

ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA
SCUOLA DI INGEGNERIA E ARCHITETTURA
CORSO DI LAUREA IN INGEGNERIA BIOMEDICA

MODELLI DI ANALISI
PER
L'INTEGRAZIONE MULTISENSORIALE

Tesi in
Strumentazione Biomedica

Relatore
Cristiano Cuppini

Presentata da
Sofia Poma

Sessione Terza
Anno Accademico 2015 - 2016

*Dovrei chiedere scusa a me stessa
per tutte le volte in cui
ho creduto
di non essere abbastanza
A. Merini*

Indice

Introduzione.....	7
1. LE PERCEZIONI SENSORIALI.....	11
1.1 Cos'è la percezione sensoriale?.....	11
1.2 Proprietà dello stimolo.....	12
1.3 Proprietà generali dei recettori.....	14
1.4 Come viene processato lo stimolo: bottom-up e top-down.....	16
1.5 Le vie sensoriali.....	18
1.5.1 La percezione uditiva.....	18
1.5.2 La percezione visiva.....	19
1.5.3 La propriocezione.....	20
1.6 Le illusioni unisensoriali.....	22
2. L'INTEGRAZIONE SENSORIALE E IL TEST DELLA SOUND-INDUCED FLASH ILLUSION.....	25
2.1 L'integrazione sensoriale.....	25
2.1.1 Aspetti generali.....	25
2.1.2 Anatomia.....	27
2.1.3 I principi alla base del fenomeno dell'integrazione.....	28
2.2 Il test della sound-induced flash illusion.....	31
2.2.1 Esperimento 1.....	32
2.2.2 Esperimento 2.....	34
2.2.3 Conclusioni.....	36
2.3 L'integrazione sensoriale deficitaria: l'autismo.....	37
2.3.1 Cos'è l'autismo?.....	38
2.3.2 La percezione nell'autismo.....	41
2.3.3 L'integrazione audio-visiva.....	45
2.4 Il test della sound-induced flash illusion su soggetti autistici.....	49
2.4.1 L'esperimento.....	50
2.4.2 Conclusioni.....	51
2.4.3 L'aspetto temporale.....	52
3. LE BODY OWNERSHIP ILLUSIONS.....	55
3.1 La somatoparafenìa e le illusioni di appartenenza.....	55
3.2 Le BOIs nello studio della percezione del proprio corpo.....	61
3.3 Trigger multimodali nelle BOIs.....	63

3.3.1 Stimoli visuo-tattili.....	64
3.3.2 Stimoli visuo-motori.....	65
3.3.3 Stimoli visuo-proprioceettivi.....	67
3.4 La somatoparafenria e i test sulle BOIs.....	69
4. I MODELLI DI ANALISI DEI SEGNALI SENSORIALI.....	71
4.1 Il modello bayesiano.....	71
4.1.1 Il Teorema di Bayes.....	72
4.1.2 I modelli bayesiani applicati alla percezione multisensoriale.....	73
4.1.3 I modelli bayesiani applicati alle BOIs.....	77
4.2 Neuroimaging ed elettroencefalografia.....	81
4.2.1 L'elettroencefalografia.....	82
4.2.2 Applicazione del modello EEG all'integrazione multisensoriale.....	83
4.3 Le reti neurali.....	89
4.3.1 I modelli di reti feedforward.....	91
4.3.2 I modelli a feedback.....	92
Conclusioni.....	99
Bibliografia.....	103
Ringraziamenti.....	107

Introduzione

La percezione unitaria della realtà è il risultato di un complesso processo d'integrazione delle informazioni provenienti da differenti canali sensoriali. La capacità del sistema nervoso centrale di utilizzare sinergicamente queste molteplici sorgenti sensoriali è definita Integrazione Multisensoriale (Stein & Meredith, 1993). Essa agisce a diversi livelli del sistema nervoso: nelle strutture subcorticali (es. il Collicolo Superiore), a elevati livelli delle corteccie associative (es. le regioni parietali posteriori) e anche nelle regioni corticali primarie (come la corteccia visiva primaria), in passato ritenute puramente unisensoriali. Come conseguenza della sua eccezionale importanza sul comportamento, lo studio dei processi che regolano l'integrazione multisensoriale è da tempo riconosciuto come un aspetto cruciale nell'ambito delle neuroscienze (Calvert, Spence, e Stein, 2004; Stein & Meredith, 1993). Numerose evidenze sperimentali hanno infatti contribuito ad individuare i processi cognitivi e le aree cerebrali responsabili dei fenomeni integrativi; tuttavia molti aspetti sui possibili meccanismi coinvolti restano ancora da chiarire. Da questo punto di vista, un importante contributo può venire dallo confronto tra soggetti sani e soggetti con deficit cerebrali, quali lesioni, alterazioni nelle risposte o nel comportamento, al fine di chiarire quali aree cerebrali sono coinvolte e per poter quindi far luce sui meccanismi neurali sottostanti. Un metodo utile di indagine sui fenomeni di interazione multisensoriale sfrutta le illusioni sensoriali, ovvero quelle situazioni in cui due stimoli di natura sensoriale differente vengono presentati più o meno in contemporanea, creando una falsa interpretazione dell'oggetto o dell'evento da cui provengono tali stimoli.

Il presente lavoro esamina una delle forme di integrazione più studiate, ovvero quella audio-visiva, in particolare, la *sound-induced flash illusion* (Shams et al., 2002): quando un singolo flash è accompagnato da molteplici segnali acustici (beep), il singolo flash è percepito come multiplo. L'aspetto interessante di questa illusione è che è stata riscontrata anche in soggetti patologici affetti da autismo (DSA, Disturbi dello Spettro Autistico), un disordine dello sviluppo caratterizzato da alterazione o mancanza totale di interazione sociale e comunicazione. Date le sue caratteristiche, per molto tempo si è pensato che i soggetti autistici potessero presentare anche deficit neurologici alle connessioni inter-compartimentali, alla base, per l'appunto, dell'integrazione multisensoriale. Grazie a studi

recenti (van der Smagt et al., 2007) hanno dimostrato che almeno alcuni collegamenti tra le aree cerebrali uditive e quelle visive sono funzionanti, poiché i DSA percepiscono l'illusione e, anzi, presentano una finestra temporale di integrazione maggiore (temporal binding window), all'interno della quale essi sono portati a processare gli stimoli multisensoriali in maniera particolare (Foss-Feig et al., 2010).

Un secondo interessante fenomeno presentato da questa tesi è quello delle *body ownership illusions* (BOIs), utili per indagare patologie quali la somatoparafrenia, una credenza delirante per cui un paziente sente che un suo arto paralizzato non appartiene realmente al suo corpo. La consapevolezza dello stato del sistema motorio e il senso di essere la causa di un'azione sono aspetti cruciali del nostro senso di "essere noi stessi". In circostanze normali, queste esperienze sono strettamente legate, se non addirittura inseparabili; in circostanze deficitarie ciò avviene in modo diverso ed è estremamente difficile capire quali meccanismi la governano. In queste illusioni, i soggetti hanno percepito oggetti non-corporei (ad es. elementi artificiali) come fossero appartenenti al proprio corpo, quando essi sono stati presentati con gli stimoli cross-modalità applicati alla parte reale ma nascosta del corpo e, più o meno in contemporanea, alla sua controparte falsa. In questo modo, le BOIs offrono un potente strumento sperimentale per esaminare il modo in cui il senso di appartenenza al corpo emerge dall'elaborazione multisensoriale operata dal cervello.

Nonostante questi problemi complessi e queste differenze, si possono delineare alcune regole generali in materia di integrazione multisensoriale e alcuni principi per guidare gli sforzi di modellazione. In particolare, i risultati che sono stati raccolti negli ultimi due decenni hanno ispirato alcune considerazioni teoriche precise e hanno permesso la formulazione di alcuni modelli computazionali che possono essere testati con i dati esistenti. Nel presente elaborato vengono presentati alcuni dei più utilizzati modelli di analisi dei dati sensoriali.

I modelli bayesiani permettono di applicare la teoria della probabilità alla percezione e alla combinazione di stimoli dei sistemi biologici (Liu et al., 1995). Essi presuppongono che il cervello esegua un'integrazione ottimale, calcolando le probabilità a posteriori dell'evento, dato un insieme di informazioni incerte.

Le immagini funzionali (Neuroimaging) hanno come obiettivo principale la comprensione delle funzioni specifiche delle diverse aree cerebrali e come queste interagiscono in un

sistema integrato di reti neuronali. In particolare, l'elettroencefalografia (EEG) misura il potenziale elettrico delle oscillazioni derivanti dalla corrente ionica nei neuroni, raccogliendo informazioni sull'attività dello strato superficiale della corteccia cerebrale. Di rilevante interesse è l'analisi che può essere svolta, grazie a questa tecnica, dei potenziali evento-correlati (ERP), che consistono in un'elaborazione complessa degli stimoli sensoriali e dipendono dal contesto psicologico nel quale avviene la stimolazione (Regan, 1989).

Gli approcci appena descritti hanno contribuito a fornire un'enorme mole di dati sulla caratterizzazione delle proprietà dell'integrazione multisensoriale e a identificare le aree cerebrali maggiormente coinvolte in questo fenomeno. Tuttavia la comprensione dei meccanismi neurali che coinvolge il sistema nervoso è ancora insufficiente. Questa limitazione potrebbe essere in parte imputata alla complessità dei meccanismi coinvolti; infatti l'integrazione multisensoriale è dovuta a popolazioni di neuroni interconnessi, nelle quali diversi fattori come le caratteristiche dei singoli neuroni, le disposizioni delle connessioni e la topologia della rete contribuiscono a determinare i fenomeni osservati. Come in molti altri campi delle neuroscienze, modelli di reti neurali, ispirate alla neurofisiologia, possono svolgere un ruolo fondamentale per migliorare la nostra comprensione dei complessi meccanismi coinvolti nella interazione sensoriale, e per formulare nuove previsioni verificabili per confermare o respingere le ipotesi. La presente tesi si è concentrata su un modello di rete, implementato di recente, che coinvolge l'integrazione audio-visiva, e il test sulla sound-induced flash illusion (Cuppini et al., 2014). Il modello sviluppato simula due catene neuronali, una visiva e una acustica, e le loro reciproche connessioni, e, attraverso una equazione del primo ordine, è stata in grado di simulare diversi aspetti cruciali delle moderne teorie sull'integrazione multisensoriale, confermando le leggi integrative fondamentali che regolano il funzionamento dei neuroni multisensoriali (legge spaziale, temporale e dell'efficacia inversa).

1. LE PERCEZIONI SENSORIALI

1.1 Cos'è la percezione sensoriale?

Una delle funzioni principali della mente umana è quella di conoscere ed entrare in rapporto con l'ambiente circostante, e per fare ciò ricorre agli organi di senso, che gli permettono di percepire la realtà.

È possibile operare una prima distinzione tra la sensazione e la *percezione*.

La sensazione è un concetto più fisico, in quanto essa è legata agli effetti immediati del contatto dei recettori sensoriali con i segnali provenienti dall'esterno. Questi segnali sono in grado di suscitare una risposta sia nel sistema sensoriale coinvolto, sia sull'intero corpo.

La percezione, invece, è un concetto più psichico, perché corrisponde all'organizzazione e all'elaborazione dei dati sensoriali in un'esperienza unitaria più complessa. Con la

percezione assorbiamo informazioni dal mondo esterno, le coordiniamo a livello sia conscio che inconscio, e infine le utilizziamo per prendere decisioni e per concepire

intenzioni in grado di produrre le azioni che ci fanno affrontare la nostra vita quotidiana.

È evidente, quindi, come sia inevitabile una certa dose di soggettività nella percezione, a volte anche molto marcata. Lo stesso stimolo può avere una rappresentazione mentale diversa e tempistiche diverse per ognuno di noi; basti pensare a quella che abitualmente è definita “sensibilità”, per cui percepiamo un suono più o meno forte, o un colore più o meno intenso, o abbiamo una soglia di percezione del dolore diversa.

Inoltre, lo stimolo può essere processato anche in base ai nostri stati mentali. Il modo in cui la nostra mente percepisce uno stimolo in un momento di depressione o di malinconia sarà sicuramente diverso dal modo in cui si percepirà lo stesso stimolo se lo stato emotivo è di euforia. Un discorso simile può essere fatto pensando alle alterazioni psichiche dovute a malattie nervose o, più banalmente, all'abuso di alcol o droghe. Lo stimolo è quindi modulato in base allo stato psico-fisico del momento.

Inoltre, la percezione dipende non solo dagli stimoli nel presente, elaborati dai recettori sensitivi, ma anche dalle esperienze precedenti a cui ogni individuo è sottoposto. Tra gli esempi più banali, si può citare il risveglio dei ricordi associati a un particolare odore o al sapore di un particolare cibo.

Se guardiamo il concetto dal punto di vista fisiologico, possiamo definire la sensazione come un cambiamento di stato del nostro sistema neurologico causato dal contatto con l'ambiente esterno, che avviene grazie all'intervento degli organi di senso.

Il sistema sensoriale monitorizza l'ambiente interno e quello esterno e successivamente invia informazioni ai centri integratori situati nel sistema nervoso centrale. Essi, a loro volta, danno origine alle risposte più appropriate. I riflessi più semplici possono essere integrati direttamente nel midollo spinale, senza ricorrere all'influenza dei centri superiori. La maggior parte delle informazioni sensoriali, però, prosegue fino alla corteccia cerebrale, in cui si distinguono cinque principali aree sensoriali.

I canali sensoriali sono: udito, vista, olfatto, gusto, tatto-pressione, propriocezione (consapevolezza della posizione del proprio corpo nello spazio e dello stato di contrazione dei propri muscoli), termocezione (capacità di percepire la temperatura e i suoi cambiamenti), equilibrio e sensazione di dolore. Ognuno di essi si avvale per la ricezione dei segnali di specifici organi di senso.

Generalmente, ogni variazione del mondo fisico viene percepita da tutti in una maniera tale che la descrizione di tale variazione risulta molto simile, nonostante ci sia una buona parte di soggettività.

1.2 Proprietà dello stimolo

I recettori sono unità fisiche microscopiche costituite da terminazioni nervose, caratterizzate da un'anatomia diversa a seconda del ruolo svolto e quindi del fine predisposto per il recettore stesso.

Indipendentemente dalla sua natura e dai meccanismi implicati, il recettore svolge la propria funzione, trasducendo la forma d'energia propria dello stimolo (sia essa meccanica, chimica, luminosa, etc.) in un segnale nervoso trasmissibile ai centri superiori.

In genere, un recettore risponde a un solo tipo particolare di stimolo, che viene quindi denominato “stimolo adeguato”. Per esempio, se si stimolano i recettori della retina, è necessario che lo stimolo sia di tipo luminoso. Nonostante ciò, si può ottenere una risposta di tipo luminoso anche se lo stimolo è non adeguato, per esempio meccanico: se si preme sull'occhio, si vede comunque qualcosa che ricorda delle immagini visive.

Una volta che il segnale elettrico è stato inviato al Sistema Nervoso Centrale (SNC), i

centri nervosi preposti all'analisi sensoriale prevedono un'organizzazione complessa ma molto specifica. Assistiamo infatti ad un'organizzazione in serie, per cui man mano che l'informazione procede nel suo percorso all'interno del SNC è sottoposta ad un'elaborazione sempre più complessa. Contemporaneamente, abbiamo anche un'organizzazione in parallelo, per cui le informazioni concernenti le varie proprietà dello stimolo viaggiano divise per gran parte del loro percorso, per essere poi sintetizzate solo ai livelli più alti, dove la sorgente di stimolazione viene finalmente percepita come un tutt'uno.

Ora analizzeremo le principali proprietà dello stimolo, poiché esse caratterizzano il primo gradino della modulazione della percezione; infatti, a seconda della qualità di ogni singola proprietà, lo stimolo produce un effetto diverso sulla stimolazione nervosa, e quindi avremo un risultato diverso sulla percezione finale dello stimolo stesso.

La *natura* (o *modalità*) dello stimolo è individuata, più che dal recettore, dalle aree corticali attivate.

Un impulso energetico diventa stimolo solo quando è capace di eccitare un recettore e generare un potenziale d'azione nelle fibre nervose afferenti al SNC. In altre parole, lo stimolo è definito dalla sua percepibilità a livello nervoso/cerebrale. Si distinguono, quindi, cinque modalità differenti: visiva, acustica, olfattiva, gustativa e somatica; se le prime quattro si riferiscono chiaramente a un singolo canale sensoriale, l'ultima è valida per tatto, temperatura, dolore e propriocettività.

Il perché e come determinate aree corticali riescano ad evocare specifiche sensazioni sensoriali è ancora sconosciuto e oggetto di dibattito.

Altro parametro fondamentale dello stimolo è la *localizzazione* nello spazio da cui lo stimolo proviene. Essa è codificata in base al campo recettivo che viene attivato.

I campi recettivi sono porzioni di spazio, corporeo o extra-personale, in cui l'applicazione di uno stimolo di qualità, intensità e durata sufficienti, determina l'eccitazione di una struttura nervosa. Tutta l'organizzazione del SNC si basa su questo concetto. Dalle radici del midollo spinale alle aree corticali, i neuroni sono altamente organizzati in funzione della provenienza dei segnali; se essi codificano porzioni di spazio adiacenti occupano posizioni adiacenti del tessuto nervoso.

Terzo parametro fondamentale dello stimolo è la sua *durata*, indicata dal tempo di scarica

dei recettori e delle strutture ad essi collegate. A tal proposito è cruciale una specifica proprietà delle cellule nervose: l'adattamento. Con questo termine si indica quel fenomeno per cui i neuroni, dopo un certo tempo dalla loro eccitazione, ritornano ad essere silenti o, comunque, ad un livello di attività talmente bassa da poter definirli in uno stato di quiescenza. Questa proprietà è particolarmente spiccata a livello dei recettori, tra cui si distinguono quelli a rapido adattamento (detti recettori fasici), che si attivano solo in corrispondenza dell'inizio della sollecitazione e smettono di rispondere se l'intensità dello stimolo resta costante, e recettori a lento adattamento (detti recettori tonici), attivi per tutta la durata d'applicazione dello stimolo, anche nel caso in cui la frequenza diminuisca. Se ne deduce che i recettori a rapido adattamento sono particolarmente efficaci nel segnalare le proprietà temporali dello stimolo, mentre quelli a lento adattamento sono indicati per l'analisi delle caratteristiche spaziali, ovvero all'identificazione della forma dello stimolo. L'ultimo, ma non certo meno importante, parametro è la sua *intensità*: pur di natura idonea, uno stimolo non induce risposta se al di sotto di una certa intensità, perciò questa soglia è definita sperimentalmente come intensità che evoca risposta nel 50% dei casi. L'intensità è codificata dal SNC in due modi: in termini di frequenza di scarica del recettore e dei neuroni ad esso collegati (codifica in frequenza), oppure con il numero di neuroni attivati (codifica di popolazione). All'aumentare dell'intensità dello stimolo, aumenta in proporzione l'ampiezza del potenziale recettoriale e aumenta anche la frequenza dei potenziali d'azione nel neurone sensoriale primario.

1.3 Proprietà generali dei recettori

Volendo analizzare i meccanismi alla base dell'esperienza sensoriale, dobbiamo innanzitutto riconoscere che essa non costituisce un fenomeno semplice ed unitario, ma rappresenta il culmine di una serie di processi che, partendo dalla stimolazione a livello dei recettori, si concludono a livello corticale, realizzando l'esperienza percettiva vera e propria.

Sebbene differenti sotto numerosi aspetti, tutte le vie sensoriali contengono diversi elementi in comune.

I sistemi sensoriali sono concepiti per analizzare gli stimoli, ovvero delle forme di energia

agenti sui recettori sensoriali, e le loro proprietà fondamentali (qualità, intensità, durata, posizione nello spazio), per poi sintetizzarle in una rappresentazione coerente e unitaria, che permetta, infine, una precisa conoscenza dello stimolo stesso.

Elemento cardine, nonché primo motore del processo percettivo, è il recettore, ovvero un trasduttore che converte lo stimolo, spesso di natura meccanica o chimica, in un segnale intracellulare elettrico, di solito una variazione locale del potenziale di membrana, che viene inviato ai centri nervosi specializzati.

Ogni recettore permette di percepire solo alcuni eventi fisici e solo se si verificano entro un certo range di ampiezza o di frequenza. Per esempio, i recettori della vista sono sensibili alla luce mentre quelli dell'udito percepiscono le vibrazioni delle onde sonore sul timpano. Per ogni recettore sensoriale è possibile stabilire i limiti inferiore e superiore entro i quali l'evento fisico viene percepito. Se parliamo di soglie, è importante sottolineare la differenza tra quella assoluta e quella differenziale (S. Stevens, 1950).

Con il termine *soglia assoluta* si definisce quel particolare livello d'energia che discrimina tra la percezione di un determinato evento fisico e l'assenza di essa; ad un livello inferiore a tale soglia assoluta, il soggetto non percepisce lo stimolo, mentre ad un livello pari o superiore, il soggetto lo percepisce. Prendendo i soliti due sensi come esempio, per la vista la soglia assoluta è rappresentata dalla luce di una candela a 48 chilometri di distanza in una notte limpida, mentre per l'udito da un orologio meccanico a 6 metri di distanza all'interno di una stanza silenziosa.

Con il termine *soglia differenziale*, invece, si indica la differenza minima di intensità che uno stimolo deve avere rispetto a un secondo stimolo affinché vengano percepiti come diversi. Sin dai primi studi, è stato chiaro che la nostra sensibilità è maggiore alle basse intensità rispetto che a quelle alte; basti pensare che riusciamo a distinguere la differenza di rumore tra un brusio e un aspirapolvere (tra i 60 e gli 80 dB), piuttosto che quello di un martello pneumatico rispetto alla musica da discoteca (tra i 100 e i 120 dB).

Per quel che concerne l'ambito della fisiologia, i recettori variano moltissimo nel loro grado di complessità, e possono essere classificati in base alla modalità di trasmissione. I recettori più semplici sono costituiti da neuroni con terminazioni periferiche libere. Essi sono detti di I tipo, e ne sono un esempio i recettori somatosensoriali. Il potenziale generatore ha origine direttamente da queste terminazioni e viene poi propagato

elettronicamente. Se superiore ad una certa ampiezza soglia, determina la comparsa del potenziale d'azione nella fibra sensoriale.

I sensi speciali (vista, udito, gusto, olfatto) sono dotati di recettori più specializzati, detti di II e III tipo. Da essi origina il potenziale dei recettori che determina la liberazione di un mediatore chimico, il quale induce la comparsa del potenziale d'azione.

1.4 Come viene processato lo stimolo: bottom-up e top-down

Un importante problema teorico che divide i neuroscienziati è la misura in cui la percezione si basa direttamente sulle informazioni presenti all'interno dello stimolo. Ciò è importante per capire in che modo lo stimolo viene processato all'interno del SNC, al fine di produrre una risposta adeguata.

Alcuni sostengono che i processi percettivi non sono diretti, ma dipendono dalle aspettative e dalle esperienze precedenti di ogni individuo, mentre altri credono dipendano esclusivamente dalle informazioni disponibili nello stimolo stesso.

Tra gli esempi più calzanti, si può citare la differenza di visione sull'argomento tra Gibson (1966), che ha proposto una teoria diretta della percezione, detta "bottom-up", e Gregory (1970), che ha invece proposto una teoria costruttivista (indiretta) sulla percezione, denominata "top-down".

Da quel momento, gli psicologi e i neuroscienziati hanno distinto i processi percettivi in questo modo.

L'elaborazione bottom-up è anche conosciuta come "data-driven processing" (trattamento guidato dei dati), perché la percezione comincia dallo stimolo stesso. Il processo è diretto dalla retina alla corteccia visiva, e ogni singolo stadio successivo nel percorso percettivo visivo porta a un'analisi sempre più complessa dell'input.

L'elaborazione top-down si riferisce all'uso delle informazioni contestuali per il riconoscimento dei modelli. Per esempio, la comprensione della scrittura di una persona risulta più semplice se vengono lette intere frasi piuttosto che singole parole. Questo perché il significato delle parole circostanti fornisce il contesto ai fini di una migliore comprensione.

Entrambi i processi, però, sono inevitabilmente soggetti a criticità, che dimostrano come nessuno dei due può rappresentare una spiegazione univoca per la modalità in cui avvengono le percezioni sensoriali.

Una semplice critica che può essere fatta all'elaborazione top-down è la seguente: se la nostra percezione dipende dalle esperienze pregresse, come avviene la percezione nei neonati? È evidente che ci dev'essere una spiegazione ulteriore e meno netta su come il nostro cervello elabora le informazioni sensoriali.

Probabilmente, l'aspetto più critico della visione dei costruttivisti è da ricercare nel fatto che essi hanno sottovalutato la ricchezza di evidenze sensoriali disponibili nel mondo reale (in contrapposizione al laboratorio, luogo in cui la maggior parte delle teorie costruttiviste sono nate). Essi utilizzano spesso l'esempio della costanza della dimensione, ovvero che noi percepiamo correttamente la dimensione di un oggetto anche se l'immagine retinica si restringe quando l'oggetto si allontana. Essi affermano che le evidenze sensoriali provenienti da altre fonti devono essere disponibili per permetterci di fare ciò. In ogni caso, le immagini retiniche sono difficilmente disponibili nel mondo reale (e non in laboratorio). Vi è una ricca gamma di informazioni sensoriali, compresi gli oggetti, il background, l'orizzonte e il movimento.

Questa moltitudine di informazioni è importante per l'approccio di Gibson, il bottom-up.

Gibson, infatti, criticò fortemente le teorie di Gregory, in quanto basate su esempi artificiali e non sulla realtà dei nostri ambienti visivi normali/reali.

La percezione diretta potrebbe spiegare la percezione dell'ambiente, che avviene solitamente in maniera rapida e precisa. Per contro, questa teoria non può spiegare perché le percezioni sono talvolta imprecise, come nel caso delle illusioni. Ad esempio, la teoria di Gibson non può tenere conto di errori percettivi, come la tendenza generale delle persone a sopravvalutare le dimensioni in senso verticale piuttosto che in senso orizzontale, né delle illusioni naturali, come quando dopo qualche tempo che si fissa il flusso di una cascata e si trasferisce lo sguardo su un oggetto statico, esso sembrerà muoversi in direzione opposta.

Né il processo diretto né quello costruttivista sembrano essere in grado di spiegare tutti i generi di percezioni che ogni singolo individuo sperimenta nel corso della propria vita. Per

questo, un'ulteriore teoria, denominata "ciclo percettivo", è stata proposta da Neisser nel 1976 e cerca di spiegare come i due processi possano essere fusi insieme per produrre una migliore interpretazione dello stimolo.

1.5 Le vie sensoriali

Come già detto precedentemente, i principali canali sensoriali sono: udito, vista, olfatto, gusto, tatto-pressione, propriocezione, termocezione, equilibrio e sensazione di dolore, e ognuno di essi si avvale di specifici organi di senso.

Nel seguito del capitolo, analizzeremo i sistemi utili alla trattazione del nostro argomento, al fine di descriverne il funzionamento e di specificarne le caratteristiche.

1.5.1 La percezione uditiva

Il suono consiste in onde pressorie condotte attraverso l'aria o l'acqua. Il sistema uditivo può essere definito come l'insieme delle strutture nervose che partecipano alla conversione del suono, percepito inizialmente come energia meccanica, in energia elettrica utilizzabile dal sistema nervoso per elaborare le sensazioni e le percezioni dell'udito.

Il suono è l'interpretazione delle onde sonore che arrivano al nostro orecchio attraverso l'analisi di frequenza (numero di cicli generati in un secondo), ampiezza (grandezza del profilo di variazione di pressione dell'onda sonora) e durata.

La frequenza delle onde sonore è misurata in hertz (Hz). L'orecchio umano medio è in grado di percepire suoni compresi nell'intervallo tra i 20 e i 20000 Hz, con la sensibilità maggiore tra i 1000 e i 3000 Hz. Il cervello trasduce la frequenza delle onde sonore nell'altezza del suono: le onde a bassa frequenza sono sentite come suoni bassi, per esempio un tuono in lontananza. Le onde ad alta frequenza, invece, sono interpretate come suoni acuti, come quelli provocati da un unghia che griffia la lavagna.

Un aspetto psicologico legato all'udito è il volume, ovvero l'interpretazione dell'intensità del suono, e dipende dalla sensibilità del singolo individuo. L'intensità è funzione dell'ampiezza dell'onda e viene misurata in decibel (dB) su scala logaritmica. Una conversazione a un volume normale vale circa 60 dB, mentre il volume di un concerto può arrivare anche a 120 dB.

Il modello del sistema uditivo centrale più utilizzato negli ultimi anni è costituito da un modello a rete (network model), secondo cui l'elaborazione dell'informazione non avviene solo in specifici centri del cervello, ma in un gran numero di centri interconnessi e diffusi all'interno di tutto il Sistema Nervoso Centrale (Masterton, 1992). Secondo tale modello, le risposte percettive agli stimoli sensoriali, di qualsiasi natura essi siano, sono mediate da numerose regioni del cervello attraverso un fitto sistema di reti neurali disposte in serie e in parallelo (Ungerleider, 1995).

La combinazione di queste reti opera in maniera differente a seconda della tipologia di informazione da processare, in modalità ascendente (bottom-up) o discendente (top-down). Entrambe le modalità operano un feed-back in misura variabile ma continuo, a seconda delle indicazioni e delle correzioni fornite dai diversi sensi.

1.5.2 La percezione visiva

Il sistema visivo può essere definito come l'insieme delle strutture nervose che partecipano alla conversione della luce nelle sensazioni e nelle percezioni della visione.

In maniera banale, per spiegare come avviene la percezione visiva, possiamo paragonare l'occhio a una macchina fotografica. L'occhio fa arrivare la luce su una superficie sensibile ai raggi luminosi (la retina) per mezzo di una lente (il cristallino) e un'apertura (la pupilla), il cui diametro può essere modificato per modulare l'intensità della luce in ingresso.

