avambraccio ecc.).



## Illusione di Aristotele e l'illusione giapponese



### illusione “giapponese”



Noi sappiamo già che questa questione di distinguere uno punto da due punti è particolarmente strana avendo studiato *l'illusione di Aristotele*: in alcuni casi si può avere l'impressione che un oggetto singolo si sdoppi. Lo possiamo fare sul naso incrociando le dita, o anche sulla bocca, stirando la bocca con le dita, e mettendo un oggetto singolo tra le labbra ci sembrerà di sentire due oggetti.

Questo accade per certi versi anche nella cosiddetta “*illusione giapponese*”, un giochino in cui si chiudono le mani incrociando le dita e poi si portano le braccia incrociate contro il torace in maniera che le dita (gli avambracci sono stirati) diventino molto difficili da localizzare (se ti chiedono di muovere un dito senza toccarlo, lo si indica e magari si muove un altro dito perché non si riesce bene a mettere a registro l'impressione del corpo e i comandi motori).

Si può provare anche *l'illusione della “forchetta deformata”*: prendere un forchettone per tirare su la pasta, fare aderire delicatamente la punta della lingua sui rebbi del forchettone, in maniera da fargli fare una specie di onda (sopra sotto, sopra sotto) e poi chiedete alla persona come sente la forchetta. Lei vi dirà che gli sembra che la forchetta non abbia i rebbi allineati e che siano deformati, uno più alto ed uno più basso.

Tutte queste illusioni suggeriscono che nella somatosensazione non c'è solo una lettura del segnale che arriva in S1, ma c'è qualcos'altro che poi aggiorna questo segnale, lo riprocesso alla luce di informazioni sullo stato del nostro corpo, quindi c'è un'interazione tra propriocezione e somatosensazione (il segnale che arriva dai recettori) che ha a che fare con il cambio di coordinate di cui parlavamo prima e questo si vede benissimo nella cosiddetta *illusione di Weber*.

Oltre ad aver formulato la legge di Weber, lui studiava il tatto e nel suo saggio è presente appunto l'illusione di Weber, che è *l'altra faccia della medaglia della soglia dei due punti*: **un oggetto che ha una lunghezza fissa, se lo viene premuto contro i polpastrelli sembrerà più lungo mentre se viene premuto contro altri distretti corporei (avambraccio o schiena) sembrerà più corto**. Questo sembra una conseguenza del fatto che la soglia dei due punti è più bassa, c'è una sensibilità migliore sul polpastrello e quindi l'oggetto sembra più grande.

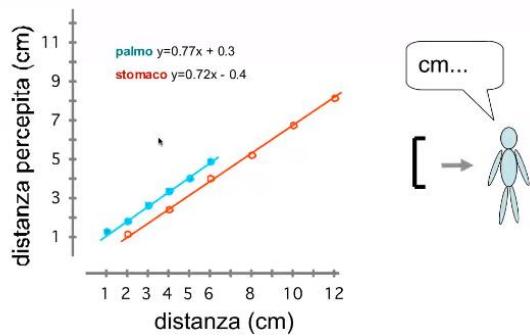
Ci si è chiesti però se questo dipende soltanto dal fatto che ci siano più recettori o recettori più densi e quindi più corteccia. Uno potrebbe dire che l'illusione di Weber dipende dal fatto che in corteccia i polpastrelli sono sovra rappresentati mentre l'avambraccio praticamente non è rappresentato e quindi è normale che la sensibilità è migliore o peggiore rispettivamente. Qualcuno ha provato a testare questa idea in maniera più quantitativa.

Sappiamo che in psicologia spesso si testano ipotesi qualitative, tipo il gruppo A deve essere maggiore del gruppo B, però, altrettanto interessante è chiedersi maggiore di quanto, quanto grande deve essere la differenza in base al mio modello. In questo caso ci si chiede *quanto dovrebbe essere diversa la stima di grandezza se quello che conta è la grandezza relativa della corteccia dedicata a quel distretto somatico*.

Si possono fare previsioni e confrontarle con i dati.

