

I LINGUAGGI DELLE
SCIENZE COGNITIVE

SANDRO NANNINI

Mente e cervello

La seconda svolta cognitiva in filosofia della mente e altri saggi



I LINGUAGGI DELLE SCIENZE COGNITIVE

I LINGUAGGI DELLE SCIENZE COGNITIVE

© 2020 .. Corisco Edizioni . Marchio Editoriale ..
Roma-Messina

Proprietà artistica e letteraria riservata.
È vietata qualsiasi riproduzione totale o parziale ai sensi
della L. N. 633 del 22/04/1941, L. N. 159 del 22/05/1993,
L. N. 248 del 18/08/00 e successive modificazioni.

ISBN: 978-88-898-138-364

a cura di

Sandro Nannini

(corisco)

Sandro Nannini

**Mente e cervello:
La seconda svolta cognitiva
in filosofia della mente
e altri saggi**

Indice

Introduzione.....	7
-------------------	---

Ringraziamenti	19
----------------------	----

Capitolo primo

La seconda svolta cognitiva in filosofia della mente	23
1. La svolta linguistica in filosofia della mente.....	23
2. La prima svolta cognitiva	25
3. La seconda svolta cognitiva	29
4. Riduzionismo e anti-riduzionismo dopo la seconda svolta cognitiva	31

Capitolo secondo

Il problema mente-corpo tra filosofia e neuroscienze	35
1. Filosofia e neuroscienze.....	35
2. Un esempio di azione volontaria.....	37
3. Il dualismo ontologico interazionistico	38
4. Il dualismo parallelistico e l'epifenomenismo	41
5. Il fisicalismo	42
6. Il funzionalismo e altre alternative al dualismo e al fisicalismo.....	45
7. L'emergentismo	47
8. Intenzionalità, autocoscienza e coscienza	50
9. Le rappresentazioni mentali	52
10. La riduzione ontologica dell'Intenzionalità	56
11. La riduzione ontologica della coscienza e dell'autocoscienza.....	64

Capitolo terzo

Mente, corpo, società e ontologia del mentale	73
1. Riduzionismo e antiriduzionismo	73
2. Verso un'ontologia filosofica di tipo naturalistico.....	74
3. Ordini di discorso e livelli d'analisi	77
4. Fatti socio-culturali, fenomeni mentali e processi cerebrali.....	83

Capitolo quarto

Tempo e coscienza	89
1. Lo Hard Problem e l'eliminativismo	89
2. Tempo, spazio e gravità nella teoria della relatività	90
3. Coscienza e gravità	95
4. Tempo, coscienza e sincronizzazioni cerebrali	99

Capitolo quinto

Realismo scientifico e ontologia del mentale	105
1. Il realismo scientifico ed il problema mente-corpo	105
2. Obiezioni e risposte sul realismo scientifico-cognitivo	108

Capitolo sesto

Il concetto di verità in una prospettiva naturalistica	113
1. La naturalizzazione dell'Intenzionalità e il concetto di verità.....	113
2. Rappresentazioni mentali.....	115
3. Concezioni della verità.	119
4. Conclusione.....	138

Capitolo settimo

Kant e le scienze cognitive sulla natura dell'io	141
1. Coscienza e autocoscienza da Cartesio a Hume.....	141
2. Il Self oggi tra filosofia e scienze cognitive	142
3. L' "io-penso" di Kant	148
4. L'io come Direttore Assente	153

Capitolo ottavo

Freud e il naturalismo cognitivo	157
1. L'inconscio freudiano e l'Intenzionalità brentaniana.....	157
2. Coscienza e Intenzionalità in Cartesio e F. Brentano	158
3. Intenzionalità e naturalismo in Freud.....	161
4. Freud e l'odierna filosofia della mente	164
5. Psicoanalisi e naturalismo cognitivo oggi.....	166

Riferimenti	169
-------------------	-----

Introduzione

Raccolgo qui dei saggi di filosofia della mente scritti tra il 2006 ed il 2018, quasi tutti profondamente rielaborati ed alcuni anche fusi insieme. Il filo rosso che li unisce è il mio tentativo di sviluppare e approfondire quella soluzione naturalistico-fisicalistica del problema mente-corpo già abbozzata in *Naturalismo cognitivo: per una teoria materialistica della mente* (Nannini 2007a).

Il primo capitolo ricostruisce per sommi capi la storia della filosofia della mente – ed in particolare il dibattito intorno al problema mente-corpo – dal secondo dopoguerra ad oggi. Negli anni Cinquanta del secolo scorso, sulla scia della «svolta linguistica» (Rorty 1967) impressa alla filosofia per un verso dall'empirismo logico e per un altro dall'ultimo Wittgenstein, la filosofia della mente diviene essenzialmente un capitolo della filosofia del linguaggio. In particolare per Ryle (1949) e con lui per tutti i comportamentisti filosofici non ci si deve più porre la domanda metafisica di che cosa sia la mente e che cosa la distingua dal corpo a cui è legata, bensì ci si deve chiedere quale sia l'uso corretto, nel linguaggio ordinario, della parola “mente” e di tutti gli altri termini psicologici. Attraverso questa analisi filosofica del linguaggio mediante il quale si esprime la *folk psychology* sarà possibile liberarsi del dualismo cartesiano e, con esso, di tutti gli pseudo-problemi metafisici legati al concetto di anima. Gli stati mentali vengono intesi nella prospettiva del comportamentismo filosofico non più come qualcosa di “interno” e soggettivo, accessibile solo introspettivamente al soggetto interessato, bensì come disposizioni comportamentali pubblicamente osservabili.

Questo approccio comportamentistico al mentale viene criticato e in larga parte abbandonato negli anni Sessanta e Settanta del Novecento ad opera dei sostenitori sia del fisicalismo (o “teoria dell'identità” tra mente e cervello che dir si voglia) sia del funzionalismo (o cognitivismo). Né i fisicalisti né i funzionalisti vogliono però ritornare al dualismo cartesiano. Ritengono piuttosto che l'essere anti-cartesiani non implichi l'accettazione del comportamentismo. Distinguendo il dualismo dal mentalismo, essi pensano sì, in polemica con i comportamentisti, che gli stati mentali siano delle cause interne del comportamento, ma ritengono anche che tali cause vadano identificate

non, come vogliono i dualisti, con presunti misteriosissimi eventi spirituali, bensì con dei normali processi cerebrali descrivibili o direttamente in termini neurologici (fisicalismo) o, più astrattamente, in termini funzionali, paragonandoli ad un *software* implementato da un'attività cerebrale difficilmente ricostruibile in modo diretto (funzionalismo e cognitivismo). In entrambi i casi la filosofia della mente si pone in un rapporto di continuità con le nascenti scienze cognitive ed in particolare con l'intelligenza artificiale. Da qui anche quella "analogia mente-computer" che è la concezione del mentale di maggior successo offerta da funzionalisti e cognitivisti negli anni Sessanta e Settanta, gli anni appunto di quella che può essere chiamata la "prima svolta cognitiva" in filosofia della mente.

A partire dagli anni Ottanta il funzionalismo viene aspramente criticato da più parti. Non solo si assiste ad un ritorno del dualismo cartesiano, ma più in generale viene ridata dignità scientifica a quello studio della coscienza che era stato rifiutato dai comportamentisti e messo sostanzialmente in parentesi anche dai funzionalisti. Occorre però distinguere entro questo grande alveo del ritorno allo studio della coscienza l'approccio di chi ripropone in sostanza, sia pur in modo aggiornato e sensibile alla lezione delle scienze cognitive, un ritorno al dualismo cartesiano o alla fenomenologia di E. Husserl dal punto di vista di chi, filosofo o neuroscienziato che sia, intenda invece rinnovare completamente la concezione tradizionale del mentale alla luce dei nuovi dati sperimentali e delle nuove teorie sul cervello offerte dalle neuroscienze (soprattutto a partire dagli anni Novanta). Questo nuovo approccio al mentale, a cui si può dare il nome di "seconda svolta cognitiva", è alla base oggi della maggior parte degli studi condotti sui processi cognitivi e sulla coscienza sia da molti neuroscienziati sia da alcuni filosofi. Tra questi ultimi occorre distinguere inoltre coloro che cercano di sviluppare quella forma aggiornata di fisicalismo che essi stessi hanno chiamata "eliminativismo" (si pensi ai coniugi Churchland) da coloro che cercano di elaborare una nuova concezione della mente appoggiandosi al neo-darwinismo (si pensi soprattutto a D.C. Dennett).

Detto in breve, questa succinta ricostruzione storico-filosofica offerta nel primo capitolo intende mostrare che la seconda svolta cognitiva è alla base della parte più feconda degli studi condotti oggi intorno al mentale. Tutti gli altri capitoli, scritti con un taglio più

marcatamente filosofico-sistematico, vogliono approfondire singoli aspetti di tale svolta.

Il secondo capitolo cerca di offrire una cornice ontologico-filosofica ai tentativi, psicologico-cognitivi e/o neuroscientifici, di naturalizzare il mentale. Entro tale cornice viene sostenuto che le tre caratteristiche fondamentali delle percezioni coscienti – ossia il loro essere degli stati Intenzionali¹, coscienti e autocoscienti – sono identiche da un punto di vista ontologico a tre distinte *higher order properties* della dinamica cerebrale. Più in dettaglio, limitando la mia analisi alle percezioni visive come esempio privilegiato le cui conclusioni posso essere estese a tutti gli altri stati di coscienza, sostengo in primo luogo che il contenuto di una percezione visiva cosciente risulta essere identico ad una neuro-modulazione di certi processi cerebrali. Ad esempio “vedere rosso” viene ridotto in termini funzionali ad un “vedere rossamente”, in coerenza con la “teoria avverbiale” della percezione. In secondo luogo, dopo aver distinto nettamente la coscienza dall’autocoscienza, identifico il carattere autocosciente della maggior parte degli stati di coscienza con il fatto che il cervello, allorché monitorizza il suo *Self* – ossia allorché monitorizza l’attività di coordinamento che esso esercita su parte dei suoi stessi processi cerebrali – si auto-percepisce come un Io. In terzo luogo sostengo che la coscienza fenomenica, la cui base neurologica è da individuarsi anzitutto (ma forse non esclusivamente) in processi di sincronizzazione tra circuiti neuronali oscillanti, può essere considerata in termini funzionali come il “formato” assunto dai processi cerebrali che implementano l’Intenzionalità e l’autocoscienza quando tali processi, proprio grazie all’aver assunto tale “formato coscienziale”, divengono reciprocamente coerenti e quindi idonei a guidare in modo più flessibile e sensibile al contesto il coordinamento senso-motorio.

Nel terzo capitolo cerco anzitutto di estendere, per analogia, alla relazione tra fatti socio-culturali e stati mentali individuali il riduzio-

¹ Seguendo l’esempio di Dennett (1987), scrivo da ora in poi “Intenzionale” (o “Intenzionalità” ecc.) con la lettera maiuscola quando voglio indicarne il peculiare senso filosofico introdotto da Brentano (1874), ossia l’aver un contenuto. Su Brentano cfr. anche *infra* capitolo ottavo).

nismo ontologico sostenuto nel capitolo precedente riguardo al problema mente-corpo. Per giungere a questa conclusione occorre tuttavia chiarire preliminarmente l'ontologia del mentale già abbozzata in precedenza. Seguendo l'epistemologia naturalizzata di W.v.O. Quine, si deve ammettere in generale che l'esistenza di tutto ciò di cui parla una teoria scientifica è relativa all'ordine di discorso a cui la teoria stessa appartiene. Tuttavia, per non cadere nel relativismo cognitivo, è opportuno distinguere tra esistenza relativa ed esistenza assoluta quando ci si interroghi sulla relazione tra gli stati mentali ed i loro correlati cerebrali. Ad esempio la psicologia implica l'esistenza degli stati mentali, mentre le neuroscienze implicano l'esistenza dei processi cerebrali. Ma – in una prospettiva naturalistica – appare sensato sostenere che gli stati mentali non potrebbero esistere se non fossero implementati da dei processi fisico-chimici quali sono i processi cerebrali. I processi fisico-chimici viceversa possono esistere anche in assenza di qualsiasi loro correlato mentale (come avviene nell'universo per la quasi totalità dei fenomeni). Ciò che esiste nell'ordine di discorso della fisico-chimica (e di conseguenza anche della biologia) gode di esistenza assoluta, mentre gli stati mentali godono, in quanto stati mentali, solo di una esistenza relativa indicizzata per così dire all'ordine di discorso della psicologia e delle scienze umane e sociali in generale. Essi sono degli stati virtuali implementati da processi cerebrali. Pertanto una esistenza assoluta è loro attribuibile solo in modo indiretto grazie all'esistenza assoluta dei processi cerebrali che li implementano. Ed è solo grazie a questa loro implementazione fisica che gli stati mentali possono essere causa di movimenti del nostro corpo o di qualsiasi altro effetto nel mondo reale. Ciò non significa però che l'attribuire una esistenza relativa e virtuale agli stati mentali (o ai loro equivalenti funzionali in psicologia cognitiva) sia inutile al fine di spiegare e prevedere il comportamento umano. Da un punto di vista ontologico gli stati mentali sono identici a certi processi cerebrali, ma da un punto di vista epistemologico la psicologia non è completamente riducibile alle neuroscienze.

Considerazioni analoghe possono essere fatte anche a proposito dei fatti sociali e culturali, che godono di esistenza relativa e virtuale entro l'ordine di discorso delle scienze umane e sociali, ma possono godere indirettamente di una esistenza assoluta (e quindi anche di poteri

causali) solo se vengono implementati dalle azioni di individui in carne ed ossa: una lingua esiste solo se ci sono degli esseri umani capaci di parlarla (o quanto meno di capirla). Ma questa identità ontologica tra stati mentali (o azioni individuali) e fatti socio-culturali non implica affatto che questi ultimi possano essere spiegati, tutti e completamente, senza l'apporto fondamentale delle scienze umane e sociali.

Inoltre – e questo è il secondo nodo problematico affrontato nel terzo capitolo – non sempre gli stati virtuali che godono di esistenza relativa entro la psicologia e le altre scienze cognitive, umane e sociali sono implementati completamente da processi fisico-chimici. In altre parole alcuni fatti socio-culturali o mentali, essendo privi di una completa realizzazione da parte di stati o processi fisico-chimici, non godono di alcuna esistenza assoluta, neppure in modo indiretto, e sono perciò privi di qualsiasi potere causale. Essi solo in apparenza sono stati virtuali; in realtà sono delle finzioni (sono dei *ficta*), ossia sono delle idealizzazioni o addirittura sono il contenuto di vere e proprie illusioni; illusioni talvolta praticamente utili, anzi in molti casi indispensabili, e tuttavia illusioni che devono essere eliminate da una concezione scientifica del mondo.

Nel quarto capitolo torno sulla seconda svolta cognitiva in filosofia della mente. Essa si configura come un mutamento di paradigma nello studio del mentale così profondo che è possibile paragonarlo alla rivoluzione che investì la fisica un secolo fa ad opera di A. Einstein, allorché questi nella teoria della relatività cambiò tutte le leggi fondamentali della fisica grazie al rifiuto del concetto newtoniano di tempo assoluto. L'analogia tra i due mutamenti di paradigma – l'uno affermatosi in fisica un secolo fa e l'altro in corso oggi nello studio scientifico della mente – diviene infatti quanto meno sostenibile quando si rifletta sul fatto che gli odierni studi neurologici sul modo nel quale noi esseri umani percepiamo lo scorrere del tempo e per ciò stesso ci sentiamo svegli (ossia acquisiamo coscienza fenomenica) rivelano che il tempo assoluto della meccanica classica, lungi dall'essere confermato dai dati sperimentali, è in effetti solo una costruzione del nostro cervello. Ciò spiega perché la rivoluzione concettuale operata da Einstein, negando essa ciò che in virtù delle modalità di funzionamento del nostro cervello siamo portati spontaneamente a ritenere indubitabile, sia apparsa così difficile da comprendere ai suoi contem-

poranei e appaia ancor oggi molto strana per il senso comune.

Più in dettaglio gli studi del neuroscienziato E. Pöppel mostrano che due eventi esterni (ad esempio due lampi di luce), a causa del limitato potere di risoluzione temporale del cervello umano, appaiono simultanei quando si susseguano con un intervallo inferiore ai trenta millisecondi. Poiché in trenta millisecondi la luce percorre circa novemila chilometri, essa sembra muoversi, per quanto è rilevabile sulla Terra ad occhio nudo, ad una velocità infinita. Di conseguenza i fenomeni relativistici di rallentamento degli orologi e accorciamento dei regoli non sono percepibili ad occhio nudo. E ciò spiega perché l'esistenza di un tempo assoluto che scorra allo stesso ritmo in tutto l'universo per tutti i possibili osservatori, indipendentemente dai loro moti reciproci, appaia del tutto ovvia, mentre i fenomeni relativistici previsti da Einstein, sebbene siano ormai scientificamente provati al di là di ogni ragionevole dubbio, non cessano di apparire fortemente contro-intuitivi.

Per ragioni analoghe, sempre legate al credere nell'esistenza di un tempo e di uno spazio assoluti, l'intera meccanica classica è ancor oggi giudicata in genere più facilmente comprensibile della teoria della relatività. Di conseguenza, per quanto controversa sia sempre apparsa ai fisici la concezione newtoniana della gravità come una forza che agisce misteriosamente a distanza tra tutti i corpi (si pensi al riguardo alla celebre polemica tra G.W. Leibniz e i newtoniani), ancora più incomprensibile è apparsa ed ancor oggi appare non solo al senso comune ma anche a molti filosofi l'idea, propria della teoria della relatività, che la gravità sia una forza solo apparente e che tutti i moti, ad esempio, dei pianeti del sistema solare che la meccanica classica interpreta come moti accelerati, causati in uno spazio tridimensionale euclideo dall'attrazione gravitazionale siano invece spiegabili con molta maggiore precisione dalla teoria generale della relatività come moti inerziali entro uno spazio-tempo tetradimensionale non-euclideo reso incurvato dalla presenza, in primo luogo, della grande massa del Sole.

Ora, se si tiene presente che secondo molti studi neurologici certi processi di sincronizzazione cerebrale sono alla base (o per lo meno sono condizione necessaria) non solo della percezione dello scorrere del tempo ma anche dell'emergere di quella coscienza fenomenica mediante la quale il cervello, oltre a costruirsi una "scena unitaria"

dell'ambiente esterno, costruisce anche, percependosi come un Io, un'immagine idealizzata della sua stessa attività, ci si può chiedere se il sentirsi i "timonieri" del proprio corpo (secondo il noto paragone aristotelico) non sia una (pur utilissima) illusione analoga a quella che porta la meccanica classica a vedere nella gravità una misteriosa azione a distanza. In altre parole la meccanica classica interpreta la gravità come una misteriosa forza di attrazione a causa del dogma dell'esistenza di un tempo assoluto e del conseguente carattere euclideo dello spazio, mentre la teoria della relatività vede in essa solo una conseguenza della curvatura dello spazio-tempo in uno spazio non-euclideo. Non potrebbe darsi allora che anche l'Io cosciente, sebbene in realtà sia riconducibile a delle proprietà sistemiche della dinamica cerebrale, appaia in modo illusorio, in virtù del limitato potere di risoluzione temporale del cervello, come una "entità spirituale" che, grazie al suo essere – come si suppone – dotata di libero arbitrio, sarebbe capace di modificare per così dire dall'esterno quella attività cerebrale della quale è invece in realtà parte integrante? È questa un'ipotesi ardita e per ora non provata, ma per nulla implausibile alla luce dei dati sperimentali già forniti dalla seconda svolta cognitiva negli ultimi trent'anni.

Nel quinto capitolo torno a sottolineare, riprendendo l'ontologia del mentale già abbozzata nel secondo e nel terzo capitolo, che tanto la prima quanto la seconda svolta cognitiva trovano la loro origine nell'epistemologia naturalizzata di Quine. Più precisamente il naturalismo cognitivo, se non vuol cadere nel relativismo cognitivo, deve presupporre la validità non solo del realismo scientifico, anche – entro limiti da precisare – del realismo metafisico; ossia deve presupporre l'esistenza assoluta di una realtà esterna che sia indipendente dal nostro conoscerla e dal modo nel quale essa a noi appare. Ma, se vogliamo fare ipotesi attendibili, sebbene sempre rivedibili, su come tale realtà effettivamente sia, possiamo formulare queste ipotesi basandoci solo su quanto ci suggeriscono le scienze empiriche e non sulla speculazione filosofica.

Il realismo scientifico, la cui sede di elezione è la fisica, richiede tuttavia degli aggiustamenti e delle riflessioni ulteriori quando venga applicato alle scienze cognitive. In primo luogo, per evitare che esso apra la strada al relativismo cognitivo, è necessario distinguere

l'esistenza assoluta dall'esistenza relativa e introdurre entro ogni ordine di discorso le finzioni (o *ficta*) accanto agli *illata* e agli *abstracta*, seguendo le linee dell'ontologia del mentale già abbozzata nel terzo capitolo.

In secondo luogo occorre far notare – a chi, negando la continuità tra scienza e filosofia, attribuisca alla sola epistemologia il compito di valutare la validità logico-formale delle teorie proposte dalle scienze empiriche – che Quine ha portato argomenti convincenti a favore del crollo della distinzione tra giudizi analitici e giudizi sintetici e che pertanto in una epistemologia naturalizzata non è possibile stabilire una rigida divisione del lavoro secondo la quale il filosofo dovrebbe occuparsi solo della validità logico-formale delle teorie scientifiche e lo scienziato solo della verità empirica del loro contenuto.

In terzo luogo è opportuno notare che intendere il realismo scientifico come una teoria filosofica che considera illegittimo ed inutile parlare di una realtà metafisica indipendente dal modo nel quale la descriviamo e spieghiamo condanna l'empirismo a slittare verso l'idealismo.

Il sesto capitolo è un'appendice del quinto. In esso difendo una concezione pragmatistica del concetto di verità adeguata al realismo scientifico presentato nel capitolo precedente. Se esiste una realtà metafisica indipendente dalla rappresentazione che ogni forma umana di conoscenza può darne, inclusa quella offerta da una concezione scientifica del mondo, allora intendere la verità delle teorie scientifico-empiriche come completamente identica alla loro coerenza interna è sbagliato. Intendere la verità come coerenza interna può andar bene per la matematica e la logica, ma non per le scienze empiriche. Quando una teoria scientifica afferma qualcosa riguardo alla realtà, la sua affermazione è vera se e solo essa corrisponde alla realtà stessa. Ma in quale senso può una teoria “corrispondere” ad una realtà in sé indipendente dal modo nel quale la percepiamo e descriviamo? Si può naturalmente ignorare questa domanda, come fa il realismo ingenuo del senso comune, e continuare ad accettare come cosa ovvia che le teorie vere rispecchiano la realtà. Ma, se si prende quella domanda sul serio, per rispondere sensatamente ad essa non ci sono a ben guardare che due possibilità: o si nega, come fanno tutte le epistemologie anti-realistiche, che il concetto di realtà metafisica ricopra

alcun ruolo nella ricerca della verità e si sostiene che qualsiasi affermazione riguardo al mondo esterno (sia essa una teoria scientifica, una semplice opinione o una percezione) prende a proprio oggetto – per dirla in termini kantiani – non la “cosa in sé” bensì il fenomeno, ossia l’insieme delle nostre esperienze; oppure, volendo evitare la deriva idealistica che è implicita nell’anti-realismo, bisogna ammettere che qualsiasi nostra rappresentazione mentale può “corrispondere” alla realtà esterna non perché assomigli ad essa, bensì perché offre uno strumento indispensabile per agire con successo su quella realtà stessa, così come lo scontrino del guardaroba aiuta la guardarobiera a trovare rapidamente il cappotto corrispondente.

Questa concezione pragmatistica della verità va tuttavia soggetta a varie obiezioni. In particolare essa non sembra corrispondere a ciò che abbiamo intuitivamente in mente quando riteniamo che ciò che pensiamo sia vero. In questo sesto capitolo cerco di rispondere a queste obiezioni, mostrando anzitutto che la teoria della verità come corrispondenza con la realtà propria del realismo ingenuo, difficilmente evitabile nella vita di tutti i giorni, è giustificabile (nella limitata misura in cui lo è) solo in base ad un concetto pragmatistico di verità: l’unico soddisfacente in una concezione scientifica del mondo.

Nel settimo capitolo torno sulla distinzione già introdotta nel secondo capitolo tra il *Self*, in quanto insieme delle attività cerebrali mediante le quali il cervello auto-monitorizzandosi e auto-rappresentandosi rende anzitutto più efficace il coordinamento senso-motorio, e l’Io, in quanto contenuto idealizzato e in larga parte illusorio (per quanto praticamente indispensabile) di quella auto-rappresentazione quando essa venga codificata dal cervello stesso, per renderla più efficace, nel “formato” della coscienza fenomenica. In questo capitolo aggiungo però che quella distinzione tra il *Self* e l’Io era già presente, sebbene in un contesto teorico estraneo all’odierno naturalismo cognitivo, in Kant sotto la forma di una distinzione tra l’unità analitica e l’unità sintetica dell’appercezione trascendentale. Secondo Kant qualsiasi mia rappresentazione cosciente contiene come parte del suo contenuto un riferimento a me quale suo autore. E questo riferimento è ciò che egli chiama unità analitica dell’appercezione trascendentale. Ma, poiché – sostiene Kant – ogni analisi presuppone una prece-

dente sintesi, l'unità analitica dell'appercezione trascendentale presuppone una sua precedente unità sintetica. Ossia ogni mia rappresentazione mentale è cosciente (e solo per questo è una mia rappresentazione) solo perché è stato il mio intelletto, in quanto unità sintetica dell'appercezione trascendentale, a costruirla mettendo insieme certe mie singole intuizioni o percezioni grazie agli "stampi" offerti *a priori* dai suoi concetti puri (o categorie che dir si voglia). In altre parole – secondo Kant – io sono l'agente che costruisce le mie rappresentazioni coscienti; e proprio per questo mi ritrovo nel contenuto di ciascuna di esse come una sorta di marchio di fabbrica.

Ora, in questo settimo capitolo avanzo l'ipotesi che da un punto di vista puramente teorico (non storico!) la distinzione kantiana tra l'unità analitica e l'unità sintetica dell'appercezione trascendentale aiuti ad introdurre una distinzione simile nell'odierna filosofia della mente tra il *Self* e l'Io. Il *Self* infatti è ciò che dà unità e coerenza all'intera attività cerebrale, essendone esso stesso parte; e quindi svolge nel cervello un ruolo di "sintesi" paragonabile a quello attribuito da Kant all'unità sintetica dell'appercezione trascendentale nel complesso delle nostre esperienze. L'Io invece, in quanto contenuto idealizzato e illusorio della "rappresentazione" che il cervello dà del suo stesso *Self* affinché esso possa svolgere correttamente il suo ruolo, è simile a quell'unità analitica dell'appercezione trascendentale che secondo Kant compare nel contenuto di ogni mio stato di coscienza proprio come risultato e testimonianza di quel processo di unificazione delle nostre esperienze svolto dall'unità sintetica.

L'ottavo e ultimo capitolo si muove all'interno del dialogo che da alcuni anni è in corso tra psicoanalisti, da un lato, e filosofi della mente e scienziati cognitivi dall'altro. Basandomi su alcuni scritti di Freud sostengo che, se si confronta il suo pensiero con le teorie oggi prevalenti in filosofia della mente, non è difficile notare che egli, in coerenza con un ateismo e un determinismo di stampo positivisticò che non abbandonò mai nell'arco di tutta la sua vita (salvo una breve parentesi da studente), difese in gioventù sulla natura della mente teorie che oggi verrebbero definite come naturalistico-fisicalistiche e solo negli scritti più tardi approdò a posi-

zioni di tipo funzionalistico; ma probabilmente lo fece solo perché le neuroscienze del suo tempo non offrivano prove sufficientemente solide a quel riduzionismo materialistico che egli in effetti non cessò mai di vagheggiare.

Siena, 21 marzo 2020

Ringraziamenti

Ringrazio anzitutto Pietro Perconti per aver accettato di pubblicare questo mio volume presso la Casa Editrice Corisco, nonché Mario Velluso e Francesco Parisi per averne curata l'impaginazione e la definitiva pubblicazione. Analogo ringraziamento esprimo anche a Luigi Pastore per l'interessamento che mi ha sempre mostrato affinché questo libro venisse pubblicato.

Ringrazio inoltre i direttori delle riviste e i curatori dei volumi (nonché le rispettive case editrici) dove sono comparsi i saggi qui sotto elencati per avermi autorizzato a ripubblicarli in tutto o in parte in forma spesso molto rimaneggiata.

Capitolo primo

Ringrazio Antonino Pennisi per avermi autorizzato a ripubblicare in forma modificata il mio saggio

La seconda svolta cognitiva in filosofia della mente, «Reti, saperi, linguaggi: Italian Journal of Cognitive Sciences», 3(2), 2014, pp. 319-339.

Capitolo secondo

Ringrazio Antonio Federico per avermi autorizzato a ripubblicare in una versione italiana ampliata e modificata il mio saggio *The mind-body problem in the philosophy of mind and cognitive neuroscience: a physicalist naturalist solution*, «Neurological Sciences», 2018. Ringrazio inoltre Christoph Lumer, Giacomo Romano, Pietro Perconti e la Casa Editrice Corisco per avermi consentito di rifondere qui il mio saggio *Realismo scientifico e ontologia naturalizzata del mentale* comparso in C. Lumer e G. Romano (a cura di), *Dalla filosofia dell'azione alla filosofia della mente*, Roma-Messina, Corisco 2018, pp. 13-41. Per un analogo motivo ringrazio anche Fabio Paglieri per avermi autorizzato a riutilizzare parte del mio saggio *La naturalizzazione delle rappresentazioni mentali*, «Sistemi intelligenti», 23(1), 2011, pp. 41-58.

Capitolo terzo

Ringrazio l'ing. Roberto Camporesi per avermi autorizzato, per conto dell'Associazione Nuova Civiltà delle Macchine, a ripubblicare in forma modificata il mio saggio *Mente, corpo e società nel naturalismo forte*, «Nuova Civiltà delle Macchine», 24(2), 2006, pp. 112-128.

Capitolo quarto

Ringrazio Luigi Pastore per avermi autorizzato, nella sua veste di *Managing Editor* di RiFP, a ripubblicare in una versione italiana ampliata e ampiamente modificata il mio saggio *Time and Consciousness in Cognitive Naturalism*, «Rivista internazionale di filosofia e psicologia» 6(3), 2015, pp. 458-473 (Il saggio è stato pubblicato prima che io divenissi direttore di RiFP, quando ne era ancora direttore Massimo Barale, collega ed amico carissimo, al quale purtroppo posso rivolgere solo un ringraziamento alla memoria).

Capitolo quinto

Ringrazio Roberta Lanfredini, Alberto Peruzzi e la Casa Editrice ETS per avermi autorizzato a ripubblicare in forma modificata il mio saggio *Realismo scientifico e naturalismo cognitivo*, in Lanfredini R., Peruzzi A. (a cura di), *A plea for balance in philosophy: Essays in honour of Paolo Parrini*, Pisa, ETS, 2013, pp. 113-127.

Capitolo sesto

Ringrazio Cristina Amoretti, Michele Marsonnet e la Casa Editrice Giuffrè per avermi autorizzato a ripubblicare in forma modificata il mio saggio *Il concetto di verità in una prospettiva naturalistica*, in Amoretti M.C., Marsonnet M. (a cura di), *Conoscenza e verità*, Milano, Giuffrè, 2007, pp. 45-69.

Capitolo settimo

Ringrazio Edoardo Amoroso, Alfredo Ferrarin, Claudio La Rocca e la Casa Editrice ETS per avermi autorizzato a ripubblicare in forma modificata il mio saggio *Kant e le scienze cognitive sulla natura dell'Io*, in Amoroso L., Ferrarin A., La Rocca C. (a cura di), *Critica della ragione e forme dell'esperienza: Studi in onore di Massimo Barale*, Pisa, Edizioni ETS, 2011, pp. 415-432.

Capitolo ottavo

Ringrazio nuovamente Fabio Paglieri per avermi autorizzato a ripubblicare in forma modificata il mio saggio *Inconscio, coscienza e intenzionalità nel naturalismo cognitivo*, «Sistemi intelligenti», 25(3), 2013, pp. 453-467.

Capitolo primo

La seconda svolta cognitiva in filosofia della mente

1. *La svolta linguistica in filosofia della mente*

La filosofia della mente, in quanto disciplina filosofica autonoma, nasce nel XX secolo come prodotto della cosiddetta «svolta linguistica» (Rorty 1967). Quest'ultima – com'è noto – coincide nella sostanza con l'affermarsi, nel Novecento, della filosofia analitica e consiste nell'intendere la filosofia stessa non come una speculazione metafisica sulla struttura ultima della realtà e della nostra conoscenza di essa, bensì come una riflessione critica sul modo nel quale ogni nostra idea viene espressa linguisticamente, partendo dal «convincimento che una spiegazione filosofica del pensiero sia conseguibile attraverso una spiegazione filosofica del linguaggio» (Dummett 1988, p. 13)².

L'analisi filosofica del linguaggio è stata condotta a sua volta, a partire dai primi del Novecento sino ai giorni nostri, in due modi molto diversi: da un lato essa si è identificata con la ricostruzione della forma logica sottesa sia alle teorie scientifiche sia, in modo meno rigoroso, al linguaggio ordinario (si pensi in particolare a B. Russell, al primo Wittgenstein e a tutti gli empiristi logici); dall'altro l'analisi filosofica si è trasformata con il secondo Wittgenstein in una dissoluzione dei problemi filosofici, considerati meri «crampi mentali»³, mediante esempi tratti dal linguaggio ordinario di frasi intuitivamente corrette.

È a questo secondo modo di intendere l'analisi filosofica che appartiene *The Concept of Mind* (Ryle 1949), libro dal quale si suole datare l'inizio della filosofia della mente nella sua accezione contemporanea⁴. Infatti il problema che si poneva G. Ryle in quel testo non era stabilire, in sede metafisica, che cosa fosse la mente, bensì chiarire in quali contesti fosse corretto, nel linguaggio ordinario, usare sia la parola “mente” sia tutti gli altri termini della *folk-*

² In tutte le citazioni formulate usando il sistema “autore-data” l'anno è sempre quello della prima edizione dell'opera citata, ma i numeri delle pagine, se nei riferimenti viene indicata una edizione successiva o una traduzione italiana, si riferiscono invece a quella edizione o traduzione.

³ Per l'espressione wittgensteiniana *mental cramp* cfr. Malcom (1966, p. 43).

⁴ Per maggiori dettagli cfr. Nannini (2011, pp. 90-97).

psychology. In questa prospettiva Ryle riteneva che gli stati mentali, a differenza di quanto aveva sostenuto Cartesio, non fossero misteriose cause mentali interne del comportamento osservabile, bensì semplicemente tendenze a comportarsi in un certo modo. Quando nella lingua di tutti i giorni parliamo di ciò che gli altri pensano, vogliono, temono o amano, le nostre parole hanno il senso che hanno non perché si riferiscano a misteriosi stati interni privati, bensì in virtù di ciò che gli altri effettivamente fanno o dicono (cioè in virtù del loro comportamento osservabile) (Ryle 1949, p. 193).