La visione è il processo attraverso cui la luce riflessa dall'ambiente viene trasformata in immagine mentale, e può essere suddivisa in tre fasi principali:

- la luce entra nell'occhio e viene messa a fuoco sulla retina dal cristallino
 - i fotorecettori della retina trasducono il segnale luminoso in segnale elettrico
 - i segnali elettrici vengono inviati attraverso le vie nervose al cervello per essere elaborate.
- Così come l'orecchio umano non percepisce tutte le frequenze delle onde sonore, i nostri occhi non intercettano tutte le frequenze dello spettro elettromagnetico. Non siamo sensibili neanche a tutte le frequenze della luce solare; le onde elettromagnetiche più lunghe di 700 nanometri e quelle più corte di 400 nm sono invisibili all'uomo.

Come abbiamo già detto per l'udito, anche la vista può essere associata a un modello di rete neurale che opera un feed-back continuo.

Nell'osservazione visiva di una scena, infatti, il cervello risolve in continuazione ambiguità, verifica ipotesi e confronta le informazioni attuali con i ricordi e le aspettative.

Durante una percezione visiva, i dati grezzi arrivano al SNC sotto forma di unità dell'informazione visiva, o pixel. Attraverso una successione di aree visive, i pixel subiscono un'analisi sempre più sofisticata, fino al riconoscimento finale dell'oggetto.

Questo modello di visione, però, ignora le massicce proiezioni di feedback che ogni area visiva superiore rinvia alle inferiori.

Sarebbe più corretto, quindi, affermare che ad ogni stadio dell'elaborazione visiva si genera un'ipotesi parziale sui dati in arrivo (bottom-up) e che questa sia poi rinviata alle aree inferiori per imporre una piccola tendenza all'elaborazione successiva (top-down).

1.5.3 La propiocezione

La propiocezione (o sensibilità propriocettiva) è un meccanismo molto sofisticato che caratterizza il nostro sistema sensoriale. Esso ha lo scopo di fornire al Sistema Nervoso Centrale (SNC) informazioni della massima precisione, e in tempo reale, sia a proposito dei parametri del movimento biomeccanico (quindi velocità, forza, direzione, accelerazione), sia di parametri fisiologici, quali lo stato e il cambiamento biologico che si verifica nei muscoli, nei tendini e nelle articolazioni a causa di un movimento compiuto dal nostro corpo.

Sulla base di queste informazioni, il cervello forma un'immagine della posizione del corpo e delle sue parti ed è in grado di correggere o modificare il movimento in corso.

In altre parole, la propiocezione è fondamentale nel controllo del progetto motorio, sia per un meccanismo di controllo sulla corretta esecuzione del movimento, sia per un meccanismo di correzione di eventuali errori causati da fenomeni esterni, che vanno ad alterare il movimento programmato.

Si può quindi dire che la propiocezione è controllata da meccanismi a feedback negativo: l'azione eseguita da un sistema viene confrontata con l'azione programmata e qualsiasi differenza (errore) viene segnalata al sistema, in modo che questo attivi le opportune correzioni.

Tra i vari canali sensoriali, la propiocezione è senza ombra di dubbio uno dei più difficili

da studiare, in particolare perché la maggior parte delle informazioni non sono percepite a livello conscio, essendo responsabili del controllo dell'elaborazione del movimento e della sua esecuzione.

Al di là di questi numerosi compiti, il sistema propriocettivo, nel suo insieme, fornisce informazioni anche alle strutture nervose in grado di elaborare i processi di consapevolezza e coscienza. Esso cioè manda informazioni anche alla corteccia cerebrale, costituendo quella che viene definita "percezione propriocettiva cosciente".

La percezione propriocettiva cosciente è una costruzione elaborata dalla corteccia cerebrale sulla base delle informazioni provenienti dai recettori propriocettivi periferici.

A livello corticale, avviene un'integrazione complessa tra le afferenze provenienti dai canali sensoriali e le informazioni immagazzinate nella memoria grazie all'esperienza. La memoria, infatti, porta un bagaglio informativo sulle esperienze passate, che permettono ad ognuno di noi di filtrare le sensazioni provenienti dal mondo esterno, facendole proprie attribuendo loro dei valori personali.

Tutto ciò che è stato appena descritto porta alla formazione di quella che comunemente viene definita "immagine corporea", ossia la consapevolezza dell'esistenza, della posizione e del movimento del nostro corpo.

Alla base della propriocezione vi è un complesso sistema sensoriale. La propriocettività elabora le informazioni sensoriali grazie a particolari recettori periferici, detti, appunto, propriocettori, posizionati nelle articolazioni e nei tessuti circostanti. I propriocettori sono organi sensitivi specializzati, la cui funzione principale è quella di fornire informazioni in retroazione sia sulle posizioni assunte dal corpo in condizioni di quiete (propriocezione propriamente detta), sia sui parametri dinamici del movimento (cinestesia). Essi segnalano, istante per istante, quali siano i movimenti che il nostro corpo sta compiendo; proprio sulla base di queste informazioni i centri superiori sono, all'occorrenza, in grado di correggere o modificare il movimento in corso.

1.6 Le illusioni unisensoriali

“Un'illusione è una distorsione di una percezione sensoriale, causata dal modo in cui il cervello normalmente organizza ed interpreta le informazioni che riceve. Le illusioni possono coinvolgere tutti i sensi, ma quelle visive sono le più famose e conosciute, dal momento che la vista spesso prevarica gli altri sensi” (Wikipedia).

Cos'è un'illusione percettiva? Le illusioni sono fenomeni percettivi anomali in cui le informazioni derivanti da stimoli esterni e reali, di natura sensoriale diversa, portano a una falsa interpretazione dell'oggetto o dell'evento da cui proviene il suddetto stimolo. In sostanza, le illusioni percettive sono il risultato di interpretazioni errate di una serie di dati sensoriali. Essa è dovuta ad un errore che si verifica nella catena di elaborazione dell'informazione sensoriale in ingresso a carico del Sistema Nervoso Centrale (SNC). Abbiamo recettori sensoriali nel cervello che rilevano luce, suono, temperatura, etc, e la maggior parte di questi recettori sono sulla superficie del nostro corpo (occhi, orecchie, bocca, etc). Vi sono, tuttavia, recettori interni, come quelli che registrano il dolore. Indipendentemente dal tipo, ciascun recettore ha il compito di rilevare la forma di energia specifica, che viene poi convertita in segnale elettrico nervoso e trasmessa al SNC. Quando il cervello riceve questi segnali, il più delle volte interpreta i segnali correttamente, ma quando non è così, sperimentiamo un'illusione sensoriale.

Una delle origini di questi disturbi della normale percezione sensoriale potrebbe essere individuata nella presenza di stimoli in competizione tra loro, che andrebbero ad influenzare il significato dello stimolo stesso; ad esempio, quando il conducente di un'auto percepisce i propri fari riflessi nella vetrina di un negozio, sperimenta l'illusione che un altro veicolo stia procedendo verso se stesso, nonostante sia pienamente cosciente che non ci sia alcuna strada di fronte.

Tuttavia, un'illusione non è sempre un errore di interpretazione; talora, l'illusione è talmente profonda che ci convince che la versione falsa di un oggetto sia in realtà quella reale.

Un'illusione può verificarsi in seguito a una stimolazione visiva prolungata, come osservare per molto tempo una fonte luminosa. L'immagine che rimane impressa sulla retina quando si distoglie lo sguardo dalla fonte viene definita “illusione fisiologica”.

Grazie a questo esempio riguardante la vista, possiamo comprendere come la percezione

può essere modificata a causa di uno squilibrio dovuto a una iper- o ipo-stimolazione dei recettori retinici.

Se prendiamo altre illusioni percettive, capiamo anche che non sempre esse derivano da fattori propriamente controllabili, come quando le onde luminose rendono percepibile una matita immersa in un bicchiere come piegata, o quando in condizioni di scarsa illuminazione alcuni oggetti ci appaiono ad una distanza maggiore o minore di quanto non siano nella realtà.

Un'illusione percettiva può essere di tre tipi: ambigua, distorta, e paradossale.

Le illusioni percettive ambigue sono immagini che posso avere due interpretazioni valide di ciò che rappresentano (Fig. 1). L'osservatore, infatti, è solitamente in grado di visualizzare mentalmente una delle due interpretazioni immediatamente, mentre la seconda gli appare visibile solo dopo un certo tempo. Tuttavia, queste due rappresentazioni mentali non possono apparire visibili nello stesso istante in quanto ciò interferirebbe con la piena percezione di una delle due, e il cervello semplicemente non consente questa sovrapposizione.

Un esempio è il cubo di Necker, in cui è difficile dire se l'angolo rappresentato emerge dalla figura o ne è alla base.

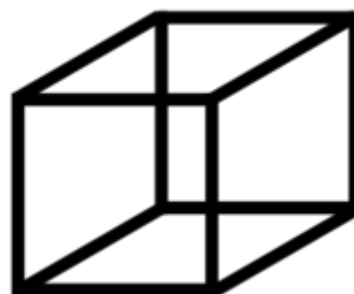


Fig. 1.

Le illusioni percettive di distorsione sono immagini distorte nella loro geometria:

dimensioni, lunghezza, posizione, curvatura (Fig. 2). Un esempio di illusione è quella di Muller-Lyer, dove due linee separate con due frecce alle estremità di ogni riga, rivolte in direzioni opposte, sembrano essere di lunghezza diversa, quando in realtà sono esattamente identiche.

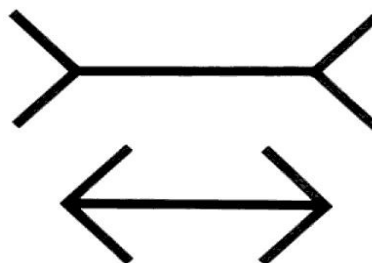


Fig. 2

Infine, un'illusione di finzione (o paradosso) è un'immagine che è semplicemente impossibile da rappresentare tridimensionalmente, ma diventa tale raffigurandola bidimensionalmente (Fig. 3). Uno dei migliori esempi di questo genere di illusione è la famosissima scala di Penrose. Questa illusione nasce perché nella figura emerge una dimensione irreali, grazie alla falsificazione della prospettiva angolare. Essa ci fa percepire una scala costruita su quattro lati come infinita, quando nella realtà ciò non può essere costruito. Quindi, l'illusione è un travisamento di un vero stimolo sensoriale, cioè una contraddizione della realtà oggettiva.

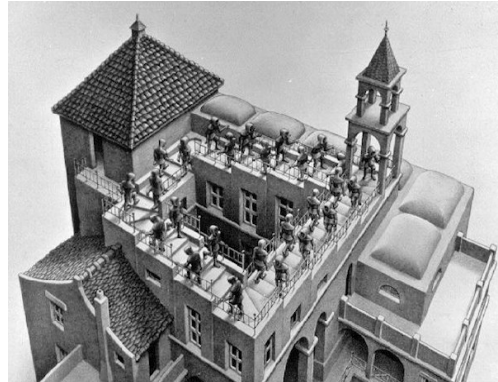


Fig. 3

Nei prossimi capitoli vedremo come queste illusioni sono state utilizzate nelle neuroscienze, e successivamente ingegnerizzate, per studiare molteplici fenomeni percettivi e, di conseguenza, anche i casi in cui questi fenomeni percettivi sono deficitari.

2. L'INTEGRAZIONE SENSORIALE E IL TEST DELLA SOUND-INDUCED FLASH ILLUSION

2.1 L'integrazione sensoriale

Come già ampiamente discusso nel primo capitolo, la nostra percezione dell'ambiente circostante e degli eventi che si verificano in esso è resa possibile grazie al nostro cervello e dal fatto che esso è in grado di ricevere informazioni dai differenti canali sensoriali. Al fine di produrre una percezione unica e coerente, queste informazioni devono essere integrate tra di loro.

Una delle domande più frequenti delle moderne neuroscienze cognitive riguarda il modo in cui il nostro cervello sia in grado di realizzare questa integrazione multisensoriale, per dar luogo alla percezione dell'ambiente esterno ed interagire opportunamente con esso.

Nessun sistema di elaborazione delle informazioni, tecnico o biologico, è abbastanza potente da poter effettuare accuratamente una percezione completa utilizzando le informazioni provenienti da un'unica modalità sensoriale, ma piuttosto cercano di utilizzare delle informazioni multisensoriali, che andranno inevitabilmente combinate e integrate ad un qualche livello cerebrale per produrre la vera e propria percezione.

Per molti anni, le diverse modalità sensoriali sono state studiate separatamente, in quanto considerate indipendenti e deputate al processamento di forme di energia differenti, quali la luce per la vista, la pressione cutanea per il tatto, o le onde sonore per l'udito.

Negli ultimi decenni è cambiata notevolmente la visione del problema, e si è iniziato a parlare di integrazione multisensoriale, definita come la capacità del Sistema Nervoso Centrale di combinare sinergicamente ed efficacemente gli stimoli cross-modali (molteplici stimoli provenienti da diversi canali) per produrre, come risposta finale, una percezione unitaria capace di sintetizzare in maniera significativa i diversi stimoli percepiti (Stein & Meredith, 1993).

2.1.1 Aspetti generali

Una delle teorie più interessanti che ha portato alla formulazione delle moderne teorie dell'integrazione multisensoriale è quella dello scienziato cognitivo Jerry Fodor, il quale sosteneva l'esistenza di un "incapsulamento informativo". Fodor basava la sua teoria sul

precetto dell'isolamento delle modalità percettive, ovvero sul credo che le informazioni provenienti da ogni singolo canale sensoriale venissero elaborate in maniera indipendente tra loro in aree neocorticali di alto livello, e che quindi gli output di una singola modalità entrassero in relazione con gli output delle altre modalità ai fini di una rappresentazione generale di un evento esterno o dello svolgimento di un compito cognitivo.

Il processo di trasformazione degli input in rappresentazioni implica una teoria computazionale della mente e quindi una realizzazione dei processi cognitivi caratterizzata da un'elaborazione di tipo sequenziale.

L'aspetto di maggior rilievo negli scritti di Fodor risiede nella descrizione che fornisce dell'architettura alla base del funzionamento della mente.

L'analisi degli input viene svolta da strutture verticali, denominate moduli. Questi moduli hanno il compito di mediare tra l'output degli organi sensoriali e i sistemi centrali di elaborazione e, perciò, sono determinati geneticamente: ciascuno di essi si occupa di un dominio specifico ed è collocato in una particolare regione del cervello. Da ciò che è stato appena detto, si deduce che ogni sistema modulare risulta autonomo per le computazioni, cioè non scambia informazioni con gli altri moduli e nemmeno con le strutture centrali.

Le teorie funzionaliste di Fodor si contrappongono a quelle connessioniste, che si basano su un'elaborazione in parallelo delle informazioni e sono tipiche delle reti neurali.

Negli ultimi decenni l'attenzione degli studiosi si è spostata sulla percezione cross-modale, poiché si è notato che in molti casi l'elaborazione delle informazioni provenienti da una modalità sensoriale viene influenzata dalle informazioni disponibili nelle altre modalità, con effetti di facilitazione o talvolta di interferenza con i compiti cognitivi.

Quindi, anche se la prima parte dell'elaborazione degli stimoli avviene in regioni neurali separate, poiché segnali diversi derivano spesso da eventi comuni, il nostro cervello mette in pratica una serie di meccanismi di sinergia e integra il contenuto informativo di tali segnali, realizzando in tal modo un'integrazione tra le informazioni prodotte dalle diverse modalità sensoriali (Meredith, 2002).

Infatti, è importante considerare il fatto che nella quotidianità raramente ci troviamo di fronte a eventi o ad oggetti le cui caratteristiche siano percepibili attraverso un unico canale sensoriale; nella maggioranza dei casi le informazioni percettive provengono da molteplici canali e non si sommano banalmente nel nostro cervello, ma interagiscono tra loro in maniera complessa.

2.1.2 Anatomia

Affinché possa aver luogo la sintesi delle informazioni provenienti dalle diverse modalità sensoriali in favore dell'economia del nostro sistema cognitivo, risulta indispensabile che i dati provenienti da specifiche modalità sensoriali convergano nella stessa regione cerebrale (fenomeno di fusione sensoriale). Tale capacità integrativa determina percezioni più accurate degli stimoli e degli eventi, consente di risolvere ambiguità in caso di stimoli conflittuali, e genera risposte più pronte.

Nel tentativo di spiegare questo fenomeno, negli ultimi decenni, numerosi studi scientifici neurofisiologici si sono concentrati sulla convergenza multimodale di input provenienti da modalità sensoriali differenti su un singolo neurone multisensoriale (ad esempio, Meredith, 2002).

A livello neurale, è stato rilevato che l'interazione tra stimoli di diversa natura avviene a diversi stadi di elaborazione corticale. La corteccia, infatti, è composta di numerose aree che sono distinguibili sulla base di vari criteri. Uno dei più utilizzati distingue le aree primarie da quelle superiori. Le aree primarie sono strettamente legate ai recettori sensoriali periferici (la retina, la coclea o la pelle) attraverso nuclei sensoriali specifici, situati nel talamo. Le aree superiori, di tipo associativo, stabiliscono connessioni indirette con le aree primarie. I neuroni in queste aree tendono ad avere proprietà di risposta più astratte, come quelle che potrebbero riguardare, in alcune aree parietali, i contesti spaziali o, in alcune regioni della corteccia frontale e temporale, le funzioni legate alla memoria. Questi neuroni, proprio grazie alla loro capacità integrativa, sono denominati "neuroni multisensoriali".

Presenti in diverse aree cerebrali e propri di tutti i mammiferi, essi sono particolarmente abbondanti nel collicolo superiore (CS). Questa struttura è contenuta nel mesencefalo ed è dotata di un'organizzazione a strati. Nello specifico, gli strati superficiali sono incaricati della rappresentazione della sola informazione visiva, mentre gli strati più profondi rappresentano le informazioni multimodali visive, uditive e tattili. Posto che il numero di neuroni multisensoriali contenuti in questa area cerebrale risulta molto maggiore di quello di altre aree, diversi studi hanno ipotizzato che il CS possa evocare una risposta agli stimoli multisensoriali significativamente maggiore rispetto a quella prodotta dalle altre aree, e quindi che possa essere riconosciuta come la zona deputata all'integrazione multisensoriale (Stein & Meredith, 1993; Meredith, 2002).

2.1.3 I principi alla base del fenomeno dell'integrazione

Nel sistema nervoso, la rappresentazione sensoriale è basata su regole organizzative sistematiche, che hanno come intento quello di creare una mappa che rispecchi la diretta ricostruzione spaziale della superficie recettoriale, tenendo conto delle relazioni di vicinanza esistenti, per cui recettori di parti adiacenti proiettano a neuroni adiacenti del SNC.

Sia a livello anatomico che a livello funzionale, le rappresentazioni dello spazio visivo, uditivo e somatosensoriale occupano regioni spazialmente distinte del sistema nervoso. Nonostante questa divisione, nel cervello esistono aree deputate a ricevere informazioni da diversi canali sensoriali, particolarmente rilevanti all'interno del collicolo superiore.

Come già ampiamente dibattuto in precedenza, il fenomeno dell'integrazione ha effetti positivi rispetto a diversi aspetti del comportamento, in particolare migliorando la capacità di rilevare stimoli bimodali e facilitando la creazione di esperienze percettive unitarie (Stein & Stanford, 2008). Queste qualità portano a specificare le tre leggi integrative fondamentali che regolano il funzionamento dei neuroni multisensoriali del CS:

1. la regola spaziale;
2. la regola temporale;
3. la regola dell'efficacia inversa.

La REGOLA SPAZIALE sottolinea il ruolo svolto dal CS nell'orientamento spaziale. Essa afferma che esiste una stretta corrispondenza tra le posizioni nello spazio dei singoli input sensoriali, tanto più lo stimolo multimodale darà luogo alla valorizzazione della risposta del sistema (enhancement).

Difatti, solo gli stimoli spazialmente coincidenti vengono integrati e aumentano così la risposta neuronale, mentre gli stimoli spazialmente distanti evocano risposte di depressione (riduzione della probabilità di percepire un segnale) o addirittura non determinano un'integrazione (Stein & Meredith, 1993).

La proprietà spaziale dipende dai campi recettivi, e da come essi sono disposti in zone centrali (eccitatorie) o periferiche (inibitorie). Per esempio, i campi recettoriali acustici presentano regioni eccitatorie più ampie rispetto a quelli visivi. Dato che i campi recettivi acustici e visivi dei neuroni bimodali hanno una parziale sovrapposizione, gli stimoli visuo-acustici spazialmente coincidenti vengono a cadere all'interno delle zone eccitatorie,

producendo l'enhancement della risposta. Se, al contrario, gli stimoli provengono da posizioni spazialmente lontane, tali che uno stimolo cade all'interno della regione inibitoria, esso potrebbe essere processato come evento separato o potrebbe provocare la depressione della risposta (Stein & Meredith, 1993; Stein & Stanford, 2008).

In maniera analoga a ciò che è stato appena detto per la regola spaziale, la REGOLA TEMPORALE afferma che esiste una sincronizzazione temporale relativa alla comparsa delle componenti unimodali, ovvero che i due input devono essere presentati simultaneamente, per poter ottenere l'enhancement della risposta, ma aggiunge che non sempre tale sincronizzazione corrisponde alla comparsa contemporanea dei singoli stimoli (Stein & Meredith, 1983). Nello specifico, affinché si possa ottenere l'enhancement, l'intervallo temporale tra la comparsa di due stimoli sensoriali differenti deve essere in un intervallo tra i 50 e i 150 ms (l'intervallo ottimale è attorno ai 100 ms) (Meredith, Nemitz & Stein, 1987).

Un esempio calzante della legge temporale è rappresentato dall'Effetto McGurk, che riesce a dimostrare la forte interazione tra vista e udito nel riconoscimento linguistico. Nella nostra vita di tutti i giorni, le componenti visive e uditive del linguaggio sono contemporanee. La tecnologia moderna ci permette di fornire stimoli manipolati artificialmente, come nel caso dello studio in questione, il quale dimostra che nel processo di riconoscimento di una parola o di un singolo fonema, il canale visivo influenza in maniera importante quello uditivo.

Nell'esperimento del 1976, il dottor McGurk montò un video che riprendeva una donna che, muovendo le labbra, produceva un determinato fonema, il quale venne doppiato, però, con il suono registrato di un diverso fonema sovrapposto. Ciò che si osservò fu come spesso il fonema percepito fosse intermedio tra i due; per esempio, il video che mostra /ga/ combinato con l'audio /ba/, spesso viene percepito /ta/ o /da/. In altre parole, ciò che viene percepito è il fonema 'ta-ta' oppure 'da-da' nonostante il labiale della donna riproduca il fonema 'ga-ga' e il suono ascoltato fosse 'ba-ba'.

Si tratta, quindi, di una situazione di conflitto tra la vista e l'udito, dal momento che, non osservando le labbra chiudersi all'inizio della sillaba, il cervello ritiene, tramite un'inferenza cognitiva, di non poter udire il suono 'ba'. Esso risolve, perciò, questa ambiguità grazie a un compromesso, ritenendo di percepire una via di mezzo tra i due

suoni, ovvero 'ta' o 'da'. Per poter percepire il fonema corretto si deve riascoltarlo chiudendo gli occhi. In questo caso, saremo in grado di udire il suono reale, poiché non ci sarà l'interferenza del canale visivo su quello uditivo (McGurk & MacDonald, 1976). Tale effetto è massimo nel caso in cui lo stimolo acustico e quello visivo vengano somministrati in modo contemporaneo, dato che non ci accorgiamo a livello percettivo dell'errore commesso, ma perde efficacia se la sincronizzazione non è ottimale, arrivando a scomparire se i due stimoli sono troppo distanti a livello temporale.

In conclusione, quindi, è possibile osservare come il sistema multisensoriale sia organizzato affinché le risposte di eccitazione e depressione dei neuroni siano ottenute grazie alle relazioni spaziali e temporali tra gli stimoli, e non grazie alle loro modalità specifiche di appartenenza.

Ultima ma non meno importante, la REGOLA DELL'EFFICACIA INVERSA precisa una relazione tra il livello di efficacia degli stimoli e la risposta neurale da loro evocata.

Alcuni studi elettrofisiologici hanno dimostrato come la stimolazione di un solo canale sensoriale produca un'attività di scarica più debole rispetto alla stimolazione simultanea di più canali sensoriali (Meredith & Stein, 1983).

Altri studi hanno dimostrato che la maggior parte delle integrazioni multisensoriali avvenga tramite sommazione lineare degli input specifici per modalità, anche se è interessante notare come la risposta multisensoriale sia massima combinando stimoli unimodali deboli, piuttosto che combinando stimoli unimodali potenti.

La regola dell'efficacia inversa approfondisce e unisce questi concetti. Tale principio afferma che l'enhancement multisensoriale generato da due stimoli deboli risulta molto maggiore rispetto all'enhancement prodotto da coppie di stimoli forti.

Ciò significa che due stimoli unimodali, presi singolarmente, non sarebbero in grado di evocare un effetto significativo sulla risposta neuronale, ma la loro combinazione può aumentare consistentemente la risposta dei neuroni multisensoriali (Stein & Meredith, 1993).

2.2 Il test della sound-induced flash illusion

La nostra percezione del mondo beneficia chiaramente delle informazioni portate dalle modalità sensoriali multiple. Una strategia tipica nell'esaminare il peso relativo di ogni singola modalità sensoriale rispetto alla percezione totale è quella di rendere le informazioni veicolate da modalità diverse, in conflitto tra loro. I risultati di questi studi identificano nella vista la modalità più importante, o dominante, e spesso suggeriscono che i segnali provenienti da modalità conflittuali sono ignorati, in favore della vista.

Due ben noti esempi di questo paradigma sono il ventriloquismo e la cattura visiva. Il primo implica un conflitto tra localizzazione spaziale quando vengono presentati uno stimolo acustico e uno visivo in posizioni leggermente diverse. La localizzazione percepita degli eventi nel complesso è determinata in modo predominante dalla localizzazione dello stimolo visivo; in altre parole, l'osservatore percepirà i due segnali come se provenissero entrambi dalla posizione dell'input visivo (Howard & Templeton, 1966). In maniera simile, la cattura visiva implica dei compiti di localizzazione spaziale quando l'informazione visiva è in conflitto con quella proveniente da un'altra modalità, ovvero quando si tratta di un'informazione propriocettiva (Hay, 1965). La localizzazione percepita è determinata nuovamente da quella visiva.

Ci sono anche paradigmi conflittuali in cui non è la vista a predominare, ma tuttavia, essa modifica la percezione delle altre modalità. Un ben noto esempio è l'effetto McGurk, di cui abbiamo parlato precedentemente; quando si verifica questo fenomeno, l'informazione visiva altera in modo significativo la percezione audiovisiva del fonema (McGurk, 1976). Uno studio successivo ha dimostrato che la modifica della percezione audiovisiva dovuta a uno stimolo visivo conflittuale non è valida solo per il linguaggio, ma accade anche nel riconoscimento delle note musicali (Saldana, 1993).

Mentre i più noti esempi di interazione cross-modale coinvolgono la modifica da parte della vista delle altre modalità sensoriali, esistono numerosi studi in letteratura che riportano risultati opposti. La maggior parte studia la modifica della percezione delle caratteristiche temporali (come durata, frequenza o sincronizzazione) da parte dello stimolo acustico. In ogni caso, le caratteristiche temporali non sono gli unici attributi dello stimolo visivo soggette a modifica. Stein et al. (1996) riportano che l'intensità percepita degli stimoli visivi aumenta in presenza di uno stimolo acustico (enhancement). Uno studio ha mostrato che uno stimolo acustico può alterare la direzione percepita del movimento di

uno stimolo visivo. L'effetto del suono è quello di polarizzare l'interpretazione dello stimolo in favore di una delle due alternative (Sekuler, 1997).

Sono innumerevoli gli studi che negli ultimi decenni si sono proposti di analizzare, attraverso i fenomeni illusori, le interazioni cross-modali. Uno dei più rilevanti venne svolto nel 2002 dalla professoressa Ladan Shams, la quale contribuì a definire il fenomeno noto come "*sound-induced flash illusion*". Si tratta di un'illusione visiva indotta dal suono: quando un singolo flash è accompagnato da molteplici segnali acustici (beep), il singolo flash è percepito come multiplo.

Di seguito, presenterò lo svolgimento del primo studio pubblicato dalla Shams, che si divide in due esperimenti, diversi ma strettamente correlati, che potranno mettere in evidenza l'interazione visuo-acustica che avviene durante la percezione del fenomeno.

2.2.1 Esperimento 1

L'obiettivo è quello di investigare un fenomeno recentemente scoperto: quando un singolo stimolo visivo (flash) è accompagnato da stimoli acustici multipli (beep), anch'esso è percepito come multiplo.

All'esperimento hanno partecipato 8 volontari (6 femmine e 2 maschi). La loro età era compresa tra i 24 e i 41 anni. In ogni prova, un disco bianco uniforme è stato proiettato, da una a quattro volte, sullo sfondo nero dello schermo di un computer (Fig. 1). Nelle prove con un singolo flash, il flash era accompagnato da un segnale acustico (beep) ripetuto dalle 0 alle 4 volte, mentre nelle prove con flash multipli, i flash erano accompagnati da uno o nessun beep.