Questo è un di **Green** che, utilizzando un **magnitude estimation** alla Stevens, ha chiesto a soggetti di stimare la distanza percepita tra due punti che venivano premuti (tatto passivo) su distretti corporei diversi per oggetti di diversa grandezza, per poi stimarne la forma della funzione psicofisica. Dato che si tratta di distanze (esponente 1) ci aspettiamo delle rette e vediamo che *maggior è la distanza sulla pelle, maggior è la distanza percepita, più o meno con pendenza unitaria*.

### **Green (1982)** **Percept. & Psychophysics**



Quello che ci interessa, però, non è la pendenza, ma l'**intercetta** di questa retta: sul palmo della mano (dati in azzurro) la retta si colloca più in alto rispetto alla retta dello stomaco (nella quale possiamo mettere anche delle distanze maggiori), e vediamo che l'intercetta è più alta. **Questa è l'illusione di Weber.**

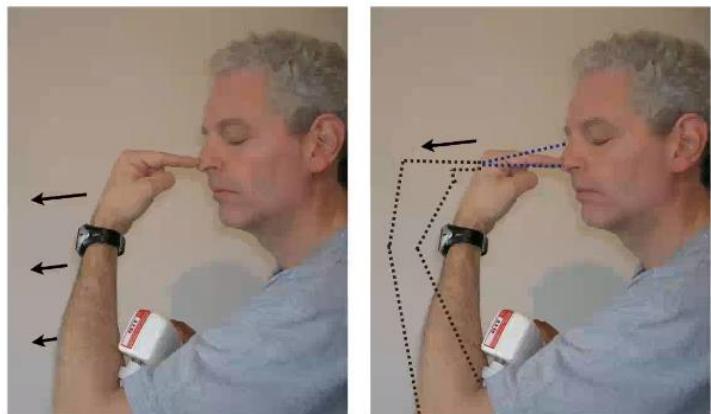
Tuttavia, la differenza tra le intercette non è così grande.

Weber faceva proprio questo ragionamento: se prendiamo l'area di corteccia dedicata al palmo della mano e quella dedicata allo stomaco, dovremmo aspettarci una differenza molto più grande tra le due rette. L'illusione di Weber c'è ma non è così grande come ci si aspetterebbe se contasse solo l'area relativa.

Sono presenti quindi altri processi che intervengono e sono quasi sicuramente processi di natura multisensoriale. **Se prendessimo le rappresentazioni corticali questa differenza dovrebbe essere molto più grande, ipotizzando una proporzionalità diretta tra rappresentazione corticale e percezione.** Quindi non bastano le rappresentazioni corticali per spiegare questo, c'è qualcosa di più che ci serve per comprendere al meglio la percezione tattile.

**L'illusione di pinocchio:** si somministra una vibrazione sul tendine del bicipite; questa vibrazione stimola i fusi neuromuscolari, pertanto il cervello riceve un segnale che è coerente con il fatto che il bicipite si sta stirando. Se esso si stira, vuol dire che il muscolo antagonista (tricipite) si sta contraendo quindi il cervello riceve un'informazione sul fatto che il braccio si sta allungando anche se in realtà resta fermo. Questa è già un'illusione (coerente con la stimolazione data in precedenza). Ora immaginate di toccarvi il naso con la punta del dito indice mentre vi stimolate al bicipite. In questo caso la stimolazione tattile al tendine del bicipite, dice al cervello "il braccio si sta allungando", però, ci sono altre informazioni, in particolare la punta del naso e la punta e del polpastrello che dicono al cervello "*il dito indice sta toccando il naso*".

Il cervello in questo caso non si limita a registrare questi due eventi, ma cerca di **organizzarli in un'interpretazione coerente** di quello che sta succedendo al mio corpo nello spazio e paradossalmente quello che può succedere è l'*impressione che il naso si allunghi*. È una soluzione per questo insieme di stimoli, il braccio si apre, ma rimane costante questa pressione del dito sul naso.