Ad esempio sono a tavola e chiedo al mio vicino il sale; questi, educatamente, me lo passa. Se qualcuno mi domanda perché me lo abbia passato, posso rispondere che lo ha fatto per educazione. Ma, con ciò, non intendo affatto dire che la sua educazione è una cosa spettrale depositata (non si sa bene come) nella sua mente (un luogo “interno” a lui solo accessibile) che, in maniera altrettanto misteriosa, ha causato il movimento della sua mano verso la saliera. Più semplicemente intendo spiegare il suo gesto attribuendo al mio vicino di tavola la tendenza (la disposizione, l’*habitus*) a comportarsi in modo educato quando si trovi in determinate circostanze. Le spiegazioni del comportamento umano in termini psicologici non sono causali, bensì disposizionali. Ad esempio, se spiego la rottura di un vetro con il fatto che il vetro stesso è stato colpito da un sasso, la mia spiegazione è causale; se invece spiego quella rottura dicendo che il vetro era fragile (ossia aveva la tendenza a rompersi, se colpito con forza sufficiente), la mia spiegazione è disposizionale. Ora, i termini psicologici – secondo la concezione che ne proponeva Ryle nel 1949 – assomigliano più al concetto di fragilità che non a quello di urto di un sasso. Non si riferiscono a misteriosi stati o eventi interni accessibili solo introspettivamente, bensì esprimono tendenze, o meglio disposizioni comportamentali pubblicamente osservabili (Ryle 1949, pp. 109-110); disposizioni che possono talvolta durare tutta la vita, come un tratto fondamentale della personalità, e talvolta possono durare invece pochi secondi, come l’intenzione di bere un bicchier d’acqua.

La soluzione data dal comportamentismo filosofico di Ryle al problema mente-corpo andò incontro tuttavia già nella seconda metà degli anni Cinquanta e negli anni Sessanta a molte critiche, soprattutto ad opera di fisicalisti e funzionalisti: il dolore ad esempio –

sostennero entrambi questi gruppi di filosofi – non è la tendenza a gemere, è la causa interna del gemere. Hilary Putnam, in particolare, propose un esempio immaginario contro il comportamentismo. Facciamo un po' di fantascienza – scrisse Putnam (1963, pp. 355-364) – e immaginiamo che esista una comunità di adulti, che chiameremo Superspartani o Superstoici, soggetti a provare dolore come tutti noi, ma dotati di uno straordinario controllo delle proprie reazioni emotive e abituati, per ragioni ideologiche a loro avviso importantissime, a reprimere ogni manifestazione osservabile del dolore stesso. Anche se atrocemente torturato un superspartano non geme, non urla e non suda, sebbene soffra le pene dell'inferno; e, se è un Super-superspartano, neppure ammetterà di conoscere che cosa significhi la parola “dolore” o di aver mai provato dolore in vita sua, sebbene ne stia patendo uno acutissimo in quel momento. L'esempio è ovviamente quanto di più implausibile si possa immaginare, ma il fatto che possiamo concepirlo senza cadere in una contraddizione logica prova a sufficienza – secondo Putnam – che il significato della parola “dolore” nel linguaggio ordinario non è riducibile ad un insieme di disposizioni comportamentali. Se lo fosse, allora dovremmo concludere che i Superspartani non possono provare dolore, dal momento che non manifestano mai di provarlo. Ma, per ipotesi, abbiamo detto che in effetti lo provano. L'esempio non sarebbe perciò logicamente possibile, se il dolore non fosse normalmente concepito come qualcosa di diverso dal comportamento mediante il quale lo manifestiamo. La parola “dolore”, nell'accezione comune, si riferisce non ad un insieme di comportamenti o di disposizioni comportamentali, bensì ad uno stato di cose interno e nascosto che è responsabile del nostro comportamento manifesto.

2. La prima svolta cognitiva

I filosofi che negli anni Cinquanta e Sessanta condividevano questa critica al comportamentismo di Ryle erano divisi in due gruppi. I primi – si pensi soprattutto agli australiani U.T. Place (1956), F. Smart (1963) e D. Armstrong (1968), sostenitori della “teoria dell'identità tra mente e cervello” (o “materialismo dello stato centrale”) –, volendo evitare che l'ammettere delle cause mentali interne del comportamento osservabile riportasse in auge il dualismo cartesiano tra mente e corpo, identificarono tali cause mentali semplicemente

con dei processi cerebrali. Putnam invece, seguito poi (nonostante la sua successiva ritrattazione) dalla maggioranza dei filosofi della mente negli anni Sessanta e Settanta, propose d'intendere sì gli stati mentali come una ridescrizione di processi cerebrali, ma in termini astratti e puramente funzionali (Putnam 1960). In tal modo la relazione tra mente e corpo veniva ad assomigliare a quella tra *software* e *hardware* in un computer. Con il primo Putnam era sorta l'"analogia mente-computer", che rimase fino alla metà degli anni Ottanta il nocciolo duro del funzionalismo (o cognitivismo) in filosofia della mente e in psicologia cognitiva.

In sostanza, sebbene i fisicalisti, sostenitori della teoria dell'identità, e i funzionalisti, fautori dell'analogia mente-computer, condividessero con Ryle il rifiuto del dualismo cartesiano, tuttavia entrambi questi gruppi motivavano il loro rifiuto in modo diverso da Ryle stesso, perché identificavano gli stati mentali non con delle disposizioni comportamentali, bensì con delle cause mentali interne, salvo poi dividersi sul modo di concepire tali cause, identificandole o direttamente con dei processi cerebrali oppure con degli stati funzionali-virtuali implementati da processi cerebrali.

La prima identificazione portava i filosofi della mente a interessarsi di neuroscienze; la seconda li induceva invece a rivolgersi agli allora nascenti studi di intelligenza artificiale, che identificavano il pensare con la struttura matematica astratta di un algoritmo la cui esecuzione poteva essere affidata indifferentemente al cervello umano o ad una macchina. In entrambi i casi, seguendo l'esempio o delle neuroscienze o dell'intelligenza artificiale, alla svolta linguistica si aggiungeva e in parte subentrava in filosofia della mente una "prima svolta cognitiva"; svolta favorita anche, da un lato, dalla nascita delle scienze cognitive intorno agli anni Cinquanta e Sessanta (Gardner 1985) e, dall'altro, dalla critica che in quei medesimi anni W.v.O. Quine rivolse, in sede filosofica, all'empirismo logico. Il crollo, ad opera di Quine (1951), della distinzione tra giudizi analitici e giudizi sintetici rese difficile infatti distinguere nettamente la scienza dalla logica della scienza; e quindi assottigliò anche il confine tra scienze cognitive e filosofia della mente. Nella stessa direzione andò inoltre la lezione di filosofi e storici della scienza come Sellars (1956) e Kuhn (1962).

In conclusione la chiave per ricostruire la struttura del pensiero umano, nonostante grandi filosofi analitici come Dummett (1991) si

muovessero in direzione opposta, andò spostandosi nel corso degli anni Sessanta e Settanta dall'analisi logica del linguaggio agli studi empirici condotti sui processi cognitivi da psicologi, linguisti e informatici. In questo contesto emerse sempre più, ad esempio, lo scarto vistoso esistente tra il modo nel quale noi esseri umani dovremmo pensare, seguendo le regole della logica matematica, ed il modo nel quale di fatto ragioniamo⁵.

Tuttavia l'ancora scarso sviluppo, negli anni Sessanta, di una vera e propria neuroscienza cognitiva impedì che il dibattito sulla teoria dell'identità potesse avvalersi più di tanto degli studi sui correlati neurali degli stati mentali. Critici e sostenitori del fisicalismo discussero a lungo in quegli anni se l'identità tra stati mentali e processi cerebrali fosse logicamente possibile (ad es. Hook 1960); ma i fisicalisti stessi, quando volevano dare degli esempi concreti di tale identità, non sapevano andare molto oltre la generica identificazione del dolore con l'attività delle fibre C.

Molto più forte fu invece il contributo dato alla prima svolta cognitiva dall'intelligenza artificiale (Boden 1990) e dalla psicologia cognitiva (Neisser 1967). Dato che gli stati mentali erano descrivibili per i cognitivisti come stati funzionali, allora divenne per loro possibile, sulla scia del primo Putnam, paragonare la relazione che intercorre tra la mente ed il cervello alla relazione di implementazione che lega in un computer il *software* all'*hardware*: l'analogia mente-computer si affermò come la base del cognitivismo.

Convergente con l'analogia mente-computer, ma da essa distinta, fu inoltre l'influenza operata su psicologi e filosofi della mente dalla "linguistica generativa" di Chomsky (1957, 1975 e 1992, per limitarsi ai volumi più noti) e in particolare dalla sua difesa dell'innatismo (Piattelli-Palmarini 1980). Secondo N. Chomsky tutte le lingue naturali funzionano in base a strutture sintattiche che sono varianti di un'unica grammatica universale innata. Mediante l'estensione dell'innatismo di Chomsky dal linguaggio al pensiero Fodor (1975, 1981, 1983 e 1987) giunse a formulare l'ipotesi che la struttura combinatoria delle lingue naturali – che sono capaci di generare un numero potenzialmente infinito di frasi mediante un reper-

⁵ Cfr. ad es. Johnson-Laird (1983) e Legrenzi (1980) per questo scarto tra logica e psicologia cognitiva.

torio finito di simboli combinati attraverso un numero ristretto di regole sintattiche – costituisca il “sistema operativo” che consente al cervello umano di generare dei processi cognitivi (*Language of Thought Hypothesis*)⁶. Per altro verso D.C. Dennett combinò negli anni Settanta il comportamentismo di Ryle con il funzionalismo del primo Putnam, giungendo per questa via a dare un’originale interpretazione naturalistica del concetto brentaniano d’Intenzionalità (Dennett 1969, 1978 e 1987).

Ma già negli anni Settanta (e con ancora maggior frequenza a partire dagli anni Ottanta) si moltiplicarono le critiche al funzionalismo e al cognitivismo. Queste critiche provennero da tre direzioni diverse. In primo luogo molti filosofi – in un clima culturale sempre più ostile, tra gli psicologi stessi, al comportamentismo e favorevole alla riscoperta della coscienza – obiettarono che il funzionalismo, al pari del fisicalismo, mette in ombra l’intrinseca irriducibilità della coscienza fenomenica all’organizzazione funzionale del cervello, per quanto complessa tale organizzazione funzionale possa essere (ad es. Searle (1980 e 1992), Levine (1983), Chalmers (1996 e 2010)).

In secondo luogo nacque all’interno stesso dell’informatica un’alternativa alla “macchina universale di Turing”, fino ad allora considerata come modello unico per qualsiasi sistema che sia capace, elaborando informazione in modo seriale, di simulare l’intelligenza umana. Vennero costruite infatti, per lo più come macchine virtuali simulate su computer digitali, le prime reti neurali artificiali: al cognitivismo si contrapponeva ormai il connessionismo (Rumelhart e McClelland (1986)).

In terzo luogo nuove tecniche d’indagine sul funzionamento del cervello *in vivo* (PET, fMRI ecc. insieme ad un miglioramento dell’EEG) hanno consentito, soprattutto a partire dagli anni Novanta, uno sviluppo straordinario di quella parte delle neuroscienze che studia i processi cognitivi: ad es. il coordinamento senso-motorio, la percezione, la memoria, le emozioni (nel loro rapporto con il ragionamento), il linguaggio, la capacità di comprendere la matematica, l’empatia, l’amore, l’aggressività e più in generale la base biologica

⁶ Fodor stesso tuttavia ha ridimensionato notevolmente questa sua ipotesi in scritti successivi (Fodor 1998 e 2000), avvicinandosi notevolmente al neomisterianesimo (cfr. *infra* secondo capitolo).

del comportamento sociale, della morale e della cultura. In altre parole molti neuroscienziati, dedicandosi allo studio scientifico di argomenti un tempo riservati alla riflessione filosofica, hanno fondato un nuovo ambito di ricerca in neurologia che va sotto il nome di neuroscienza cognitiva⁷.

3. La seconda svolta cognitiva

Da questi nuovi apporti dell'intelligenza artificiale, della robotica (soprattutto Brooks 1990 e 1991) e della neurobiologia, insieme al contributo dato dal neo-darwinismo in tutte le sue versioni⁸ ad una naturalizzazione dell'immagine che noi esseri umani abbiamo di noi stessi, sono sorte all'interno della filosofia della mente a partire dalla fine degli anni Ottanta nuove tendenze di stampo naturalistico, i cui sviluppi sono tuttora in corso; tendenze che vengono a configurarsi, in campo filosofico, come una ancor più radicale "seconda svolta cognitiva".

Fra gli autori che più hanno contribuito a questa seconda svolta cognitiva in filosofia della mente ha contribuito e continua a contribuire in modo essenziale D.C. Dennett (già protagonista della prima svolta), che, a partire dagli anni Novanta, ha fondato sulla teoria darwiniana dell'evoluzione una nuova immagine della coscienza e dell'autocoscienza (Dennett 1991, 1995, 1998, 2003, 2005, 2013a e 2017). Ma eguale importanza hanno avuta ed hanno, insieme a molti altri filosofi, i coniugi Paul e Patricia Churchland, i quali, riprendendo idee di W. Sellars, R. Rorty e W.v.O. Quine, hanno dato vita, in polemica con Fodor ed il cognitivismo in generale, ad una nuova

⁷ Cfr. ad es. Crick (1994), Edelman (1987, 1989, 1992, 2004, 2006), Damasio (1994, 1999, 2003, 2010), Roth (1994 e 2001), Roth e Grün (2006), Pauen e Roth (2008), Roth e Strüder (2014), Oliverio (1995 e 2009), Stamenov e Gallese (2002), W. Singer (2003 e 2004), Rizzolatti e Sinigaglia (2006), Benini (2009 e 2012), Ammaniti e Gallese (2013), Dehaene (2014).

⁸ Cfr. in particolare, per il darwinismo nella sua forma più classica, Dawkins (1976, 1982, 1986 e 2009, per limitarsi ad alcuni dei suoi libri più noti). Per un neo-darwinismo più "liberale" cfr. Gould (2002). Contro corrente invece, contrari alla teoria della selezione naturale pur senza dirsi creazionisti, Fodor e Piattelli-Palmarini (2010). Per quanto concerne gli studi sul neo-darwinismo, cfr. inoltre Minelli (2007) sulle cosiddette teorie EVO-DEVO (*Evolutionary Developmental Biology*) e soprattutto Pievani (2006) per un'equilibrata introduzione generale al neo-darwinismo.

forma di fisicalismo, l'“eliminativismo”, che evita molte delle difficoltà connesse alla teoria dell'identità (P.M. Churchland 1988, 1989, 1995, 2007; P.S. Churchland 1986, 2002, 2011 e 2013). La teoria dell'identità infatti, pur ancor oggi difesa da alcuni filosofi (ad es. Gozzano 2007), pretende che ogni stato mentale sia riducibile ad un corrispondente processo cerebrale. E ciò non sempre è in accordo con i dati empirici delle neuroscienze e della psicologia cognitiva. L'eliminativismo invece ammette che gli stati mentali, come sono descritti dalla *folk-psychology*, molto spesso non sono direttamente riducibili a processi cerebrali; ma spiega tale mancata riduzione, senza fare alcuna concessione al dualismo cartesiano, con il carattere prescientifico dei concetti usati dalla *folk-psychology*, che parla spesso di entità e proprietà che semplicemente non esistono. Ciò non significa però che gli eliminativisti neghino l'esistenza della coscienza, come molti rimproverano loro in modo affrettato e fuorviante. Significa solo che essi ritengono che il concetto corrente di coscienza debba essere sostituito da uno o più concetti scientificamente più adeguati (Churchland e Churchland 1996)⁹.

Soprattutto attraverso l'eliminativismo il rapporto tra filosofia della mente e neuroscienze si è fatto sempre più stretto; e, insieme alla neurofilosofia dei Churchland, sono nate in campo filosofico anche la “neuroetica” e la “neuroestetica”¹⁰, mentre si sono moltiplicate le collaborazioni tra neuroscienze e scienze sociali con la nascita di nuove discipline come la “neuroeconomia”¹¹.

Non sono mancati e non mancano certo eccessi, giustamente denunciati da alcuni come una vera e propria «neuromania» (Legrenzi e Umiltà 2009)¹², in questo tentativo di riconsiderare *brain-wise* non solo

⁹ Ho messa in luce e fatta mia questa replica, in polemica con la maggior parte degli interpreti dei Churchland, in Nannini (2011, p. 203).

¹⁰ Sulla neuroetica esiste ormai una bibliografia sterminata. Cfr. ad es. Racine (2010), Churchland (2011), Cerroni e Rufo (2009), Corbellini e Sirgiovanni (2013). Lo stesso dicasi per la neuroestetica. Cfr. ad es. Ramachandran e Hirstein (1999) e Zeki (1999 e 2008).

¹¹ Cfr. ad es. Bourguin e Walliser (1992), Bourguin e Nadal (2004) e Innocenti (2009).

¹² Il sottotitolo di *Neuromania*, ossia *Il cervello non spiega chi siamo*, è in contrasto ad esempio con Swaab (2010), *Noi siamo il nostro cervello*, ed è in sintonia invece con Noë (2004 e 2009). A Legrenzi e Umiltà hanno però replicato Aglioti e Berlucchi (2013) rimproverando loro di coltivare una vera e propria «neurofobia».

tutte le scienze cognitive e la filosofia della mente, ma anche l'intera tradizione storico-filosofica. Si vengono a creare in tal modo cortocircuiti fuorvianti tra complesse problematiche filosofico-culturali e pretese localizzazioni cerebrali. Del resto P.S. Churchland stessa, fondatrice della neurofilosofia, ha chiarito quanto sia errato ed ingenuo credere che i neuroscienziati siano ormai in grado di leggere i nostri pensieri con le loro neuroimmagini e, quel che sarebbe ancor peggio, manipolarli con i loro strumenti (Churchland 2013, pp. 261-271)!

Tuttavia, mano a mano che si diraderà il polverone sollevato da queste superficiali identificazioni tra stati mentali e processi cerebrali, è auspicabile che emerga con chiarezza che i nuovi dati sperimentali e le nuove elaborazioni teoriche offerte dalla neuroscienza cognitiva non possono non essere di centrale importanza per chi, in sede filosofica, conduca un'analisi di concetti quali la coscienza, l'autocoscienza e l'Intenzionalità (intesa in senso brentanoiano).

4. Riduzionismo e anti-riduzionismo dopo la seconda svolta cognitiva

La seconda svolta cognitiva mostra perciò come sia possibile ai filosofi che fondino le loro riflessioni principalmente sugli studi offerti dalla neuroscienza cognitiva tentare di naturalizzare i fenomeni mentali¹³. Tuttavia il punto più controverso di tali tentativi rimane ancor oggi la riduzione a processi neurologici anche degli stati di coscienza. Perché ad esempio, quando i coni delle nostre retine reagiscono ad onde elettromagnetiche aventi una lunghezza di circa 440 nanometri, noi vediamo una macchia blu? Che cosa ha in comune una certa lunghezza d'onda con un certo colore? E inoltre che cosa hanno in comune le proprietà fisiche della dinamica cerebrale con i *qualia* della nostra esperienza cosciente e soggettiva? Tutto quello che si può provare empiricamente è una certa correlazione tra stati mentali e processi cerebrali. Ma correlazione non necessariamente vuol dire identità.

È questo, in ultima analisi, secondo tutti gli anti-riduzionisti

¹³ Un ottimo esempio dell'utilizzo di studi neuroscientifici per affrontare una più vasta problematica filosofica è reperibile sia in Pastore e Dellantonio (2009 e 2017) sia in Romano (2006 e 2009). In modo analogo anche il modo nel quale riusciamo a comprendere le menti altrui è venuto a basarsi – nel contesto di una epistemologia naturalizzata che si avvale della teoria dei “neuroni specchio” (Rizzolatti e Sinigaglia 2006) –, sull'empatia e la “teoria della simulazione” (Romano 2002).

l'irrisolubile *Hard Problem* che Chalmers (1996 e 2010) e molti altri filosofi e scienziati ritengono che impedisca, in linea di principio, la riduzione della coscienza a qualsiasi tipo di attività cerebrale. Tra la coscienza stessa e l'attività cerebrale esiste insomma, secondo gli anti-riduzionisti, un *gap* fenomenologico, cognitivo e ontologico: la coscienza si presenta come qualcosa di diverso dall'attività cerebrale (*gap* fenomenologico) ed è conoscibile in modo diverso da quest'ultima (*gap* cognitivo – Levine 1983 e 2001), perché è intrinsecamente qualcosa di diverso da essa (*gap* ontologico) (Nagel 1974, Jackson 1982 e 1986)¹⁴.

Ma è davvero inevitabile concludere che il *gap* fenomenologico tra mente e cervello ne implica uno che è anche cognitivo ed ontologico? Un naturalista riduzionista può negare questa implicazione. Ma poi ha il dovere di spiegare in altro modo perché l'esistenza del *gap* fenomenologico stesso sembri essere *prima facie* insormontabile. Di fronte a questa sfida lanciata dagli anti-riduzionisti egli può o negare *tout court* l'esistenza di un *gap* fenomenologico tra coscienza e cervello (ad es. Dennett 2005) oppure può sostenere – in una direzione prospettata da P.M. Churchland (1995, pp. 195-200) – che tale *gap* fenomenologico esiste sì, ma esso non ne implica affatto uno ontologico e cognitivo perché la coscienza, lungi dall'essere una sostanza o una proprietà non fisica, è solo il modo nel quale il cervello presenta a se stesso parte della sua attività quando la monitorizza. In altre parole la coscienza è una sorta di percezione endocerebrale (cfr. *infra* capitolo secondo).

Gli anti-riduzionisti possono certo obiettare che anche la soluzione di Paul Churchland va incontro allo *Hard Problem*: uno Zombi resterebbe uno Zombi anche se il funzionamento del suo cervello includesse tutti i possibili e immaginabili processi di auto-monitoraggio interno e fosse perciò uno Zimbo (Dennett 1999 e 2013b)! Si può tuttavia replicare a questa obiezione che, se si riuscisse davvero a chiarire nel dettaglio quali processi cerebrali siano sottesi alla comparsa della coscienza, sarebbe del tutto irragionevole continuare a credere che la coscienza stessa non sia una funzione biologica implementata, al pari di molte altre, da normali processi cerebrali.

In conclusione, se cerchiamo di collocare nel contesto della storia

¹⁴ F. Jackson ha però successivamente rivisto la sua posizione (Jackson 1998).

delle idee sul problema mente-corpo la discussione degli ultimi quarant'anni tra riduzionisti e anti-riduzionisti riguardo alla relazione coscienza-cervello, possiamo notare che al fondo di molte posizioni anti-riduzionistiche sulla coscienza si trova, in forma più o meno esplicita e consapevole, la tesi di Husserl e di tutti i fenomenologi secondo la quale, prima di indagare da un punto di vista empirico la relazione tra noi ed il mondo, occorre sottoporre ad un'analisi filosofico-trascendentale la nostra semplice esperienza cosciente e soggettiva, facendo astrazione da qualsiasi conoscenza di tipo scientifico noi possiamo avere sul mondo fisico e su noi stessi. È questa, com'è noto, la celebre *epochè* fenomenologica che sta a fondamento delle critiche di Husserl (1913) al naturalismo e che è ripresa da tutti i sostenitori della fenomenologia anche quando coniughino la loro adesione a quest'ultima con un genuino interesse per le scienze cognitive in generale e le neuroscienze in particolare (ad es. Petitot *et al.* 1999, Gallagher e Zahavi 2008)¹⁵. Tutti costoro sostengono infatti non solo che l'esperienza soggettiva e cosciente è un punto di partenza obbligato per qualsiasi ricerca empirica (tesi questa che anche un naturalista non può non condividere), ma anche – e questo invece è il punto dove la fenomenologia ed il naturalismo prendono strade opposte – che la sua analisi filosofica disegna *a priori* la cornice trascendentale entro i cui limiti soltanto potranno svilupparsi le scienze empiriche.

L'eliminativismo dei Churchland al contrario, ispirandosi alla continuità prospettata da Quine (1969) tra scienza e filosofia, offre un punto di vista alternativo ed opposto a quello proposto dalla fenomenologia: quando affrontiamo il problema della natura della coscienza, lungi dal mettere in parentesi tutto quello che sappiamo sul funzionamento del cervello, dobbiamo al contrario ricercare proprio in tale funzionamento la chiave per spiegare perché la coscienza stessa ci appaia secondo le caratteristiche (talvolta ingannevoli) che essa intuitivamente presenta (cfr. *infra* capitolo quarto). Secondo questa ipotesi ciascun essere umano finisce per identificarsi con un Io-persona che, sebbene sia solo una finzione prodotta dalla dinamica cerebrale e dal contesto sociale, sembra sottomettere a sé

¹⁵ Cfr. però al riguardo anche la critica di Lanfredini (2010). Per un approccio fenomenologico alla filosofia della mente cfr. anche Biuso (2008).

tutto il resto dell'attività mentale e riesce a porsi al centro di quel sistema di relazioni e di ruoli che è per ciascuno di noi la propria vita sociale (cfr. *infra* capitolo settimo).

Capitolo secondo

Il problema mente-corpo tra filosofia e neuroscienze

1. Filosofia e neuroscienze

Da più di venti anni è in corso un intenso dialogo tra filosofi e neuroscienziati intorno alla natura della mente ed al suo rapporto con il cervello, ossia anzitutto intorno alle soluzioni di quel problema che i filosofi chiamano di solito il problema mente-corpo¹⁶. Se – come dirò tra poco – si eccettua la possibilità di considerare insolubile questo problema, le soluzioni che fin dall’antichità classica ne sono state date sono fondamentalmente tre (Nannini 2011):

1. Il “dualismo mente-corpo” (nell’antichità anima-corpo) – Mente e corpo (in particolare mente e cervello) sono due sostanze o due proprietà ontologicamente e epistemologicamente irriducibili. Lo studio della mente appartiene alle scienze umane, lo studio del cervello appartiene invece alle scienze della natura¹⁷.
2. Il “fisicalismo” (nell’antichità “atomismo”, in età moderna “materialismo”, oggi giorno anche *Mind-Brain Identity Theory* oppure, con alcune importanti differenze, “eliminativismo”) – Tutti i fenomeni mentali sono riducibili a processi fisici in larga parte cerebrali o comunque sono riconducibili a proprietà e disposizioni della dinamica cerebrale. (Quando non lo sono, ciò significa secondo gli eliminativisti che essi in realtà non esistono e che i loro concetti devono essere eliminati dal linguaggio della scienza)¹⁸.

¹⁶ Alcuni filosofi e neuroscienziati hanno anche scritto libri o saggi a quattro mani: ad esempio Gallese e Goldman (1998), Pauen e Roth (2001 e 2008), Bennett e Hacker (2003). Cfr. anche le repliche di J.R. Searle e D.C. Dennett a Bennett e Hacker in Bennett *et alii* (2007).

¹⁷ Il dualismo, sia pur in versioni molto diverse, annovera tra i suoi sostenitori Platone, Descartes, Popper (1994) e molti filosofi contemporanei. Per una introduzione generale e ampia bibliografia cfr. H. Robinson (2016) e, con una antologia di testi, Lavazza (2008).

¹⁸ Tra i sostenitori del fisicalismo (atomismo, materialismo, ecc.), se esso viene inteso nella varietà delle sue accezioni e versioni, si possono menzionare Democrito ed Epicuro, T. Hobbes, D. Diderot, P.H.T. D’Holbach, J.O. de La Mettrie e altri

3. Il “funzionalismo” (noto soprattutto tra gli psicologi anche come “cognitivismo”, implicitamente lontano erede dell’“ilomorfismo” di Aristotele) – Gli stati mentali sono stati funzionali (detti anche virtuali) del cervello che vengono implementati dall’attività cerebrale così come, in un computer, il *software* viene implementato dall’*hardware*. Poiché uno stesso stato mentale può essere implementato in modo imprevedibile da processi cerebrali diversi (teoria delle “molteplici realizzazioni”), la psicologia cognitiva è autonoma dalle neuroscienze. Lo stesso dicasi, più in generale, di ogni studio della mente, non importa se filosofico o psicologico, rispetto allo studio del cervello umano (e più in generale del sistema nervoso centrale nella sua relazione con il resto del corpo umano)¹⁹.

Neuroscienziati e filosofi si dividono attualmente riguardo al problema mente-corpo in due grandi gruppi: molti optano, sia pur con sfumature diverse, per una soluzione naturalistico-fisicalistica basata su una collaborazione tra lo studio della mente (psicologia e filosofia) e lo studio del cervello (neuroscienza cognitiva). Altri optano invece per il funzionalismo oppure si oppongono in varie forme al naturalismo²⁰, ossia ritengono che lo studio della mente non possa essere condotto con i metodi propri delle scienze naturali. Alcuni di loro anzi optano per un ritorno al dualismo cartesiano. Nel primo gruppo prevalgono i neuroscienziati (ma ci sono anche eminenti filosofi)²¹, nel secondo i filosofi e gli psicologi²².

illuministi, molti positivisti dell’Ottocento, nel Novecento taluni neo-positivisti, Feigl (1967), Place (1956), Smart (1963), Armstrong (1968) e oggi, ad esempio, i già menzionati coniugi Churchland e, entro certi limiti, anche D.C.Dennett. Per una introduzione generale al fisicalismo con ampia bibliografia cfr. Stoljar (2015).

¹⁹

Il funzionalismo ha costituito la corrente principale in filosofia della mente e in psicologia cognitiva negli anni Sessanta e Settanta del Novecento: si pensi soprattutto al suo fondatore, il già menzionato H. Putnam (che lo ha poi rinnegato), a J.A. Fodor, anch’egli già menzionato, e a molti altri. Sebbene sia in crisi fin dagli anni Ottanta, il funzionalismo è ancor oggi tutt’altro che morto. Per una introduzione generale con ampia bibliografia cfr. Levin (2013).

²⁰

Sul naturalismo cfr. Papineau (2016), Nannini e Sandkühler (2000).

²¹

Cfr. ad esempio tra i neuroscienziati i già menzionati F. Crick, A.R. Damasio e G.

Vorrei qui difendere, da un punto di vista filosofico, le ragioni del primo gruppo, ossia vorrei mostrare mediante l'analisi concettuale di una semplicissima azione volontaria – fermarsi davanti ad un semaforo rosso – che la riduzione degli stati mentali a *higher order properties* della dinamica cerebrale è la soluzione più ragionevole del problema mente-corpo e che, per giungere a questa conclusione, un filosofo della mente può oggi servirsi di molte teorie e molti dati empirici offerti dalla neuroscienza cognitiva.

2. Un esempio di azione volontaria

Sono in macchina, vedo che il semaforo è rosso e allora premo il freno e mi fermo. La mia percezione cosciente, nella quale il semaforo mi appare rosso (*[Rosso*]*) – percezione causata dal fatto che il semaforo, secondo il senso comune, è effettivamente rosso (*[Rosso]*) – causa a sua volta, insieme ad altri miei stati e disposizioni mentali *M* (molti dei quali inconsci), la mia azione *A* di premere il pedale del freno. Se – seguendo un orientamento largamente diffuso tra i filosofi della mente – si opta per una rappresentazione degli stati mentali come “atteggiamenti proposizionali”²³, l'intero processo può essere schematizzato mediante una “ricostruzione filosofica” (in grigio gli stati interni alla mia mente accessibili solo a me per via introspettiva, salvo gli elementi inconsci di *M* non accessibili direttamente neppure a me; le relazioni di causa ed effetto sono espresse mediante il simbolo “→”):

Edelman, nonché Edelman e Tononi (2000); tra i filosofi i già menzionati J.R. Searle, D.C. Dennett e i coniugi Churchland.

²² Cfr. ad esempio tra i filosofi apertamente dualisti il già menzionato D. Chalmers, Swinburne (1986 e 2013), Lowe (2000) e W. Robinson (2004); tra i filosofi genericamente di orientamento fenomenologico cfr. soprattutto Dreyfus (1999).

²³ In conformità con questa teoria descrivo qui ogni stato di coscienza accessibile solo introspettivamente come una relazione evidenziata in grigio del tipo *R(a*,b*)* tra un soggetto agente (o pensante) e un fatto, per come esso appare al soggetto stesso. Inoltre mi riferisco mediante [...] al contenuto dell'enunciato dichiarativo “...” e aggiungo un asterisco a come le cose appaiono al soggetto stesso. Ad esempio, se il fatto che il semaforo è rosso (contenuto dell'enunciato dichiarativo “Il semaforo è rosso”) viene abbreviato mediante *[Rosso]*, allora il contenuto della mia percezione di quel fatto (ossia il modo nel quale a me il semaforo appare) viene rappresentato mediante *[Rosso*]*; e quindi il fatto che io vedo che il semaforo è rosso viene espresso mediante la relazione *[Vedere(Io*,[Rosso*])]*. (Perché anche *Io** sia asteriscato verrà chiarito in seguito, cfr. anche *infra* capitolo settimo).

Ricostruzione filosofica

$$[Rosso] \rightarrow ([\text{coscientemente}(\text{Vedere}(\text{Io}^*, [Rosso^*]))] \ \& \ M) \rightarrow A]$$

Questa è nel linguaggio della *folk-psychology* una buona spiegazione del perché ho fermato la macchina: il semaforo era rosso! In essa – si noti – si presume che io veda che il semaforo è rosso perché esso è effettivamente rosso: *[Rosso]* per il senso comune è identico o quanto meno simile a *[Rosso*]* (“trasparenza” degli stati di coscienza (Metzinger 2009)). Salvo il caso di percezioni ingannevoli, si presume nella vita di tutti i giorni che vi sia somiglianza tra il contenuto dei nostri stati di coscienza (ossia l’“oggetto fenomenico” interno alla mente, contrassegnato qui con un asterisco) e il fatto reale esterno che li ha causati.

Ma, se passiamo dal linguaggio della *folk psychology* al linguaggio della fisica e di altre scienze naturali, vediamo bene che, quando diciamo che il semaforo è rosso, intendiamo dire che una lampadina interna al semaforo emette onde elettromagnetiche aventi una lunghezza compresa tra 620 e 750 nanometri. Questo è il vero stimolo distale *[SD-Rosso]* della mia percezione: un evento fisico incolore che è molto diverso dal modo nel quale io lo percepisco (*[SD-Rosso] ≠ [Rosso*]*). Nel linguaggio delle scienze *hard* questo stimolo è la causa di quei processi cerebrali che a loro volta causano la risposta motoria *R* di premere il freno, una implementazione fisica dell’azione *A*:

$$[SD-Rosso] \rightarrow \text{processi neuronali} \rightarrow R$$

Che relazione c’è tra le due descrizioni (e implicitamente spiegazioni) della mia azione volontaria, l’una formulata nel linguaggio della *folk psychology*, l’altra nel linguaggio della fisica, della chimica e della biologia? Il problema mente-corpo in fondo è tutto qui!

