

DALLA FILOSOFIA ALLE NEUROSCIENZE E RITORNO

Appunti per una prospettiva neurofilosofica

Alfredo PATERNOSTER

In this paper I shall discuss the question whether philosophy and neurosciences could collaborate on a common research project. I shall argue that neurosciences are relevant to philosophy only to the extent that they have something to say on mind. This is the case of cognitive neuroscience, the discipline whose object is the study of the neural bases of mental processes and capacities. Through the discussion of two case studies, I shall show how cognitive neurosciences could be of advantage to philosophy and vice-versa.

Il rapporto tra filosofia e scienza è un tema su cui i filosofi si sono interrogati spesso negli ultimi cento anni o giù di lì. Due grandi filosofi del Novecento, Wittgenstein e Quine, hanno impersonato le due posizioni canoniche che si possono nutrire al riguardo, quella separatista e quella continuista.

In breve, secondo Wittgenstein la scienza si occupa del mondo empirico, mira cioè a scoprire dei fatti e a proporne certe spiegazioni, mentre la filosofia si occupa di concetti (quindi, nella visione del grande filosofo viennese, di linguaggio): la filosofia non scopre né spiega nulla e l'unico compito che può assumersi è quello di mostrare da quali fraintendimenti linguistici originano i problemi filosofici. Una volta mostrata l'origine linguistica del problema, la confusione concettuale che ne è alla base, il problema si dissolverà. Le proposizioni della filosofia sono a priori e hanno natura "grammaticale", esprimono cioè norme d'uso del linguaggio; e sono verità soltanto per modo di dire, perché, non potendo essere false, di tali proposizioni non ha senso nemmeno dire che sono vere: la proprietà di essere vero o falso non si applica alle norme.

All'opposto Quine nega che si possa tracciare una distinzione tra verità analitiche (= proposizioni vere meramente in virtù del loro significato) e verità sintetiche, o fattuali: fatti e concetti sono intrecciati inestricabilmente, da cui l'impossibilità di distinguere sedicenti problemi puramente filosofici da problemi scientifici. Questo apre la strada alla contaminazione tra scienza e filosofia, in ambedue le direzioni: la scienza incorpora più o meno esplicitamente assunzioni filosofiche; e le sedicenti

proposizioni filosofiche non sono immuni da revisioni sollecitate dalla scoperta di fatti empirici.

Ho fatto questa premessa perché le considerazioni che mi accingo a proporre in questo articolo presuppongono una cornice continuista o, come molti amano dire, una forma di naturalismo filosofico; se invece si pensa come Wittgenstein, allora l'idea stessa che le neuroscienze possano dire qualcosa di interessante sui problemi filosofici sarà giudicata come l'esito di un'irrimediabile confusione. Assumo dunque senza discussione ulteriore il punto di vista continuista, che appare oggi, pur con alcuni *caveat* su cui possiamo qui sorvolare, decisamente maggioritario. Ciò che dirò in questo articolo porterà altra acqua al mulino del continuismo.

Ci sono (almeno) due modi diversi in cui filosofia e neuroscienze possono incontrarsi. Un modo consiste semplicemente nel fatto che alcune scoperte delle neuroscienze hanno rilievo filosofico. Questo tipo di rapporto è del tutto estrinseco, nel senso che filosofia e neuroscienze non solo restano due pratiche epistemiche ben distinte (questo è ovvio), ma nemmeno intrattengono una cooperazione di qualche tipo. La filosofia prende atto, per così dire, di certi fatti e ragiona su quali conseguenze ne risultano per la nostra immagine del mondo ed eventualmente per i nostri comportamenti. Il secondo tipo di rapporto è più intrigante e problematico e consiste nella possibilità che filosofia e neuroscienze collaborino in una ricerca comune, beninteso ognuna con i propri metodi e dalla sua propria prospettiva; quindi, nel caso della filosofia, principalmente con l'analisi delle argomentazioni e dei concetti usati dagli scienziati, l'interpretazione di certi risultati sperimentali, il confronto non acritico col senso comune. Per comodità chiamerò "neurofilosofia" questo tipo di rapporto¹.

In questo articolo mi occuperò soltanto di questo secondo tipo di relazione, quella neurofilosofica. Riguardo al primo tipo, mi limiterò a fare il seguente esempio.

¹ Questa espressione è stata introdotta da Patricia Smith Churchland principalmente per indicare una certa concezione del rapporto tra concetti mentalistici e concetti neuroscientifici, e di conseguenza un certo modo di fare filosofia, empiricamente informato dalla ricerca neuroscientifica. Si tratta di un'accezione più ristretta di quella che uso qui. Cfr. Patricia S. CHURCHLAND, *Neurophilosophy. Toward a Unified Understanding of the Mind/Brain*, MIT Press, Cambridge (MA) 1986. Un'accezione più simile alla mia si trova nel recentissimo Michele DI FRANCESCO, *Spiegazione neurocentrica e mente personale. Tra filosofia e neuroscienze*, "Rivista di Filosofia", vol. CVIII, n. 2, Anno 2017, pp. 234-247.

