13

La coscienza come problema scientifico tra filosofia e neuroscienze

di *Paolo Pecere*

13.1
Rinascita di un'indagine multidisciplinare: neuroscienza, neuroentusiasmo, filosofia

La relazione tra cervello e coscienza costituisce un tema sempre più presente nella ricerca neuroscientifica contemporanea. Il problema è rimasto a lungo implicito a margine delle ricerche su specifiche prestazioni cognitive o fenomeni patologici, in cui l'attivazione di "moduli" funzionalmente isolati veniva separata dal generale contesto dell'esperienza cosciente unitaria. Ma a partire dagli anni Novanta, anche grazie al rapido sviluppo di tecniche di osservazione non invasiva dell'attività cerebrale, sono state avanzate numerose ipotesi empiriche sui «correlati neurali della coscienza» (NCC)[1]. Questa indagine, condotta inizialmente da una ristretta cerchia di scienziati, ha sollecitato l'istituzione di centri di ricerca, associazioni e riviste dedicate allo studio interdisciplinare della coscienza, che coinvolgono anche filosofi e psicologi[2]. L'affermazione anche politico-culturale delle ricerche sul cervello ha favorito, parallelamente, la diffusione (anche popolare) dell'idea di un paradigma neuroscientifico onni-

1. Già secondo Kandel (2000, pp. 16-7), i progressi della neuroscienza «dell'ultimo decennio» stanno cominciando a svelare «l'enigma più profondo», cioè «la rappresentazione neurale della coscienza».

2. Il "Journal of Cognitive Neuroscience", fondato nel 1989 sotto la direzione di Michael Gazzaniga, dedica spazio a queste ricerche interdisciplinari. La rivista più esplicitamente dedicata al tema è il "Journal of Consciousness Studies", fondato nel 1994. Nello stesso anno è stata istituita la "Association for the Scientific Study of Consciousness", che ha anche una sua rivista open access, "Psyche".

comprensivo, in base al quale la comprensione del funzionamento del cervello promette di contribuire allo sviluppo di discipline come l'etica, il diritto, l'estetica, l'economia. Nel caso della coscienza, questo ha rilanciato speranze (e paure) sulla possibilità di manipolare – in base alla conoscenza dei meccanismi cerebrali – il contenuto stesso della nostra esperienza. A proposito della distanza tra questo orizzonte, talvolta vago, e le effettive capacità delle conoscenze disponibili lo storico Michael Hagner (2002) ha parlato di una "ciberfrenologia", che prometterebbe – aggiornando la frenologia del XIX secolo – di leggere nelle neuroimmagini "che cosa pensi" e "chi sei". La stessa infondata trasvalutazione dei dati ha fatto parlare psicologi e neuroscienziati di "neuromania" (Legrenzi, Umiltà, 2009) e di "neuromitologia" (Hasler, 2012). Mentre sono in atto nuovi progetti di ricerca neuroscientifici su larga scala, insomma, il bisogno di definire esattamente l'ambito e la portata di queste indagini appare sempre più associato a una responsabilità filosofica.

Prima di tutto bisogna chiedersi: che cos'è la "coscienza", di cui queste indagini vorrebbero scoprire i correlati neurali? Alcuni scienziati propongono una definizione di coscienza come proprietà unitaria, suscettibile come tale di indagine scientifica. Per esempio Giulio Tononi, uno dei massimi scienziati impegnati in questo campo, scrive:

> Tutti sanno che cosa sia la coscienza. È ciò che svanisce ogni notte quando cadiamo in un sonno senza sogni e riappare quando ci svegliamo o quando sogniamo. Così, la coscienza è sinonimo di esperienza – qualsiasi esperienza – di forme o suoni, pensieri o emozioni, riguardo al mondo o a noi stessi (Tononi, 2012a, p. 293).

Che la coscienza possegga un'omogeneità essenziale, che accomuna tali diverse componenti, è una tesi ben radicata nella filosofia moderna, da quando Descartes scrisse che la cosa pensante «dubita, intende, afferma, nega, vuole, non vuole, immagina, inoltre, e sente» (Descartes, 1641, trad. it. p. 28); e molti filosofi contemporanei si sono esplicitamente richiamati a Descartes nel sostenere il carattere unitario della coscienza. Per esempio John Searle, tra i protagonisti del dibattito contemporaneo, ha sostenuto che i processi coscienti si possono caratterizzare in genere per unità di contenuti eterogenei, specifica qualità soggettiva di questi ultimi, intenzionalità (riferimento a un oggetto, reale

o astratto), causalità. Sarebbe compito delle teorie neuroscientifiche spiegare come queste caratteristiche risultino dall'attività cerebrale.

Tuttavia molti filosofi hanno sostenuto che questioni molto diverse vanno sotto il titolo della "coscienza" e che ciò renderebbe inopportuno servirsi di un singolo concetto. Per esempio, come ha mostrato Block (1995), diverse sono la "coscienza fenomenica" – il "che cosa si prova a essere" (*what it is like to be*) un determinato essere cosciente, di cui parlava Nagel in un suo celebre articolo (1974) – e la "coscienza di accesso", cioè la disponibilità cognitiva di contenuti mentali, che rende possibile il ragionamento e l'azione controllata. Non si tratta di una semplice distinzione concettuale: casi clinici ed esperimenti neuropsicologici mostrano che vi sono stati di accesso a informazioni sull'ambiente e sullo stesso corpo del soggetto che non si accompagnano a nessun contenuto fenomenico: un esempio è la nota sindrome della "visione cieca" (*blindsight*) studiata da Humphrey (1970; 2006)[3]. D'altra parte, secondo Block (2011), si danno anche stati di coscienza fenomenica, come la percezione di particolari del campo visivo, che non vengono integrati nelle informazioni che il soggetto sa riportare e usare. Perciò anche i correlati neurali di questi diversi generi di stati dovrebbero essere di diverso genere. Dato, però, che l'osservazione di stati fenomenici presuppone il resoconto cosciente del soggetto, e dunque una qualche forma di accesso, si pone quello che Block chiama "problema della misura", cioè la ricerca di un protocollo sperimentale per aggirare questa sovrapposizione di stati e distinguere i correlati dei due tipi di coscienza.

Un problema più generale, messo in luce da molti filosofi, è che difficilmente i diversi tipi di processi mentali si possono far corrispondere uno-a-uno a specifici correlati nell'attività cerebrale, a causa del carattere distribuito e variabile (anche a livello individuale) di quest'ultima[4].

3. Si tratta di un fenomeno tipico di soggetti che hanno subito lesioni nelle aree della corteccia occipitale, che sono necessarie per l'elaborazione cosciente dei segnali visivi, mentre hanno un apparato visivo altrimenti intatto. Queste persone non hanno coscienza di vedere nulla, ma sono capaci di evitare ostacoli e svolgere altri comportamenti che presuppongono un'elaborazione inconscia dell'informazione visiva. Il *blindsight* è dunque un fenomeno in cui è massimamente vistosa la dissociazione tra disponibilità di informazioni e coscienza.

4. Sui problemi metodologici riguardo alla ricerca di NCC cfr. i bilanci di Chalmers (2000) e Noë, Thompson (2004).

Queste divergenze non devono far pensare che tra i filosofi non sussista alcuna unanimità. Per esempio, è tesi condivisa pressoché unanimemente che la coscienza dipenda dall'attività di un sistema fisico (come un organismo immerso nell'ambiente) e che essa non implichi l'esistenza di alcuna entità immateriale. Anche questa tesi chiama in causa la filosofia cartesiana, stavolta colpevole – secondo un'accusa che non manca quasi mai nei testi introduttivi di filosofi e neuroscienziati – di aver professato un dualismo metafisico privo di possibili conferme sperimentali. Tuttavia, come vedremo, la tensione tra materialismo ed eredità cartesiana continua ad animare le discussioni sul programma di ricerca della "scienza della coscienza", rimandando a un antefatto filosofico non ancora superato. In generale, a differenza di altre vicende filosofico-scientifiche della modernità, la storia del problema della coscienza non appare come un progresso lineare, ma – per usare un'espressione di Daniel Dennett (1985, p. 3) – come l'oscillazione di un pendolo, in cui si continua a tornare su alternative già discusse. Pertanto, sarà utile ricordare alcuni momenti storici del lungo percorso che ha portato al costituirsi di questo campo di indagine multidisciplinare.

13.2
Origini moderne: organo dell'anima, localizzazione e limiti della neuroscienza

In filosofia il problema del rapporto coscienza-cervello ha una lunghissima tradizione, che per secoli si è intrecciata allo sviluppo della medicina e della fisiologia, e alcune delle tesi storicamente avanzate per definire il rapporto tra coscienza e cervello fanno ancora da sfondo al dibattito contemporaneo.

Se si vuole individuare il momento storico in cui le questioni ancora aperte hanno la loro origine bisogna indicare senz'altro la comparsa della filosofia cartesiana. Per un verso, infatti, Descartes fu tra i primi ad applicare le idee del meccanicismo allo studio dell'uomo: considerando il corpo come «macchina» e mettendo in rilievo le analogie tra i movimenti degli automi meccanici e quelli involontari dell'uomo (Descartes, 1662, trad. it. p. 120), attraverso uno studio anatomico di prima mano. In generale, Descartes diede formulazione esemplare a un ideale esplicativo meccanicistico che negli stessi anni alcuni filosofi – come Hobbes – ritennero sufficiente alla comprensione dell'in-

tera attività mentale, sviluppandolo in senso materialistico[5]. D'altra parte, sostenendo che la mente – intesa come «tutto ciò che ha luogo in noi e di cui siamo coscienti» (Descartes, 1644, trad. it. p. 7) – sia rigorosamente distinta dal corpo, egli rese possibile la posizione del problema: come è possibile l'unità di mente e corpo? Questo problema, che naturalmente era già presente nella filosofia platonica, e che la filosofia aristotelica risolveva con il concetto di forma, si doveva porre ormai nei termini della nuova scienza della natura.

Mentre il problema ebbe ben note elaborazioni metafisiche – si pensi a Leibniz, Spinoza e Malebranche, che ebbero peraltro forti interessi per l'aspetto fisiologico della questione (cfr. Scribano, 2015) – gli anatomisti e i fisiologi lo affrontarono da subito, e poi per tutto il XVIII secolo, nei termini della ricerca di un "organo dell'anima", cioè di quella parte del cervello in cui la mente avrebbe agito. Le posizioni variarono a seconda delle diverse versioni del meccanicismo, che potevano comportare l'inerzia ma anche l'attività della materia (cfr. Clericuzio, 2015). Un elemento filosofico decisivo di queste ricerche fu la tesi lockiana della concepibilità di una "materia pensante", che sollevò una discussione capace di prolungarsi dalle dispute teologiche sull'immortalità dell'anima fino al vitalismo materialistico francese[6]. Il suo correlato scientifico era il concetto di *active power*, molto diffuso nella filosofia naturale inglese, che Newton aveva adottato per spiegare fenomeni come l'attrazione e la vita, e che da allora rimase un tema centrale del newtonianismo (cfr. Heimann, McGuire, 1971; Gabbey, 2002). In questo contesto – tra scienziati e filosofi come Maupertuis, Diderot e La Mettrie – fu ripresa con decisione l'elaborazione di alternative al dualismo metafisico e si definì la questione, ancora discussa nelle neuroscienze contemporanee, se la coscienza possa ritenersi una «qualità generale della materia» o «un prodotto dell'organizzazione» della materia stessa[7].

5. Non a caso fu proprio la diffusione del meccanicismo cartesiano in Inghilterra a dare occasione, in un contesto originariamente metafisico, alla nascita del termine "materialismo", per designare la tesi di chi ammette la sola esistenza di sostanze corporee, e alle polemiche che esso non ha smesso di innescare. Il primo a usarlo sarebbe stato il platonico Henry More nel 1668 (Bloch, 1995).

6. Sul passaggio tra dibattito sull'anima e materialismo, tra XVII e XVIII secolo, cfr. Thomson (2008). Sulle neuroscienze del XVIII secolo cfr. Whitaker, Smith, Finger (2007).

7. Cfr., per esempio, Maupertuis (1768, pp. 147, 149) e Diderot (1987, p. 105), e le discussioni in Thomson (2008) e Wolfe (2010).

Le indagini sull'organo dell'anima costituirono un campo di ricerche e riflessioni disciplinarmente anfibio, come conferma il fatto che due interventi indipendenti, uno filosofico e l'altro anatomico, ne segnarono la fine (Hagner, 2008, pp. 25-118). Il primo fu quello di Kant che, in uno scritto comparso in appendice a un'opera fisiologica sull'organo dell'anima, considerò «non soltanto insolubile», ma «contraddittorio» il compito di localizzare la mente, sottolineando la differenza irriducibile tra esperienza interiore ed esperienza della realtà fisica nello spazio (Kant, 1796, p. 35). L'intervento di Kant fu anche il primo a rilevare un conflitto disciplinare tra «facoltà», cioè tra la psicologia (che apparteneva all'epoca alla facoltà filosofica), e, dall'altra parte, l'anatomia e la fisiologia (che appartenevano alla facoltà medica), avanzando un argomento originale per risolverlo: nonostante il successo delle indagini sui correlati fisiologici dell'attività mentale – sosteneva Kant –, si dà un genere di principi, collegati all'attività della «coscienza pura», che resteranno sempre eterogenei rispetto alle leggi empiriche individuate dalla scienza della natura. Ne sono esempi i principi del diritto puro e quelli della logica, che regolano l'attività con cui la coscienza unifica i dati dei sensi (ivi, pp. 31-2). Si trattava di un argomento epistemologico, del tutto indipendente da qualsiasi forma di dualismo delle sostanze (che Kant rifiutava) e diverso da quello – già diffuso nel XVIII secolo – secondo cui le analisi psicologiche precederebbero logicamente le indagini fisiologiche[8]. In base alla sua insistenza sull'irriducibilità delle leggi, dunque al suo carattere normativo e non descrittivo, esso si può considerare il prototipo di un'intera famiglia di argomenti ancora diffusi contro il riduzionismo[9].

Sul piano anatomico, negli stessi anni, fu l'opera di Gall (1798) a proporre di rinunciare all'indagine sulla sede dell'anima, in quanto inadatta a una trattazione scientifica, proponendo invece una sistematica ricerca sulla localizzazione cerebrale di diversi «organi» responsabili delle diverse funzioni in cui si scompone l'attività mentale. Quella della localizzazione delle "facoltà" nel cervello non era un'idea nuova, e la specifica tassonomia adottata da Gall (che localizzava, per esem-

8. Cfr. Hatfield (1995), che sottolinea come la psicologia empirica settecentesca fosse già largamente non materialistica.

9. Su questo argomento di Kant, il suo contesto di origine e la sua fortuna, rimando a Pecere (2015).

pio, la tendenza a uccidere e l'orgoglio), fu da subito oggetto di forti critiche; inoltre il programma di Gall prevedeva la possibilità di misurare le capacità mentali in base all'esame del cervello, e fu perciò subito oggetto di accuse di materialismo e determinismo. Ma proprio perché la teoria di Gall fece discutere in tutta Europa (e in America) essa si può considerare un momento chiave – se non l'evento inaugurale – di un programma di ricerca che avrebbe celebrato con l'elettrofisiologia e l'elettroencefalografia, nell'ultimo quarto del secolo, i suoi grandi successi. Si cominciò a parlare di "funzioni del cervello" piuttosto che di funzioni spirituali o mentali e si andò realizzando quello che è stato chiamato un processo di "cerebralizzazione" dell'uomo[10].

In questo arco di tempo si andarono anche formando molte delle categorie filosofiche, generalmente antidualistiche, che ancora orientano il dibattito contemporaneo: la metafisica monistica della *Naturphilosophie* (almeno nei primi decenni del secolo) e il materialismo ebbero larga parte nell'ispirare le ricerche empiriche sulle basi fisiologiche della coscienza. Una tendenza pragmaticamente materialistica era già dominante tra gli scienziati dei secoli precedenti, ma molti esitavano ad abbracciare le teorizzazioni più estreme di chi, come Cabanis (1802, p. 138), paragonava il pensiero a una «secrezione» del cervello. Il materialismo era infatti una posizione pubblicamente scomoda per ragioni politiche, soprattutto dopo il 1848. Inoltre restava viva, particolarmente tra gli scienziati tedeschi, la memoria delle cautele avanzate da Kant sulle capacità esplicative della fisiologia, che tornarono in discussione negli anni in cui la psicologia cominciava ad acquisire una sua collocazione accademica e lo "psicologismo" veniva discusso criticamente dai filosofi (Kusch, 1995).

Espressione teorica di tutte queste tensioni fu la cosiddetta "disputa sull'*Ignorabimus*" (*Ignorabimus-Streit*), inaugurata dal discorso sui limiti della conoscenza della natura tenuto nel 1872 dal massimo fisiologo tedesco dell'epoca (e pioniere dell'elettrofisiologia), Emil Du Bois-

10. Cfr. ancora Hagner (2008). Secondo Vidal (2006, p. 377), già tra XVII e XVIII secolo si sarebbe costituita la nozione di un "soggetto cerebrale", secondo cui l'identità individuale è stabilita dal cervello. Laddove Vidal segnala l'attestazione di questa nozione, che peraltro è stata spesso associata a idee materialistiche e meccanicistiche, Hagner insiste sul suo dispiegamento in un programma di ricerca anatomico-fisiologico dominato dalla localizzazione delle funzioni, che è sostanzialmente ancora in corso. Sulla presenza centrale di temi neurologici nella cultura settecentesca si veda anche Rousseau (1990; 2004).