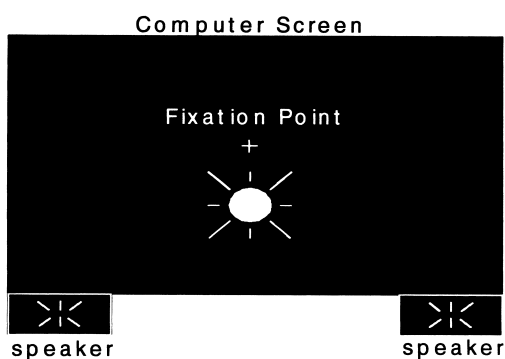


Fig. 1. Un disco bianco uniforme è disposto su uno sfondo nero, leggermente al di sotto del punto di fissazione, che è posto al centro dello schermo. Approssimativamente nello stesso istante, alcuni beep sono emessi da due casse acustiche poste direttamente al di sotto e ai fianchi dello schermo

I dati di seguito riportati fanno riferimento a prove con un singolo flash accompagnato da 1 fino a 4 beep. Il primo segnale acustico ha sempre preceduto il primo flash di 23 ms. Ogni beep ha avuto una durata di 7 ms, mentre i beep consecutivi sono stati distanziati di 57 ms

(Fig. 2). Si è scelto di mantenere i flash successivi strettamente distanziati al fine di ottenere l'illusione di percepire flash multipli.

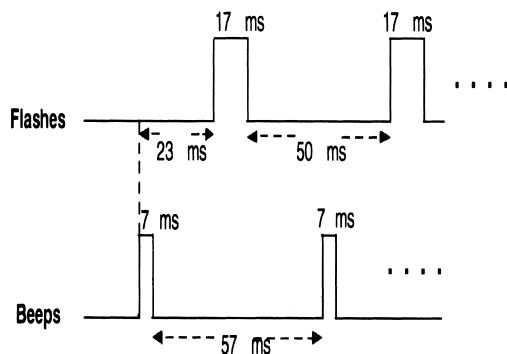


Fig. 2. Profili temporali dell'esperimento 1. Possiamo notare la relazione che esiste tra il timing dei beep e quello dei flash.

I partecipanti si sono seduti a una distanza di 57 cm dallo schermo. Per tutta la durata delle prove c'era un punto di fissazione esattamente al centro dello schermo. L'obiettivo per i partecipanti era quello di giudicare il numero di flash osservati. L'esperimento consisteva in 5 prove per ogni configurazione possibile, per un totale di 60 prove. Si è scelto di porre in ordine casuale le 15 prove che contenevano l'illusione tra le altre 45 che non implicavano l'illusione. Tale impostazione serviva a garantire che i partecipanti impiegassero la stessa strategia (per giudicare il numero di flash) sia per i test con l'illusione che per i test senza.

I risultati sono mostrati in Fig. 3, che rappresenta i dati delle prove in cui è stato presentato un singolo flash. Il numero di flash percepiti è il dato sull'asse delle ordinate, mentre il numero di beep emessi è quello sull'asse delle ascisse; le barre verticali rappresentano l'errore medio standard.

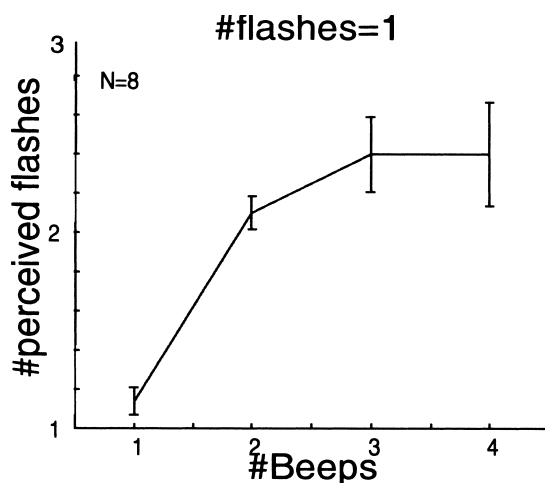


Fig. 3. Flash illusori. È rappresentata la media dei flash percepiti dagli 8 partecipanti all'esperimento, e messo in contrasto con i beep emessi dagli altoparlanti. Ricordiamo che il caso preso in esame è quello in cui è stato emesso un unico flash. I partecipanti hanno detto di aver visto 2 o più flash, quando il singolo flash era accompagnato da 2 o più beep.

I partecipanti hanno riportato di aver visto un flash, vale a dire il valore reale, quando il numero dei beep era 1, ma hanno riportato di aver visto 2 o più flash quando il flash era accompagnato da 2 o più beep. Gli studiosi si riferiscono a questo fenomeno come “*sound-induced flash illusion*”. I risultati appena citati suggeriscono che i beep multipli possano modificare la percezione del singolo flash.

Una possibile osservazione è che il numero di flash riportati sia in risposta alla percezione acustica, contrapposta a quella visiva, dato che dalla Fig. 3 si nota che il numero di flash percepiti corrisponde al numero di beep emessi.

Per investigare su questa possibilità, gli studiosi hanno utilizzato i cosiddetti “test di cattura”, in cui c'era una discrepanza tra il numero di flash e quello di beep: a fronte di un singolo beep, il numero di flash variava da 2 a 4. Dato che precedentemente le risposte degli osservatori erano state determinate dal numero di segnali acustici, ci si aspettava di ottenere, come numero di flash percepiti, un numero in accordo con quello dei beep emessi (0 oppure 1). Ciò non è avvenuto: le risposte degli osservatori sono state in linea con il numero di flash a cui sono stati sottoposti, e quindi in conflitto con il numero di beep. In questo caso, quindi, la risposta dei partecipanti è basata sulla percezione visiva; in altre parole, non esiste una determinazione del fenomeno da parte di un'eventuale deriva cognitiva della percezione acustica.

Prendendo un unico esempio per spiegare ulteriormente il fenomeno, la risposta dei partecipanti è stata la medesima, sia quando essi sono stati esposti a un flash accompagnato da due beep, sia nel caso di due flash accompagnati da uno o da nessun beep: i partecipanti hanno sempre riportato di aver visto due flash. È anche interessante far notare che i partecipanti, una volta concluso l'esperimento, hanno ammesso di non essere stati in grado di distinguere i test illusori da quelli senza l'illusione.

2.2.2 Esperimento 2

Il primo esperimento stabilì che lo stimolo acustico è in grado di alterare la percezione visiva. Il secondo, invece, venne utilizzato per investigare su quale dev'essere la finestra temporale massima tra i beep per cui avvenga l'illusione visiva. Tale finestra è detta “*Temporal Binding Window*”(TBW).

All'esperimento hanno partecipato 8 volontari (5 femmine e 3 maschi). La loro età era

compresa tra i 19 e i 27 anni e nessuno aveva precedentemente partecipato al primo esperimento.

La configurazione degli stimoli era la precedente, ma il numero di flash e beep era lo stesso per tutte le prove: in ogni prova un singolo flash era accompagnato da due beep. Un beep era sempre fisicamente simultaneo al flash, mentre la sincronizzazione dell'altro beep variava da una prova all'altra, con un'asincronia nell'insorgenza dello stimolo (stimulus onset asynchronies, SOA): 25, 70, 115, 160, 205, 250 ms tra un beep e l'altro, sia prima che dopo il flash (Fig. 4). L'obiettivo per i partecipanti era quello di giudicare il numero di flash osservati. L'esperimento consisteva in 5 prove per ogni configurazione possibile, per un totale di 60 prove, disposte in un ordine casuale.

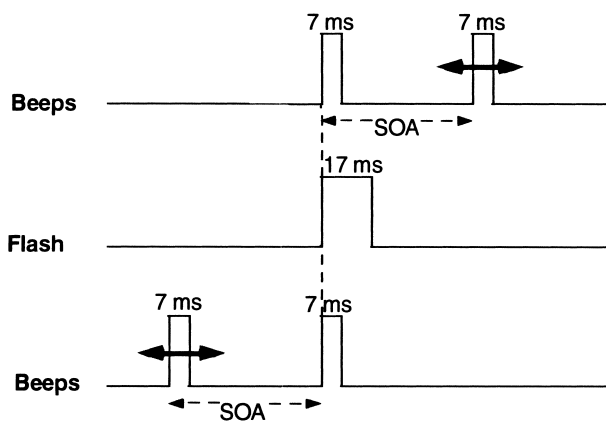


Fig. 4. Profilo temporale. In ogni prova, un flash era accompagnato da due beep. Il primo beep era sempre simultaneo alla comparsa del flash, mentre il secondo poteva presentarsi prima o dopo il flash, come raffigurati rispettivamente nel profilo superiore e inferiore. Il timing del beep non simultaneo variava da 25 a 250 ms.

I risultati sono riportati in Fig. 5. L'asse delle ordinate rappresenta la percentuale di prove in cui gli osservatori hanno visto più di un flash, mentre l'asse delle ascisse segna il timing dei beep. Lo zero indica il momento in cui viene emesso il flash; indica, quindi, la simultaneità tra flash e beep. I numeri positivi e negativi segnano la sincronizzazione dei beep variabili occorsi, rispettivamente, dopo o prima del flash.

L'illusione inizia a diminuire intorno ai ± 70 ms. Tuttavia, è ancora significativa (tra il 33% e il 23%) fino ai ± 115 ms. Questi risultati vogliono suggerire che il fenomeno illusorio tende a svanire se il secondo beep è troppo lontano dal primo ed esce, quindi, dalla TDW.

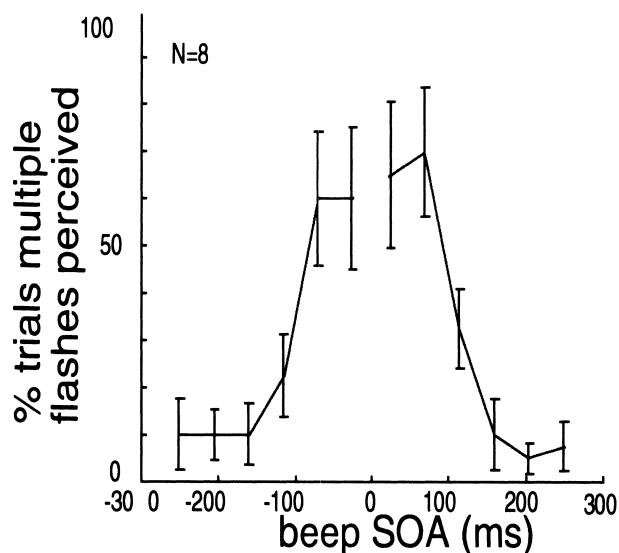


Fig. 5.

2.2.3 Conclusioni

I risultati dei due esperimenti descritti sopra respingono qualsiasi spiegazione alternativa all'effetto del flash illusorio che è stato osservato. Questo fenomeno non sembra essere dovuto a un enhancement attentivo nei confronti degli stimoli acustici, dato che le prove con un singolo beep non hanno suscitato alcun fenomeno illusorio significativo, ad eccezione di una debole fusion illusion nel caso di singolo beep (Esperimento 1). Ciò non è causato dal movimento oculare, poiché l'effetto è più forte con flash di durata minore (dati non mostrati in questi esperimenti); inoltre l'effetto persiste in presenza di dischi di dimensioni maggiori, mentre diminuisce alla diminuzione del contrasto tra il disco e lo sfondo.

Altri risultati, inoltre, respingono una possibile origine cognitiva di tipo “top down”: l'illusione sparisce quando il secondo beep cade al di fuori della finestra di interazione (Esperimento 2) e diventa più forte all'aumento dell'eccentricità del disco all'interno del campo visivo.

Perciò, l'unica spiegazione per i risultati ottenuti è che gli stimoli uditivi (beep) alterino la percezione degli stimoli visivi (flash) in caso di interazioni percettive bimodali, anche quando non sono presenti ambiguità negli stimoli visivi. Questa alterazione è più rilevante nel caso in cui un singolo flash, accompagnato da due beep, sia percepito come multiplo. La modulazione inversa, denominata fusion illusion, ossia la fusione di due flash fisici in

un unico flash percepito, quando essi sono accompagnati da un singolo beep, anche se presente, è del tutto trascurabile (Esperimento 1). Questa asimmetria nella modulazione è interessante, poiché non può essere spiegata dalle ipotesi di “modality appropriateness” (Welch & Warren, 1980), una ben nota teoria che sostiene che la direzione delle interazioni cross-modali dipenda dall'adeguatezza delle modalità coinvolte in un determinato compito; la modalità più adatta a portare a termine un compito dominerà inevitabilmente sulle altre. La teoria appena citata non può essere la spiegazione per l'asimmetria dei dati, né per la prova (giudicare il numero di flash sullo schermo), né tanto meno può essere il motivo per cui le modalità sensoriali coinvolte (vista e udito) cambiano a seconda della prova somministrata. I risultati, anzi, suggeriscono che la direzione delle interazioni cross-modali dipenda almeno parzialmente dalle caratteristiche degli stimoli; alcuni studi hanno rilevato che gli stimoli discontinui in una modalità alterano la percezione di uno stimolo continuo in una seconda modalità, ma non il viceversa (Saldana & Rosenblum, 1993). Concludendo, possiamo notare come l'illusione trattata abbia mostrato un certo grado di robustezza nei confronti dei cambiamenti di molteplici parametri. Moderate manipolazioni della sincronizzazione assoluta e relativa, sia del flash che del beep, la forma, il colore, la luminosità e la dimensione del flash, la frequenza e l'intensità del beep, la disparità spaziale tra la fonte sonora e quella visiva, nessuna di queste modifiche è riuscita ad alterare la percezione dell'illusione.

Un tale grado di robustezza, e nel contempo di semplicità, suggerisce che questa illusione possa riflettere il funzionamento di uno dei principali circuiti neuronali, in opposizione all'attività neurale marginale. Suggerisce, inoltre, che l'interazione cross-modale sia la regola, piuttosto che l'eccezione, nella nostra percezione degli eventi esterni.

Inoltre è interessante analizzare la finestra temporale, che abbiamo definito nell'intorno di 100 ms, dato che è coerente con la finestra di integrazione dei neuroni multisensoriali del cervello umano. Molti studi si sono dedicati allo studio di tale finestra temporale di integrazione, soprattutto perché essa non è uguale per tutti i soggetti, e può essere modificata dalla presenza di patologie o anomalie.

2.3 L'integrazione sensoriale deficitaria: l'autismo

I bambini e gli adulti con autismo, così come tutti coloro che sono soggetti ad altre

disabilità dello sviluppo, possono avere un sistema sensoriale non adeguatamente funzionante. A volte, per esempio, uno o più sensi sono ipo- o iper-reattivi agli stimoli. L'integrazione sensoriale è un processo neurobiologico innato e riguarda l'integrazione e l'interpretazione della stimolazione sensoriale dall'ambiente al cervello. Una disfunzione sensoriale integrativa, quindi, rappresenta un disordine in cui l'input sensoriale non può essere organizzato, o integrato, appropriatamente dal sistema nervoso, e ciò può produrre vari gradi di problemi nello sviluppo, nell'elaborare le informazioni e nel comportamento.

Per analizzare a fondo come l'integrazione multisensoriale sia deficitaria in patologie come l'autismo, è opportuno introdurre l'argomento spiegando in cosa consistono i Disturbi dello Spettro Autistico (DSA).

2.3.1 Cos'è l'autismo?

Stime recenti indicano che circa 60 milioni di persone al mondo sono affette da Disturbi dello Spettro Autistico (DSA). Tutte le fonti più autorevoli, dall'ICD 10 (International Classification of Diseases) al DSM IV (Manuale diagnostico e statistico dei disturbi mentali), concordano nel definire questo gruppo di sindromi come Pervasive Developmental Disorders, ovvero disturbi dello sviluppo mentale, i cui sintomi hanno un'insorgenza precoce, prima dei tre anni d'età, e permangono per tutta la durata della vita. Che cosa possa causare l'insorgenza di questi disturbi è ancora oggetto di ricerca; sappiamo però che esse sono in aumento. Confrontando gli studi epidemiologici dal 1966 al 1997, si è notato che, mentre il rapporto maschi/femmine si è mantenuto pressoché uguale 4/1, l'aumento dei casi riscontrato sia stato di quasi il 4% annuo. Oggi si stima che nella sua forma marcata l'autismo coinvolga un bambino ogni 300 nati, mentre sintomi dello spettro autistico possano manifestarsi quasi nell'1% della popolazione (Wing, 1996).

I DSA sono disturbi nel neurosviluppo, caratterizzati da una triade sintomatologica.

La triade sintomatologica descrive le principali aree interessate dalle difficoltà autistiche:

- 1) alterazione e compromissione grave e generalizzata della capacità di interazione sociale
- 2) alterazione e compromissione della qualità della comunicazione, verbale e non-verbale (nei casi più estremi, il bambino può persino non sviluppare mai il

linguaggio)

3) modelli di comportamento e interessi limitati, stereotipati e ripetitivi.

La persona autistica, come si è detto, ha gravi deficit nella capacità di comunicazione. Ciò che manca non è solo il linguaggio verbale, che nella maggioranza dei casi compare in ritardo rispetto allo sviluppo normale, o può perfino non comparire affatto, ma è la capacità di utilizzare la comunicazione spontanea nell'interazione. Per esempio, partendo da un gesto molto semplice, il bambino autistico mostra difficoltà nell'indicare con il dito indice per chiedere o indicare qualcosa; tale gesto, infatti, presuppone che il bambino posseda la consapevolezza che l'adulto con il quale sta interagendo comprenda il suo messaggio, ma purtroppo i deficit relazionali degli autistici precludono loro anche queste semplici comunicazioni non verbali. Egli reputa molto più semplice afferrare il braccio dell'adulto e condurlo nei pressi dell'oggetto che desidera (Mundy, Sigman, Ungerer e Sherman, 1986; Baron-Cohen, 1989, 1998).

Generalmente, il bambino autistico non possiede lo sguardo condiviso, evita il contatto visivo, non risponde al suo nome, evita il contatto umano. Inoltre, nei casi in cui sia presente il linguaggio, spesso può parlare in seconda o terza persona e difficilmente utilizza il linguaggio per le funzioni classiche, come il chiedere o il raccontare (Baron Cohen, 1995).

Nelle relazioni con gli altri manca la consapevolezza delle regole sociali, cosa che noi utilizziamo continuamente e che ci aiuta a destreggiarci soprattutto nelle situazioni nuove o improvvise, dalle più semplici, come salutare o non alzare la voce in un ambiente pubblico, alle più sofisticate, come il valutare la tempistica di una conversazione e quindi sapere quando conviene interrompere il proprio interlocutore (Baron Cohen et al., 1999).

L'incapacità di utilizzare i canali comunicativi "normali" e la mancanza di intuito sociale possono rendere la vita di una persona autistica molto complicata; l'incomprensione del mondo che li circonda può, in alcuni casi, causare anche episodi di autolesionismo o di aggressività ingiustificata verso gli altri.

Le crisi che le persone autistiche vivono non sono generate solo dall'incomprensione del mondo circostante e dalle difficoltà comunicative, ma spesso anche da reazioni verso stimoli sensoriali che sono trattati in modo molto alterato a livello del Sistema Nervoso

Centrale.

Innanzitutto, sembra vi sia un'iperselettività sensoriale, per cui certi stimoli vengono percepiti in maniera opposta a quanto avviene negli individui neurotipici. Questo si pensa che avvenga in quanto i singoli stimoli sono scollegati dall'insieme e dal contesto, mostrando una chiara mancanza di coerenza centrale, funzione indispensabile, che ci permette di unire in una visione coerente stimoli anche molto diversi tra loro (Uta Firth, 1989). Ciò si manifesta, per esempio, anche in bambini molto piccoli, quando riescono a comporre un puzzle a rovescio, ovvero dalla parte del cartone, in quanto il loro cervello è in grado di discriminare la forma degli incastri molto meglio del cervello di un bambino con profilo tipico di sviluppo.

Anche dal punto di vista uditivo, possiamo osservare questa ipersensibilità. In genere possono percepire suoni quasi impercettibili o, al contrario, essere disturbati dai rumori forti. Questo non aiuta l'apprendimento linguistico, né la partecipazione ai momenti sociali, come la scuola e, successivamente, il lavoro.

Se per quel che riguarda la sintomatologia la scienza è riuscita a fornire un quadro piuttosto dettagliato, la stessa cosa non si può dire per l'eziologia, ovvero le cause scatenanti dei disturbi dello spettro autistico; questo è l'obiettivo primario degli studiosi al momento, per poter intervenire in maniera più efficace sul problema e, nel caso, anche trovare un modo per prevenirlo.

Per molto tempo l'autismo è stato studiato seguendo teorie psicogenetiche, che imputavano l'origine dell'autismo a un problema di tipo relazionale e psicologico tra il bambino e la madre, considerata incapace di manifestare il suo amore, tanto da bloccarne la naturale evoluzione sociale e comunicativa. Chiaramente, queste teorie non erano supportate dallo studio scientifico del fenomeno, ma piuttosto dalla convinzione che tutti i problemi di carattere psico-relazionale corrispondessero a un'eziologia di tipo psico-relazionale.

Se queste teorie erano giustificate alla fine del 1800 in quanto la tecnologia non era ancora adatta allo studio della mente, già nella seconda metà del 1900 emergeva l'esigenza di trovare basi più scientifiche su cui studiare l'autismo.

Diversi sono i filoni di ricerca, da quello cognitivo a quello neurofisiologico, che hanno cercato, a partire dagli anni '70, di spiegare le cause alla base dell'insorgenza delle disfunzioni cognitive riscontrate sui soggetti affetti da Disturbi dello Spettro Autistico

(DSA), ma purtroppo c'è ancora tanta strada da fare.

2.3.2 La percezione nell'autismo

Nell'analisi dei Disturbi dello Spettro Autistico (DSA) è fondamentale considerare attentamente il ruolo che riveste l'aspetto della percezione, poiché alla base dei comportamenti problematici sono state riconosciute non solo cause psicologiche, ma anche cause biologiche e ambientali, tra cui, per l'appunto, delle alterazioni nella percezione sensoriale.

Secondo numerosi studiosi, infatti, un'esperienza sensoriale anomala è una caratteristica primaria per poter spiegare i sintomi alla base dell'autismo (Delacato, 1974).

Delle percezioni anomale possono dunque scatenare dei livelli elevati di ansia, a cui seguirebbero comportamenti ossessivi o compulsivi, rendendo i criteri diagnostici più comunemente accettati solo una conseguenza secondaria connessa allo sviluppo sensoriale.

Uno degli aspetti più complicati nello studio delle percezioni relativamente al mondo dell'autismo è l'osservazione diretta dei fenomeni e, nella maggioranza dei casi, è anche difficile interloquire con i soggetti, quindi ricevere risposte alle domande che vengono poste loro. Inoltre, ogni bambino autistico, come d'altronde nel caso dei bambini sani, ha la sua propria sensibilità e quindi anche i tratti caratteristici di ipo- e iper-sensibilità cambiano da caso a caso, rendendone ulteriormente difficile la classificazione.

Per ovviare a tutta questa serie di problematiche ma, nel contempo, continuare a lavorare su queste patologie, la scienza è ricorsa alle testimonianze di adulti autistici ad alto funzionamento (è presente una diagnosi di autismo, ma con lo sviluppo essi hanno dimostrato un livello più elevato di capacità intellettiva nelle prove formali di valutazione delle abilità cognitive, con abilità comportamentali sociali e adattive superiori a quelle usuali per i bambini con autismo) e, in maniera minore, all'osservazione da vicino di bambini autistici. In questo modo, è riuscita a distinguere alcune caratteristiche della percezione autistica del mondo.

Il mondo percettivo delle persone autistiche risulta essere sorprendentemente diverso da quello delle persone non autistiche; esse hanno esperienze sensoriali reali quanto quelle dei non autistici, ma possono essere viste-udite-sentite diversamente, o interpretate in maniera insolita.

I bambini autistici spesso tendono a mostrare una capacità maggiore nel cogliere e ricordare i dettagli. Per esempio, alcuni riescono a riconoscere con esattezza la nota prodotta dal tintinnio fra due calici di cristallo o a identificare decine di marche di aspirapolvere semplicemente dal loro suono. Queste straordinarie abilità percettive, purtroppo, non sono requisiti idonei nella vita quotidiana, in quanto al minimo cambiamento ambientale spesso risultano a disagio e mal disposti nei confronti dell'adattamento.

Numerosi studi (Volkmar, Cohen e Paul, 1986; O'Neill, 1995) forniscono prove della presenza di disturbi sensoriali nel 70-80% delle persone autistiche.

La risposta atipica agli stimoli sensoriali è una caratteristica clinica importante per la diagnosi dello Spettro Autistico. Essa può manifestarsi come una ipo- o iper-reattività ad alcune o a tutte le modalità sensoriali: vista, tatto, olfatto, gusto, sensibilità cinestesica e udito, classificato come il più frequente tra i disturbi. Ovviamente, esistono delle differenze individuali nella gamma e nella severità di questi problemi, anche se le principali anomalie sono comuni (O'Neill e Jones, 1997; Crispiani 2002).

La principale anomalia sensoriale può essere semplificata definendola come una ipo- e/o iper-sensibilità agli stimoli ambientali che spesso fluttua tra i due poli (Baranek GT, Foster LG, Berkson G., 1997; O'Neill M, Jones RS, 1997). Tali stimoli possono essere di natura uditiva, olfattiva e tattile. La causa deriverebbe da un'apertura eccessiva del canale sensoriale, la quale porterebbe all'abbassamento della soglia percettiva, per cui anche un input di lieve entità potrebbe innescare il processo di percezione. Conseguentemente, una persona autistica può annusare le persone e gli oggetti che lo circondano, ricercare il contatto in maniera eccessiva, fino all'autolesione, oppure può rifiutare le fonti di rumori, di odori e di contatti. Un comportamento frequente è quello di coprirsi le orecchie con le mani, nel chiaro tentativo di attutire quelli che per loro sono rumori assordanti; per contro, in alcuni casi sembrano non rispondere affatto ai suoni.

Dal punto di vista tattile, in genere i bambini tendono a schivare il contatto fisico, ma ci sono casi in cui assistiamo a una ricerca di autostimolazione sensoriale specifica (Crispiani, 2002). Emblematica è l'esperienza riportata da Temple Grandin, una persona autistica ad alto funzionamento, professore di Scienze del Comportamento all'Università del Colorado; grazie ai libri da lei scritti, in cui si prefiggeva l'obiettivo di spiegare come

una persona affetta da DSA affronta la quotidianità, ha contribuito a migliorare le metodologie di intervento per supportare le persone colpite da questa sindrome. Tra i racconti più interessanti, la scienziata, a soli 18 anni, ha costruito una macchina da stretta (“hug machine”) per soddisfare il proprio desiderio e bisogno di sentirsi abbracciata molto forte e che può regolare in base alle proprie necessità.

Uta Frith (1989), psicologa tedesca e docente presso l'University College di Londra, ha il merito di aver cercato di spiegare le disfunzioni sociali nell'autismo, ipotizzando un danno specifico della capacità di integrare l'informazione a differenti livelli.

Il sistema cognitivo di un individuo normodotato possiede la capacità di elaborare le singole informazioni percepite dall'ambiente esterno e di integrarle ai fini di costruire livelli sempre più alti di contesto del significato di quanto osservato.

La persona autistica, invece, è incapace di cogliere il significato dello stimolo nel suo complesso; si parla, infatti, di Deficit nella Coerenza Centrale (DCC), che potrebbe, almeno in parte, spiegare le problematiche sociali e percettive tipiche di queste sindromi. Secondo questa teoria (detta, per l'appunto "Teoria della Coerenza Centrale"), il profilo cognitivo dell'autismo è caratterizzato da un'elaborazione segmentata dei frammenti di un'esperienza, ovvero da una tendenza a dare più importanza al singolo particolare e da una difficoltà ad accedere da questo particolare al suo senso più generale. Infatti, spesso chi soffre di autismo si focalizza su un solo, spesso irrilevante, aspetto di un oggetto. Per esempio, possono concentrarsi sul colore di un utensile ed ignorarne altri aspetti, quali la forma o lo scopo. In questo caso, concentrandosi solo sul colore, può essere difficile per un bambino autistico distinguere una forchetta da un cucchiaio.

Quindi, l'individuo autistico presenta un deficit nella capacità di integrare una serie di caratteristiche per ricreare la struttura globale di uno stimolo e pertanto avrebbe conseguenti difficoltà ad attribuire un significato allo stimolo stesso.

Il nostro cervello è da sempre abituato a scomporre e ricomporre gli stimoli al fine di darci una comprensione il più ottimale possibile della realtà che ci circonda, ma questo deficit che si evidenzia nel cervello autistico porta le persone che ne soffrono a non sapersi relazionare non solo con gli altri esseri umani, ma anche con l'ambiente e con gli stimoli percettivi.

Questo deficit è in grado di spiegare sia le carenze che si rilevano, sia le isole di abilità a

volte sorprendenti (Firth, Happè, 1994). A causa del fallimento dei processi centrali di pensiero, infatti, i soggetti autistici esprimono le sensazioni come percezioni frammentarie, come pure in forma frammentaria pianificano ed eseguono le azioni. I successi che si evidenziano nella percezioni di singole parti che compongono una realtà più generale (“vedere l’albero anziché la foresta”) può essere attribuito a questa loro specifica abilità di individuare in modo preferenziale le singole parti di un oggetto piuttosto che una totalità. Questa evidente incapacità di raggiungere una coerenza centrale o un significato ha come inevitabile conseguenza il distacco o la frammentazione delle attività, e questo potrebbe essere anche la causa del deficit sociale.

Questa tendenza a frammentare le informazioni si ritrova, solitamente, anche nel metodo di elaborazione delle percezioni sensoriali.

I DSA tendono a concentrarsi su uno stimolo sensoriale alla volta, trascurando l'insieme e il contesto. Questo fa diventare la percezione di stimoli complessi, provenienti da più canali contemporaneamente, difficoltosa, se non addirittura impossibile. Essa viene anche definita “Multichannel perception”(O'Neil, 1997).

Come conseguenza di questa percezione frammentaria, gli individui autistici mostrano ripetitività, resistenza al cambiamento, ansia in luoghi non familiari. Molti bambini diventano estremamente insistenti sulle routine. Alcuni esempi comuni sono: mangiare o bere lo stesso cibo ad ogni pasto, vestire certi abiti o insistere che altri indossino sempre gli stessi abiti, andare a scuola percorrendo sempre la stessa strada.