È stato sostenuto, in modo particolare dagli studiosi che hanno promosso la nascita di una vera e propria disciplina, la neuroetica², che alcune scoperte delle neuroscienze mettono in discussione gli assunti di senso comune –e implicitamente incorporati nel diritto– concernenti la responsabilità individuale e il libero arbitrio. Esperimenti, peraltro controversi, come quelli di Libet³ mostrerebbero infatti che l'intenzione deliberata di eseguire una certa azione è successiva di oltre un centinaio di millisecondi all'istante in cui il cervello avvia l'azione. Almeno a uno sguardo superficiale sembra seguirne che, invece di essere consapevolmente responsabili delle nostre azioni, siamo “agiti” dal nostro cervello, quasi che sia stato il cervello, non *noi*, a decidere l'azione da intraprendere e successivamente a creare l'illusione che la decisione sia stata nostra. Non vi è dubbio che, assumendo che questa interpretazione dei risultati sperimentali sia corretta, ci sia molto da fare per un filosofo morale; ma appunto si tratterebbe di una riflessione sulle conseguenze *etiche* di certi fatti, non un esempio di collaborazione tra filosofia e neuroscienze. Diversamente starebbero le cose se il filosofo discutesse criticamente l'interpretazione dei dati, entrando così nel territorio delle neuroscienze. Come avremo modo di vedere, la discussione degli esperimenti è proprio uno dei compiti principali che un neurofilosofo nella mia accezione si deve assumere.

Ciò basti per l'impatto delle neuroscienze sull'etica. Veniamo a quello che qui ci interessa: discutere quale tipo di fecondazione reciproca, se c'è, sussiste tra la filosofia e le neuroscienze; altrimenti detto, discutere le prospettive della neurofilosofia nell'accezione qui utilizzata. Per cominciare, pongo una domanda preliminare. Alla luce dell'esistenza di aree di ricerca in cui filosofi e neuroscienziati collaborano, alcuni direbbero che in questi ultimi anni è sorta una nuova disciplina: la filosofia delle neuroscienze. Si tratta di un giudizio fondato?

Si osservi, in primo luogo, che nel corso della seconda metà del Novecento c'è stata una grande fioritura delle cosiddette filosofie speciali delle singole scienze, come la filosofia della fisica, la filosofia della biologia, la filosofia della psicologia. Pertanto

² Andrea LAVAZZA, Giuseppe SARTORI, *Neuroetica. Una nuova prospettiva di ricerca*, “Giornale Italiano di Psicologia”, vol. 4, anno 2010, pp. 755-778; Andrea LAVAZZA, Giuseppe SARTORI (a cura di), *Neuroetica*, il Mulino, Bologna 2011.

³ Benjamin LIBET, *Mind Time. The temporal factor in consciousness*, Harvard University Press, Cambridge MA 2004; tr. it. Pier Daniele NAPOLITANI, *Mind Time* (a cura di Edoardo BONCINELLI), Raffaello Cortina, Milano 2007.

sarebbe tutt'altro che sorprendente se nascesse una filosofia delle neuroscienze; e alcuni direbbero che già c'è. Indubbiamente vi sono problemi filosofici che potrebbero essere rubricati sotto questa categoria. Per esempio Friedemann Pulvermüller, uno dei più noti neuroscienziati impegnati nella ricerca delle basi neurali del significato, ritiene che il cosiddetto problema del *symbol grounding*, vale a dire il problema di spiegare in che modo le espressioni del linguaggio naturale sono connesse al loro significato referenziale –a ciò che ad esse corrisponde nel mondo reale–, trovi una soluzione nelle cosiddette teorie “*embodied*”, secondo le quali la comprensione del significato di una parola viene realizzata tramite l’attivazione delle aree sensorimotorie⁴. Si tratta quindi di una tesi che istituisce una connessione esplicita tra una questione filosofica e certi fatti neurali: viene avanzata una giustificazione neurale di una tesi filosofica. A loro volta, i filosofi che (come chi scrive) si occupano delle basi naturali del significato e della comprensione sono oggi costretti a misurarsi con l’immensa letteratura neuroscientifica relativa ai pattern di attivazione sottostanti ai processi di comprensione e di immaginazione. Si noti come la discussione di problemi di questo genere comporti una conoscenza almeno superficiale delle tecniche sperimentali oggi disponibili (risonanza magnetica funzionale, stimolazione magnetica transcranica, EEG ecc.). Dunque questo è sicuramente un caso in cui un problema in prima istanza scientifico (come funziona la comprensione del linguaggio, declinato nella forma “come il cervello rende possibile la comprensione del linguaggio”) mette in gioco questioni filosofiche; e queste, a loro volta, non possono trovare una risposta soltanto in una riflessione “in poltrona”. Si tratta quindi di un caso in cui la collaborazione tra neuroscienziati e filosofi potrebbe essere fruttuosa, come vedremo ulteriormente in seguito.