Reymond, che presentava la spiegazione scientifica della coscienza come un problema insolubile. Du Bois-Reymond (1872, p. 122) scriveva che la conoscenza perfetta dei meccanismi del cervello «non ci svela altro che materia in movimento. Nessuna disposizione o movimento di particelle materiali ci permette di gettare un ponte nel regno della coscienza». In particolare egli si riferiva alla impossibilità di spiegare mediante processi fisico-chimici esperienze come dolore, piacere, sapori, odori, suoni e colori (ivi, p. 123), insomma quella che oggi si chiama "coscienza fenomenica". Du Bois-Reymond ricordava che simili tesi circolavano in filosofia almeno dal XVII secolo, citando in proposito le celebri pagine della *Monadologia* di Leibniz sui limiti del meccanicismo[11]. Ma egli non intendeva affatto riaprire la via verso concezioni metafisiche del passato, riabilitando per esempio il vecchio problema del commercio tra anima e corpo, ma piuttosto additare uno specifico «iato» epistemologico (ivi, p. 117) che riguardava (e circoscriveva) la ricerca scientifica di cui era portavoce: bisognava concedere senz'altro che la coscienza fosse inseparabile dall'attività del cervello, ma restava incomprensibile *come* il contenuto della coscienza potesse formarsi su questa base (Du Bois-Reymond, 1880, p. 395).

Veniva così formulato quello che circa un secolo dopo è stato presentato, in termini sostanzialmente identici, come «iato esplicativo» (*explanatory gap*) o come "problema difficile" della coscienza[12]. E in effetti, nelle discussioni che seguirono l'intervento di Du Bois-Reymond, si riconosce un orizzonte problematico ancora attuale[13]. Al

11. Un *ignoramus* sul nesso tra cervello e coscienza era addirittura la posizione dominante nella filosofia del XVIII secolo. Vi si poteva giungere riprendendo ancora i termini cartesiani, per esempio nella grande *Encyclopédie* francese, dove Diderot scrive: «come possono degli oggetti che producono soltanto movimento nei nervi imprimere delle idee nella nostra anima? Per risolvere tale questione bisognerebbe [...] penetrare nel mistero inesplicabile che consiste nell'unione meravigliosa di queste due sostanze» (Diderot, 1987, p. 489). Ma altrettanto diffuso era un *ignoramus* empiristico, legato all'ipotesi lockiana della materia pensante. Du Bois-Reymond stesso ricordava come questa tesi fosse stata sostenuta da La Mettrie, presentandolo come precursore dell'*Ignorabimus* (1875, p. 197).

12. Rispettivamente Levine (1983) e Chalmers (1996). Cfr. *infra* nota 28 per alcune prove di questa affermazione. Che molti dei dibattiti contemporanei della *philosophy of mind* siano una ripetizione (*replay*) delle discussioni di quest'epoca è stato già sostenuto da Tennant (2007).

13. Sull'*Ignorabimus-Streit* e le sue appendici novecentesche cfr. almeno i saggi raccolti in Bayertz, Gerhard, Jaeschke (2007).

problema posto dal grande fisiologo si ricollegarono subito psicologi come Wilhelm Wundt (1873, II, p. 331) e filosofi come Wilhelm Dilthey (1883, pp. 10, 13), interessati a sottolineare – in linea con una tesi già presente nella cultura tedesca dell'epoca, di cui abbiamo già rilevato la derivazione kantiana[14] – l'eterogeneità tra i «domini» dell'esperienza psichica e quello delle scienze della natura. D'altra parte, molti criticarono la tesi della limitazione considerandola scientificamente infondata. Tra questi vi fu Haeckel, che accusò Du Bois-Reymond di intenzioni reazionarie e filoclericali, rispondendo con il suo *Die Welträthsel* (1899) – un pamphlet piuttosto divulgativo che ebbe una enorme diffusione – dove affermava che il problema della coscienza era stato *già* risolto dalla conoscenza scientifica: la coscienza e ogni altra attività mentale complessa non erano altro che «*funzioni delle cellule gangliari o neuroni della corteccia cerebrale*» (ivi, p. 11).

Sullo slancio dell'attività di Haeckel prese vita la "società monistica", che ebbe notevole rilievo nella vita culturale tra i due secoli. Tra i giovani filosofi ispirati dal monismo vi fu Rudolf Carnap (1963a, p. 6) che, pur abbandonando questa metafisica, avrebbe mantenuto la tesi secondo cui il metodo scientifico è il solo adatto a ottenere una conoscenza fondata e che non vi possono essere «enigmi» scientifici propriamente insolubili. Queste idee sarebbero rimaste centrali nel movimento filosofico che dalla filosofia scientifica del Circolo di Vienna giunse, con l'emigrazione di Carnap e altri, negli Stati Uniti. Proprio il suo allievo Quine avrebbe sostenuto la necessità di riprendere in considerazione il «mistero» della coscienza, affermando che non esisteva ancora una descrizione della coscienza «in termini scientifici soddisfacenti» e prendendo posizione, contro il riduzionismo, a favore di un dualismo di concetti di derivazione kantiana (Quine, 1987, pp. 132-3)[15].

14. Per esempio, Wundt sottolineava la possibilità di coniugare il monismo ontologico già sostenuto da Theodor Fechner – e dunque il rifiuto del «dualismo dell'accadere psichico e dell'accadere naturale» – con il riconoscimento della differenza di «metodo» tra «analisi fisica» della sensazione e «scomposizione logica» dei nessi tra i fenomeni, concludendo: «necessità meccanica e necessità logica sono diverse non per essenza, ma per modo di considerare i fenomeni» (1863, pp. 199-200; cfr. Poggi, 1977, pp. 567-78).

15. Il riferimento diretto di Quine era il "monismo anomalo" di Davidson (1970), che ripresenta un argomento di tipo kantiano contro la riducibilità degli eventi mentali a leggi naturali. La riproposizione quineana del problema, d'altra parte, avrebbe ispirato la teoria biologica della coscienza di Edelman (cfr. PAR. 13.4.2).

Non si tratta di un esempio isolato. In generale, le tensioni problematiche manifestatesi con l'*Ignorabimus-Streit* ebbero molteplici sbocchi nel laboratorio della filosofia tedesca tra la fine del XIX e l'inizio del XX secolo. Diverse versioni di monismo non riduzionistico ebbero origine dalla combinazione di un realismo monistico con l'idea – originariamente kantiana – del "dualismo di concetti" irriducibile che separa mente e cervello. Tra i sostenitori di simili tesi vi furono il filosofo Alois Riehl (1879) e Moritz Schlick (1925). Quest'ultimo – contro l'*Ignorabimus* (ivi, § 35, p. 364) – avanzò un «monismo epistemologico», secondo cui i «sistemi di concetti» psicologici (riferiti al «processo cosciente vissuto») e i «sistemi di concetti» fisici si riferiscono alla medesima realtà, che nel vissuto viene colta immediatamente e nella fisica è rappresentata mediatamente (ivi, § 32, pp. 335-6). In base a questa ipotesi, con il progresso delle ricerche empiriche sui «correlati fisiologici delle grandezze psichiche», anche l'unità della coscienza sarebbe apparsa un giorno come uno tra i tanti nessi che collegano i fenomeni naturali (ivi, § 35, p. 371). Questa teoria sarebbe stata ripresa nella *identity theory* con cui Herbert Feigl, in Minnesota, diede un nuovo inizio al dibattito americano su coscienza e cervello (Heidelberger, 2003; cfr. PAR. 13.3.1).

Un ultimo aspetto da ricordare è l'ipotesi che, verso la fine del secolo, divenne dominante riguardo ai correlati neurali della coscienza. I problemi di una localizzazione esclusiva delle funzioni in aree delimitate del cervello, evidenziati dalle discussioni sull'area di Broca e altre analoghe scoperte, avevano gradualmente imposto l'immagine di un'attività distribuita del cervello, che venne presto recepita anche da filosofi e psicologi. Così Friedrich Lange nella sua celebre *Geschichte des Materialismus*, traendo un bilancio delle ricerche neurologiche sulla localizzazione delle funzioni cerebrali, riferì che la «coscienza» doveva dipendere da un «meccanismo centrale» che coinvolgesse «l'intero cervello» (Lange, 1875, II, pp. 343-4). A questa tesi si ricollegava il problema, che venne messo in rilievo, per esempio, da Wundt (1873, III, pp. 320-31) e James (1890, p. 180), di ricavare l'unità della coscienza da una molteplicità di «unità elementari»[16].

16. Si può appena notare, qui, l'importanza che per le indagini sul rapporto mente-corpo ebbe il parallelismo psicofisico di Fechner (a dispetto del suo monismo fortemente speculativo). Contro la monadologia ancora diffusa nella filosofia dei suoi tempi (e adottata anche, per esempio, da Mach), Fechner scrisse che «la mente non si

Riassumendo quanto emerso da questi rapidi cenni, molti elementi teorici fondamentali del dibattito contemporaneo si erano già formati in Europa alla fine del XIX secolo: 1. il rifiuto del dualismo ontologico in nome del materialismo o di diverse versioni di monismo; 2. la scomposizione delle "funzioni" cerebrali e la ricerca di una loro localizzazione; 3. la distribuzione dei correlati fisici della coscienza in una rete complessa di interazioni neurali, non limitate a una singola area cerebrale; 4. la ricerca di un'epistemologia capace di abbracciare gli aspetti qualitativi e intellettivi della coscienza. Il dibattito successivo – cui ci apprestiamo a dedicarci – ha ripreso e raffinato queste direzioni di ricerca, integrandole con nuove analisi, come quella su coscienza fenomenica e coscienza di accesso. Di recente, di fronte all'enorme accumulo di *big data* neuroscientifici, il ricordo di questo antefatto filosofico rivive nei richiami all'esigenza di interpretare i dati della neuroscienza alla luce di categorie filosofiche e psicologiche senza le quali quei dati rimarrebbero "ciechi" (Block, 2014).

13.3
La riscoperta della coscienza nella *philosophy of mind*

È un luogo comune storiografico affermare che la filosofia della mente di tradizione analitica abbia avuto un nuovo inizio negli anni Cinquanta del secolo scorso, con la teoria dell'identità tra mente e processi cerebrali sviluppata dai filosofi australiani Ullin Place (1956) e John Smart (1959) e, parallelamente, da Herbert Feigl in America (1967 – la prima edizione era del 1958)[17]. Si trattò in effetti, almeno nel caso di Feigl, della ripresa della discussione europea sul parallelismo psicofisico (Heidelberger, 2003; cfr. PAR. 13.2). Ma è vero che il dibattito successivo si è sviluppato largamente per contrasto con questo nuovo inizio, tentando di offrire versioni teoricamente meno vulnerabili di una concezione fondamentalmente materialistica, che era condivisa da quasi

trova negli atomi, ma nei sistemi» (Fechner, 1854, p. 47). Sulla dottrina e l'eredità di Fechner cfr. almeno Heidelberger (1993).

17. Cfr., per esempio, Marraffa (2002). Tra le tante ricostruzioni del dibattito americano fornite dai suoi protagonisti cfr. Kim (1998, pp. 1-27) e Searle (2004, pp. 41-106).

tutti i protagonisti del dibattito e che era stata condivisa del resto, nella generazione precedente, anche dai sostenitori del comportamentismo.

In questo contesto, il rapporto tra coscienza e cervello ha costituito una questione di forte divergenza tra gli studiosi, alcuni dei quali – come Putnam, Dennett, Paul e Patricia Churchland – hanno ritenuto quello della coscienza un concetto sostanzialmente estraneo all'approccio della conoscenza scientifica e carico di ingiustificati pregiudizi ontologici. Parallelamente, a partire dall'inizio degli anni Settanta, un ampio fronte antiriduzionista di filosofi – come Nagel, Block, Searle, Levine, McGinn, Chalmers – ha tratto una conclusione molto diversa dalla difficoltà di descrivere l'aspetto qualitativo della coscienza fenomenica con gli strumenti concettuali finora adottati nelle scienze cognitive, sottolineando (con diversi esiti) i limiti della descrizione scientifica della realtà.

Per comprendere la specificità di questo dibattito sulla coscienza, occorre sottolineare due aspetti. Per un verso esso è caratterizzato, fin dalle prime discussioni sulla teoria dell'identità, da un marcato uso di argomenti logico-metafisici, che spesso ha allontanato dal terreno delle indagini empiriche attraverso un abbondante uso di esperimenti mentali su mondi possibili e bizzarre anomalie cognitive; per un altro verso, l'appartenenza dei protagonisti di questo dibattito alla tradizione empiristica e naturalistica, che impone l'adozione di un'ontologia e metodologia coerenti con le teorie scientifiche più efficaci, ha determinato fin dall'inizio una particolare attenzione alla possibilità di una specificazione e selezione empirica delle ipotesi messe in campo. Vogliamo ora fornire alcuni esempi di questa tensione tra metafisica e apertura sperimentale, che ha portato i filosofi a invocare una risposta dalle neuroscienze – risposta che, come vedremo poi (PAR. 13.4), non è stata ancora sufficiente a risolvere la questione.

13.3.1. L'IDENTITÀ TRA COSCIENZA E PROCESSI CEREBRALI

Nell'articolo *Is Consciousness a Brain Process?* Place scriveva che l'analisi di concetti come conoscere, credere e ricordare in termini di disposizioni di comportamento, svolta da Wittgenstein e Ryle, è «fondamentalmente solida», mentre sembra esservi una costellazione di concetti, come «coscienza, esperienza, sensazione, immagini mentali, per cui un resoconto in termini di processi interiori è inevitabile» (Place, 1956, p. 44). L'obiettivo della teoria dell'identità era di resistere a questa tentazione e perfezionare il programma di un'analisi dei processi mentali che

facesse a meno del linguaggio interiore, laddove l'approccio comportamentista non sembrava in grado di farlo. Place condivideva ampiamente la prospettiva dei filosofi come Ryle e Wittgenstein, che avevano criticato l'esistenza di un «ambiente interiore» a cui si riferirebbero le osservazioni introspettive, denunciando gli errori a cui conduce questa «fallacia fenomenologica» e analizzando la teoria dell'interiorità in termini di comportamento linguistico. Place rettificava questa prospettiva con una semplice apertura verso nuove possibili indagini empiriche: «la tesi che la coscienza sia un processo nel cervello non può essere esclusa in base a ragionamenti logici». Anche Smart avrebbe ribadito questa impostazione logica della questione. Replicando all'obiezione secondo cui processi fisici e mentali avrebbero individuato proprietà i cui rispettivi termini non sono reciprocamente interscambiabili in proposizioni (*salva veritate*), e che dunque non sono identiche, Smart ricorreva alla distinzione fregeana tra il significato e il riferimento di un enunciato, che avrebbe permesso di considerare identico il referente dei due diversi discorsi (Smart, 1959, p. 146). Ma per il momento questa possibilità non era confortata da alcun dettaglio empirico, e Smart continuava a parlare di una generica identità *possibile* tra C-fibre neurali e percezione del dolore – rimandando dunque a un'evidenza empirica acquisita più di cinquant'anni prima, che era già stata alla base delle riflessioni di Wittgenstein sulla psicologia[18].

Negli anni successivi, con l'avanzare dell'alternativa funzionalista (PAR. 13.3.2), l'ipotesi dell'identità sarebbe stata discussa ancora in termini logico-modali. Alla possibilità di identificare i «ruoli causali» assunti dai diversi stati mentali e fisici (Lewis, 1966), veniva opposto – in stile cartesiano – che la stessa contingenza di questa identificazione basterebbe a escludere la validità dell'ipotesi (Kripke, 1972, pp. 144-55). Armstrong difese la teoria materialistica definendo lo stato mentale come quello «adatto a essere causa di certi effetti» e, per escludere ogni specificità ontologica del mentale, riduceva le qualità della sensazione a proprietà degli oggetti corporei (Armstrong, 1981, p. 20). Egli parlava in proposito di un'ipotesi «metafisica» sul «potere» della materia, riprendendo un termine filosofico tipica-

18. Forse proprio per la scelta di Wittgenstein le C-fibre, i nocicettori individuati da Sherrington all'inizio del XX secolo, sono state a lungo le più popolari unità neurali nelle discussioni filosofiche sulle sensazioni. Cfr. Grahek (2007, pp. 141-66).

mente lockiano. Come si vede, argomenti tipici della filosofia moderna venivano ripensati senza tener conto della ricerca scientifica. Armstrong del resto avanzava la seguente considerazione metodologica:

> Il filosofo non ha la competenza professionale per difendere positivamente la causa del materialismo. In proposito deve basarsi sull'evidenza prodotta dallo scienziato, in particolare dal fisico. Ma può almeno neutralizzare le obiezioni al materialismo proposte dai suoi colleghi filosofi (ivi, p. 31).

Per questa sua separazione teorica dalla ricerca fisiologica l'approccio della scuola australiana è stato definito efficacemente un «materialismo disincarnato» (Wolfe, 2013). D'altra parte, l'approccio basato sull'identità del ruolo causale non è risultato capace di spiegare i contenuti fenomenici che non sembrano possedere una funzione, portando alla sostituzione del fisicalismo riduzionistico con la più modesta tesi della sopravvenienza tra stati mentali e fisici[19].

La questione della verifica empirica della teoria dell'identità venne maggiormente approfondita da Feigl. Per Feigl, come abbiamo visto, si trattava di riproporre la teoria realistica e monistica sull'identità mente-corpo di Schlick[20]. Anche per Feigl la questione aveva un aspetto puramente logico: adottando la distinzione tra senso e denotazione, si potevano identificare i «raw feels», i puri vissuti dell'esperienza, con i processi neurofisiologici corrispondenti: «i dati dell'esperienza sono la realtà denotata da una classe molto ristretta di concetti neurofisiologici» (ivi, p. 83). Ma già questo passaggio, a causa dell'impostazione realistica di Feigl, comportava specifiche considerazioni epistemologiche: secondo Feigl – che riprendeva una concezione avanzata da Russell – la scienza consiste essenzialmente nella conoscenza di relazioni e non stabilisce dunque la qualità dei termini che queste relazioni collegano. Nulla impedisce, dunque, di identificare questi termini, cui si riferiscono le teorie neurofisiologiche, con le qualità che sono date immediatamente nell'esperienza (ivi, p. 84). Un'ipotesi tan-

19. Cfr. Kim (2005, p. 167) e il bilancio critico di Block (2015). Su riduzionismo e sopravvenienza cfr. il saggio di De Caro e Marraffa (CAP. 14).