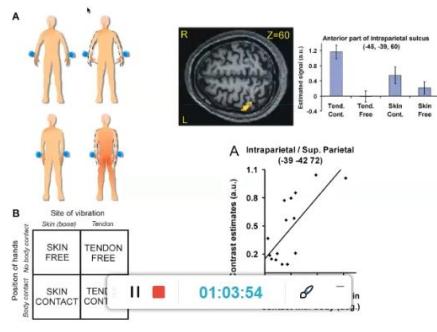


Se vengono fatti test in maniera sistematica (articolo dell'88), dove viene stimolato il tendine del bicipite, le persone possono riportare varie cose: possono riportare l'illusione di pinocchio (il naso si allunga), ma possono anche riportare un altro tipo di illusione, il naso rimane normale, ma le dita si allungano, molto spesso succede che le persone riportino che entrambe le cose accadono. **Questo suggerisce che il tatto aggiorna, sottopone ad ulteriore elaborazione le informazioni somatosensoriali che vengono dai recettori, alla luce anche di una rappresentazione interna del corpo, di come il corpo si sta muovendo nel caso dell'illusione di Pinocchio.**

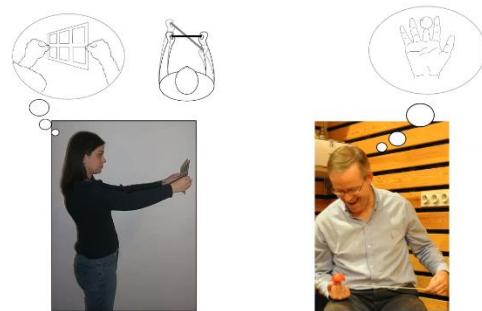
Esistono *varianti multisensoriali* dell'illusione di Pinocchio: ad esempio con la pallina di ping-pong sul dito, che essendo vista dall'alto sembra una sfera completa e quindi si crea la rappresentazione della sfera sul dito, per cui si ha la percezione che il dito sia più corto.

Anche l'*illusione della mano di gomma* è sulla stessa linea di illusioni e “deformità percettive”.

Dell'illusione di Pinocchio esiste anche una versione più recente che potremmo chiamare *l'illusione del vitino di vespa*: si può mettere in atto stimolando con lo stesso metodo i tendini del tricipite del braccio mentre un soggetto si sta stringendo le mani sulla vita. Il soggetto avrà l'impressione che le mani si muovano verso la parte centrale del corpo e quindi la vita diventi più stretta, oppure, nel caso opposto, quindi stimolando il tendine del bicipite, le braccia si aprono e quindi il giro vita si allarga.



Ehrsson et al. (2005), PLOS Biol.



Bruno, Dell'Anna & Jacorossi (2004) Perception  
Ehrsson, Syms, Van der Hulst & Wagemans (2016) Current Biology

Questo è un lavoro molto elegante di **Ehrsson**, uscito 15 anni fa, il quale ha usato la risonanza per mostrare che quando le persone sono soggette *all'illusione del vitino di vespa*, si trovano delle attivazioni specifiche in alcune parti del **lobo parietale** dei soggetti.

In particolare, del **solco intraparietale**, quindi diverso dall'area somatosensoriale primaria. Tra l'altro *un'area parietale non è solo legata alla somatosensazione, ma coinvolta anche nella pianificazione dei movimenti e ai processi di natura multisensoriale*.

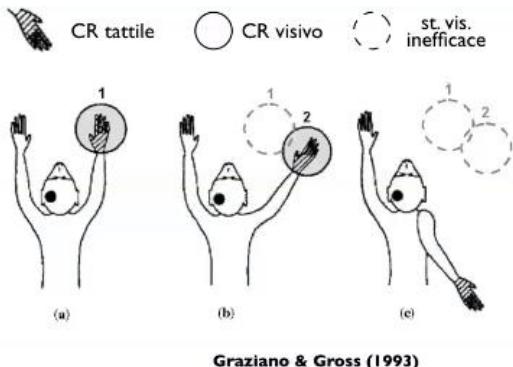
Quindi, ciò è coerente con quello che noi sappiamo dagli studi sull'**agnosia tattile**, (esiste anche l'**agnosia visiva** per gli oggetti), che **di solito è associata con dei danni focali nel lobo parietale ma non nella parte anteriore, nella parte un po' più posteriore del lobo parietale**.