Tuttavia questo ancora non dimostra che esista una vera e propria filosofia delle neuroscienze, per due ragioni non del tutto indipendenti l’una dall’altra. La prima è che, tipicamente, una filosofia di una scienza X ha tra i suoi scopi principali quello di indagare i fondamenti filosofici della scienza X, ma non mi pare di scorgere l’esistenza di una riflessione di questo genere nel caso delle neuroscienze. Le questioni fondazionali discusse in discipline attigue come la filosofia della psicologia o

⁴ Cfr. ad es. Friedemann PULVERMÜLLER, *How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics*, “Trends in Cognitive Sciences”, Volume 17, n. 9, anno 2013, pp. 458-470.

la filosofia della scienza cognitiva⁵ lambiscono soltanto le neuroscienze. Queste discipline si occupano infatti di cervello solo marginalmente o solo sotto quegli aspetti che hanno rilievo per lo studio della mente. La seconda ragione è che una filosofia di una scienza X si costituisce nel momento in cui emergono con chiarezza certi problemi filosofici canonici legati all'ambito in questione. È possibile, forse anche probabile, che questo accada; ad esempio, un buon candidato al titolo di problema canonico della filosofia delle neuroscienze è quello dell'“inferenza inversa”, che riguarda la legittimità di trarre certe conclusioni relative a quale processo cognitivo realizza un certo compito (dunque conclusioni che riguardano la *mente*) dalla conoscenza delle aree cerebrali coinvolte da un lato nell'esecuzione di tale compito, dall'altro nella realizzazione del processo in questione (torneremo più avanti sull'inferenza inversa). Al momento, tuttavia, non abbiamo una massa critica di problemi del genere; non si sente il bisogno di parlare di una filosofia delle neuroscienze in aggiunta alla categoria già solidamente attestata di filosofia della scienza cognitiva. Non vorrei d'altra parte dare l'impressione di attribuire un'importanza eccessiva alla domanda circa l'esistenza di una filosofia delle neuroscienze. Infatti, al di là delle etichette disciplinari, che notoriamente rispondono anche a logiche di “marketing” della produzione scientifica, quello che conta è che indubbiamente esiste un ambito di studi, come che lo si voglia chiamare esattamente, nel quale la filosofia –una certa filosofia– e le neuroscienze possono cooperare fruttuosamente; nel quale il filosofo può suggerire allo scienziato ipotesi sperimentali e interpretazioni di dati e lo scienziato può mostrare al filosofo che certe concettualizzazioni risultano inappropriate alla luce delle osservazioni.

Detto questo, non dobbiamo dimenticare che ci sono domande che hanno un genuino interesse filosofico e domande che non ce l'hanno⁶. Chiedersi quale area

⁵ La scienza cognitiva è lo studio interdisciplinare dei fenomeni mentali incentrato sull'idea che i processi mentali sono processi di elaborazione di informazioni. Si vedano ad esempio: Diego MARCONI, *Filosofia e scienza cognitiva*, Laterza, Roma-Bari 2003; Massimo MARRAFFA, Alfredo PATERNOSTER (a cura di), *Scienze cognitive. Un'introduzione filosofica*, Carocci, Roma 2011; Edoardo DATTERI, *Che cos'è la scienza cognitiva*, Carocci, Roma 2017.

⁶ Mi arrischio a fare questa affermazione a dispetto della grande difficoltà di spiegare di che cosa si occupa la filosofia, di individuare quali siano i tratti caratteristici di un problema specificamente filosofico. Difficoltà che si rispecchia nell'assenza di un consenso ampio su quali siano i compiti della filosofia. Penso, e in qualche misura dovrebbe emergere in questo articolo, che ci siano problemi specificamente filosofici, problemi specificamente scientifici, e problemi all'intersezione,

cerebrale realizza un certo compito cognitivo di per sé non è filosoficamente molto interessante, sebbene informazioni di questo genere possano poi svolgere un ruolo in contesti teorici più ampi e in qualche senso filosofici; ma chiedersi, ad esempio, se e in che misura nel cervello le operazioni percettive e le operazioni concettuali siano separate spazialmente e, pur sulla scala di pochi millisecondi, temporalmente è una questione di grande interesse filosofico, perché ciò comporta conseguenze sulla natura della percezione e dei concetti e su come dobbiamo intendere la relazione tra percezione e cognizione. Ho deliberatamente scelto questi due esempi per far vedere come due fatti indubbiamente neurologici, di natura non troppo diversa, abbiano un ben diverso peso specifico dal punto di vista filosofico.

Ho appena sostenuto che solo *alcuni* aspetti relativi al funzionamento del cervello hanno rilievo filosofico. Ma che cosa rende *filosofico* un problema? Che cosa rende un fatto neurobiologico filosoficamente interessante? In prima istanza e in termini molto generali, penso che “neuro-fatti” e “neuro-questioni” diventino filosoficamente pregnanti quando fanno una differenza per la *mente*, per la nostra comprensione dei fenomeni mentali⁷. Fenomeni come la percezione, il ragionamento, la pianificazione dell’azione, la comprensione del linguaggio, la metacognizione, la coscienza, l’identità personale, l’io. Senza dimenticare che *il* problema fondamentale della filosofia della mente è proprio quello della relazione tra mente e cervello o più in generale tra mente e mondo fisico. Che lo si voglia porre in termini strettamente metafisici, come problema della relazione tra proprietà mentali e proprietà neurobiologiche, piuttosto che caratterizzarlo in termini epistemologici, ossia come problema di quale tipo di relazione sussista tra teorie che si collocano a un livello esplicativo differente, esso resta il nocciolo duro alla radice di tutte le altre questioni. Per quanto ci è dato finora di scorgere, questo problema non è nemmeno scalfito in superficie dalle conoscenze che abbiamo maturato sul cervello, e in non pochi “circoli filosofici” è diffuso il sospetto che nemmeno un progresso assai significativo in tali conoscenze possa spostare le cose. Nondimeno, mi sembra difficile negare che, se qualche progresso verso la soluzione di questo problema è possibile, esso non potrà che venire –pur

per i quali non ha molto senso chiedersi se siano filosofici piuttosto che scientifici. Sono proprio questi ultimi i problemi che ci interessano qui.