20. Feigl, 1967, pp. 79-80. In una nota Feigl rimanda a Schlick (1925, §§ 22-35), come alla «prima formulazione genuinamente perspicace, lucida e convincente del punto di vista realistico-monistico che è qui difeso», auspicando una traduzione in inglese della *Allgemeine Erkenntislehre* (che sarebbe comparsa soltanto nel 1974).

to speculativa – di cui Feigl (seguendo Schlick) attribuiva una prima formulazione a Kant – richiedeva però una verifica empirica, sulla cui possibilità egli riteneva opportuno che la filosofia si pronunciasse. Richiamandosi al vecchio problema dell'*Ignorabimus*, Feigl concedeva non vi fosse ancora una teoria scientifica della coscienza (ivi, p. 105). Ma in contrasto con Carnap, che – commentando la teoria di Feigl – insisteva nel sostenere che i criteri per scegliere tra dualismo e monismo fossero puramente linguistici e pragmatici, piuttosto che fattuali e scientifici (Carnap, 1963a, pp. 884-5), Feigl osservava che la corrispondenza puntuale tra attività cerebrale e attività psichica si sarebbe potuta verificare sperimentalmente, confutando almeno il dualismo. A tal fine Feigl – con un esperimento mentale che anticipava quanto sarebbe stato tecnicamente possibile da lì a pochi anni con il *brain imaging* – immaginava un «autocerebroscopio», cioè uno strumento che permettesse di osservare su uno schermo, simultaneamente all'esperienza mentale, «una "immagine" molto ingrandita delle correnti nervose cerebrali» (Feigl, 1967, p. 89)[21]. Una volta confermato puntualmente il parallelismo tra stati mentali e fisici, sarebbe rimasta aperta l'alternativa tra l'identitismo e l'emergentismo, secondo cui sarebbe impossibile ridurre le proprietà mentali alle leggi fisico-chimiche. Feigl, pur propendendo per il primo in base a ragioni di semplicità e coerenza del sapere scientifico, non escludeva la seconda ipotesi[22].

Feigl era molto chiaro nel sottolineare che la sua teoria, senza verifica empirica, sarebbe rimasta un mero «esercizio logico» (ivi, p. 160). Egli riconosceva in particolare che, alla luce delle neuroscienze, l'identità tra processi mentali coscienti e processi materiali non avrebbe riguardato specifiche «fibre» cerebrali, ma piuttosto «aspetti globali» e «configurazionali dei processi neurali» (ivi, pp. 79, 138). Richiamandosi con

21. L'ipotesi di una osservazione esattissima dei processi cerebrali in soggetti coscienti in base a «raffinati strumenti di osservazione» si trovava già in Schlick (1925, § 33, p. 348).

22. L'emergentismo era stato da poco riconsiderato da Meehl e Sellars (1956) nel primo volume dei "Minnesota Studies in the Philosophy of Science" – curato da Feigl e Scriven – a testimonianza delle discussioni che stavano avendo luogo all'Università del Minnesota. Feigl conveniva con Carnap che questa tesi avrebbe costituito una correzione del fisicalismo, che a lungo tra gli eredi del Circolo di Vienna aveva rappresentato un criterio di adesione alla scienza; ma suo obiettivo era proprio legittimare la posizione di alternative empiricamente discriminabili rispetto agli interdetti logico-filosofici del passato.

ciò a un'ipotesi già acquisita dalle neuroscienze del secolo precedente, Feigl rimandava la piena soluzione della questione allo sviluppo della ricerca empirica.

Anche Sellars prese posizione in favore della teoria dell'identità come ipotesi plausibile e verificabile in una scienza futura. Già nel 1956, in *Empiricism and the Philosophy of Mind*, Sellars aveva sostenuto che una teoria scientifica potesse spiegare i dati dei sensi e che, dunque, era filosoficamente sbagliato prendere questi ultimi come punti di partenza di ogni conoscenza. Successivamente Sellars si spinse a ipotizzare che la teoria dell'identità avrebbe potuto essere quella vera, a condizione però di assumere specifici «campi sensoriali» per giustificare il passaggio dalla struttura particellare della materia alla fenomenologia unitaria della coscienza. Ciò implicava il riconoscimento che la teoria neurofisiologica fosse «incompleta» e dovesse essere integrata con «nuovi oggetti» (Sellars, 1962, pp. 36-7).

L'ipotesi originale di Feigl di includere i vissuti nell'ontologia elementare della scienza della natura venne in seguito difesa da Maxwell (1974) e, ancora di recente, è stata valutata positivamente da Searle (2004, pp. 57-8). In questi approcci restava peraltro aperta la questione se e con quali enti tale ripensamento concettuale dovesse condurre in futuro a una nuova scienza della natura: un'ambiziosa ipotesi che venne a lungo tenuta a margine del dibattito filosofico-scientifico dalla diffusione del paradigma funzionalistico-computazionale.

13.3.2. COMPUTAZIONE E COSCIENZA FENOMENICA

La tesi che identificava gli stati mentali con stati funzionali di un organismo, elaborata da Hilary Putnam negli anni Sessanta, si presentava come una versione raffinata di materialismo capace di superare i limiti del behaviorismo e della teoria dell'identità (Putnam, 1975, p. 450). Le tesi critiche di Putnam erano argomentate soprattutto attraverso esperimenti mentali, basati sul confronto tra le facoltà mentali umane e quelle di altre specie (reali o fittizie). I casi ipotetici di uomini "superspartani", educati a non manifestare mai le proprie sensazioni, o abitanti di un mondo X, in cui non fossero ammessi neanche resoconti linguistici sugli stati mentali, risultavano nella impossibilità di definire gli stati mentali in base ai comportamenti (ivi, pp. 332-3). Per contro, sarebbe stato possibile verificare l'esistenza di stati mentali «esaminando i loro cervelli» (ivi, p. 337). D'altra parte, il problema di definire specifici stati cerebra-

li corrispondenti agli stessi stati mentali in diverse specie di organismi, dunque la «realizzabilità multipla» degli stati mentali, portava alla definizione di stato mentale come «stato funzionale dell'organismo». L'ipotesi funzionalista, per quanto «meccanicistica nella sua ispirazione», era compatibile con il dualismo, poiché in teoria un sistema dotato di un'anima avrebbe potuto riprodurre le medesime prestazioni funzionali di uno solamente materiale (ivi, p. 436). Il chiarimento di tali questioni richiedeva un approfondimento empirico del modello funzionalista.

Nonostante gli occasionali riferimenti ai meccanismi cerebrali, la *pars construens* della teoria di Putnam si basava sul modello della macchina di Turing, che portava in questi anni alla costituzione del programma di ricerca delle scienze cognitive noto come "funzionalismo computazionale"[23]. L'impostazione stessa di questo programma di ricerca muoveva dal riconoscimento dell'importanza dei processi "subpersonali" (cioè non coscienti) nell'elaborazione delle informazioni e dal riconoscimento del carattere vago e ingannevole dell'introspezione, in base alla quale il soggetto elaborerebbe resoconti largamente confabulatori delle proprie percezioni e dei propri moventi. Il problema della coscienza rimase pertanto a margine delle ricerche, quando non fu considerato senz'altro insolubile.

Anche la concezione funzionalistica, per realizzarsi in una formulazione teoricamente compiuta, rimandava agli sviluppi empirici delle neuroscienze. Secondo Putnam il funzionalismo forniva uno «schema per ipotesi», che una teoria compiuta (forse «utopistica») avrebbe dovuto realizzare ed eventualmente correggere (ivi, p. 433). La maggiore plausibilità di questo diverso tipo di ipotesi rispetto a quella identitista, nel frattempo, era stabilita sufficientemente dall'argomento della realizzabilità multipla degli stati mentali e dalla fecondità del modello computazionale. Negli anni successivi, Putnam avrebbe parzialmente modificato le proprie tesi: riconoscendo l'importanza degli approcci biologico-evoluzionistici, egli ha sostenuto che il funzionalismo non debba essere necessariamente computazionale e (anche tenendo conto della teoria esternalista del significato) che la definizione delle funzioni cognitive non può prescindere dall'ambiente; ha anche concesso che la

23. Per un'ottima introduzione cfr. Marraffa, Paternoster (2012, pp. 18-25). Sui modelli computazionali come simulazioni di funzioni cognitive cfr. i saggi di Datteri (CAP. 11), De Caro e Marraffa (CAP. 14).

nozione di "capacità intenzionali" – che secondo molti scienziati cognitivi deve essere eliminata dal vocabolario scientifico – possa risultare ineliminabile nelle descrizioni scientifiche, pur sottolineando l'esigenza di una spiegazione naturalistica della emergenza di tali capacità: a proposito dell'intero programma di ricerca, Putnam ha ricordato che «i dettagli li consider[o] qualcosa che in larga parte è demandato al lavoro degli scienziati nei diversi ambiti» (Putnam, 2012, p. 111). Ma a dispetto delle diverse correzioni del suo punto di vista originario, il problema della coscienza fenomenica gli appare pur sempre risolto attraverso l'analisi dell'organizzazione funzionale (ivi, p. 627). Putnam contesta infatti la stessa posizione del problema della coscienza, le cui basi filosofiche sarebbero ancora troppo influenzate dal dualismo cartesiano: una volta isolate delle rappresentazioni mentali interiori, che come le idee cartesiane fungerebbero da medio tra soggetto e mondo, si porrebbe il problema di "naturalizzarle" in termini materialistici. Putnam contrappone a questo «cartesianesimo-con-materialismo» (ivi, p. 281), a suo avviso ormai dominante nella filosofia anglo-americana, una concezione che recupererebbe lo spirito realistico e non dualistico della filosofia di Aristotele (ivi, p. 99; cfr. *infra*, nota 43, PAR. 13.5.2).

L'approccio di Putnam è rimasto esemplare per le scienze cognitive. Uno dei massimi risultati in questo campo è stata la teoria dello «spazio globale di lavoro» (*global workspace theory*) di Baars che, pur non riferendosi direttamente a Putnam, implica una posizione funzionalista:

> La coscienza è associata a uno spazio di lavoro globale o a *un suo equivalente funzionale*. Resta da vedere come questo sistema sia realizzato in dettaglio [...]. I contenuti coscienti sono internamente coerenti e globalmente distribuiti. Se affermiamo le cose a questo livello di astrazione, possiamo evitare di impegnarci a favore di una particolare realizzazione di "hardware" (Baars, 1988, p. 104).

A sostegno del suo modello Baars menzionava le ricerche di neuroscienziati come Gazzaniga sul cervello in quanto sistema di elaborazione parallela organizzato per moduli (ivi, p. 107). L'evidenza sperimentale di connessioni distribuite tra i circuiti in cui si trasmette l'informazione avrebbe dunque fornito sufficiente conferma scientifica del modello. La qualità soggettiva della coscienza rimaneva al di fuori di questo approccio (anche stavolta con un riferimento alla filosofia di Aristotele: il modello dello «spazio globale di lavoro» è presentato come una versione contemporanea della teoria aristotelica del senso comune).

I sostenitori del funzionalismo, in generale, non hanno considerato il concetto di coscienza fenomenica adeguato a una trattazione scientifica e hanno spesso avanzato il sospetto che esso produca un falso problema. Un'articolata giustificazione di questa scelta si trova negli scritti di Dennett. Nella sua tesi di dottorato (che aveva scritto sotto la guida di Ryle) Dennett muoveva dall'esigenza di interrompere l'«infruttuosa oscillazione del pendolo» tra le teorie metafisiche come dualismo e teoria dell'identità (Dennett, 1985, p. 3), ma al tempo stesso riteneva inadeguate le principali alternative proposte a tal fine, il comportamentismo e la fenomenologia. Il primo non sarebbe riuscito a spiegare l'apprendimento e il carattere finalizzato a scopi del comportamento; la seconda – intesa non nel senso specifico della tradizione filosofica iniziata da Husserl, ma in genere come la concezione che considera epistemicamente primari e irriducibili i contenuti fenomenici – avrebbe fatto uso di spiegazioni psicologiche estranee all'apparato concettuale della scienza e indipendenti, in particolare, da qualsiasi conoscenza biologica (ivi, pp. 31-9). Dennett contestava a entrambe le prospettive di aver trascurato le condizioni biologiche del comportamento animale e riproponeva un'indagine fondata su queste domande: «Qual è il rapporto tra la vita mentale umana e gli eventi nel cervello? Come si possono mettere in corrispondenza [*map*] le nostre osservazioni di senso comune sul pensare, il credere, il vedere, il provare dolore con le scoperte della cibernetica e della neurobiologia?» (ivi, p. 13).

La prima mossa di Dennett è consistita nell'elaborazione di una teoria del contenuto mentale che escludesse la presunta evidenza del significato nell'introspezione, ricavandolo invece dalla «funzione» svolta dai segni nel contesto delle interazioni con l'ambiente e con gli altri individui (Dennett, 1987, pp. 287-321; cfr. Millikan, 1984). Analogamente, per affrontare il problema della coscienza fenomenica, Dennett ha allestito un metodo "eterofenomenologico": si tratta di considerare i resoconti sugli stati interiori come delle credenze, la cui verità non è messa in discussione dai soggetti che le posseggono, ma la cui giustificazione e spiegazione devono essere cercate dal filosofo e dallo scienziato con lo stesso distacco di un antropologo che esamini le credenze di un popolo primitivo, considerandole «finzioni» (Dennett, 1991a, pp. 66-98). La critica degli oggetti "eterofenomenologici", come i *qualia* e il sé, prende le mosse dall'argomento secondo cui non è possibile stabilire empiricamente in quale momento, nell'intervallo spazio-temporale che separa gli stimoli dal comportamento di un organismo, la rappresentazione coscien-

te sia esperita dal soggetto. Dennett chiama il luogo ipotetico di questa rappresentazione il "teatro cartesiano" e la teoria fallace che pretende di localizzarlo nel cervello il "materialismo cartesiano". In realtà, secondo Dennett – che interpreta così i risultati delle scienze cognitive (Baars) e delle neuroscienze contemporanee – il cervello è sede di una organizzazione parallela e modulare delle informazioni, che non ha nessun luogo di coordinamento. Il fondamento della vita cosciente è dunque da considerarsi una «macchina virtuale implementata nell'architettura parallela del cervello» (ivi, p. 91), che elabora in parallelo resoconti oggettivi e protocolli motori, producendo «versioni molteplici» in competizione tra di loro (ivi, pp. 102-38). Il fatto che di volta in volta una sola di queste versioni – attraverso un meccanismo selettivo – risulti nell'elaborazione di un resoconto linguistico o di un comportamento determinati non significa che un «significatore centrale» abbia svolto la funzione di coordinamento. Dennett corrobora questa ipotesi con diversi esperimenti mentali, da cui si ricava la difficoltà di distinguere il dato percettivo dalla sua elaborazione posteriore. In conclusione, critica dell'intuizione e ipotesi meccanicistica formano una teoria secondo cui la qualità dell'esperienza cosciente è risolta in un complesso mutevole di credenze, e il sé è esso stesso un costrutto sociale, un «centro di gravità narrativa» (ivi, pp. 412-30).

La teoria di Dennett costituisce un tentativo di sviluppare le idee di Wittgenstein (e Ryle) sul contenuto ingannevole dell'introspezione attraverso i risultati delle scienze cognitive, in particolare del funzionalismo e del connessionismo (il campo di ricerche sulle reti neurali distribuite e parallele). Occorre però sottolineare che, come abbiamo visto, Dennett intravedeva anche nella biologia evoluzionistica e nelle neuroscienze campi disciplinari che avrebbero giocato un ruolo per una teoria scientifica della coscienza (Dennett, 1969, p. 40). Dennett proponeva in effetti una teoria evoluzionistica sullo sviluppo e la selezione delle funzioni cerebrali, rimandando alla scienza per una conferma specifica di questo tipo di ipotesi (ivi, pp. 48, 71). In più occasioni, negli anni successivi, egli avrebbe constatato come lo sviluppo delle neuroscienze confortasse le sue ipotesi (cfr. PAR. 13.5.1).

Rispetto a questi sviluppi scientifici Dennett ha sostenuto che il suo lavoro filosofico riguarda solo «la possibilità» di una teoria, e che il suo «schizzo di teoria o modello di come il cervello *potrebbe* fare qualcosa» costituisce un «programma di ricerca» che eventualmente potrà condurre a una «teoria empiricamente confermata» (Dennett, 1991a,

p. 41). Quel che importa è prevenire l'illusione che la coscienza sia un «mistero» – come porterebbe a credere il dualismo a partire da Descartes – mostrando la via per una sua spiegazione naturalistica fondata sul repertorio già disponibile della conoscenza scientifica (ivi, p. 37). Così Dennett, per evitare di entrare nei dettagli di «troppe cose misteriose» (ivi, p. 193), procede a un livello «più generale e astratto» di quello di una teoria empirica, rimandando al lavoro dei neuroscienziati.

Un simile connubio di critica del senso comune e ipotesi materialistica si trova nei lavori di Paul e Patricia Churchland[24]. Secondo Patricia Churchland la spiegazione della coscienza in termini di «meccanismi» latenti è l'unica opzione plausibile: «poiché non può essere la magia, devono essere meccanismi» (Churchland, 1986, pp. 460-1). Lo sviluppo delle neuroscienze più recenti mostra che una scienza della «mente-corpo» costituisce un risultato rivoluzionario prossimo venturo, la cui importanza sarà «almeno uguale alle rivoluzioni copernicana e darwiniana» (ivi, p. 481). In generale filosofi come Dennett e i Churchland hanno riposto grande fiducia nelle potenzialità degli approcci scientifici sviluppati negli ultimi decenni: perciò essi non hanno potuto accogliere le tesi dei tanti filosofi che, sottolineando l'esigenza di una teoria scientifica alternativa, hanno fatto a loro avviso il gioco dei difensori del «mistero».