3. Il dualismo ontologico interazionistico

La soluzione data al problema dal dualismo ontologico interazionistico di origine cartesiana, secondo il quale mente e corpo appartengono a due distinti ordini di realtà capaci di interagire causalmente, può essere rappresentata così:

Ciò che la rende volontaria, a loro avviso, è il fatto che la mia decisione cosciente di premere il pedale del freno (un atto mentale secondo loro estraneo al mondo fisico) ne è condizione necessaria. Se bastasse l'attività cerebrale da sola a produrre il movimento del mio piede verso il freno, quel movimento sarebbe un riflesso automatico, non un'azione compiuta di propria libera scelta. La mia libertà di scelta sarebbe annullata anche se la mia risposta motoria avesse due cause concomitanti e entrambe singolarmente sufficienti, l'una mentale e l'altra cerebrale ("causalità sovraderminata"). Anche in questo caso infatti premerei comunque il pedale del freno, indipendentemente dal mio volerlo premere o meno. Anche in questo caso il mio cervello deciderebbe per me! Pertanto, affinché la mia decisione di premere il pedale del freno sia condizione necessaria di quel movimento, occorre non solo che l'attività della mia mente possa interferire con l'attività del cervello, ma anche che possa farlo in modo tale da provocare una interruzione della catena causale di eventi fisici che connette l'*input* sensoriale all'*output* motorio.

Ma quali studi neurologici attestano l'esistenza di questa interferenza della mente cosciente sull'attività del cervello nel processo di coordinamento senso-motorio quando venga eseguita un'azione volontaria? Gli studi esistenti sembrano attestare l'esatto contrario! In particolare un famoso esperimento di B. Libet risalente al 1977 ha dato vita ad una enorme discussione e controversia non solo tra i neuroscienziati ma anche tra i filosofi della mente sull'esistenza o meno del libero arbitrio: secondo Libet ed altri neuroscienziati la presa di coscienza di muovere un dito è successiva alla comparsa nel cervello di *readiness potentials* che sembrano esserne la vera causa, puramente fisica ed inconscia, antecedente alla decisione cosciente di compiere quel movimento (Libet 2004, Haggard e Libet 2001, Soon *et alii* 2008, Matsushashi e Hallet 2008; ma tra i tanti contributi critici cfr. soprattutto Dennett 2003 e Swinburne 2013).

L'interpretazione da dare all'esperimento di Libet è molto difficile e controversa; ma è indubbio che – sebbene Libet stesso abbia optato, in modo in genere giudicato poco convincente, per un'interpretazione in parte favorevole al dualismo, sostenendo che la mente, se non può decidere che cosa fare, può tuttavia porre

un veto alla “decisione” presa in precedenza dal cervello – tale esperimento pone un problema molto serio ai dualisti cartesiani, perché, se si accetta l’identificazione cartesiana tra mente e coscienza, non c’è più spazio per sostenere che, sebbene io divenga consapevole in ritardo della “decisione” presa dal cervello di compiere una certa azione, quella “decisione”, essendo essa identica ad una attività del mio cervello, sarebbe nondimeno una mia decisione. Per un cartesiano infatti la mia decisione di compiere una certa azione o è cosciente oppure non è affatto una mia decisione, è solo un processo fisico che, pur avvenendo nel mio cervello, nulla ha a che fare con me!

Ad ogni modo, si giudichi l’esperimento di Libet come si vuole, il dualismo ontologico interazionistico viene oggi rifiutato dalla maggioranza dei filosofi analitici e dei neuroscienziati. La ragione fondamentale di questo rifiuto risiede nel fatto che tale teoria è incompatibile con una concezione scientifica del mondo. Infatti il dualismo ontologico interazionistico richiede che la mente, in quanto sostanza o proprietà non-fisica, possa interferire nel cervello con la catena causale di eventi fisici che connette lo stimolo sensoriale alla risposta motoria in violazione del “principio di chiusura causale del mondo fisico” (Yalowitz 2012) e quindi in violazione anche del principio fisico di conservazione dell’energia.

4. Il dualismo parallelistico e l’epifenomenismo

Nella storia della filosofia sono reperibili delle forme non interazionistiche di dualismo ontologico che consentono di evitare la violazione della chiusura del mondo fisico. Mi riferisco al “dualismo parallelistico” e all’“epifenomenismo”. Secondo il dualismo parallelistico, sebbene mente e corpo vengano concepiti come due sostanze o proprietà metafisicamente distinte e incapaci di agire causalmente l’una sull’altra, gli atti mentali e gli eventi fisici si accordano perfettamente gli uni agli altri o per volontà divina (si veda l’“armonia prestabilita” di Leibniz 1710) o perché essi sono in realtà manifestazioni parallele e reciprocamente indipendenti di una medesima sostanza in se stessa né fisica né mentale (si pensi alle versioni del “monismo neutrale”, pur così diverse tra loro, proposte in epoche diverse da Spinoza 1677, Russell 1921 e Chalmers 1996).

Di fronte al fatto che il dualismo parallelistico – per rendere conto della perfetta sincronia tra ciò che accade nella mente e ciò che accade nel cervello – debba ricorrere o ad ipotesi metafisico-teologiche come la summenzionata “armonia prestabilita” di Leibniz o ad una ipotesi così speculativa ed oscura come il monismo neutrale riesce difficile sfuggire all’impressione che il dualismo parallelistico, sebbene le sue varie formulazioni siano molto sofisticate e alcune di esse siano state proposte da grandi filosofi, sia in realtà privo di sufficienti riscontri empirici e che quindi non sia recepibile in una concezione scientifica del mondo.

L’epifenomenismo sfugge invece a questa critica, perché dal punto di vista empirico esso è indistinguibile dal fisicalismo: secondo entrambe le teorie il comportamento è guidato interamente dal cervello. Infatti secondo gli epifenomenisti l’attività cerebrale produce sì stati mentali, ma la mente è poi incapace di retroagire causalmente sul cervello, che quindi resta l’unico “autore” di tutto il nostro comportamento. Ma l’epifenomenismo, sebbene sia indistinguibile empiricamente dal fisicalismo e quindi non violi il principio di chiusura causale del mondo fisico, incorre in un’altra grave difficoltà. Non concedendo agli stati mentali, in quanto eventi estranei al mondo fisico, alcun ruolo nella spiegazione del comportamento, si espone – a differenza del fisicalismo, come chiarirò tra poco – al “rasoio di Occam”: gli enti non devono essere moltiplicati oltre il necessario. Perché ipotizzare l’esistenza di stati mentali non fisici se tale ipotesi è del tutto inutile da un punto di vista empirico-scientifico? In conclusione il dualismo ontologico in tutte e tre le sue principali versioni – interazionismo, parallelismo e epifenomenismo – è incompatibile con una concezione scientifica del mondo o quanto meno è ad essa estraneo.

5. Il fisicalismo

La principale alternativa al dualismo mente-corpo è il fisicalismo. Facendo riferimento al mio esempio precedente, esso può essere così rappresentato:

della teoria dell'identità, a partire da Place (1956), ritengono che almeno in linea di principio tutti gli stati mentali siano riducibili a processi cerebrali. Gli eliminativisti (o sostenitori del "materialismo eliminativo"), a partire da Paul Churchland (1989), pensano invece che molti stati mentali non siano riducibili a processi cerebrali neppure in linea di principio. Ma ciò dipende a loro avviso semplicemente dal fatto che tali processi non esistono. I loro concetti devono essere perciò eliminati dal linguaggio scientifico al pari di quanto è avvenuto ad esempio in chimica alla fine del Settecento con il concetto di flogisto. In modo analogo vecchi concetti della *folk psychology* come ad esempio "credenza" o "desiderio" dovranno essere sostituiti da nuovi concetti generati da un processo di co-evoluzione tra psicologia cognitiva e neuroscienze (cfr. la "neurofilosofia" di Patricia Churchland 1986 e 2002; cfr. anche Nannini 2010).

Il fisicalismo è stato fatto proprio, in modo più o meno esplicito, negli ultimi venticinque anni da quasi tutti quei neuroscienziati che, a cavallo tra scienza e filosofia, hanno dato vita ad una «epistemologia basata sul cervello» (Edelman 2006, p. XVI). Ad esempio A. Damasio, in polemica con il dualismo cartesiano, ha intitolato un suo libro *Descartes' Error* (Damasio 1994) e F. Crick ha difeso il fisicalismo con queste parole: «L'ipotesi straordinaria è che proprio "Tu", con le tue gioie e i tuoi dolori, i tuoi ricordi e le tue ambizioni, il tuo senso di identità personale e il tuo libero arbitrio, in realtà non sei altro che il *comportamento* di un vasto insieme di cellule nervose e delle molecole in esse contenute. Come avrebbe detto l'Alice di Lewis Carroll "Non sei altro che un pacchetto di neuroni"» (Crick 1994, trad. it. modificata p. 17, corsivo dell'autore).

Tuttavia i fisicalisti devono fronteggiare un'obiezione che li ha quasi sempre condannati, dall'antichità ad oggi, ad essere in minoranza sia tra i filosofi sia tra gli scienziati: il *gap* fenomenologico esistente tra gli stati mentali ed i loro correlati neuronali sembra rendere del tutto implausibile la riducibilità ontologica degli uni agli altri. Perché ad esempio, quando certi coni della retina reagiscono ad onde elettromagnetiche aventi una lunghezza compresa tra i 400 e i 450 nanometri, noi vediamo una macchia blu? Che cosa ha in comune una certa lunghezza d'onda con un certo colore? E, più in generale, che cosa hanno in comune le proprietà fisiche di qualsiasi correlato neuronale con quei *qualia* (colori, sapori, suoni ecc., ma anche dolore,

piacere ed altre emozioni) che sono oggetto della nostra esperienza cosciente e soggettiva? È mai possibile che cose che appaiono così diverse siano in realtà la stessa cosa? Chi fosse capace di rispondere a questa domanda saprebbe risolvere quel problema che è stato chiamato da Chalmers (1996), appunto, lo *Hard Problem* (cfr. *supra* capitolo primo).

6. *Il funzionalismo e altre alternative al dualismo e al fisicalismo*

Per non incorrere nelle difficoltà poste dallo *Hard Problem* molti filosofi, psicologi e neuroscienziati hanno cercato nel corso del Novecento e cercano tuttora delle soluzioni del problema mente-corpo che consentano loro di liberarsi del fisicalismo senza ricadere nel dualismo ontologico. Alcuni – detti in genere “neomisteriani”, o semplicemente scettici (ad es. McGinn 1982 e 1991) – ritengono di potersi liberare dello *Hard Problem* asserendo che l’emergere della mente (e più precisamente della coscienza) dall’attività cerebrale è un mistero insondabile: la scienza si occupa di fatti oggettivi descrivibili in terza persona; e quindi non potrà mai comprendere la coscienza, ossia le esperienze soggettive, che sono accessibili solo introspettivamente e sono descrivibili solo in prima persona (Nagel 1974). Altri (ad es. Smith *et alii* 2005) – appoggiandosi ora alla fenomenologia husserliana ora al buddismo o ad altre religioni o filosofie speculative – sostengono, detto in breve, che la fisica e le altre scienze naturali non godono di alcun privilegio nel rappresentare il mondo. La scienza è una «forma simbolica» tra le altre, al pari del mito o dell’arte (Cassirer 1923-1929). Le teorie della psicologia e delle scienze umane in generale, formulate in prima persona come propongono soprattutto i fenomenologi, godono di una propria autonoma validità su un piano di parità con le teorie in terza persona delle neuroscienze e delle altre scienze naturali (cfr. al riguardo anche Putnam 2016 e il “naturalismo liberalizzato” di De Caro e Mcarthur 2004). Analoga è la tesi di quei neurologi e filosofi che sono concordi nel sostenere che psicologia e neuroscienze si occupano di oggetti incommensurabili e che è un “errore categoriale” confondere le attività della mente con quelle del cervello. Ad esempio è insensato dire che un cervello pensa. Solo le persone possono pensare, i neuroni del cervello possono unicamente depolarizzarsi, ma non pensare (Bennett e Hacker 2003).

Ciascuna di queste alternative sia al fisicalismo sia al dualismo

cartesiano meriterebbe un maggiore approfondimento. Tuttavia esse nel complesso si basano su una separazione della filosofia e delle scienze umane dalle scienze naturali estranea a quella continuità tra scienza e filosofia che è all'origine dell'incontro odierno tra filosofia della mente e neuroscienza cognitiva.

Considerazioni analoghe possono essere fatte anche a proposito dell'idealismo, una delle correnti filosofiche dominanti dalla fine del Settecento fino alla prima metà del Novecento. In estrema sintesi secondo l'idealismo, quando si mettano in parentesi le profondissime differenze che pur esistono tra le sue varie versioni, ciò che è veramente reale è solo lo spirito; i corpi esistono solo in quanto sono oggetto di percezioni, di pensieri o teorie scientifiche. Questa teoria filosofica è oggi in profondo declino perché, sebbene certe interpretazioni della meccanica quantistica tornino implicitamente a rivalutarla (cfr. di nuovo Chalmers 1996 e 2010), è anch'essa, al pari delle teorie prima menzionate, in contrasto con almeno uno dei presupposti di tutta la scienza moderna: il realismo scientifico (cfr. *infra* capitolo quinto). Lo scienziato deve infatti andare in cerca di una descrizione e spiegazione del mondo che sia la più obiettiva possibile, ossia deve andare in cerca di una teoria sul mondo che sia la più indipendente possibile da ogni punto di vista particolare che per ragioni personali, culturali o specie-specifiche si possa avere su di esso.

Di conseguenza l'unica vera "terza via" ancor oggi vitale tra dualismo ontologico e fisicalismo è il funzionalismo (o cognitivismo). Questa soluzione del problema mente-corpo pone in parentesi la questione se gli stati mentali siano riducibili o meno a processi cerebrali e afferma che, anche ammesso che lo siano, essi possono essere descritti e spiegati solo in termini psicologico-funzionali e non in termini neurologici. I funzionalisti cercano di conciliare il monismo ontologico con il dualismo epistemologico. In altre parole secondo i funzionalisti per identificare uno stato mentale è sufficiente ridescriverlo funzionalmente mediante le sue cause ed i suoi effetti, considerandolo come un passo intermedio tra lo stimolo sensoriale e la risposta motoria in un diagramma di flusso che ricostruisca il modo nel quale viene elaborata l'informazione contenuta nell'*input* sensoriale (oppure endo-psichico) per determinare causalmente la risposta motoria (inclusi ovviamente i movimenti dell'apparato fonatorio). In genere i funzionalisti non sono contrari in linea di principio all'idea che gli

stati funzionali a cui essi riducono gli stati mentali siano implementati dall'attività del cervello, ma pensano anche che scoprire quale sia questa implementazione sia per un verso superfluo e per un altro comunque praticamente impossibile, dato che, a causa delle “molteplici realizzazioni”, ogni stato mentale può essere implementato da un numero indefinito di processi cerebrali diversi (cfr. in particolare lo studio ormai classico di Fodor 1975).

Il contributo dato dal funzionalismo alle scienze cognitive non può essere sottovalutato. E nessuna soluzione del problema mente-corpo che sia minimamente credibile può permettersi di accettare accanto al riduzionismo ontologico anche un completo riduzionismo epistemologico: anche chi sostiene che gli stati mentali sono processi cerebrali non può difendere seriamente, almeno per ora, la tesi che tutta la psicologia sarebbe integralmente traducibile nel linguaggio delle neuroscienze (Legrenzi e Umiltà 2009). Tuttavia una completa separazione tra le due scienze è altrettanto inaccettabile. Ridescrivere uno stato mentale semplicemente come uno stato capace di svolgere nel cervello una determinata funzione non è sufficiente per giustificare l'esistenza. Altrimenti dei chimici che volessero riproporre oggi le teorie di G.E. Stahl, abbandonate grazie a A.L. Lavoisier già alla fine del Settecento, potrebbero rivendicare l'esistenza del flogisto semplicemente attribuendogli lo svolgimento di tutte quelle funzioni che attualmente vengono riconosciute all'ossigeno, senza però essere in grado di spiegare come il flogisto stesso, per come veniva inteso da Stahl, possa mai essere capace di assolvere ad un tale compito! In conclusione ciò che serve è una forma di fisicalismo che proponga una soluzione dello *Hard Problem* scientificamente fondata.

7. L'emergentismo

Si può essere tentati dal credere che una soluzione di tal genere sia offerta oggi dall'“emergentismo” (ad es. Beckermann *et alii* 1992, Stephan 1999, O'Connor e Wong 2015). Secondo gli emergentisti il cervello produce grazie alla sua attività fisica una entità di nuovo tipo, la coscienza, che gode di una sua autonoma realtà rispetto al mondo fisico da cui emerge e su cui può perciò retroagire causalmente. Ma in realtà questa soluzione al problema mente-corpo è ambigua, perché può essere interpretata in tre modi differenti corrispondenti rispettivamente al dualismo, al funzionalismo e al fisicalismo:

- “Emergentismo dualistico”: la coscienza, pur emergendo dall’attività cerebrale, dopo essere emersa viene a costituire un livello di realtà non appartenente al mondo fisico. Pertanto essa, quando retroagisce sul cervello, viola il principio di chiusura del mondo fisico stesso e va incontro a tutte le obiezioni che possono essere rivolte al dualismo ontologico interazionistico.
- “Emergentismo funzionalistico”: la coscienza “emerge” dall’attività cerebrale solo nel senso che essa è una realtà virtuale che si presume sia implementata in qualche modo dall’attività cerebrale stessa, ma si ignora (almeno per ora o forse per sempre) come ciò sia possibile. Questa versione debole dell’emergentismo coincide di fatto con il funzionalismo (e in parte con il neomisterianesimo) e va incontro alle medesime obiezioni. In particolare essa separa eccessivamente la psicologia cognitiva dalle neuroscienze.
- “Emergentismo fisicalistico”: la mente, inclusa la coscienza, è un insieme di *higher order properties* della dinamica cerebrale che emergono dall’attività di singoli neuroni così come il calore di un gas emerge dal moto delle sue singole molecole in quanto energia cinetica media dell’insieme dei loro movimenti. Secondo questa interpretazione la mente, pur essendo un insieme di proprietà sistemiche del cervello distinguibili dall’attività locale dei singoli neuroni, non viola il principio di chiusura del mondo fisico allorché retro-agisce su di essi perché è essa stessa un insieme di proprietà fisiche. L’emergentismo così inteso è perciò perfettamente compatibile con il fisicalismo e contiene anzi un’idea di grande interesse che cercherò in seguito di chiarire e sviluppare: gli stati mentali sono ontologicamente riducibili non direttamente all’attività di singoli neuroni, bensì a certe proprietà sistemiche della loro dinamica complessiva. Ciò permette di ipotizzare che la relazione tra l’attività dei singoli neuroni e la dinamica cerebrale nel suo complesso sia interpretabile come una interazione di “causalità sistemica” nelle due direzioni: *bottom-up causation* dagli elementi del sistema (i neuroni) al sistema nel suo complesso (dinamica cerebrale) e *top-down causation* nella direzione opposta. Tuttavia occorre precisare che questa forma

di emergentismo è molto controversa e si presta ad interpretazioni molto diverse. Molti studiosi dei sistemi complessi interpretano la relazione *bottom-up* come una relazione concettuale di ridescrizione da concreto a astratto: la dinamica cerebrale secondo loro non è un effetto dell'interazione tra i singoli neuroni, è semplicemente tale interazione ridescritta in termini più astratti. Pertanto – essi concludono – non esiste alcuna differenza ontologica tra l'attività dei singoli neuroni e la dinamica cerebrale complessiva e dunque non è concepibile alcuna *top-down causation* dalle proprietà sistemiche della dinamica cerebrale all'attività dei singoli neuroni (ad es. Lindemann 2014, section 3c). Questi studiosi, riguardo al problema mente-corpo, sono dei riduzionisti tanto ontologici quanto epistemologici e rifiutano di fatto ogni forma di emergentismo. All'opposto altri fisici o biologi, basandosi sull'interazione tra *hardware* e *software*, interpretano entrambe le relazioni – degli elementi con il sistema e del sistema con gli elementi – come relazioni causali intercorrenti tra due distinti livelli di realtà, venendo con ciò a sostenere una forma forte di emergentismo che coincide di fatto con l'emergentismo dualistico (Ellis 2016), con tutti i problemi che questo comporta. Esiste tuttavia una interpretazione intermedia – almeno implicitamente già presente in Edelman e Tononi (2000) – secondo la quale l'interazione tra un sistema ed i suoi elementi può essere autenticamente causale in entrambe le direzioni se si ammette la possibilità di una “causalità immanente” compatibile con il riduzionismo ontologico. Si pensi ad esempio al processo graduale di assestamento che può avvenire nelle “macchine di Boltzmann”, reti neurali artificiali nelle quali l'attività di ogni singolo neurone del sistema provoca dei mutamenti in tutti gli altri e viceversa, finché mediante successive oscillazioni il sistema stesso giunge provvisoriamente ad uno stato di equilibrio (Hinton e Sejnowski 1986). Dato che l'assestamento avviene gradualmente nel tempo, se ne deve concludere che l'interazione tra ogni singolo neurone e l'intero sistema è effettivamente di tipo causal-sistemico in entrambe le direzioni, ma ciò non impedisce che ontologicamente il sistema sia identico all'insieme dei suoi elementi. Infatti i neuroni inte-

ressati sono sempre gli stessi, ma gli stati in cui essi si trovano nelle successive tappe del processo di assestamento sono invece diversi. In conclusione solo questa terza interpretazione dell'emergentismo fisicalistico consente di ammettere un fenomeno di emergenza dall'attività dei singoli neuroni alle proprietà sistemiche della dinamica cerebrale senza ricadere nel dualismo mente-corpo e permette così di evitare che si divenga vittime del trucco su cui si basa l'odierno ingannevole successo dell'emergentismo antiriduzionistico: confondere la versione dualistica di questa teoria con la sua versione fisicalistica, credendo di essere riusciti in tal modo a conciliare il dualismo, comprese le sue implicazioni teologiche (evidenti ad esempio in Ellis), con una concezione scientifica del mondo.

8. Intenzionalità, autocoscienza e coscienza

Tenendo conto delle critiche che possono essere rivolte alle soluzioni date al problema mente-corpo in alternativa al fisicalismo, le escludo tutte e propongo qui un abbozzo di soluzione che sia viceversa naturalistico-fisicalistica. Essa si basa su un'analisi concettuale tesa a mettere in luce quali siano i problemi epistemologici che, se non chiariti, rischiano di rendere impossibile ogni soluzione dello *Hard Problem* fondata sulla continuità tra scienza e filosofia. Non pretendo con ciò di proporre una soluzione scientifica definitiva di tale problema. Mi riprometto soltanto di rimuovere alcuni ostacoli concettuali che sembrano rendere inconcepibile anche il solo andarne in cerca.

La prima difficoltà che deve affrontare un naturalista fisicalista è chiedersi anzitutto quali mai possano essere i correlati neurali degli stati di coscienza a cui questi stati stessi sono ontologicamente riducibili. Esiste una qualche attività del mio cervello che può essere ontologicamente identica, ad esempio, alla mia percezione cosciente del fatto che il semaforo davanti a me è rosso? Questa percezione è una esperienza vissuta nella quale vengono a fondersi tre sue differenti caratteristiche: l'Intenzionalità brentaniana (ossia l'avere un contenuto), l'autocoscienza e la coscienza. Vedo che il semaforo è rosso (e non verde ad esempio), so che sono io a vederlo e, essendo sveglio e attento, sono consapevole di tutto ciò. Ora, queste tre caratteristiche sono fuse nel mio cervello come lo sono nella mia esperienza soggettiva oppure

sono invece implementate da processi cerebrali distinti e in linea di principio separabili?

La seconda ipotesi sembra essere la più probabile, dato che in pazienti che soffrano di certe patologie neurologiche queste tre caratteristiche compaiono in modo separato. Ad esempio gli ormai notissimi casi di *blindsight* ci dicono che dei soggetti che abbiano subito una commisurotomia, pur non essendo consapevoli (a differenza dei soggetti normali) della comparsa nel loro emi-campo visivo sinistro di una barra o verticale o orizzontale, riescono a indovinare l'orientamento della barra stessa – se forzati a pronunciarsi al riguardo – con una percentuale nettamente superiore al caso: prova evidente questa dell'esistenza, nel loro emisfero destro, di una percezione “visiva” (ossia visiva dal punto di vista funzionale) non accompagnata da coscienza fenomenica, o per lo meno non accompagnata da una percezione cosciente completa e normale (Weiskrantz 2009). All'inverso coloro che soffrono della sindrome di Anton-Babinski, pur essendo di fatto ciechi, hanno l'impressione di non esserlo; ossia si sentono coscienti di percezioni che in realtà non hanno (coscienza fenomenica questa, ovviamente ingannevole, di stati Intenzionali in realtà inesistenti da un punto di vista funzionale) (Prigatano e Schacter 1991, Carvajal *et alii* (2012)). Intenzionalità e coscienza fenomenica, se questi casi clinici sono – come sembra – attendibili, costituiscono dunque fenomeni neurologicamente distinti e, talvolta, separati.

In modo analogo la sindrome della “mano aliena” mostra come, per colpa di una specifica patologia psico-neurologica, dei soggetti possano comportarsi come se le loro due mani fossero guidate da due Io diversi aventi talvolta intenzioni opposte, sebbene questo strano comportamento compaia in soggetti che in quel momento sono pienamente coscienti (svegli e attenti) in modo del tutto normale (Biran e Chatterjee 2004, Biran *et alii* 2006). In questo caso è l'Io a presentarsi come un fenomeno cerebrale almeno in parte distinto dalla coscienza fenomenica. Pertanto, volendo andare in cerca di una riduzione funzionale di stati mentali che siano Intenzionali, autocoscienti e coscienti, converrà avanzare ipotesi funzionali distinte per queste tre caratteristiche, in modo da poter trovare poi loro implementazioni neurologiche altrettanto distinte, sebbene non prive di interazioni.

9. *Le rappresentazioni mentali*

Prima di avanzare qualsiasi ipotesi sulla naturalizzazione dell'Intenzionalità, conviene preliminarmente cercare di definire che cosa sia uno stato Intenzionale o rappresentazione mentale che dir si voglia. E, per raggiungere questo obiettivo, è opportuno chiedersi che cosa sia una rappresentazione in generale. In prima approssimazione una rappresentazione, al pari di un segno, è – come dicevano gli scolastici – *aliquid quod stat pro aliquo*, qualcosa che sta per qualcos'altro. Pertanto una rappresentazione è costituita da un Rappresentante che sta per un Rappresentato, reale o immaginario, sia quest'ultimo un oggetto materiale, una persona, uno stato di cose, un evento o un'astrazione. In ogni caso si pone anzitutto il problema della natura di quello "sta per", ossia della relazione semantica che lega il Rappresentante al Rappresentato.

La teoria più vicina al senso comune riguardo alla natura delle relazioni semantiche è quella che le considera come delle relazioni per somiglianza: la foto del mio gatto rappresenta il gatto perché gli assomiglia. Ma anche in un caso privilegiato come questo sembra difficile sostenere che il Rappresentante si riferisce al Rappresentato senza alcuna mediazione. Mi sarebbe davvero possibile notare la somiglianza tra il mio gatto e la sua foto se il vedere questi due oggetti reali esterni non evocasse in me una medesima immagine interna, più in generale una medesima rappresentazione mentale?

Tuttavia il concetto di rappresentazione mentale non è esente da oscurità e inoltre, se si cerca di risolvere il problema di come i Rappresentanti si riferiscano ai loro Rappresentati tramite l'introduzione delle rappresentazioni mentali in un ruolo di mediazione, si corre il rischio di cadere in un regresso all'infinito: che cosa medierà infatti l'immagine interna del gatto con il gatto stesso se non una ulteriore meta-rappresentazione (anch'essa interna) e così via? È per evitare questa evidente fallacia che molti filosofi del linguaggio hanno ritenuto che, per chiarire che cosa sia una relazione semantica, si possa e si debba fare completamente a meno di ogni mediazione interna alla mente tra Rappresentante e Rappresentato, senza al contempo fare ricorso ad alcuna relazione di somiglianza tra di essi. Alcuni di loro hanno perciò pensato che il Rappresentante si riferisca al Rappresentato non perché ne sia un'immagine, bensì perché ne è un sostituto e un indicatore (come lo è lo scontrino per ritrovare il cappotto al

guardaroba, sebbene non gli assomigli affatto). Da qui un insieme di teorie, di origine prevalentemente comportamentistica e logico-empiristica, sulla natura della relazione semantica che vengono etichettate in genere come “esternalistiche”, la principale delle quali è la “teoria del riferimento diretto” (Kripke 1971); una teoria secondo la quale il Rappresentato, in quanto riferimento esterno della rappresentazione, conferisce un significato al Rappresentante in virtù di una convenzione arbitraria senza bisogno di alcuna mediazione da parte di qualsivoglia rappresentazione mentale. In tal modo si evita, per spiegare come possano instaurarsi delle relazioni semantiche tra un segno ed il suo riferimento, ogni ricorso alla mediazione di entità o platonizzanti, come i celebri “sensi” di Frege (1892), e quindi troppo speculative, o viceversa sì introspettivamente esperibili in modo diretto, ma mal definite e troppo soggettive per dare ai termini del linguaggio un riferimento intersoggettivamente stabile²⁴.

Tuttavia, quali che siano i pregi dell’esternalismo in filosofia del linguaggio ed in epistemologia, tale teoria, quando venga usata in filosofia della mente per negare il ruolo svolto dalle rappresentazioni mentali nella comprensione delle relazioni semantiche, diviene poco plausibile, soprattutto allorché venga applicata alle immagini. È infatti nel linguaggio e in altri sistemi semantici artificiali che i segni devono il loro significato unicamente a convenzioni che o sono un prodotto storico-culturale variabile da epoca a epoca e da popolo a popolo, come avviene per le lingue naturali, oppure sono il frutto di scelte arbitrarie delle autorità competenti, come avviene per i cartelli stradali. Ma il fatto che, quando riconosco in una foto il mio gatto, vedo quella certa cosa che poi io chiamo “gatto” (mentre altri lo chiamano *cat*, *Katze*, *chat* o in altro modo), e non, che so io, una certa altra cosa che io chiamo “cammello” (una cosa che appare a tutti comunque diversa dalla prima cosa indipendentemente da come la chiamano),

²⁴ Applicando alla filosofia della mente concetti nati entro la filosofia del linguaggio, identifico sostanzialmente l’esternalismo (alcuni preferiscono dire esternismo) con le dottrine antirappresentazionalistiche che o negano l’esistenza delle rappresentazioni mentali o le considerano comunque irrilevanti per chiarire come un Rappresentante si riferisca al suo Rappresentato, mentre chiamo internalistiche (al solito alcuni dicono internistiche) le teorie che presentano al riguardo un’ipotesi opposta. Non è questo ovviamente l’unico modo possibile per caratterizzare l’opposizione tra esternalismo e internalismo: molto utile per chiarire tutta la discussione Dellantonio (2007).

non sembra essere né un tratto culturale variabile da popolo a popolo né tanto meno il frutto di qualche scelta arbitraria. In altre parole non bisogna confondere ciò che si vede (determinato dall'interazione fisica tra la realtà esterna e l'apparato visivo umano) con ciò che si dice di vedere (determinato invece dall'interpretazione che una certa cultura e una certa lingua inducono a dare di ciò che si vede).

Data l'esistenza di queste difficoltà della teoria del riferimento diretto, si può pensare allora di difendere l'esternalismo identificando le relazioni semantiche con delle relazioni non per convenzione, bensì per contiguità o causalità. Nel primo caso si può ritenere che, per insegnare ad uno straniero il significato della parola "mela", sia sufficiente pronunciarla mentre si punta con il dito ad una mela a noi vicina (contiguità). Nel secondo caso si può sostenere che una certa foto è la foto del mio gatto perché certe macchie di colore si sono formate su un certo pezzo di carta a causa del fatto che proprio il mio gatto, e non un'altra cosa, si trovava davanti alla macchina fotografica quando è stata scattata la foto stessa (causalità). Tuttavia neppure queste altre forme di esternalismo appaiono convincenti nel loro tentativo di fare a meno di qualsiasi rappresentazione mentale per connettere un Rappresentante al suo Rappresentato. Le relazioni semantiche per contiguità o causalità infatti, oltre ad essere poco adatte per spiegare come un segno possa riferirsi ad un oggetto astratto o inesistente, non sfuggono, al pari di quelle per somiglianza, alla necessità di un processo di selezione dei possibili Rappresentati di un medesimo Rappresentante; processo che richiede sempre l'intervento attivo di un osservatore teso a guardare alla realtà esterna in un certo modo tramite una qualche sua rappresentazione mentale interna. Come potrebbe uno straniero che sta imparando l'italiano capire che è proprio la mela e non il suo colore o la sua buccia ciò a cui mi riferisco quando, puntando il dito verso di essa, dico «Questa è una mela!» se non condividesse con me, almeno in parte, un certo sistema di rappresentazioni mentali, alla luce del quale soltanto appare ovvio che, puntando il dito verso la mela, mi sto riferendo alla mela nella sua interezza e non alla sua buccia²⁵? E, per ragioni analoghe, il fatto che la presenza del mio

²⁵ Si pensi al classico esempio analogo fornito da Quine (1960, p. 41 sgg.): come potrebbe capire l'etnografo – sulla sola base del fatto che l'indigeno australiano, messo in presenza di un coniglio, dice "gavagai" – che per l'indigeno "gavagai" si-

gatto abbia contribuito a causare il formarsi di un certo insieme di macchie colorate su un certo pezzo di carta sarebbe sufficiente a fare di quel pezzo di carta una foto del mio gatto (e non, ad esempio, una foto della macchina fotografica o dell'attrez-zatura per lo sviluppo, anch'esse concause della foto stessa al pari della presenza del gatto) se io non possedessi una rappresentazione mentale interna (qualunque cosa essa sia) del mio gatto che mi consente di riconoscerlo in quelle macchie colorate?

È per evitare queste ed altre consimili difficoltà incontrate dagli esternalisti che altri filosofi e scienziati cognitivi hanno difeso una concezione internalistica delle relazioni semantiche; una concezione che richiede l'esistenza di rappresentazioni mentali come mediatrici indispensabili, in ogni rappresentazione, della relazione semantica tra Rappresentante e Rappresentato. Gli internalisti devono però guardarsi dal rischio di riproporre, nel tentativo di chiarire il rapporto intercorrente tra la rappresentazione mentale interna e l'oggetto a cui essa si riferisce, tutte le difficoltà precedentemente illustrate a proposito delle relazioni semantiche in generale. Se infatti una rappresentazione mentale, quale che sia il suo supporto "interno" – vale a dire sia essa una certa modificazione di una misteriosa sostanza immateriale o sia invece un processo cerebrale –, si riferisse al suo oggetto in virtù di una relazione per somiglianza, contiguità o causalità, si riproporrebbero tutte le difficoltà precedentemente esaminate e si incorrerebbe in più in un regresso all'infinito, dato che ogni rappresentazione mentale richiederebbe una meta-rappresentazione anch'essa mentale e così via.