⁷ Questa affermazione è giustificata dalla considerazione che la questione della natura della mente e della conoscenza è sempre stata al centro dell’agenda dei filosofi, come osserva ad esempio José L. BERMÚDEZ nel suo *Philosophy of Psychology*, Routledge, London 2005.

rendendomi conto che vi è qui una certa qual aria di circolarità– dallo studio di che cosa accade nel cervello quando siamo impegnati in certe attività mentali e dunque dal perseguimento della *neuroscienza cognitiva*, la disciplina che indaga il modo in cui i processi cerebrali realizzano i processi psichici in base ai principi metodologici dell'*integrazione* e della *coevoluzione* di psicologia e neuroscienza⁸.

Posto dunque che, in quanto filosofi, siamo interessati ai fatti neurobiologici solo quando questi intersecano il piano mentale, la domanda che ci dobbiamo porre sarà: a quali condizioni un fatto neurobiologico fa una differenza per la mente?

Un po' sorprendentemente, qualche scienziato ha espresso scetticismo in generale sulla possibilità di istituire connessioni chiare tra fatti neurobiologici e fatti mentali. Il neuropsicologo Max Coltheart, ad esempio, ha sostenuto che gli esperimenti di neuroimmagine (tipicamente basati sulla risonanza magnetica funzionale: fMRI) non sono in grado di decidere quale tra due teorie psicologiche di un certo processo cognitivo è quella giusta⁹. Coltheart fa diversi esempi, traendone la morale generale che non c'è alcuna corrispondenza sistematica tra un compito cognitivo e un certo sottosistema cerebrale sufficientemente ben definito. In assenza di tale corrispondenza, la domanda di come il cervello realizza i processi e le operazioni mentali è destinata a restare senza risposta.

Se è vero che il bersaglio polemico di Coltheart è ristretto agli esperimenti di neuroimmagine, non ai metodi della neuroscienza cognitiva in generale, si tratta nondimeno di una critica devastante, perché da oltre vent'anni a questa parte il *neuroimaging* costituisce la fonte primaria di dati in neuroscienza cognitiva. D'altra parte la tesi di Coltheart appare troppo radicale, ed è probabilmente da intendersi come un'utile provocazione contro –come non notarla?– una fastidiosa propensione a credere (quanto in buona fede?) che collezionando sempre più fMRI arriveremo a dire qualcosa di risolutivo sui fatti mentali. Che le cose non stiano così è comprovato dal fatto che molto spesso le ricerche dei neuroscienziati, e specificamente quelle

⁸ Cfr. ad es. William BECHTEL, *Cognitive neuroscience: Relating neural mechanisms and cognition*, in Peter K. MACHAMER, Peter MCLAUGHLIN e Rick GRUSH (a cura di), *Philosophical Reflections on the Methods of Neuroscience*, University of Pittsburgh Press, Pittsburgh (PA) 2006, pp. 81-111; Massimo MARRAFFA, Alfredo PATERNOSTER, *Persone, menti, cervelli*, Mondadori Education, Milano 2012, cap. 3.

⁹ Max COLTHEART, *What has functional neuroimaging told us about the mind (so far)?*, "Cortex", vol. 42, anno 2006, pp. 323-331.

basate sulla fMRI, sono volte a rispondere alla domanda: quale area del cervello realizza la tale funzione? Ebbene, al di là dell'imprecisione con cui si può attualmente rispondere a questa domanda, dovuta a limiti tecnici di risoluzione che sicuramente saranno via via rimossi, il punto è che la risposta a questa domanda non ci dice nulla o quasi nulla di interessante sulla funzione mentale¹⁰. Quello a cui un filosofo è interessato è semmai qualcosa di simile al *come* viene realizzata la tale funzione mentale, o a che rapporto ci sia tra quella funzione e altre capacità cognitive. Domande nelle quali il cervello c'entra eccome, ma più come condizione estrinseca. L'idea, espressa in modo un po' spiccio, è che per capire la mente dobbiamo, oltre a guardare al cervello, avere buoni modelli di come funziona la mente stessa. L'integrazione/coevoluzione di psicologia e neuroscienza cui alludevo poc'anzi è una buona strategia di ricerca ma, per gli scopi della filosofia, la neuroscienza resta ancillare rispetto alla psicologia. O almeno, è la psicologia a guidare la neuroscienza più di quanto la neuroscienza guidi la psicologia. In caso contrario, avremo scoperte interessantissime sul cervello che però ci dicono poco sui problemi che ci interessano filosoficamente.