13.3.3. NATURALISMO E INCLUSIONE DELLA COSCIENZA FENOMENICA

Tra i primi e più conseguenti sostenitori dell'esigenza di un approccio scientifico alla coscienza va annoverato, tra i filosofi contemporanei, John Searle. Anche Searle condivide la liquidazione del dualismo metafisico, in quanto prospettiva empiricamente infondata, e difende pertanto una concezione naturalistica della coscienza, ma al contrario di filosofi come Putnam e Dennett, egli sostiene con forza che intenzionalità e coscienza siano proprietà naturali, che nessuna spiegazione materialistica, funzionalistica o riduzionistica è stata finora in grado di spiegare adeguatamente.

La critica di Searle al programma delle scienze cognitive è stata avviata dal celebre esperimento mentale della "stanza cinese" (Searle,

24. Per maggiori dettagli su questi autori cfr. il saggio di De Caro e Marraffa (PAR. 14.3).

1980). In questo esperimento si immagina un uomo chiuso in una stanza, che riceve diverse strisce di testo scritte in lingua cinese (senza conoscerla), e insieme a queste un protocollo di regole, scritte in una lingua che conosce, con istruzioni per comporre dei testi in cinese e inviarli al di fuori della stanza. L'argomento sostiene che l'uomo, seguendo le regole, potrebbe produrre messaggi in cinese che apparirebbero sensati a beneficio di un osservatore esterno, senza per questo comprenderne affatto il significato. La conclusione è che l'applicazione di un programma di regole sintattiche non può sostituire la comprensione del significato. Questo celebre argomento, rivolto criticamente contro il modello computazionale della coscienza, ha costituito la premessa di un'indagine sulle basi naturali dell'intenzionalità, che è indissolubile per Searle dalla questione della coscienza, poiché tutti i vissuti intenzionali sono a suo avviso intrinsecamente coscienti (o potenzialmente tali). Così, Searle difende la realtà di «stati qualitativi di coscienza, soggettivi e interiori» e «stati mentali intrinsecamente intenzionali come credenze e desideri», che devono essere descritti in prima persona e non potranno mai essere spiegati da un resoconto in terza persona che non tenga presente questa dimensione soggettiva (Searle, 1992, pp. XI-XIII). Per questa ragione egli ha condotto una lunga polemica contro gli approcci comportamentistici, fisicalistici e funzionalistici, inclusa la strategia "eterofenomenologica" di Dennett[25].

In generale, la tesi di Searle è che la conoscenza scientifica debba rendere conto di coscienza e intenzionalità in quanto proprietà biologiche (Searle stesso ha definito la sua posizione un "naturalismo biologico"):

> Nella mia prospettiva i fenomeni mentali sono biologicamente fondati; essi sono sia causati dalle operazioni del cervello, sia realizzati nella struttura del cervello. In questa prospettiva sia la coscienza, sia l'intenzionalità sono parti della biologia umana quanto la digestione e la circolazione del sangue. È un dato di fatto oggettivo che il mondo contenga certi sistemi, i cervelli, con stati mentali soggettivi, ed è un dato di fatto fisico relativo a questi sistemi che essi abbiano aspetti mentali (Searle, 1983, p. IX; cfr. Id., 2004, pp. 111-5).

Secondo Searle gli stati coscienti soggettivi devono dipendere causalmente da specifici «poteri» (*powers*) del cervello. Riconoscendo la po-

25. Per la critica a Dennett, con uno scambio tra i due, cfr. Searle (1997, pp. 97-131).

sizione eterodossa di questo suo naturalismo biologico, egli ha svolto fin dai suoi primi interventi una «diagnosi» del pregiudizio contro la realtà del «mentale», che domina nella filosofia analitica dai tempi del primo fisicalismo (con l'eccezione di una versione del monismo, come abbiamo visto) fino al funzionalismo, divenuto dominante negli anni Settanta. Questo pregiudizio dipenderebbe da «tre principali ragioni» (Searle, 1981, pp. 421-3): 1. il verificazionismo, che ha seminato un persistente sospetto nei confronti della prospettiva in prima persona; 2. l'adesione al paradigma delle scienze cognitive, che ha suscitato la speranza di ridurre i problemi filosofici a «questioni tecniche» su processi computazionali di livello subpersonale, ignorando la specificità degli stati mentali coscienti tipici dell'uomo; 3. la paura del cartesianesimo, per cui l'ammissione delle entità mentali condurrebbe necessariamente ai peggiori eccessi metafisici del dualismo cartesiano. Secondo Searle, al contrario, ammettere diversi «livelli di descrizione» – come nel caso di proprietà emergenti quali la solidità di un corpo e l'esplosione di un combustibile – non comporta alcun dualismo ontologico. È possibile, per esempio, riconoscere che esperienze visive e intenzioni di agire causino il comportamento al livello macroscopico, anche se al livello microscopico termini come "intenzione di agire" sono «inappropriati» (Searle, 1983, pp. 267-9). Con queste tesi Searle si schierava a favore di un emergentismo che avrebbe presto raccolto sostenitori tra i neuroscienziati impegnati nel problema della coscienza.

Ma il problema del naturalismo di Searle è stato proprio il fatto di non trovare un soddisfacente punto di riferimento scientifico per i suoi concetti di causazione e realizzazione degli stati mentali. Se «la natura logica di questi *tipi* di relazioni tra mente e cervello» è perfettamente comprensibile, e si può affermare che stati mentali e fisici esistono «nella stessa sostanza», tuttavia «i problemi empirici e concettuali» nella descrizione di queste relazioni sono «incredibilmente complessi e il progresso, a dispetto di molti discorsi ottimistici, è stato lento come un'agonia» (ivi, p. 267). Per esempio, scriveva Searle, «sappiamo come i meccanismi neurali stimolano i movimenti muscolari», ma «sappiamo poco su come l'azione intenzionale ha origine nel cervello» (ivi, pp. 269-70). Questo atteggiamento di sfiducia è stato mantenuto negli anni successivi, anche quando le neuroscienze sono divenute il cardine delle scienze cognitive e sono comparse nuove teorie neuroscientifiche specificamente mirate alla spiegazione della coscienza (cfr. PAR. 13.5.1).

L'aporia di Searle è esito esemplare di un dilemma affrontato da tutti i sostenitori di un naturalismo che si proponga di includere la coscienza fenomenica tra i fenomeni naturali descritti dalle scienze. Un modo per farlo è affermare che l'autonomia del mentale appartiene a una stagione circoscritta delle scienze cognitive e che è opportuno tornare a un «naturalismo inclusivo» (Hatfield, 2009), già sostenuto da eminenti filosofi e scienziati del XIX e XX secolo, nel quale l'intuizione psicologica risulta un momento irriducibile delle spiegazioni scientifiche di stati coscienti qualitativi, senza che ciò comporti una modifica delle teorie scientifiche attuali. In questa prospettiva, per esempio, le qualità cromatiche (e in genere i *qualia*) possono essere considerate come stati soggettivi reali, che come segni designano proprietà disposizionali degli oggetti fisici. Questo tipo di approccio postula nuove proprietà fondamentali, senza richiederne una spiegazione riduttiva, e talvolta si coniuga a una metafisica delle qualità intrinseche della materia. Un approccio diverso – benché talvolta intrecciato al precedente – consiste nel sottolineare i limiti di tutti gli approcci scientifici disponibili e guardare al futuro per una teoria scientifica nuova nelle sue leggi fondamentali.

Una tale posizione è sostenuta da David Chalmers con il suo «naturalismo dualistico». Secondo Chalmers nessuna spiegazione funzionale di processi e capacità mentali, quali quelle fornite dalle scienze cognitive, risolve il «problema difficile della coscienza», cioè «perché tutta questa elaborazione è accompagnata dall'esperienza di una vita interiore?» (Chalmers, 1996, p. XII). Per introdurre la sua tesi, Chalmers ha presentato l'esperimento mentale dello zombie, secondo cui sono pensabili esseri fisicamente indistinguibili dagli umani, dotati delle medesime capacità e lo stesso comportamento, ma incapaci di "sentire" coscientemente come noi (ivi, pp. 94-9). Con questo esperimento mentale (di cui è stata vivacemente discussa la pertinenza)[26] e altri analoghi[27],

26. Alcuni critici sostengono, per quanto riguarda lo zombie di Chalmers, che si tratti di un essere solo apparentemente concepibile: posta la sua indistinguibilità fisica e comportamentale da un uomo, la coscienza gli apparterrebbe necessariamente (Dennett, 1991a, capp. 10-12). Dennett paragona la resistenza di Chalmers ad ammettere la sufficienza del funzionalismo alle tesi dei vecchi vitalisti sulla presunta insufficienza delle spiegazioni biologiche della vita (ivi, pp. 314-5).

27. Argomenti analoghi sono lo "spettro invertito" – secondo cui in due mondi fisicamente identici si potrebbero percepire diversamente i colori – e i "*qualia* assen-

Chalmers e altri filosofi vogliono mettere in evidenza una eterogeneità tra gli stati coscienti e le proprietà fisiche biologiche e computazionali che caratterizzano gli approcci scientifici contemporanei, attualizzando così in un nuovo contesto la vecchia questione dei limiti della scienza[28]. Secondo Chalmers, in particolare, nessuna teoria disponibile di tipo computazionale (Baars, Dennett, P. S. Churchland), neurobiologico (Crick, Koch, Edelman), fisico (speculazioni sulla fisica quantistica di Penrose e altri) o evoluzionistico ha attualmente risolto il «problema difficile» (ivi, pp. 111-22). Per soddisfare entrambi i principi – "prendere la coscienza sul serio" e "prendere la scienza sul serio" – Chalmers ritiene che occorrerebbe una riforma della scienza della natura, che faccia spazio a un'ontologia dualistica introducendo principi e leggi psicofisiche relative a come gli stati coscienti sono collegati a quelli materiali (ivi, pp. 213-46). In questa prospettiva Chalmers non esita a rivalutare ipotesi metafisiche del passato, che potrebbero fornire uno sfondo adeguato alla nuova teoria scientifica della coscienza, tra cui una ipotesi "panpsichista" secondo cui la stessa "informazione" avrebbe due aspetti, uno fisico e uno fenomenico – con la conseguenza che anche semplici macchine, come un termostato, avrebbero un certo grado di coscienza fenomenica (ivi, pp. 276-310).

I lavori di Chalmers sono casi esemplari di come, nella discussione degli ultimi venti anni, la promessa di una teoria neuroscientifica della coscienza si sia accompagnata alla riscoperta e riconsiderazione di ipotesi metafisiche del passato (Chalmers, 2010, pp. 103-39). Simili congetture mostrano il grado di speculazione a cui è giunta la ricerca di una nuova scienza ancora mancante, che ha contribuito a radicalizzare ulteriormente il dissenso tra i filosofi naturalisti (cfr. PAR. 13.5.1). Ma un tratto comune a tutti i filosofi considerati in questo capitolo resta

ti", in cui si immagina un duplicato fisico della Terra privo di percezioni qualitative. Esistono diverse versioni di entrambi gli argomenti: una sintesi di tutte le discussioni è in Tye (2013). Per alcuni bilanci critici cfr. Chalmers (1996, cap. 7) e Dennett (2005), che considera l'intera discussione "reazionaria" e difende il funzionalismo.

28. Anche il "problema difficile", come la tesi monistica, era infatti ben noto nella filosofia e nella scienza del XIX secolo. Un esperimento mentale sui "*qualia* assenti" è già accennato in Du Bois-Reymond (1872, p. 124), che parla di un "dormiente senza sogni"; Lange (1875, II, p. 156) ne trae spunto per allestire un esperimento su due mondi fisicamente indistinguibili, uno dei quali senza vissuti, che presenta sostanzialmente la stessa idea degli esperimenti mentali oggi in discussione.

la tesi di una continuità tra indagine filosofica e teoria scientifica della coscienza. Lo stesso Chalmers passa in rassegna le diverse ipotesi metafisiche possibili, distinguendo tra esse quelle più plausibili, ma rimanda in ultima analisi al lavoro degli scienziati per una soluzione della questione. Si tratta di un criterio formulato fin dai primi anni della filosofia scientifica, in alternativa alle tesi di discontinuità avanzate da filosofi come Husserl e Wittgenstein (cfr. PAR. 13.5.2), e che è stato adottato innumerevoli volte. Come già Schlick, anche Reichenbach sosteneva la possibilità di definire gli stati mentali coscienti in base a un «processo interno» materiale, ritenendo però necessario mantenere il linguaggio psicologico per via dei limiti della fisiologia dell'epoca (Reichenbach, 1938, p. 245). Ma se Feigl parlava di una «cambiale» che la scienza avrebbe pagato in futuro, rimandando a quando «la neurofisiologia del futuro (3000 d.C.?)» avrebbe fornito una derivazione deduttiva dei comportamenti dagli stati cerebrali (Feigl, 1967, pp. 75, 102), e molti altri protagonisti di queste ricerche hanno ripetuto che, un giorno, una teoria scientifica avrebbe risolto il problema di identificare stati mentali e stati neurali, oggi molti filosofi ritengono che questa soluzione sia imminente e provenga dalle neuroscienze.

13.4
Le teorie neuroscientifiche: evidenze empiriche e ipotesi metafisiche

13.4.1. COSCIENZA E CERVELLO: VECCHI PROBLEMI, NUOVE EVIDENZE EMPIRICHE

Come abbiamo visto, il problema della localizzazione degli stati coscienti nei meccanismi del cervello è stato posto fin dal principio delle indagini filosofico-scientifiche moderne, risultando dalla crisi dell'ilemorfismo aristotelico, dall'emergere del paradigma meccanicistico e dalla diffusione del concetto filosofico di coscienza. Con la crescita delle indagini empiriche, nell'anatomia e fisiologia del XIX secolo, si è consolidata l'ipotesi che l'interazione di diverse aree cerebrali fosse responsabile dell'emergere della coscienza. Mancavano ancora i dettagli di questo meccanismo (su cui si concentravano le ricerche sperimentali) e molti scienziati e filosofi credevano che anche la sua conoscenza non avrebbe risolto completamente il problema della coscienza. Ma in generale moltissimi aspetti del dibattito più recente erano già presen-

ti (cfr. PAR. 13.2)[29]. All'inizio del XX secolo, tuttavia, la diffusione del comportamentismo in psicologia e in filosofia determinò un rallentamento di questo tipo di ricerca, che è stata poi ripresa con crescente intensità soltanto negli ultimi decenni del secolo.

Ricordare questa lunga storia della questione è utile anche per distinguere quali siano le *nuove* evidenze empiriche che hanno suscitato l'intenso dibattito attuale. Lo studio comparato dell'anatomia e del comportamento ha determinato un progresso costante e graduale nell'individuazione dei correlati neurali di particolari stati coscienti e, a partire dalla fine degli anni Sessanta, alcune specifiche indagini sperimentali hanno incoraggiato la formulazione di teorie generali. Tra queste indagini bisogna ricordare prima di tutto gli studi sul "cervello diviso" di Sperry e Gazzaniga che, rilevando l'elaborazione dissociata dell'informazione in persone con resezione del corpo calloso, diedero inizio a intense riflessioni sulle basi neurali dell'unità della coscienza[30]. L'ipotesi avanzata da Gazzaniga di un "modulo interprete", responsabile della integrazione "narrativa" delle informazioni distribuite, richiedeva a sua volta una puntuale indagine sui meccanismi di questa integrazione[31].

29. Un utile resoconto delle discussioni dell'epoca, a cui abbiamo già rimandato, è Lange (1875, II), dove sono discusse diverse tesi sostenute da scienziati come Fechner, Helmholtz, Du Bois-Reymond, Hitzig, Wundt e altri, poi ricomparse nel dibattito contemporaneo, tra cui: 1. il carattere globale dei correlati della coscienza (ivi, pp. 343-4); 2. la tesi secondo cui la base fisica della coscienza è paragonabile, più che a un singolo omuncolo cartesiano, a «un Parlamento di piccoli uomini» (ivi, pp. 344, 418); 3. l'ipotesi che la coscienza potrebbe dipendere da «un rapporto tra tutte le intensità delle eccitazioni date nel dominio della sensazione» (ivi, p. 439n.); 4. la tesi della realizzabilità cerebrale multipla di medesime funzioni cognitive (ivi, p. 365); 5. la critica alla trasparenza dell'introspezione (ivi, p. 386); 6. la tesi della non riducibilità del contenuto qualitativo delle sensazioni e dell'agire libero al dominio fisiologico (ivi, pp. 396, 404); 7. il collegamento di questo problema con la questione dei «limiti della conoscenza della natura» (ivi, pp. 409-10), con particolare riferimento all'unità dell'esperienza cosciente (ivi, pp. 417-9).

30. Come è noto, i pazienti studiati da Gazzaniga potevano mostrare di riconoscere un oggetto percepito con l'emisfero destro (nel lato sinistro del campo visivo), mentre riferivano di non aver percepito nulla. Questo resoconto dipendeva dall'emisfero sinistro, sede dell'elaborazione linguistica, che non aveva accesso visivo all'oggetto. Esperimenti come questo sollevarono la questione se in questi pazienti vi fossero due distinte coscienze.

31. Una prima sintesi teorica si trovava in Gazzaniga, Le Doux (1978). Tra i primi tentativi di interpretare filosoficamente questi fenomeni cfr. Nagel (1971), che pone-

La stessa questione si è posta a partire da altre osservazioni cliniche sulla disarticolazione tra padronanza di informazioni e coscienza, come gli studi sulla visione cieca (*blindsight*) di Humphrey (1970; cfr. nota 3). Infine, a partire dagli anni Settanta, gli studi di Libet e altri hanno permesso di misurare il tempo di elaborazione degli stati coscienti e hanno posto nuovamente in questione il rapporto tra processo cosciente e suoi correlati neurali, mostrando come un "potenziale di preparazione" osservabile nel cervello preceda di una frazione di secondo la deliberazione cosciente degli stessi movimenti[32]. La comparsa di diversi modelli teorici sul nesso tra processi neurali e coscienza, negli anni successivi, è stata certamente influenzata dall'esigenza di collegare e interpretare risultati sperimentali come questi.