Si comprende facilmente perciò perché l'internalismo, per sottrarsi a tali rischi, tenda spontaneamente a trovare il suo fondamento nel concetto brentano d'Intenzionalità. Secondo quest'ultimo – com'è noto – gli atti mentali sono irriducibili a processi fisici, perché ciò che è mentale possiede la proprietà intrinseca e primitiva di riferirsi ad altro da sé. È questa intrinseca tendenza del mentale a riferirsi ad altro da sé che spiega immediatamente, senza bisogno di ricorrere a nessuna meta-rappresentazione, perché una percezione sia sempre

gnifica proprio "coniglio" e non "parte non distaccata di coniglio" o "sezione temporale di coniglià" se non presupponesse che lui e l'indigeno condividono le medesime «ipotesi analitiche» di sfondo (*ivi*, pp. 89-94)?

percezione di qualcosa e un pensiero abbia sempre un contenuto. In altre parole, secondo gli “internalisti Intenzionalisti” la relazione semantica che in una rappresentazione mentale lega il Rappresentante al suo Rappresentato non è né per convenzione né per somiglianza né per contiguità; e neppure coincide con una relazione di causa-effetto. È semplicemente una relazione primitiva e irriducibile; una relazione che non è ulteriormente spiegabile, ma che non ha neppure bisogno di spiegazione, perché è intuitiva ed evidente. È una relazione intrinseca al mentale nella sua differenza ontologica rispetto a tutto ciò che è meramente fisico.

Tale internalismo Intenzionalistico si fonda perciò sul dualismo ontologico – cartesiano e, insieme, brentano – tra mente e corpo. E questa implicazione metafisica è un prezzo che nessun filosofo della mente o scienziato cognitivo di orientamento naturalistico dovrebbe essere disposto a pagare. Per uscire da questa *impasse* – o esternalismo o internalismo Intenzionalistico –, occorre perciò delineare una forma di internalismo compatibile con il naturalismo, ossia occorre tracciare una strada che porti alla naturalizzazione dell’Intenzionalità degli stati mentali, considerata provvisoriamente per le ragioni sopra esposte in modo isolato rispetto alle altre due caratteristiche fondamentali di ogni stato mentale: il suo essere uno stato di coscienza e di autocoscienza.

10. La riduzione ontologica dell’Intenzionalità

Per poter fare delle ipotesi sulla naturalizzazione dell’Intenzionalità, conviene perciò isolare nell’esempio precedente del mio fermare la macchina davanti ad un semaforo rosso l’Intenzionalità della mia percezione, separandola dalla sua soggettività e dal suo essere uno stato di coscienza. In altre parole, mettendo provvisoriamente in parentesi nella ricostruzione filosofica della mia percezione il suo carattere cosciente (*coscientemente*) e autocosciente (*Io**), si ottiene come residuo uno stato meramente Intenzionale, ossia un “vedere” (non cosciente e impersonale) che ha come proprio contenuto il fatto che il semaforo appare essere rosso (*[Rosso*]*) e che proprio per ciò si riferisce al (ossia ha una relazione di Intenzionalità ($\Rightarrow$) col) fatto esterno che il semaforo è effettivamente rosso (ossia *[Rosso]* secondo il senso comune):

$$[Vedere([Rosso*])] \Rightarrow [Rosso]$$

La messa in parentesi, provvisoriamente, del carattere cosciente e soggettivo della mia percezione consente dunque di isolare in essa uno stato interno che presenta notevoli somiglianze con gli stati interni, ad esempio, dei congegni che guidano i prototipi sperimentali delle auto senza pilota. Queste auto, pur non essendo ovviamente né coscienti né autocoscienti, nondimeno per poter funzionare (per quanto ciò, almeno per ora, è possibile) devono essere dotate di un sistema elettronico di guida che consenta loro di fermarsi quando si trovino davanti ad un semaforo rosso. Il sistema, ricevuto un segnale da una telecamera o da un sensore di altro tipo puntato sul semaforo, è in grado (o almeno deve essere in grado, se la macchina è davvero *self-driving*!) di trasformarlo in uno stato interno che a sua volta attiverà i freni. Comunque lo si voglia chiamare e descrivere, questo stato interno del sistema che guida l'auto è dal punto di vista funzionale un "vedere" che il semaforo è rosso.

Ora, nella prospettiva riduzionistica qui adottata e limitatamente all'azione di fermarsi di fronte ad un semaforo rosso, non c'è alcuna differenza sostanziale tra la funzione svolta dal sistema di guida della macchina e quella svolta dal cervello umano (salvo il fatto che per ora quest'ultimo, in quanto cosciente e autocosciente, è molto più affidabile e versatile, ossia è capace di adattarsi rapidamente a situazioni ambigue o imprevedibili). In entrambi i casi la "percezione funzionale" che il semaforo è rosso è implementata da processi fisici (o fisico-chimici). Pertanto la ricostruzione filosofica della mia percezione come mero stato Intenzionale può essere così formulata:

Ricostruzione filosofica della relazione di Intenzionalità

$$[Rosso] \rightarrow [Vedere([Rosso*])] \Rightarrow [Rosso]$$

Ma – ecco una prima cruciale difficoltà – in che modo uno stato fisico può implementare uno stato Intenzionale (ossia uno stato che in virtù della sua stessa definizione si riferisce a qualcosa d'altro da sé)? Qui il paragone con la macchina senza pilota può aiutarci molto. Se si prende in considerazione non il vedere in quanto atto cosciente

e autocosciente, bensì solo quel “vedere” funzionale (*f-Vedere*) che è comune ad un essere umano ed al sistema di guida di un’auto senza pilota, allora si può dire che qualsiasi stato fisico che sia capace di trasformare l’informazione veicolata dalle onde elettro-magnetiche emesse dal semaforo in un comando atto ad attivare i freni può essere considerato come l’implementazione di un *f-Vedere*. Pertanto, seguendo Dretske (1981 e 1995), si può dare una ridefinizione contestuale di *[Vedere([Rosso*])]* e della sua relazione di Intenzionalità ($\Rightarrow$) con *[Rosso]* mediante una catena di cause ed effetti nella quale *[f-Vedere]* e *f-M* siano delle “riduzioni funzionali” (Kim 1998, pp. 97-103) rispettivamente di *Vedere* e *M*:

**Riduzione funzionale della relazione di Intenzionalità
(prima insufficiente approssimazione)**

$$[Rosso] \rightarrow ([f-Vedere] \& f-M) \rightarrow A$$

In tal modo la relazione d’intenzionalità ($\Rightarrow$) scompare e viene sostituita da una sequenza di relazioni causali ($\rightarrow$) fisicamente implementabili. In questa prima approssimazione della riduzione funzionale della mia percezione scompare però il suo contenuto (ossia *[Rosso*]*) insieme all’effetto che tale contenuto ha sul mio comportamento. Ma ciò non va bene: io mi sono fermato proprio perché il semaforo mi è apparso rosso! Anche se non lo fosse stato ma io lo avessi visto così, mi sarei fermato lo stesso. Le mie azioni volontarie dipendono non da come il mondo effettivamente è, ma da come a me appare. Dall’altro lato, tuttavia, sarebbe inopportuno riassegnare a *[f-Vedere]* questo contenuto, trasformandolo in un *[f-Vedere([Rosso*])]*, perché in tal caso si verrebbe a reintrodurre nella riduzione funzionale di uno stato Intenzionale un *quale* come *[Rosso*]* a cui, a causa dello *Hard Problem*, è difficile assegnare una implementazione fisica. La riduzione funzionale di una percezione propone perciò il seguente dilemma: se non si reintroduce in essa il contenuto percepito, la riduzione stessa manca di qualcosa di essenziale; ma se tale contenuto viene reintrodotta, compare allora nella riduzione funzionale un *quale* la cui implementazione fisico-cerebrale mediante delle relazioni di causa-effetto tra eventi fisici sembra essere impossibile: il fatto che

la mia percezione sia causata da *[Rosso]* non garantisce che il suo contenuto sia *[Rosso*]*. Altrimenti le percezioni ingannevoli e le allucinazioni sarebbero in linea di principio impossibili. Questa difficoltà diviene ancora più evidente quando si cerchi di ridurre funzionalmente non una percezione bensì un pensiero. Infatti il mio pensare a Zeus non può sicuramente essere causato da Zeus, dato che Zeus non esiste!

La soluzione a questo problema è offerta dalla teoria avverbiale della percezione (Sellars 1975, Tye 1984). Nel mio esempio, essendo *[f-Vedere]* uno stato meramente funzionale né cosciente né autocosciente, del suo contenuto si deve tener conto solo in vista della funzione che esso svolge nel processo che collega causalmente lo stimolo distale all'azione. Ma per svolgere tale funzione non c'è bisogno che *[f-Vedere]* abbia un contenuto in senso proprio del quale io sia cosciente; è sufficiente che *[f-Vedere]* sia quel certo modo di “vedere” che, in questo caso, è causato dalla presenza di un semaforo rosso e che causa a sua volta l'azione di frenare. Pertanto, applicando la teoria avverbiale della percezione, si può sostituire *[Rosso*]* con l'operatore avverbiale “rossamente” e trasformare *[Vedere([Rosso*])]* in *[rossamente(f-Vedere)]*. Con ciò il contenuto della percezione visiva viene sostituito da un certo modo di vedere. La riduzione funzionale della percezione che il semaforo è rosso viene così a coincidere (in riferimento sia al mio cervello sia al sistema di guida dell'auto senza pilota) con la seguente catena di cause ed effetti (Nannini 2004 e 2007b):

**Riduzione funzionale della relazione di Intenzionalità
(versione corretta)**

$$[Rosso] \rightarrow ([rossamente(f-Vedere)] \& f-M) \rightarrow A$$

Normalmente gli psicologi cognitivi sono in grado di offrire, mediante dei diagrammi di flusso, riduzioni funzionali ben più raffinate di questa riguardo al modo nel quale il cervello umano elabora l'informazione contenuta nello stimolo sensoriale (oppure endopsichico) per produrre la risposta motoria. Ma lo schema proposto sopra, basato sulla teoria avverbiale della percezione, è un primo passo, grossolano dal punto di vista scientifico ma utile dal punto di vista

filosofico, per sapere in quale direzione possa essere cercata l'implementazione cerebrale dell'Intenzionalità.

Tale passo è compiuto *top down* da uno stato Intenzionale alla sua riduzione funzionale. Per giungere successivamente alla implementazione cerebrale di questa riduzione funzionale, è necessario inoltre che venga compiuto un secondo passo nella direzione opposta, *bottom up*, dall'attività dei singoli neuroni alla dinamica complessiva del cervello. Infine, se, compiendo un ulteriore passo, risulterà possibile identificare le varie caratteristiche della riduzione funzionale degli stati mentali con corrispondenti *higher order properties* della dinamica cerebrale, ci si sarà avvicinati non poco alla individuazione di una soluzione naturalistico-fisicalistica del problema mente-corpo.

Indicazioni importantissime sul modo di compiere il secondo passo dall'attività dei singoli neuroni alla dinamica complessiva del cervello possono essere reperite nelle ipotesi avanzate al riguardo da Edelman e Tononi (2000) (cfr. anche Zeppi, Plebe e Perconti 2016). Applicando queste ipotesi all'esempio summenzionato del vedere che il semaforo è rosso, si può sostenere che tale percezione, al pari della percezione di qualsiasi altro *quale* (visivo, acustico, tattile ecc.), è implementata mediante un complesso sistema di processi di modulazione dell'attività cerebrale ottenuti grazie al rilascio in tutto il cervello, in modo appropriatamente distribuito, di neuro-modulatori prodotti dai gangli basali e, più in generale, grazie a quel complesso di interazioni tra la corteccia, il talamo e i gangli basali stessi che Edelman chiama *Dynamic Core*. In questa ottica occorre presumere che la riduzione funzionale della percezione che il semaforo è rosso, ossia [*rossamente(f-Vedere)*], sia implementata da un processo di neuro-modulazione di parte della dinamica cerebrale; processo qui indicato simbolicamente mediante [*nm3(d)*], dove *d* è la dinamica cerebrale che implementa le percezioni visive e *nm3* è quella neuro-modulazione di tale dinamica che compare di solito in maniera distribuita in certe parti del cervello umano allorché quest'ultimo, esposto ad uno stimolo distale del tipo di [*SD-Rosso*], comandi l'esecuzione di una risposta motoria del tipo di *R* (o altre risposte motorie che richiedano il riconoscimento di oggetti rossi nell'ambiente esterno).

In altre parole, dopo aver ridotto funzionalmente il contenuto della percezione ad un certo modo di vedere – trasformando [*Vedere([Rosso*])*] in [*rossamente(f-Vedere)*] – non esiste più alcuna diffi-

coltà ad ipotizzare che questo stato funzionale sia implementato da una certa dinamica complessiva dell'apparato visivo. Mentre è impossibile comprendere in qual modo uno stato fisico come d possa avere una relazione di Intenzionalità con lo stato fisico esterno $[SD-Rosso]$ tramite quel misterioso oggetto fenomenico che è il *quale* $[Rosso^*]$ e solo grazie a questa altrettanto misteriosa relazione possa causare la risposta motoria R , è invece pienamente plausibile che il *proxy stimulus* retinico, causato dallo stimolo distale $[SD-Rosso]$, provochi nel *Dynamic Core* tramite il rilascio opportunamente distribuito di certi neuromodulatori la neuromodulazione $nm3$ della dinamica cerebrale d e che quest'ultima, così neuromodulata, scateni la risposta motoria R . In altre parole, mentre un *quale* come $[Rosso^*]$, in quanto oggetto della mia percezione cosciente, non può essere implementato fisicamente, un certo modo di *f-Vedere*, ossia il “vedere rossamente”, può invece essere implementato da un certo modo di modulare quella dinamica cerebrale che implementa il mio “vedere” stesso in quanto stazione intermedia della catena causale che connette lo stimolo distale alla risposta motoria (Nannini 2007a, pp. 199-229).

Di conseguenza la dinamica cerebrale che implementa la riduzione funzionale della mia percezione che il semaforo è rosso, in quanto mero stato Intenzionale, può essere rappresentata simbolicamente così (C è l'implementazione cerebrale di $f-M$, mentre $[SD-Rosso]$ e R sono le ridescrizioni, in termini scientifici, rispettivamente di $[Rosso]$ e di A):

Implementazione cerebro-sistemica della relazione d'Intenzionalità

$$([SD-Rosso] \rightarrow ([nm3(d)] \& C) \rightarrow R$$

Infine, questa dinamica cerebro-sistemica può essere considerata, seguendo Edelman e Tononi, una *higher order property* emergente nel *Dynamic Core* dall'attività dei singoli neuroni che partecipano al coordinamento senso-motorio:

Implementazione neuronale della relazione di Intenzionalità

$$[SD-Rosso] \rightarrow \text{processi neurali} \rightarrow R$$

Si noti inoltre che la stessa dinamica cerebro-sistemica può essere implementata da processi neuronali distinti, che Edelman chiama “degenerati”. Ciò consente di difendere il fisicalismo dall’obiezione sollevata contro di esso dai funzionalisti, allorché questi ultimi obiettano, mediante la teoria delle “molteplici realizzazioni” prima menzionata, che uno stato mentale, essendo implementabile mediante processi neuronali diversi, non può essere identico a nessuno di essi (se lo fosse, sarebbe simultaneamente identico a cose diverse e quindi sarebbe diverso da se stesso). Un singolo stato mentale, in quanto stato funzionale, è identico infatti ad una sola *higher order property* della dinamica cerebrale; ma è quest’ultima che può emergere – secondo la teoria che ho chiamata prima “emergentismo fisicalistico” – da processi neuronali degenerati di tipo diverso.

Concludendo, la riduzione ontologica dell’Intenzionalità di una percezione è un processo articolato in tre passi. Primo passo (*top down*): riduzione funzionale della ricostruzione filosofica dell’Intenzionalità di quella percezione. Secondo passo (*bottom up*): chiarire, in termini fisico-chimici, come dall’implementazione neuronale dell’Intenzionalità emerga la sua implementazione cerebro-sistemica. Terzo duplice passo (sia *top-down* sia *bottom up*): mostrare che, almeno in certi esempi privilegiati, la riduzione funzionale e l’implementazione cerebro-sistemica convergono verso una comune meta-rappresentazione matematica di tipo vettoriale.

Un abbozzo dei primi due passi è stato qui già illustrato. Come esempio del terzo passo si possono citare gli studi sulla rappresentazione vettoriale della visione dei colori. Questi studi mostrano che non è possibile percepire alcun colore singolo se non come un elemento in un sistema di colori. In tale sistema ogni atto mentale del tipo [*coloratamente(f-Vedere)*] – al pari della riduzione funzionale di qualsiasi altro *quale* appartenente a qualsiasi altra modalità sensoriale (come mostrato da Paul Churchland 1995, pp. 21-55) – è meta-rappresentato geometricamente come un vettore in uno spazio vettoriale tridimensionale avente come sue dimensioni tre assi cartesiani: l’asse del giallo-blu, l’asse del rosso-verde e l’asse del bianco-nero (P.S. Churchland 2002, pp. 185-189). È in questo spazio vettoriale che viene a formarsi il “doppio cono” entro il quale sono collocati tutti i colori dello spettro del visibile:

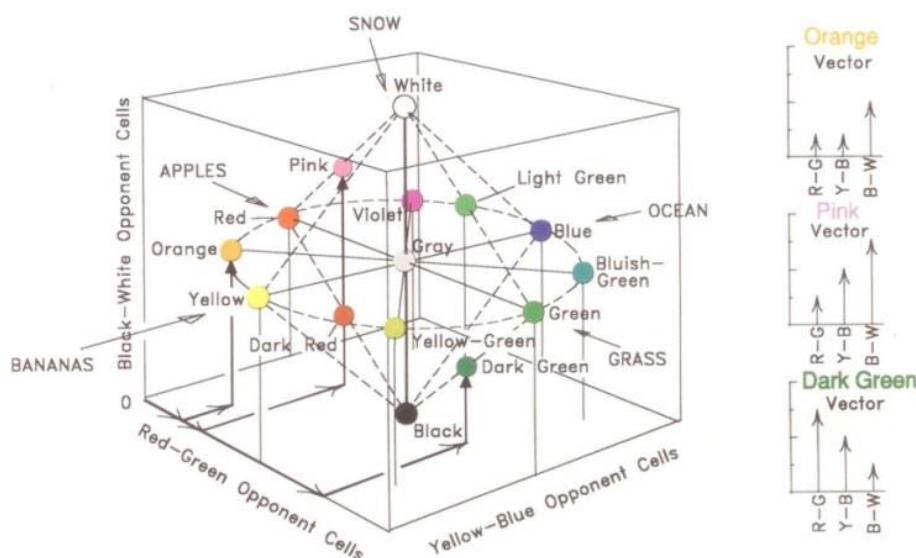


Fig. 3
(Da P.S. Churchland (2002, p. 186))

Inoltre anche la dinamica cerebrale emergente dall'attività dei singoli neuroni – studiata in particolare allo *Swartz Center for Computational Neuroscience* dell'Università di San Diego in California (Makeig 2001) – può essere meta-rappresentata come una traiettoria di *pattern* neuronali in un iperspazio vettoriale. Ora, se la meta-rappresentazione vettoriale della riduzione funzionale della percezione di un certo colore viene a coincidere con la meta-rappresentazione vettoriale della dinamica cerebrale emergente da quei processi neuronali che vengono osservati (o che comunque sono empiricamente ricostruibili) quando un soggetto riconosca quel colore, allora si può dire che quella percezione è stata ontologicamente ridotta ad una certa attività cerebrale.

Più in dettaglio tre gruppi di neuroni del corpo genicolato laterale (qui simbolizzati mediante tre cerchi e etichettati rispettivamente come “blue vs. yellow”, “green vs. red” e “white vs. black”) ricevono e variamente combinano in controfase i segnali inibitori e eccitatori, relativi ad un certo punto del campo visivo, provenienti da tre gruppi di coni della retina qui rappresentati mediante tre triangoli ed etichettati come “blue”, “green” e “red” a seconda delle lunghezze delle onde elettromagnetiche a cui reagiscono depolarizzandosi. I segnali inibitori inviati dalla retina verso il corpo genicolato laterale sono qui simbolizzati mediante frecce punteggiate, mentre quelli eccitatori sono simbolizzati mediante frecce continue. In tal modo questi tre gruppi di neuroni del corpo genicolato

laterale (ossia i tre cerchi) implementano – come suggerito dalle loro rispettive denominazioni, quando la loro attività venga astrattamente meta-rappresentata mediante uno spazio vettoriale tridimensionale – i tre assi cartesiani del giallo-blu, del verde-rosso e del bianco-nero. Il primo gruppo, combinando i segnali inibitori provenienti da “blue” con quelli eccitatori provenienti sia da “green” sia da “red”, implementa l’asse del giallo-blu. Il secondo gruppo, combinando invece i segnali inibitori provenienti da “green” con quelli eccitatori provenienti da “red”, implementa l’asse del verde-rosso. Il terzo gruppo infine – combinando i segnali eccitatori provenienti da tutti e tre i gruppi di coni che reagiscono al punto in questione del campo visivo con i segnali inibitori provenienti dai coni che reagiscono ai punti circostanti – implementa l’asse del bianco-nero. In tal modo tutta questa complessa attività neuronale che, interessando la retina ed il corpo genicolato laterale, viene a implementare da un punto di vista funzionale la visione dei colori, può essere così meta-rappresentata:

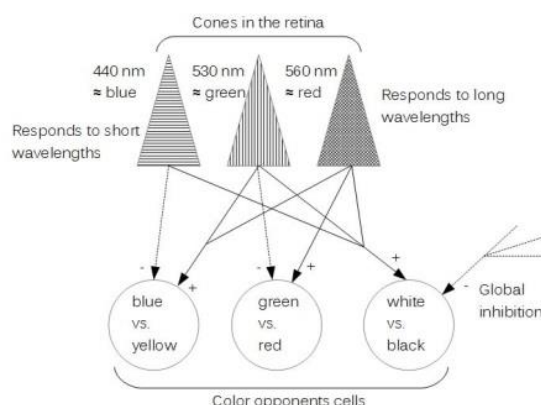


Fig. 4

(Rielaborazione da Churchland (2002, p. 186)

Con ciò può considerarsi concluso lo schema di riduzione ontologica di una percezione, in quanto mero stato Intenzionale (privo di coscienza e autocoscienza), ad un certo insieme di processi cerebrali. Ovviamente si tratta di un semplice abbozzo filosofico.

11. La riduzione ontologica della coscienza e dell'autocoscienza

Si obietterà che, avendo io messo in parentesi il carattere cosciente e soggettivo della mia percezione, non ho per ora neppure sfiorato lo *Hard*

Problem. Questo è vero, ma la riduzione ontologica dell'Intenzionalità ad una *higher order property* della dinamica cerebrale meta-rappresentabile in termini vettoriali spiana in effetti la strada, almeno in linea di principio, ad una analoga riduzione della coscienza e dell'autocoscienza²⁶. Per reintrodurre queste due ultime caratteristiche, torniamo alla ricostruzione filosofica della mia percezione. Se in essa, per analogia con il modo nel quale $[Vedere([Rosso*])]$ è stato funzionalmente ridotto a un $[rossamente(f-Vedere)]$, si sostituiscono *coscientemente* (coscienza fenomenica) e *Io** (autocoscienza fenomenica) rispettivamente con gli operatori avverbiali *f-coscientemente* (coscienza funzionale) e *soggettivamente* (autocoscienza funzionale), si ottiene la riduzione funzionale della mia percezione concepita in modo completo come stato Intenzionale, cosciente e soggettivo (ossia autocosciente):

Riduzione funzionale completa

$$[Rosso] \rightarrow ([f-coscientemente(soggettivamente(rossamente(f-Vedere)))] \& f-M) \rightarrow A$$

In modo analogo è facile passare, in abbozzo, alla implementazione cerebro-sistemica completa di tale riduzione funzionale:

Implementazione cerebro-sistemica completa

$$[SD-Rosso] \rightarrow ([nm1(nm2(nm3(d)))) \& C) \rightarrow R$$

In essa *nm1*, *nm2* e *nm3* rappresentano simbolicamente le neuromodulazioni che fanno di *d* l'implementazione cerebro-sistemica di una percezione cosciente, soggettiva e Intenzionale. In altre parole, la dinamica cerebrale *d*, opportunamente modulata da *nm3* in modo da implementare lo stato funzionale $[rossamente(f-Vedere)]$, viene ulteriormente modulata da *nm2*, che rende la percezione del semaforo rosso funzionalmente connessa alla "rappresentazione" che il cervello, automonitorandosi, si dà della sua stessa attività e dello stato del resto del corpo. Si ottiene così una connessione, nel cervello, tra la percezione complessiva

²⁶ Per una introduzione generale ai concetti di coscienza e auto-coscienza cfr. in particolare Di Francesco (2000) e Perconti (2008 e 2011).

del mondo esterno e le informazioni di cui il cervello stesso dispone sulla sua propria attività e sul resto del corpo. Questa connessione è essenziale per ottenere un buon coordinamento senso-motorio (su tutto questo si veda in Edelman l'interazione tra la «scena unitaria» ed il «Self» (Edelman 2004, pp. 41-50)).

Tuttavia tale coordinamento non potrebbe raggiungere quel grado di precisione, duttilità e sensibilità al contesto che è tipico degli esseri umani e di altri animali se la percezione che il cervello ha dell'ambiente esterno e di sé, insieme al resto del corpo, non fosse in parte cosciente. Rendere cosciente una parte dell'attività cerebrale è appunto la funzione della neuromodulazione qui indicata simbolicamente con *nml*, una neuromodulazione coincidente probabilmente sia con processi di sincronizzazione cerebrale tra circuiti neuronali oscillanti (Engel e Singer 2001, Engel 2003 e 2005, W. Singer 2004) sia con altri processi cerebrali non ancora completamente conosciuti ma comunque indagati dai neuroscienziati con crescente successo (P.S. Churchland 2014, pp. 259-266; Dehaene 2014), come è stato provato ad esempio riguardo alle percezioni visive (Tacca 2010). La coscienza fenomenica sembra essere infatti anzitutto, come suggerito dal *blindsight* e dalla sindrome di Anton-Babinski (cfr. *supra*), una funzione di monitoraggio che il cervello opera sulla sua stessa attività per trovare una sintonia fine tra i movimenti del corpo e l'ambiente esterno. Sembra quindi che essa sia una preconditione di quell'attività di coordinamento tra i vari processi cerebrali che è stata ricostruita, in due modi diversi ma convergenti, da Baars (1988 e 1997) e dal già più volte menzionato Edelman (2004) rispettivamente come *Global Workspace Memory* e come *Dynamic Core*. Più in dettaglio mi sembra che si possa avanzare l'ipotesi che la coscienza fenomenica, lungi dall'essere una sostanza (o una proprietà) non fisica, sia invece una sorta di “codice-macchina” interno al cervello che conferisce all'informazione veicolata da certe parti della dinamica cerebrale il “formato” necessario per accedere alla *Global Workspace Memory* (Nannini 2013 e 2016).

Per chiarire quale rapporto vi sia tra la coscienza fenomenica e la sua implementazione cerebro-sistemica si può ricorrere ad una analogia: se si prende un *file* di immagine e lo si apre come se fosse un *file* di testo, si può vedere che l'informazione contenuta nel *file*, pur essendo sempre la stessa, si presenta in due modi completamente

diversi: come l'immagine, ad esempio, di Marilyn (Fig. 5) o come una sequenza inintelligibile di lettere e numeri (Fig. 6 – L'informazione contenuta in una sola immagine occupa decine di pagine in formato testo; ne viene qui proposto perciò solo un frammento iniziale):



Fig. 5

ÿØÿá67Exif6767II*676767676767676767676767ÿì67Ducky67
67676767<6767ÿî67

Fig. 6

La ricostruzione dell'attività del cervello nel linguaggio delle neuroscienze è paragonabile alla foto presentata in formato testo. La medesima attività cerebrale, quando venga monitorata e formattata dal cervello stesso nel "formato" della coscienza, è invece paragonabile all'immagine di Marilyn. In altre parole il *gap* fenomenologico esistente tra l'avere un'esperienza soggettiva e la sua implementazione cerebro-sistemica è indubitabile e paragonabile, sul piano emotivo, allo scarto che percepisco tra l'immagine della affascinante Marilyn ed una insignificante sequenza di segni. In modo analogo c'è una bella differenza tra il sentire un terribile mal di denti e l'ascoltare un neurologo che mi informa sulla attività delle fibre *C* nel mio cervello. Ma è un errore credere che quel *gap* fenomenologico possa essere spiegato mediante l'ammissione di un *gap* ontologico

(come pensano i dualisti). Il mio dolore non è cosa diversa dalla sua implementazione cerebro-sistemica. È la stessa cosa sotto due descrizioni diverse: l'una nel linguaggio delle neuroscienze, l'altra nel codice-macchina del cervello; un codice che si è formato per selezione naturale nel corso dell'evoluzione biologica.

Si obietterà forse che mediante questa ipotesi il *gap* esplicativo (Levine 1983) tra i *qualia* e i loro correlati neuronali viene semplicemente trasferito all'interno del "codice coscienziale", ma non viene cancellato. Perché – si obietterà – gli oggetti la cui superficie riflette delle onde elettromagnetiche aventi una lunghezza compresa tra 620 e 750 nanometri vengono percepiti proprio come rossi e non, ad esempio, come verdi o gialli? L'ipotesi che la coscienza fenomenica sia una sorta di codice-macchina del cervello – si obietterà – non spiega questa connessione tra lunghezze d'onda e colori e, più in generale, tra *qualia* e attività cerebrale. E dunque – hanno concluso in molti (ad es. Bieri 1995) – lo *Hard Problem* non è stato risolto!

Ebbene, ciò in un certo senso è vero; ma la ragione di questa mancata spiegazione è semplice e non ha nulla a che vedere con l'esistenza di una differenza metafisica tra mente e corpo (come credono invece i dualisti): l'ipotesi summenzionata non spiega la connessione tra i *qualia* e i loro correlati neuronali per la banalissima ragione che... non c'è nulla da spiegare! Quella connessione è arbitraria e quindi non è razionalmente giustificabile. Essa è frutto di una sorta di convenzione, di un codice appunto introdotto in modo cieco da Madre Natura per selezione naturale nel corso dell'evoluzione biologica: i genotipi che producono fenotipi dotati di cervelli cablati in modo tale da adottare quella convenzione si sono replicati in numero maggiore di altri, perché i loro fenotipi – essendo capaci di costruire, mediante il loro cervello, un'immagine accurata del mondo esterno che sfrutti, ad esempio, il modo nel quale gli oggetti riflettono la luce o, più in generale, il modo nel quale certe loro proprietà fisiche eccitano certi recettori – godono di un vantaggio adattativo e, vivendo in media più a lungo, hanno una prole più numerosa. Non c'è nessuna somiglianza tra le proprietà degli oggetti fisici e i processi neuronali che esse scatenano (ad esempio tra la lunghezza di certe onde elettromagnetiche riflesse dalla loro superficie e la visione di certi colori); ma, una volta stabilitasi una connessione empirica costante tra certe proprietà fisiche degli oggetti esterni e certi stati mentali (identici onto-

logicamente a certe proprietà della dinamica cerebrale), l'animale dispone di un trucco efficacissimo per ricostruire un'immagine dell'ambiente esterno che possa guidarlo nei suoi movimenti. In fondo si tratta di trucco analogo a quello usato dalla guardarobiera di un teatro per ritrovare rapidamente un cappotto sulla base della semplice presentazione del relativo scontrino (cfr. *infra* capitolo sesto). Ovviamente Madre Natura ha trovato molti altri trucchi, oltre alla invenzione della coscienza fenomenica come codice-macchina interno al cervello, per migliorare la sintonia fine dell'interazione tra gli animali e il loro ambiente; ma il trucco della coscienza fenomenica, una volta presentatosi per caso nel corso dell'evoluzione delle specie e affermatosi per la sua efficienza, ha svolto e svolge un ruolo decisivo nella determinazione del comportamento almeno degli esseri umani e probabilmente di molti altri animali (Edelman 1989).

Tornando all'esempio della percezione del semaforo rosso, la coscienza compare nella sua ricostruzione filosofica come coscienza fenomenica (*coscientemente*), nella riduzione funzionale come coscienza funzionale (*f-coscientemente*) e nella implementazione cerebro-sistemica come coscienza cerebrale (*nm1*). Se si sceglie la descrizione neurologica come quella che, in una concezione scientifica del mondo, deve essere considerata per *default* come la migliore rappresentazione della realtà, allora la coscienza cerebrale è l'unica veramente reale e causalmente efficace (cfr. *infra* capitolo quinto). La coscienza funzionale è invece una ridescrizione della coscienza cerebrale nel linguaggio della psicologia cognitiva. La coscienza fenomenica è infine il "formato" che assumono le altre neuromodulazioni – *nm3* (ossia *rossamente*) e *nm2* (ossia *soggettivamente*) – quando, ulteriormente rimodulate dal codice-macchina *nm1*, appaiono al cervello stesso (ossia a me) come degli oggetti della mia percezione (ossia come [*Rosso**] ed *Io**) distinti dall'atto di percepirli. In conclusione la chiave della soluzione dello *Hard Problem* in termini naturalistico-fisicalistici sta nel comprendere che la coscienza fenomenica – sebbene appaia al senso comune come un qualcosa di esterno e di "superiore" all'attività del cervello – in realtà, lungi dall'essere un qualcosa di non fisico, è una modalità di presentazione di parte di quella stessa attività cerebrale nel "formato coscienziale".

Esiste tuttavia una differenza tra la ridescrizione della coscienza cerebrale come coscienza funzionale e come coscienza fenomenica.

La riduzione funzionale di uno stato di coscienza è implementata completamente dalla dinamica cerebrale, mentre lo stato di coscienza stesso, per come viene vissuto (ossia soggettivamente esperito), presenta due caratteristiche neurologicamente non implementabili. In primo luogo esiste una relazione tra l'atto percettivo e il suo "percelto" (ad esempio tra *Vedere* e *[Rosso*]*) a livello della coscienza fenomenica che non corrisponde alla relazione intercorrente sia tra la riduzione funzionale dell'atto percettivo stesso ed una sua proprietà (ossia tra *f-Vedere* e *rossamente*) sia tra la dinamica cerebrale ed una sua modulazione (ossia tra *d* e *nm3*). La prima relazione intercorre tra l'atto percettivo ed il suo oggetto; la seconda tra la riduzione funzionale di quell'atto (o la sua implementazione cerebro-sistemica) ed un suo modo di essere. Ad esempio nel primo caso vedo qualcosa (vedo rosso o vedo verde), nel secondo caso "vedo" in un certo modo (ossia "vedo rossamente" o "verdemente" così come potrei vedere chiaramente o indistintamente). Pertanto nel passaggio dalla percezione cosciente alla sua riduzione funzionale e implementazione cerebro-sistemica va perduta la distinzione apparentemente ontologica tra la percezione stessa ed il suo oggetto fenomenico, il suo *quale*. Mentre sembra quanto meno plausibile ridurre ontologicamente una percezione ad una certa attività cerebrale, non sembra possibile ridurre ciò che è oggetto di una percezione ad una proprietà della percezione stessa. La distinzione apparentemente ontologica tra soggetto e oggetto, propria del formato coscienziale, va perduta nella riduzione funzionale e nella conseguente implementazione cerebro-sistemica.