Come dicevo, tuttavia, la critica di Coltheart è eccessivamente radicale. Se è vero che la neuroimmagine non riesce a discriminare tra teorie in competizione, è tuttavia altrettanto vero che ci ha permesso di tracciare macro-mappe funzionali, di avere un quadro, per quanto grossolano, dell'organizzazione cerebrale, che soltanto qualche decennio prima sarebbe stato impensabile. Sapere che cosa si attiva quando è un'informazione fondamentale per la già citata inferenza inversa, la pratica argomentativa più diffusa, sebbene spesso implicita, in neuroscienza cognitiva. L'inferenza inversa consiste nel seguente schema argomentativo¹¹:

(P1) Nello studio in oggetto l'area *A* si attiva durante l'esecuzione del compito *C*;

(P2) In altri studi (indipendenti dal precedente e possibilmente relativi a compiti diversi da *C*) si è osservata l'attivazione dell'area *A* contestualmente all'esecuzione del processo cognitivo *P*;

¹⁰ Paolo LEGRENZI, Carlo UMLTA, *Neuromania*, il Mulino, Bologna 2009.

¹¹ Cfr. ad es. Russell A. POLDRACK, *Can cognitive processes be inferred from neuroimaging data?*, "Trends in Cognitive Science", vol. 10, n. 2, anno 2006, pp. 59-63; Guillermo DEL PINAL, Marco J. NATHAN, *The Future of Cognitive Neuroscience: Reverse Inference in Focus*, "Philosophy Compass", volume 12, n. 7, anno 2017, pp. 1-19.

→(quindi) (C) Il compito C viene realizzato tramite il processo P (ovvero, da un punto di vista leggermente diverso: P è l'ingrediente fondamentale di una teoria di C).

Ad esempio (semplificando un po'), l'area premotoria si attiva quando i soggetti sono impegnati in un compito di riconoscimento d'intenzione (P1); in diversi altri esperimenti si è riscontrata attivazione dell'area premotoria quando è in atto un processo di simulazione dell'azione di afferramento di un oggetto (P2); conclusione: le intenzioni soggiacenti alle azioni vengono riconosciute simulando le azioni in questione.

È plausibile affermare che l'inferenza inversa è la pratica argomentativa standard in neuroscienza cognitiva, in quanto definiamo come scopo di quest'ultima la spiegazione delle capacità *mentali* sulla base di dati sul cervello: nell'inferenza inversa si deduce una conclusione sulla mente (riguardante fatti psicologici) a partire da osservazioni correlate su cervello e fenomeni mentali. Ciò non toglie che una combinazione di questo schema inferenziale con il suo complementare, l'inferenza diretta, nella quale la prima premessa e la conclusione sono scambiate (= si deduce che l'area A è investita del compito C a partire dalla premessa che il compito C è realizzato dal processo P e da una premessa analoga a P2, in cui area e processo sono correlati), possa costituire un modo proficuo di fare neuroscienza cognitiva, tramite un reciproco rinforzo delle premesse: la conclusione di un'inferenza diretta può diventare una premessa di un'inferenza inversa e viceversa.

Secondo alcuni autori l'inferenza inversa è la risposta al problema messo in luce da Coltheart, consentendo di decidere quale tra due o più teorie psicologiche in competizione è quella giusta. Infatti, posto che ciò che distingue due teorie psicologiche sono le diverse predizioni riguardo al tipo di processi cognitivi coinvolti, se si hanno due processi cognitivi candidati per il compito C (siano essi $P1$ e $P2$) e si dà il caso che, durante l'esecuzione di un altro compito diverso da C , l'area A è attivata quando viene eseguito $P1$ ma non quando viene eseguito $P2$, sembra seguirne che è $P1$ a realizzare C . Questa valutazione è tuttavia eccessivamente ottimistica: ci sono alcune difficoltà non da poco che impediscono di saltare a piè pari alla conclusione. A dispetto dell'apparenza di validità deduttiva dell'inferenza inversa, la sua affidabilità è limitata considerevolmente da un problema: la *multifunzionalità* dell'area cerebrale. Che cosa vuole dire? Vuole dire che l'inferenza inversa

metterebbe capo a conclusioni affidabilmente vere soltanto se l'area cerebrale in questione (A) fosse selettiva, cioè dedicata a un solo processo. Ma questa condizione non capita quasi mai: tipicamente un'area cerebrale è coinvolta in differenti processi/funzioni e una funzione/processo coinvolge diverse aree. Lo scetticismo di Coltheart sopra descritto deriva essenzialmente da questo.

Sono state avanzate alcune proposte per rendere più affidabile l'inferenza inversa. Non è il caso di illustrarle, perché ciò richiederebbe un livello di "tecnicità" fuori luogo in questa sede. Ma l'idea generale si può esprimere, spero, con accettabile chiarezza: non è l'area cerebrale il dato su cui dobbiamo basarci, bensì la forma del pattern di attivazione del circuito cerebrale coinvolto in un processo, ovverosia la "geometria" della configurazione costituita dall'insieme dei neuroni che scaricano quando ha luogo un dato processo cognitivo; e i pattern di attivazione sono associabili selettivamente e affidabilmente ai processi cognitivi perché disponiamo di algoritmi sofisticati capaci di "leggere" direttamente i pattern nelle neuroimmagini. Tuttavia non mi sento di azzardare valutazioni sulle prospettive di questo metodo. Quello che a noi interessa è che la questione dell'inferenza da fatti neurali a fatti psichici può essere vista come un classico problema di filosofia della scienza di natura metodologica. In quanto tale, è un caso in cui i filosofi possono aiutare i neuroscienziati a meglio concepire il lavoro che stanno facendo. Il problema, filosofico non meno che scientifico, è quale tipo di ragionamento a partire da premesse empiriche che collegano fatti neurali a fatti psichici ci consente di derivare conclusioni affidabili relative all'organizzazione della mente.