Un'altra condizione di questa proliferazione teorica è stato senza dubbio lo sviluppo di tecniche di osservazione non invasiva, vecchie, come l'elettroencefalografia (EEG), e nuove, come la tomografia a emissione di positroni (PET) e la fMRI (risonanza magnetica funzionale). Queste tecniche, misurando cambiamenti nell'attività elettrica o nel flusso ematico, permettono di osservare l'attivazione di aree cerebrali durante l'attività cosciente. Accanto a queste sono risultate preziose anche tecniche di stimolazione di aree ridotte e di singoli neuroni, come la stimolazione magnetica transcranica (TMS). Grazie al perfezionamento di simili tecniche, negli ultimi decenni è stata raccolta una grandissima quantità di dati, che richiedono un lavoro di raccordo e interpretazione. Esiste tuttavia un consenso abbastanza diffuso su alcuni risultati generali raggiunti a proposito della coscienza. Sul piano anatomico-funzionale la coscienza:

– è un processo distribuito e globale, che non è limitato all'attivazione di una singola area cerebrale, ma non coinvolge nemmeno l'intero cervello;
– coinvolge sempre (ma non solo) il sistema talamo-corticale;
– coinvolge interazioni rientranti tra "assemblee" di neuroni corticali e sottocorticali;

va già l'accento sulla «integrazione» di contributi disomogenei quale caratteristica fondamentale della coscienza.

32. I risultati di Libet riguardavano l'esecuzione di semplici movimenti e la loro pertinenza per la questione del libero arbitrio è stata da subito vivacemente discussa. Cfr. Libet (2004) per il bilancio dell'autore e De Caro (2013) per una messa a punto critica.

– è una proprietà che viene elaborata in un tempo determinato (di 100-200 msec).

Sul piano fenomenologico, d'altra parte, neuroscienziati cognitivi, psichiatri e psicologi hanno messo in rilievo che la coscienza:
– è unitaria e inseparabile nei suoi elementi (come confermano i casi patologici di lacune e sdoppiamenti, in cui non viene esperita una distinzione in parti del contenuto coscienziale);
– si può sovrapporre a, ma non coincide sempre con, altri stati: la veglia, l'attenzione, specifiche modalità di percezione sensoriale, la memoria. Rispettivamente, infatti, si danno stati coscienti: nel sonno (fase REM), in mancanza di attenzione verso un determinato oggetto (che successivamente può venire portato "in primo piano"), in mancanza di varie specifiche percezioni sensoriali (come vista e propriocezione), in casi di amnesia.

Nonostante queste assunzioni, c'è grande diversità di opinioni su come risolvere teoricamente il problema del rapporto tra coscienza e cervello.

13.4.2. I MODELLI TEORICI: TRA DATI E IPOTESI METAFISICHE

Il riproporsi del problema della coscienza nella comunità neuroscientifica – come mostreremo attraverso alcuni esempi – è stato certamente stimolato dalle discussioni che hanno avuto luogo nelle scienze cognitive e nella filosofia della mente tra gli anni Cinquanta e Settanta. Ma una caratteristica singolare di questa nuova attività di ricerca è che le proposte teoriche più compiute hanno determinato la rimessa in gioco di questioni metafisiche, che filosofi e scienziati cognitivi – anche in virtù della loro impostazione filosofica empiristica e naturalistica – avevano ritenuto a lungo mal poste o insolubili.

Questa affermazione va subito precisata, puntualizzando che sussistono nella comunità scientifica orientamenti molto diversi e che questo fenomeno appartiene solo a un ristretto gruppo di scienziati teoricamente impegnati. La maggior parte dei ricercatori punta più modestamente sull'accumulo sistematico di nuovi dati e sulla possibilità di costruire simulazioni dell'attività cerebrale, piuttosto che a tentare nuove sintesi teoriche. A tal fine sono in corso progetti su grande scala come la simulazione dell'attività computazionale del cervello umano (*Human Brain Project*) e la mappatura dettagliata del cervello di un topo (il *Brain Atlas* dello Allen Institute for Brain

Science)[33]. Gli scienziati più impegnati sulle basi della coscienza sottolineano che la raccolta di dati non può risolvere il problema senza una teoria, ma a loro volta si dividono tra chi propende per l'individuazione empirica – attraverso osservazioni selettive – di specifici correlati neurali della coscienza e chi ritiene questo approccio insufficiente, mentre considera necessaria una elaborazione concettuale più generale che spieghi come mai determinate specie di interazioni cerebrali diano luogo alla coscienza (per un bilancio cfr. Koch, Tononi, 2008).

Tra i sostenitori del primo approccio si può annoverare Francis Crick, che con la collaborazione di Christoph Koch ha proposto in diversi lavori pionieristici una serie di considerazioni empiriche sul possibile concorso delle oscillazioni neurali a 40hz per l'emergenza della coscienza, nonché (a partire dagli anni Novanta) sull'attivazione di specifici sottoinsiemi di neuroni, proponendo dunque diversi candidati per la funzione dei NCC. Oltre a ipotizzare specifiche caratteristiche dei correlati neurali in genere, Crick e Koch si sono concentrati sui circuiti di una particolare forma di coscienza, la visione, tentando poi di generalizzare i risultati delle loro ipotesi. Il punto di vista fermamente materialistico e riduzionistico di Crick è particolarmente evidente nell'ormai celebre affermazione programmatica: «"Tu", le tue gioie e le tue tristezze, le tue memorie e le tue ambizioni, il tuo senso dell'identità personale e del libero arbitrio, di fatto non sei altro che il comportamento di un'ampia assemblea di cellule nervose» (Crick, 1994, p. 3). Si tratta di un orientamento radicato nella tradizione materialistica, di cui le controversie più recenti hanno messo in luce il carattere metafisico e non meramente empirico.

Un caso analogo sono i lavori di Jean-Pierre Changeux, in cui questa tradizione filosofica diviene esplicita. Insieme al suo allievo Dehaene e altri ricercatori, Changeux ha elaborato, a partire dalla fine degli anni Ottanta, un modello neurale della coscienza come spazio di lavoro globale. Questi lavori hanno sviluppato esplicitamente il modello di Baars (PAR. 13.3.2), fornendo evidenza neurologica alle ipotesi sull'accesso globale e sui meccanismi di selezione delle informazioni. Tuttavia, prima ancora di incontrare questo modello, Changeux aveva già esposto

33. <http://www.humanbrainproject.eu/discover/the-project/overview> e <http://www.alleninstitute.org/>. Alle interpretazioni dei dati della mappatura è dedicato il programma di ricerca *MindScope*, diretto da Christof Koch.

la sua filosofia materialistica ne *L'homme neuronal* (1983). Secondo Changeux – che è stato allievo di Jacques Monod – i tempi sarebbero maturi per una spiegazione deterministica del comportamento umano in termini biologico-molecolari (Changeux, 1983, pp. 122-4). Questa spiegazione comporterebbe una «teoria biologica degli oggetti mentali», che sono «identificati con lo stato fisico prodotto dall'attivazione correlata e transitoria (sia elettrica sia chimica) di una vasta popolazione o "assemblea" di neuroni distribuiti al livello di differenti aree corticali» (ivi, p. 174). La formazione di concetti stabili nella memoria, a sua volta, dipenderebbe da un meccanismo di epigenesi ispirato alla teoria darwiniana e fondato sulla selezione ambientale di popolazioni di neuroni, la cui attivazione risulta efficace dal punto di vista adattivo. La selezione di «tracce» corrispondenti a comportamenti efficaci avrebbe condotto spontaneamente alla formazione di una «grammatica degli oggetti mentali» (ivi, p. 175). Questo approccio, in generale, avrebbe permesso di realizzare finalmente il programma materialistico della biologia moderna, collegandolo con le ambizioni delle scienze cognitive. Changeux concludeva trionfante: «Darwin permette di riconciliare Fodor e Epicuro» (ivi, p. 176).

È importante sottolineare l'aspetto metafisico – e pertanto controverso – di questo programma di ricerca. Changeux attingeva a risultati condivisi delle neuroscienze dell'epoca – dalla teoria del rinforzo delle connessioni neurali di Hebb all'uso del paradigma selezionista per lo studio delle popolazioni di neuroni di Edelman (ivi, p. 173) – ma li sviluppava nel senso di un ambizioso riduzionismo biochimico, giustificato da esempi ancora rudimentali, che voleva essere la realizzazione di programmi filosofici moderni come quello di La Mettrie (ivi, p. 146): rilevando elettronicamente gli impulsi neurali corrispondenti ai comportamenti si sarebbero potuti ridurre questi ultimi a dei «grafici», rendendo possibile una matematica del comportamento (ivi, pp. 125-9) e la rappresentazione «ideografica» del pensiero attraverso le neuroimmagini (ivi, p. 209). Questa interpretazione dei dati non sarebbe stata poi approvata da molti neuroscienziati come Edelman, che – come vedremo tra poco – pur condividendo le medesime evidenze sperimentali hanno sostenuto, in base alla considerazione dell'esperienza cosciente, l'esigenza di riconoscere il carattere emergente e irriducibile dei processi mentali rispetto a quelli fisici e fisiologici (cfr. da ultimo Gazzaniga, 2011).

Il problema di spiegare la formazione delle «immagini mentali» coscienti ha condotto a un esito diverso, ma altrettanto metafisico, nelle

ricerche di Antonio Damasio. Damasio si è occupato primariamente del problema di come emerga il sentimento del sé, che egli definisce come «il sentimento di ciò che accade quando vieni modificato dagli atti di apprendere qualcosa» (Damasio, 1999, p. 10) e che dunque coinvolge sempre una rappresentazione del corpo e della sua interazione con l'ambiente. Egli sostiene che la condizione della «coscienza di base», cioè del sentimento di essere qui e ora, consiste in «schemi neurali che rappresentano [*map*] entrambi gli agenti [l'organismo e l'oggetto], e il modo in cui interagiscono» (ivi, pp. 133, 169). In un recente lavoro, Damasio affronta il problema di come queste mappe, per quanto materiali, possano dare luogo a un contenuto qualitativo, le immagini mentali, e lo risolve con una mossa schiettamente metafisica: identifica mappe neurali e immagini mentali (Damasio, 2010, p. 65). Tra le diverse (ed eterogenee) fonti di ispirazione filosofica, che Damasio adduce occasionalmente per presentare questa tesi, la principale è senz'altro la metafisica di Spinoza. In un libro specificamente dedicato al parallelismo tra idee spinoziane e ricerca neuroscientifica contemporanea, Damasio commenta con trasporto la definizione spinoziana della mente come idea del corpo (Damasio, 2004, pp. 209-17), che egli interpreta liberamente come una sorta di teoria materialistica del «doppio aspetto». Traendo rigorosamente le conclusioni di questa mossa, Damasio non esita a ipotizzare che finanche i singoli neuroni potrebbero avere un aspetto «protofenomenico»: l'emergenza dei *qualia* dall'interazione del corpo con il sistema nervoso suggerisce una «fusione funzionale di stati corporei e percettivi». Si può parlare dunque di «protocognizione» e «protosentimenti» a livello cellulare (Damasio, 2010, pp. 252, 258). Quella di interpretare la filosofia spinoziana in termini materialistici, ridimensionando la portata metafisica del concetto spinoziano di sostanza, era una mossa tipica del materialismo moderno. E già materialisti del XVIII secolo, come abbiamo accennato (PAR. 13.2), discutevano se considerare la coscienza come una «qualità della materia» o un «prodotto dell'organizzazione» della materia. In questa prospettiva si può dire che, mentre Changeux rivaluta quest'ultima ipotesi, Damasio si riavvicina alla prima[34].

34. Maupertuis aveva attribuito alle molecole di materia «un certo principio d'intelligenza, qualcosa di simile a quel che chiamiamo desiderio, avversione, memoria» (Maupertuis, 1768, pp. 147, 149). Rispondendo a un'obiezione di Diderot

Una posizione diversa è stata sviluppata da Gerald Edelman, ideatore della prima articolata teoria neuroscientifica della coscienza come proprietà emergente del cervello. Dopo aver elaborato la sua teoria della selezione dei gruppi neuronali, Edelman ha presentato la teoria della coscienza come il frutto di un'esigenza scientifica – «completare il programma [naturalistico] di Darwin» (Edelman, 2004, pp. 1-3) – ma anche come contributo a un dibattito filosofico che, all'inizio degli anni Ottanta, aveva trovato in Putnam e Searle due critici dell'originario programma funzionalistico e pertanto due interlocutori privilegiati (Edelman, 1989, pp. 17, 49-52). La teoria di Edelman, sviluppata successivamente con la collaborazione di Giulio Tononi (Edelman, Tononi, 2000), individua il correlato neurale della coscienza in un gruppo distribuito di neuroni chiamato «nucleo dinamico». Esso non corrisponde all'attività di un'area sempre identica del cervello, ma di un complesso di neuroni, collegati da connessioni «rientranti» (cioè tali da interagire con un meccanismo di feedback), che varia continuamente, in corrispondenza con diversi stati di coscienza qualitativi e diverse intensità di coscienza. L'ipotesi muove dalla constatazione fenomenologica che la coscienza è un fenomeno unitario, in cui avviene l'integrazione di un'informazione: in termini intuitivi, il grado di integrazione aumenta al crescere del numero di stati possibili, quindi del grado di differenziazione, che il cervello può assumere nella sua interazione con l'ambiente. Il modello fornisce dunque un metodo matematico per misurare il grado di integrazione dell'informazione, facendola corrispondere con il grado di coscienza. Così, per esempio, quando il grado di integrazione dell'informazione è molto basso – come negli stati di sonno senza sogni, in cui l'attivazione dei neuroni è molto poco differenziata – il grado della coscienza diviene evanescente.

La teoria di Edelman, come abbiamo detto, è una forma di emergentismo: i *qualia*, gli stati coscienti (soggettivi e qualitativi) "esistono" e corrispondono alle differenziazioni tra stati possibili di informazione nel cervello. Per illustrare la relazione di fondazione e dissomiglianza tra sistema nervoso e sistema cosciente, Edelman usa l'immagine

egli aveva poi ammesso che questa ipotesi non spiegava ancora l'unificazione di queste proprietà degli elementi nel singolo organismo. Analogamente Damasio incorre nuovamente nel "problema della combinazione" di James (cfr. PAR. 14.4.2). Dopo la comparsa dell'opera di Damasio, il problema è stato ricordato da molti filosofi, come Dennett (2005, pp. 10-1) e Chalmers (2010).

del combustibile e dell'esplosivo, la stessa con cui Searle ha illustrato il rapporto di dipendenza causale e indipendenza ontologica della mente cosciente rispetto al cervello. Sussiste in effetti una certa affinità tra la teoria di Edelman e quella di Searle[35]. Tuttavia Searle, pur definendo la teoria di Edelman come «il più accurato e profondo tentativo di affrontare il problema della coscienza che ho visto in letteratura», ha dichiarato che essa non spiega veramente «come» i meccanismi neurali causino la coscienza e se essi siano indispensabili perché un sistema sia cosciente: «per cui il mistero rimane» (Searle, 1997, pp. 48, 50). Edelman, rispondendo all'obiezione di non aver spiegato «il vero e proprio sentimento di un *quale*», ha successivamente affermato che la sua spiegazione rispetta uno standard tipico della metodologia scientifica. Paragonando la funzione che nel suo modello collega stati neurali e stati coscienti all'equazione newtoniana F=ma, Edelman sostiene che la conoscenza nomologica empiricamente fondata costituisce tutto quanto si può chiedere a una spiegazione scientifica:

> Queste sono le proprietà del fenotipo e qualsiasi fenotipo che sia cosciente fa esperienza dei propri qualia differenziali poiché tali qualia *sono* le distinzioni operate. È sufficiente spiegare le basi di tali distinzioni – proprio come in fisica basta descrivere materia ed energia, senza spiegare perché esiste qualcosa e non il nulla – e la nostra teoria lo può fare mettendo in rilievo le differenze nelle strutture neurali e nella dinamica alla base delle diverse modalità e funzioni cerebrali (Edelman, 2004, p. 121).

D'altra parte, Edelman ha considerato il «cosiddetto problema difficile» (di Chalmers e altri) come «mal posto, perché sembra formulato nell'aspettativa che, per un osservatore, un costrutto teorico possa condurre per descrizione al fare esperienza della qualità fenomenica che esso descrive». Rettificando questa erronea posizione del problema, possiamo concludere che «quando avremo compreso i suoi meccanismi in maggior dettaglio, la coscienza perderà il suo mistero» (Edelman, 2003, p. 5524).

Tononi ha continuato a lavorare all'elaborazione matematica ed empirica di questa teoria della coscienza, sviluppandola nella sua teoria dell'integrazione dell'informazione (IIT), con esiti parzialmente

35. Un'importante differenza è che Edelman nega la causalità autonoma degli stati coscienti (2006, p. 92).

diversi sul piano concettuale. Per introdurla Tononi compie un'analisi dell'operazione svolta dall'insieme di neuroni con la loro dinamica, confrontandola con quella di qualsiasi sistema fisico: al contrario di un semplice meccanismo dotato di sensori, che seleziona uno tra due stati possibili di ogni suo elemento (come avviene pure per ogni singolo neurone), il cervello umano, attraverso le numerosissime connessioni tra i neuroni (le sinapsi), è in grado di assumere un numero enorme di stati possibili. Sappiamo che la riduzione delle connessioni attive – per cause fisiche o chimiche, naturali o indotte –, mentre comporta la riduzione del repertorio degli stati possibili del sistema, impoverisce anche il contenuto della coscienza o lo annulla. Così un meccanismo semplice e un cervello in cui siano neutralizzate le connessioni sono ugualmente privi di coscienza. In generale, Tononi cerca di collegare le proprietà fenomenologiche della coscienza alle transizioni di stato del particolare sistema fisico che è il cervello umano, in quanto queste corrispondono all'integrazione dell'informazione. Ecco dunque l'ipotesi fondamentale: la coscienza è informazione integrata. Essa è pertanto quantificabile, assumendo un grado maggiore al crescere di questa integrazione.