Se invertiamo i termini del problema, questa selettività negli stimoli porta spesso le persone autistiche a subire un sovraccarico percettivo: generalmente, le situazioni caratterizzate da un eccesso di stimoli visivi (luoghi affollati o con immagini e luci molto stimolanti), o di stimoli uditivi (luoghi rumorosi o con suoni inconsueti o sgradevoli), possono suscitare disagio e insofferenza che possono dar luogo a reazioni di rabbia e di aggressività.

Il sovraccarico di informazioni può dissiparsi in tempo, prima di condurre ad un black-out temporaneo, in cui il bambino si chiude in se stesso e non interagisce col resto del mondo. Si può incorrere in una grave ipersensibilità sensoriale (i colori diventano insopportabili, certi motivi prendono a mettersi in rilievo in modo invasivo, il tatto può dare sensazioni di “pizzicore” o di “solletico”, o provocare “choc”) e tuttavia essere in grado di continuare a

trattare le informazioni. In altri casi, un sovraccarico non distribuito nel tempo può dare luogo ad agnosia sensoriale temporale, una temporanea incapacità di trattare informazioni tattili, sonore o visive (Williams).

2.3.3 L'integrazione audio-visiva

L'UDITO

L'udito è uno dei sensi analizzabili facilmente anche nel neonato. Subito dopo la nascita, nonostante l'im maturità dell'apparato uditivo, il neonato riesce ad avere una risposta a stimoli acustici di una certa intensità: nel caso di un suono debole ruota gli occhi nella direzione da cui proviene il suono, mentre se questo è molto intenso li ruota nella direzione opposta (questo comportamento mostra la stretta relazione tra l'apparato acustico e quello visivo). In caso di stimolo di frequenza elevata la reazione ottenuta è di difesa (pianto o spavento), a basse frequenze invece si ha reazione di attenzione e orientamento. Come nel caso del volto, anche per quanto riguarda la voce della mamma, il bambino mostra un riconoscimento immediato, diventando presto capace di discriminare anche le modulazioni qualitative per un'opportuna interazione emotiva.

Nonostante nella maggioranza dei casi la comparsa dei sintomi dell'autismo avvenga dopo 12-18 mesi di sviluppo normale, alcuni comportamenti che evidenziano l'insorgenza dei sintomi sono osservabili anche nei neonati. Per esempio, tendono ad incurvare la schiena allontanandosi dalla persona che li accudisce in modo da evitare il contatto fisico; in altri casi, i neonati non mostrano nessuno dei comportamenti appena descritti in relazione ai suoni, quindi non rispondono alla voce della mamma ed evitano il contatto visivo (S. Edelson, 2004).

Generalmente i suoni che disturbano maggiormente le persone autistiche sono striduli e acuti (come quelli prodotti da frullatori, aspirapolvere, trapani elettrici, seghe). Alcune persone non tollerano i rumori della quotidianità; per qualcuno può essere insopportabile perfino il rumore della pioggia o il sottofondo della televisione. In questo caso, è come se il bambino autistico percepisse costantemente un brusio di fondo e fosse distratto da questa interferenza proveniente dall'esterno. Naturalmente, come succede per i soggetti a sviluppo tipico, esistono differenze individuali e un suono che disturba una persona può essere piacevole per un'altra. Inoltre è ipotizzabile che i problemi uditivi di comprensione possano

essere alla base dei problemi di linguaggio.

L'ipersensibilità uditiva nei DSA è stata descritta nella letteratura, ma ci sono ancora molte controversie per quel che riguarda le possibili cause, le diagnosi e le conseguenze.

Anormalità nell'ambito dell'ipersensibilità uditiva sono presenti nel 90% dei casi; non ci sono teorie per spiegare questo fatto. Sebbene le cause rimangano tuttora sconosciute, è l'anormalità più comune per i DSA (E. Gomes, 2008).

Le teorie su questo argomento differiscono in relazione alla natura dell'alterazione, strutturale o funzionale, alla modalità interessata, sensoriale o cognitiva, i processi coinvolti, integrazione o modulazione tra le aree corticali, ma tutte concordano sul fatto che è sintomo dell'autismo la processazione atipica degli stimoli acustici (Iarocci G, McDonald, 2006).

Ciò che si può capire da quanto è stato appena discusso è che al momento non si ha una spiegazione univoca ai deficit di percezione uditiva nell'autismo. Nonostante ciò, un gruppo di scienziati ha potuto notare delle anomalie in una zona del sistema limbico, l'ippocampo, che sembra essere neurologicamente immaturo nei soggetti autistici (Bauman & Kemper, 1994). L'ippocampo è responsabile degli input sensoriali, così come dell'apprendimento e della memoria. In parole semplici, l'informazione è trasferita dagli organi sensoriali all'ippocampo, dov'è processata, e poi trasferita alle aree della corteccia cerebrale specializzate nell'immagazzinamento a lungo tempo. Dato che le informazioni uditive sono processate proprio nell'ippocampo, esse potrebbero essere trasferite in maniera errata alla memoria a lungo termine.

LA VISTA

La percezione visiva è meglio descritta come un'esperienza soggettiva, la quale deriva dalla stimolazione sensoriale del nostro sistema nervoso.

L'attenzione per alcuni particolari, quali il colore, il movimento e la forma, sono innate; altri tipi di discriminazione visiva, come una distinzione più fine all'interno di una categoria di oggetti, continuano a svilupparsi nel corso della vita.

La vista può ugualmente prestarsi a distorsioni percettive. Alcuni DSA possono essere attratti da certi tipi di colore, da oggetti in movimento, da particolari forme, mentre altri possono esserne spaventati, spesso a causa di distorsioni delle misure e del moto. Alcuni si comportano come se fossero ciechi quando si trovano in luoghi sconosciuti, altri hanno

momenti in cui vedono tutto bianco o tutto nero o, addirittura, in modo bidimensionale; altri ancora manifestano problemi nella percezione della profondità. Inoltre, si riscontrano difficoltà nello stabilire il contatto oculare o nel riconoscere le espressioni facciali e tale limite sembra dovuto a specifici deficit funzionali di un'area cerebrale, l'area fusiforme lobo temporo-ventrale, che potrebbe spiegare le implicazioni in ambito sociale.

Nel cervello di un bambino normodotato le aree visive inferiori creino sofisticate rappresentazioni tridimensionali, per esempio di un cavallo. Incrementando le sue nozioni sul mondo, le aree corticali superiori del bambino generano descrizioni più astratte e teoriche del cavallo: un animale con un lungo muso, quattro zampe ed una coda... Col tempo, la visione che il bambino ha del cavallo è dominata da queste astrazioni superiori. Egli è maggiormente motivato dai concetti ed ha meno accesso alle rappresentazioni fatte in precedenza, di carattere maggiormente visivo (Snyder, 2003).

Negli ultimi dieci anni, numerose ricerche si sono concentrate sul problema di come le persone soggette da autismo percepiscono il movimento.

Come ogni processo intrinsecamente dinamico, la percezione visiva del movimento richiede l'integrazione di informazioni sia spaziali che temporali. Per di più, la percezione visiva di un oggetto in movimento richiede l'integrazione di informazioni provenienti da regioni disconnesse dello spazio retinico (Wallach, 1976). In questo senso, gli studi (Johansson, von Hofsten, & Jansson, 1980) si sono concentrati sulla percezione di uno stimolo locale, come un singolo punto in movimento, in opposizione a uno stimolo globale, più punti che si muovono rispetto ad un punto fisso. Dato che gli osservatori con DSA mancano di alcuni processi di movimento locale, essi possono differire dai soggetti tipici nella loro percezione di questo tipo di stimolo.

Alcuni studi hanno osservato che i deficit nella sensibilità visiva al movimento nei soggetti con DSA siano funzione della velocità, direttamente proporzionali all'aumento della velocità (Gepner & Mestre, 2002). Altri studi, invece, hanno proposto che la durata dello stimolo potesse essere un fattore critico perché la performance con uno stimolo breve dipende molto di più dalla memoria e dai processi attentivi, aspetti compromessi nell'autismo (de Jonge et al., 2007).

Le difficoltà sociali sono largamente considerate il deficit fondamentale dell'autismo. Numerose ricerche hanno investigato se i deficit percettivi potessero essere la causa ultima di molti dei problemi sociali e cognitivi dell'autismo (Schultz, 2005). In generale, il movimento del corpo umano porta con sé moltissime informazioni di carattere sociale. Per esempio, dall'osservazione di una sagoma di una persona in movimento, il soggetto può capirne lo stato d'animo (Atkinson, Dittrich, Gemmell & Young, 2004), il sesso (Pollick, Kay, Heim, & Stringer, 2005) e l'età.

I soggetti autistici possono avere accesso a queste informazioni sociali? Essi possono percepire accuratamente il movimento a livello visivo, ma poi falliscono nell'analisi cognitiva di queste informazioni percettive. Al contrario, possono fallire nella percezione del movimento, e, come risultato, l'analisi sociale e cognitiva è resa pressoché inutile. Lo studio in questione è utile in quanto i soggetti affetti da autismo tendono ad analizzare gli stimoli visivi sempre a livello locale (Behrmann et al., 2006; Happé & Frith, 2006; Mottron et al., 2006).

Dato che la percezione visiva del movimento richiede normalmente l'integrazione globale delle informazioni spaziali e temporali, numerose ricerche si sono concentrate sulla sensibilità visiva rispetto al movimento nei soggetti con Disturbi dello Spettro Autistico.

L'INTEGRAZIONE

Partendo dalla Teoria della Coerenza Centrale (Frith, 1994), secondo la quale l'autismo potrebbe implicare un'abilità alterata nel percepire il significato dello stimolo nel suo complesso, diversi studi si sono interessati al problema di elaborazione da parte dei soggetti affetti da questo deficit degli stimoli multisensoriali.

I bambini con DSA, infatti, mostrano rilevanti difficoltà nell'orientamento visivo rispetto agli stimoli interattivi, ad esempio, sentirsi chiamare per nome o sentire il suono delle mani che applaudono (Dawson et al., 1998).

Un importante studio in merito venne eseguito nel 2008 da Elizabeth Mongillo ("*Audiovisual Processing in Children with and without Autism Spectrum Disorders*"), la quale cercò di analizzare la reazione percettiva all'effetto McGurk di un gruppo di soggetti autistici ad alto funzionamento, posto a diretto confronto con un gruppo di soggetti normodotati.

Ciò che è emerso dallo studio è che i bambini con DSA mostrano difficoltà in questo

specifico tipo di test, in quanto viene richiesta un'elaborazione audio-visiva che coinvolge stimoli "umani", quali le voci e i volti. Al contrario, la loro performance nelle prove che riguardavano gli stimoli "non-umani", per esempio vedere/sentire una pallina che rimbalza, è risultata essere pressoché paritaria rispetto a quella dei soggetti normodotati. I risultati potrebbero suggerire delle forti implicazioni cognitive nello svolgere i compiti di integrazione; in questo caso particolare, una delle spiegazioni potrebbe essere che i bambini autistici sono meno influenzati dalle informazioni di linguaggio visivo, soprattutto perché prestano, in generale, meno attenzione ai volti e alle espressioni.

2.4 Il test della sound-induced flash illusion su soggetti autistici

I Disturbi dello Spettro Autistico (DSA) sono disordini dello sviluppo, di cui l'autismo rappresenta la forma più rilevante e severa. Recenti studi suggeriscono che vi siano diffuse anomalie nello sviluppo neurologico nei soggetti autistici che potrebbero essere correlate all'integrazione delle informazioni tra le diverse regioni del cervello, che in questi soggetti risulta deficitaria (Barnea-Goraly et al., 2004; Bertone, Mottron, Jelenic, & Faubert, 2003; Cherkassky, Kana, Keller, & Just, 2006).

L'integrazione audio-visiva è particolarmente rilevante nelle situazioni "sociali", quali la percezione delle emozioni e il linguaggio, entrambe deficitarie nell'autismo. Inoltre, esistono notevoli evidenze sperimentali sul fatto che l'integrazione sensoriale avvenga in aree cerebrali specifiche, le quali sono sensibili alle informazioni provenienti dalle varie modalità sensoriali (Stein & Meredith, 1993).

In uno studio del 2007, van der Smagt et al. ("*Brief report: can you see what is not there?*") sono stati usati pazienti affetti da autismo ad alto funzionamento, a cui è stato sottoposto il sound-induced flash illusion; i risultati sono stati poi confrontati con quelli ottenuti testando un secondo gruppo simile al primo per età e IQ-level, ma che non presentavano diagnosi di DSA.

L'integrazione multimodale anomala nei soggetti con DSA potrebbe tradursi in una diminuzione dell'occorrenza o della forza dell'illusione, rispetto al gruppo di controllo. D'altra parte, delle prestazioni in linea con quelle dei soggetti sani potrebbero indicare che i problemi di integrazione audio-visiva nei DSA possano provenire da un livello di elaborazione più elevato, come, ad esempio, quello cognitivo.

2.4.1 L'esperimento

Allo studio di van der Smagt hanno partecipato un gruppo di 15 soggetti autistici e un gruppo di controllo di 15 soggetti sani, simile al primo per età e IQ-level.

Come nell'esperimento della Shams che abbiamo ampiamente descritto in precedenza, sono stati utilizzati beep provenienti da altoparlanti e flash formati da un disco bianco proiettato su uno schermo nero.

Nello specifico, sono state presentate 12 condizioni differenti: 3 condizioni visive (1, 2 o 3 flash), combinate con 4 condizioni acustiche (0, 1, 2 o 3 beep). Talvolta il numero di flash e beep presentati corrispondeva, altre volte no. La condizione più significativa era quella di un unico flash, esattamente come nello studio della Shams.

La presentazione del disco era di 17 ms. Nel caso dei flash multipli, essi avevano una distanza temporale (SOA) di 50 ms. I beep, invece, avevano una durata di 9 ms, separati da una SOA di 50 ms. Infine, la distanza tra il primo beep e il primo flash era di 17 ms (Fig. 6).

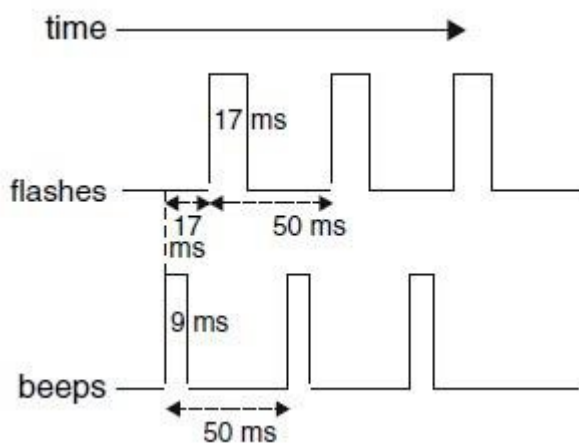


Fig. 6. Profilo temporale del caso in cui sono stati presentati 3 flash e 3 beep. Le altre condizioni erano identiche, ma presentavano un numero minore di flash e/o di beep.

I risultati di entrambi i gruppi sono presentati in Fig. 7, lo schema a destra per i DSA e quello a sinistra per il gruppo di controllo. Nel grafico il numero di flash percepiti è rappresentato in funzione del numero di beep presentati.

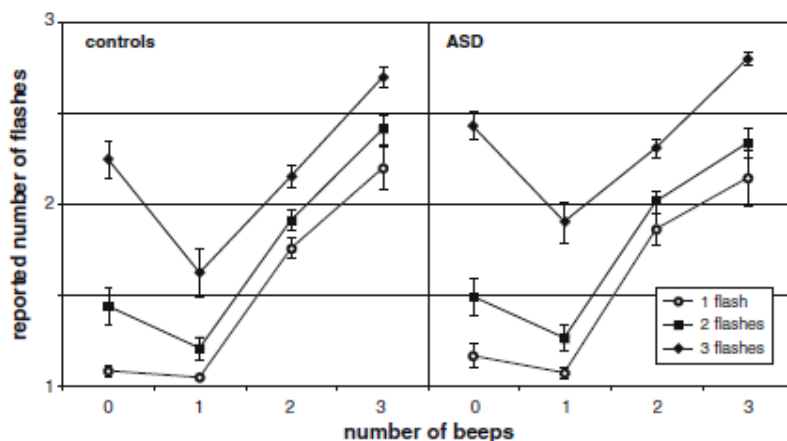


Fig. 7. Effetti del numero di beep sul numero di flash riportati, per soggetti sani (sinistra) e autistici (destra).

2.4.2 Conclusioni

L'obiettivo principale di questo studio era quello di mostrare se i soggetti con DSA mostrino un profilo di integrazione audio-visiva normale o anormale rispetto ai soggetti di controllo.

Se osserviamo la Fig. 7, è immediatamente evidente che un aumento del numero di beep o i flash comporta un aumento dei flash percepiti.

In particolare, si possono osservare due fenomeni diametralmente opposti. Il primo viene definito "fusion illusion" e avviene quando i soggetti tendono ad associare più flash ad un singolo beep, percependo quindi un numero inferiore di flash rispetto al numero effettivamente presentato. Il secondo fenomeno è definito "fission (o flash) illusion" e avviene quando i soggetti percepiscono un numero di flash superiore a quello realmente presentato, a causa di un numero maggiore di beep.

In generale, possiamo affermare che l'effetto illusorio è stato riscontrato in entrambi i gruppi studiati: il numero di beep presentati ha influenzato il numero di flash percepiti. Questi risultati indicano che i soggetti con DSA sono in grado di integrare le informazioni uditive e visive, probabilmente ad un livello iniziale di processazione sensoriale.

Ciò implica che, anche se sono state trovate anomalie nell'anatomia cerebrale coinvolta nell'integrazione delle informazioni multimodali (Barnea-Goraly et al., 2004), almeno alcuni collegamenti tra le aree cerebrali uditive e quelle visive sembrano funzionare in modo appropriato nei DSA.

2.4.3 L'aspetto temporale

Un lato della questione che ancora non è stato affrontato è l'importanza che ha l'aspetto temporale nell'integrazione multisensoriale. Esso, peraltro, è l'aspetto che più differenzia i soggetti sani e quelli autistici. Se, infatti, l'interazione tra due stimoli temporalmente vicini hanno un'elevata probabilità di essere percepiti e processati come se provenissero dalla stessa fonte (oggetto o evento), d'altro canto due stimoli che a livello temporale risultano molto separati devono essere processati come tali, dato che è molto più alta la probabilità che provengano effettivamente da oggetti o eventi distinti tra loro.

Questi aspetti temporali dell'integrazione risultano essere alterati nei soggetti autistici. Studi recenti, per esempio, hanno mostrato differenze nella capacità di combinare insieme informazioni provenienti da canali sensoriali diversi, come per l'effetto McGurk, il quale risulta essere meno evidente nei soggetti autistici (Stevenson et al., 2014).

Un ulteriore studio recente ha cercato di porre l'attenzione, sempre nell'ambito degli aspetti temporali, sulle eventuali differenze esistenti nella "temporal binding window" tra DSA e un gruppo di controllo formato da soggetti sani (TD, typical development), ovvero nella finestra di integrazione degli stimoli multisensoriali; per fare ciò è ricorso anch'esso al sound-induced flash illusion, in condizioni molto simili a quelle del secondo esperimento della Shams (Foss-Feig et al., 2010).

L'obiettivo era, quindi, quello di verificare se un beep fornito prima o dopo ad una coppia flash-beep sincrona potesse causare l'illusione di un ulteriore flash. I parametri fondamentali per tale studio sono la sincronizzazione (timing) con cui il secondo beep viene fornito e la distanza temporale tra esso e la coppia flash-beep.

Per ogni condizione di SOA (asincronia nell'insorgenza dello stimolo, stimulus onset asynchronies) è stata verificata l'occorrenza dell'illusione, e i risultati mostrano una forte presenza di tale fenomeno illusorio (Fig. 8).

Come si può notare dal grafico, per entrambi i gruppi in esame il fenomeno illusorio diminuisce all'aumentare della distanza temporale tra i due beep, e questo è sintomo della grande influenza che ha la SOA sull'insorgenza dell'illusione percettiva.

Ma si può altresì notare una profonda differenza tra i due gruppi: al crescere della distanza temporale tra i due segnali acustici, i soggetti sani mostrano un decrescere più rapido del fenomeno illusorio rispetto a quello dei soggetti autistici, nonostante anche per quest'ultimi il fenomeno tenda a diminuire. Ciò indica, quindi, che i soggetti autistici mostrano una

maggior propensione nel riportare l'illusione rispetto ai soggetti sani.

Questa differenza non è di poco conto; da un'analisi statistica è risultato che la finestra temporale dei DSA arriva ad essere doppia (da -300 ms a +300 ms) rispetto a quella del gruppo TD (da -150 ms a +150 ms).

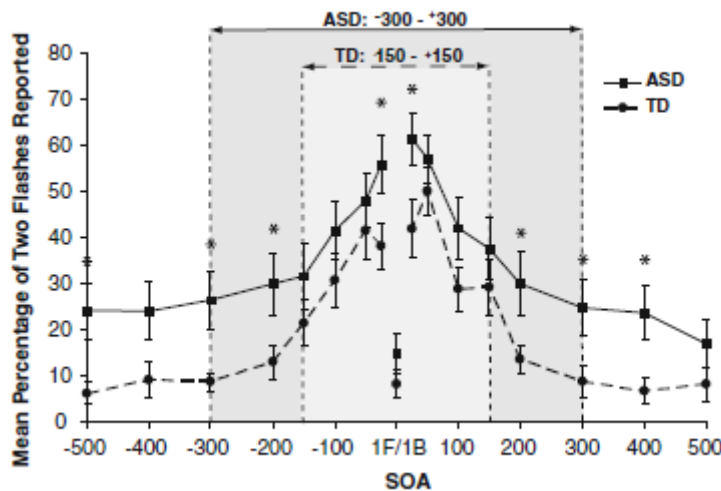


Fig. 8. Nel grafico viene riportata l'insorgenza del fenomeno illusorio in funzione della SOA.

Tale esperimento suggerisce dunque che i soggetti autistici presentino una temporal binding window maggiore, all'interno della quale essi sono portati a processare gli stimoli multisensoriali in maniera particolare.

Nel complesso, quindi, i risultati riportati dai differenti studi sopra descritti mostrano che l'integrazione multisensoriale degli stimoli audio-visivi avviene anche negli autistici, anche se esiste una sorprendente differenza nelle finestre temporali di integrazione tra questi soggetti e i soggetti sani.

3. LE BODY OWNERSHIP ILLUSIONS

Nel capitolo precedente abbiamo analizzato e discusso un test su un'illusione sensoriale specifica, la sound-induced flash illusion, e di come esso possa anche essere utilizzato per indagare alcune patologie nervose che coinvolgono la percezione multisensoriale, quali l'autismo, nella speranza che la ricerca si concentri su questi aspetti e continui ad indagare sulle cause, al fine di trovare, se non una cura, almeno un aiuto per le persone affette da questi disturbi.

In questo capitolo, invece, ci concentreremo su un altro genere di test percettivi, rivolti alla rappresentazione di noi stessi da parte del cervello.

Rappresentare a livello cerebrale il nostro corpo è un compito assai importante, poiché ciò è alla base, per esempio, della percezione dei nostri arti nello spazio, e ciò serve, a sua volta, per la programmazione dei movimenti.

Il senso di personificazione è un'esperienza complessa, che coinvolge più di un componente di sensazione e azione: noi siamo normalmente coscienti che "il corpo che abitiamo è il nostro" e che i nostri arti ci appartengono; questa sensazione onnipresente è detta *appartenenza del corpo*.

La consapevolezza dello stato del sistema motorio e il senso di essere la causa di un'azione sono aspetti cruciali del nostro senso di "essere noi stessi".

In circostanze normali, queste esperienze sono strettamente legate, se non addirittura inseparabili; in circostanze deficitarie, come ad esempio nella sindrome della somatoparafrenia, ciò avviene in modo diverso ed è estremamente difficile capire quali meccanismi la governano.

3.1 La somatoparafrenia e le illusioni di appartenenza

“Giuro su Dio, su quello che vuole, che io non... Uno dovrebbe saper riconoscere il proprio corpo, cos'è e cosa non è suo. Ma questa gamba, questa cosa..non mi convince, non la sento vera... E poi non mi sembra una parte di me.”

Un giovane Oliver Sacks, chiamato urgentemente per un consulto, arriva in una camera di ospedale e trova questo signore, terrorizzato e disgustato, che gli rivolge queste parole. Il

paziente afferma di aver trovato nel suo letto una gamba recisa; dopo essersi dato come unica spiegazione plausibile quella di uno stupido scherzo fatto da qualcuno appartenente allo staff dell'ospedale, scocciato, decide di buttarla giù dal letto. Ma insieme alla gamba cade dal letto anche lui. La gamba, infatti, era la sua.

Nel 1942, il neurologo Josef Gerstmann riportò i casi di due pazienti che avevano subito una lesione all'emisfero cerebrale destro, i quali presentavano credenze deliranti riguardanti il lato sinistro del loro corpo, caratterizzate da un'alterazione patologica del senso di appartenenza dei propri arti. Gerstmann suggerì il termine *somatoparafrenia* per indicare queste illusioni, o distorsioni, concernenti la percezione e i deliri riferiti agli arti o al lato colpito.

Il senso di noi stessi, il modo in cui il corpo vede se stesso, è mediato dalla propriocezione. Insieme al sistema visivo e al sistema vestibolare, essa collabora a restituirci il senso del corpo. Da un punto di vista neuroanatomico tale fenomeno integra componenti che vanno dal sistema nervoso periferico alla corteccia cerebrale passando per strutture sottocorticali. La parte sinistra della corteccia somatosensoriale sembra avere un ruolo maggiore per quanto riguarda l'orientamento del corpo, mentre quella destra è più legata a tutti quei fenomeni relativi alla consapevolezza corporea, e ciò spiegherebbe anche perché i sintomi si manifestano maggiormente nei pazienti che hanno riportato danni all'emisfero destro.

La somatoparafrenia, in altre parole, è una credenza delirante per cui un paziente sente che un suo arto paralizzato non appartiene realmente al suo corpo.

Il sintomo è spesso associato a un'altra sindrome neurologica, denominata Negligenza Spaziale Unilaterale (NSU).

La NSU rappresenta una condizione nella quale, in seguito a una lesione cerebrale, il paziente perde la capacità di esplorare l'emicampo visivo controlaterale alla lesione. Il paziente con neglect personale perde la consapevolezza del lato del corpo opposto a quello in cui si trova la lesione cerebrale (Bisiach, 1996).

Le caratteristiche principali della sindrome sono le seguenti:

1. mancata risposta agli stimoli presentati controlesionalmente;
2. significativa diminuzione dei movimenti di esplorazione verso lo spazio controlesionale;

3. presenza della sintomatologia anche in assenza di deficit sensoriali o motori: il comportamento non può essere spiegato esclusivamente in termini percettivi o motori.

Qualora la lesione sia meno grave, oppure nella fase subacuta, il soggetto può essere capace di interagire con gli oggetti presenti nello spazio controlesionale, ma solo in alcune circostanze; più precisamente, il paziente si dimostra consapevole di tali oggetti solo in assenza di stimoli distrattori collocati ipsilesionalmente. Se, viceversa, due stimoli sono presentati contemporaneamente, uno nell'emicampo sano e uno in quello controlesionale, quest'ultimo verrà negletto.

Nonostante la connotazione delirante, la somatoparafrenia non si associa ad altri sintomi o disturbi psichiatrici, e inoltre risulta essere, nella maggioranza dei casi, acuta e fluttuante. Ciò che viene percepito dal paziente può essere definito “non-belonging feeling” in quanto egli non riconosce come proprio l'arto plegico, che viene solitamente attribuito a un parente, a un familiare, al medico.

Tuttavia non sono stati ancora approfonditi i motivi per cui la somatoparafrenia è un sintomo così fluttuante: probabilmente entrano in gioco meccanismi di plasticità corticale.

Le manifestazioni di somatoparafrenia sono molteplici, e possono, certamente semplificando notevolmente, essere riassunte come segue:

1. sensazione di estraneità delle parti del corpo interessate, di separazione dal corpo del paziente, o entrambi (Anton 1893, Roth 1949);
2. convinzioni deliranti di rinnegamento delle parti del corpo colpite;
3. convinzioni deliranti che le parti del corpo interessate appartengano ad un'altra persona;
4. più complessi errori di identificazione deliranti della parte del corpo (Garcin et al. 1938).

Nel singolo paziente, questi sintomi possono variare nel tempo. Feinberg et al. (2005) stabiliscono una distinzione tra i semplici errori di identificazione (i pazienti rinnegano le parti del corpo, o attribuiscono la proprietà ad un'altra persona), che possono essere facilmente corretti quando l'errore viene indicato dall'esaminatore, e i veri e propri errori di

identificazione deliranti, che ostinatamente resistono alle dimostrazioni dell'esaminatore. Feinberg et al. (2005) suggeriscono un ruolo principale ai danni frontali nel determinare questi disturbi. I dati attuali suggeriscono che una vasta rete corticale fronto-temporo-parietale costituisce la base neurale del senso di appartenenza del corpo. Tuttavia, la questione è ancora oggetto di dibattito nella comunità scientifica.

Risulta evidente la stretta associazione tra neglect personale ed extra-personale da un lato, e la somatoparafrenia dall'altro, poiché la maggior parte dei pazienti mostra entrambe le manifestazioni della sindrome. Tuttavia, i casi sporadici in cui ciò non avviene possono portare ad ipotizzare che la somatoparafrenia possa essere indipendente dalla consapevolezza del lato controlesionale del corpo, valutata, per esempio, richiedendo ai pazienti di identificare le loro parti del corpo (Bisiach et al., 1986). Inoltre, una delle caratteristiche principali della somatoparafrenia è la negazione ostinata dell'appartenenza delle parti controlesionali del corpo, anche quando esse sono poste nel lato ipsilesionale, vale a dire, in una porzione non-neglect di spazio (Bisiach et al. 1990).