Il secondo esempio concreto di interazione neurofilosofica può essere proficuamente introdotto partendo da una possibile obiezione al quadro fin qui proposto. Sopra ho affermato che i fatti neurobiologici sono filosoficamente pertinenti solo se fanno una differenza per la mente; ma non tutte le differenze sono uguali. Vale a dire: non tutti i fatti psichici sono filosoficamente interessanti. Dopotutto, così come non consideriamo filosoficamente rilevante conoscere quale area è preposta a un certo processo cognitivo, perché dovremmo considerare filosoficamente interessante sapere che la percezione o la comprensione del linguaggio sono realizzate da certi processi psichici? Il filosofo che si interroga sulla natura della percezione o della comprensione non sarà illuminato dallo scoprire che, ad esempio,

il processo percettivo che dà luogo alla nostra esperienza visiva ordinaria è articolato in più stadi computazionali, il primo dei quali è volto a rilevare gli sbalzi di luminanza presenti nell'immagine retinica. A questo riguardo, è il caso di ricordare come la critica di Wittgenstein all'idea che i meccanismi o processi del comprendere siano pertinenti per il concetto ordinario nonché filosofico di comprensione colpisce egualmente i processi psichici e quelli cerebrali (si vedano in particolare i paragrafi §148, §151, §158 delle *Ricerche filosofiche*).

Quando ho assunto il punto di vista continuista ho implicitamente espresso il mio dissenso verso questo genere di obiezioni. Il dominio della percezione visiva è uno di quelli che meglio ne evidenzia l'infondatezza. Basti pensare a come un grande psicologo sperimentale, l'inventore del paradigma ecologista James J. Gibson, abbia elaborato una teoria psicologica della percezione che si opponeva allo stesso tempo al realismo indiretto dei filosofi e al costruttivismo degli psicologi. Nel dominio della percezione distinguere tra una teoria filosofica ed una psicologica è difficile, al di là degli aspetti idiosincratici dei rispettivi linguaggi. La discussione del caso seguente offrirà un contributo ulteriore alla posizione continuista, fornendo al tempo stesso un secondo esempio di integrazione tra filosofia e neuroscienza.

Lo scopo ultimo della neuroscienza cognitiva potrebbe essere descritto come il perseguimento di una mappa psico-neurale completa, ossia una descrizione compiuta, esaustiva e ad alta risoluzione di quale funzione mentale viene eseguita da ciascuna area del cervello, una descrizione che dia accuratamente conto anche della multifunzionalità di cui si parlava sopra. Tuttavia, come osserva Ned Block¹², persino mappe ad alta risoluzione del cervello, cioè descrizioni delle attività di ogni *singolo* neurone (!), sarebbero di scarsa utilità ai fini della spiegazione di come funziona la mente; e nemmeno disporre di un'enorme quantità di dati sul cervello sarà utile *in assenza di un quadro concettuale nel quale fatti psicologici e fatti neurobiologici siano collegati in modo chiaro*. E questo rende evidente il fatto che i filosofi possano ritagliarsi un ruolo: il quadro concettuale cui Block allude ancora non c'è –o ce ne sono di alternativi ciascuno dei quali ha pregi e difetti– e la messa a punto di un quadro concettuale continua ad essere pane per i denti dei filosofi.

¹² Ned BLOCK, *Consciousness, Big Science and Conceptual Clarity*, in Jeremy FREEMAN, Gary MARCUS (eds.), *The Future of the Brain*, Princeton University Press, Princeton (NJ) 2014.

Più specificamente, l'esempio che mi accingo a discutere evidenzia molto bene come una persistente confusione concettuale renda di difficilissima lettura risultati empiricamente molto interessanti. Non sappiamo bene come ricomprendere questi dati in un quadro teorico chiaro e coerente. Il problema che ci si pone è quello di quali condizioni neurobiologiche (più specificamente neurofisiologiche) sono richieste affinché uno stimolo venga percepito coscientemente. In prima battuta si potrebbe procedere in questo modo (metodo "sottrattivo" o "contrastivo"): scegliere una serie di stimoli che sono ai limiti della soglia di coscienza; registrare i pattern cerebrali che hanno luogo quando lo stimolo è percepito consciamente e quelli che hanno luogo quando lo stimolo non è percepito (ovvero è percepito solo subliminalmente); "sottrarre" i secondi dai primi: quel che resta dovrebbe essere l'ingrediente neuronale della coscienza.

Il guaio è, prosegue Block, che per determinare la condizione in cui il percetto è conscio oppure no dobbiamo chiedere al soggetto se ha percepito lo stimolo¹³, e questa richiesta innesca una serie di processi cognitivi (comprensione della richiesta, decisione di rispondere, focalizzazione dell'attenzione sullo stimolo ecc.) che interferiscono con ciò che stiamo cercando: noi vogliamo sapere che cosa capita nel cervello quando il percetto viene presentato al soggetto e questi ne è conscio; ma non si riesce a isolare questa condizione perché i criteri standard sui quali ci basiamo per attribuire al soggetto uno stato cosciente richiedono il coinvolgimento di processi cognitivi che, si direbbe, vanno al di là della coscienza dello stimolo.