Come spiegare questa perdita? Si torna qui allo *Hard Problem*. Se non si vuole ricadere nel dualismo, occorre ammettere che i *qualia*, in quanto oggetti fenomenici distinti dagli atti percettivi mediante i quali sono percepiti, non sono riducibili all'attività del cervello semplicemente perché... non esistono (Dennett 1988 e 2016)! Essi sono delle illusioni create dal cervello, sono dei *ficta*, ossia sono un suo trucco per migliorare il coordinamento senso-motorio (del resto anche nelle scienze è opportuno parlare talvolta di idealizzazioni che a rigore non esistono, come in fisica i gas perfetti o il moto perpetuo) (cfr. *infra* capitolo terzo). La dinamica cerebrale che implementa la mia percezione, quando venga neurologicamente modulata in modo opportuno attraverso *nm1*, assume un formato coscienziale grazie al quale compare – nel cervello, ad opera del cervello e per il cervello – una triplice

distinzione tra l'attività del vedere, il soggetto vedente e l'oggetto visto; distinzione che è invece assente sia nel formato avverbiale della riduzione funzionale sia nel formato cerebro-sistemico della implementazione neurologica di una percezione cosciente. Mi pare che questo sia un modo più chiaro di formulare ciò a cui si riferisce Edelman quando parla, in maniera non priva di oscurità, della «trasformazione fenomenica» dei processi neuronali in *qualia* (Edelman 2004, pp. 65-73).

Considerazioni analoghe possono essere fatte riguardo all'autocoscienza (cfr. *infra* capitolo settimo). Quando, nel vedere che il semaforo è rosso, divengo consapevole del fatto che sono io a vederlo rosso, con ciò il mio cervello monitorizza e quindi rappresenta a se stesso il suo *Self* (implementato da *nm2*) come un *Io**. E dunque io, in quanto appaio al mio cervello (ossia a me stesso) nel formato coscienziale come un *Io**, sono una illusione al pari di *[Rosso*]*. Infatti è sì vero che quel mio essere consapevole del fatto che sono io a vedere che il semaforo è rosso è in termini funzionali un *soggettivamente(Vedere)*, ossia è uno stato virtuale implementato da *nm2(d)*; ma questo aspetto della dinamica cerebrale, quando venga presentato dal cervello a se stesso nel formato coscienziale, diviene, in quanto *Io**, l'agente di un atto percettivo dotato in apparenza di autonomia ontologica rispetto all'atto stesso. Pertanto è in virtù del formato coscienziale che io mi sento come il soggetto che ha quella percezione e che agisce di conseguenza, allorché premo il pedale del freno; ma in realtà io sono solo una parte del contenuto idealizzato e ingannevole della percezione stessa. Mi sento sì, per riprendere la celebre immagine aristotelica, come il timoniere del mio corpo, ma in realtà sono solo un servomeccanismo di un sofisticato timone automatico! La sensazione di *agency* (Wegner 2002) che provo ogni volta che compio un'azione volontaria – una sensazione che ha la sua radice nella percezione endopsichica dello scorrere del tempo (cfr. *infra* capitolo quarto) – è tuttavia talmente radicata nella mia normalità psichica che tale sensazione permane anche se mi convinco che è illusoria, un po' come avviene per molte illusioni ottiche (Nannini 2007a, pp. 153-156).

In conclusione occorre ammettere, per quanto ciò sia lontano dal senso comune, che i *qualia* non esistono e che io stesso a rigore non esisto (Metzinger 2003). Quando sono sveglio, mi sento esistente

ovviamente; ed è questa sensazione che Descartes sfruttò per concepire il suo celebre *Cogito, ergo sum*. Ma si tratta in realtà di una sensazione ingannevole, sebbene utilissima, che il mio cervello ricrea di istante in istante. Quando mi sveglio al mattino, la mia impressione di essere lo stesso che si è addormentato la sera prima è basata solo sui ricordi presenti delle mie esperienze soggettive passate. Io proverei la stessa impressione anche se fossi un robot costruito durante la notte e, durante la costruzione, fossero stati inseriti artificialmente nei circuiti del mio cervello dei falsi ricordi di una vita mai vissuta (è questa la fantasia che, com'è noto, si trova alla base del celebre film di fantascienza *Blade Runner*).

Capitolo terzo

Mente, corpo, società e ontologia del mentale

1. Riduzionismo e antiriduzionismo

Fa parte del senso comune ritenere che l'Italia esista solo perché esistono gli italiani e che, se in Italia ed in Germania calano i consumi, è perché la gente compra meno. Più in generale nessuno, a meno che non abbia già in testa qualche teoria filosofica precostituita, trova nulla da ridire se si afferma che i fatti sociali esistono solo perché la gente in carne ed ossa si comporta in un certo modo. Naturalmente le cose cambiano quando dal piano ontologico si passa a quello epistemologico: il fatto che l'andamento della borsa di Milano in un certo giorno dipenda da una miriade di azioni individuali di compravendita non significa che tale andamento possa essere spiegato direttamente ricostruendo le motivazioni psicologiche dei singoli compratori e venditori. L'economia non è riducibile alla psicologia. Tuttavia pochi, quando la distinzione sia spiegata loro con chiarezza, sono disposti a vedere in tale irriducibilità epistemologica una irriducibilità ontologica ed a sostenere che l'andamento della borsa ha un'esistenza autonoma rispetto alle azioni individuali di coloro che hanno comprato e venduto azioni. Pochi credono, insomma, nell'esistenza della *conscience collective* di durkheimiana memoria (ammesso e non concesso che Durkheim (1894) stesso intendesse questo concetto nel modo ingenuo che gli viene in genere attribuito). Dal punto di vista ontologico il rapporto tra la società e gli individui, o meglio tra la società e le menti che stanno dietro alle azioni individuali, sembra non presentare particolari difficoltà. Per il senso comune non esiste alcun problema società-individui.

Diverse sono invece le intuizioni dell'uomo della strada se si passa al rapporto mente-corpo. Anche in questo caso nessuno dubita oggi che, quando pensiamo a che cosa faremo domani, qualcosa succede nel nostro cervello. L'idea aristotelica che l'intelletto non si appoggi ad un organo è ormai tramontata anche per il senso comune. Tuttavia – quando un fisicalista sostiene che gli atti mentali sono identici a processi cerebrali e un funzionalista, per correggerlo, afferma che in realtà essi sono dei programmi implementati da quell'*hardware* particolare che è il cervello – l'uomo della strada (e insieme a lui più di un filosofo) pensa che hanno torto entrambi, dato che gli

atti mentali emergono certo dall'attività del cervello, ma sono qualcosa di irriducibile a qualsiasi fenomeno fisico-chimico! Come ed in quale senso tale emergenza della mente dal brulichio dell'attività dei neuroni sia possibile – pensa l'uomo della strada (insieme a molti filosofi e scienziati) – rimane un problema insoluto, forse addirittura un mistero insondabile.

Riassumendo, il senso comune non trova nulla di strano nel dire che i fatti sociali esistono perché si realizzano tramite atti mentali ed azioni di individui in carne ed ossa; viceversa, sebbene tutti sappiano che in assenza di un cervello funzionante non si dà nessun pensiero e nessuna attività mentale di alcun genere, la maggior parte della gente trova addirittura assurdo il credere che gli atti mentali stessi siano riducibili a processi cerebrali.

Ma in effetti questo diverso atteggiamento del senso comune verso il rapporto società-individui, da un lato, e mente-corpo, dall'altro, non ha alcun fondamento: il problema ontologico è analogo in entrambi i casi e suscettibile delle medesime soluzioni. Per mettere in luce questa somiglianza ho però bisogno di tracciare le linee essenziali di un'ontologia filosofica che sia coerente con quella prospettiva naturalistica secondo la quale l'*homo sapiens* è un prodotto dell'evoluzione biologica al pari di tutti gli altri animali e pertanto anche la sua mente può e deve essere spiegata come un qualsiasi altro fenomeno naturale.

2. Verso un'ontologia filosofica di tipo naturalistico

L'ontologia nella tradizione filosofica è, insieme alla teologia, parte essenziale della metafisica. Poiché il naturalismo, nato all'interno della filosofia analitica, è tendenzialmente avverso ad ogni forma di metafisica, sembra che un filosofo naturalista non possa lasciare alcuno spazio a domande intorno all'ontologia. Le domande che un filosofo legittimamente può porre a chi faccia una qualsiasi affermazione empirica – pensano ancora molti filosofi analitici – sono soltanto due: “Che cosa intendi dire?” e “Come lo sai?”; ossia domande di tipo logico-semantico e epistemologico, ma non ontologico. Qualsiasi questione intorno all'esistenza di questa o quella entità o viene posta all'interno di una determinata teoria (e quindi è riducibile al significato di questo o quel termine entro la teoria stessa) oppure è insensata. Ad esempio chiedersi se esiste un numero più grande di

qualsiasi altro è perfettamente legittimo all'interno dell'aritmetica (e la risposta ovviamente è "No!" in questo caso), ma chiedersi se i numeri esistono nello stesso senso in cui esistono le pietre è assurdo: quest'ultima domanda non ammette nessuna risposta sensata, poiché non c'è una teoria che riguardi, allo stesso tempo, sia i numeri sia le pietre.

Questo rifiuto dell'ontologia non è tuttavia condiviso da tutti quei naturalisti che, rifacendosi all'epistemologia di W.v.O. Quine (1953, 1960 e 1969), ritengono che l'ontologia stessa sia un capitolo rispettabile della filosofia analitica. Essi aggiungono perciò alle due domande summenzionate una terza domanda che il filosofo analitico può e deve legittimamente porsi: "Che cosa esiste realmente?". Nel tentare di rispondere a questa terza domanda in direzione della costruzione di un'ontologia filosofica di tipo naturalistico che s'ispiri largamente (anche se non completamente!) a Quine, do per scontati alcuni presupposti (non discussi qui dettagliatamente, ma che è sperabile che risultino in seguito quanto meno plausibili alla luce della solidità dell'edificio teorico di cui fanno parte) (Nannini 2007a, pp. 159-174).

In primo luogo ogni teoria scientifica, escluse la matematica e la logica, è una teoria empirica. Pertanto essa può essere vera solo se è confermata da prove empiriche. Ma per altro verso una teoria scientifica deve anche rispettare alcuni criteri formali di validità individuati da una epistemologia normativa che, pur muovendosi nell'ambito della continuità quiniana tra scienza e filosofia, rivendichi tuttavia per la filosofia stessa un certo margine di autonomia rispetto alla scienza (Koppelberg 1999) a differenza di quanto avviene nella "epistemologia naturalizzata" dei quiniani ortodossi; un margine di autonomia che il filosofo dovrebbe sfruttare per costruire un'epistemologia multidisciplinare nella quale tutte le scienze possano ritrovarsi.

In secondo luogo i criteri di validità individuati da una epistemologia che, pur essendo normativa nel senso sopra chiarito, voglia essere di tipo naturalistico devono essere applicabili a tutte le teorie scientifiche indipendentemente dal loro oggetto e dai loro metodi di ricerca. A differenza di quanto pensano gli storicisti, gli idealisti, i dualisti e i relativisti cognitivi, le scienze naturali e le scienze umane e sociali, sebbene possano servirsi di metodi molto diversi di ricerca

e di conferma dei risultati ottenuti, condividono i medesimi criteri formali di validità. Ad esempio, qualora si ritenga, seguendo Hempel (1965), che una spiegazione sia formalmente corretta solo se *l'explanandum* è logicamente deducibile o induttivamente inferibile da un *explanans* che contenga non solo condizioni iniziali ma anche leggi generali o probabilistiche, allora tale condizione di validità vale non solo per la spiegazione degli eventi naturali, ma anche per la spiegazione delle azioni umane e, più in generale, per la spiegazione di tutti i fenomeni storico-sociali. In altre parole, il naturalismo – come viene inteso qui – rifiuta la distinzione storicistica tra scienze della natura e scienze dello spirito o tra scienze nomotetiche e scienze idiografiche: il naturalismo include il monismo metodologico (Nannini 1992).

In terzo luogo ogni teoria scientifica definisce il proprio oggetto interno. Tale oggetto è infatti definibile solo come ciò che la teoria descrive, spiega e prevede nel suo proprio linguaggio. Ad esempio, come potremmo comprendere che cosa è un elettrone al di fuori della meccanica quantistica? Ma ciò non implica che, quando si passi a chiedersi come si possa parlare dell'esistenza di ciò che è oggetto di una teoria scientifica, si debba concludere che non avrebbe alcun senso distinguere l'esistenza assoluta dall'esistenza relativa e che l'esistenza di una qualsiasi cosa dovrebbe essere intesa sempre e soltanto come esistenza relativa ad un certo ordine di discorso. Infatti è evidente che gli impegni ontologici di alcune teorie godono di esistenza assoluta, mentre ciò che è oggetto di altre teorie, pur godendo di esistenza relativa entro il proprio ordine di discorso, non ha in quanto tale esistenza assoluta e può essere considerato come assolutamente esistente solo in modo indiretto quando risulti essere implementato da altre entità (o proprietà o stati di cose o eventi) che godono invece di esistenza assoluta in modo diretto. Ad esempio è implausibile che l'esistenza dei sassi venga posta sullo stesso piano dell'esistenza dei numeri. I sassi infatti godono di esistenza assoluta, perché il mondo fisico a cui appartengono basta a se stesso. È presumibile che i sassi continueranno ad esistere sulla Terra anche dopo l'estinzione della specie umana. I numeri viceversa – a meno che non si sia dei platonici – possono esistere solo come contenuto dei pensieri degli esseri umani che pensano ad essi.

In altre parole gli oggetti interni delle varie teorie scientifiche

appartengono a ordini di discorso stratificati l'uno sull'altro. Di conseguenza gli oggetti appartenenti ai livelli superiori possono godere di esistenza assoluta solo indirettamente se vengono implementati da oggetti esistenti direttamente in senso assoluto in quanto appartenenti ad un ordine di discorso che fino a prova contraria è da considerarsi, per ragioni pragmatiche e per *default*, come “di base”. Ad esempio, mentre la fisico-chimica ha il privilegio di descrivere un mondo chiuso in se stesso entro il quale tutto ciò che è fisico-chimico è spiegabile solo in termini fisico-chimici, la biologia non può rivendicare un'analoga chiusura del suo oggetto interno: è infatti ormai un'opinione diffusissima anche tra i non addetti ai lavori, dopo il tramonto del vitalismo, che i fenomeni biologici sono riducibili in termini fisico-chimici. Pertanto i fenomeni fisico-chimici e i fenomeni biologici appartengono a due ordini di discorso sovrapposti l'uno all'altro. Di conseguenza i fenomeni biologici sono reali in senso assoluto – e quindi sono anche dotati di effettivi poteri causali – solo se vengono implementati da fenomeni fisico-chimici.

Ciò non è ovviamente sufficiente per definire con rigore che cosa sia un ordine di discorso. Tuttavia le intuizioni abbastanza chiare che, anche senza essere degli scienziati, abbiamo oggi sulle relazioni che intercorrono tra i fenomeni fisico-chimici e i fenomeni biologici possono essere trasferite per analogia anche alle relazioni corpo-mente e individui-società. Certo, stabilire quali relazioni ontologiche intercorrano tra i fenomeni neurologici e quelli mentali o tra i fenomeni mentali e quelli socio-culturali è molto più controverso di quanto non lo sia stabilire quali relazioni ci siano tra i fenomeni fisico-chimici e i fenomeni biologici. Tuttavia non si dice nulla d'incomprensibile se in via preliminare si afferma che ogni teoria scientifica – non importa se si tratti di scienze naturali, cognitive, umane o sociali – nel definire il suo oggetto interno definisce, almeno implicitamente, anche l'ordine di discorso a cui esso appartiene e perciò stesso suggerisce una qualche sua relazione con gli oggetti di altre scienze appartenenti a ordini di discorso diversi.

3. Ordini di discorso e livelli d'analisi

Per poter definire in modo chiaro la relazione che intercorre tra oggetti interni appartenenti ad ordini di discorso gerarchicamente sovrapposti, occorre distinguere gli ordini di discorso stessi dai livelli

d'analisi interni a ciascun ordine²⁷. In particolare la relazione che gli stati mentali, oggetto della *folk psychology* (e le loro “riduzioni funzionali”, oggetto della psicologia cognitiva), hanno con la dinamica cerebrale non deve essere confusa con la relazione, tutta interna all'ordine di discorso dei fenomeni fisico-chimici, che intercorre tra questa stessa dinamica cerebrale e l'attività dei singoli processi neuronali da cui emerge (nel senso precisato dall'emergentismo fisicalistico e dal concetto di causalità sistemica – cfr. *supra* capitolo secondo). La prima relazione è una relazione di implementazione, la seconda è una relazione di causalità sistemica – entro quel sistema fisico complesso che è il cervello umano – tra processi neuronali locali e proprietà sistemiche di alto livello della dinamica cerebrale. La differenza tra le due relazioni è profonda. La relazione di implementazione è una relazione concettuale: ciò che nel linguaggio della *folk-psychology* viene descritto come uno stato mentale viene ridescritto in termini scientifici o come uno stato funzionale intermedio tra l'*input* sensoriale e l'*output* motorio nel linguaggio della psicologia cognitiva oppure come una *higher order property* della dinamica cerebrale nel linguaggio delle neuroscienze. La relazione intercorrente tra i singoli processi neuronali e la dinamica cerebrale che risulta dalla loro interazione è invece una relazione empirica di tipo causal-sistemico.

La prima relazione, ponendo una distinzione puramente concettuale tra gli stati mentali (o le loro riduzioni funzionali) e la dinamica cerebrale che li implementa, presuppone una identità ontologica tra gli oggetti interni della psicologia e gli oggetti interni delle neuroscienze: gli stessi fenomeni vengono descritti ora come mentali ora come cerebrali. Psicologia e neuroscienze appartengono a ordini di discorso differenti, ma i loro rispettivi oggetti interni, in termini di esistenza assoluta, sono coincidenti. E, in una concezione scientifica del mondo, conviene per ragioni pragmatiche considerare fondamentale per *default* la descrizione neurologica.

La seconda relazione invece, essendo una relazione empirica tra singoli processi neuronali e la dinamica cerebrale complessiva che risulta dalle loro interazioni, è una relazione micro-macro tra fenomeni che, appartenendo a livelli d'analisi diversi, possono entrare in

²⁷ Faccio qui uso in modo libero della distinzione introdotta da Kim (1998 e 2005) fra ordini e livelli.

una relazione di causalità sistemica. Tuttavia, poiché entrambi i livelli interessati appartengono all'ordine di discorso dei fenomeni fisico-chimici, questa relazione causale tra livelli diversi non implica alcuna violazione del principio di chiusura del mondo fisico.

Distinguere le relazioni tra ordini di discorso dalle relazioni tra livelli d'analisi interni a ogni ordine di discorso non deve però oscurare il fatto che le relazioni tra ordini e livelli non sono indipendenti le une dalle altre. Per comprendere questo intreccio è necessario chiarire preliminarmente quali siano i livelli d'analisi che, in modo trasversale, possono comparire entro ogni ordine di discorso. Sulla scia di una nota coppia di concetti originariamente introdotta in epistemologia generale da H. Reichenbach e poi riutilizzata da Dennett (1998, pp. 95-120) in filosofia della mente, conviene distinguere gli *illata* dagli *abstracta* (Nannini 2007a, p. 170)²⁸:

- 1) I *T*-oggetti di una teoria scientifica *T* sono dei *T-illata* se e solo se essi sono considerati in *T* come esistenti indipendentemente dal loro essere oggetto di *T* stessa. In altre parole sono *T-illata* quegli oggetti di *T* che *T* stessa pone come esistenti indipendentemente da *T*.
- 2) Tutti i *T*-oggetti, la cui esistenza è considerata in *T* come dipendente dall'esistenza di *T-illata*, sono dei *T-abstracta*.

Ad esempio, per chi accetti il realismo scientifico, è sì vero che solo entro la meccanica quantistica possiamo dire che cosa sia un elettrone, ma l'esistenza degli elettroni non dipende dal fatto che tale teoria li descrive in un certo modo. Gli elettroni esistevano ben prima che i fisici ne scoprissero l'esistenza e li definissero in un certo modo! Gli elettroni sono perciò degli *illata* della meccanica quantistica. In termini più precisi per qualsiasi teoria *T*, quando *T* sia stata formalizzata, i *T-illata* sono i valori delle variabili di *T* che ricadano entro il dominio di un quantificatore. Ovviamente – poiché, a seconda del tipo di logica prescelto per la formalizzazione di *T*, sia le variabili individuali sia quelle predicative possono ricadere nel dominio di un quantificatore esistenziale – tanto degli individui

²⁸ Le definizioni che do degli *illata* e degli *abstracta*, pur riprendendo quelle di Reichenbach e di Dennett, non le rispecchiano pienamente.

quanto delle proprietà o delle relazioni possono essere degli *illata*. T potrà avere cioè, a seconda di come venga formalizzata, degli impegni ontologici o di tipo nominalistico o di tipo essenzialistico. Potrà anzi darsi il caso che si preferisca essere nominalisti se T appartiene ad un certo ordine di discorso e essenzialisti se T appartiene ad un altro ordine di discorso.

Comunque sia, se riteniamo che in T un certo concetto possa essere espresso solo mediante variabili predicative o relazionali alle quali non è possibile applicare un quantificatore, allora l'estensione di tali predicatori o relazioni è in T solo una collezione d'individui creata dal concetto stesso e non un tutto ontologicamente preesistente; è il concetto che, nel definire tale tutto, lo costituisce. In tal caso il riferimento di questo concetto è un *abstractum* che esiste solo indirettamente tramite gli *illata* che lo realizzano.

Dennett, nell'applicare la distinzione tra *illata* e *abstracta* al problema mente-corpo, la identifica sostanzialmente con quella tra realtà fisica e realtà mentale (concependo la mente, funzionalisticamente, come una "macchina virtuale" implementata dall'attività cerebrale). Credo però che una tale identificazione porti a confondere gli ordini di discorso con i livelli d'analisi e che convenga perciò considerare la distinzione tra *illata* e *abstracta* non come una distinzione inter-teorica, bensì come una distinzione intra-teorica interna ad un medesimo ordine di discorso, sebbene essa si ritrovi trasversalmente in ogni ordine di discorso.

Inoltre gli oggetti interni di un ordine di discorso di grado L_n (non importa se sono L_n -*illata* o L_n -*abstracta*), qualora vengano implementati da oggetti di grado inferiore, sono delle realtà virtuali. Per rendere plausibile questa conclusione è opportuno distinguere, a differenza di Dennett, ciò che è virtuale da ciò che è astratto. Ad esempio un mal di denti, al livello di descrizione proprio della *folk psychology*, è qualcosa di purtroppo molto concreto! Tuttavia, se i fisicalisti hanno ragione, il mal di denti è la ridescrizione in termini psicologici di uno stato ontologicamente riducibile a certi processi neurologici. E quindi il mal di denti, in quanto fenomeno psicologico, è qualcosa, al tempo stesso, di concreto e di virtuale. Per questo non conviene identificare ciò che è virtuale con ciò che è astratto. In termini più generali l'esser virtuale si riferisce ad una relazione concettuale di

ridescrizione tra due diversi ordini di discorso; la relazione *illatum-abstractum* intercorre invece tra termini che, pur distribuiti tra livelli d'analisi distinti e gerarchicamente ordinati, appartengono allo stesso ordine di discorso.

Distinguere tra *illata* e *abstracta* non è tuttavia sufficiente per chiarire a che cosa si riferiscano tutti i termini di una teoria scientifica. Già in fisica compaiono dei termini o delle espressioni che parlano di cose a rigore inesistenti, come ad esempio i gas perfetti, il moto perpetuo, i sistemi fisici perfettamente chiusi ecc. Considerarli degli *abstracta* è perciò una forzatura, perché non si vede da quali *illata* essi possano essere mai implementati.

Si potrà certo obiettare a questa osservazione che in effetti quei termini della fisica che si riferiscono a cose inesistenti sono delle idealizzazioni che, pur non essendo realizzabili in modo completo, possono esserlo tuttavia con una certa approssimazione. Ma questa risposta, ammesso e non concesso che valga in fisica, certamente non vale nelle scienze umane e sociali, perché lì troviamo dei concetti, come ad esempio quello di libero arbitrio, che, se i “deterministi duri” o gli eliminativisti hanno ragione, designano non delle idealizzazioni, ma delle vere e proprie illusioni; illusioni molto utili peraltro nella vita quotidiana, che una teoria scientifica della mente non può certo ignorare. E tuttavia illusioni (Nannini 2007a, pp. 135-158)! Conviene perciò, per rendere sensato l'uso di termini che si riferiscano a entità (o proprietà, capacità ecc.) inesistenti, introdurre accanto agli *illata* ed agli *abstracta* la categoria dei *ficta*, ossia delle finzioni, delle costruzioni teoriche che designano qualcosa d'inesistente, siano esse idealizzazioni o vere e proprie illusioni:

- 3) Un *T-fictum* è un *T*-oggetto che non è completamente implementabile mediante oggetti appartenenti all'ordine di discorso della fisico-chimica.

La differenza tra gli *abstracta* ed i *ficta* è che gli *abstracta*, essendo implementati da *illata* del loro medesimo ordine di discorso che sono a loro volta implementati da oggetti ontologicamente riducibili a fenomeni fisico-chimici, sono essi stessi riducibili in ultima analisi a fenomeni fisico-chimici; i *ficta* invece sono delle idealizzazioni o delle illusioni che, pur essendo ontologicamente irriducibili a fenomeni

fisico-chimici e quindi non godendo in alcun modo di esistenza assoluta, tuttavia è pragmaticamente utile considerare come dotati di esistenza relativa entro il loro ordine di discorso; e questa finzione è giustificata dall'esistenza assoluta a livello fisico-chimico di un "trucco" che la genera. Ad esempio, sempre se i "deterministi duri" o gli eliminativisti hanno ragione, il libero arbitrio – come già detto – non esiste; ma l'illusione di essere liberi è invece un sentimento vissuto come realmente esistente in termini psicologici, il cui contenuto, pur illusorio, viene però parzialmente giustificato dal fatto che la dinamica di un sistema complesso fisico qual è il cervello umano non è in linea di principio prevedibile.

Si obietterà che introdurre la categoria dei *ficta* è qualcosa che ripugna allo spirito antimetafisico della filosofia analitica. Perché mai un'ontologia scientifica dovrebbe ammettere delle diafane entità non esistenti? Se si è costretti a parlarne, lo si faccia in modo tale da evitare di introdurre nel linguaggio scientifico dei nomi il cui riferimento sia inesistente! Lo si faccia ad esempio, come propose B. Russell nel suo celebre *On denoting* (Russell 1905), trasformando i nomi di oggetti inesistenti in predicati la cui estensione sia nulla. In tal modo, com'è noto, l'enunciato "L'attuale re di Francia è calvo", pronunciato in un'epoca nella quale la Francia è una repubblica e quindi non esiste alcun attuale re di Francia, diviene "Esiste una ed una sola cosa x tale che x è calva ed è attualmente re di Francia". E questo è un enunciato falso, perché implica l'esistenza di una cosa che è re di Francia in un periodo in cui la Francia è una repubblica. Oppure si potrebbe sostenere, come fece P.F. Strawson (1950), che qualsiasi enunciato che parli di qualcosa di inesistente è semplicemente insensato. In entrambi i casi, seguendo o Russell o Strawson, si riuscirà a liberare la nostra ontologia dagli scomodi *ficta*!

Tuttavia nessuna di queste mosse è realmente soddisfacente. Infatti secondo la soluzione di Russell è falso non solo che l'attuale re di Francia è calvo, ma anche che non è calvo (dato che anche quest'ultima affermazione ne implica l'esistenza), mentre entrambi gli enunciati sono per Strawson semplicemente privi di senso. Pertanto, visto che i gas perfetti, al pari dell'attuale re di Francia, a rigore non esistono, sarà necessariamente falso (o insensato) asserire che essi rispettano (o non rispettano) la legge di Boyle-Mariotte. Il che appare molto contro-intuitivo, dal momento che qualunque fisico pensa che i

gas perfetti rispettino quella legge! Meglio perciò ammettere la possibilità che alcuni nomi, non solo nel linguaggio ordinario ma anche in una teoria scientifica, si riferiscano a cose inesistenti, ossia a *ficta*, a finzioni. Il che non significa ovviamente che qualunque finzione possa figurare come un *fictum* di una qualche teoria scientifica. Saranno ammissibili in una teoria scientifica perciò solo quei *ficta* che risultino pragmaticamente utili per prevedere o spiegare dei fenomeni empiricamente accertabili. Ad esempio è molto dubbio che un concetto di origine teologica come quello di anima immortale possa essere usato legittimamente in alcuna scienza, comprese le scienze umane e sociali.

4. *Fatti socio-culturali, fenomeni mentali e processi cerebrali*

Quanto sopra detto riguardo agli ordini di discorso, ai livelli d'analisi, alle definizioni di *illatum*, *abstractum* e *fictum* nonché alla distinzione tra reale e virtuale fornisce un primo abbozzo di un'ontologia filosofica interdisciplinare valida sia per le scienze *hard* (fisica, chimica e biologia) sia per le scienze cognitive, umane e sociali. Tale ontologia è naturalistico-fisicalistica. Volendone riassumere i tratti essenziali si può dire che in essa solo ciò che è oggetto della fisico-chimica è veramente reale, mentre tutti gli altri fenomeni sono o delle realtà virtuali o dei *ficta*.

L'ontologia naturalistico-fisicalistica sopra abbozzata costituisce inoltre una cornice teorica entro la quale è più facile valutare se il senso comune ci inganna o meno quando induce ancor oggi molte persone a considerare la relazione tra mente e corpo come profondamente diversa, per ragioni metafisiche, dalla relazione che intercorre tra fatti sociali e azioni individuali.

Cominciamo da un esame di quest'ultima relazione. Il senso comune ci suggerisce qualcosa di strano quando ci induce a pensare ad esempio che non ci sono eserciti senza soldati o che non ci sarebbe l'Italia se non ci fossero gli italiani? Per il naturalista fisicalista assolutamente no! Inoltre, se gli individui implementano fatti sociali attraverso le loro azioni, è perché essi sono dotati di una mente. Ciò non vuol dire ovviamente che i fatti o le istituzioni sociali, per essere implementati, non richiedano anche oggetti materiali o processi fisici: ad esempio la religione cristiana senza le chiese non è concepibile; ed una qualsiasi azione non può essere eseguita senza muovere qualche parte del corpo. Un movimento corporeo diviene tuttavia un'azione

solo quando è compiuto volontariamente. E certi oggetti materiali fanno parte di istituzioni sociali solo perché degli esseri umani attribuiscono loro un certo significato: ad esempio, se i cassieri ed i clienti delle banche non fossero persone dotate di una mente, certi pezzi di carta non potrebbero essere degli assegni. Pertanto ciò che è essenziale affinché possano esistere i fenomeni sociali e culturali è che degli esseri umani abbiano determinate intenzioni, credenze, emozioni ecc. Le società e le culture sono, in ultima analisi, delle realtà virtuali implementate completamente (o, in parte, dei *ficta* implementati solo parzialmente) dall'attività di menti individuali.

Si obietterà a tutto ciò che le società funzionano solo perché esistono le culture (in senso antropologico). Ad esempio l'Italia non potrebbe esistere se gli italiani non potessero comunicare tra loro; e non potrebbero comunicare se non condividessero la conoscenza di una medesima lingua. Ma le culture – proseguono i critici del naturalismo fisicalistico – presentano una dimensione normativa che non si vede come possa essere implementata da processi mentali empiricamente rilevabili. Ad esempio la lingua italiana include una grammatica e, più in generale, degli *standard* di correttezza che hanno una loro oggettività indipendentemente dal modo, più o meno corretto, nel quale l'italiano viene effettivamente parlato. Se qualcuno afferma che l'italiano si sta degradando perché pochi italiani usano ormai il congiuntivo tutte le volte che sarebbe necessario, si potrà forse pensare che questi è un vecchio cruscante, ma non si potrà sostenere che dice qualcosa di contraddittorio, come invece accadrebbe se tutto ciò che è grammaticalmente corretto coincidesse per definizione con ciò che è di uso corrente. In modo analogo qualcuno può sostenere che una certa azione in certe circostanze è l'unica moralmente giusta, pur nella consapevolezza che quasi nessuno al mondo, posto in quelle circostanze, la compie davvero. Anche in questo caso si potrà sostenere che l'etica di chi la pensa così è eccessivamente rigoristica, ma non che questi dice qualcosa di assurdo.

Più in generale molti tratti culturali – valori e norme morali, criteri estetici, regole linguistiche, regole di galateo ecc. – sembrano sottrarsi in linea di principio alle smentite che possono venire dal comportamento effettivamente tenuto dagli uomini in carne ed ossa. L'esistenza di tali tratti culturali empiricamente non *defeasible* sembra creare perciò un grosso problema al fisicalista. Per un platonico non

c'è alcuna difficoltà a sostenere che il Bene in sé è cosa diversa da qualsiasi azione effettivamente compiuta; forse nessuna azione è mai davvero completamente buona. Ma, come può giustificare una simile distinzione chi pensi che i fatti sociali e culturali, se esistono, devono essere implementati da atti individuali compiuti in virtù di certe intenzioni e credenze?

È qui che al naturalista fisicalista torna utile la categoria dei *ficta*. I valori e le norme morali, i principi giuridici, i criteri estetici, le regole linguistiche ecc. sono delle costruzioni teoriche, sono delle convenzioni umane; sono in una parola dei *ficta* e perciò, al pari di Faust o Renzo e Lucia, possiamo amarli o odiarli, possiamo dedicare ad essi volumi di critica che occupano intere biblioteche, possiamo trovarli utili o perniciosi, possiamo combattere e persino morire per la diffusione di atteggiamenti di accettazione o di rifiuto nei loro confronti, ma a rigore essi non esistono e perciò non possono essere causa di alcunché: il nulla non causa nulla. Esistono però gli atti mentali di coloro che credono in essi. Pertanto, quando sembra che qualche valore o qualche norma influenzi il comportamento umano, occorre precisare che non il valore o la norma è ciò che influenza il relativo comportamento, bensì il fatto che tale valore o tale norma è l'oggetto interno di qualche atto mentale; un atto che, in quanto implementato da processi cerebrali, è qualcosa di fisico e quindi può essere, esso sì, la causa di eventi nel mondo fisico senza lederne la chiusura.

Si noti inoltre che la distinzione tra ciò che è reale e ciò che è virtuale è relativa agli ordini di discorso che vengono presi in considerazione. Pertanto è sensato dire che gli atti mentali incorporati nelle azioni individuali sono delle realtà implementanti, in tutto o in parte, rispetto ai fatti sociali e culturali, ma che quegli stessi atti mentali sono delle realtà virtuali rispetto ai processi cerebrali da cui sono implementati (di nuovo in tutto o in parte). I fatti sociali e culturali sono perciò per il naturalista fisicalista delle realtà virtuali o dei *ficta* per così dire alla seconda potenza, perché sono implementati in tutto o in parte da atti mentali che, a loro volta, sono implementati in tutto o in parte da processi cerebrali. E ciò è coerente ovviamente con l'assunto di fondo del fisicalismo, secondo il quale solo il mondo fisico è veramente reale e non abbisogna di alcuna ulteriore implementazione.