In ogni caso, l'interpretazione unitaria di certe sindromi non implica necessariamente che tutti i sintomi debbano sempre coesistere, né implica che possano essere spiegati da un singolo meccanismo implementato in una singola regione cerebrale o da una singola rete nervosa. Infatti, due disordini possono essere menomazioni funzionalmente indipendenti che accadono contemporaneamente, semplicemente perché il danno cerebrale tende a non essere sufficientemente specifico.

La più stretta associazione tra la somatoparafrenia e il neglect extra-personale può riflettere il fatto che una caratteristica principale della somatoparafrenia è la distinzione netta tra oggetti corporei (le parti del corpo), e quelli extracorporei (le parti del corpo di altre persone, o altri oggetti non-corporei, come le protesi), e ciò porta come risultato un rinnegamento delirante delle parti controlesionali del corpo, e, anche se meno frequentemente, una sostituzione dell'arto interessato con altri oggetti (Ehrenwald 1931), fino ad arrivare ad allucinazioni controlesionali.

È interessante notare lo studio sperimentale riportato da Daprati et al. (2000), nel quale si è osservato che il paziente, nel caso in cui gli fosse richiesto di riconoscere la sua mano eseguire un movimento tramite l'osservazione di uno schermo, non era in grado di portare a termine il compito, non riconoscendo come sua la mano rappresentata nel video.

In altri casi riportati, invece, pazienti somatoparafrenici sono stati in grado di riconoscere come propria la mano osservata nello schermo, quindi da un'osservazione indiretta, mentre hanno persistito nel giudicare di qualcun altro la propria mano osservata in maniera diretta (Jenkinson et al., 2011).

Analizzando le caratteristiche appena descritte riguardanti situazioni in cui sono presenti deficit cerebrali, si può facilmente capire l'importanza che ha assunto negli ultimi anni l'indagine nei confronti di tutte le illusioni sull'appartenenza del corpo, siano esse derivanti da danni cerebrali o comuni anche alle persone sane.

Una differenza fondamentale tra la percezione del proprio corpo e il corpo altrui o altri oggetti nell'ambiente esterno sta nel tipo di input sensoriali disponibili nel cervello. Nel processare il nostro corpo, il cervello ha accesso a una serie di informazioni – quali somatosensoriali, termocezioni, nocicezioni e segnali vestibolari – che non sono disponibili durante la percezione di altri oggetti o corpi.

Tuttavia, le prove su pazienti somatoparafrenici suggeriscono che l'elaborazione intatta sensoriale da modalità diverse può non essere di per sé sufficiente per la nascita della sensazione di possesso del corpo. Infatti, è stato proposto che la somatoparafrenia possa essere dovuta ad una perdita nel processo del segnale multisensoriale (Vallar and Ronchi, 2009). Pertanto, il senso di appartenenza del corpo dovrebbe essere considerato come il risultato di processi del cervello che integrano diversi segnali sensoriali nella percezione unitaria del “mio corpo”.

Un'altra differenza fondamentale riguarda il fatto che nel percepire il proprio corpo tutti i segnali sensoriali e motori che convergono nella percezione del “mio corpo” sono strettamente vincolati alle leggi fisiche.

Per esempio, quando si colpisce un pugno sul tavolo, la vista del contatto è sempre accompagnata alla sensazione tattile del colpo sulla nostra mano. Questo non accade nella percezione di oggetti esterni o di corpi altrui che, nonostante la natura della multisensorialità, non è soggetta a vincoli rigidi. Basta prendere in considerazione questo esempio: si vede un cane in un parco e si sente l'abbaiato all'incirca nello stesso momento e proveniente dalla stessa direzione. Anche se è probabile che il cane visto sia lo stesso che abbaia, vi è anche la possibilità che vi sia un secondo cane appena dietro l'albero. Così, mentre nella percezione di eventi esterni possiamo contemplare la possibilità di fonti

diverse, nella percezione del nostro corpo ciò non accade, poiché tutti i segnali multimodali coinvolti provengono dalla stessa fonte: il corpo fisico.

A causa del fatto che i segnali multimodali relativi al corpo sono strettamente legati insieme e non indipendenti, è difficile investigare sperimentalmente con i metodi adottati nella ricerca multisensoriale. Per contro, quando si studia la percezione multisensoriale degli oggetti esterni, inclusi i corpi altrui, gli esperimenti di solito coinvolgono la presentazione simultanea di segnali sensoriali che sono indipendenti. Ciò permette l'introduzione di ritardi tra le occorrenze degli stimoli, di presentarli da differenti posizioni, o anche di manipolare le informazioni in modo che esse possano riferirsi allo stesso contesto o meno.

Pertanto, è possibile indagare su come la percezione multisensoriale sia influenzata dalle relazioni spaziali, temporali e semantiche tra gli stimoli manipolati.

Lo stesso metodo non può essere applicato direttamente allo studio del possesso del proprio corpo. Per esempio, non è possibile introdurre un ritardo temporale tra il vedere il pugno colpire il tavolo e sentire la sensazione nella mano, se non attraverso l'uso di dispositivi, quali una videocamera o uno schermo.

Date le limitazioni inerenti allo studio del corpo fisico, le ricerche sperimentali sono state ampiamente condotte sfruttando le illusioni di possesso del corpo (BOIs, *Body Ownership Illusions*). In queste illusioni, i soggetti hanno percepito oggetti non-corporei (ad es. elementi artificiali) come fossero appartenenti al proprio corpo, quando essi sono stati presentati con gli stimoli cross-modali applicati alla parte reale ma nascosta del corpo e alla sua controparte falsa. In questo modo, le BOIs offrono un potente strumento sperimentale per esaminare il modo in cui il senso di appartenenza al corpo emerge dall'elaborazione multisensoriale operata dal cervello.

Nel seguito del capitolo ci concentreremo su questo genere di illusioni di appartenenza corporea, distinguendo queste dalle altre classi di illusioni corporee che, insieme alle precedenti, hanno fornito spunti essenziali su come la percezione del proprio corpo è costruita attraverso l'informazione multisensoriale e sensorimotoria.

3.2 Le BOIs nello studio della percezione del proprio corpo

Le Body Ownership Illusions (BOIs), o illusioni corporee, si riferiscono a quei fenomeni psicologici in cui la percezione del proprio corpo devia in maniera importante dalla configurazione del corpo fisico, ad esempio in termini di dimensioni, posizione o proprietà. Poiché la loro induzione è ottenuta attraverso la stimolazione multisensoriale e/o sensorimotoria, le illusioni corporee forniscono informazioni essenziali su come la propria percezione corporea sia costruita in tempo reale, sulla base degli stimoli disponibili nel cervello.

Una classe interessante di illusioni corporee è quella delle *body distortions illusions*, in cui le persone possono percepire che la dimensione o la posizione della parte (o delle parti) del corpo sono cambiate drasticamente, senza soddisfare necessariamente i normali vincoli anatomici. Un ben noto esempio è la *Pinocchio illusion* (Fig. 1A). Il soggetto, bendato, riceve una vibrazione sul bicipite mentre tocca la punta del naso col dito. L'estensione illusoria del braccio genera l'illusione che il suo naso, il suo dito, o entrambi, siano allungati. In alternativa, un'illusione di distorsione simile può essere indotta attraverso correlazioni temporali tra informazione propriocettiva e input tattile, come nella *phantom nose illusion* (Ramachandran and Hirstein, 1998) (Fig. 1B). Il ricercatore muove il dito di un soggetto bendato per fargli toccare il naso di un secondo soggetto, mentre contemporaneamente tocca il naso del primo soggetto. Dato che il movimento del soggetto e il tocco del naso del secondo soggetto sono sincroni rispetto alla percezione del tocco del proprio naso, il soggetto sperimenta l'illusione di toccare il suo naso allungato.

Indipendentemente dal metodo impiegato, le illusioni distorsive dimostrano che il cervello computa la posizione e la forma del corpo in maniera dinamica e flessibile, senza la necessità di soddisfare i vincoli anatomici del corpo umano.

Un'altra notevole classe di illusioni sono indotte sperimentalmente *out of body experiences*, in cui le persone percepiscono loro stesse come dislocate rispetto all'origine spaziale della prospettiva visiva (Lenggenhager et al., 2007). Uno dei primi studi su queste illusioni è stato svolto attraverso l'uso di uno specchio (Stratton, 1896). Nelle recenti sperimentazioni, i partecipanti sono stati posti davanti a un display, che mostrava la loro schiena ripresa da una certa distanza. Attraverso questo espediente, i partecipanti sono stati in grado di vedere la loro schiena dalla prospettiva di una terza persona (3 PP_third person perceptive), in

contrasto con la normale prima prospettiva (1 PP) che abbiamo del nostro corpo. Attraverso l'applicazione di una stimolazione tattile sulla schiena del soggetto mentre esso vede lo strumento toccare la sua schiena attraverso lo schermo, i soggetti possono sperimentare una deriva illusoria della loro posizione spaziale verso il corpo visto nel video (Lenggenhager et al., 2007) (Fig. 1C). Altri risultati sperimentali mostrano che la posizione percepita di sé può essere separata da quella del corpo fisico, in modo simile a quanto riferito da pazienti sottoposti a esperimenti di origine neurologica “out of body” (Blanke and Mohr, 2005). Questa classe di illusioni fornisce quindi un importante punto di riferimento per lo studio del ruolo dell'integrazione multisensoriale per ciò che concerne la coscienza di sé e della propria posizione.

A differenza della distorsione del corpo e delle illusioni esterne al corpo, le illusioni BOIs si riferiscono alla percezione illusoria di oggetti non corporei (ad esempio, arti artificiali) come parte del proprio corpo e come fonte delle sensazioni corporee ad esso collegate, come il tatto. Ad esempio, si può arrivare a pensare che la mano di un manichino sia la propria mano, e si hanno forti risposte fisiologiche quando viene visto l'accoltellamento della stessa (Ehrsson et al., 2007). Botvinick e Cohen (1998) fornirono la prima relazione sull'illusione della mano di gomma (RHI), in cui i soggetti percepirono una mano di gomma come fosse la propria. In questa configurazione sperimentale, i soggetti hanno la loro mano nascosta alla vista, mentre possono vedere una mano di gomma collocata in una posizione simile. Se la mano di gomma viene colpita contemporaneamente a quella reale, la maggioranza dei partecipanti affermerà di percepire il tocco come se provenisse dalla mano di gomma, e quindi come se essa fosse parte del proprio corpo (Fig. 1D).

A parte le derive propriocettive, la RHI ha dimostrato di diminuire la temperatura e rallentare l'elaborazione dell'input tattile rispetto alla mano reale (Moseley et al, 2008) e di innescare risposte automatiche dei partecipanti quando vedono la mano di gomma "in pericolo" (Armel e Ramachan-Dran, 2003).

È infatti interessante notare che è stato dimostrato che le aree cerebrali associate all'ansia e alla consapevolezza intero-cettiva (ossia a livello viscerale) si attivano quando l'arto falso è in pericolo e ad un livello simile a quello che si attiverebbe se fosse in pericolo la mano reale (Ehrsson et al., 2007).

L'induzione dell'illusione può anche essere dimostrata in assenza di qualsiasi input visivo. Nella cosiddetta "RHI somatica", ai partecipanti bendati, è stato mosso passivamente il dito indice sinistro al fine di toccare una mano di gomma, mentre il ricercatore toccava la mano destra del soggetto. Dopo alcuni secondi per correlare le informazioni tattili e propriocettive, i partecipanti hanno riportato di aver toccato la propria mano (Ehrsson et al., 2005a). Estendendo i risultati alle altre illusioni corporee, le BOIs rivelano che il nostro cervello calcola in modo dinamico quali sono le parti del nostro corpo sulla base delle informazioni multisensoriali e sensorimotorie disponibili.

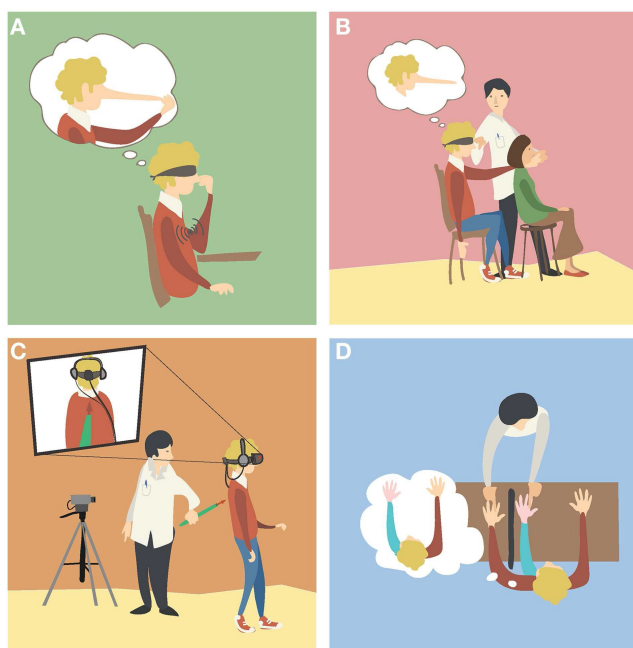


Fig. 1. Esempi di body illusions.
(A) The Pinocchio illusion.
(B) The phantom nose illusion.
(C) Out of body illusion.
(D) The rubber hand illusion.

3.3 Trigger multimodali nelle BOIs

La stimolazione delle illusioni di appartenenza del corpo (BOIs) è stata mostrata in diversi studi sperimentali che hanno fornito stimoli crossmodali differenti (Fig. 2). Questi studi possono essere classificati in base ai principali stimoli multisensoriali forniti, cioè, visivo e tattile, visivo e propriocettivo o, infine, visivo e motorio. Di seguito analizzeremo l'importanza della corrispondenza spaziale e temporale per ogni coppia di stimoli crossmodali, nonché il ruolo dell'informazione semantica che la vista degli oggetti non corporei comporta.

3.3.1 Stimoli visuo-tattili

L'esperienza del nostro corpo è plasmata in maniera importante dai segnali visivi e somatosensoriali. Uno dei principali contributi di integrazione visuo-tattile riguarda le informazioni sugli episodi di contatto tra il nostro corpo e l'ambiente circostante. Per esempio, quando una mosca si posa sulla nostra mano, sia la vista che il tatto informano il cervello circa il tempo e la posizione del contatto. Ma anche prima che si verifichi il contatto puramente fisico, la visione è in grado di fornire informazioni in anticipo circa il dove e il quando la mosca si poserà, e ciò ci permette di agire di conseguenza.

Lo studio dei neuroni bimodali visuo-tattili ha confermato l'importanza che essi rivestono nel percepire il nostro corpo e lo spazio circostante; essi sono stati usati per costruire una rappresentazione ingrandita e flessibile dello spazio che circonda il corpo (spazio peripersonale) che serve principalmente per guidare i nostri movimenti (Lloyd et al, 2003; Gentile et al., 2011).

Per quanto riguarda lo studio dell'integrazione visuo-tattile attraverso le BOIs, il primo test sulla mano di gomma (RHI) è stato indotto attraverso l'applicazione dello stimolo tattile sia alla mano vera nascosta che alla mano falsa visibile (Botvinick e Cohen, 1998) (Fig. 2A, D).

L'evidenza sperimentale ha dimostrato che la RHI è indotta quando la mano reale e la controparte falsa sono toccate allo stesso tempo e in regioni omologhe. Ad esempio, quando sono introdotti disallineamenti spazio-temporali tra il tocco visto e percepito, la RHI era significativamente inibita (Slater et al, 2008).

Per contro, altri studi hanno dimostrato che quando il corpo falso è realistico e ha sovrapposizioni nello spazio con la vera controparte corporea, risultati positivi possono essere segnalati anche in presenza di stimolazione visuotattile asincrona (Maselli e Slater, 2013).

Per ciò che concerne l'importanza dell'allineamento temporale, possiamo affermare che, quando sono stati introdotti ritardi inferiori a 300 ms tra le stimolazioni, i partecipanti hanno percepito il tocco sulla propria mano anche se in realtà la stimolazione era avvenuta sulla mano di gomma, mentre per i ritardi più grandi queste sensazioni illusorie sono state significativamente attenuate (Shimada et al., 2014).

Oltre all'accoppiamento temporale, una congruenza spaziale è risultata essere essenziale per l'induzione dell'illusione. Accarezzare la mano reale e quella di gomma in sincronia

temporale, ma in luoghi diversi (ad esempio, l'indice contro il mignolo oppure il palmo contro l'avambraccio) ha eliminato l'illusione (Kammers et al, 2009; Limanowski et al. 2013), a dimostrazione che la corrispondenza temporale da sola non è sufficiente perché avvenga la RHI.

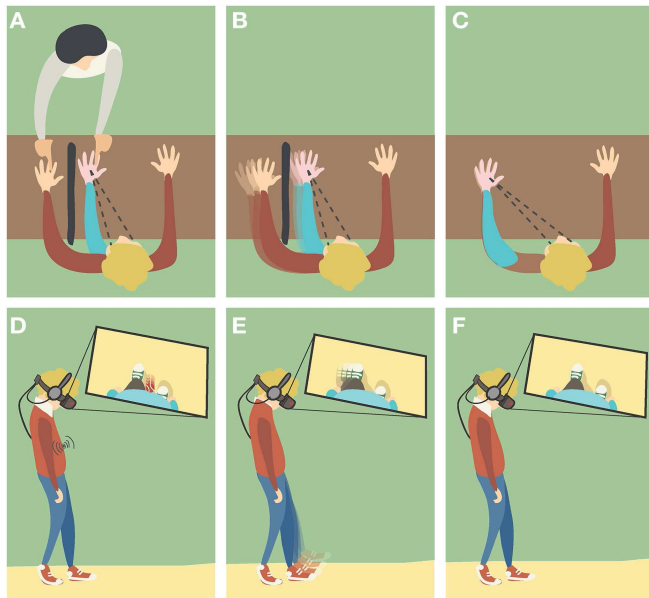


Fig. 2. Metodi differenti di induzione delle BOIs.

(A, D) attivazione visuo-tattile: il soggetto sta guardando il corpo falso posto in una posizione plausibile per essere toccato, durante la ricezione di una stimolazione tattile sincrona nella controparte reale che rimane fuori dalla visuale.

(B, E) attivazione visuo-motoria: il soggetto sta eseguendo movimenti con la mano reale che rimane fuori dal campo visivo, mentre guarda la controparte falsa compiere un movimento sincrono.

(C, F) attivazione visuo-proprioceppiva: il partecipante sta guardando il falso corpo posto in una posizione sovrapposta con la controparte reale che rimane fuori dalla visuale.

3.3.2 Stimoli visuo-motori

La nostra percezione del corpo è fortemente determinata dai nostri movimenti, dal momento che questi forniscono informazioni che facilitano la distinzione tra il sé e l'altro. Ad esempio, quando un pianista suona il pianoforte a quattro mani con un collega, egli usa, tra gli altri dati, le sue informazioni motorie insieme al feedback visivo, tattile e proprioceppivo, per riconoscere quale delle due mani destre che vede sia la sua.

L'esperienza di muoversi e di agire nello spazio comporta un contenuto estremamente ricco di informazioni relative al corpo, che va ben al di là dell'integrazione multisensoriale di due o più modalità sensoriali. Ciò può essere meglio compreso considerando la distinzione tra i movimenti passivi e quelli attivi. In un movimento passivo non vi è alcuna intenzione motoria: una forza generata esternamente sposta il nostro corpo e, di conseguenza, una serie di informazioni sensoriali, tra cui la visione e la proprioceppività, vengono aggiornate.

Al contrario, quando si esegue un movimento attivo, il cervello genera e implementa un piano di movimento che si basa su simulazioni interne del nostro sistema motorio: il piano motorio viene eseguito, monitorato e corretto confrontando la copia dell'efferenza con il feedback sensoriale generato.

L'implicazione di simulazioni interne del corpo rende il trattamento dei segnali visuo-motori durante il movimento attivo più ricco di contenuti informativi relativi al nostro corpo. È stato dimostrato, per esempio, che si riesce meglio a discriminare input visuo-motori sincroni contro input asincroni durante l'esecuzione di movimenti attivi rispetto ai passivi (Shimada et al., 2010). Il monitoraggio visivo e il riconoscimento delle proprie azioni contribuisce in modo critico anche al senso di possesso del corpo e svolge un ruolo importante nella discriminazione tra sé e l'altro (Jeannerod, 2004).

Diversi studi hanno dimostrato che le BOIs sono indotte quando entrambi i corpi, reale e falso, muovono parti del corpo omologhe allo stesso tempo (Fig. 2B, E). Ad esempio, in uno studio ai partecipanti è stato chiesto di eseguire un movimento guardando uno schermo in cui una mano, propria o del ricercatore, era visualizzata nella stessa posizione della loro mano nascosta. Quando è stato chiesto loro di indicare se la mano vista era la propria, le prestazioni dei partecipanti hanno mostrato un alto tasso di auto-attribuzione errata quando vedevano la mano del ricercatore fare il medesimo movimento della propria (Daprati et al., 1997). In una configurazione simile, i partecipanti hanno visto una mano virtuale compiere un movimento uguale a quello della propria mano nascosta, sotto diversi orientamenti spaziali e ritardi temporali. Quando viene chiesto di decidere se il movimento visualizzato corrisponde al proprio, i soggetti hanno commesso in maniera significativa più errori nel riconoscere loro stessi per ritardi temporali inferiori a 150 ms e per deviazioni angolari minori di 15° (Franck et al., 2001). Una variante della stessa impostazione è stata utilizzata per capire come l'azione e gli stimoli visivi sulla configurazione spaziale del corpo vengono utilizzati nei processi di riconoscimento del proprio corpo. I risultati hanno mostrato che, quando le mani eseguono movimenti diversi, i partecipanti sono quasi sempre in grado di riconoscere la propria mano. Al contrario, quando le informazioni visuo-motorie sono ambigue (stessi movimenti) i partecipanti sono stati meno precisi e le loro risposte sono state influenzate dalla configurazione spaziale in cui le mani erano visualizzate (van den Bos e Jeannerod, 2002).

Ciò indica che, analogamente a quanto già affermato per gli stimoli visuo-tattili,

l'induzione delle BOIs con stimoli visuo-motori dipende criticamente dalla congruenza spazio-temporale tra i movimenti sentiti e quelli visti.

3.3.3 Stimoli visuo-propriocezionali

La nostra percezione del corpo è fortemente influenzata dalle informazioni visive e propriocezionali. Il principale contributo di integrazione visuo-propriocezionale si basa sulla definizione che il nostro corpo ha nello spazio. Entrambe le modalità informano il cervello su dove una parte specifica del corpo si trova in un determinato momento.

L'evidenza sperimentale ha dimostrato che quando la visione e la propriocezione forniscono informazioni contraddittorie, la posizione percepita delle nostre parti del corpo può deviare significativamente da quella veritiera. Ad esempio, quando la mano è vista in una posizione non veritiera attraverso specchi o prismi, i soggetti percepiscono la loro mano in un unico luogo, da qualche parte tra il visibile e la posizione percepita, e più vicina a quella visibile (van Beers et al., 1999). Questa cattura visiva della propriocezione ha mostrato di influenzare le prestazioni di raggiungimento del movimento e di svolgere un ruolo fondamentale nel monitoraggio dell'esecuzione del movimento (Bagesteiro et al., 2006; Boulinguez & Rouhana, 2008).

Nel complesso, evidenze sperimentali suggeriscono che le informazioni visive e propriocezionali contribuiscono in modo significativo alla percezione del proprio corpo e, più in particolare, nel determinare la percezione della sua configurazione spaziale.

Le manipolazioni della posizione e dell'orientamento del corpo falso, relativamente al corpo vero, hanno permesso di indagare fino a che punto le discrepanze visuo-propriocezionali influenzino l'induzione e la forza delle BOIs.

Inoltre, è stato possibile verificare se esiste una stretta corrispondenza tra le configurazioni spaziali viste e percepite, e se possa innescare le BOIs senza ulteriori stimolazioni crossmodali (ad esempio, visuo-tattili o visuo-motorie) (Fig. 2C, F). Finché la mano finta è stata posta vicino alla linea mediana del corpo, la RHI è stata indotta in presenza di stimolazione visuo-tattile con congruenza spazio-temporale, per diversi disallineamenti testati tra le posizioni delle mani reali e false, sia sul piano orizzontale che su quello verticale (Zopf et al., 2010), suggerendo che la distanza tra le mani non è un fattore cruciale per l'insorgenza dell'illusione.

È stato invece scoperto che la distanza orizzontale tra le mani influenza in modo significativo i risultati soggettivi per il riconoscimento delle condizioni di asincronia visuo-tattile (Zopf et al, 2010): i punteggi positivi sono significativamente più elevati quando le due mani sono più vicine di 15 cm rispetto al caso in cui la distanza sia più grande (45 cm). Rispetto al piano verticale, invece, non sono state riscontrate differenze significative nei punteggi quando la distanza tra la mano reale e quella falsa è tra i 12 e i 27,5 cm, pur mantenendo entrambi alla stessa altezza, vicino alla linea mediana dei soggetti.

Mentre in tutti gli studi già menzionati la mano vera e quella falsa differivano per posizione, ma erano per lo più allineate in termini di orientamento, altri esperimenti hanno dimostrato che le BOIs possono essere indotte in presenza di un ulteriore disallineamento sull'orientamento, a condizione che il corpo falso sia stato posto in una posizione anatomicamente plausibile. Ad esempio, ruotare la mano sinistra finta di 44° in senso orario non impedisce ai partecipanti di sperimentare il contatto fisico come se provenisse dalla posizione della mano di gomma (Butz et al, 2014).

In contrasto con il gran numero di studi che utilizzano input visuo-tattili, ce ne sono stati meno che hanno studiato se la congruenza spaziale dei segnali visuo-proprioceettivi possa essere di per sé sufficiente per l'induzione delle BOIs.

Per quel che riguarda gli stimoli visuo-motori, invece, si può affermare che, accanto alle informazioni sui tempi di movimento e le parti del corpo interessate, essi forniscono informazioni su come la posizione relativa e l'orientamento dei corpi, reali e falsi, cambino nel tempo. Alcuni degli studi che hanno impiegato la stimolazione visuo-motoria di parti artificiali del corpo hanno seguito il protocollo RHI e hanno presentato l'arto artificiale vicino alla linea mediana e in posizione laterale rispetto a quella reale (Sanchez-Vives et al, 2010), mentre altri studi hanno messo l'arto finto sopra quello vero (Kalckert e Ehrsson, 2012). Per quanto riguarda la distanza orizzontale, nessuna differenza significativa è stata trovata nei rapporti soggettivi quando lo spostamento dei due arti avveniva in modo sincrono (Yuan e Steed 2010), mentre per quanto riguarda la distanza verticale, solo le piccole distanze, ad esempio 12 cm, hanno consentito l'induzione di una RHI robusta sotto una stimolazione visuo-motoria congruente (Kalckert e Ehrsson, 2014).

Il fatto che sia stato trovato che le BOIs si verificano quando si ha una visione statica di

un corpo virtuale spazialmente coincidente, altamente realistico, e sotto stimolazione visuo-tattile asincrona, suggerisce che gli input visuo-proprioceettivi congruenti da soli potrebbero essere sufficienti per indurre l'illusione, e ciò si può sostenere anche nel caso di discrepanze visuo-tattili (Maselli e Slater, 2013). Inoltre, studi in cui la BOI è stata valutata in condizioni "puramente visive", suggeriscono che quando non abbiamo stimolazione visuo-tattile né visuo-motoria, la posizione relativa e l'allineamento del corpo vero e di quello falso contano.

Tutti questi risultati rivelano che la congruenza spaziale degli stimoli visuo-proprioceettivi non è necessaria per l'evidenza dell'illusione, a condizione che il corpo falso sia visto in una configurazione anatomica plausibile e in presenza di una stimolazione visuo-motoria o visuo-tattile congruente. In altre parole, per gli stimoli visuo-proprioceettivi è fondamentale la congruenza temporale, ma può essere considerata superflua quella spaziale.

Tuttavia, diversi gradi di disallineamento spaziale modulano significativamente sia l'intensità che il tempo di inizio dell'illusione.

3.4 La somatoparafrenia e i test sulle BOIs

Tutti i dati discussi fanno riferimento a casi di pazienti sani. Recenti studi, invece, hanno iniziato ad esplorare l'induzione delle BOIs sui pazienti somatoparafrenici (Jenkinson et al., 2013).

Si è scoperto, infatti, attraverso l'illusione che utilizza uno specchio, che esiste una differenza a seconda della prospettiva dell'osservatore: nel caso in cui al paziente è stato chiesto di riconoscere il proprio arto con un'osservazione diretta (1 PP, first-person visual perspective), egli non è stato in grado di portare a termine il compito; al contrario, nel caso di un'osservazione indiretta (3 PP, third-person visual perspective), come ad esempio attraverso uno specchio, il paziente è riuscito a identificare come proprio il braccio che stava vedendo.

Inoltre, il suddetto studio ha mostrato per la prima volta come questo cambiamento sia stato modificato con una semplice manipolazione di attenzione; in particolare, anche se il paziente aveva mostrato un rifiuto quasi totale del braccio durante la visualizzazione diretta, ha mostrato un notevole miglioramento del senso di proprietà degli arti (arrivando

al 100% di riconoscimento) durante la visualizzazione di se stesso in uno specchio, dirigendo quindi la sua attenzione allo spazio extra-personale. In questo caso specifico, lo spazio extra-personale è rappresentato dallo spazio attorno/vicino allo specchio; infatti, nel caso in cui il paziente ha diretto la sua attenzione allo spazio peri-personale, ovvero guardando vicino al proprio corpo, ha mostrato un miglioramento meno pronunciato (50%). Quindi, dirigere l'attenzione selettiva a diversi settori di spazio riesce a suscitare un cambiamento nel senso di proprietà del corpo.