Questi processi sappiamo essere multisensoriali perché sono presenti dimostrazioni convincenti: per esempio giochi in cui si ha l'impressione che si allunghi un braccio o che un dito si accorci. Si può prendere una pallina da ping-pong, tagliarla a metà e metterla a modo di fungo su un dito, guardandola dall'alto si ha l'impressione che il dito si accorci perché la pallina da ping pong vista dall'alto si ha l'impressione che sia una sfera completa. Sentendo il contatto del dito contro la superficie della sfera vuol dire che il dito è più corto.

Un'altra illusione interessante visuo-tattile è *l'illusione della mano di gomma*, che succede quando nascondo alla vista di un partecipante la sua mano vera e posiziono in prossimità della sua mano vera una replica che può essere di gomma, di gesso (l'importante è che sia simile ad una mano) e successivamente, utilizzando un apparato (due pennellini), lo sperimentatore somministra simultaneamente uno stimolo visivo ed uno tattile. Lo stimolo tattile viene somministrato alla mano vera del partecipante e lo stimolo visivo è il fatto che il partecipante vede che il pennellino stimola la mano finta. Quando queste due stimolazioni sono a registro dal punto di vista spazio-temporale dopo alcuni secondi, **il soggetto ha l'impressione che la sensazione di venir toccato sul dito avviene nel punto in cui c'è la mano di gomma e non dove c'è realmente la mano**. C'è quindi una **riorganizzazione della percezione del corpo nello spazio e di dove la persona percepisce**, nello spazio, il fatto che è avvenuto questo evento meccanico (sono stato toccato sulla mia pelle). Se la stimolazione è fatta molto bene, **può succedere che il soggetto faccia esperienza che la mano di gomma sia la propria mano**, una specie di scambio tra mano vera e mano di gomma. È un po' la stessa cosa che succede quando noi *utilizziamo uno strumento*, quando giochiamo a tennis, utilizzando la racchetta, oppure quando tagliamo un foglio di carta con forbici: abbiamo la sensazione di sentire la forbice che taglia nel punto in cui sta tagliando (chiaramente nella forbice non ci sono recettori, la sensazione avviene a livello della mano), però è come se la forbice diventasse parte del nostro corpo.

Un aspetto affascinante è la possibilità di far diventare parte della nostra rappresentazione corporea delle cose che non sono parte del nostro corpo, **ma lo diventano dal punto di vista funzionale**. Tutto questo ha delle applicazioni ad esempio sulla riabilitazione dei portatori di protesi, che devono reimparare ad utilizzare la protesi come se fosse parte del loro corpo.

## Applicazioni



Graziano & Gross (1993)

Questo discorso lo possiamo cominciare facendo riferimento a certe ricerche classiche di **Charlie Gross** che è stato il primo a mostrare che esistono unità neurali nella corteccia parietale della scimmia con **campo recettivo multisensoriale visuo-tattile**. Queste cellule scaricano sia quando viene presentato uno stimolo visivo in una certa zona dello spazio esterno, sia quando un certo distretto corporeo

viene stimolato, quindi hanno una risposta sia visiva che tattile.

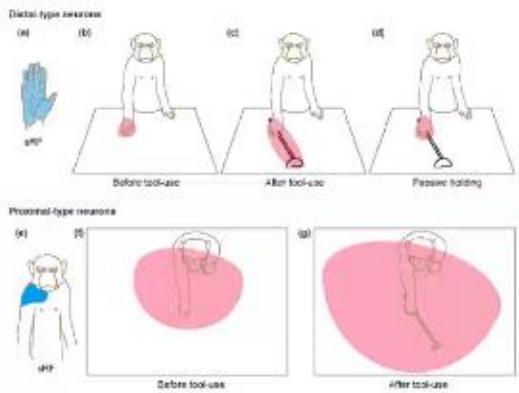
Ma la cosa più interessante è il rapporto tra queste due componenti del campo recettivo perché **il campo recettivo visivo non è retinotopico, ma è ancorato al corpo della scimmia**. Se io metto la mano della scimmia in un certo punto dell'ambiente e poi la stimolo con un evento meccanico (tattile), l'unità neurale scarica e fa lo stesso se presento uno stimolo visivo in quello stesso punto (campo recettivo visivo e tattile sono riferiti al campo visivo e alla mappa somatotopica, riguardano la mano da un lato e una certa zona del campo visivo dall'altro)