Tononi individua un complesso procedimento matematico per misurare l'integrazione dell'informazione in un sistema fisico, introducendo una grandezza che chiama Φ[36]. Proviamo a vedere intuitivamente in che cosa consiste questo procedimento. A ogni sensazione si può associare una data configurazione di stati di un determinato insieme di neuroni interconnessi. Ogni transizione tra diversi stati di coscienza corrisponde a una selezione in un repertorio di possibili stati successivi (condizionata dalla struttura anatomica delle connessioni). Attraverso specifiche condizioni matematiche sulla probabilità di queste transizioni, la teoria definisce il grado di integrazione di un sistema, in quanto proprietà irriducibile a quelle delle sue parti, e individua il particolare complesso di elementi a cui corrisponde il massimo grado di integrazione, associandolo alla coscienza. In questa prospettiva matematica, l'esperienza soggettiva di un animale corrisponde a una specifica struttura geometrica nel sistema di riferimento definito dalle

36. Tononi dà un'elegante spiegazione di questa scelta notazionale: la I di informazione si sovrappone al cerchio O, che simboleggia l'integrazione. Si può ricordare che invece Feigl indicava con Φ gli stati neurofisiologici e con Ψ gli stati psichici (Feigl, 1967, p. 104).

variabili di stato dei neuroni, che si evolve con il variare degli stati di attivazione di questi ultimi.

Secondo Tononi (2004; 2012a; 2015) la teoria così impostata permette di risolvere due problemi. Il primo è il problema di fissare «le condizioni che determinano in che misura un sistema abbia coscienza». La teoria della misura della integrazione dell'informazione offre una risposta a questo problema, che si accorda con le evidenze empiriche finora disponibili ed è suscettibile di ulteriori verifiche (o falsificazioni) empiriche. Una caratteristica fondamentale di questa soluzione è che essa considera la coscienza come una proprietà «intrinseca» di un sistema, che non dipende dalle interazioni tra corpo e ambiente (nonostante il fatto che l'informazione provenga dall'ambiente e che gli stati di coscienza svolgono evidentemente un ruolo nell'adattamento)[37].

Il secondo problema è di stabilire «che cosa determina il tipo di coscienza che un sistema possiede», cioè le irriducibili «qualità», che sono esemplificate da diverse modalità (come visione, udito, dolore), sottomodalità (colore e movimento) e dimensioni (per esempio, blu e rosso), e sembrano differire a seconda che – per esempio – il soggetto che ascolta una orchestra sia un musicista esperto, un semplice appassionato di musica o un pipistrello. Secondo Tononi entrambe le questioni – quella della "quantità" e quella della "qualità" dell'esperienza – dipendono da caratteristiche del loro «sostrato fisico». La quantità e la conformazione delle connessioni neurali coinvolte – che come è noto non sono soltanto diverse tra le specie, ma anche tra individui della stessa specie, e possono mutare con l'apprendimento – determinano grado e qualità dello stato di coscienza. In questo modo Tononi ritiene che soltanto la sua teoria, tra quelle attualmente disponibili, affronti il «problema difficile» di Chalmers, piuttosto che evitarlo o liquidarlo in quanto mal posto.

L'impossibilità di ridurre la coscienza, in quanto proprietà di un sistema, alle proprietà dei suoi elementi, e dunque l'antiriduzionismo, permette di superare il principale ostacolo concettuale del "problema

37. Sul piano empirico, non a caso, Tononi ha lavorato a lungo sulla discriminazione tra stati coscienti e incoscienti nelle diverse fasi del sonno, che corrisponde al passaggio da stati con elevata differenziazione e integrazione dell'informazione neurale (nella fase REM) a stati di attività largamente omogenea e poco integrata, muovendo dal caso del sonno. Analoghe ricerche sperimentali sono state svolte confrontando, con la stimolazione transcranica, l'attività cerebrale durante la veglia e in condizione di anestesia (Ferrarelli *et al.*, 2010).

difficile", cioè lo iato tra processi fisici (distribuiti in parti) e processi coscienti (fenomenologicamente unitari e coesi). Al tempo stesso, essa comporta rilevanti e impegnative conseguenze filosofiche. In primo luogo, la coscienza (Φ) è caratterizzata come una nuova quantità fondamentale della scienza della natura:

> Al suo livello più generale, la teoria ha implicazioni ontologiche. Prende le mosse dalla fenomenologia e, facendo un uso critico di esperimenti mentali, sostiene che l'esperienza soggettiva sia una cosa sola con la capacità di un sistema di integrare l'informazione. In questa prospettiva, l'esperienza, cioè l'integrazione dell'informazione, è una quantità fondamentale, come lo sono la massa, la carica e l'energia (Tononi, 2004, p. 19).

Per apprezzare la portata filosofica di questa mossa è utile osservare che Tononi, esaminando alcuni precedenti storici della sua teoria (cfr. Tononi, 2012b, pp. 153-5), attribuisce a Kant il merito di aver sottolineato l'importanza dell'«unità sintetica» della coscienza. La teoria dell'integrazione dell'informazione, da questo punto di vista, subordina il problema della qualità degli stati coscienti a quello dell'unità sintetica della coscienza – proprio come aveva fatto Kant[38] – ma presenta una spiegazione della sintesi in termini delle proprietà globali dei sistemi fisici che la realizzano. Pertanto, il concetto della sintesi è sviluppato attraverso un modello che risale piuttosto alla filosofia di Leibniz, in quanto ricava l'attività della coscienza come una proprietà emergente dal "meccanismo" causale della natura, integrando (senza però contraddirlo o considerarlo esplicativamente inadeguato) il repertorio concettuale della scienza della natura contemporanea: così il «complesso» cosciente, definito come una «struttura che è un massimo di irriducibilità intrinseca Φ rispetto ai suoi elementi» ha una causalità intrinseca e un'unità irriducibile alla somma delle parti ed è in ciò analogo a una monade leibniziana; ma è pur sempre un «sistema fisico» (Tononi, 2015).

Questa teoria rimette pertanto in discussione su nuove basi questioni tradizionali della filosofia moderna (come l'emergere della coscienza dalla materia e la giustificazione del libero arbitrio)[39]. Al tempo

38. Si ricordi che, per Kant, la semplice rappresentazione di un concetto come "rosso" è possibile solo presupponendo l'«unità sintetica della coscienza», cioè l'Io penso (Kant, 1787, p. 109, B 133-134).

39. La più recente e aggiornata esposizione della teoria (Tononi, 2015) mette in particolare rilievo gli aspetti filosofici. Il libero arbitrio è difeso qui nei termini di una

stesso, essa presenta un ripensamento radicale – se non un rovesciamento – di alcune intuizioni del funzionalismo: ha infatti la conseguenza di estendere la coscienza a sistemi non biologici. A proposito dell'introduzione di Φ nel linguaggio scientifico Tononi scrive:

> Ne consegue che ogni sistema fisico ha esperienza soggettiva nella misura in cui è capace di integrare l'informazione, indipendentemente da ciò di cui è composto. Così, un'affascinante implicazione della teoria è che dovrebbe essere possibile costruire artefatti coscienti dotandoli di un complesso con alto valore di Φ. Inoltre, dovrebbe essere possibile progettare la qualità della loro esperienza sensibile strutturando appropriatamente la loro effettiva matrice delle informazioni (Tononi, 2004, p. 19).

Introducendo la nuova grandezza fondamentale, dunque, la teoria dell'integrazione dell'informazione realizza una vecchia ipotesi del funzionalismo di Putnam, quella della possibilità di attribuire un grado di coscienza a sistemi fisici non biologici, come gli artefatti umani, ma lo fa con un diverso concetto di coscienza, in base a cui le funzioni intellettive sono inseparabili dalle proprietà qualitative. D'altra parte, per prevenire applicazioni empiricamente abusive di questa possibilità, come il panpsichismo, la teoria pone dei parametri restrittivi alla scala spaziale e temporale entro cui l'informazione può considerarsi integrata (Tononi, Koch, 2014)[40].

Questi aspetti filosofici della teoria dell'integrazione dell'informazione, che è attualmente tra le teorie della coscienza più discusse in campo neuroscientifico, suggeriscono un commento generale: rispetto alla generazione precedente di neuroscienziati, come Changeux ed Edelman, studiosi come Damasio e Tononi – tenendo conto dell'ampio movimento filosofico antiriduzionistico degli ultimi tre decenni del secolo scorso – pongono in maggior rilievo la possibilità che, affrontando il problema della coscienza nella sua piena ricchezza fenomenologica, la ricerca scientifica possa approdare alla conferma o

teoria compatibilista, secondo cui un sistema cosciente non è soggetto a una determinazione esterna, ma effettua una «determinazione intrinseca», in cui a maggiore coscienza corrisponde una selezione tra più stati possibili e in tal senso maggiore potere causale. Questa tesi è una conseguenza dell'antiriduzionismo della teoria.

40. Per esposizioni più aggiornate della teoria cfr. ora Oizumi, Albantakis, Tononi (2014) e Tononi (2015).

all'aggiornamento di ipotesi metafisiche tipiche dell'età moderna, entrando così direttamente nel campo della filosofia.

13.5
Scienza e filosofia della coscienza: continuità o discontinuità?

13.5.1. CONTROVERSIE NATURALISTICHE

La ricerca di una "scienza della coscienza", come abbiamo visto, comporta un impegno metafisico che è stato variamente interpretato da filosofi e scienziati, portando in luce forti conflitti teorici. D'altra parte l'atteggiamento di cauto empirismo, che ispira molti ricercatori impegnati in questo campo, non sembra tanto capace di risolvere questi conflitti, quanto piuttosto di rimandarli. Così, la difesa del paradigma classico delle scienze cognitive contro le sfide dei "reazionari", messa in atto da Dennett (2005, pp. 8-13), non costituisce evidentemente una presa di posizione neutrale. In effetti, il presupposto – che Dennett condivideva con Quine – di una continuità tra filosofia e neuroscienze per l'elaborazione di una concezione naturalistica è stato tra le condizioni dei conflitti teorici, come è diventato evidente con la comparsa di teorie neuroscientifiche che non si sono limitate a integrare i dati empirici, ma hanno posto in luce il problema di interpretarli.

Si confronti per esempio il modo in cui due filosofi come Dennett e Searle hanno interpretato gli sviluppi neuroscientifici contemporanei. Nella seconda edizione di *Content and Consciousness*, comparsa nel 1985, Dennett constatava con soddisfazione che «Edelman [...], Changeux, i "nuovi connessionisti" dell'intelligenza artificiale e altri stanno ora sviluppando dei modelli "evoluzionistici" a un livello di dettaglio e sofisticazione empirica che io non potevo immaginare negli anni '60» (Dennett, 1985, p. xi). La sua spiegazione della coscienza, come abbiamo visto, rimandava ai «dettagli» forniti dai neuroscienziati, e ancora una volta il riferimento era a Edelman (Dennett, 1991, p. 184n). Più di recente Dennett ha riconosciuto ancora che lo straordinario sviluppo delle neuroscienze ha permesso una fondamentale integrazione degli approcci computazionali (Dennett, 2005, p. 155).

Da parte sua, Searle ha esaminato le più recenti teorie neuroscientifiche della coscienza, affermando di volta in volta che l'indagine sui

«poteri causali» del cervello non ha ancora trovato risposte adeguate, al punto che il problema della coscienza rimane avvolto nel «mistero». Così Searle ha ritenuto promettenti, ma insufficienti, le teorie di Crick ed Edelman (Searle, 1997), obiettando a quest'ultimo – come abbiamo visto – una carenza esplicativa; ha criticato Damasio per aver confuso concettualmente immagine mentale depositata nel cervello e coscienza, accusandolo di non aver posto il problema di quest'ultima (Searle, 2011); ha contestato a Tononi (e Koch, che nel frattempo si è avvicinato alla teoria dell'integrazione dell'informazione) il panpsichismo (Searle, 2013).

È molto istruttivo, in proposito, lo scambio di opinioni che ha avuto luogo tra Dennett e Searle intorno alla corretta interpretazione della teoria di Edelman. Mentre Dennett approvava il lavoro di Edelman sulla «evoluzione intracerebrale» (Dennett, 1985, p. XI), egli lasciava a margine la teoria del «nucleo dinamico», probabilmente perché già considerava i «qualia» – su cui si è appuntato il lavoro successivo di Edelman – come un «dono avvelenato» della filosofia alle neuroscienze (cfr. Dennett, 2005, p. 75). Riprendendo la questione, dunque, Dennett ha potuto affermare che Edelman «approva tacitamente» il suo approccio eterofenomenologico (in Searle, 1997, p. 118) e che Searle non avrebbe «nessun programma di ricerca», perché «nessuna teoria scientifica» potrà mai realizzare la sua esigenza di includere la prima persona. Searle ha correttamente replicato che Edelman non condivide l'impostazione di Dennett (ivi, p. 125). Ma è esemplare della dipendenza della filosofia dall'autorità degli scienziati il fatto che, per stabilire chi dei due si collochi a fianco di una avanguardia scientifica, entrambi abbiano fatto riferimento a «comunicazioni personali» con Edelman – trascurando il fatto che Edelman avrebbe potuto non essere infallibile nel valutare le implicazioni filosofiche della sua stessa teoria.

Il dissidio tra i due filosofi, come abbiamo visto, è radicato invece in una diversa valutazione delle scienze cognitive in generale. Dennett aderisce al paradigma e apprezza la crescita empirica graduale delle conoscenze, ritenendo inadeguata, perché troppo generica, la posizione del problema della coscienza, che rischia di fingere nuovi problemi laddove alcune proprietà del cervello sono state «già spiegate» funzionalmente e altre devono ancora esserlo (Dennett, 2005, pp. 151-7). Al contrario Searle definisce l'orientamento empirico e pragmatico, volto alla scomposizione delle nozioni della psicologia comune e alla

loro analisi in termini di singoli moduli computazionali – tipico della scienza cognitiva – come approccio a «mattoni da costruzione», e lo ritiene inadeguato a rendere conto scientificamente di proprietà globali della coscienza come l'unità, l'intenzionalità e il contenuto qualitativo, che egli raccoglie nel concetto di «campo unificato» della coscienza (Searle, 2007, p. 174).

A partire dal caso di Searle, tuttavia, si può vedere che non sussiste un accordo nemmeno tra i sostenitori della tesi secondo cui ci sarebbe il bisogno di nuova teoria scientifica del «campo unificato» della coscienza fenomenica. Per esempio Block (2009, pp. 1111-2) afferma che la teoria di Tononi è ancora una forma di funzionalismo, incapace di distinguere sperimentalmente tra coscienza fenomenica e intelligenza. D'altra parte Chalmers, che ha avanzato riserve sulle capacità esplicative delle teorie sui NCC di Crick ed Edelman (Chalmers, 2010, p. 11), considera quella di Tononi promettente – benché non conclusiva – descrivendola come «panpsichismo emergentista» (Chalmers, 2013, p. 13); e Tononi del resto – pur rifiutando questa interpretazione panpsichista – si riferisce esplicitamente alla posizione del "problema difficile" di Chalmers e tiene certamente presente la sua ipotesi metafisica dell'informazione quando attribuisce coscienza a semplici macchine, come un termostato, e quando considera la possibilità di innovare il lessico scientifico e introdurre l'esperienza come «caratteristica fondamentale del mondo, accanto a massa, carica e spazio-tempo» (le parole sono di Chalmers, 2010, p. 17). Ma considerare la coscienza come una proprietà intrinseca, reale e misurabile, ponendola dunque sullo stesso livello di altre proprietà fisiche e biologiche, costituisce un'assunzione fortemente ipotetica, la cui portata esplicativa e il cui accordo con il resto del sapere scientifico sono ancora da dimostrare. È stato ancora una volta Searle a toccare il punto dolente, presentando severamente l'ipotesi metafisica di Chalmers come frutto di «disperazione» teorica e considerandola quale semplice combinazione del funzionalismo con assunzioni «tanto implausibili da togliere il fiato» (Searle, 1997, pp. 135-76).

Come mostrano queste controversie, il problema della coscienza ha contribuito a mettere in questione la stessa omogeneità del naturalismo filosofico, che è servita in passato a stabilire, attraverso il ricorso alle scienze, una generale concezione antimetafisica della realtà, legata alla continuità della filosofia con le scienze. Pertanto, negli ultimi anni, hanno trovato voce tentativi di elaborare un naturalismo "liberalizza-

to" non riduzionista (De Caro, MacArthur, 2004), ma anche diverse critiche filosofiche all'intero programma naturalistico, che presentano i conflitti interni al naturalismo anglo-americano come insolubili alternative metafisiche e ripropongono una irriducibile discontinuità tra discorso filosofico e teoria scientifica.

13.5.2. DISCONTINUITÀ E TRADIZIONI FILOSOFICHE

Come abbiamo visto, la constatazione di una disomogeneità tra le proprietà dell'esperienza e le descrizioni neuroscientifiche dei rispettivi correlati ha spinto molti scienziati – tra cui Edelman e Gazzaniga – verso l'emergentismo e il riconoscimento di diversi livelli di descrizione. Ma affermare l'irriducibilità delle descrizioni significa anche considerare mal poste alcune delle questioni delle nuove teorie naturalistiche della coscienza, con le loro implicazioni filosofiche e metafisiche. Questo ha comportato un più frequente richiamo a genealogie filosofiche del passato alternative al naturalismo, inteso come posizione filosofica compiuta e onnicomprensiva.