Una possibile spiegazione del fenomeno è che quando ci osserviamo allo specchio vediamo come la nostra immagine si presenta dall'esterno (3PP), quindi il paziente può per un momento vedere i suoi arti come collocati nello spazio peri-personale. Focalizzare l'attenzione sullo specchio può quindi migliorare ulteriormente l'elaborazione dei segnali visivi relativi al proprio corpo, che può a sua volta migliorare la body ownership. La misura in cui questo effetto comporta una competizione con l'integrazione delle diverse rappresentazioni del corpo, o una più precisa elaborazione dei segnali di predizione degli errori rimane oggetto di studio (Fotopoulou, 2012).

Lungo tutta la durata della nostra vita, siamo soliti osservare il nostro corpo attraverso gli specchi; questa è forse un'ulteriore ragione per cui anche i pazienti somatoparafrenici sono in grado di riconoscersi da un'osservazione 3PP, mentre falliscono in quella 1PP.

Inoltre, utilizzando il test della RHI si può notare come il compito viene portato a termine con più frequenza quando l'attenzione spaziale è stata rivolta alla mano di gomma, mentre spostare l'attenzione alla propria mano ha diminuito la frequenza del riconoscimento.

Quest'ultimo dato suggerisce che focalizzare l'attenzione nello spazio peri-personale potrebbe migliorare la propria rappresentazione del corpo in 1PP.

I dati appena discussi sono, purtroppo, soggetti a notevoli limitazioni. La raccolta di dati sistematici sui pazienti somatoparafrenici, infatti, presenta diverse difficoltà; per esempio, la quantificazione e la caratterizzazione dei dati illusori (deliranti) sono problematiche data la variabilità del comportamento tra i vari pazienti, e anche nello stesso paziente, e la variabilità dell'intensità dei sintomi, che rende ogni caso diverso dagli altri. Tuttavia, descrizioni dettagliate di casi specifici, in combinazione con manipolazioni sperimentali attentamente controllate e ripetute, rimangono importanti per caratterizzare la variabilità clinica di questa rara sindrome, e costituiscono la base per la comprensione dei meccanismi soggettivi e neurali del senso di appartenenza del corpo.

4. I MODELLI DI ANALISI DEI SEGNALI SENSORIALI

Nei precedenti capitoli abbiamo cercato di analizzare diverse tipologie di test basati sulle illusioni sensoriali che prendono in esame l'integrazione sensoriale audiovisiva e la sensazione di appartenenza del corpo, puntando l'attenzione su alcune forme deficitarie di tali fenomeni. Come abbiamo già affermato, l'integrazione multisensoriale opera a diversi livelli cerebrali: in strutture sottocorticali (in particolare il collicolo superiore), in corteccie associative di livello superiore (ad esempio, regioni parietali posteriori), e anche nelle aree corticali primarie. A causa dei complessi meccanismi non lineari dei fenomeni integrativi del cervello, uno strumento chiave per la loro comprensione è rappresentato dai modelli neuro-computazionali.

Nel presente capitolo andremo a riassumere alcune tecniche di analisi dei dati che sono state approfondite negli ultimi decenni e che vengono utilizzate dalle moderne neuroscienze.

4.1 Il modello bayesiano

Gli eventi percettivi assumono importanza nel momento in cui entrano in relazione con un contesto complesso come quello dell'ambiente esterno. Per utilizzare le informazioni sensoriali in modo efficiente, al fine di dare giudizi e guidare le nostre azioni, il sistema nervoso è costantemente impegnato nel combinare le informazioni incerte provenienti dalle diverse modalità sensoriali. Questa incertezza deriva da una serie di fattori, tra cui il rumore ambientale o l'intrinseca variabilità neurale. Numerosi studi hanno analizzato il modo in cui le persone utilizzano e combinano gli stimoli per la percezione (Warren et al., 1981), mettendo in evidenza la ricca serie di effetti che si verificano nella percezione multimodale. Negli ultimi decenni, anche grazie agli sviluppi della statistica e dell'intelligenza artificiale, i ricercatori hanno cominciato ad applicare i concetti della teoria delle probabilità ai problemi di percezione biologica e di combinazione degli stimoli (Liu et al., 1995). Una sorprendente osservazione di questo lavoro è che i metodi bayesiani si sono dimostrati efficaci nella costruzione di teorie computazionali per la percezione e il controllo sensorimotorio; inoltre, la psicofisica sta fornendo un crescente numero di evidenze che ci portano ad affermare che i processi percettivi umani sono "bayes-ottimali", ovvero hanno un comportamento in linea con il teorema di Bayes, il quale viene impiegato

per calcolare le probabilità della causa che ha scatenato l'evento che si vuole analizzare. Nell'accezione riguardanti i processi cerebrali, esso porta ad affermare che il cervello rappresenta le informazioni sensoriali in maniera probabilistica, in forma di distribuzioni di probabilità; i modelli bayesiani presuppongono che il cervello esegua una integrazione ottimale, calcolando la probabilità a posteriori dell'evento dato un insieme di informazioni incerte.

Questa osservazione, insieme al lavoro comportamentale e computazionale su cui si basa, ha implicazioni fondamentali per le neuroscienze, in particolare nel modo in cui concepiamo i processi neurali e la natura delle rappresentazioni neurali delle variabili percettive e motorie.

4.1.1 Il Teorema di Bayes

I modelli bayesiani, in generale, rappresentano un concetto di probabilità con una concezione soggettivista, poiché vanno ad esprimere il “grado di affidabilità” che lo sperimentatore ha nel fatto che si possa o meno verificare un dato evento (nel nostro contesto, una percezione sensoriale), e dipende quindi dal grado di conoscenza di tale esperienza che egli ha acquisito in una situazione precedente. Tale incertezza viene codificata da una funzione di distribuzione di probabilità.

Il processo di apprendimento, nel contesto bayesiano, consiste nell'aggiornamento delle condizioni iniziali, definite “prior”, disponibili prima rispetto ai dati x del campione e che dipendono dalle esperienze precedenti, riguardo al parametro d'interesse dell'evento in esame, definito ϑ e rappresentato dalla distribuzione di probabilità $p(\vartheta)$, alla luce dei dati osservati. Una volta registrati i dati x , si otterrà una nuova distribuzione di probabilità per ϑ , detta distribuzione di probabilità a posteriori $p(\vartheta|x)$. I parametri ϑ , quindi, non sono più fissi e incogniti, ma vengono considerati come variabili, mentre i dati x risulteranno costanti, e la loro analisi potrebbe apportare delle modifiche ai parametri ϑ .

Il teorema di Bayes permette di tenere conto di opinioni e conoscenze, eventualmente esistenti, a priori sul fenomeno oggetto di studio. La formulazione del teorema è

$$p(\vartheta|x) = \frac{p(\vartheta)p(x|\vartheta)}{p(x)}$$

dove si evidenziano le tre distribuzioni fondamentali:

- $p(\vartheta)$ è la probabilità a priori, che descrive l'informazione che lo sperimentatore possiede del fenomeno prima ancora di aver raccolto i dati, e ciò si basa normalmente sulle esperienze precedenti;
- $p(x|\vartheta)$ è detta funzione di verosimiglianza, ed esprime la distribuzione di probabilità che il ricercatore, messo a conoscenza del parametro di interesse ϑ , assegnerebbe ai dati campionati;
- $p(\vartheta|x)$ è la distribuzione a posteriori, che esprime la probabilità che lo sperimentatore, una volta analizzati i dati, aggiorni le proprie conoscenze a priori.

4.1.2 I modelli bayesiani applicati alla percezione multisensoriale

Ogni sistema che effettua la stima delle variabili non osservabili sulla base delle variabili osservate esegue un'inferenza.

In natura, ci sono quasi sempre più suoni che raggiungono le nostre orecchie nello stesso momento. Ad esempio, mentre si cammina nella foresta, potremmo ascoltare contemporaneamente i cinguettii di un uccello, il suono di un ruscello, i nostri passi, il vento che soffia e così via. Le orecchie ricevono un'onda complessa che è la somma di più onde sonore prodotte da fonti diverse. Per dare un senso agli eventi e agli oggetti nell'ambiente, il sistema uditivo deve dedurre quali componenti sono stati causati dalla stessa fonte e dovrebbero, quindi, essere integrati, e quali componenti sono stati causati da fonti diverse e dovrebbero essere tenuti separati.

Allo stesso modo, le informazioni visive sono sempre frammentate e rumorose, e il sistema visivo deve determinare quali parti dell'immagine retinica corrispondono allo stesso oggetto e devono essere raggruppate insieme e quali parti corrispondono a oggetti diversi e non devono essere raggruppate. Inoltre, per determinare come interpretare alcuni segnali, il sistema deve dedurre che tipo di fonte ha dato origine al segnale. Questi tipi di problemi di inferenza esistono in ogni modalità sensoriale.

In qualsiasi momento, un individuo riceve tipicamente più segnali sensoriali dalle diverse modalità e deve determinare quali di questi segnali originino dallo stesso oggetto e debbano quindi essere combinati, e quali scaturiscano da oggetti diversi e non debbano

essere integrati.

Recentemente, per interpretare gli esperimenti sull'integrazione multisensoriale, sono stati adottati modelli bayesiani. Questi modelli mirano a formalizzare come diversi segnali sensoriali con diversi livelli di affidabilità riescano a combinarsi a livello percettivo. Nella maggior parte degli studi effettuati, i soggetti sono stati esposti a stimoli cross-modali presentati a vari gradi di incongruenza nello spazio (ad esempio, in diverse posizioni) o nel tempo (ad esempio, il numero o il tasso di stimoli diversi) e veniva richiesto loro di esprimere un giudizio sulle caratteristiche degli stimoli esterni proposti (ad esempio, localizzarli spazialmente ecc). In queste condizioni, effetti percettivi cross-modali sono spesso accompagnati da illusioni sensoriali, come l'illusione di fissione / fusione nel dominio temporale (Shams et al, 2000; Shams & Beierholm, 2005).

Il modello bayesiano interpreta questi effetti percettivi supponendo che le caratteristiche di una stimolazione esterna, chiamata ϑ (ad esempio, la posizione spaziale o il numero degli stimoli) si trasformino in segnali sensoriali, chiamati x .

All'osservatore bayesiano sarà permesso di conoscere la migliore stima di ϑ a partire dalle esperienze a priori e grazie all'analisi delle caratteristiche di x , massimizzando la probabilità a posteriori $p(\vartheta | x)$ calcolata con la regola di Bayes, ovvero tenendo alta la probabilità di cambiare la propria percezione a priori in base all'analisi dei dati forniti dall'esperienza, quindi:

$$p(\vartheta|x) = \frac{p(\vartheta)p(x|\vartheta)}{p(x)}$$

dove sono presenti le distribuzioni analizzate precedentemente: $p(x|\vartheta)$ è la funzione probabilità di verosimiglianza, che specifica come vengono generati i segnali sensoriali (e tiene conto delle incertezze), e $p(\vartheta)$ rappresenta il conoscenza a priori circa il parametro ϑ .

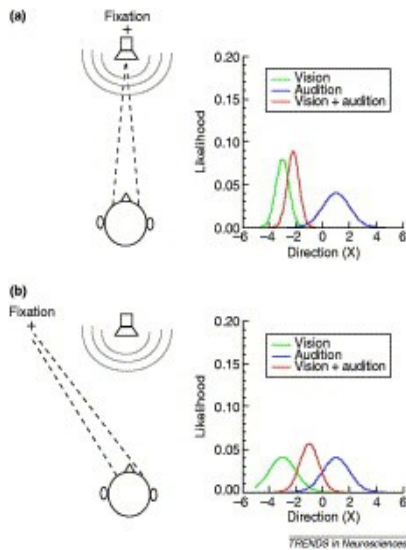


Fig. 1.

Per illustrare la struttura di base dei calcoli bayesiani, prendiamo in considerazione il problema dell'integrazione multisensoriale circa alcune proprietà di un determinato contesto. La Fig. 1 illustra la formulazione bayesiana di un problema (stimare la posizione di un oggetto X tramite segnali visivi e uditivi V e A). L'obiettivo di un osservatore bayes-ottimale sarebbe calcolare la funzione di densità condizionale $p(X | V, A)$. Utilizzando la regola di Bayes, questo è dato da

$$p(X | V, A) = p(V, A | X) p(X) / p(V, A) \quad (\text{Equazione 1})$$

dove $p(V, A | X)$ indica la probabilità relativa di percepire i dati forniti per diversi valori di X , e $p(X)$ è la probabilità a priori di avere diversi valori di X . Poiché le sorgenti di rumore nei meccanismi uditivi e visivi sono statisticamente indipendenti, possiamo scomporre la funzione probabilità nel prodotto di funzioni di probabilità associate ai segnali visivi e uditivi, rispettivamente:

$$p(V, A | X) = p(V | X) p(A | X) \quad (\text{Equazione 2})$$

$p(V | X)$ e $p(A | X)$ rappresentano le informazioni fornite dai dati visivi e uditivi sulla posizione del bersaglio. La funzione di densità a posteriori è quindi proporzionale al prodotto di tre funzioni: le funzioni di probabilità associate a ciascun segnale e la funzione di densità a priori, che rappresenta la probabilità relativa del bersaglio di trovarsi in

qualsiasi posizione (Knill & Pouget, 2004).

Quando un segnale è meno certo rispetto ad un altro, la stima integrata dovrebbe essere sbilanciata verso quello più affidabile. Supponendo che un sistema possa con precisione calcolare e rappresentare funzioni di probabilità, il calcolo descritto nelle Equazioni 1 e 2 applica implicitamente questo comportamento (Fig. 1).

Fino a poco tempo fa, i modelli di percezione multisensoriale, e la combinazione di segnali in generale, davano per scontato che i diversi segnali fossero tutti causati dalla stessa fonte (Knill & Pouget, 2004), e quindi modellavano il modo in cui il sistema nervoso combinerebbe i diversi segnali per stimare una proprietà fisica attraverso questa unica condizione. Tali modelli, però non consideravano la condizione generale in cui più segnali sensoriali potessero avere molteplici cause. Un modello che ha proposto segnali indipendenti, nonché un segnale comune, per due segnali sensoriali è stato quello di Shams et al. (2005). Questo modello bayesiano rappresenta interazioni audio-visive che vanno dalla fusione alla parziale integrazione in una varietà di compiti percettivi. In questo modello, sono state prese in considerazione due fonti, sA e sV, rispettivamente acustica e visiva; la probabilità a priori congiunta delle due sorgenti, cioè $p(sA, sV)$ ha colto l'interazione tra le due modalità, e ha portato come risultato l'intero spettro di interazioni. Kording et al. (2007) hanno dimostrato che un modello bayesiano gerarchico che esegue in modo esplicito compiti di inferenza causale si adatta bene ai test spaziali audio-visivi. Nel contesto della localizzazione spaziale, Alais e Burr (2004) hanno utilizzato un modello bayesiano per interpretare il giudizio di localizzazione degli stimoli audiovisivi; per realizzare ciò, all'osservatore è stato chiesto di considerare ogni presentazione bimodale come evento singolo.

La formulazione bayesiana di questo problema è:

$$p(\vartheta | x_v, x_a) = p(x_v, x_a | \vartheta) p(\vartheta) / p(x_v, x_a)$$

dove x_v e x_a sono rispettivamente, gli stimoli di posizione visivo e uditivo, e ϑ è la posizione dell'evento bimodale che deve essere stimato.

Supponendo che il rumore associato a ciascun segnale sensoriale sia indipendente e ipotizzando una distribuzione a priori uniforme, la massima stima a posteriori coincide con la stima di massima verosimiglianza ed è la somma degli stimoli uditivi e visivi pesati

ognuno per la loro affidabilità (Alais & Burr, 2004).

Secondo quanto appena detto, dunque, i modelli bayesiani sono potenti strumenti per prevedere la combinazione degli stimoli a livello comportamentale. Tuttavia, gran parte di questi, considerano il cervello come una scatola nera e non gli forniscono le adeguate basi meccanicistiche.

Riassumendo, le caratteristiche più interessanti dei dati sopra descritti per quanto riguarda l'ipotesi di codifica bayesiana sono che:

- i soggetti implicitamente regolano il peso degli stimoli in modo bayes-ottimale, sulla base dello stimolo stesso e sui parametri di visualizzazione;
- il comportamento percettivo e motorio riflette un sistema che tiene conto dell'incertezza di entrambi i segnali, sensoriale e motorio;
- gli esseri umani si comportano quasi in modo ottimale anche quando l'informazione sensoriale è caratterizzata da funzioni di densità altamente non-gaussiane, che porta a modelli complessi di previsione del comportamento;

4.1.3 I modelli bayesiani applicati alle BOIs

Recenti studi hanno dimostrato che la percezione umana della proprietà del corpo è altamente malleabile. Intuitivamente, il nostro senso di appartenenza del nostro corpo e parti del corpo appare intrinseca, stabile e immutabile. Tuttavia, recenti ricerche hanno dimostrato un incredibile grado di malleabilità nel nostro senso di possesso del corpo e della percezione. Mentre i protocolli e le regioni del cervello coinvolte in queste alterazioni di proprietà del corpo sono state studiate recentemente, le regole e i meccanismi computazionali rimangono poco conosciuti (Ehrsson et al., 2004).

Un esempio ben noto, come abbiamo ampiamente discusso nel capitolo precedente, è l'illusione della mano di gomma (RHI), in cui una mano fittizia viene attribuita erroneamente a se stessi quando essa viene posizionata in maniera anatomicamente e posturalmente plausibile, vicino alla mano vera nascosta, e mentre viene accarezzata in modo sincrono con quella vera. La RHI ovviamente comporta interazioni tra modalità visiva, tattile e propriocettiva. Inoltre, la percezione dell'illusione può essere descritta come l'inferenza di una causa comune per sensazioni propriocettive, tattili e visive, mentre

l'assenza di illusione può essere descritta come la percezione di fonti indipendenti per le sensazioni visive (mano di gomma) e quelle propriocettive e tattili (vera e propria mano). Pertanto, la percezione della RHI sembra dipendere da un processo di inferenza causale operante sui tre stimoli sensoriali.

Il modello che andremo a descrivere è stato presentato da Samad & Shams (2014). Esso realizza un'inferenza sulla struttura causale delle sensazioni, cioè se hanno una causa comune o cause indipendenti, in base alla somiglianza dei segnali sensoriali e alla probabilità a priori che abbiano una causa comune. Le proprietà dello stimolo vengono poi valutate in base alla struttura causale dedotta, che comporta l'integrazione dei sensi solo se giustificato dall'origine causale (Fig. 2). Questo modello è stato utilizzato per l'integrazione multisensoriale di informazioni spaziali (Kording et al., 2007) e temporali (Wozny et al., 2008), così come per la ricalibrazione sensoriale crossmodale (Wozny & Shams, 2011).

I segnali spaziali (X) e temporali (τ) provenienti dalla modalità visiva (X_v, τ_v) e da quella somatosensoriale (più esattamente, dalla propriocezione X_p e dal tatto τ_t) sono integrati o separati, a seconda che il cervello deduca una causa comune o cause indipendenti per le sensazioni.

La probabilità a posteriori che una struttura causale data da segnali sensoriali visivi (v), tattili (t), e propriocettivi (p) è calcolata utilizzando la regola di Bayes come segue:

$$p(C|X_v, X_p, \tau_v, \tau_t) = \frac{p(X_v, X_p, \tau_v, \tau_t|C)p(C)}{p(X_v, p, \tau_v, \tau_t)}$$

dove C è una variabile binaria che rappresenta la struttura causale (1 o 2 cause), quindi $C=1$ è la probabilità di avere una causa comune, mentre $C=2$ è la probabilità di avere cause distinte e indipendenti. X_v e X_p sono le sensazioni di localizzazioni, rispettivamente visive e propriocettive, τ_v e τ_t sono le sensazioni temporali, rispettivamente visive e tattili.

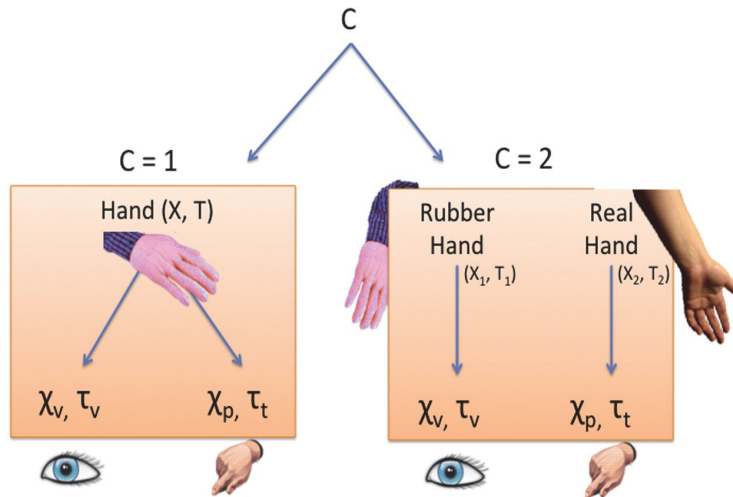


Fig. 2.

Una manipolazione semplice e breve della nostra esperienza sensoriale può indurre l'alterazione radicale della nostra proprietà del corpo e della percezione. L'illusione della mano di gomma (RHI) è stata scoperta 16 anni fa, ed è stata ampiamente studiata da quel momento in avanti. Tuttavia, i meccanismi computazionali di questa illusione, che offrirebbero una spiegazione del perché essa si verifichi, non sono stati ancora indagati. È ormai generalmente accettato che la percezione multisensoriale in ambienti naturali comporti due problemi computazionali, il problema dell'inferenza causale - determinare quali segnali siano causati dalla stessa fonte e quali siano causati da fonti esterne - e il problema dell'integrazione - come integrare i segnali sensoriali provenienti dalla stessa sorgente.

Per utilizzare il modello bayesiano, sono state incluse tre modalità (propriocezione, visione, tatto) e sia le informazioni spaziali, ricavate dalla propriocezione e dalla vista, che le informazioni temporali, ricavate dalla vista e dal tatto. Le simulazioni hanno riprodotto i risultati classici del caso in cui la RHI si verifica, vale a dire quando carenze sincronizzate tra la mano vera e quella di gomma provocano la percezione di una causa comune per gli stimoli visivi e tattili, e anche del caso in cui la RHI non si verifica, ovvero quando carenze asincrone producono la percezione di cause indipendenti. Inoltre, il modello fa previsioni circa il limite spaziale dell'illusione, cioè si prevede che l'illusione sarà indebolita aumentando la distanza tra la mano di gomma e quella reale, e inizierà a scomparire se la distanza supera i 30 cm (vedi Fig. 3).

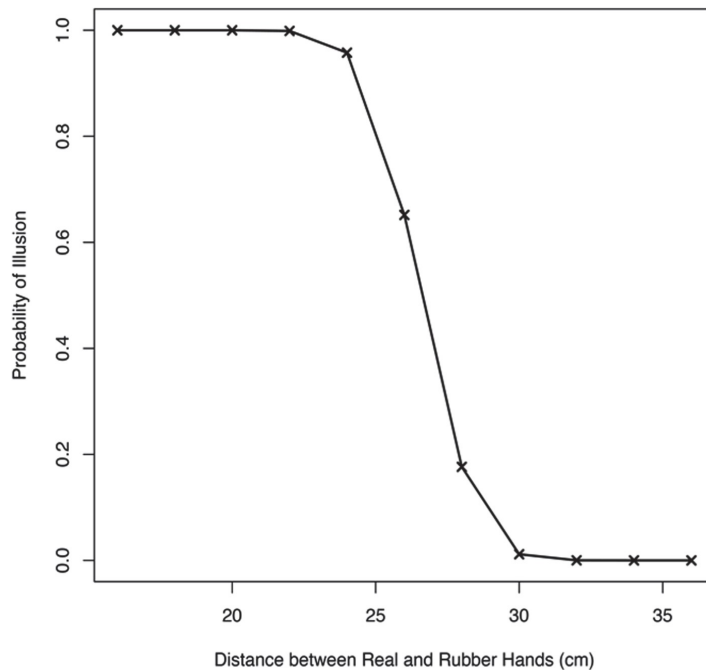


Fig. 3. La probabilità di sperimentare l'illusione viene tracciata come funzione della distanza (in centimetri) tra la mano di gomma e la mano reale. Poiché la distanza tra i due aumenta, l'illusione diventa più debole e alla fine non riesce a verificarsi. Questi risultati sono qualitativamente e quantitativamente in linea con i risultati empirici ottenuti nei precedenti studi (Lloyd, 2007).

Il modello di inferenza causale bayesiana mostra che diversi fattori contribuiscono alla percezione di una causa comune e, di conseguenza, all'illusione della RHI. Questi includono:

- la sovrapposizione tra le stime spaziali propriocettive e visive (che dipende sia dalla prossimità spaziale della mano di gomma alla mano reale, nonché dal grado di rumore propriocettivo);
- la congruenza tra le sensazioni tattili e visive;
- la tendenza a priori di integrare gli stimoli crossmodali.

Questo modello prevede che l'illusione sia tanto più forte quanto più le due mani sono vicine tra loro, quanto più rumorosa è la modalità propriocettiva, quanto più congruente il pattern temporale dei tracciati risulta attraverso le modalità visiva e tattile, e maggiore è la tendenza a integrare i segnali. Se uno di questi fattori è debole, tuttavia, non causerà necessariamente la rottura dell'illusione, dato che gli altri fattori possono compensare e collettivamente fornire sufficienti prove per una causa comune. È la forza dell'evidenza complessiva che determina la probabilità di inferire una causa comune e quindi di percepire l'illusione, e non un singolo fattore di per sé. Inoltre, il modello non solo fornisce una descrizione quantitativa delle condizioni che danno luogo all'illusione, ma spiega

anche che la RHI si verifica a causa di un'inferenza statistica ottimale sulla struttura causale e sulle proprietà spazio-temporali delle sensazioni (Samad & Shams, 2014).

In conclusione, un modello Bayesiano che fa una inferenza sulla struttura causale degli stimoli sensoriali, cioè dei segnali visivi, propriocettivi e tattili, basata sulla somiglianza degli stimoli e sulle conoscenze a priori, può spiegare l'illusione della mano di gomma. Più precisamente, questi risultati suggeriscono che quando le informazioni spazio-temporali ricavabili dai sensi sono sufficientemente congruenti, una causa comune per le sensazioni viene dedotta dal sistema nervoso, portando all'esperienza di una fonte unificata e, infine, alla proprietà del corpo.

Quello che i risultati attuali dimostrano è che questo processo può essere modellato come una regola sofisticata e statisticamente ottimale di inferenza (inferenza causale bayesiana), che sembra possa governare anche altri processi percettivi. Pertanto, sembra che la nostra percezione e la coscienza di sé non siano diverse, in linea di principio, dalla percezione del mondo esterno: seguono le stesse regole, e possono essere modificate allo stesso modo.

4.2 Neuroimaging ed elettroencefalografia

Le immagini funzionali hanno lo scopo di fornire informazioni qualitative e quantitative utili a comprendere meglio la fisiologia e le funzioni delle cellule, dei tessuti e degli organi analizzati.

Il NeuroImaging ha come obiettivo principale la comprensione delle funzioni specifiche delle diverse aree cerebrali e come queste interagiscono in un sistema integrato di reti neuronali.

È uno strumento di primaria importanza nelle neuroscienze cognitive e in neuropsicologia. Oltre alle classiche applicazioni di ricerca sperimentale sui processi neurocognitivi, le tecniche di neuroimaging funzionale, a livello di ricerca, l'obiettivo principale dello studio delle funzioni cerebrali è la comprensione delle funzioni delle diverse aree e i collegamenti reciproci in un sistema integrato di reti neuronali, mentre, a livello clinico, lo studio funzionale è utile per la rilevazione, comprensione ed esecuzione di diagnosi differenziali in vari quadri patologici, quali traumi, tumori o malattie neurodegenerative.

Tale metodo, spesso non invasivo, permette di visualizzare, oltre alla posizione e

all'estensione dei danni cerebrali (importantissime nell'ambito medico-neurologico), anche l'impatto che i processi mentali hanno sull'attività cerebrale e le modifiche che questa subisce ogniqualvolta il cervello svolge un compito.

Tre sono le principali tecniche di visualizzazione cerebrale: la EEG (Elettroencefalografia), la PET (Tomografia a Emissione di Positroni) e la fMRI (Risonanza Magnetica funzionale). Nel seguito ci concentreremo sull'EEG, ponendo l'attenzione sul contributo di questo metodo d'analisi alla comprensione dei meccanismi di integrazione multisensoriale.

4.2.1 L'elettroencefalografia

L'EEG è un metodo di monitoraggio elettrofisiologico, messo a punto già negli anni '20 da Hans Berger, per registrare l'attività elettrica cerebrale. È tipicamente non invasivo, con gli elettrodi posizionati sullo scalpo, anche se esistono elettrodi invasivi (corticali o di profondità), talvolta utilizzati in applicazioni specifiche.

L'EEG misura il potenziale elettrico delle oscillazioni derivanti dalla corrente ionica nei neuroni, raccogliendo così informazioni sull'attività dello strato superficiale della corteccia cerebrale. Esiste uno stretto rapporto fra il segnale EEG misurato e lo stato di attivazione neuronale. In contesti clinici, infatti, quando la corteccia è impegnata nell'elaborazione di informazioni, come può essere la risoluzione di un compito (sensoriale, motorio, cognitivo, problem solving...), la differenza di potenziale è generata da una grande quantità di neuroni, i quali però si attivano in maniera asincrona. Il segnale elettrico misurato sullo scalpo, che risulta dalla somma di tutte le attivazioni, in questo caso, ha una frequenza maggiore ma un'ampiezza minore. Nel caso in cui, invece, la corteccia non è particolarmente implicata nei processi di pensiero, ad esempio durante il sonno, i neuroni mandano un input più lento, da cui deriva un segnale con frequenza minore ma ampiezza notevolmente maggiore.