Questa situazione rende, se non proprio impossibile, difficilissimo individuare il correlato neurale della coscienza, rendendo quest'ultimo un concetto empiricamente vuoto. Ora, questo problema potrebbe essere superato se, a differenza di quanto pensa Block, uno stimolo percettivo diventasse cosciente solo quando il soggetto ha un accesso a tale stimolo, vale a dire se egli sfrutta in un processo, quale è ad esempio il riferire lo stato di coscienza, l'informazione relativa allo stimolo. È chiaro infatti che in questa prospettiva ciò che prima veniva giudicato un fattore interferente diviene ora parte della coscienza stessa. Questa è ad esempio la posizione di Stanislas Dehaene, neuroscienziato di primissimo piano, che, sulla base di questa idea di coscienza, ritiene affidabili le procedure sperimentali basate sul metodo sottrattivo. Diviene in tal modo possibile individuare le condizioni (o correlati) neurali della

¹³ O eventualmente escogitare un altro indicatore comportamentale.

coscienza; nelle parole di Dehaene, siamo del tutto legittimati a parlare di “firme della coscienza”¹⁴. Quando uno o più di tali eventi cerebrali ha luogo (tipicamente se ce n'è uno, ci sono tutti gli altri), il soggetto è cosciente di qualche stimolo. Secondo Dehaene le firme della coscienza sono (i) l'attivazione repentina dei circuiti parietali e frontali, (ii) la presenza all'EEG di un'onda lenta ritardata (la cosiddetta P3), (iii) un picco ritardato ed improvviso di oscillazioni ad alta frequenza e (iv) la sincronizzazione di aree distanti, cioè l'oscillazione alla stessa frequenza di gruppi neuronali distanti.

Dehaene ha individuato le firme della coscienza andando a vedere che cosa capitava nel cervello in una serie di situazioni in cui i soggetti affermano di essere coscienti o comunque esibiscono comportamenti inequivocabilmente mediati da stati coscienti. Ciò presuppone quindi, lo ripetiamo, un'idea di coscienza diversa da quella di Block: se per quest'ultimo la coscienza è fondamentalmente il mero apparire di qualcosa a un soggetto o organismo, del tutto indipendentemente da come il soggetto sfrutta l'informazione contenuta in questa apparenza, per Dehaene la coscienza sale alla ribalta solo quando quell'informazione è disponibile ai processi cognitivi “di alto livello”. Nella felice immagine del filosofo Daniel C. Dennett, che condivide e ha contribuito a ispirare il modello di Dehaene, la coscienza è la “fama nel cervello”, intendendo con questo che quanto più un'informazione è condivisa da diversi processi, tanto più ne avremo coscienza.

A questa differenza corrisponde, sul piano cerebrale, una diversa ipotesi. Per il primo modello, quello di Block, l'attivazione delle aree visive è (o almeno può essere) condizione sufficiente per il prodursi di coscienza, laddove per il secondo modello, quello di Dehaene, la coscienza si manifesta soltanto quando l'informazione sensoriale raggiunge, tramite connessioni a lunga distanza, le aree fronto-parietali.

Qual è la morale principale di questa storia dal punto di vista di questo articolo? Che c'è un'interdipendenza tra ciò che noi definiamo coscienza, i criteri linguistico-comportamentali per l'attribuzione di coscienza e i metodi sperimentali di cui ci serviamo per individuare i correlati neurali, le firme cerebrali della coscienza. Se non ci mettiamo d'accordo su che cosa è la coscienza, i risultati sperimentali non

¹⁴ Stanislas DEHAENE, *Consciousness and the Brain. Deciphering how the Brain Codes our Thoughts*, Viking Press, New York 2014; tr. it. Pier Luigi GASPA, *Coscienza e cervello. Come i neuroni codificano il pensiero*, Raffaello Cortina, Milano 2017.

potranno funzionare come tribunale delle teorie; ma, d'altra parte, la definizione della coscienza non può essere fatta completamente a tavolino; i moltissimi risultati sperimentali relativi, ad esempio, ai disturbi della coscienza o ai più drammatici casi degli stati vegetativi, sono una fonte di dati inestimabile, che ci obbliga ad affinare gli strumenti concettuali di cui ci serviamo per spiegare quale sia la natura della coscienza. Sindromi sorprendenti, come la “vista cieca” (*blindsight*) o l'eminegligenza spaziale unilaterale (*neglect*), –ma ve ne sono diverse altre– danno luogo a situazioni alle quali i nostri concetti ordinari non si applicano, costringendoci a introdurre nuovi concetti o almeno a ripensare quelli esistenti. Nel *blindsight*, ad esempio, i soggetti sono capaci di eseguire compiti visivi con risultati abbastanza buoni pur non avendo alcuna coscienza dichiarata degli stimoli ad essi proposti. Come dobbiamo descrivere questa situazione? Come un caso di dissociazione tra due tipi di coscienza (c'è una coscienza cognitiva, o d'accesso, ma non una coscienza fenomenica)? Come un caso di processamento del tutto inconscio di informazione? Impossibile leggere i fatti neurologici indipendentemente da una teoria di sfondo; impossibile dare per scontata l'una o l'altra interpretazione dei dati sperimentali. In casi come questo, clinica, neuropsicologia, neuroimmagini, EEG e ancora altre fonti di dati pongono vincoli ben precisi a una caratterizzazione ragionevole della coscienza, senza, d'altra parte, essere in grado di dare risposte definitive riguardo a ciò che i dati mostrano.