Se io sposto la mano della scimmia come nella figura B, a questo punto la cellula non risponde più quando presento uno stimolo visivo nella posizione A, che è quella che prima produceva una risposta, ma adesso risponde se lo stimolo visivo viene presentato nella posizione B. *In un certo senso il campo recettivo tattile si trascina dietro il campo recettivo visivo*. Entrambe le posizioni A e B diventano totalmente inefficaci cioè non producono una risposta nella cellula se la mano viene posta in una posizione molto lontana da quella posizione nello spazio (c). **Quindi è proprio un campo recettivo tattile-visivo, ma ancorato al corpo.**

Quindi spostando il corpo si sposta anche il campo visivo:

- a. Determino il campo visivo-tattile;
- b. Sposto la mano e il campo tattile si sposta quello visivo;
- c. Se la mano viene messa molto lontano il neurone non scarica.

Rappresentazioni visive ancorate sul corpo, quindi **campi recettivi visivi somatocentrici**.



Questo è un altro risultato classico portato a termine da **Hiriki**: se si prova a studiare questi campi recettivi visuo-tattili utilizzando un paradigma in cui la scimmia deve **imparare ad utilizzare uno strumento**, questo utilizzo dello strumento può modificare la forma del campo recettivo visuo-tattile. Si prende in considerazione un campo recettivo che riguarda la zona vista intorno alla mano e il tatto sulla mano. Si prendono le misure iniziali. Successivamente si addestra la scimmia a utilizzare un rastrello (per esempio da croupier) per poter raggiungere degli oggetti che sono più lontani della sua capacità di allungare il braccio (di solito per procurarsi una ricompensa). Dopo questo addestramento con lo strumento, il campo recettivo visuo-tattile si modifica, cioè si allunga includendo anche la parte del rastrello. Questo non avviene se la scimmia tiene semplicemente il rastrello in mano, ma non fa l'**attività intenzionale** di addestrarsi ad utilizzare lo strumento. La stessa cosa avviene con altri tipi di neuroni visuo-tattili, ad esempio quelli che si trovano sulla spalla, il cui campo recettivo visivo include lo spazio peripersonale davanti alla scimmia che si amplifica nel momento in cui la scimmia impara ad utilizzare lo strumento.

**In sostanza l'esperimento di Hiriki ci dice che a seguito di un allenamento con un oggetto viene aumentata la propriocezione di se stessi perché l'oggetto viene interiorizzato alla propria rappresentazione corporea.**

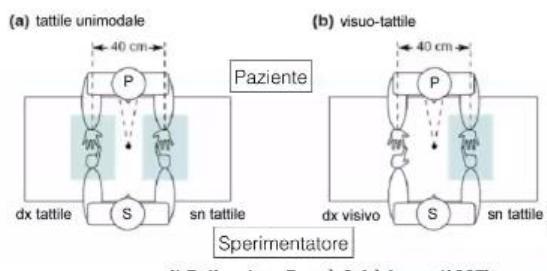
## Estinzione visuotattile nei pazienti neglect

**Neglect:** non si riesce a prestare attenzione ad una parte di emicampo.

Il fenomeno dell'estinzione visuotattile si può sfruttare per vedere se anche nell'uomo c'è questa riorganizzazione plastica come avviene nelle scimmie di Hiriki.

Noi sappiamo che questo fenomeno non si limita solo alla scimmia ma si verifica anche nell'uomo. Lo sappiamo grazie agli esperimenti di **Elisabetta Ladavas** e il suo gruppo di Bologna, che ha avuto un'idea molto elegante, cioè utilizzare un fenomeno che si trova in pazienti affetti da neglect che si chiama **estinzione**.