Il confronto tra l'attività spontanea neuronale e la sua variazione durante l'attività indotta ci permette di individuare, anche in tempo reale, le aree di maggiore attività elettrica. Inoltre, l'attuale possibilità di registrare più canali contemporaneamente ci permette di generare una mappa cerebrale (EEG topografico) della distribuzione spaziale di tale attività elettrica. Nonostante una risoluzione spaziale limitata, l'EEG continua ad essere un valido strumento per la ricerca e la diagnosi, soprattutto quando è richiesta una risoluzione temporale nella gamma dei millisecondi.

Attualmente, dal punto di vista cognitivo e funzionale, uno degli ambiti di impiego dell'EEG di maggior interesse è l'analisi dei potenziali evocati (EP), che consistono in una media dell'attività EEG time-locked alla presentazione di uno stimolo esterno. Ad esempio stimoli visivi innescano EP visivi (VEP), mentre stimoli uditivi determinano EP acustici (AEP).

L'EP è caratterizzato da una specifica onda avente una determinata ampiezza e latenza. L'ampiezza, o valore di picco, è correlata con l'estensione delle aree corticali coinvolte e può assumere sia polarità positiva, sia negativa. La latenza è la distanza temporale tra il momento di applicazione dello stimolo ed il momento di comparsa di deflessioni positive o negative, dette componenti.

Una sottocategoria importante è rappresentata dai potenziali evento-correlati (ERP), che consistono in un'elaborazione più complessa degli stimoli e dipendono dal contesto psicologico nel quale avviene la stimolazione (Regan, 1989).

Un ERP è, quindi, "qualsiasi risposta elettrofisiologica a uno stimolo interno oppure esterno" (Wikipedia). Mentre gli EP riflettono l'elaborazione dello stimolo fisico da parte di alcune strutture anatomiche ben conosciute, gli ERP sono causati da processi di ordine superiore, che potrebbero coinvolgere aspetti come la memoria o l'attenzione, oppure cambi nello stato mentale.

4.2.2 Applicazione del modello EEG all'integrazione multisensoriale

La tecnica EEG descritta in precedenza, trova largo impiego sia nella pratica clinica che in ampi settori della ricerca medica. Infatti, le indagini tradizionali sui meccanismi di elaborazione sensoriale si erano sempre focalizzate sull'attività nelle corteccie sensoriali in funzione dei rispettivi ingressi sensoriali primari (ad esempio la modulazione di attività nella corteccia visiva in risposta a stimoli visivi). Solo più di recente, ci si è concentrati sugli ingressi cross-sensoriali e sul modo in cui essi influenzino l'elaborazione sensoriale precoce nella cosiddetta corteccia unisensoriale. Ciò ha comportato una notevole riconcettualizzazione sul modo in cui i sistemi sensoriali interagiscono per influenzare la percezione, ottenendo prove sempre più univoche sul fatto che l'attività neuronale in una data regione sensoriale corticale sia modulata non solo dai suoi ingressi sensoriali primari, ma anche dalla stimolazione degli altri sistemi sensoriali (Foxy e Schroeder, 2005;).

Meredith et al, 2009). Un particolare interessante è che, mentre un ingresso uditivo della corteccia visiva potrebbe non necessariamente farci ottenere una risposta rilevabile quando viene presentato isolato, si è visto che può modulare la risposta evocata quando viene presentato in concomitanza ad uno stimolo visivo (Allman et al., 2008; Allman e Meredith, 2007). Perciò, si può affermare che gli ingressi cross-sensoriali servono anche a modulare la risposta neurale agli stimoli sensoriali primari,

Nel seguito verrà presentato uno studio sperimentale del 2013, condotto da Mercier e collaboratori, che rappresenta una buona applicazione di tale modello.

Basandosi su studi precedenti che, attraverso registrazioni elettrofisiologiche, avevano rivelato che la fase di attività oscillatoria nella corteccia uditiva primaria e in quella secondaria può essere considerata un "reset" per gli input somatosensoriali o visivi (Kayser et al., 2008), Mercier et al. hanno cercato di valutare se e in quale quantità la stimolazione uditiva avrebbe influenzato le risposte visive in condizioni di stimolazione multisensoriale.

Per fare ciò, essi si sono avvalsi di pazienti affetti da epilessia, a cui erano stati impiantati degli elettrodi subdurali. I punti di posizionamento (COI, "contact of interest") convergevano sulla corteccia posteriore, per poter verificare direttamente la risposta uditiva nella regione visiva.

Ai soggetti sono stati forniti stimoli solo uditivi, solo visivi, e audiovisivi, in ordine casuale; l'intervallo temporale tra ogni stimolo era anche esso casuale, tra 750 e 3000 ms. Lo stimolo acustico, ad una frequenza di 1000 Hz e con una durata di 60 ms, è stato presentato ad un livello di ascolto tra i 60 e i 70 dB, attraverso delle cuffie, mentre lo stimolo visivo, rappresentato da un disco rosso centrato, è stato presentato su un monitor ad una distanza di osservazione di 75 cm e per 60 ms. I pazienti avevano il compito di rilevare gli stimoli, di qualsiasi natura essi fossero (uditivo, visivo, o audiovisivo).

Utilizzando tali registrazioni intracraniche, questo studio ha scoperto che gli stimoli puramente uditivi modulano l'attività neuronale della corteccia visiva umana. Le risposte sono state generalmente coerenti con un modello oscillatorio caratterizzato da piccole variazioni di ampiezza, in contrasto con la maggiore ampiezza dei classici ERP evocati dalla stimolazione audiovisiva e da quella puramente visiva (vedi Fig. 4).

Questo tipo di risposta uditiva nella corteccia visiva può essere considerato nel contesto dei meccanismi sottostanti la generazione di ERP. L'opinione prevalente prevede che, a seguito

di input sensoriali in ingresso, un ERP emerge a causa della sincronizzazione dei gruppi neuronali. Tale sincronizzazione risulta dal ripristino della fase di oscillazione guidata dallo stimolo, che può essere accompagnata da un aumento dell'attività neuronale (Becker et al., 2008; Shah et al., 2004). Le risposte uditive registrate nella corteccia visiva suggeriscono che la stimolazione cross-sensoriale influenzi l'attività corticale sensoriale, dato che la maggioranza dei COI ha dimostrato un profilo tipico di ripristino di fase.

A supporto di queste affermazioni, alcuni ulteriori studi hanno rivelato che ciò riorganizza le fluttuazioni dei potenziali di membrana sottosoglia (oscillazioni in corso), rendendo l'insieme neuronale suscettibile alla scarica della risposta alla stimolazione della modalità sensoriale primaria, in questo caso quella visiva (Lakatos et al., 2007).

Anche se il metodo attuale non consente di determinare il percorso attraverso il quale si è verificata l'influenza guidata dall'udito, ci sono diverse possibilità altamente plausibili da considerare. Questi includono una linea uditiva corticale verso la via visiva (Cappe & Barone, 2005), o una regione di mediazione multisensoriale di ordine superiore, quale il solco intra-parietale (Leitao et al., 2012).

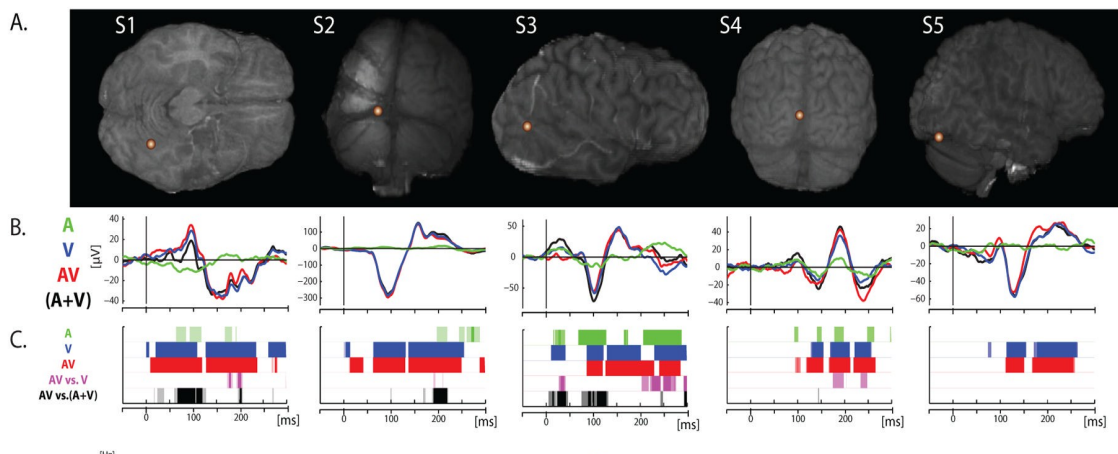


Fig. 4. Esempi di forme d'onda ERP. (A) la posizione dei COI per ciascuno dei cinque partecipanti, con i COI selezionati nelle diverse regioni della corteccia visiva, che rappresentano una vasta gamma di risposte osservabili; (B) le risposte ERP per condizioni puramente uditive (A, verde), puramente visive (V, blu) e audio-visive (AV; rosso). La somma delle risposte unisensoriali (A + V) è raffigurata con una linea nera; (C) i risultati statistici che rappresentano il confronto di ogni condizione con il livello base e, in particolare, tra le condizioni che seguono il maximum model (AV vs V; magenta) e il modello additivo (AV vs. (A + V); nero).

Per determinare se gli ERP rappresentato una modulazione statisticamente significativa, le ampiezze post-stimolo (da 0 ms a 300 ms) sono state confrontate con i valori di ampiezza di base/istante pre-stimolo (da -100ms a 0 ms). Ciò è stato fatto abbinando una prova casuale di permutazione per ogni istante temporale: per ogni prova, un istante di tempo è stato scelto a caso all'interno del periodo di riferimento. Dopo aver determinato la coppia di valori, rappresentata dall'istante pre- e quello post-stimolo, è avvenuta la permutazione, ed è stata calcolata la differenza tra i due valori. Infine, un valore medio è stato calcolato in ogni prova.

Particolarmente rilevante è che il segnale pre-stimolo può trasportare informazioni legate temporalmente alla stimolazione, soprattutto nel contesto dei disegni sperimentali multisensoriali e ciò può riflettere anticipazione, trascinarsi, o fluttuazioni dell'attenzione (Besle et al., 2011; Fiebelkorn et al., 2013 Lakatos et al., 2009).

Dopo queste misurazioni, Mercier e collaboratori hanno potuto parlare di statistiche.

In totale, il 69% dei COI ha mostrato un potenziale evocato visivo (VEP) (74 dei 108 COI) sia per la condizione puramente visiva, sia per quella audiovisiva.

Questo è stato valutato visivamente considerando l'aspetto transitorio e fasico della risposta che caratterizza un ERP tipico, e confermato da una significativa risposta post-stimolo.

Piuttosto indicativo per questo contesto, il 20% dei COI ha anche mostrato risposte significative alla condizione puramente uditiva (22 su 108, vedi Fig. 4).

La natura dell'attività evocata uditiva nelle regioni corticali visive era notevolmente diversa dai classici potenziali evocati uditivi (AEP) osservati nella corteccia uditiva (Fig. 4, S1 e S2, Molholm et al., 2006). Invece, la risposta è stata caratterizzata da un andamento oscillatorio lento che ha raggiunto valori significativi solo per brevi periodi di tempo, quelli corrispondenti ai picchi, positivi e/o negativi (Fig. 4). Ciò dimostra che le risposte puramente visive differiscono da quelle audiovisive per forma, ampiezza e latenza, in funzione dei COI.

La Fig. 4 mostra anche la presenza di risposte alla condizione puramente uditiva in alcuni dei COI, anche se, come ci si aspetterebbe, questi sono di ampiezza molto minore rispetto a quella delle risposte agli stimoli contenenti un elemento visivo.

Inoltre si è scoperto che, anche se generalmente la direzione del AEP segue quella del VEP,

ciò non è sempre accaduto in questo studio. Le differenze nella polarità della risposta per i diversi tipi di stimolazione probabilmente riflettono le diverse popolazioni neurali sottostanti.

In Fig. 5 sono rappresentate le collocazioni dei COI che presentano risposte ERP significative per la stimolazione puramente uditiva. Fatta eccezione per la regione posteriore del lobo occipitale, dove la rappresentazione era relativamente scarsa, la distribuzione spaziale delle risposte uditive è apparsa relativamente omogenea per quel che riguarda la copertura dei COI.

Per valutare se la stimolazione puramente uditiva ha modulato l'ERP visivo, sono state confrontate le risposte ERP audiovisive con quelle puramente visive. Questo ha rivelato differenze significative per 28 dei 79 COI, i quali mostravano un ERP visivo (ovvero il 35% del totale), con la risposta audiovisiva che tendeva ad essere maggiore rispetto a quella puramente visiva (nel 61% dei casi, ovvero 17 dei 28 COI).

Il 53,6% dei COI con una significativa integrazione multisensoriale (MSI) aveva anche ERP significativi in risposta alla stimolazione puramente uditiva (piccolo quadrato rosso con bordi bianchi in Figura 5).

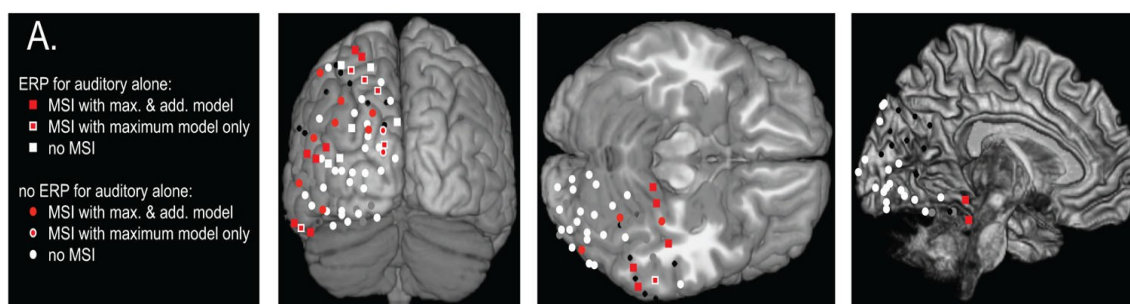


Fig. 5. Sintesi dei risultati ERP osservati per tutti i partecipanti. Per ogni COI, la forma quadrata indica che è stato osservato un ERP per uno stimolo puramente uditivo, mentre il cerchio indica che non è stato osservato alcun ERP uditivo. Il codice di colore rappresenta la presenza (rosso) o l'assenza (bianco) di integrazione multisensoriale (MSI). Infine, i COI senza alcuna risposta ERP o contaminati da artefatti sono stati raffigurati rispettivamente in nero e grigio.

Per ciò che concerne le statistiche multisensoriali, il *maximum criterion model* misura quanto la risposta multisensoriale differisce dalla risposta massima unisensoriale. Nel contesto della presente analisi, dato che i contatti di interesse (COI) sono limitati alle cosiddette regioni unisensoriali visive, è stata fatta un'ipotesi "zero", che afferma che qualsiasi risposta ad uno stimolo uditivo, anche nel caso in cui sia presentato insieme allo stimolo visivo, dovrebbe essere limitata al rumore. Qualsiasi risposta ad uno stimolo audiovisivo dovrebbe pertanto corrispondere alla risposta ad uno stimolo puramente visivo. In questo caso l'ipotesi è risultata essere falsa, e ciò significa che alcune informazioni circa lo stimolo uditivo vengono elaborate all'interno della regione visiva; infatti si è potuto osservare variazioni significative nell'attività in risposta alla stimolazione puramente uditiva.

Per valutare l'impatto della stimolazione uditiva sul trattamento di uno stimolo visivo, abbiamo applicato il *maximum criterion model* ha rivelato che gli ERP per la stimolazione audiovisiva erano spesso più elevati rispetto a quelli per la stimolazione puramente visiva. Tali effetti MSI, però, non implicano sistematicamente che una stimolazione puramente uditiva evochi un ERP rilevante sullo stesso COI.

La preponderanza di risposte MSI osservata qui può riflettere ciò che afferma la regola dell'efficacia inversa (Stein e Meredith, 1993). Le registrazioni EEG di superficie, infatti, hanno ripetutamente dimostrato che il miglioramento multisensoriale tende ad essere più grande quando gli ingressi unisensoriale sono minimamente efficaci se presentati isolatamente (Kayser et al., 2008; Meredith e Stein, 1986).

Un altro aspetto che è stato preso in esame è il modo in cui la modulazione di fase eseguita dagli ingressi uditivi cross-sensoriali interagiscono con la modulazione di fase evocata dall'ingresso primario visivo. Valutando la MSI a livello di fase, possiamo affermare che il *maximum criterion model* ha rivelato una maggiore concentrazione di fase per la condizione audiovisiva, con lo stimolo audiovisivo che porta a una maggiore concentrazione di fase rispetto allo stimolo puramente visivo.

In sintesi, il presente studio rappresenta uno dei tanti esempi di come vengono utilizzate le immagini EEG dalle moderne neuroscienze. In particolare, esso ha utilizzato l'analisi EEG per valutare gli effetti della stimolazione uditiva nella corteccia sensoriale visiva. Mercier et al. hanno potuto affermare, attraverso lo studio degli ERP e grazie all'applicazione del

maximum criterion model, che esistono effetti netti sul miglioramento delle MSI, con la risposta ERP agli stimoli audiovisivi, in media, maggiore in ampiezza rispetto alla risposta ERP allo stimolo puramente visivo, fornendo prove chiare sull'importanza a livello neurologico dell'integrazione multisensoriale.

4.3 Le reti neurali

Sebbene la conoscenza del cervello sia ancora limitata sotto molti punti di vista, si conoscono molte informazioni circa l'anatomia e la fisiologia delle reti neurali biologiche. L'anatomia di base di una singola cellula nervosa, ovvero il neurone, è ben nota e sono state identificate le principali reazioni biochimiche che governano la sua attività. Un neurone può essere considerato come l'unità computazionale elementare del cervello; esso valuta l'intensità di ogni input, somma i diversi input confrontando il risultato con una soglia opportuna, e infine determina il valore dell'output. Nel cervello umano ne sono presenti circa 100 classi differenti.

Il sistema nervoso comunica attraverso segnali elettrici. Il compito principale del neurone è quello di generare un potenziale elettrico, superata una certa soglia, il quale si propaga lungo l'assone (l'output del neurone). L'input al neurone, invece, è rappresentato da un insieme di fibre, i dendriti, che sono in contatto con gli assoni degli altri neuroni per riceverne i potenziali elettrici. Infine, il punto di connessione fra l'assone di un neurone e il dendrite di un altro neurone viene chiamato sinapsi.

Tutti questi aspetti prettamente biologici sono alla base degli studi sulla computazione neurale.

Una rete neurale artificiale si ispira ai sistemi neurali biologici, dei quali cerca di modellare la struttura e di simulare le funzioni di base. La filosofia delle reti considera un numero elevato di processori che hanno una capacità computazionale elementare, i neuroni artificiali o nodi, connessi ad altre unità dello stesso tipo. Ad ogni connessione è assegnato un peso, cioè può essere più o meno debole, in modo tale che un neurone possa influenzarne un altro in funzione della "forza" della connessione fra i due, proprio come nei sistemi neurali biologici in cui un neurone (pre-sinaptico) influenza un secondo neurone (post-sinaptico) in funzione del potenziale post-sinaptico (Fig. 6).

La simulazione della struttura dei sistemi neurali biologici potrebbe portare allo sviluppo di sistemi artificiali che conservino le caratteristiche del comportamento intelligente dei sistemi biologici. Ciò rappresenta un cambiamento di prospettiva piuttosto radicale rispetto alla computazione convenzionale.

L'obiettivo di costruire tali macchine è molteplice. Dal punto di vista prettamente ingegneristico, l'interesse è rivolto alle applicazioni potenzialmente utili per scopi pratici, quali l'intelligenza artificiale e la robotica; un punto di vista più generale, ovvero quello che interessa le moderne neuroscienze, è interessato alle reti neurali per le informazioni che si possono trarre sul funzionamento, generale e particolare, delle reti neurali, sia per ciò che riguarda i normali processi psicologici e cognitivi, sia nei casi di patologie nervose, difficili da studiare in altro modo.

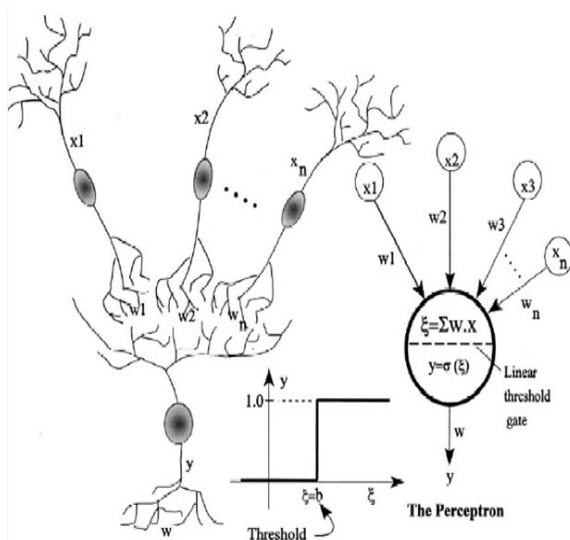


Fig. 6 Confronto tra rete neurale biologica e artificiale:
 n = neuroni in entrata;
 x_i = intensità del segnale;
 w_i = forza sinapsi;
 b = soglia del neurone postsinaptico
 x = combinazione segnali in entrata
 s = funzione di attivazione
 y = output

I modelli precedentemente descritti, in particolare le inferenze bayesiane, hanno un grosso limite, in quanto non incorporano il funzionamento dei neuroni realistici, ma derivano le probabilità direttamente dai dati sperimentali. Tuttavia, è estremamente importante capire dove e come questi calcoli possono essere realizzati da una rete neurale biologicamente ispirata.

4.3.1 I modelli di reti feedforward

I modelli neurali basati sul feedforward partono dal presupposto che i neuroni in una determinata regione (quale può essere il collicolo superiore (CS), che abbiamo già precedentemente analizzato) ricevono input afferenti solo dalle aree unisensoriali. Quando tali modelli vengono applicati alla corteccia, sono utili per studiare le proprietà delle aree associative di ordine superiore (che rispondono a più di una modalità sensoriale), le quali ricevono input di ordine inferiore, convergenti dalle aree unisensoriali.

Un modello di integrazione multisensoriale che cerca di spiegare diverse proprietà multisensoriali del collicolo superiore (CS) in conseguenza ad un singolo neurone computazionale è stato presentato da Rowland e colleghi (2007).

Esso considera un neurone principale e un interneurone inibitorio. L'assunzione del modello principale considera il modo in cui le afferenze ascendenti e discendenti sono disposte sui dendriti del neurone principale. Gli input uditivi e visivi discendenti convergono sullo stesso dendrite; i due input ascendenti (non-corticali) convergono su dendriti diversi. In questo modo, solo le afferenze discendenti possono interagire sinergicamente, dato che solo loro hanno un bersaglio comune.

Nella sua forma algebrica, il modello assume che ogni dendrite sommi e poi elevi al quadrato gli input. Tutti gli ingressi convergono anche su un interneurone, che li somma. Il soma calcola la somma degli output dendritici e la divide per l'output dell'interneurone. Infine, un logaritmo è utilizzato per calcolare l'uscita, al fine di spiegare la non linearità nel rapporto ingresso/uscita del neurone. Quindi, l'equazione è:

$$m = \tau \ln \left[\sum (\alpha (Va^2 + Aa^2 + (Vd + Ad)^2) / (Va + Aa + Vd + Ad)) \right]$$

dove m è l'output del neurone multisensoriale, α e τ sono parametri costanti, e i restanti parametri indicano gli input visivi e acustici, ascendenti e discendenti.

Questo modello può spiegare diversi comportamenti dei singoli neuroni del CS, quali: l'effetto della disattivazione corticale; la valorizzazione multisensoriale, e l'efficacia inversa; le finestre temporali di integrazione. Lo svantaggio principale è che esso considera la computazione in una singola posizione spaziale. Tuttavia, al fine di attuare alcune delle

idee dei modelli bayesiani (ad esempio per rilevare la posizione di uno stimolo), sono necessarie catene di neuroni unisensoriali con campi recettivi posti in diverse posizioni. Per ovviare a tale problematica, un esempio di modello che considera il problema spaziale nel CS è quello presentato da Ursino, Cuppini e Magosso (2008, 2009), costituito da due strati di neuroni unisensoriali e da uno strato di neuroni multisensoriali, i quali eseguono la somma pesata degli input provenienti da campi recettivi spazialmente sovrapposti. I neuroni unisensoriali mostrano una saturazione non lineare. La presenza di una sigmoide, ossia una funzione non lineare, può spiegare l'efficacia inversa, provocando uno spostamento dalla valorizzazione dell'enhancement a seconda della forza degli stimoli cross-modali. Per ciò che riguarda la soppressione multisensoriale, tale modello è in grado di riprodurre il principio spaziale, secondo cui, quando due stimoli cross-modali sono presentati in qualsiasi punto all'interno del campo recettivo del neurone bersaglio, indipendentemente dal loro allineamento spaziale, producono la valorizzazione multisensoriale, mentre se uno dei due stimoli è al di fuori del campo recettivo specifico, l'attività viene soppressa rispetto alla stimolazione unisensoriale (Meredith & Stein, 1986).

4.3.2 I modelli a feedback

Le reti a feedforward precedentemente descritte analizzano quelle situazioni in cui due segnali provenienti da modalità sensoriali diverse (ad esempio, un segnale visivo e uno acustico) convergono al fine di formare un'unica percezione integrata.

Ora andremo ad analizzare quelle circostanze in cui i segnali, presi singolarmente nelle loro modalità, si influenzano l'un l'altro. D'altra parte, una delle proprietà più interessanti della corteccia è che l'integrazione multisensoriale è già presente nelle aree unisensoriali primarie (nel senso che la risposta ad uno stimolo nella modalità specifica è influenzata da uno stimolo in una modalità diversa, come già visto per l'analisi con EEG).

Per indagare questo aspetto, sono stati presentati modelli di reti neurali utilizzate per simulare la rappresentazione multisensoriale dello spazio peripersonale (Magosso, 2010) intorno alla mano. Il modello comprende due sotto-reti, una per ciascun emisfero, ciascuna delle quali rappresenta lo spazio peripersonale della mano controlaterale ed è in grado di riprodurre diversi risultati in vivo sulla rappresentazione dello spazio "peri-mano". È importante sottolineare che, grazie ai collegamenti in feedback, il modello imita la

facilitazione cross-modale nelle aree unisensoriali: la combinazione di uno stimolo tattile sottosoglia, spazialmente congruente con lo stimolo visivo, aumenta l'attivazione dell'area unimodale tattile in maniera significativa.

Un secondo interessante modello è stato proposto da Cuppini et al. (2014), e si è dedicato ad analizzare l'integrazione cross-modale nelle aree unisensoriali.

Il modello assume la presenza di due aree unisensoriali (una uditiva e una visiva) con un'organizzazione topologica abbastanza fine per l'area visiva, e un'organizzazione spaziale più ampia per quella uditiva. Un punto chiave del modello è la presenza di collegamenti bidirezionali diretti tra i neuroni visivi e quelli uditivi. Inoltre, il modello assume una diversa dinamica temporale per i neuroni uditivi e per quelli visivi, in modo da simulare campi recettivi spazio-temporali più affidabili per le due modalità: l'elaborazione uditiva presenta una migliore risoluzione temporale, mentre l'elaborazione visiva una migliore acutezza spaziale. Con questo miglioramento, il modello può essere in grado di spiegare anche il fenomeno della sound-induced flash illusion (ampiamente descritto nel Capitolo 2) usando un singolo set di parametri.

Nella sound-induced flash illusion, infatti, le informazioni uditive determinano l'esperienza fenomenica del numero di stimoli visivi percepiti. Questi effetti funzionali devono essere attuati da collegamenti fisici tra le diverse regioni, sia sensoriali che associative, del sistema nervoso. Diversi risultati psicofisici e di Neuroimaging mostrano che l'alterazione uditiva indotta della percezione visiva è associata ad una modulazione dell'attività neurale nella corteccia visiva (Shams et al., 2001): quando un singolo flash è percepito erroneamente come doppio, l'attività dell'area visiva primaria (V1) è aumentata; al contrario, tale attività è ridotta quando un doppio flash è percepito come fosse singolo (Watkins et al., 2006, 2007). Quindi, l'attività in V1 sembra riflettere la percezione soggettiva del partecipante, piuttosto che lo stimolo visivo presentato fisicamente. Più specificamente, negli studi illusori, i potenziali evocati visivi (VEP) sono modulati dal suono, suggerendo che questi effetti cross-modali si verificano a livello della corteccia occipitale (Shams et al., 2001). È importante sottolineare che, i VEP associati ai flash illusori sono qualitativamente molto simili a quelli associati ai flash fisici, il che farebbe pensare che i meccanismi cerebrali alla base delle due percezioni generate nella corteccia

visiva siano simili.

Le affermazioni precedenti supportano l'idea che il nostro sistema nervoso sia organizzato in compartimenti sensoriali specifici, che giungono alla massima performance solo grazie all'integrazione multisensoriale. Essa inizia nell'elaborazione percettiva precoce e arriva a compimento solo più tardi, nelle aree corticali associative di livello superiore (Fuxe & Schroeder, 2005). Inoltre, le influenze di feedback su ogni corteccia sensoriale specifica da parte delle zone di convergenza multisensoriale possono svolgere un ruolo importante nel modulare le interazioni crossmodali nelle aree sensoriali primarie.

Il modello di Cuppini e collaboratori è rappresentato in Fig. 7, e consiste in due catene, costituite ciascuna da $N=180$ neuroni, una rappresentante l'area visiva e l'altra l'area acustica. I cerchi grigi schematizzano i neuroni; ogni linea invece rappresenta una connessione sinaptica eccitatoria (linee che terminano con una freccia) o inibitoria (linee che terminano con un punto).