Non mi resta che tentare di abbozzare qualche conclusione.

Attraverso la discussione di un paio di casi ho fatto vedere in che modo filosofia e neuroscienze possano collaborare proficuamente. Non ho discusso invece l'altro grande tema, quello delle conseguenze etiche delle scoperte delle neuroscienze, tema sicuramente assai dibattuto, ma, almeno a mio giudizio, ascrivibile totalmente all'ambito della filosofia morale.

Ho cominciato questo articolo introducendo la nozione di continuismo relativamente al rapporto tra scienza e filosofia (naturalismo filosofico). Quasi tutto quello che ho detto in questo articolo presuppone il continuismo; ma sposare il continuismo non è obbligatorio, sebbene sia oggi la posizione significativamente maggioritaria, o almeno questa è la mia impressione. Anche una volta preso per buono, il continuismo lascia tuttavia totalmente liberi riguardo a come realizzarlo nella pratica del filosofo di professione. Il contributo della filosofia alla scienza cognitiva e specificamente alle neuroscienze non è chiaramente riconducibile a

schemi prestabiliti, né dopotutto deve esserlo. Nella pratica filosofica effettiva, l'analisi dei concetti praticata essenzialmente a priori (= mi lascio guidare dal mio discernimento filosofico anche nella disamina di nozioni sulle quali la scienza ha autorità), le revisioni metafisiche ispirate alle ontologie scientifiche (= correggo l'ontologia del senso comune a partire dalla scienza), le cosiddette ricostruzioni razionali delle teorie scientifiche e della loro evoluzione (= una sorta di sintesi filosofica che mira a cogliere quello che è ritenuto l'essenziale dal punto di vista teorico di quella posizione) coesistono sfumando l'una nell'altra¹⁵. Insomma il contributo della filosofia alle scienze della mente, nel nostro caso alle neuroscienze, è vitale e fecondo facendo un po' tutte queste cose. Ed è vitale e fecondo anche se può succedere che i filosofi propongano tesi erranee, esattamente come avviene per le ipotesi scientifiche¹⁶.

Non sorprendentemente, i neuroscienziati tendono a pensare che i fatti mentali siano *completamente* determinati da fatti neurobiologici. Anche supponendo che sia così, la complessità di questa dipendenza rende problematico dedurre teorie eliminazioniste, secondo le quali l'ontologia mentale del senso comune è del tutto priva di fondamento, o riduzioniste, secondo le quali i processi mentali non sono altro che processi cerebrali visti da una particolare ma limitata prospettiva, la nostra in quanto soggetti di esperienza. In realtà, ciò che le neuroscienze hanno scoperto sul cervello (relativamente molto) e sulla relazione tra mente e cervello (poco) lascia aperte molte opzioni metafisiche alternative; ed è anche per questa ragione che la metafisica della mente resta in diversi casi un ambito abbastanza indipendente. Qualcosa di simile può essere detto sul piano epistemologico: come osserva Michele Di Francesco in un recente articolo¹⁷, uno dei rischi dell'enfasi oggi posta sul cervello nelle spiegazioni dei fenomeni mentali è quello di promuovere una forma di monismo esplicativo che riduce le spiegazioni di livello personale e le spiegazioni di livello psicologico del comportamento a spiegazioni esclusivamente neuroscientifiche. Come in qualche misura anche i casi descritti qui dovrebbero aver mostrato, questo genere di riduzionismo è ben lontano dall'aver trovato una sua legittimazione in ciò che abbiamo fin qui scoperto. Al contrario, la complessità dei fenomeni in oggetto e

¹⁵ Cfr. Alfredo PATERNOSTER, *La ricostruzione filosofica della scienza cognitiva: una lente deformante?*, "Sistemi Intelligenti", vol. XXIII, n. 1, anno 2011, pp. 7-23 (spec. §3).

¹⁶ *Ivi*, p. 21.

¹⁷ DI FRANCESCO, *Spiegazione neurocentrica e mente personale. Tra filosofia e neuroscienze*.

forse la loro stessa natura suggeriscono l'opportunità di «una pluralità di livelli di indagine – biologico, neurale, psicologico, antropologico, sociale e filosofico – caratterizzati da stili esplicativi profondamente diversi, che possono proficuamente interagire»¹⁸.

Sembra dunque che ci sia ancora molto lavoro per il filosofo che sia disponibile alla fatica di cimentarsi, con tutta la prudenza e l'umiltà del caso, con una letteratura difficile e specialistica. E che parte di questo lavoro consiste nel discriminare pazientemente quello che è filosoficamente importante da quello che non lo è.

¹⁸ *Ivi*, p. 245.