Il naturalista fisicalista non vede perciò nessun ostacolo, a differenza di un senso comune molto diffuso, a trasferire per analogia sul rapporto tra mente e corpo quel riduzionismo ontologico che appare più che plausibile quando venga applicato al rapporto tra fatti sociali e azioni individuali. Inoltre questa ipotesi è compatibile con il credere che, sebbene tutti gli atti mentali (quando non siano dei *ficta*) da un punto di vista ontologico siano solo realtà virtuali implementate da processi cerebrali, può darsi benissimo che essi non siano tutti ugualmente spiegabili in termini neurologici e che dunque la psicologia continui a mantenere ancora per molto tempo (o forse per sempre) la sua autonomia rispetto alle neuroscienze: il riduzionismo ontologico non deve essere confuso con quello epistemologico.

In conclusione il naturalista fisicalista, appoggiandosi all'ontologia filosofica sopra delineata, può ragionevolmente sostenere che tutti gli atti o stati mentali della *folk psychology*, dopo essere stati opportunamente analizzati e eventualmente disarticolati in varie componenti, possono essere ripartiti in tre gruppi:

- Quelli che sono direttamente identici a processi cerebrali o a proprietà della dinamica del cervello nel suo complesso: ad esempio il mal di denti.
- Quelli che sono analizzabili come stati funzionali (e quindi virtuali) variamente implementati dall'attività cerebrale: ad esempio il vedere un colore.
- Quelli inesistenti, ossia i *ficta* mentali, i cui concetti, pur indispensabili nella vita di tutti i giorni, forse potranno essere espunti dalla neuro-psicologia del futuro e potranno essere sostituiti con i concetti di processi cerebrali capaci di generare l'illusione dell'esistenza dei *ficta* stessi: ad esempio i *qualia*, in quanto distinti dall'atto mentale del percepirli, o – come già detto – il libero arbitrio).

Se viene chiesto infine quale sia lo *status* epistemologico di una tale ipotesi sull'ontologia del mentale, il naturalista fisicalista può rispondere, nello spirito quiniano della continuità tra filosofia e scienza, che essa non è né una verità *a priori* né un'ipotesi scientifica direttamente falsificabile mediante qualche *experimentum crucis*: è

piuttosto una cornice teorica costruita per favorire lo sviluppo della ricerca scientifica stessa. Solo nel lungo periodo, a seconda dei risultati ottenuti, si potrà dire se l'indirizzo generale che essa propone è fecondo oppure sterile.

Capitolo quarto

Tempo e coscienza

1. *Lo Hard Problem e l'eliminativismo*

Spiegare come la coscienza fenomenica possa emergere dall'attività cerebrale è – secondo una posizione oggi molto diffusa non solo i tra filosofi della mente ma anche tra molti scienziati cognitivi e neuroscienziati – lo *Hard Problem* per eccellenza di chi si occupi a vario titolo della mente e del suo rapporto con il cervello (cfr. *supra* capitolo primo e capitolo secondo). Molti ritengono anzi che quel problema sia non solo *hard*, ma addirittura insolubile in linea di principio. A costoro i naturalisti (e in particolare gli eliminativisti) possono obiettare che, se c'è uno scarto incolmabile tra due modi diversi di riferirsi alle nostre esperienze soggettive, ora colte introspettivamente nei termini della *folk psychology* e ora spiegate nei termini delle neuroscienze, occorre abbandonare i concetti pre-scientifici con i quali sono stati descritti sinora i processi mentali (inclusi gli stati di coscienza) e sostituirli con concetti nuovi attinti dalle neuroscienze stesse. Queste ultime – sostengono gli eliminativisti – stanno provocando un vero e proprio mutamento di paradigma nella scienza della mente e, di conseguenza, nell'immagine che noi esseri umani abbiamo di noi stessi (cfr. *supra* capitolo primo).

Tuttavia, per quanto grandi siano stati negli ultimi venti anni i progressi delle neuroscienze, gli eliminativisti continueranno ad essere poco convincenti finché non riusciranno a ricostruire un'immagine scientifica della mente umana all'interno della quale sia possibile formulare una spiegazione plausibile del perché tale immagine scientifica ci appaia introspettivamente, sotto molti aspetti, come del tutto inaccettabile. Quale strada prendere allora? La soluzione che qui propongo è la seguente:

- a) riesaminare brevemente un caso esemplare (forse *il* caso esemplare per antonomasia) di mutamento di paradigma: quello che ha segnato il passaggio dalla meccanica classica alla teoria della relatività di Einstein;
- b) trovare alcune analogie, accanto ovviamente anche a profonde differenze, tra quel mutamento di paradigma e quello che, in corso sotto i nostri occhi, ci sta portando da una visione dell'uomo e del suo mondo ancora incentrata sul dualismo

cartesiano tra mente e corpo ad una nuova concezione scientifica della mente come attività del cervello nella sua interazione con il mondo esterno;

- c) desumere da tutto ciò, in combinazione con gli studi sulle basi neurologiche del senso del tempo, una spiegazione quanto meno plausibile del perché l'immagine dell'autocoscienza come di un Io pilota del corpo ci appaia da un lato introspettivamente come più adeguata (e più utile nella vita di tutti i giorni) di una teoria neurologica della coscienza e dell'autocoscienza che le riduca ad essere proprietà della dinamica cerebrale, ma dall'altro, sul piano scientifico, ci si presenti come fondamentalmente errata.

2. Tempo, spazio e gravità nella teoria della relatività

La teoria della relatività è talmente nota che sarebbe ridicolo e presuntuoso offrirne una ennesima esposizione divulgativa²⁹. Mi limiterò perciò a mettere in luce – molto sinteticamente ed in modo facilmente comprensibile per un lettore che, come me, abbia una formazione prevalentemente filosofica e umanistica – che l'epicentro del terremoto che Einstein ha provocato nella fisica contemporanea è dipeso essenzialmente da un mutamento del concetto di tempo; mutamento che poi ha prodotto, a cascata, analoghi cambiamenti in tutti i concetti fondamentali della fisica.

Chiariamo questo punto mediante il *Gedankenexperiment* proposto da Einstein stesso nella sua esposizione divulgativa della teoria della relatività (Einstein 1917). Cadono due fulmini lungo una linea ferroviaria a circa tre chilometri di distanza l'uno dall'altro. Due osservatori – l'uno (*Y*) fermo accanto ai binari nel punto intermedio tra i due fulmini, l'altro (*X*) in un treno che si muove lungo i binari ad una velocità pari ai $3/5$ della velocità della luce – vogliono stabilire se i due fulmini siano caduti simultaneamente o no. Non potendo accertare questo fatto in modo diretto, dal momento che non possono rivolgere il loro sguardo simultaneamente in direzioni opposte, si servono entrambi di due specchi orientati a 45 gradi, l'uno rivolto verso il primo fulmine e l'altro verso il secondo. Supponiamo ora che

²⁹ La letteratura sulla teoria della relatività è sterminata. Molto utile è stata per me l'introduzione di Boniolo e Dorato (1997).

l'osservatore a terra rilevi che la luce emessa dai due fulmini è stata riflessa dai suoi due specchi nello stesso istante. Dato che i due specchi si trovano alla stessa distanza dai due rispettivi fulmini e, come provato da J. Maxwell già nella seconda metà dell'Ottocento, la velocità della luce è la stessa in tutte le direzioni, egli, se vede che gli specchi hanno riflesso la luce emessa dai fulmini nello stesso istante, ne concluderà che i fulmini stessi sono caduti, una frazione di secondo prima, nello stesso istante. L'osservatore sul treno invece vedrà riflessa nei suoi specchi prima la luce emessa dal fulmine a cui va incontro e dopo la luce emessa dal fulmine dal quale si sta allontanando. Pertanto i due fulmini appariranno come simultanei solo al primo osservatore (cfr. fig. 7).

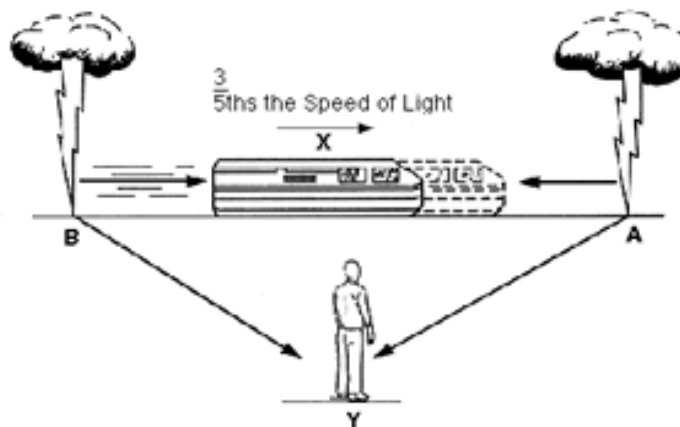


Fig. 7

Immagine scaricata da: www.quora.com

Come interpretare questa differenza tra i due osservatori? Secondo la meccanica classica newtoniana i due fulmini saranno o non saranno in realtà simultanei a seconda della loro collocazione nello spazio e nel tempo assoluti indipendentemente dal modo nel quale essi appaiono a questo o a quell'osservatore. Esiste insomma un sistema di riferimento privilegiato nell'universo. Di conseguenza la velocità della luce, che si muove a velocità costante in tutte le direzioni solo nel sistema di riferimento privilegiato (ossia, si diceva alla fine dell'Ottocento, nell'etere), varierà invece per ciascun osservatore a seconda della velocità del moto dell'osservatore in tale sistema. Nell'esempio dato da Einstein – poiché per il senso comune sembra

ovvio ammettere che è il treno a muoversi sui binari, mentre sarebbe molto strano pensare che il treno sta fermo ed i binari insieme a tutto l'ambiente circostante gli scorrono sotto come un *tapis roulant* – l'osservatore fermo lungo i binari penserà che i fulmini sono effettivamente caduti simultaneamente e che l'osservatore sul treno dovrebbe correggere la sua osservazione tenendo conto del suo moto rispetto ai binari. In conclusione, se il tempo è assoluto, la velocità della luce non può essere la stessa per osservatori che siano in moto l'uno rispetto all'altro: all'osservatore sul treno il raggio di luce che gli viene incontro apparirà più veloce del raggio di luce che invece lo insegue, mentre per l'osservatore a terra i due raggi di luce si muoveranno alla medesima velocità. Secondo la meccanica classica l'osservatore a terra, per il quale la velocità della luce è la medesima nelle due opposte direzioni (in accordo con la teoria di Maxwell), osserva la realtà per come essa effettivamente è, mentre l'osservatore sul treno è ingannato dal suo moto rispetto al sistema di riferimento privilegiato nel quale i binari sono fermi.

Tuttavia questa conclusione è fallace per due ordini di motivi. In primo luogo, anche ammesso che esista un sistema di riferimento privilegiato dello spazio e del tempo assoluti, l'osservatore a terra non sarà fermo rispetto ad esso (moto di rotazione e di rivoluzione della Terra rispetto al Sole, moto del Sole rispetto alle altre stelle della nostra galassia, moto della nostra galassia rispetto alle altre galassie ecc.). Non è quindi mai possibile stabilire qui sulla Terra quale sia il sistema di riferimento privilegiato rispetto al quale la velocità della luce è identica in tutte le direzioni. Ma a questa obiezione il sostenitore della meccanica classica può facilmente rispondere che, sebbene sia molto difficile stabilire in pratica quale sia il sistema di riferimento privilegiato ("punto di vista di Dio") rispetto al quale soltanto la velocità della luce è costante, nondimeno è concepibile che tale sistema esista.

Questa replica va però incontro – secondo Einstein – ad una nuova ed insuperabile obiezione: le leggi dell'elettromagnetismo scoperte e formulate matematicamente da Maxwell stabiliscono che la velocità della luce è costante in tutte le direzioni per tutti gli osservatori indipendentemente dai loro eventuali movimenti gli uni rispetto agli altri. Sembra esserci dunque un contrasto insanabile tra la meccanica classica di Newton e la teoria dell'elettromagnetismo di Maxwell.

Come conciliare questo contrasto? Questo era il grande problema che si ponevano i fisici tra la fine dell'Ottocento ed i primi del Novecento. Il giovane Einstein lo risolse semplicemente smentendo Newton! L'etere – sostenne Einstein – non esiste; nessun sistema di riferimento è privilegiato! Non ha senso chiedersi, nell'esempio summenzionato, se i due fulmini siano caduti simultaneamente o no rispetto ad un tempo assoluto che scorrerebbe in modo uniforme per tutti gli osservatori. I due fulmini sono caduti simultaneamente per il primo osservatore, ma non per il secondo. Questa conclusione appare assurda solo a chi fatichi a liberarsi del pregiudizio – in effetti non sostenuto da alcun dato empirico – secondo il quale esiste un tempo assoluto che scorre allo stesso ritmo per tutti gli osservatori.

Più in dettaglio, se due osservatori sul treno sincronizzano i loro orologi scambiandosi dei segnali luminosi e lo stesso fanno due osservatori a terra, gli orologi sul treno e quelli a terra segneranno ore diverse. E – ciò che più stupisce – dal punto di vista degli osservatori a terra gli orologi sul treno andranno indietro, ma per gli osservatori sul treno saranno gli orologi a terra ad andare indietro. Inoltre, risulterà che per ciascun gruppo di osservatori il rallentamento degli orologi dell'altro gruppo sarà accompagnato, sempre nell'altro gruppo, da un accorciamento delle distanze nello spazio. E queste modifiche nelle misure del tempo e dello spazio avverranno in modo tale che la velocità della luce risulterà essere la medesima in entrambi i gruppi.

Queste conclusioni valgono solo se gli osservatori interessati si muovono gli uni rispetto agli altri di moto rettilineo uniforme, ossia se si muovono di moto inerziale. Per questo la prima formulazione della teoria di Einstein, risalente al 1905, va sotto il nome di teoria speciale (o ristretta) della relatività. Dopo dieci anni di studi Einstein riuscì nel 1915 a estendere la sua teoria anche ad osservatori che si muovano gli uni rispetto agli altri di moto accelerato ed a formulare la cosiddetta teoria generale della relatività. In essa Einstein, grazie al “principio di equivalenza” da lui introdotto tra accelerazione e gravità, sostenne, facendo uso della geometria non euclidea di G.F.B. Riemann, che il moto accelerato di un corpo entro un campo gravitazionale è equivalente ad un moto inerziale descritto entro un sistema di riferimento non-euclideo di tipo riemanniano. In altre parole la presenza di una massa deforma intorno a sé, per così dire, le coordinate spazio-temporali di tutti i sistemi di riferimento:

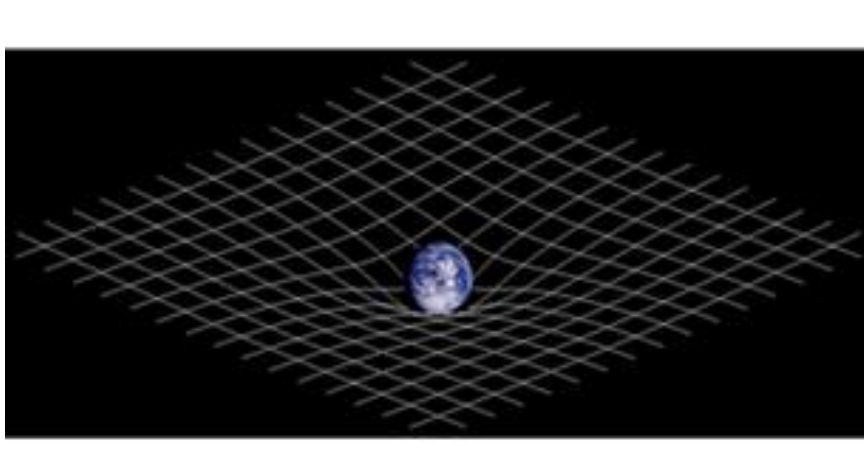


Fig. 8

La Terra, ad esempio, gira intorno al Sole non perché attratta da quest'ultimo, bensì perché si muove di moto inerziale in uno spazio-tempo reso incurvato dalla presenza della massa solare. Inerzia e gravità sono riconciliate entro una teoria unica. Ma il concetto di gravità risulta profondamente modificato rispetto alla meccanica classica: la gravità secondo Einstein non è più una forza, bensì un indice di curvatura dello spazio-tempo nelle sue quattro dimensioni. La gravità newtoniana nella teoria generale della relatività in un certo senso scompare e viene sostituita dalla curvatura di uno spazio-tempo non euclideo, entro il quale i corpi, in assenza di altre forze diverse dalla gravità, si muovono sempre per pura inerzia lungo la loro geodetica, ossia lungo la linea più breve tra due punti in uno spazio incurvato).

In conclusione tutti i concetti fondamentali della meccanica classica, dal tempo allo spazio e alla gravità, vengono sostituiti nella teoria della relatività da altri concetti che, a differenza dei primi, sono molto lontani dal senso comune. Non sono più i fenomeni direttamente osservabili che danno senso ai concetti teorici della scienza; al contrario sono questi ultimi che ci costringono a considerare come ingenua e solo apparente l'immagine del mondo fornita direttamente dai sensi. Nel contrasto tra visione scientifica del mondo e senso comune è quest'ultimo che deve faticosamente e contro-intuitivamente adeguarsi alla prima e non viceversa.

3. Coscienza e gravità

Instaurare un paragone tra il cambiamento di paradigma che investì la fisica un secolo fa ad opera di Einstein e la svolta cognitiva oggi in corso impressa dalle scienze cognitive in generale e dalle neuroscienze in particolare all'immagine che noi esseri umani abbiamo di noi stessi può sembrare molto ardito. E tuttavia esistono, quanto meno, alcune analogie tra il modo nel quale nacque e fu recepita un secolo fa la teoria della relatività e le obiezioni sollevate oggi contro la possibilità di naturalizzare la coscienza. Ad esempio il dibattito suscitato alla fine dell'Ottocento da esperimenti, come quello di A.A. Michelson e E.W. Morley, che rendevano evidente l'incompatibilità della meccanica classica con la costanza della velocità della luce per tutti gli osservatori – dibattito da cui, come abbiamo detto, prese le mosse il giovane Einstein per elaborare la teoria speciale della relatività – presenta una certa somiglianza con la discussione accesa oggi intorno all'esistenza del libero arbitrio dopo gli esperimenti neurologici di B. Libet degli anni Ottanta del XX secolo; esperimenti che ne dimostrerebbero l'inesistenza, dato che l'agente prende coscienza di aver deciso di compiere volontariamente e liberamente proprio in quell'istante l'azione di piegare un dito con circa 300 millisecondi di ritardo rispetto alla comparsa nella sua corteccia cerebrale di “potenziali di prontezza” correlabili con quell'azione stessa (Libet 2004) (cfr. *supra* capitolo secondo).

Alla fine dell'Ottocento, prima della soluzione radicale proposta da Einstein, si cercò di rendere conto del fallimento dell'esperimento di Michelson e Morley o con l'inadeguatezza della strumentazione con cui l'esperimento stesso era stato realizzato o mediante ipotesi aggiuntive *ad hoc* che non negassero la validità complessiva delle teorie fisiche allora dominanti. Molto successo ebbe infatti proprio per questo l'ipotesi di H. Lorentz e G.F. Fitzgerald sull'accorciamento fisico degli oggetti materiali nella direzione del moto. In modo analogo oggi molti si rifiutano di accettare l'importanza degli esperimenti di Libet per una nuova concezione della mente e della coscienza o sostenendo la loro inattendibilità metodologica oppure contestando la loro rilevanza filosofica grazie ad una revisione neo-compatibilistica del concetto di libero arbitrio (ad es. Dennett 1984 e 2003); revisione che si presenta, al pari dell'ipotesi di Lorentz e Fitzgerald, come un'aggiunta filosofica *ad hoc* ai dati empirici disponibili.

Inoltre, come molti sostengono oggi l'irriducibilità della coscienza ai suoi correlati neurali, così dopo il 1905 un filosofo allora notissimo qual era all'epoca H. Bergson scrisse addirittura un libro contro la teoria della relatività per sostenere la differenza, a suo avviso incolumabile, tra il tempo spazializzato della fisica e il tempo vissuto (o durata) del senso comune: l'uno semplice costrutto teorico della scienza, l'altro dato empirico introspettivamente indubitabile, coincidente con quel divenire interiore alla nostra coscienza del quale abbiamo diretta ed incontrovertibile certezza (Bergson 1922).

Se queste somiglianze di tipo storico tra i due mutamenti di paradigma – il primo del secolo scorso in fisica, il secondo odierno nelle scienze cognitive – apparissero troppo superficiali per essere prese seriamente in considerazione, si deve aggiungere che tra le due svolte concettuali esistono analogie teorico-formali ben più profonde. Certo, chi voglia oggi naturalizzare la mente e la coscienza non può avvalersi purtroppo di teorie psicologico-cognitive o neuro-scientifiche già pienamente sviluppate come lo è in fisica la teoria della relatività. Tuttavia c'è qualcosa in comune tra la trasformazione del concetto di tempo (e, a cascata, di molti altri concetti della fisica) che dobbiamo alla teoria della relatività, da un lato, e la svolta cognitiva che riduce la coscienza ad essere il formato comune alla dinamica dei processi cerebrali che vanno a costituire il *Dynamic Core*, dall'altro (cfr. *supra* capitolo secondo)? Insieme a molte e indubbie differenze colpiscono quanto meno alcune somiglianze.

Anzitutto – come ho già detto detto – la forza di gravità nella teoria generale della relatività viene, per così dire, ad essere assorbita dalla curvatura di uno spazio-tempo non più euclideo. Di conseguenza le leggi fondamentali della fisica non possono essere formulate mediante equazioni lineari (Einstein 1949, p. 46). Ma che cosa succede se – volendo dare una rappresentazione semplificata, a scala locale e per basse velocità, del moto di un corpo – torniamo alla meccanica classica? Ecco che le coordinate spazio-temporali tornano ad essere euclidee, ossia la curvatura dello spazio-tempo scompare, e in compenso compare di nuovo la forza di gravità.

Ora, dato che, come prima accennato a proposito della visione dei colori (cfr. *supra* capitolo secondo), la dinamica cerebrale che implementa il *Dynamic core* è suscettibile di essere meta-rappresentata come una "traiettoria" di *pattern* di attività cerebrale in

un iper-spazio vettoriale, si può azzardare l'ipotesi che – quando il cervello si auto-monitorizza e crea una “rappresentazione neuronale” di parte della sua stessa attività, rappresentandosi come un *Io** per rendere più efficace il coordinamento senso-motorio (cfr. *infra* capitolo settimo) – esso dia della sua stessa dinamica una rappresentazione semplificata che è, in un certo senso, formalmente analoga alla rappresentazione che la meccanica classica dà del moto dei corpi in un campo gravitazionale per effetto della sola forza di gravità in assenza di altre forze. In altre parole, come la meccanica classica, ignorando la curvatura dello spazio-tempo, separa le leggi del moto dalla struttura del campo gravitazionale (Einstein 1949, p. 41), in modo analogo può darsi che il cervello, nella rappresentazione che implicitamente (e nel formato coscienziale che è ad esso proprio) dà a sé della sua stessa attività, tenda a isolare la coscienza funzionale ed il *Self* dalla dinamica cerebrale che li implementa. In tal modo quella dinamica cerebrale viene ad essere rappresentata cartesianamente come una coscienza fenomenica ed un *Io* che agiscono sul cervello, per così dire, dall'esterno.

Perché il cervello stesso si auto-inganni fino a questo punto non è facile a dirsi. Ma l'ipotesi più probabile è che la dinamica cerebrale che implementa la coscienza sia utile – al pari di altri trucchi informatici operanti anche nei singoli neuroni, come ad esempio la digitalizzazione dei segnali elettrici che corrono lungo l'assone (Kandel *et al.* 1995, pp. 35-36) – per ridurre al minimo le perdite della quantità d'informazione che viene trasferita da un'area cerebrale all'altra. Ebbene, è poi così improbabile che questa stessa dinamica richieda, per essere efficace, di essere monitorata e internamente coordinata e che il modo trovato dalla selezione naturale per monitorarla e coordinarla sia di darle una “rappresentazione” interna nel formato della coscienza fenomenica, un formato nel quale tale dinamica appare (al cervello stesso) come isolata dal resto dell'attività cerebrale e ad essa funzionalmente sovra-ordinata?

Se è così, allora noi – come già detto –, sebbene abbiamo una spontanea inclinazione a sentirci i timonieri del nostro corpo (secondo la celebre immagine aristotelica), siamo in realtà solo un servo-meccanismo del nostro cervello. Ma siamo un servo-meccanismo che, quando il nostro cervello venga programmato mediante l'esperienza della vita sociale (apprendimento di una lingua, di una

cultura, di una morale ecc.), può divenire un “servo-meccanismo ribelle”. Madre Natura infatti, sebbene abbia dotato noi esseri umani del cervello che ancor oggi abbiamo essenzialmente affinché i nostri progenitori potessero un tempo sopravvivere prima nella foresta e poi nella savana, ha fatto di noi – mediante il medesimo meccanismo della selezione naturale – anche animali che vivono in società (P.S. Churchland 2011) e che sono capaci, in virtù di un gigantesco processo di esattazione, di parlare e di trasmettere le proprie esperienze da una generazione all'altra. Ed è così che quel servo-meccanismo cerebrale che è la coscienza, sviluppatosi un tempo presumibilmente solo per realizzare un miglior coordinamento senso-motorio, ha finito, collocato in un contesto sociale e culturale, non solo per creare l'illusione di un *Io** cosciente di sé e del mondo che lo circonda, ma anche per fare di esso una persona che, da un lato, ritiene di mantenere la propria identità nell'arco di tutta la propria vita, pur nella pluralità e talvolta mutevolezza dei ruoli sociali che essa viene ad assumere, e, dall'altro, interagisce con gli altri esseri umani riconoscendo in essi altrettante persone dotate di una mente cosciente simile alla sua. Ciascun essere umano finisce in tal modo per identificarsi con un *Io-persona** che, sebbene sia solo una finzione prodotta dalla dinamica cerebrale e dal contesto socio-culturale, sembra sottomettere a sé tutto il resto dell'attività mentale e riesce a porsi al centro di quel sistema di relazioni e di ruoli che è per ciascuno di noi la propria vita sociale. Ed è così che noi diveniamo gli esseri umani che siamo (cfr. *infra* capitolo settimo).

In conclusione il modo nel quale fu recepito agli inizi del secolo scorso il mutamento prodotto dalla teoria della relatività nei concetti fondamentali della meccanica classica ci insegna che, quando qualche nuova teoria scientifica sconvolge nel profondo concezioni fortemente radicate, è normale che, sebbene la nuova teoria sia empiricamente ben confermata, molti l'accettino con riluttanza. E dunque, se è vero che noi esseri umani abbiamo una forte inclinazione naturale a rifiutarci di ammettere rivoluzioni concettuali che sconvolgano le nostre convinzioni più profonde, non può darsi forse che anche la convinzione che la coscienza e l'*Io* siano qualcosa di ontologicamente irriducibile a caratteristiche dell'attività cerebrale, convinzione ancor oggi pervicacemente diffusa, sia in radice, oltre che parte essenziale della cultura prevalente da secoli in Occidente, il prodotto di un

inganno del nostro cervello? Un inganno benefico, intendiamoci, selezionato dall'evoluzione biologica a causa dei vantaggi adattativi che esso presenta. Un inganno, comunque, al quale emotivamente non possiamo sottrarci. Infatti, quando compiamo un'azione volontaria, non possiamo fare a meno di sentirci gli autori di quell'azione; ed è solo perché abbiamo quel sentimento di *agency* che siamo capaci di compiere azioni volontarie (Wegner 2002). Ma ciò non dovrebbe impedirci di ammettere che, sul piano teorico per chi aspiri a conoscere la verità (cfr. *infra* capitolo settimo), l'immagine che di noi offrono le neuroscienze è preferibile.

4. Tempo, coscienza e sincronizzazioni cerebrali

Le somiglianze di tipo storico e teorico tra il mutamento di paradigma che consentì alla fisica di passare dalla meccanica classica alla teoria della relatività ed i processi di trasformazione che investono oggi l'immagine che noi esseri umani abbiamo di noi stessi non sono inoltre solo di tipo epistemologico-formale, ma investono a ben guardare anche i rispettivi contenuti della teoria della relatività e della naturalizzazione del mentale. La teoria della relatività è nata, come si è detto, da un profondo mutamento concettuale che ha reso la durata di un intervallo temporale dipendente dal moto relativo di due osservatori che confrontino le loro misurazioni. In un certo senso la teoria della relatività ha sostituito il tempo assoluto newtoniano con una miriade di "tempi fenomenici" relativi alle misurazioni prese da differenti osservatori. E questa sostituzione ha fatto sì che la gravità venga intesa nella teoria della relatività non più come una forza, bensì come un indice di curvatura dello spazio-tempo.

Ebbene, anche l'analogia sopra proposta tra coscienza fenomenica e gravità è legata, sia pur in modo molto più indiretto, al tempo fenomenico; è legata più precisamente al modo nel quale il cervello umano accomuna l'essere coscienti ed il percepire lo scorrere del tempo. Anzitutto, se la coscienza è il formato mediante il quale il cervello monitorizza e coordina parte della sua attività, allora tale formato molto probabilmente o coincide con un processo di sincronizzazione tra circuiti neuronali oscillanti o quanto meno trova in tale sincronizzazione una sua condizione necessaria. Il legame tra coscienza fenomenica e sincronizzazione cerebrale, proposto inizialmente solo per la percezione visiva (Crick e Koch 1990), è attualmente preso in seria

considerazione per spiegare la comparsa anche di altri stati di coscienza. Si pensi in particolare alla spiegazione di certi fenomeni di *binding* come ad esempio la spiegazione del meccanismo cerebrale mediante il quale *input* sensoriali di diverso tipo, quali sono i colori, le forme ed i movimenti, sono fusi nella percezione cosciente di un oggetto avente un certo colore, mentre si muove su uno sfondo di colore diverso (W. Singer 2001 e 2004, Engel 2003, Tacca 2010).

Se ora ammettiamo che non è possibile sentirsi svegli e coscienti senza percepire, secondo dopo secondo, lo scorrere del tempo, non si può non essere colpiti dal fatto che, almeno stando a quanto dice il neurologo tedesco E. Pöppel, noi esseri umani percepiamo lo scorrere del tempo mediante due meccanismi di sincronizzazione cerebrale analoghi a quelli che sono alla base della comparsa della coscienza fenomenica (Pöppel 1997a e 1997b; Pöppel and Logothetis, 1986).

Il primo meccanismo individuato da Pöppel crea una finestra di trenta millisecondi: tutti gli stimoli visivi che cadono entro questa finestra sono percepiti come simultanei. La seconda finestra, della durata di due o tre secondi, viene costruita mettendo insieme circa cento finestre del primo tipo: tutti gli stimoli visivi che cadono entro questa seconda finestra sono percepiti in sequenza, ma al tempo stesso sono vissuti come un singolo istante di “presente ricordato” (Edelman 1989; cfr. *supra* capitolo secondo) ossia di *specious present* (James 1891, vol. 2, p. 609). In altre parole il presente psicologico, lungi dall’essere puntiforme e quindi inesteso, ha una durata di due o tre secondi circa. Pertanto la percezione della simultaneità (prima finestra) e del presente (seconda finestra) sono due differenti costruzioni del cervello che si basano entrambe su dei processi di sincronizzazione cerebrale.

Pertanto i due meccanismi cerebrali con cui percepiamo il tempo (simultaneità e presente) sono meccanismi di sincronizzazione di circuiti cerebrali oscillanti analoghi a quelli che sono, almeno in parte, responsabili dell’emergere della coscienza. E dunque – in accordo, potrebbe notare uno storico della filosofia, con l’“estetica trascendentale” di Kant – le neuroscienze confermano che l’intuizione (secondo Kant “pura”) del tempo è la forma del senso interno, ossia è il formato comune a tutti gli stati di coscienza. Più precisamente, la percezione di essere svegli coincide con la sensazione interna del passare del tempo.

Ora, se questa conclusione che deriva dalla neurobiologia viene combinata con la teoria della relatività, si può finalmente spiegare perché quella teoria appaia inevitabilmente contro-intuitiva per il senso comune. Normalmente si dice che il rapporto tra la velocità dei corpi il cui movimento sia direttamente osservabile e la velocità della luce è troppo piccolo per produrre dilatazioni temporali e contrazioni dello spazio che possano essere rilevate a occhio nudo. Ma questa risposta diventa fuorviante se non viene specificato l'ordine di grandezza rispetto al quale il rapporto tra la velocità dei corpi direttamente osservabili e la velocità della luce è da considerarsi piccolo. La piccolezza è un concetto comparativo. Niente è piccolo in sé. Ciò che è piccolo lo è sempre in confronto a qualcosa che è più grande. Se consideriamo il problema da questo punto di vista, è chiaro che la scala secondo la quale il rapporto tra la velocità di un corpo macroscopico in movimento e la velocità della luce è troppo piccolo per essere percepito ad occhio nudo è una scala umana. Più precisamente, è la scala alla quale lavora il cervello umano per effettuare il coordinamento senso-motorio (gli orologi atomici ovviamente lavorano ad una scala molto diversa!). Pertanto la ragione per la quale il tempo assoluto della meccanica classica sembra intuitivamente accettabile, mentre il tempo dilatabile della teoria della relatività è contro-intuitivo, deve essere cercata non nella struttura della realtà esterna oggettiva, bensì nel modo in cui il cervello umano la percepisce³⁰.

Infatti secondo le teorie di Pöppel tutti gli stimoli visivi che cadano in un intervallo di trenta millisecondi sono percepiti come simultanei. Di conseguenza, per il cervello umano (e quindi per tutti gli esseri umani nella loro vita quotidiana) la velocità della luce è praticamente infinita in un raggio di novemila chilometri (la distanza coperta dalla luce in trenta millisecondi). Ma se la velocità della luce viene considerata come tendente all'infinito, allora in virtù di un calcolo matematico elementare le equazioni della teoria della relatività tendono a coincidere con le equazioni della meccanica classica. In altre parole, è il potere limitato della risoluzione temporale del cervello umano che fa sì che il tempo assoluto della meccanica classica sia intuitivamente accettabile mentre il tempo dilatabile della

³⁰ Di parere opposto Benini (2017), che identifica (a mio avviso confonde) il tempo reale con il tempo fenomenico “costruito” dal cervello.

teoria della relatività ci appare contro-intuitivo.

Più in generale le teorie psico-neurologiche della percezione suggeriscono che il tempo fenomenico non è identico al tempo reale. Inoltre, se la percezione del tempo coincide con l'emergere della coscienza, allora è plausibile ipotizzare che la relazione tra il tempo fenomenico e il tempo reale sia analoga alla relazione tra la coscienza fenomenica e i suoi correlati neurali. Di conseguenza, se si riesce a scoprire nel funzionamento del cervello umano una spiegazione della differenza tra il tempo fenomenico e il tempo reale, perché mai non si dovrebbe poter spiegare con certe proprietà della dinamica cerebrale la ragione per cui, quando si è svegli, ci si sente liberi e padroni di se stessi, mentre si è in realtà solo un servo-meccanismo del proprio cervello?