Ogni neurone ha il compito di codificare le informazioni provenienti da una specifica posizione dello spazio e sono topologicamente allineati, perciò i neuroni che codificano posizioni vicine nello spazio sono posti vicini gli uni agli altri. La distanza tra ciascun neurone è stata assunta pari a 1° , perciò, posto $N=180$, ogni strato copre un'area di 180° . Per descrivere la risposta neurale, ci si è avvalsi di un'equazione differenziale del primo ordine, che simula le proprietà integrative della membrana cellulare, e di una caratteristica statica sigmoideale, che stabilisce la presenza di una soglia inferiore e di una saturazione superiore per l'attivazione neurale. Il valore di saturazione è stato imposto a 1, quindi tutte le uscite sono normalizzate al valore massimo.

Così l'ingresso totale ricevuto da ogni neurone risulta la somma di tre contributi: un input esterno, un input proveniente dai neuroni della stessa area unisensoriale e un input cross-modale proveniente dai neuroni dell'altra modalità sensoriale.

Ciascun input visivo e acustico esterno è rappresentato con una componente spaziale e una componente temporale. Gli aspetti spaziali sono descritti con una funzione Gaussiana che simula uno stimolo locale visivo o acustico, quindi un flash o un beep, che origina da una posizione limitata dello spazio, filtrato poi dai campi recettivi dei neuroni (blu quello visivo e rosso quello acustico). Il punto centrale di tale funzione Gaussiana (p_a e p_v come mostrati in Fig. 1) corrisponde al punto di applicazione dello stimolo, mentre la deviazione

standard (σ_a e σ_v) è relazionata alla larghezza del campo. Grazie alle osservazioni sperimentali precedenti (Meredith & Stein, 1983; Shams et al., 2002), è stato possibile applicare al modello l'ipotesi che il sistema visivo sia spazialmente più preciso di quello acustico, per cui si è assunto $\sigma_v < \sigma_a$.

Gli aspetti temporali degli input sono descritti con una cinetica del secondo ordine, caratterizzata da due costanti di tempo, τ_a e τ_v , rispettivamente la modalità acustica e per la modalità visiva; in particolare, la modalità uditiva ha una risposta più rapida agli stimoli esterni e crossmodali (quindi $\tau_a < \tau_v$).

L'input cross-modale, invece, è rappresentato da sinapsi eccitatorie che connettono in maniera simmetrica le due aree sensoriali. In questo modo, un neurone in una modalità sensoriale può essere eccitato solo dal neurone corrispondente dell'altra modalità sensoriale, posto nella stessa posizione spaziale.

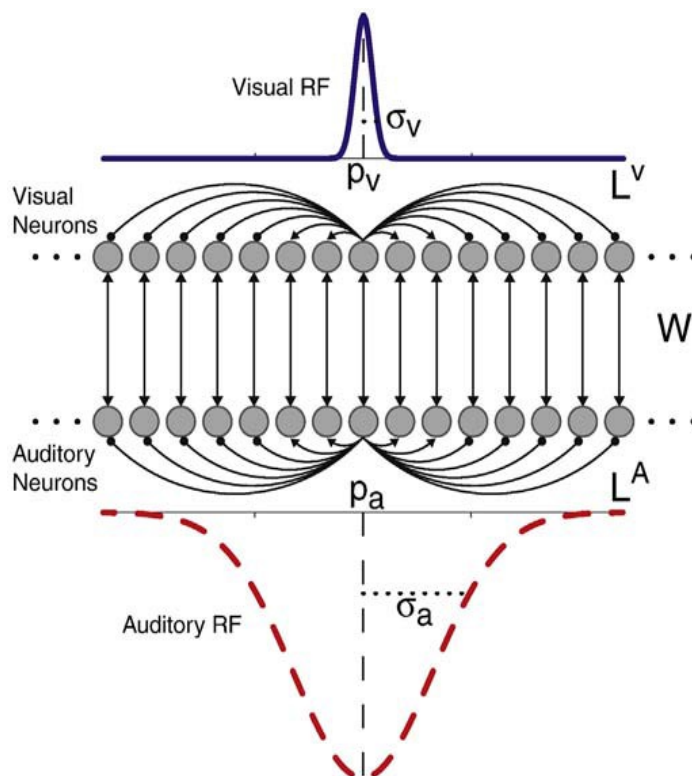


Fig. 7. Schema del modello della rete neurale di Cuppini et al. (2014)

Infine, il modello assume l'inibizione laterale tra i neuroni nella stessa area. All'interno

dello stesso strato unimodale, sono presenti delle connessioni laterali tra gli input. Queste connessioni includono sia le sinapsi laterali eccitatorie sia quelle inibitorie, che sono state disposte secondo un modello a cappello messicano: una zona eccitatoria centrale viene circondata da una corona circolare di inibizione. In tal modo, ogni neurone eccita (e viene eccitato da) i suoi neuroni prossimali, e inibisce (e viene inibito da) i neuroni più distali. Quindi, gli stimoli distali unimodali tendono a sopprimersi reciprocamente (Fig. 8).

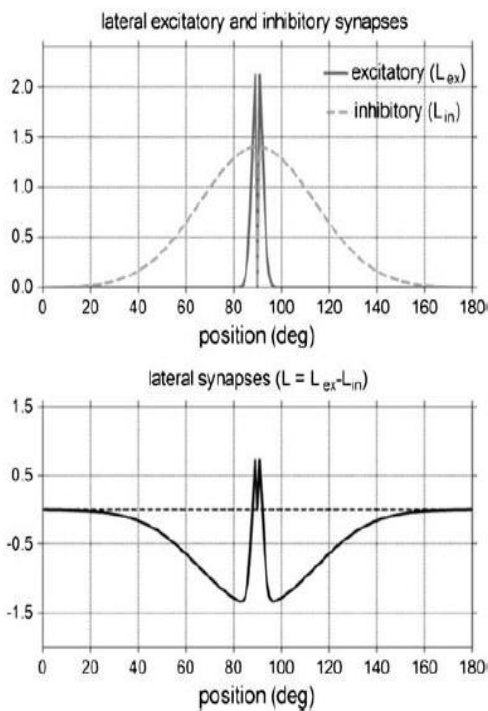


Fig. 8. Distribuzione a “cappello messicano” delle sinapsi laterali di un neurone.

L'equazione del primo ordine che descrive, per ogni j -esima posizione, l'andamento della rete è:

$$\tau \frac{dy_j^m(t)}{dt} = -y_j^m(t) + F(u_j^m(t))$$

dove τ è la costante di tempo del neurone e m specifica l'area corticale ($m = a$ oppure v , a seconda che sia acustica o visiva); $F(u)$ rappresenta una relazione di tipo sigmoideale.

Nel presente modello, l'incertezza dello stimolo è codificata nella risoluzione (spaziale o temporale) della risposta neurale. Di conseguenza, l'input sensoriale più preciso può

esercitare la sua influenza eccitatoria sull'altro (Fig. 9). A causa della sua scarsa accuratezza, la modalità meno precisa presenta una moderata eccitazione sotto-soglia in corrispondenza dello stimolo più preciso in un intervallo fino a 100 ms. Questi valori si addicono all'ipotesi applicata dal modello, secondo cui la percezione illusoria si verifica se i due segnali acustici vengono presentati con una finestra temporale paragonabile ai 100 ms.

Un altro effetto che il modello è in grado di simulare è il cambiamento percepito nella durata (Scheier et al., 1999) e nell'intensità (Bolognini et al., 2007) di uno stimolo visivo indotto da un ingresso uditivo simultaneo. Questo effetto è evidente, per esempio, confrontando i casi B e C in Fig. 10; inoltre sembra simmetrico, nel senso che uno stimolo visivo può anche prolungare la durata di un segnale sonoro (confrontare A e C in Fig. 10). Infine, l'analisi di sensitività effettuata sul parametro σ mostra che modellare i risultati sia in gran parte indipendente dal valore di questo parametro, a condizione che $\sigma_a \gg \sigma_v$, un presupposto che è fortemente sostenuto dai dati sperimentali, sia per A1 che per il CS.

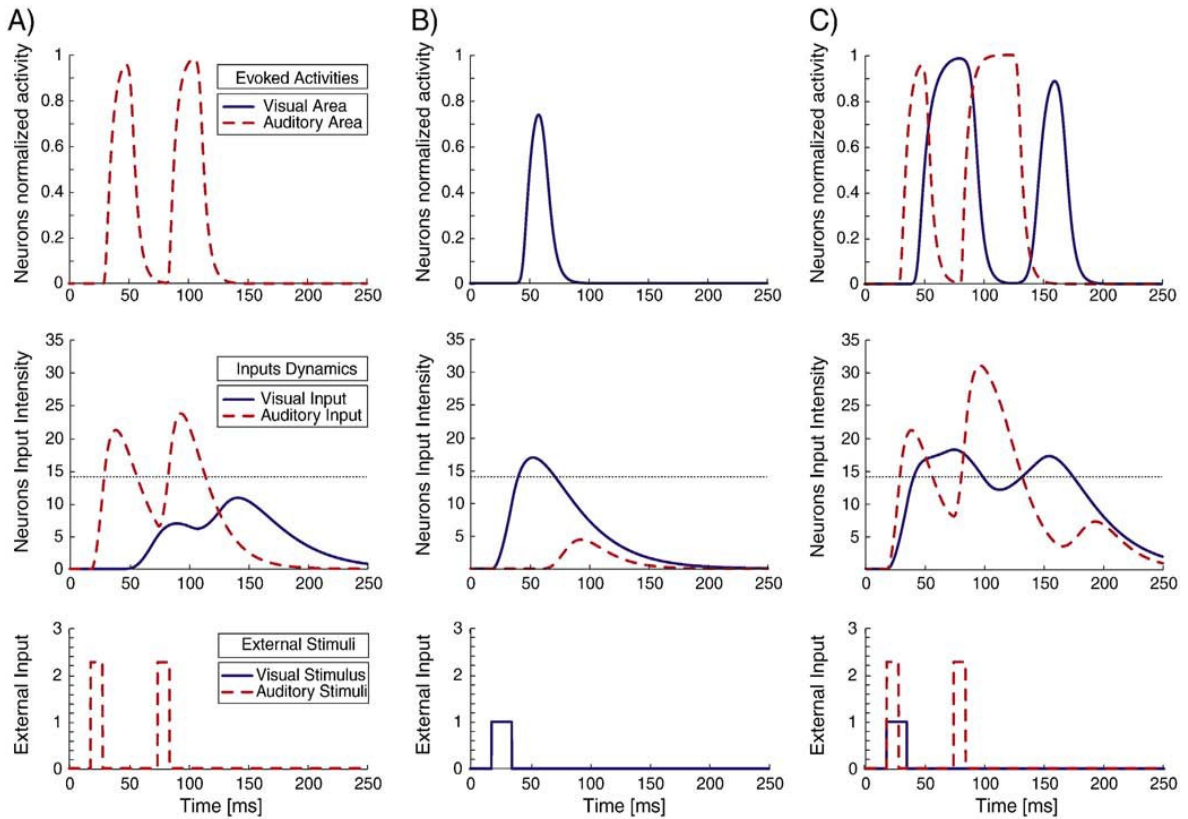


Fig. 9. A) la rete viene stimolata con uno stimolo uditivo doppio (due beep). B) la rete viene stimolata con un singolo stimolo visivo (un flash). C) la rete viene stimolata con un ingresso cross-modale, implementato come la combinazione dei due stimoli unisensoriali precedenti (due beep e un flash). La prima riga mostra l'attività nelle aree uditive (linea rossa) e in quelle visive (linea blu) del modello, evocati dalla diversa configurazione degli stimoli raffigurati nelle ultime righe. Le righe centrali mostrano la dinamica temporale degli ingressi complessivi rispetto alla soglia di attivazione (linea tratteggiata nera). Le due configurazioni unisensoriali, di per sé, sono in grado di suscitare un'attività rilevante solo nella relativa area specifica, ma, se combinate, come in C), il risultato è la nascita di un ulteriore picco "illusorio" dell'attività visiva, indotta dall'effetto del secondo beep, che aggiunge una componente cross-modale all'input eccitatorio nell'area visiva.

Grazie a questa breve analisi del modello di Cuppini et al., siamo riusciti a dimostrare l'importanza che le reti neurali artificiali stanno acquisendo negli ultimi decenni, con la speranza che il progresso delle tecnologie possa aiutarci ulteriormente nella comprensione dei meccanismi neurali sottostanti l'integrazione multi-sensoriale.

Conclusioni

In molteplici regioni cerebrali, sono stati individuati neuroni in grado di rispondere a stimoli provenienti da diverse modalità sensoriali (Stein & Meredith, 1993; Meredith, 2002; Stein et al., 2004). L'integrazione multisensoriale è la capacità del nostro cervello di integrare le informazioni provenienti dalle diverse modalità sensoriali, ed è fondamentale per un'accurata esperienza sensoriale e per l'interazione efficace con l'ambiente: migliora il rilevamento degli stimoli esterni, risolve le situazioni di conflitto, accelera la reattività, facilita i processi di recupero della memoria e il rilevamento degli oggetti.

Come conseguenza della sua eccezionale importanza sul comportamento, il problema dell'integrazione multisensoriale è da tempo riconosciuto come un aspetto cruciale delle neuroscienze (Calvert, Spence, e Stein, 2004; Stein & Meredith, 1993). Un punto di vista tradizionale, riportato dalla maggior parte dei testi delle neuroscienze fino a pochi decenni fa, assume che i singoli sensi siano prima trattati separatamente nelle aree unisensoriali per estrarre le loro informazioni, e solo successivamente combinati nelle fasi di analisi successive, nelle aree di associazione multisensoriali del cervello.

Dati recenti, tuttavia, soprattutto per quanto riguarda le corteccie primarie, mettono in discussione questa visione tradizionale, mostrando che anche le prime aree percettive corticali (come la corteccia visiva primaria, V1, o la corteccia uditiva primaria, A1) ricevono input da altre aree unisensoriali o da aree associative multisensoriali e mostrano alcuni comportamenti multisensoriali (Schroeder & Foxe, 2005).

Un altro aspetto importante è che l'integrazione multisensoriale può esibire proprietà diverse, a seconda del compito specifico a cui i neuroni multisensoriali partecipano; ad esempio, per eseguire compiti spaziali, le informazioni vengono integrate per gestire il proprio spazio peripersonale, oppure per raggiungere e manipolare gli oggetti, o ancora per spostare la testa e gli occhi.

Un metodo recente di indagine sulle interazioni cross-modali, siano esse su soggetti sani o patologici, è quello che utilizza le illusioni sensoriali, in cui una singola informazione è resa veicolata da due modalità, poste in conflitto tra loro. Sono innumerevoli gli studi che negli ultimi decenni si sono proposti di analizzare, attraverso i fenomeni illusori, queste interazione cross-modali.

Nel presente lavoro abbiamo trattato alcuni esempi di questi test, quali la sound-induced flash illusion per gli stimoli audio-visivi, o le body ownership illusions (BOIs), che coinvolgono anche la propriocezione e il senso di appartenenza corporeo. Nel primo caso parliamo di un'illusione visiva indotta dal suono: un singolo flash se accompagnato da molteplici segnali acustici (beep), viene percepito come multiplo (Shams, 2002). Nel secondo caso, invece, i soggetti hanno percepito oggetti non-corporei (ad es. elementi artificiali) come fossero appartenenti al proprio corpo, quando essi sono stati presentati con gli stimoli crossmodali applicati alla parte reale ma nascosta del corpo e alla sua controparte falsa.

Tutti i dati ricavabili da questi test devono essere analizzati, nel tentativo di comprendere e spiegare i meccanismi neurali che governano queste percezioni erranee. L'integrazione multisensoriale, infatti, opera a diversi livelli cerebrali: in strutture sottocorticali (in particolare il collicolo superiore), in cortecce associative di livello superiore (ad esempio, regioni parietali posteriori), e anche nelle prime aree corticali (cortecce primarie), tradizionalmente considerate come puramente unisensoriali.

A causa dei complessi meccanismi non lineari dei fenomeni integrativi presenti nel cervello, può essere utile applicare diverse tipologie d'analisi per una loro più completa e corretta comprensione. Qui ci siamo concentrati su tre tipologie distinte.

Il NeuroImaging ha come obiettivo principale la comprensione delle funzioni specifiche delle diverse aree cerebrali e come queste interagiscono in un sistema integrato di reti neurali. È uno strumento di primaria importanza nelle neuroscienze cognitive e in neuropsicologia. Tale metodo permette di visualizzare l'impatto che i processi mentali hanno sull'attività cerebrale e le modifiche che questa subisce ogniqualvolta il cervello svolge un compito. In particolare, lo studio dell'elettroencefalografia (EEG) e dei potenziali evento-correlati (ERP) di Mercier et al. (2014) ha permesso di valutare gli effetti della stimolazione uditiva nella corteccia sensoriale visiva. I ricercatori hanno potuto affermare che esistono effetti netti sul miglioramento dell'integrazione multisensoriale. Le teorie bayesiane affermano che il cervello rappresenta le informazioni sensoriali in maniera probabilistica, in forma di distribuzioni di probabilità. I modelli bayesiani si sono dimostrati efficaci nella costruzione di teorie computazionali per la percezione e il

controllo sensori-motorio; inoltre, la psicofisica sta fornendo un crescente numero di evidenze che ci portano ad affermare che i calcoli percettivi umani sono bayes-ottimali. Essi presuppongono che il cervello esegua un'integrazione ottimale, calcolando le probabilità a posteriori dell'evento, dato un insieme di informazioni incerte.

Questa osservazione, insieme al lavoro comportamentale e computazionale su cui si basa, ha implicazioni fondamentali per le neuroscienze, in particolare nel modo in cui concepiamo i calcoli neurali e la natura delle rappresentazioni neurali delle variabili percettive e motorie.

Più precisamente, questi risultati suggeriscono che quando le informazioni spazio-temporali ricavabili dai sensi sono sufficientemente congruenti, il sistema nervoso riesce a dedurre una causa comune per le sensazioni, portando all'esperienza di una fonte unificata. Ciò è particolarmente evidente negli studi che combinano segnali uditivi e visivi (Bresciani et al., 2006; Shams et al., 2005). Anche se è chiaro che l'integrazione multisensoriale realizza una performance quasi-ottimale nel trattare l'incertezza degli stimoli, molto poco si sa circa i meccanismi neurali impiegati.

Ultimo, ma non meno importante, il modello delle reti neurali artificiali guarda al problema con uno sguardo più ingegneristico, simulando i meccanismi neurali biologici in un algoritmo computazionale. In particolare, il modello analizzato in questo lavoro (Cuppini et al., 2014) è stato costruito per suggerire un'implementazione neurale in grado di imitare alcuni aspetti dell'integrazione multisensoriale e ha fornito ulteriori prove sulla sound-induced flash illusion, senza però utilizzare una connessione diretta con l'inferenza bayesiana.

Questa panoramica apre la strada alle sfide future, come quella di conciliare i modelli neurofisiologici e quelli bayesiani in una teoria unificante, e di stimolare la futura ricerca sia nei domini teorici che in quelli applicativi.

Nel presente lavoro, i parametri del modello (dimensioni del campo recettivo e costanti di tempo) sono stati forniti senza preoccuparsi del loro rapporto con l'incertezza dello stimolo (vale a dire, con le probabilità di verosimiglianza) e con le informazioni a priori. Shams et al. (2005), per esempio, hanno stimato i priori dai dati e hanno dimostrato che un modello che non tiene conto di alcuna probabilità a priori risulta avere una performance scarsa. Al fine di raggiungere un'inferenza bayesiana, questi aspetti del modello devono riflettere le

statistiche dell'ambiente esterno, cioè il modo in cui vengono generati i dati.

Una possibile applicazione futura del modello di Cuppini è quella di indagare come i parametri possono essere regolati automaticamente, per fornire un'inferenza bayesiana, data una certa probabilità condizionale dei dati e data un'aspettativa a priori dello stimolo. Se si raggiungerà tale obiettivo, potremmo essere in grado di verificare se alcune regole di plasticità delle sinapsi possono regolare automaticamente i propri pesi, per realizzare una stima ottimale.

Per affrontare il problema dell'inferenza causale, una classe innovativa di modelli bayesiani è stato proposto negli ultimi anni (Shams & Beierholm 2010). In questi modelli l'osservatore non solo stima le variabili sconosciute, ma fa anche una stima sulla possibile struttura causale che ha prodotto tali variabili. Forse, nello studio della percezione multisensoriale, il modello più avanzato di questo tipo è stato utilizzato da Kording et al. (2007). In questo lavoro, l'osservatore ideale non solo stima la posizione dei segnali, ma deduce anche se i due stimoli provengono dalla stessa fonte, fornendo così accurate previsioni dei risultati in un compito di localizzazione audio-visiva.

Tutti questi studi possono considerarsi di buon auspicio per la ricerca futura, che dovrà essere sempre più mirata, al fine di unificare definitivamente le teorie computazionali a quelle bayesiane.

La fusione multisensoriale, con i diversi aspetti correlati ad essa, rappresenta un problema di crescente importanza, e ne acquisirà sempre di più nei prossimi anni, sia per quel che concerne l'implementazione delle reti neurali, sia nel campo delle neuroscienze teoriche, sia nei suoi domini maggiormente applicativi, quali la robotica.

Bibliografia

Magnee MJ, de Gelder B, van Engeland H, Kemner C, *Audiovisual speech integration in pervasive developmental disorder: evidence from event-related potentials*, J Child Psychol Psychiatry, 2008

Stevenson R, Wallace A. *Multisensory temporal integration: task and stimulus dependencies*. Exp Brain Res, 2013

Shams L, Kamitani Y, Shimojo S. *Visual illusion induced by sound*. Cognitive Brain Research, 2002

Mongillo EA, Irwin JR, Whalen DH, Klaiman C, Carter AS, Schultz RT. *Audiovisual processing in children with and without autism spectrum disorders*. J Autism Dev Disord, 2008

Stevenson RA, Siemann JK, Schneider BC, Eberly HE, Woynaroski TG, Camarata SM, Wallace MT. *Multisensory temporal integration in autism spectrum disorders*. The Journal of Neuroscience, 2014

Ryan A. Stevenson, Justin K. Siemann, Tiffany G. Woynaroski, Brittany C. Schneider, Haley E. Eberly, Stephen M. Camarata, Mark T. Wallace, *Evidence for Diminished Multisensory Integration in Autism Spectrum Disorders*, J Autism Dev Disord, 2014

Temple Grandin, *Pensare in immagini e altre testimonianze della mia vita di autistica*, Edizioni Erickson, Trento, 2001

Gabriel Levi, *Linee guida per l'autismo, raccomandazioni tecniche-operative epr i servizi di neuropsichiatria dell'età evolutiva*, documento fornito dalla Società Italiana di Neuropsichiatria dell'infanzia e dell'adolescenza

Dee Unglaub Silverthorn, *Fisiologia umana, un approccio integrato*, ed. Pearson

Giuseppe Costantino Budetta, *Sensazione – Percezione – Consapevolezza*

Bogdashina O., *Le percezioni sensoriali nell'autismo e nella sindrome di Asperger*, ed. Uovonero, 2012

Ozdem Erturk, *Startle and blink reflex in high functioning autism*, Neurophysiologie Clinique/Clinical Neurophysiology, marzo 2016

Martha D. Kaiser, Maggie Shiffrar, *The visual perception of motion by observers with autism spectrum disorders: a review and synthesis*, Psychonomic Bulletin & Review 2009

Gomes E, Pedroso FS, Wagner MB, *Auditory hypersensitivity in the autistic spectrum disorder*, Pró-Fono Revista de Atualização Científica. 2008

Jerry A. Fodor, *The Modularity of Mind*, Cambridge: MIT Press, 1983

Stein, B., & Meredith, *Merging of senses*, Cambridge: MIT Press, 1993

I.P. Howard, W.B. Templeton, *Human Spatial Orientation*, Wiley, London, 1966

H. McGurk, J.W. MacDonald, *Hearing lips and seeing voices*, Nature 264, 1976

J.C. Hay, H.L. Pick, K. Ikeda, *Visual capture produced by prism spectacles*, Psychonomid. Sci., 1965

Van der Smagt MJ, Van Engeland H, Kemmer C., *Brief report: can you see what is not there? Low-level auditory-visual integration in autism spectrum disorders*. J Autism Dev Disord, 2007

Foss-Feig JH, Kwakye LD, Cascio CJ, Burnette CP, Kadivar H, Stone WL, Wallace MT. *An extended multisensory temporal binding window in autism spectrum disorders*, Exp Brain Res, 2010

Paola Invernizzi, Martina Gandola, *What is mine? Behavioral and anatomical dissociations between somatoparaphrenia and anosognosia for hemiplegia*, Behavioural Neurology, 2013

Giuseppe Vallar, Roberta Ronchi, *Somatoparaphrenia: a body delusion*, Exp Brain Res, 2009

Jenkinson et al., *Body ownership and attention in the mirror: Insights from somatoparaphrenia and the rubber hand illusion*, 2013

Beierholm U., Shams L., *Bayesian priors are encoded independently from likelihoods in human multisensory perception*, Journal of Vision, 2009

Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, et al., *Causal Inference in Multisensory Perception*, 2007

Wei Ji Ma, Jeffrey M Beck, Peter E Latham & Alexandre Pouget, *Bayesian inference with probabilistic population codes*, Nature Neuroscience, 2006

David C. Knill & Alexandre Pouget, *The Bayesian brain: the role of uncertainty in neural coding and computation*, Neurosciences, 2004

Majed Samad, Albert Jin Chung, Ladan Shams, *Perception of Body Ownership Is Driven by Bayesian Sensory Inference*, PlosOne, 2015

Ladan Shams & Ulrik R. Beierholm, *Causal inference in perception*, Cognitive Sciences, 2010

Fiebelkorn et al., *Ready, Set, Reset: Stimulus-Locked Periodicity in Behavioral Performance Demonstrates the Consequences of Cross-sensory Phase Reset*, J Neurosci, 2011

Mercier et al., *Auditory-driven phase reset in visual cortex: Human electrocorticography reveals mechanisms of early multisensory integration*, Neuroimage, 2013

Mauro Ursino, Cristiano Cuppini, Elisa Magosso, *Neurocomputational approaches to modelling multisensory integration in the brain: A review*, Neural Networks, 2014

Cristiano Cuppini, Elisa Magosso, Nadia Bolognini, Giuseppe Vallar, Mauro Ursino, *A neurocomputational analysis of the sound-induced flash illusion*, NeuroImage, 2014

Gianluca Smeraldi, *Introduzione alle reti neurali*, EMERNet

Rowland, B. A., Stanford, T. R., & Stein, B. E., *A model of the neural mechanisms underlying multisensory integration in the superior colliculus*, Perception, 2007

Magosso, E., Zavaglia, M., Serino, A., di Pellegrino, G., & Ursino, M., *Visuotactile representation of peripersonal space: a neural network study*, Neural Computation, 2010

wikipedia

specialeautismo.it

treccani.it

2001agsoc.it

cliccascienze.it

mypersonaltrainer.it

simplypsychology.org

Ringraziamenti

Eccoci qui, finalmente, a mettere il punto in fondo a questo lavoro, che è un po' come mettere il punto in fondo a un periodo lungo, forse troppo, della mia vita. È durata tanto, ma pensare che ogni sacrificio mi ha fatto arrivare a questo momento, mi fa capire che ne è valsa la pena.

Ringrazio innanzitutto la mia famiglia, per avermi supportato, per aver compreso che, per cause di forza maggiore, non ho mai potuto dare la priorità allo studio, per non avermi mai fatto pesare che, per l'ennesimo anno, dovevo rimandare la laurea, per avermi incoraggiato ad andare avanti, a non guardare il resto, ma ad essere orgogliosa di ciò che stavo facendo, perché se sono come sono è merito loro e se ho fatto tutto questo è anche per loro.

Ringrazio chi c'è sempre stato, chi c'è stato solo un po', e anche chi c'era ma ora non c'è più, perché in quasi 7 anni di università le amicizie cambiano, ci sono persone che purtroppo ho perso e altre che per fortuna ho trovato.

Ringrazio profondamente quel meraviglioso mondo che è la ginnastica artistica, e tutte le persone che ho conosciuto tramite essa, che mi accompagna da ormai 20 anni, e che mi ha insegnato ad essere caparbia e determinata, a sapere che con il sacrificio si raggiunge qualsiasi obiettivo.

Ringrazio la Beba, forse l'amica più vecchia che ho, che con la sua spontaneità mi insegna ogni giorno a vedere le cose un po' più facili, a stringere i denti ma a sorridere nei momenti difficili, che una soluzione c'è sempre, che vede il mondo a cuoricini (e anche questo ogni tanto ci vuole!), e con lei ringrazio tutta la sua famiglia, che in fin dei conti è diventata anche un po' la mia.

Ringrazio l'Ami che è sempre con me, per festeggiare e brindare. E ridere e scherzare, ma che è soprattutto con me quando le cose non vanno, quando si commuove più di me e quando mi capisce anche se non parlo, perché parlare a volte è una gran fatica!

E ringrazio la Je. Quando mi dicono che le amicizie non possono nascere da adulti, che se nascono è perché c'è un interesse dietro, non posso che pensare a lei, perché la considero un'anima affine e non so come avrei potuto affrontare questo percorso senza di lei; va a finire che mi tocca ringraziare anche ingegneria, perché senza di lei probabilmente non ci saremmo mai conosciute!

E ringrazio tutti, davvero tutti, perché se dovessi nominarvi non finirei più, e, che vogliate essere nominati o meno, sapete che sto parlando anche di voi! Tutti quelli che con una parola o un gesto mi hanno fatto capire di esserci, che conto per loro, che l'impegno che metto nelle cose ha un senso, se non altro quello di avere intorno delle persone meravigliose, ma anche tutti coloro che mi hanno messo i bastoni tra le ruote, perché se c'è una cosa che posso affermare è che non ho mai avuto una vita facile, c'è sempre stato un problema da risolvere, sempre qualcosa che mi ha fatto mettere me stessa in secondo piano.

E, allora, il ringraziamento più grande va a me stessa, perché questo percorso e questa Tesi mi fanno affermare, a dispetto di tutto e tutti, con serenità e convinzione, che io valgo.