## estinzione visuotattile

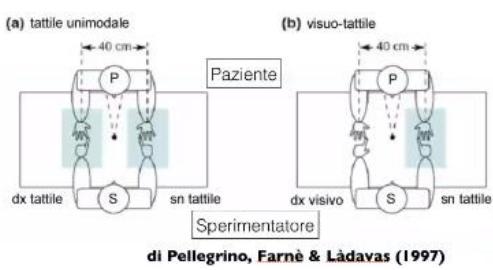


di Pellegrino, Farnè & Ladavas (1997)

**Il neglect è quella sindrome neurologica per cui le persone non riescono a prestare attenzione ad un emicampo.** L'estinzione è una condizione simile, ma meno grave del neglect. Quello che succede è che il paziente fa fatica a prestare attenzione, ma lo fa normalmente anche nell'emicampo controlesionale. Se, però, ha due stimoli, uno nell'emicampo ipsilaterale alla lesionale ed uno nell'emicampo controlesionale, quello nell'emicampo buono tende ad estinguere, cioè a rendere difficile prestare attenzione all'altro. Quindi, con una luce a destra e una a sinistra, la luce a destra fa sì che la luce a sinistra venga negletta (estinzione visiva).

Esiste anche un'**estinzione visuo-tattile** (questo può succedere infatti anche con le mani) nel quale lo stimolo alla mano destra estingue lo stimolo alla mano sinistra. Può anche avvenire in un *contesto multisensoriale* in cui lo sperimentatore nasconde la mano destra del soggetto con uno schermo e somministra uno stimolo tattile alla mano sinistra e contemporaneamente fa vedere che qualcosa si avvicina alla mano destra del soggetto.

### estinzione visuotattile



All'inizio hanno fatto il classico esperimento di **estinzione cross-modale**.

### estinzione cross-modale visuotattile

(di Pellegrino et al, 1997, Nature)



**L'estinzione è quel fenomeno per cui lo stimolo visivo tende ad estinguere lo stimolo tattile controlesionale.** Di Pellegrino, Farnet e Làdavas hanno pensato che questo fenomeno dell'estinzione visuotattile si potesse essere sfruttato per andare a vedere se, anche nell'uomo, sia presente questo fenomeno di riorganizzazione plastica dei campi recettivi come accade nella scimmia negli esperimenti di Hiriki.

1. Il paziente ha la mano sinistra coperta (che è quella soggetta ad estinzione), lo sperimentatore tocca la mano sinistra, contemporaneamente avvicina il dito alla mano destra, quindi c'è lo stimolo visivo a destra che estingue lo stimolo tattile a sinistra. Tocco della mano sinistra e avvicinamento della mano destra.

→ Estinzione

2. Dopo si fa il test cruciale (sopra): si fa la stimolazione tattile alla mano sinistra che è coperta da uno schermo e si presenta uno stimolo visivo che però non è vicino alla mano ma distante. In questo caso non c'è l'estinzione perché questo tipo di fenomeno chiama in causa delle unità neurali che hanno dei campi recettivi come quelli che ho descritto prima, che sono ancorati sulla mano e quindi lontano dalla mano non funziona.



→ **No estinzione visuo tattile;**



3. Un'altra condizione sperimentale in cui il soggetto impara ad utilizzare un rastrello (in questo caso è un mestolo, basta che serva per raggiungere oggetti che non riusciamo a raggiungere senza). Dopo l'addestramento, al soggetto viene fatto di nuovo un esperimento di estinzione visuo-tattile ma stavolta lo stimolo visivo non viene presentato solo in prossimità della mano destra ma viene presentato anche in prossimità della punta dello strumento. Succede che, di nuovo, c'è estinzione visuo-tattile, anche nella posizione lontana dal corpo che prima non produceva l'estinzione visuo-tattile. Si estingue anche se è lontano perché c'è stato un'inclusione dell'oggetto nella rappresentazione del braccio.

→ **Estinzione.**