Una prima famiglia di obiezioni si è richiamata all'approccio wittgensteiniano, il quale aveva portato a negare che gli stati coscienti siano qualcosa al di là dell'interazione ambientale, sociale e linguistica realizzata dagli organismi. Abbiamo visto che, mentre Ryle aveva parlato di un vero e proprio «errore categoriale» nell'attribuzione di stati mentali a un sistema fisico (Ryle, 1949), il suo allievo Dennett aveva deciso di sviluppare la critica al sostanzialismo di Ryle in un nuovo senso scientifico, ritenendo riduttivo l'approccio comportamentista. Prendendo spunto da un argomento di Wittgenstein, è stato invece Anthony Kenny a sostenere che lo studio scientifico della mente rischia di commettere una «fallacia mereologica», attribuendo al cervello, o a parti del cervello, stati che ha senso attribuire soltanto all'«essere umano come intero»[41]. Secondo Bennett e Hacker, che hanno applicato questa critica alle neuroscienze contemporanee, avrebbe certamente senso indagare empiricamente le «condizioni neurali» e «concomitanti» per il «possesso e l'esercizio» di capa-

41. L'argomento compare in Kenny (1971) con il titolo di «fallacia dell'omuncolo» ed è ispirato a un passo di Wittgenstein (1953, § 281). La riformulazione in termini di «fallacia mereologica» si trova in Bennett, Hacker (2003, pp. 68-72).

cità psicologiche, ma ciò non toglierebbe «la verità concettuale secondo cui queste capacità (*powers*) e il loro esercizio nella percezione, nel pensiero, nel sentimento sono attributi di esseri umani, non delle loro parti – in particolare, non dei loro cervelli» (Bennett, Hacker, 2003, p. 3). In un animato dibattito con gli autori, Dennett ha obiettato che questo approccio difende semplicemente la psicologia di senso comune, invocando regole linguistiche d'uso standard (senza peraltro elencarle) alla maniera di Wittgenstein, e così facendo assume un atteggiamento «oscurantista» rispetto alle prospettive e ai risultati che sono stati adottati per decenni nelle scienze cognitive (Bennett *et al.*, 2007, pp. 81-6, 90-5). Bennett e Hacker hanno replicato affermando di voler contestare un sottinteso "cartesianesimo", che si annida in tutti i tentativi di localizzazione e riduzione delle facoltà mentali, tentando di sostituirlo con un approccio "aristotelico" (ivi, pp. 130-3)[42]. Questa tesi ricorda quella di Putnam, che al «cartesianesimo-con-materialismo [...] dominante nella filosofia della mente anglo-americana» (Putnam, 2012, pp. 584-5, 589) opponeva il funzionalismo e il realismo empirico, attribuendo a queste tesi una origine aristotelica[43]. Al di là della correttezza storica, questi richiami ad Aristotele, se per un verso nascondono un sostanziale ritorno del vecchio sospetto wittgensteiniano contro il primato epistemico dell'introspezione, mostrano come si tratti ancora di rielaborare la frattura tra scienza meccanicistica e intuizione, che risale al XVII secolo, seguendo la via – sempre trac-

42. Il riferimento era al *De anima*, 408b, 12-15, dove Aristotele scrive: «Dire che un'anima è arrabbiata è come se si dicesse che l'anima tesse o costruisce. È invece meglio non dire che l'anima prova pietà, impara o pensa, ma che l'uomo fa queste cose con la sua anima». Bennett e Hacker rivolgono questo ragionamento contro i neuroscienziati che, con una mossa cartesiana, avrebbero sostituito all'anima il cervello (Bennett, Hacker, 2003, pp. 85-8, 111-2, 261-2).

43. In particolare l'importanza esclusiva del concetto di funzione, ai fini di una comprensione dell'attività cognitiva, riprenderebbe l'idea aristotelica di isolare la forma dalla materia; mentre l'esternalismo nella teoria del significato riprenderebbe il realismo diretto della teoria aristotelica. Per difendere questa attualizzazione funzionalistica dell'ilemorfismo aristotelico, Putnam (in un articolo scritto con Martha Nussbaum) ha sostenuto che si tratta di una «posizione sostenibile anche nel contesto della moderna teoria della materia» (Putnam, Nussbaum, 1992). Collegare l'idea del funzionalismo con filosofie del passato appare tuttavia al tempo stesso facile e poco pregnante. Cfr. Meerbote (1991) per la tesi di un "funzionalismo" di Kant. Riguardo al realismo diretto, l'attribuzione di questa tesi ad Aristotele è stata contestata da Caston (1998). Cfr. la replica in Putnam (2012, pp. 275-302).

ciata da Wittgenstein – di considerare il problema non come scientifico, ma come illusorio[44].

Un esito analogo hanno quelle critiche al programma scientifico sulla coscienza che si richiamano alla tradizione europea della fenomenologia trascendentale[45]. L'idea di fondo è sostenere che il problema "difficile" della coscienza possa essere risolto tenendo conto delle neuroscienze, ma non soltanto dalle neuroscienze, bensì mediante il contributo irriducibile dell'analisi fenomenologica. Un esempio di questo approccio è noto come "neurofenomenologia" e la sua idea fondamentale è molto semplice. Si tratta prima di tutto – come scrive Francisco Varela, che ha per primo sviluppato queste idee a partire dagli anni Settanta – di sottolineare il contributo euristico fondamentale delle descrizioni fenomenologiche per le indagini dei correlati neurali e comportamentali che a queste esperienze si riferiscono. Occorre dunque adottare un principio di «codeterminazione fenomenologica», basato sull'ipotesi di lavoro secondo cui «le analisi fenomenologiche della struttura dell'esperienza e le loro controparti nella scienza cognitiva sono correlate tra loro attraverso vincoli reciproci» (Varela, 1996, p. 343). Varela si preoccupa di distinguere questa assunzione dalle analoghe constatazioni, che più volte sono state avanzate dagli psicologi e dagli scienziati cognitivi, sul primato logico e sulla coevoluzione di psicologia e neuroscienze. Egli rimarca l'importanza del fatto che queste analisi siano condotte sul piano immanente dell'intuizione pura individuato da Husserl, senza nessun impegno di tipo ontologico e metafisico, dunque anche senza presupporre la distinzione tra soggetto e oggetto. Si tratta quindi non soltanto di un contributo euristico e metodologico – che in sé sarebbe peraltro poco originale –, ma di un antidoto contro la proliferazione metafisica (riduzionistica e non) che è in corso nelle ricerche sulla coscienza. Questo aspetto delle idee di Varela è stato recentemente ripreso e sviluppato da Bitbol, secondo il quale la soluzione fenomenologica del problema "difficile" della coscienza non coincide con una scelta tra le

44. Come è noto Wittgenstein, nel *Tractatus logico-philosophicus*, aveva affermato che non esistono questioni scientifiche insolubili, bensì problemi della vita che talvolta vengono erroneamente formulati come domande scientifiche (Wittgenstein, 1922, 4.11; 6.5; 6.52). Carnap aveva ripreso questa tesi nell'*Aufbau* (1928, pp. 260-1), cogliendone il riferimento all'*Ignorabimus-Streit*.

45. Si noti peraltro che la "fallacia mereologica", che gli autori appena menzionati hanno ritrovato partendo da Wittgenstein, era già stata fatta valere negli stessi termini dal fenomenologo Erwin Straus (1935).

diverse opzioni metafisiche in campo (dal materialismo al panpsichismo), ma piuttosto con una sua «dissoluzione» conseguente alla messa in evidenza che tutte queste posizioni presuppongono la coscienza e implicano delle prese di posizione interne all'esperienza. Nella prospettiva neurofenomenologica «post-dualista», «né i costituenti relativi all'esperienza in prima persona, né le strutture oggettive sono considerati come intrinsecamente fondamentali» (Bitbol, 2014, pp. 347-8)[46].

La medesima intenzione è stata manifestata dal filosofo Paul Ricoeur nel suo dialogo con Jean-Pierre Changeux. Ricoeur ha affermato che, nel valutare le tesi delle neuroscienze, occorre rifiutare l'alternativa tra dualismo e monismo della sostanza, accogliendo piuttosto un «dualismo semantico» del discorso sul cervello e sul sé, evitando così l'«amalgama semantico» che condurrebbe a espressioni sbagliate come: «il cervello pensa» (Changeux, Ricoeur, 1998, pp. 24-6). Contro il naturalismo di Changeux, Ricoeur ha sostenuto che la descrizione fenomenologica del «corpo vissuto» può aiutare a pensare l'unità dell'uomo in un «discorso misto»: le scoperte delle neuroscienze si possono riferire al «sostrato» del sistema nervoso, quale base materiale dell'esperienza, ma occorre rifiutare le tesi di chi – come Changeux – afferma che il cervello stesso possa essere «proiettivo», sviluppare la coscienza, contenere immagini (ivi, pp. 59-61, 68). L'esperienza fondamentale di «possesso» del «corpo vissuto», secondo Ricoeur, è irriducibile a qualsiasi spiegazione di tipo naturalistico: la conoscenza neurologica potrà offrire «spiegazioni» dei meccanismi che stanno alla base della vita cosciente (e mentale in genere), ma non potrà mai sostituire la «comprensione» fenomenologica di quest'ultima.

Dietro il dialogo tra Changeux e Ricoeur ardeva ancora il fuoco di vecchie polemiche culturali tra materialismo e teologia, che gli autori hanno tentato di spegnere (senza peraltro trovare un accordo sostanziale). Ma al di là delle vecchie questioni metafisiche, la prospettiva "incarnata" della fenomenologia francese – che peraltro Ricoeur riprendeva da Husserl e da Merleau-Ponty – riscuote notevole interesse negli ultimi anni a livello internazionale e viene perseguita in varie direzioni. Coniugandosi con le teorie sulla mente estesa, essa ha dato luogo a una

46. Questa posizione nega che il dominio fisico-fisiologico abbia un primato ontologico rispetto a quello della coscienza, ma lo considera come un contenuto stabilizzato dell'esperienza stessa. Questo permetterebbe secondo Bitbol di neutralizzare il "problema difficile" di Chalmers (Bitbol, 2014, p. 350).

posizione che è stata in seguito ripresa da diversi filosofi, secondo cui il problema della coscienza non si può ridurre a quello di elaborare una teoria scientifica[47].

Un aspetto comune di queste diverse forme di critica al programma di ricerca sulla coscienza consiste nel metterne in rilievo i pregiudizi ontologici di fondo, senza entrare nel merito delle teorie scientifiche, ma separandole dalla componente metafisica che vi è intrecciata. Già Wittgenstein, in un noto passo delle *Philosophische Untersuchungen* (1953, § 304), osservava che la sua analisi critica del linguaggio privato non aveva escluso l'esistenza dei processi mentali: «Tuttavia ritorni sempre al risultato che la sensazione in sé stessa non è nulla. – Niente affatto. Non è qualcosa, ma non è nemmeno nulla». Questa conclusione, nel tipico stile maieutico di Wittgenstein, sollevava in fondo il problema dei malintesi che si annidano nelle teorie che parlano degli stati coscienti e del persistente rischio di reificazione che essi producono – una questione che anche la tradizione fenomenologica ha a lungo sviluppato. Lo scopo di Wittgenstein era distinguere l'uso delle parole che descrivono un processo interiore, con cui conduciamo la nostra esperienza, da questo processo stesso, che invece «non [è] ancora compreso» (ivi, §§ 305, 308, 412). Parte delle controversie sulle teorie della coscienza dipendono effettivamente dal conflitto tra chi ha inteso negare la realtà in senso proprio di qualsiasi proprietà puramente fenomenica, come Putnam, P. S. Churchland e Dennett[48], e chi invece, come Searle (e i fenomenologi), ha sostenuto che «quando si tratta di coscienza l'esistenza del fenomeno è la realtà» (Searle, 1997, p. 112).

Ma il problema della continuità o discontinuità tra teorie neuroscientifiche e filosofia non si riduce al dilemma tra una nozione univoca e una plurale di "realtà", che come tale si può sciogliere distinguendo diversi livelli di descrizione. L'aspetto più scottante della questione

47. Una introduzione a questa prospettiva si trova in Gallagher, Zahavi (2008).

48. Dennett ha insistito più volte sull'esigenza di una rottura ontologica: «Poiché io nego che qualsiasi entità possa avere le caratteristiche di un dolore o un pensiero, tanto peggio per lo statuto ontologico di queste cose» (Dennett, 1978, p. 38). Egli cioè «sostiene che non ci sono cose come i dolori, anche se la gente prova dolore» (Dennett, 1980, p. 76). Queste affermazioni sembrano avere un significato eliminativistico – che Searle (1981, p. 416) considerò pertanto contraddittorio – nel contesto della sua epistemologia strettamente ancorata alle scienze cognitive classiche. Cfr. però Dennett (1991b), dove l'autore apre uno spiraglio su una realtà *sui generis*, non oggettiva, degli stati intenzionali.

consiste nel fatto che una compiuta teoria sul meccanismo della coscienza potrebbe implicare nuove tecniche di modificazione della coscienza stessa attraverso le sue basi materiali. Rispetto a questa ipotesi i chiarimenti ontologici assumono un rilievo non soltanto metafisico, ma anche medico, etico e politico. Le alternative concettuali che restano implicite nella pratica scientifica tornano, pertanto, ad assumere un valore non meramente verbale.

Anche questo aspetto della questione ha radici moderne. Come mostra il caso di Ricoeur la prospettiva discontinuista costituisce una rielaborazione della vecchia idea, fatta valere dalla tradizione kantiana contro il materialismo e lo spiritualismo, secondo cui la realtà non è riducibile a un unico modo di apparire, come quello oggettivo praticato dalla scienza della natura. Ed è noto come questa prospettiva si legasse alla difesa dell'autonomia della logica, della morale, del diritto, della religione pura rispetto ai principi e alle leggi della scienza della natura. D'altra parte, come ha mostrato la storiografia più recente, il riconoscimento della dimensione culturale della natura umana e della sua plasticità è stato ben presente anche nelle concezioni materialistiche (Wolfe, 2014). A lungo il peso di dicotomie troppo rigide, come materialismo/idealismo, meccanicismo/vitalismo, natura/cultura ha impedito la ricerca di una convergenza teorica, e per questo molti filosofi come Wittgenstein hanno deciso di tagliare il nodo alla radice, isolando i problemi della vita dai problemi della conoscenza scientifica. Il problema della nuova scienza della coscienza riapre dunque una ricerca interrotta, avviata all'epoca in cui il discorso metafisico sull'anima-sostanza fu definitivamente escluso dalla ricerca sul rapporto tra coscienza e cervello, che spesso è ignorata da chi se ne occupa. La questione filosoficamente decisiva, dunque, non è tanto se (e quando) questa ricerca porterà a un nuovo paradigma scientifico-tecnologico, ma se essa saprà riconoscere la presenza di strutture concettuali del passato, che talvolta si annidano nelle teorie scientifiche senza essere notate.

Riferimenti bibliografici

ARMSTRONG D. (1981), *The Nature of Mind*, The Harvester Press, Brighton (UK).

BAARS B. (1988), *A Cognitive Theory of Consciousness*, Cambridge University Press, Cambridge.

BAYERTZ K., GERHARD M., JAESCHKE W. (2007), *Weltanschauung, Philosophie und Naturwissenschaft im 19. Jahrhundert. Der Ignorabimus-Streit*, Meiner, Hamburg.

BENNETT M. R., HACKER P. M. S. (2003), *Philosophical Foundations of Neurosciences*, Blackwell, Malden.

BENNETT M. R. *et al.* (2007), *Neuroscience and Philosophy*, Columbia University Press, New York.

BITBOL M. (2014), *La conscience a-t-elle une origine?*, Flammarion, Paris.

BLOCH O. (1995), *Le matérialisme*, PUF, Paris (1ᵉ éd. 1985).

BLOCK N. (1995), *On a Confusion about a Function of Consciousness*, in "Behavioral and Brain Sciences", 18, pp. 227-87.

ID. (2009), *Comparing the Major Theories of Consciousness*, in M. Gazzaniga (ed.), *The Cognitive Neurosciences*, The MIT Press, Cambridge (MA) (4th ed.), pp. 1111-22.

ID. (2011), *Perceptual Consciousness Overflows Cognitive Access*, in "Trends in Cognitive Sciences", 15, 12, pp. 567-75.

ID. (2014), *Consciousness, Big Science and Conceptual Clarity*, in G. Marcus, J. Freeman (eds.), *The Future of the Brain: Essays by the World's Leading Neuroscientists*, Princeton University Press, Princeton, pp. 161-76.

ID. (2015), *The Canberra Plan Neglects Ground*, in T. Horgan, M. Sabates, D. Sosa (eds.), *Qualia and Mental Causation in a Physical World* (in press).

CABANIS P. J. G. (1802), *Rapports du physique et du moral de l'homme*, Crapart, Paris.

CARNAP R. (1928), *Der logische Aufbau der Welt*, Weltkreis Verlag, Berlin.

ID. (1963a), *Intellectual Autobiography*, in P. A. Schilpp, *The Philosophy of Rudolf Carnap*, Open Court, La Salle, pp. 3-83.

ID. (1963b), *Herbert Feigl on Physicalism*, in P. A. Schilpp, *The Philosophy of Rudolf Carnap*, Open Court, La Salle, pp. 881-5.

CASTON V. (1998), *Aristotle and the Problem of Intentionality*, in "Philosophy and Phenomenological Research", 58, 2, pp. 249-98.

CHALMERS D. (1996), *The Conscious Mind. In Search of a Fundamental Theory*, Oxford University Press, Oxford.

ID. (2000), *What Is a Neural Correlate of Consciousness*, in T. Metzinger (ed.), *Neural Correlates of Consciousness*, The MIT Press, Cambridge (MA), pp. 17-39.

ID. (2010), *The Character of Consciousness*, Oxford University Press, Oxford.

ID. (2013), *The Combination Problem for Panpsychism*, <http://consc.net/papers/combination.pdf>.

CHANGEUX J. P. (1983), *L'homme neuronal*, Fayard, Paris.

CHANGEUX J. P., RICOEUR P. (1998), *Ce qui nous fait penser. La nature et la régle*, Odile Jacob, Paris.