Riassumendo, lo *Hard Problem* non è né un vero problema scientifico in attesa di essere risolto empiricamente né un mistero irrisolvibile. Sembra essere un mistero in via di principio insolubile solo perché è in realtà uno pseudo-problema filosofico che deve essere non già risolto, bensì dissolto attraverso un'analisi concettuale. Tuttavia l'analisi concettuale in grado di dissolverlo non può essere ridotta ad una semplice analisi del linguaggio comune sulla scia di L. Wittgenstein. Tale analisi deve alimentarsi, per così dire, con le nuove idee provenienti dalle scienze cognitive in generale e dalle neuroscienze in particolare. Le teorie di Einstein costrinsero i fisici a correggere e in parte abbandonare alcuni concetti fondamentali della meccanica classica. Perché una nuova scienza della mente/cervello non dovrebbe cambiare profondamente la *folk psychology* ed il senso comune?

In altre parole, lo sforzo che è stato richiesto da più di un secolo a questa parte per comprendere ed accettare la teoria della relatività, nonostante il suo carattere fortemente contro-intuitivo, è paragonabile allo sforzo richiesto oggi per superare la riluttanza di molte persone a comprendere ed accettare il fatto che noi esseri umani siamo solo un «pacchetto di neuroni» (Crick 1994, p. 17). Ma anche il vantaggio che otteniamo in entrambi i casi grazie a questi sforzi è altrettanto grande. Le teorie neuro-psicologiche di Pöppel, se combinate con la teoria della relatività, sono infatti in grado di spiegare perché quest'ultima, sebbene sia preferibile alla meccanica classica da un punto di vista scientifico, appare però al senso comune contro-

intuitiva e meno accettabile di essa. Allo stesso modo una nuova scienza della mente/cervello permette già oggi (e permetterà sempre più in futuro) di iniziare a chiarire perché l'immagine che noi esseri umani abbiamo istintivamente di noi stessi è a favore del dualismo cartesiano, sebbene tale concezione sia fondamentalmente sbagliata da un punto di vista scientifico.

Capitolo quinto

Realismo scientifico e ontologia del mentale

1. Il realismo scientifico ed il problema mente-corpo

Secondo il “realismo scientifico” (Chakravartty 2011) esiste una realtà indipendente dal nostro conoscerla e dal modo nel quale la conosciamo, ma sono solo le scienze empiriche – e non la speculazione metafisica – che ci consentono mediante i loro impegni ontologici di fare ipotesi attendibili, sebbene sempre rivedibili, su quale sia la struttura di fondo di tale realtà. Quando questo assunto di base del realismo scientifico venga applicato agli impegni ontologici delle scienze cognitive, ne risulta quel mutamento radicale nell’approccio filosofico allo studio della mente prima indicato come “seconda svolta cognitiva” (cfr. *supra* capitolo primo); un approccio secondo il quale l’uomo è un animale tra gli altri, è un prodotto dell’evoluzione biologica e pertanto anche il suo essere dotato di una mente cosciente e di capacità cognitive superiori a quelle degli altri animali è una caratteristica puramente biologica legata al funzionamento del suo cervello (W. Singer 2003). La svolta cognitiva ed il naturalismo cognitivo sono perciò una delle manifestazioni di quella continuità tra scienza e filosofia, inaugurata da Quine, che trova la sua massima espressione nel realismo scientifico.

Chi si interroghi sul problema mente-corpo entro l’orizzonte teorico delineato dal realismo scientifico giunge perciò inevitabilmente a conclusioni che verranno sdegnosamente rifiutate da tutti i critici del naturalismo cognitivo. Sia sofisticata quanto si vuole la riduzione neuro-funzionale della coscienza fenomenica che il naturalista sia riuscito ad escogitare – diranno in coro, almeno in questo d’accordo tra loro, tutti gli avversari del naturalismo cognitivo – egli commette comunque un errore categoriale imperdonabile quando scambia la coscienza stessa con l’attività cerebrale da cui emerge! Di fronte ad un’obiezione pregiudiziale di questo genere tutto ciò che il naturalista può fare è replicare ai suoi avversari che, quando gli obiettano di commettere un errore categoriale allorché identifica la coscienza fenomenica con una proprietà della dinamica cerebrale, sono loro che non sono capaci di rivedere le loro categorie filosofiche alla luce delle novità apportate dalle scienze naturali e cognitive!

Ma le scienze – controreplicheranno gli avversari del realismo

scientifico e del naturalismo cognitivo – non parlano mai con una voce sola e si prestano, comunque, ad interpretazioni filosofiche divergenti. A tutto ciò il difensore del naturalismo cognitivo non ha null'altro da aggiungere se non che anche per lui il realismo scientifico, per come egli lo intende, non è un sistema dogmatico che pretenda di offrire, come la vecchia metafisica, verità eterne; è solo un principio regolativo di quell'inesauribile *work in progress* che è la ricerca scientifica.

È vero però che questo principio regolativo, quando venga trasferito dalla fisica, sua scienza di elezione, alle scienze cognitive ha bisogno di qualche significativo aggiustamento, soprattutto riguardo ad alcune categorie centrali di un'ontologia interdisciplinare che voglia includere anche gli stati mentali in quanto impegni ontologici delle scienze cognitive. Si approderà in tal modo ad un "realismo scientifico-cognitivo" che implichi un'ontologia del mentale in accordo con il naturalismo (cfr. *supra* capitolo terzo e Nannini 2007a, pp. 159-174).

Per giungere ad abbozzare un'ontologia del mentale di tal genere conviene partire da un riesame della celebre affermazione quiniana che è a fondamento del realismo scientifico: "Essere è essere il valore di una variabile"; affermazione questa che – contrariamente all'interpretazione che talvolta ne viene data – non deve essere intesa come un invito ad accettare acriticamente tutti gli impegni ontologici di qualsiasi teoria scientifica. Come Quine stesso si premurò di chiarire, quella affermazione infatti non serve a scoprire come realmente è fatto il mondo, bensì soltanto a rendere esplicito come una data teoria dice che è fatto (Quine 1953, p. 16). Pertanto, data la pluralità degli impegni ontologici delle varie scienze (in particolare delle varie scienze cognitive), quella formula rischia di condurci in braccio al relativismo cognitivo a meno che, in accordo con l'ideale dell'unità della scienza e con il presupposto dell'unicità del mondo reale, non introduciamo una gerarchia e delle relazioni di "sopravvenienza" (Kim 1993) tra gli impegni ontologici della fisica e quelli delle scienze cognitive e non ammettiamo come ovvio ed indubitabile che, mentre il mondo fisico continuerebbe ad esistere anche se nell'universo non vi fosse più alcuna traccia di esseri dotati di coscienza, viceversa la coscienza stessa non potrebbe esistere se non esistessero dei cervelli. Detto in breve i processi cerebrali, appartenendo al mondo fisico, esistono effettivamente, mentre gli stati mentali,

poiché esistono unicamente se vengono implementati da processi cerebrali, sono reali solo in senso virtuale o fittizio, ossia non sono che delle ridescrizioni dei processi cerebrali in un altro “formato”, sia esso il linguaggio scientifico della psicologia cognitiva che descrive tali processi in termini funzionali (ad es. mediante diagrammi di flusso) oppure sia esso il “codice-macchina” interno “Intenzional-coscienziale” che il cervello impone alla sua stessa attività quando la monitorizza per poterla più finemente controllare e rendere coerente (cfr. *supra* capitolo secondo).

Per chiarire i rapporti gerarchici che intercorrono tra gli impegni ontologici di scienze appartenenti a differenti ordini di discorso, conviene inoltre distinguere entro ciascun ordine i tre livelli d'analisi già precedentemente indicati: gli *illata*, gli *abstracta* ed i *ficta* (cfr. *supra* capitolo terzo). Rilevante è qui soprattutto chiarire ulteriormente la differenza tra ciò che è virtuale e ciò che è fittizio (ossia è un *fictum*). Ciò che è virtuale è implementato completamente in ultima analisi da *illata* fisico-chimici, mentre i *ficta* sono solo delle finzioni, l'illusione della cui esistenza è creata, in termini fisico-chimici, da un qualche “trucco scenico”. Chiariamo questo punto mediante un paragone terra terra. In *Guerre Stellari* tutto ciò che sullo schermo fa Han Solo (il personaggio) gode di una realtà virtuale (e di un “potere causale virtuale”) grazie a quello che ha fatto sul *set* Harrison Ford (l'attore). Se Han Solo alza un braccio, è Harrison Ford che lo ha alzato. Ma, ogni volta che l'astronave di Han Solo, il celebre *Millenium Falcon*, compie un balzo nell'iperspazio a velocità superiore a quella della luce, quel balzo non è qualcosa di virtuale, perché non è stato e non poteva essere implementato sul *set* mediante alcun processo fisico. Il *Millenium Falcon* e i suoi balzi nell'iperspazio non sono qualcosa di virtuale; sono piuttosto dei *ficta* generati da qualche trucco scenico.

In modo analogo, quando vedo una macchia rossa, il colore rosso (il *quale*) è il contenuto del mio vedere-rosso, ossia è il contenuto di una rappresentazione mentale cosciente implementata da una certa dinamica cerebrale. La dinamica cerebrale è ciò che realmente esiste, il vedere-rosso è una realtà virtuale, ossia è una rappresentazione di tale dinamica che il cervello fornisce a se stesso auto-monitorandosi nel formato dell'Intenzionalità e della coscienza, mentre il colore rosso, in quanto contenuto del vedere-rosso, è un *fictum*, ossia è un

qualcosa d'inesistente a livello fisico, la cui presenza viene simbolicamente evocata da un trucco escogitato dal cervello, ossia dall'aver imposto un formato Intenzional-coscientiale alla dinamica cerebrale che implementa il "vedere-rosso": un formato che, si noti bene, a differenza del contenuto fittizio da esso formattato, è una proprietà reale del cervello coincidente con la sincronizzazione e forse con altre caratteristiche della dinamica cerebrale (di nuovo cfr. *supra* capitolo secondo).

2. Obiezioni e risposte sul realismo scientifico-cognitivo

In conclusione il realismo scientifico-cognitivo, così ricostruito, si presenta come la cornice teorica più plausibile per le scienze cognitive, quando si prefiggano l'obiettivo di naturalizzare il mentale. Ciò non esime tuttavia chi lo sostenga dal cercare di rispondere ad una duplice obiezione che è stata avanzata da più parti contro il realismo scientifico in generale ed il connesso tentativo di naturalizzare l'epistemologia. In primo luogo – sembra ovvio notare – l'epistemologia non può limitarsi a descrivere i principi metodologici ed i presupposti teorici fatti propri dagli scienziati. Deve anche valutarne la validità formale. "E questo – pensano molti filosofi – è nostro compito esclusivo!". In secondo luogo – è stato rimproverato ai sostenitori del realismo scientifico e del naturalismo soprattutto da parte degli empiristi (cfr. in particolare Parrini 1995, 2011 e 2012) – chi è contrario alla metafisica, come lo sono i naturalisti, dovrebbe per coerenza essere contrario anche al realismo metafisico, ossia dovrebbe essere contrario al supporre l'esistenza di una realtà concepibile come qualcosa di completamente indipendente dal modo nel quale noi esseri umani possiamo conoscerla.

Queste obiezioni sono molto serie e non è facile per i realisti scientifici in generale – e quindi neppure per un sostenitore del realismo scientifico-cognitivo – trovare una risposta soddisfacente. Bisogna tentare tuttavia. Anzitutto è sì vero che l'epistemologia presenta una ineliminabile dimensione normativa, ma è anche vero (almeno per chi segua in questo Quine e creda nel crollo della distinzione tra giudizi analitici e giudizi sintetici) che, non essendo possibile separare la valutazione della validità logico-formale e della chiarezza concettuale di una teoria scientifica dalla valutazione del suo grado di conferma empirica, quella dimensione normativa, lungi dall'essere

riservata ad una mera analisi filosofica, è parte integrante di un lavoro comune a cavallo tra filosofia e scienza. È un grave errore ad esempio – un errore da cui deriva la presunta insolubilità del problema mente-corpo – prima definire a tavolino, basandosi su concetti offerti dalla tradizione filosofica di origine cartesiana, che cosa si debba intendere per coscienza e poi, dopo una (in genere frettolosa) ricerca condotta *top-down*, concludere che nessun processo neurologico conosciuto riesce a implementarla. Dobbiamo al contrario, procedendo *bottom-up*, costruire un nuovo concetto scientifico di coscienza partendo da un esame della dinamica cerebrale.

Veniamo ora alla critica più insidiosa per chi, essendo naturalista, non vuol certo ricadere nella vecchia metafisica e nella filosofia speculativa! Un punto alto di confronto è qui offerto, di nuovo, da Paolo Parrini³¹, il quale, difendendo «una “terza via” tra il realismo metafisico ed il relativismo radicale», approda ad un «realismo empirico» che, pur mantenendo la distinzione tra soggetto e oggetto, dà di questa distinzione stessa «una interpretazione puramente empirica, e non metafisica» (Parrini 2011, pp. 20-21). In altre parole Parrini riconosce da un lato la ineliminabilità di un concetto di verità che, in quanto idea regolativa, distingue il significato di “vero” da quello di “giustificato” e quindi garantisca sempre (a differenza del relativismo – aggiungo io – che identifica “vero” con “internamente coerente”) la rivedibilità di qualsiasi teoria scientifica e la sua apertura verso ipotesi nuove alla luce di quanto suggerito da un termine di confronto ad essa esterno, ma dall’altro nega che questo termine di confronto debba essere necessariamente una realtà completamente indipendente dal soggetto conoscente, una kantiana cosa in sé. Per svolgere tale compito basta – egli dice – «l’esperienza assunta nella sua ideale inesauribilità e interpretata e organizzata grazie alle nostre sempre modificabili e rinnovabili procedure razionali» (Parrini 2011, p. 22).

Ora, Parrini ha certamente ragione quando afferma che «non siamo mai in grado di raffrontare le nostre esperienze delle cose con le cose in se stesse o di per sé considerate» (Parrini 2011, p. 23). Ma ciò significa forse che, volendo rendere conto del processo mediante

³¹ Aggiungo che lo stimolo ricevuto dalle acute riflessioni di Paolo (sia leggendo i suoi scritti sia avendo la fortuna di poterci talvolta parlare amichevolmente) è stato per me prezioso anche laddove giungo a conclusioni diverse dalle sue.

il quale si sviluppa la ricerca scientifica, «il realismo metafisico non riesce a dare alcun contenuto di carattere operativo al suo richiamo a una realtà quale essa è in se stessa» (*ibidem*)? Questa conclusione di Parrini mette in ombra il fatto che i dati empirici sono sempre *theory-laden* e che quindi l'esperienza (anche se «assunta nella sua ideale inesauribilità») – quando venga presa come ultimo e insuperabile termine di confronto per giudicare quale tra due teorie reciprocamente incompatibili che vengano confermate dai medesimi dati empirici sia più vicina alla verità – non è sufficiente, sul piano puramente pragmatico, per indurre lo scienziato a rifiutare il relativismo e ad andare in cerca di nuovi dati empirici. In fondo nella situazione summenzionata della scelta tra due teorie reciprocamente incompatibili ma entrambe empiricamente ben confermate – si pensi alla teoria della relatività e alla meccanica quantistica – accettare il relativismo cognitivo sarebbe la soluzione più semplice alla luce dell'*evidence* disponibile e quindi sarebbe la soluzione da privilegiare, se ci basassimo unicamente su criteri pragmatici di scelta come la semplicità o l'economicità delle teorie e non decidessimo che la ricerca scientifica deve sempre prendere le mosse dal presupposto – non supportato da alcun dato empirico, eppure motore indispensabile dell'intera impresa scientifica – che c'è là fuori qualcosa la cui natura è del tutto indipendente dall'essere oggetto dei nostri tentativi di conoscerla e che quel qualcosa è l'unica ed unitaria sorgente di tutta quella quantità d'informazione che le nostre teorie trasformano in esperienza *theory-laden* sussumendola sotto questo o quel formato.

Ciò inoltre è pienamente in accordo con il realismo scientifico-cognitivo promosso dalla seconda svolta cognitiva. Alla luce di tale svolta infatti i dati percettivi e tutti i processi cognitivi che essi innescano sono il prodotto di una interazione fra il nostro cervello e la realtà fisica esterna. Non vediamo il mondo com'è, bensì come il nostro cervello ce lo presenta in funzione del perseguimento dei nostri scopi nell'interazione con il mondo stesso (ad es. Roth 2001). Solo che il cervello, da bravo primo realista ingenuo, nasconde alla nostra coscienza questo suo lavoro e ci presenta il prodotto finito nella forma di un accesso diretto alla realtà in sé (Metzinger 2009). Nel fare ciò il nostro cervello ci inganna e quindi in sede scientifica (non nella vita di tutti i giorni!) dobbiamo pazientemente smontare la sua costruzione, coincidente con il senso comune, e sostituirla con

un'immagine scientifica del mondo. Un'immagine che sicuramente non è univoca e priva di controversie e che anzi ha già conosciuto radicali mutamenti di paradigma nel corso della storia umana e probabilmente ne conoscerà ancora (cfr. *supra* capitolo quarto). Ma tutti questi cambiamenti sarebbero senza senso se non presupponessimo che grazie ad essi cerchiamo di cogliere sempre meglio la struttura di fondo di una realtà a noi esterna; una realtà che era lì molto prima della comparsa dell'*Homo Sapiens* sulla faccia della Terra e che quasi certamente sarà ancora lì pressoché immutata molto tempo dopo la sua estinzione.

Capitolo sesto

Il concetto di verità in una prospettiva naturalistica

1. La naturalizzazione dell'Intenzionalità e il concetto di verità.

La naturalizzazione della mente è possibile solo se è possibile naturalizzare la coscienza, il *Self* e l'Intenzionalità brentiana (*supra* capitolo secondo). Riguardo a quest'ultima si può pensare che la relazione semantica tra una rappresentazione mentale e l'oggetto a cui si riferisce possa essere fisicamente implementata dalla relazione causale tra l'oggetto stesso e la sua percezione: ad esempio, se c'è un gatto davanti a me, vedo un gatto proprio perché è il gatto stesso a causare con la sua presenza davanti a me quei processi cerebrali che implementano la mia percezione. Secondo questa teoria il contenuto (o oggetto interno che dir si voglia) della mia percezione è determinato dall'oggetto reale che è la causa principale della percezione stessa (Dretske 1981 e 1995).

Ma la difficoltà maggiore che presenta questa teoria è chiarire come dei processi cerebrali che intrattengano con gli eventi dell'ambiente circostante relazioni unicamente di causa ed effetto (o comunque relazioni empiriche – e quindi contingenti – spiegabili mediante delle leggi di natura), possano implementare degli stati Intenzionali, che viceversa sono connessi in modo intrinseco, ossia concettualmente necessario, all'oggetto a cui si riferiscono. Sembra essere inspiegabile infatti come una relazione d'Intenzionalità, che può intercorrere talvolta anche nei confronti di un oggetto inesistente (il mio pensare a Zeus ad esempio), possa essere implementata da una relazione causale tra dei processi cerebrali ed altri eventi fisici che accadono nel mondo esterno. Una relazione di causa ed effetto non può mai intercorrere fra degli eventi non accaduti o cose non esistenti!

Una difficoltà di questo tipo si presenta inoltre non solo riguardo alle idee di oggetti inesistenti, ma anche riguardo alle percezioni ingannevoli: ad esempio ho davanti a me uno scoiattolo, ma a causa dell'oscurità vedo un gatto; oppure vedo la post-immagine di un lampo di luce rossa, ma in effetti davanti a me non c'è alcuna luce rossa. Come è mai possibile che il contenuto delle mie percezioni sia determinato causalmente dall'oggetto percepito anche quando tale oggetto è o inesistente o in quel momento assente e quindi non può essere causa di alcunché? Come Zeus non può essere la causa del mio pensare a lui,

così, se la causa della mia percezione è uno scoiattolo e si ritiene che sia tale causa a determinare il contenuto della mia percezione, allora come è possibile che la presenza di uno scoiattolo determini l'insorgere in me della percezione di un gatto?

Il modo più convincente di risolvere problemi di tal genere consiste per un filosofo naturalista, come già è stato detto, nel ridurre funzionalmente gli stati mentali a tappe intermedie del coordinamento senso-motorio. Secondo questo approccio naturalistico infatti il mio vedere un gatto è quel modo “gattesco” di vedere che – essendo implementato da certe proprietà della dinamica del mio cervello causate di norma, attraverso il senso della vista, dalla presenza di un gatto – causa a sua volta un comportamento che tende a realizzare quelle mie intenzioni che sono evocate dalla supposta presenza di quel gatto. Questa riduzione causale dell'Intenzionalità è compatibile con le percezioni ingannevoli. Non c'è nulla di strano infatti che in circostanze particolari (ad esempio se mi trovo quasi al buio) la presenza di uno scoiattolo, in virtù di una catena causale deviante, possa far sì che compaia nel mio cervello una percezione “gattesca” e che quindi io abbia l'impressione di vedere un gatto.

Ciò sembra andare incontro però ad una seria obiezione. Se questa teoria fosse vera – potrebbe notare un suo critico – allora la verità di una percezione verrebbe a coincidere con la sua utilità nel coordinamento senso-motorio. Poniamo il caso che io voglia in certe circostanze compiere una certa azione. Ad esempio voglio prendere in collo il gatto che è davanti a me (azione *A*). Ovviamente per compiere tale azione devo prima vedere che c'è un gatto davanti a me (percezione *P*). Pertanto, se le percezioni vengono interpretate come tappe intermedie del coordinamento senso-motorio, allora la mia percezione potrà essere definita come quel processo cerebrale che è causato dalla presenza del gatto ed è causa a sua volta, insieme a molte altre concause qui messe in parentesi, della mia azione di prenderlo in collo: “*gatto* $\rightarrow P \rightarrow A$ ”). E quindi, dato che, se non vedessi che davanti a me c'è un gatto, non potrei certo prenderlo in collo, se ne deve concludere che la verità della mia percezione coincide con la sua utilità: senza di essa non potrei compiere l'azione che desidero compiere.

Ma, in molti altri casi, delle percezioni vere sono inutili o addirittura dannose, mentre una percezione ingannevole – o talvolta anche l'assenza di una percezione che, se l'avessi, sarebbe vera – può risultare

casualmente molto utile. Se ad esempio, mentre sto facendo una passeggiata lungo una strada di campagna, vedo in lontananza dei cani randagi e, preso dalla paura, comincio a urlare e a correre, questo è il modo migliore per vedermeli venire addosso! Se invece, pur essendo una persona nervosa ed inesperta che si comporterebbe nel modo sbagliato se vedesse quei cani, rimango calmo e tranquillo semplicemente perché o non li ho visti o li ho scambiati per animali meno pericolosi, può darsi che anche i cani se ne stiano calmi e tranquilli. Non conoscere la verità potrebbe essermi molto utile in questo caso!

Come può rispondere il naturalista ad un'obiezione di tal genere? Intuitivamente appare subito evidente che egli dovrà identificare la verità di una percezione non con la sua presunta utilità nel caso particolare in cui la si ha, bensì con l'utilità che percezioni simili hanno nella maggior parte dei casi. Sebbene la verità di una singola percezione non necessariamente coincida con la sua utilità, non solo è sensato presumere che l'evitare percezioni ingannevoli sia in genere vantaggioso, ma è anche plausibile definire la verità in modo tale che risultino essere vere, in base alla definizione data, quelle percezioni che sono prodotte da meccanismi cerebrali capaci di causare azioni che in genere, dato un certo *input* sensoriale, sono risultate vantaggiose per gli esseri umani ed i loro antenati animali nel corso dell'evoluzione biologica.

Ma in quale relazione sta questa definizione pragmatistica della verità con quella che intuitivamente accettiamo quando diciamo che è vero che la neve è bianca perché essa è effettivamente bianca? Per applicare il concetto di verità alle percezioni e più in generale alle rappresentazioni mentali occorre perciò analizzarlo e ridefinirlo in modo più accurato. E questa sarà anche l'occasione per una riconsiderazione più generale del concetto di verità in una prospettiva di naturalizzazione del mentale³².

2. Rappresentazioni mentali

La prima difficoltà che si pone nel parlare della verità o falsità delle percezioni è che il concetto di verità, da Aristotele in poi, è sempre stato definito in riferimento al linguaggio: solo le proposizioni possono essere, a rigore, vere o false. Per poter dire che una perce-

³² La bibliografia sul concetto di verità è sterminata. Per una prima introduzione cfr. Glanzberg (2016).

zione ingannevole è falsa occorre estendere il concetto di verità dalle proposizioni a tutte le rappresentazioni mentali. Questo può essere fatto mediante una semplice definizione stipulativa dopo aver chiarito che cosa si debba intendere per rappresentazione mentale, in generale, e per percezione in particolare.

Per rappresentazione mentale $RM(O_n)$ intendo qui un qualsiasi stato mentale che abbia una relazione d'Intenzionalità con un qualche "oggetto" O (non importa se esistente o meno, se fisico o mentale, se sostanza, proprietà, stato di cose o evento) mediante il proprio contenuto (o oggetto interno O_n , dove n è un parametro che individua un certo ordine di discorso L_n : ad esempio il riferirsi a contenuti percettivi o ai contenuti di concetti del senso comune o di concetti scientifici ecc.)³³. Per percezione intendo, più in particolare, una rappresentazione mentale il cui contenuto percettivo (o dato sensoriale che dir si voglia) O_p sia fornito dai sensi. Stabilite queste convenzioni, il concetto di verità può essere facilmente esteso dalle proposizioni alle percezioni nel modo seguente: la percezione $RM(O_p)$ di O è vera se e solo se il giudizio percettivo "Questo è un O_p ", pronunciato mentre si punta il dito verso O , è vero. Una volta chiarito in qual modo le percezioni possono essere dette vere o false, non è difficile estendere questa definizione della verità a tutte le rappresentazioni mentali (pensieri ecc.) concependole, sulla scia degli empiristi, come costruzioni a partire da dati sensoriali.

Si può dire inoltre che la percezione $RM(O_p)$ di O è vera se e solo se essa è in relazione d'Intenzionalità" ($\blacktriangleright$) con O e tale relazione è soddisfatta (ossia, convenzionalmente, il valore S di $RM(O_p) \blacktriangleright O$ è uguale a 1):

$$\begin{aligned} &RM(O_p) \blacktriangleright O \\ &S(RM(O_p) \blacktriangleright O) = 1 \end{aligned}$$

Perciò, se conveniamo che dire che una percezione è vera significa asserire che il suo valore di verità V è 1, allora, generalizzando,

³³ Sulle rappresentazioni mentali in generale cfr. ad esempio Pitt (2017) e Stich e Warfield (1994), mentre per una loro definizione da un punto di vista naturalistico, come quella qui abbozzata, cfr. Nannini (2006).

possiamo stabilire la seguente equivalenza tra verità e soddisfacimento di una rappresentazione mentale:

$$VS\text{-}equivalenza: V(RM(O_n))=1.\equiv S(RM(O_n) \blacktriangleright O)=1$$

Limitandosi alle percezioni, si può dire ad esempio che una percezione che rappresenti un certo O come un O_p è vera se e solo se la relazione d'Intenzionalità che essa intrattiene con O tramite O_p è soddisfatta.

Questa conclusione può essere letta, tuttavia, in tre modi molto diversi a seconda del significato che viene attribuito alla relazione d'Intenzionalità “ $\blacktriangleright$ ” e della relativa relazione S di soddisfacimento:

1. La relazione di Intenzionalità che la rappresentazione mentale $RM(O_n)$ intrattiene con O è soddisfatta se e solo se il suo contenuto O_n è logicamente coerente con il contenuto di tutte le altre rappresentazioni mentali appartenenti all'ordine di discorso L_n .
2. La relazione d'Intenzionalità che la rappresentazione mentale $RM(O_n)$ intrattiene con O è soddisfatta se e solo se il suo contenuto O_n assomiglia in modo sufficiente, sotto qualche riguardo e secondo un accettabile *standard* di sufficienza, a O .
3. La relazione d'Intenzionalità che la rappresentazione mentale $RM(O_n)$ intrattiene con O è soddisfatta se e solo se il suo rappresentare O come un O_n fa parte di una catena causale che contribuisce in genere al successo di qualsiasi azione A che l'agente compia su O nell'intento (ossia con il desiderio $D(G_n)$) di ottenere un qualche scopo G :

$$S(RM(O_n) \blacktriangleright O)=1.\equiv O \rightarrow RM(O_n) \& [D(G_n) \& \text{altri stati mentali}] \rightarrow A \rightarrow G.$$

Ad esempio se, visto il gatto O , voglio prenderlo in collo (azione A) in modo tale che esso poi sia tra le mie braccia (stato di cose che costituisce il mio scopo G), allora il mio rappresentarmi O come un

gatto (ossia il fatto che ho una percezione RM il cui contenuto è O_n) fa parte di una catena di cause ed effetti che inizia con la presenza di O di fronte a me, prosegue con l'insorgere nella mia mente della percezione $RM(O_n)$, ossia con il mio vedere il gatto, e infine, dopo che questa percezione si è combinata con il mio desiderio $D(G_n)$ di prenderlo in collo (più altri stati mentali), si conclude con un certo movimento delle mie mani (azione A) che, a sua volta, produce in genere il realizzarsi di G nel mondo esterno (ossia il gatto è finalmente tra le mie braccia). Ora, se l'oggetto O che dà inizio a questa catena di cause ed effetti è davvero un gatto, allora l'averlo rappresentato come un gatto ($RM(O_n)$) avrà contribuito al raggiungimento del mio scopo. Se invece l'oggetto O' che causa la comparsa di $RM(O_n)$ nella mia mente è uno scoiattolo, allora l'averlo scambiato per un gatto (ossia l'averlo rappresentato come un O_n), ad esempio a causa di una scarsa illuminazione, sarà stato causa del mio fallimento: volevo prendere in collo un gatto e mi ritrovo tra le braccia uno scoiattolo (ammesso che mi riesca di prenderlo!). In entrambi i casi $RM(O_n)$ intrattiene una relazione d'Intenzionalità o con O o con O' , ma tale relazione è soddisfatta solo nel primo caso (vedo un gatto dove c'è effettivamente un gatto) e non nel secondo (vedo un gatto dove c'è in realtà uno scoiattolo). Ciò che è cruciale notare è che, secondo questo terzo modo d'intendere l'Intenzionalità, l'essere tale relazione d'Intenzionalità soddisfatta o meno non dipende minimamente da una eventuale somiglianza tra O e O_n , bensì unicamente dal ruolo causale (e quindi, più in generale, dalla funzione pratica) che $RM(O_n)$ svolge nella catena causale che va da O a G . Se voglio prendere in collo un animale che mi faccia le fusa e mi ritrovo tra le mani un animale che non è capace di farle, la mia rappresentazione di quell'animale è ingannevole; se invece me lo rappresento in modo tale che in genere riesco ad ottenere quello che volevo (o almeno aumento sensibilmente la probabilità di poterlo ottenere), allora essa è corretta.

Se applichiamo a questi tre modi d'intendere la relazione d'Intenzionalità la relazione di equivalenza tra soddisfacimento e verità di una rappresentazione mentale ("*VS-equivalenza*"), vediamo chiaramente che a tali tre modi corrispondono le tre concezioni della verità più largamente accettate dalla tradizione filosofica: a (1) corri-

sponde quella concezione della verità che vede in essa una relazione di *coerenza* tra tutte le rappresentazioni appartenenti ad un medesimo ordine di discorso; a (2) corrisponde la concezione della verità come *corrispondenza* tra mente e mondo; a (3) corrisponde la *concezione pragmatistica* della verità: una percezione, un giudizio, una proposizione o una teoria sono veri se e solo se il meccanismo psicologico che li produce (o la cornice teorica che fa loro da sfondo) guida le nostre interazioni con il mondo in modo tale da massimizzare il raggiungimento dei nostri scopi (ad esempio consentendoci di formulare previsioni empiricamente confermabili).

3. *Concezioni della verità.*

La concezione pragmatistica della verità è quella più convincente per chi sposi un punto di vista naturalistico riguardo alle rappresentazioni mentali, perché le relazioni causali tra la mente e il mondo a cui tale concezione riduce la verità sono spiegabili mediante leggi e teorie delle scienze naturali. Tuttavia, per dissipare gli equivoci che possono renderla meno plausibile, è opportuno considerare se a partire da un punto di vista naturalistico non si possano ridefinire tutte le principali concezioni della verità, introducendo, se necessario, delle ulteriori distinzioni rispetto alle tre sopra individuate.

Cominciamo con il precisare in quale senso si possa parlare dell'oggetto di una rappresentazione mentale (o anche dell'oggetto di un concetto, di una teoria ecc.). Anzitutto chiamiamo O l'“oggetto Intenzionale” a cui la rappresentazione si riferisce. Nel caso di una percezione in senso stretto, O coincide con l'oggetto fisico (lo stato, l'evento, la proprietà ecc.) la cui presenza o il cui accadere produce lo stimolo sensoriale che a sua volta causa la comparsa di una percezione che rappresenta O stesso come, ad esempio, un O_n ($O \rightarrow RM(O_n)$). In tal caso O è qualcosa che appartiene necessariamente alla realtà esterna. E ciò è vero, si noti bene, anche nel caso delle percezioni ingannevoli. Uno scoiattolo è qualcosa di reale anche quando, nell'oscurità, provoca una “percezione-di-gatto”. Se invece la rappresentazione mentale in questione è un pensiero, un giudizio, una fantasia, un'allucinazione o una post-immagine, allora l'oggetto O “Intenzionato” dal contenuto O_n della rappresentazione può essere inesistente. Anche in tal caso, ad ogni modo, esso deve

essere concepito come qualcosa di distinto dal contenuto che lo “Intenziona” (o, se si vuole, come qualcosa di distinto da quel “senso” della rappresentazione che lo determina come il “riferimento” della rappresentazione stessa)³⁴. Tuttavia si possono dare al riguardo due casi:

a) l’oggetto Intenzionale non esiste, ma l’osservatore crede che esista (“Aiuto, i fantasmi!”, detto con sincerità);

b) colui che pensa a qualcosa d’inesistente, ad esempio a Zeus, sa che l’oggetto del suo pensiero non esiste, ma prendendolo, come si suol dire, “ad oggetto” delle sue riflessioni, lo “oggettivizza”, ossia lo tratta come se esistesse sul serio e non lo identifica affatto con il contenuto del pensiero mediante il quale lo individua e si riferisce ad esso. Ciò è particolarmente evidente nel caso dei numeri, le cui proprietà sono spesso studiate dai matematici come se essi esistessero davvero e godessero di certe proprietà in modo indipendente dal nostro pensarli e conoscerli; e ciò resta vero anche per quei matematici che, non essendo dei platonici, ritengano che i numeri siano dei meri costrutti per convenzione.