CHURCHLAND P. (1986), *Neurophilosophy. Toward a Unified Science of the Mind-Brain*, The MIT Press, Cambridge (MA).
CLERICUZIO A. (2015), *Le forme e i moti della materia. Trasformazioni del meccanicismo nel Seicento*, in P. Pecere (a cura di), *Il libro della natura*. I. *Scienze e filosofia da Copernico a Darwin*, Carocci, Roma, pp. 67-107.
CRICK F. H. (1994), *The Astonishing Hypothesis: The Scientific Search for the Soul*, Scribner, New York.
DAMASIO A. (1999), *The Feeling of What Happens. Body and Emotion in the Making of Consciousness*, Harvest Book, Orlando.
ID. (2004), *Looking for Spinoza*, Vintage Books, London.
ID. (2010), *Self Comes to Mind. Constructing the Conscious Brain*, Pantheon, New York.
DAVIDSON D. (1970), *Mental Events*, in L. Foster, J. Swanson (eds.), *Experience and Theory*, University of Massachusetts Press, Boston, pp. 79-101.
DE CARO M. (2013), *Le neuroscienze cognitive e l'enigma del libero arbitrio*, in M. Di Francesco, M. Marraffa, *Il soggetto. Scienze della mente e natura dell'io*, Bruno Mondadori, Milano, pp. 147-65.
DE CARO M., MACARTHUR D. (eds.) (2004), *Naturalism in Question*, Harvard University Press, Cambridge (MA).
DENNETT D. (1978), *Brainstorms*, Bradford Books, Montgomery.
ID. (1980), *Reply to Prof. Stich*, in "Philosophical Books", 21, 2, pp. 73-6.
ID. (1985), *Content and Consciousness*, Routledge and Kegan Paul, London (1ª ed. 1969).
ID. (1987), *The Intentional Stance*, The MIT Press, Cambridge (MA).
ID. (1991a), *Consciousness Explained*, Back Bay Books, New York.
ID. (1991b), *Real Patterns*, in "The Journal of Philosophy", 88, 1, pp. 27-51.
ID. (2005), *Sweet Dreams. Philosophical Obstacles to a Science of Consciousness*, The MIT Press, Cambridge (MA).
DESCARTES R. (1641), *Meditationes de prima philosophia*, Elsevier, Amsterdam, ora in Id., *Oeuvres*, vol. VII, Vrin, Paris 1996, pp. 1-561 (trad. it. in Id., *Opere 1637-1649*, Bompiani, Milano 2009, pp. 703-99).
ID. (1644), *Principia philosophiae*, Elsevier, Amsterdam, ora in Id., *Oeuvres*, vol. VIII, Vrin, Paris 1996, pp. 1-329 (trad. it. in Id., *Opere 1637-1649*, Bompiani, Milano 2009, pp. 1705-2211).
ID. (1662), *L'homme*, Schuyl, Leida, ora in *Oeuvres*, vol. XI, Vrin, Paris 1996, pp. 119-215 (trad. it. in Id., *Opere postume*, Bompiani, Milano 2009, pp. 361-507).
DIDEROT D. (1766), *Idée*, in D. Diderot, J. le Rond d'Alembert (éd.), E*ncyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers*, vol. VIII, Briasson, Paris, pp. 489-94.
ID. (1987), *Le rêve de D'Alembert*, in Id., *Oeuvres complètes*, vol. XVII, Hermann, Paris 1975- (il testo, comparso postumo, è del 1769).

DILTHEY W. (1883), *Einleitung in die Geisteswissenschaften*, Duncker & Humblot, Leipzig.
DU BOIS-REYMOND E. (1872), *Über die Grenzen des Naturerkennens*, in Id., *Reden*, Veit, Leipzig 1886, pp. 105-40.
ID. (1875), *La Mettrie*, in Id., *Reden*, Veit, Leipzig 1886, pp. 178-210.
ID. (1880), *Die Sieben Welträthsel*, in Id., *Reden*, Veit, Leipzig 1886, pp. 381-417.
EDELMAN G. (1989), *Remembered Present. A Biological Theory of Consciousness*, Basic Books, New York.
ID. (2003), *Naturalizing Consciousness. A Theoretical Network*, in "PNAS", 100, 9, pp. 5520-4.
ID. (2004), *Wider than the Sky. The Phenomenal Gift of Consciousness*, Yale University Press, New Haven.
ID. (2006), *Second Nature. Brain Science and Human Knowledge*, Yale University Press, New Haven.
EDELMAN G., TONONI G. (2000), *A Universe of Consciousness. How Matter Becomes Imagination*, Basic Books, New York.
FECHNER G. T. (1854), *Über die Atomistik*, in "Zeitschrift für Philosophie und philosophische Kritik", 25, pp. 25-57.
FEIGL H. (1958/1967), *The "Mental" and the "Physical". The Essay and a Postscript*, University of Minnesota Press, Minneapolis (1958[1]).
FERRARELLI F. *et al.* (2010), *Breakdown in Cortical Effective Connectivity during Midazolam-Induced Loss of Consciousness*, in "PNAS", 107, 6, pp. 2681-6.
GABBEY A. (2002), *Newton, Active Powers and the Mechanical Philosophy*, in I. B. Cohen, G. E. Smith (eds.), *The Cambridge Companion to Newton*, Cambridge University Press, Cambridge, pp. 329-57.
GALL F. (1798), *Über die Verrichtungen des Gehirns der Menschen und der Thiere*, in "Der neue Teutscher Merkur", 3, pp. 311-32.
GALLAGHER S., ZAHAVI D. (2008), *The Phenomenological Mind*, Routledge, London.
GAZZANIGA M. (2011), *Who's in Charge? Free Will and the Science of the Brain*, Harper Collins, New York.
GAZZANIGA M., LE DOUX J. E. (1978), *The Integrated Mind*, Plenum Press, New York.
GRAHEK N. (2007), *Feeling Pain and Being in Pain*, The MIT Press, Cambridge (MA).
HAECKEL E. (1899), *Die Welträthsel. Gemeinverständliche Studien über monistische Philosophie*, A. Kröner, Stuttgart.
HAGNER M. (2002), *Cyber-Phrenologie. Die neue Physiognomik des Geistes und ihre Ursprünge*, in K. P. Dencker, *Die Politik der Maschine*, Hans Bredow Institut Verlag, Hamburg, pp. 182-97.
ID. (2008), *Homo cerebralis. Der Wandel vom Seelenorgan zum Gehirn*, Suhrkamp, Frankfurt am Main (2000[1]).

HASLER F. (2012), *Neuromythologie. Eine Streitschrift gegen die Deutungsmacht der Hirnforschung*, Transcript Verlag, Bielefeld.

HATFIELD G. (1995), *Remaking the Science of Mind. Psychology as a Natural Science*, in C. Fox, R. Porter, R. Wokler (eds.), *Inventing Human Science*, University of California Press, Berkeley, pp. 184-231.

ID. (2009), *Sense Data and the Mind-Body Problem*, in R. Schumacher (ed.), *Perception, Reality: From Descartes to the Present*, Mentis, Berlin 2004, pp. 305-31, ora in Id., *Perception and Cognition*, Clarendon Press, Oxford, pp. 297-322.

HEIDELBERGER M. (1993), *Die Innere Seite der Natur. Gustav Theodor Fechners wissenschaftlich-philosophische Weltauffassung*, Klostermann, Frankfurt am Main.

ID. (2003), *The Mind-Body Problem in the Origin of Logical Empiricism: Herbert Feigl and Psychophysical Parallelism*, in P. Parrini, W. Salmon (eds.), *Logical Empiricism: Historical and Contemporary Perspectives*, Pittsburgh University Press, Pittsburgh.

HEIMANN P. M., MCGUIRE J. E. (1971), *Newtonian Forces and Lockean Powers. Concepts of Matter in Eighteenth Century Thought*, in "Historical Studies in the Physical Sciences", 3, pp. 233-306.

HUMPHREY N. (1970), *What the Frog's Eye Tells the Monkey's Brain*, in "Brain, Behavior, and Evolution", 3, pp. 324-37.

ID. (2006), *Seeing Red: A Study in Consciousness*, The Belknap Press of Harvard University Press, Cambridge (MA).

JAMES W. (1890), *The Principles of Psychology*, Holt & Company, New York.

KANDEL E. (2000), *The Brain and Behaviour*, in E. Kandel, J. H. Schwartz, T. M. Jessell, *Principles of Neural Science*, McGraw-Hill, New York (4th ed.), pp. 5-18.

KANT I. (1900–), *Kant's gesammelte Schriften*, Hrsg. Königlich Preußischen Akademie der Wissenschaften, de Gruyter, Berlin: vol. III, *Kritik der reinen Vernunft* (2ª ed. Hartknoch, Riga 1787; è indicata anche la paginazione di questa edizione preceduta da B); vol. XII, pp. 30-5, *Über das Organ der Seele* (ed. or. in T. Sömmering, *Über das Organ der Seele*, Nicolovius, Königsberg 1796, pp. 81-86).

KENNY A. (1971), *The Homunculus Fallacy*, in M. Grene (ed.), *Interpretations of Life and Mind*, Routledge, London, pp. 65-74.

KIM J. (1998), *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*, The MIT Press, Cambridge (MA).

ID. (2005), *Physicalism, or Something Near Enough*, Princeton University Press, Princeton.

KOCH C. (2004), *The Quest for Consciousness. A Neurobiological Approach*, Roberts & Company, Englewood.

KOCH C., TONONI G. (2008), *The Neural Correlates of Consciousness. An Updated Account*, in "Annals of the New York Academy of Sciences", 1124, pp. 239-61.
KRIPKE S. (1972), *Naming and Necessity*, Harvard University Press, Cambridge (MA).
KUSCH M. (1995), *Psychologism. A Case Study in the Sociology of Philosophical Knowledge*, Routledge, London.
LANGE F. (1875), *Geschichte des Materialismus und Kritik seiner Bedeutung in der Gegenwart*, Baedeker, Iserlohn, vol. II. *Geschichte des Materialismus seit Kant* (1ª ed. 1866).
LEGRENZI P., UMILTÀ C. (2009), *Neuro-mania. Il cervello non spiega chi siamo*, il Mulino, Bologna.
LEVINE J. (1983), *Materialism and Qualia. The Explanatory Gap*, in "Pacific Philosophical Quarterly", 64, pp. 354-61.
LEWIS D. K. (1966), *An Argument for the Identity Theory*, in "Journal of Philosophy", 63, pp. 17-25.
LIBET B. (2004), *Mind Time. The Temporal Factor in Consciousness*, Harvard University Press, Cambridge (MA).
MARRAFFA M. (2002), *Filosofia della mente*, in N. Vassallo, F. D'Agostini (a cura di), *Storia della filosofia analitica*, Einaudi, Torino, pp. 194-222.
MARRAFFA M., PATERNOSTER A. (2012), *Persone, menti, cervelli. Storia, metodi e modelli delle scienze della mente*, Mondadori Education, Milano.
MAUPERTUIS P. L. (1768), *Système de la nature*, in Id., *Oeuvres*, J. M. Bruyset, Lyon, vol. I, pp. 135-216.
MAXWELL G. (1974), *Unity of Consciousness and Mind-Brain Identity*, in J. C. Eccles (ed.), *Mind and Brain: The Many Faceted Problems*, Paragon House, New York, pp. 233-7.
MEEHL P. E., SELLARS W. (1956), *The Concept of Emergence*, in H. Feigl (ed.), *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, University of Minnesota Press, Minneapolis, pp. 239-52.
MEERBOTE R. (1991), *Kant's Functionalism*, in J. C. Smith (ed.), *Historical Foundations of Cognitive Science*, Kluwer Academic Publishing, Dordrecht, pp. 161-88.
MILLIKAN R. (1984), *Language, Thought and Other Biological Categories*, The MIT Press, Cambridge (MA).
NAGEL T. (1971), *Brain Bisection and the Unity of Consciousness*, in "Synthese", 22, pp. 396-413.
ID. (1974), *What It Is Like to Be a Bat?*, in "Philosophical Review", 83, 4, pp. 435-50.
NOË A., THOMPSON E. (2004), *Are there Neural Correlates of Consciousness?*, in "Journal of Consciousness Studies", 11, 1, pp. 3-28.
PECERE P. (2015), *Kant's Über das Organ der Seele and the Limits of Physiology: Argument and Legacy*, in R. Hanna *et al.* (eds.), *Critical Paths outside*

the Critiques. Kant's Shorter Writings, Cambridge Scholars Publishers, Newcastle upon Tyne (in press).
PLACE U. T. (1956), *Is Consciousness a Brain Process?*, in "British Journal of Psychology", 47, pp. 44-50.
POGGI S. (1977), *I sistemi dell'esperienza. Psicologia, logica e teoria della scienza da Kant a Wundt*, il Mulino, Bologna.
PUTNAM H. (1975), *Mind, Language and Reality. Philosophical Papers*, vol. II, Cambridge University Press, Cambridge.
ID. (2012), *Philosophy in an Age of Science*, Harvard University Press, Cambridge (MA).
PUTNAM H., NUSSBAUM M. C. (1992), *Changing Aristotle's Mind*, in M. C. Nussbaum, A. Rorty (eds.), *Essays on Aristotle's "De Anima"*, Oxford University Press, Oxford, pp. 27-56.
QUINE W. V. O. (1987), *Quiddities. An Intermittently Philosophical Dictionary*, Belknap Press of Harvard University Press, Cambridge (MA).
REICHENBACH H. (1938), *Experience and Prediction*, University of Chicago Press, Chicago.
RIEHL A. (1879), *Der philosophische Kriticismus*, vol. II, Engelmann, Leipzig.
ROUSSEAU G. S. (2004), *Nervous Acts. Essays on Literature, Culture and Sensibility*, Palgrave McMillan, Basingstoke-New York.
RYLE G. (1949), *The Concept of Mind*, Hutchinson, London.
SCHLICK M. (1925), *Allgemeine Erkenntnislehre*, Springer-Verlag, Berlin (ora Suhrkamp, Frankfurt am Main 1979).
SCRIBANO E. (2015), *Macchine con la mente. Fisiologia e metafisica tra Cartesio e Spinoza*, Carocci, Roma.
SEARLE J. (1980), *Minds, Brains and Programs*, in "Behavioral and Brain Sciences", 3, pp. 417-24.
ID. (1981), *Analytic Philosophy and Mental Phenomena*, in "Midwest Studies in Philosophy", 6, 1, pp. 405-24.
ID. (1983), *Intentionality*, Cambridge University Press, Cambridge.
ID. (1992), *The Rediscovery of the Mind*, The MIT Press, Cambridge (MA).
ID. (1997), *The Mistery of Consciousness*, The New York Review of Books, New York.
ID. (2004), *Mind. A Brief Introduction*, Oxford University Press, Oxford.
ID. (2007), *Dualism Revisited*, in "Journal of Physiology", 101, pp. 169-78.
ID. (2011), *The Mystery of Consciousness Continues*, in "The New York Review of Books", June 9.
ID. (2013), *Can Information Theory Explain Consciousness?*, in "The New York Review of Books", January 10.
SELLARS W. (1962), *Philosophy and the Scientific Image of Man*, in Id., *Science, Perception and Reality*, Ridgeview, Atascadero 1991, pp. 1-40 (ed. or. University of Pittsburgh Press, Pittsburgh 1962).

SMART J. J. C. (1959), *Sensations and Brain Processes*, in "Philosophical Review", 68, pp. 141-56.
STRAUS E. (1935), *Vom Sinn der Sinne. Ein Beitrag zur Grundlegung der Psychologie*, Springer-Verlag, Berlin.
TENNANT N. (2007), *Mind, Mathematics and the Ignorabimusstreit*, in "British Journal for the History of Philosophy", 15, 4, pp. 745-73.
THOMSON A. (2008), *Bodies of Thought. Science, Religion and the Soul in the Early Enlightenment*, Oxford University Press, Oxford.
TONONI G. (2004), *An Information Integration Theory of Consciousness*, in "BMC Neuroscience", 5, pp. 1-22.
ID. (2012a), *Integrated Information Theory of Consciousness: An Updated Account*, in "Archives Italiennes de Biologie", 150, pp. 290-326.
ID. (2012b), *PHI. A Voyage from the Brain to the Soul*, Pantheon, New York.
ID. (2015), *Integrated Information Theory*, in "Scholarpedia", 10, 1, p. 4164.
TONONI G., KOCH C. (2014), *Here, there, but not everywhere*, <http://arxiv.org/ftp/arxiv/papers/1405/1405.7089.pdf>.
TYE M. (2013), *Qualia*, in "The Stanford Encyclopedia of Philosophy", April 22, <http://plato.stanford.edu/entries/qualia/> (revisione sostanziale della 1ª ed. del 1997).
VARELA F. (1996), *Neurophenomenology: A Methodological Remedy to the "Hard Problem"*, in "Journal of Consciousness Studies", 3, pp. 330-46.
VIDAL F. (2006), *Les Sciences de l'âme XVIe-XVIIIe siècle*, Honoré Champion, Paris.
WHITAKER H., SMITH C. U. M., FINGER S. (eds.) (2007), *Brain, Mind and Medicine: Essays in Eighteenth-Century Neuroscience*, Springer-Verlag, Berlin.
WITTGENSTEIN L. (1953), *Philosophical Investigations*, Blackwell, Oxford.
WOLFE T. C. (2010), *Endowed Molecules and Emergent Organization: The Maupertuis-Diderot Debate*, in "Early Science and Medicine", 15, pp. 38-65.
ID. (2013), *Un matérialisme désincarné: la théorie de l'identité cerveau-esprit*, in "Matèrie prèmiere", 1, pp. 77-100.
ID. (2014), *Materialism*, in A. Garrett (ed.), *The Routledge Companion to Eighteeenth Century Philosophy*, Routledge, London, pp. 91-118.
WUNDT W. (1863), *Vorlesungen über die Menschen- und Tierseele*, vol. 1, Voss, Leipzig.
ID. (1873), *Grundzüge der physiologischen Psychologie*, Engelmann, Leipzig 1902⁵ (1873¹).