Occorre distinguere perciò, in tutti i casi senza eccezione, l’oggetto Intenzionale O di una rappresentazione mentale dal suo contenuto O_n (od oggetto interno che dir si voglia). Il contenuto di una rappresentazione mentale non deve mai essere confuso con l’oggetto Intenzionale a cui si riferisce. Ad esempio il numero 19 rimane lo stesso numero e mantiene invariate tutte le sue proprietà – anche per un convenzionalista che lo consideri un *fictum* a rigore inesistente, del quale non si potrebbe dire nulla se non fosse stata inventata l’aritmetica – sia che venga rappresentato mediante il numerale arabo “19” come un $\langle 1 \times 10 + 9 \times 1 \rangle$ sia che venga rappresentato mediante il numerale romano “XIX” come un $\langle 10 + 10 - 1 \rangle$.

³⁴ Applico qui alle rappresentazioni mentali in generale la celebre distinzione di G. Frege tra *Sinn* e *Bedeutung*. Mi muovo perciò in una prospettiva naturalistica sì, ma al tempo stesso internalistica, sganciando il naturalismo dall’externalismo e dal comportamentismo a cui è spesso invece connesso. Sul naturalismo internalistico cfr. Nannini (2006, p. 95).

Questo esempio tratto dall'aritmetica mette in luce inoltre che non tutte le rappresentazioni appartengono al medesimo ordine di discorso. Ad esempio una cosa è il meccanismo psicologico mediante il quale io m'immagino il numero 19 e un'altra è come esso deve essere rappresentato nel sistema decimale. Può darsi che a me, se devo sommarlo ad altri numeri, venga spontaneo rappresentarmelo, poniamo il caso, come un $\langle 10+5+4 \rangle$, mentre un'altra persona (o io stesso in un'altra occasione) se lo rappresenta come un $\langle 20-1 \rangle$ ed un terzo lo vede in un terzo modo ancora. Viceversa, se si adottano i consueti numerali arabi, quel numero può essere rappresentato solo come un $\langle 1 \times 10 + 9 \times 1 \rangle$. Perché questa differenza? È molto semplice: le prime sono rappresentazioni mentali individuali (*RM*) del numero 19, mentre la seconda è una rappresentazione culturale (*RC*) del medesimo numero condivisa da tutti coloro che conoscono l'aritmetica moderna. Per sottolineare questa distinzione conveniamo d'indicare con O_n il contenuto della rappresentazione mentale $RM(O_n)$ e con O_n^* il contenuto della rappresentazione culturale $RC(O_n^*)$, rappresentazione che peraltro può esistere solo se viene implementata da rappresentazioni mentali individuali.

Stabilite queste convenzioni, possiamo ridefinire le tre concezioni della verità sopra menzionate ed operare al loro interno ulteriori distinzioni sino ad ottenere i seguenti modi fondamentali d'intendere il concetto di verità:

- (1) La verità come coerenza;
- (2) La verità come corrispondenza:
 - (2a) Corrispondenza tra mente e mondo mediante somiglianza e identità;
 - (2b) Corrispondenza tra linguaggio e metalinguaggio;
 - (2c) Corrispondenza tra mente e mondo mediante causalità e identità.
- (3) La concezione pragmatistica ingenua della verità;
- (4) La concezione olistico-quiniana della verità;
- (5) La concezione naturalistica della verità;

Esistono ovviamente altre concezioni della verità che qui non vengono prese in considerazione, come ad esempio la concezione deflazionistica (Horwich 1998). Le sette sopra elencate sono ad ogni

modo le concezioni fondamentali per riuscire a chiarire da un punto di vista naturalistico perché l'esistenza di percezioni vere ma inutili oppure false ma casualmente utili non smentisca l'adeguatezza di una concezione naturalistica dell'Intenzionalità. Esaminiamole più nel dettaglio:

(1) *La verità come coerenza.* — La rappresentazione mentale $RM_i(O_n)$ (con $i = 1, 2, \dots$), sia essa una percezione, una fantasia, una credenza ecc., è “coerente-vera” se e solo se è logicamente coerente con il sistema delle altre rappresentazioni mentali $RM_j(O_n)$ (con $i \neq j$) che appartengono al suo stesso ordine di discorso. Lo stesso dicasi per una rappresentazione culturale $RC_i(O_n^*)$. La verità, così intesa, è una relazione logica tra tutti gli O_n (o O_n^*) appartenenti allo stesso ordine di discorso. Essa non ha nulla a che fare perciò con l'oggetto Intenzionale O , quando tale oggetto appartenga alla realtà esterna, ossia al mondo reale. Diverso è il discorso quando O sia un oggetto fittizio (come lo sono, ad esempio, i numeri secondo i convenzionalisti). Questo modo d'intendere la verità può andar bene perciò tutt'al più per le verità della logica e della matematica (quando vengano riferite a degli oggetti costruiti mediante convenzioni), ma è sicuramente insufficiente in riferimento a qualsiasi forma di conoscenza empirica.

(2) *La verità come corrispondenza.* — Sia per il senso comune sia per il pensiero scientifico il concetto di verità empirica è infatti associato indissolubilmente, almeno per chi accetti il realismo scientifico (cfr. *supra* capitolo quinto), all'idea di una qualche forma di corrispondenza tra la mente e/o il linguaggio da un lato (percezioni, pensieri, proposizioni, rappresentazioni, teorie ecc.) ed il mondo reale dall'altro: le rappresentazioni mentali del tipo $RM(O_n)$ e quelle culturali del tipo $RC(O_n^*)$ sono “corrispondenti-veri” se e solo se i loro rispettivi oggetti O_n e O_n^* corrispondono all'oggetto reale O . Come già diceva Aristotele, è vero che la neve è bianca, perché essa è effettivamente bianca.

Il problema di tutte le teorie corrispondentistiche della verità, nonostante la loro indubbia plausibilità da un punto di vista intuitivo,

consiste tuttavia nella difficoltà che si incontra a precisare in quale senso debba essere intesa tale corrispondenza tra mente e mondo. L'ipotesi più comune al riguardo è quella che identifica questa corrispondenza con una relazione di somiglianza. Ma, come vedremo subito, è questa un'ipotesi non priva di difficoltà.

(2a) *Corrispondenza mediante somiglianza e identità.* — Per il senso comune una rappresentazione è vera se e solo se il suo oggetto interno è simile all'oggetto reale: la neve è effettivamente bianca; perciò, se la vedo bianca, la vedo com'è. E dunque la mia percezione è vera. Tuttavia basta un momento di riflessione per rendersi conto che questa teoria è ingenua e sbagliata. Come può qualcosa che si trova nella mia testa, dove hanno luogo solo processi cerebrali, avere qualcosa in comune con dei cristalli d'acqua ghiacciata? Inoltre una teoria che ipotizzi una somiglianza tra la neve "fenomenica" (coincidente con l'oggetto interno della mia percezione della neve) e la neve reale può, *prima facie*, sembrare plausibile tutt'al più riguardo ad una proprietà come la forma, ma è manifestamente falsa per il colore: nel mio cervello non c'è sicuramente nulla di bianco, quando vedo un mucchio di neve. Com'è allora che, quando vediamo per alcuni secondi ed in piena luce un gatto a pochi metri da noi, non abbiamo alcun dubbio che quello che vediamo sia effettivamente un gatto e non semplicemente un "simulacro-di-gatto" o "gatto fenomenico" che dir si voglia? Per il senso comune (e anche per tutti i filosofi che non si siano lasciati corrompere troppo da G. Berkeley), quando vediamo un gatto, il riferimento della nostra percezione $RM(O_n)$ è il gatto reale O , non la sua immagine interna O_n . Riassumendo, nella vita di tutti i giorni per un verso assimiliamo O a O_n ogni volta che identifichiamo O mediante le proprietà di O_n : ad esempio siamo sicuri che ciò che vediamo è un gatto perché assumiamo che l'oggetto O che è davanti a noi abbia quelle stesse proprietà "gattesche" che caratterizzano O_n (miagolare, fare le fusa ecc.)³⁵. Per altro verso però siamo consapevoli che il gatto fenomenico non può essere

³⁵ Anzi, a rigore, O_n è costruito dalla nostra mente giustappunto come l'insieme di tali proprietà. Su tale assunto fondamentale dell'empirismo – introdotto da Locke e sulla sua scia da Berkeley e da Hume – cfr. Russell (1912, pp. 7-18).

identico a quello reale. Il gatto fenomenico, ad esempio, è più piccolo o più grande a seconda della sua distanza da noi, ma sappiamo bene che le dimensioni del gatto reale rimangono sempre le stesse. Pertanto da un lato identifichiamo O con O_n , ma, dall'altro, contraddittoriamente lo distinguiamo da esso. Come uscire dall'*impasse*? Forse riflettere su che cosa intendiamo dire quando affermiamo che una certa percezione è ingannevole può aiutarci. Se ad esempio, immerso nella semi-oscurità, mi chiedo se l'animale che è davanti a me sia proprio un gatto, ciò che voglio sapere è in effetti se ciò che vedo corrisponde a ciò che vedrei se l'animale fosse in piena luce. E, in modo analogo, quando un fisico si chiede se la fusione fredda sia possibile, vuol sapere se l'esistenza di un certo processo fisico sia coerente o no con quella teoria completa della materia che la fisica si prefigge come un ideale a cui tendere. Possiamo dire perciò che, quando definiamo la verità di $RM(O_n)$ o $RC(O_n^*)$ come una certa corrispondenza per somiglianza tra O_n (o O_n^*) e O , in realtà per un verso stiamo pensando alla somiglianza tra l'oggetto interno della nostra rappresentazione e l'oggetto interno della rappresentazione idealizzata che di O darebbe un osservatore ideale perfettamente informato e, per un altro, senza contraddirci stiamo identificando tale secondo oggetto interno, quello della rappresentazione ideale, con l'oggetto reale. Ossia la rappresentazione $RM(O_n)$ o $RC(O_n^*)$ è *simil-vera* rispetto a O se e solo se:

- i. (O_n (o O_n^*) è simile, per qualche aspetto rilevante, a O_n^{**} , oggetto interno di quella rappresentazione culturale $RC(O_n^{**})$ che di O darebbe un osservatore ideale, e
- ii. si assume per *default* che O_n^{**} sia identico a O .

In altre parole una rappresentazione mentale o culturale di un certo oggetto reale è vera, secondo questa concezione corrispondentistica ingenua della verità, se e solo se il suo oggetto interno assomiglia a sufficienza all'oggetto interno della rappresentazione che di quello stesso oggetto reale darebbe un osservatore ideale (un osservatore cioè che, essendo onnisciente, si rappresenta il mondo come esso effettivamente è e conosce integralmente la verità su tutto). In

tal modo la misteriosa relazione di somiglianza tra l'oggetto interno di una mia qualsiasi rappresentazione soggettiva e l'oggetto reale viene sostituita con una comprensibilissima relazione concettuale tra gli oggetti interni di due rappresentazioni: la mia e quella dell'osservatore ideale. Ciò però ha una conseguenza indesiderata per il sostenitore della verità come corrispondenza: la relazione di corrispondenza tra la mia rappresentazione e l'oggetto reale si rivela essere in effetti una relazione concettuale di coerenza tra l'oggetto interno della mia rappresentazione e l'oggetto interno della rappresentazione dell'osservatore ideale. La concezione corrispondentistica della verità non è che una variante della teoria coerentistica; una variante nella quale alla coerenza tra una qualsiasi rappresentazione e tutte le altre rappresentazioni del suo stesso ordine di discorso viene a sostituirsi la coerenza tra due rappresentazioni appartenenti a due differenti ordini di discorso (l'ordine di O_n o O_{n*} , da un lato, e l'ordine di O_{n**} dall'altro). Pertanto anche nella teoria corrispondentistica per somiglianza, al pari che in quella coerentistica, l'aggancio con il mondo reale andrebbe perduto se esso non venisse tacitamente e surrettiziamente reintrodotta mediante l'assunzione, priva di qualsiasi ragione di sostegno, che l'osservatore ideale, se esistesse, sarebbe capace di rappresentare il mondo come esso effettivamente è, ossia mediante l'assunzione arbitraria che O_{n**} sia identico a O . Per questa ragione la teoria corrispondentistica ingenua della verità, che può essere definita come la teoria della corrispondenza (tra O_n [o O_{n*}] e O) mediante somiglianza (tra O_n [o O_{n*}] e O_{n**}) e identità (tra O_{n**} e O), pur essendo perfettamente adeguata a cogliere che cosa s'intenda comunemente per verità, non riesce a chiarire per davvero, a causa dell'oscurità del secondo passaggio da O_{n**} a O , che cosa mai possa essere quella relazione di corrispondenza tra mente e mondo che pur è alla base di tutte le teorie corrispondentiche.

(2b) *Corrispondenza tra linguaggio e metalinguaggio.* — Tutti i problemi della teoria corrispondentistica ingenua sembrano essere superati dalla concezione tarskiana della verità. Secondo Tarski (1956) – com'è noto – la proposizione “ p ” è vera (o meglio, potremmo dire, è “Tarski-vera”) se e solo se p . In tal modo la verità viene intesa,

per ogni proposizione p appartenente ad un linguaggio L , come una relazione di equivalenza chiaramente definita tra p stessa e la meta-proposizione “La proposizione ‘ p ’ è vera” appartenente ad un metalinguaggio L' nel quale vengono descritte certe caratteristiche di L . Ad esempio la proposizione “La neve è bianca” è vera se e solo se la neve è bianca. Da misteriosa relazione tra linguaggio e mondo la verità sembra essere divenuta, grazie a Tarski, una relazione chiaramente definita tra linguaggio e metalinguaggio. Ma, a ben guardare, trarre dalla teoria di Tarski questa conclusione è errato. In primo luogo perché Tarski stesso dette la sua definizione della verità solo in riferimento a linguaggi formalizzati. Essa vale perciò solo dove sia possibile distinguere con chiarezza il linguaggio al quale appartiene la proposizione p dal metalinguaggio al quale appartengono sia il nome “ p ” di quella stessa proposizione sia il predicato “essere vero” – come avviene nella fisica, nella matematica o nella logica, scienze dotate di un linguaggio formalizzato – mentre tale definizione non è applicabile alle rappresentazioni mentali, dato che è impossibile (almeno per ora) trovare nelle scienze umane la precisione e il grado di formalizzazione necessari. In secondo luogo – e ciò è ancora più importante – la definizione tarskiana della verità è solo in apparenza corrispondentistica. In effetti essa, ponendo “ ‘ p ’ è vera ” come equivalente a “ p ” o dichiara la superfluità del predicato “essere vero”: e allora è una definizione deflazionistica (o “decitazionale”) della verità; oppure stabilisce una relazione, tutta interna alla dimensione linguistica, tra il linguaggio oggetto L ed il metalinguaggio L' : e allora da questa concezione intra-linguistica della verità scompare ogni effettiva corrispondenza tra linguaggio e mondo. Se a qualcuno sembra il contrario, è solo perché questi, sposando il “fisicalismo semantico” degli empiristi logici (ad es. Beckermann 1999, p. 63 sgg.), identifica la descrizione fisicalistica della realtà con la realtà stessa. E ciò equivale ad assumere di nuovo, senza un argomento chiaro a sostegno, l’identità tra ciò che è oggetto di un certo ordine di discorso (quello del linguaggio fisicalistico) e la realtà stessa, ossia l’identità tra un certo O_n^{**} e O .

(2c) *Corrispondenza tra mente e mondo mediante causalità e identità.* — Una definizione della verità effettivamente di tipo corri-

spondentistico non può evitare perciò d'ipotizzare una qualche relazione tra la mente ed il mondo. Ma come farlo senza ricadere nelle difficoltà in cui incappa la definizione della verità come relazione di somiglianza tra l'oggetto interno di una rappresentazione ed il suo oggetto Intenzionale (coincidente con un oggetto reale nelle rappresentazioni empiriche vere)? Il modo più semplice e chiaro sembra essere quello di limitarsi a sostituire tale relazione di somiglianza con una relazione di causa ed effetto intercorrente tra le rappresentazioni mentali e la realtà esterna. Ciò è particolarmente plausibile a livello percettivo: se la percezione che ho di O è causata da O , allora in condizioni normali tale percezione è vera. Ad esempio, la differenza tra il vedere un'oasi effettivamente esistente e l'avere un miraggio consiste nel fatto che il primo stato mentale è causato dall'eccitazione che la luce riflessa dall'oasi provoca nei neuroni delle mie retine, mentre il miraggio ha altre cause. Generalizzando, la rappresentazione mentale $RM(O_n)$ è "causal-vera" se e solo se essa è causata da O (in quanto oggetto fisico che un osservatore ideale descriverebbe come un O_n^{**}) e si assume come evidente che O_n^{**} è identico a O .

Questa teoria va incontro tuttavia a due obiezioni. In primo luogo talvolta abbiamo percezioni false: ad esempio, come già detto, il mio vedere un gatto può essere causato, se c'è poca luce, dalla presenza di uno scoiattolo. Secondo questa teoria ingenua della corrispondenza mediante causalità, invece, una percezione causata da uno scoiattolo dovrebbe essere per definizione la percezione vera di uno scoiattolo. Percezioni ingannevoli sarebbero logicamente impossibili, se questa teoria fosse corretta; ma, poiché esse esistono, la teoria è scorretta.

In secondo luogo la riduzione della corrispondenza tra mente e mondo ad una relazione causale dal mondo alla mente si basa su un presupposto non argomentato: poiché tale relazione causale è spiegabile mediante le leggi e le teorie delle scienze naturali, essa è formulabile solo se sia le rappresentazioni mentali del mondo esterno sia il mondo esterno stesso vengono descritti nel linguaggio di queste scienze, ossia solo se, per lo più in modo implicito, viene data per scontata l'identità tra la realtà (O) e quella sua descrizione scientifica idealizzata (O_n^{**}) che, per delle ragioni ritenute troppo ovvie per

essere rese esplicite, viene considerata come quella descrizione del mondo veramente completa e del tutto corretta a cui la scienza deve tendere (anche se mai riuscirà a raggiungerla pienamente). Assunzione questa che la teoria della corrispondenza mediante causalità e identità non è in grado di giustificare.

(3) *La concezione pragmatistica ingenua della verità.* — Per tentare di superare le due difficoltà summenzionate occorre perciò modificare la teoria causale precedente. Anzitutto, per evitare l'obiezione basata sull'esistenza di percezioni ingannevoli, bisogna prendere in considerazione non solo le relazioni causali che vanno dal mondo alla mente (le percezioni in primo luogo), ma anche quelle che ritornano dalla mente al mondo (ossia le azioni volontarie tese a modificare l'ambiente esterno). Pertanto, estendendo questa teoria causale dalle sole percezioni a tutte le rappresentazioni mentali, si può dire che la rappresentazione mentale $RM(O_n)$ è vera di O se e solo se il suo rappresentare O come un O_n fa parte di una catena causale che, in condizioni normali, aumenta la probabilità che qualsiasi azione A che l'agente compia su O nell'intento (o desiderio $D(G_n)$) di ottenere un qualche scopo G abbia successo. In modo analogo si può interpretare in senso causale anche la verità delle rappresentazioni culturali, quando ci si ricordi che esse possono essere causalmente efficaci solo tramite delle rappresentazioni mentali individuali che le implementino (*supra* capitolo terzo). Pertanto in questa concezione causalistica estesa la verità è ridotta ad un insieme di relazioni causali che rendono in genere più efficace il comportamento dell'agente. In altre parole, secondo questa concezione della verità che potremmo chiamare pragmatistico-ingenua, la rappresentazione mentale $RM(O_n)$ è "util-vera" se e solo se in casi normali le due seguenti condizioni sono soddisfatte:

- i. O , in quanto considerato come qualcosa di fisico che in una descrizione scientifica idealizzata sarebbe descritto come un O_n^{**} , causa $RM(O_n)$ e questa rappresentazione mentale, combinandosi con altri stati mentali come in particolare il desiderio di ottenere lo stato di cose G , causa a sua volta

- l'azione A ; azione che infine produce (o quanto meno aumenta significativamente la probabilità di produrre) G stesso;
- ii. si assume implicitamente come evidente che O_n^{**} è identico a O .

Ora, poiché le relazioni causali menzionate in (i) sono spiegabili mediante leggi di natura e teorie scientifiche, la concezione pragmatica della verità è quella che più facilmente si accorda con la naturalizzazione dell'Intenzionalità e, più in generale, con una concezione scientifica del mondo. Tuttavia anche questa definizione pragmatica della verità è ingenua, perché, al pari della concezione causalistica, continua ad identificare, senza alcun argomento a sostegno, l'oggetto reale O con l'oggetto interno O_n^{**} di quella descrizione ideale di O (ossia $RC(O_n^{**})$), che, formulata nei termini delle scienze naturali, si assume che faccia parte di quella cornice teorica di sfondo che, per *default*, si ritiene che stia alla base di qualsiasi teoria scientifica.

(4) *La concezione olistico-quiniana della verità.* — Le definizioni della verità come simil-verità o come util-verità appaiono intuitivamente corrette non solo nella vita di tutti i giorni ma anche in campo scientifico finché la concezione generale del mondo che fa ad esse da sfondo viene considerata come evidente ed indubitabile, ossia finché nessuno dubita che un osservatore ideale, se volesse rappresentarsi O , lo rappresenterebbe come un O_n^{**} e, nel rappresentarselo così, lo vedrebbe come esso effettivamente è. In quel caso infatti, in virtù della supposta identità tra O e O_n^{**} , l'auspicata ma inverificabile e misteriosa corrispondenza tra O e l'oggetto interno O_n (o O_n^*) delle nostre rappresentazioni mentali o culturali di O stesso può essere tranquillamente sostituita con la più facilmente comprensibile corrispondenza (per somiglianza) tra O_n (o O_n^*) e O_n^{**} . Essere ragionevolmente sicuri della verità della percezione di O significa, ad esempio, avere fondate ragioni per credere che qualunque altra persona normale che sia in grado di osservare O nelle migliori condizioni possibili (e quindi *a fortiori* l'osservatore ideale) percepirebbe quello che percepiamo noi. In modo analogo il risultato di un espe-

rimento scientifico viene considerato affidabile se, essendo stato l'esperimento stesso ripetuto più volte sempre con il medesimo esito, la comunità scientifica si convince che, almeno sino a prova contraria, nessun ulteriore esperimento potrà contraddirlo e che quindi il suo esito va inserito nell'immagine scientifica del mondo (equivalente, almeno per *default*, alla rappresentazione che della realtà darebbe l'osservatore ideale).

In altre parole l'uomo della strada crede di solito nell'esistenza di una verità assoluta e nella sua conoscibilità. E anche lo scienziato, quando la sua disciplina si trovi in un periodo di graduale sviluppo, non pensa nulla di diverso. Ma, quando una scienza cessi di essere "normale" ed entri in un periodo di crisi dei propri fondamenti – una crisi che in genere la porterà ad un "cambiamento di paradigma", per riprendere il notissimo concetto di Kuhn (1962) – allora diviene palese che nessun osservatore può in linea di principio essere considerato ideale. Un osservatore che, vedendo il mondo com'è, giunga alla verità assoluta non solo non esiste di fatto, ma è impossibile in linea di principio che esista. Ogni rappresentazione O_n^{**} di O è necessariamente interna ad una certa particolare teoria (e perciò ad una qualche $RC(O_n^{**})$). Nessuna rappresentazione del mondo può essere *from nowhere* e quindi nessuna può coincidere con la rappresentazione assolutamente vera di un supposto osservatore ideale. In altre parole ogni concezione generale del mondo che per un certo periodo di tempo venga accettata dalla maggior parte degli scienziati come la cornice teorica che fa da sfondo a tutte le teorie scientifiche particolari è in effetti una normale teoria scientifica come le altre e quindi è rivedibile al pari delle altre. La sua differenza rispetto alle teorie scientifiche particolari consiste solo nel fatto che, essendo essa molto più ampia e complessa, non può essere falsificata mediante un singolo *experimentum crucis* e quindi viene di solito abbandonata solo in presenza di una cornice teorica alternativa che appaia preferibile agli scienziati per ragioni pragmatiche. È in genere utile insomma, almeno per un certo periodo di tempo finché nuovi dati sperimentali non mettano in crisi le vecchie idee, che l'oggetto interno O_n^{**} di una certa rappresentazione culturale idealizzata di O (sia essa pienamente nota oppure sia essa in molti casi particolari accettata per *default* anche dove manchino chiari riscontri empirici) venga

considerato come identico a O e quindi possa fungere da metro di paragone per la verità di tutte le altre teorie scientifiche su O stesso, anche se in linea di principio O_n^{**} è un oggetto interno ad una teoria come tutti gli altri oggetti interni e quindi è rivedibile come tutti gli altri alla luce di eventuali nuovi dati empirici.

Ma le cose cambiano, quando si verifichi in qualche scienza – e soprattutto in fisica, scienza generale per eccellenza – un mutamento di paradigma. In questo caso emerge con chiarezza quella «relatività ontologica», messa in luce da Quine (1969), di ciò che è oggetto di scienza. Volendo tradurre il pensiero di Quine mediante i termini e i simboli qui adottati, si può dire che la verità di una particolare teoria scientifica $RC(O_n^*)$ non può mai essere valutata, neppure da un punto di vista puramente pragmatico, mediante un confronto diretto con O per la semplice ragione che i dati sperimentali O_p di cui uno scienziato può disporre non sono mai sufficienti per isolare nel complesso della realtà un particolare oggetto O . La critica al secondo dei «due dogmi dell'empirismo» (Quine 1953), ossia l'impossibilità di trovare, per ogni concetto teorico, un unico insieme di dati osservativi a cui il suo significato sarebbe riducibile, fa sì che il riferimento empirico di ogni concetto teorico sia in linea di principio imperscrutabile. Nessun dato sperimentale può da solo falsificare una particolare teoria, mentre il medesimo insieme di dati sperimentali può confermare teorie tra loro incompatibili alla luce di differenti cornici teoriche di sfondo (sottodeterminazione delle teorie scientifiche rispetto ai dati osservativi). Nessun singolo *experimentum crucis* può falsificare una singola teoria scientifica. La relazione tra le teorie scientifiche ed il complesso dei dati osservativi è di tipo olistico.

Ciò non significa però che i dati sperimentali non siano di cruciale importanza per lo sviluppo della scienza nel suo complesso, come è particolarmente evidente nel momento in cui venga a maturazione un cambiamento di paradigma. Infatti, quando per un certo numero di anni si moltiplicano gli insuccessi degli esperimenti che sono stati concepiti e realizzati entro la cornice teorica generale dominante, gli scienziati di solito vanno in cerca di nuove e più adeguate cornici teoriche. Ma quasi mai questa ricerca dà, almeno all'inizio e talvolta per lunghi anni, un risultato univoco accettato da

tutti gli scienziati. È più facile che questi ultimi siano costretti a scommettere chi su questa chi su quella cornice teorica basandosi su criteri puramente pragmatici come la sua eleganza, semplicità o economicità. In altre parole uno scienziato che voglia accertare la verità o meno della teoria scientifica $RC(O_n^*)$ non solo non può mai confrontarla direttamente con O , ma talvolta, se egli si trova in un periodo di mutamento di paradigma della sua scienza, purtroppo non può neppure confidare nella identità di O con un O_n^{**} che sia l'oggetto interno di una cornice teorica da tutti riconosciuta come indubitabile; deve piuttosto assumersi il rischio di scegliere fra vari O_n^{**} basandosi su criteri puramente pragmatici. E questa scelta ovviamente è sempre provvisoria e rivedibile.

Tale relatività di O_n^{**} rispetto ad una cornice teorica di sfondo sempre rivedibile in linea di principio può essere interpretata tuttavia in due modi molto diversi. Da un lato si può ritenere che la realtà sia assolutamente inconoscibile (in altre parole sia una kantiana “cosa in sé”) e che quindi non solo nessuna cornice teorica generale possa mai essere considerata, quasi fosse una “verità eterna” della vecchia metafisica, come quella assolutamente e definitivamente vera, ma che anzi non abbia alcun senso neppure chiedersi quale tra le cornici disponibili si avvicini di più alla verità, dato che non si dispone di alcun criterio per valutare tale distanza. E allora non si ha più alcuna certezza che, data una qualsiasi teoria scientifica $RC(O_n^*)$ esista un unico oggetto reale O , riferimento comune e metro di validità di tutte le teorie e rappresentazioni che lo riguardano. La teoria di Quine sulla “relatività ontologica” viene intesa, secondo questa prima interpretazione, come un completo relativismo cognitivo. Ogni teoria scientifica ha i propri peculiari impegni ontologici; e quindi, come non esiste una verità assoluta, non esiste neppure una realtà assoluta (cfr. in particolare De Caro e Mcarthur 2004 e Putnam 2016). Ma ciò fa perdere di vista l'unicità del mondo reale e l'aspirazione della scienza a dare, sia pur nella diversità degli ordini di discorso a cui appartengono le scienze naturali e le scienze cognitive, umane e sociali, una rappresentazione unitaria e coerente di tutta la realtà (Nannini 2007a, pp. 85-91).

Oppure, dall'altro lato, si può ritenere che, sebbene una rappresentazione assolutamente e definitivamente vera di O non sia raggiun-

gibile, ciò non sia un motivo sufficiente per dubitare dell'esistenza di qualcosa, sia esso o meno precisamente un O , che nella realtà in sé è ciò a cui lo scienziato si riferisce quando parla di O come di un O_n^* (o quanto meno lo identifica per *default* con un ancora non pienamente conosciuto O_n^{**}). Anzi la scienza stessa perderebbe di senso se, a differenza di quanto pensano i relativisti, non si credesse nell'esistenza di una realtà indipendente dal nostro conoscerla e non si pensasse che è tale realtà in sé a confermare o smentire le nostre previsioni su di essa quando le mettiamo alla prova mediante opportuni esperimenti (cfr. *supra* capitolo quinto).

Del resto, nonostante qualche oscillazione, questa sembra essere l'interpretazione suggerita da Quine stesso del suo concetto di relatività ontologica, allorché egli affermava di non vedere alcuna contraddizione nel sostenere un «realismo un po' ingenuo» e al tempo stesso nel «salutare l'uomo come soprattutto l'autore, piuttosto che lo scopritore, della verità». Secondo Quine infatti si potevano «sostenere i due punti [di vista] insieme perché la verità scientifica relativa agli oggetti fisici è ancora la verità, per tutti gli uomini che ne sono autori». Quine insomma non vedeva contraddizione in questo suo duplice atteggiamento per la semplice ragione che secondo lui da un lato «parliamo sempre all'interno del nostro sistema attuale quando attribuiamo la verità; e non possiamo fare altrimenti»; ma dall'altro lato, quando avvenga un cambiamento nell'immagine scientifica del mondo, «[...] noi non diciamo che la verità cambia con esso, ma che noi abbiamo prima erroneamente pensato che qualcosa fosse vero e che poi abbiamo migliorato le nostre conoscenze. Fallibilismo è la parola d'ordine, non relativismo. Fallibilismo e naturalismo» (Quine 1981, pp. 33-34)³⁶.

Pertanto si può dire che una certa concezione scientifica generale del mondo è provvisoriamente “Quine-vera”, rispetto al mondo reale stesso nella sua interezza, se gli scienziati hanno ragioni sufficienti alla luce di criteri pragmatici per considerarla come la cornice teorica maggiormente capace di selezionare le teorie scientifiche maggiormente dotate di potere predittivo, retrodittivo ed esplicativo, finché la ricerca stessa non ne suggerisca una migliore.

³⁶ Per un commento più ampio cfr. Nannini (2007a, pp.64-66).

(5) *La concezione naturalistica della verità* — Prendiamo ora le mosse dalle due concezioni della verità che sembrano essere preferibili per chi voglia muoversi entro l’orizzonte teorico di una epistemologia naturalizzata – ossia la “corrispondenza tra mente e mondo mediante causalità e identità”, da un lato, e la “concezione olistico-quiniana”, dall’altro, quest’ultima soprattutto nella sua versione contraria al relativismo – e vediamo se è possibile ricavarne un concetto unitario e coerente di “natural-verità” che possa stare a fondamento di un’ontologia del mentale in sintonia con una soluzione naturalistico-fisicalistica del problema mente-corpo (*supra* capitolo secondo e capitolo quinto).

In primo luogo il concetto di natural-verità deve essere in accordo con un’ontologia interdisciplinare che, per poter essere estesa dalle scienze naturali alle scienze cognitive, umane e sociali, deve ammettere una chiara distinzione tra l’esistenza assoluta e l’esistenza relativa ad un qualche ordine di discorso. Ad esempio gli stati mentali godono di esistenza solo relativa in quanto oggetti interni della psicologia e delle altre scienze cognitive, umane e sociali e possono acquisire esistenza assoluta unicamente in modo indiretto se vengono implementati da processi fisico-chimici (in particolare processi cerebrali). In caso contrario sono dei meri *ficta*.

La distinzione tra esistenza assoluta e esistenza relativa ne produce una analoga anche riguardo al concetto di verità. Ciò è particolarmente evidente riguardo ad asserzioni relative ai *ficta*. Ad esempio non solo per il senso comune ma anche nei tribunali è molto importante sapere se è vero o no che un certo crimine è stato commesso di propria libera volontà. Rispondere a domande di questo tipo implica però che preventivamente ci si sia interrogati sull’esistenza o meno del libero arbitrio come capacità peculiare della mente umana. Per il senso comune (e in genere anche per i giudici) la risposta è ovviamente positiva, senza addentrarsi troppo nello spinoso problema metafisico della natura della mente e del libero volere.

E, tra i filosofi, tranquillamente positiva è anche la risposta a quella domanda da parte di tutti i fautori del dualismo mente-corpo. Ma per i naturalisti, avversari al dualismo anche perché ritengono che il concetto dualistico di libero arbitrio sia incompatibile con una concezione scientifica del mondo, le cose si complicano.

Alcuni naturalisti, i compatibilisti, definiscono come libera qualsiasi azione che, pur essendo determinata da cause che operano secondo leggi di natura, sia causata da un atto di volontà interno alla mente stessa dell'agente. Insomma un'azione è libera se è volontaria. Ma questa tesi va incontro ad una grave obiezione: come può un'azione essere libera se l'atto di volontà che ne è la causa non è a sua volta libero? Rispondere – come i compatibilisti, in ultima analisi, sono costretti a fare – che solo l'agire può essere libero, non il volere, suona molto strano. Come posso dire di aver agito liberamente se l'atto di volontà che mi ha spinto ad agire come ho agito è stato causalmente determinato, ad esempio, dai processi fisico-chimici che a mia insaputa si sono svolti nel mio cervello? Inoltre come può essere libero un atto di volontà se la libertà è stata definita come volontarietà? Non si finisce in tal modo per affermare (o tautologicamente o insensatamente) che la volontà è libera quando è volontaria? E tuttavia anche negare *tout court* l'esistenza del libero arbitrio sembra molto difficile. Se agire di propria libera volontà è impossibile, allora sembra essere insensato, almeno *prima facie*, parlare di responsabilità morale o giuridica!

Il naturalista ha tuttavia a disposizione una soluzione per evitare questo dilemma grazie ai concetti di *fictum*, di esistenza relativa e di “verità relativa”. Infatti, se si assume che il libero arbitrio, in quanto capacità peculiare della mente umana, sia un *fictum*, allora risulterà che esso entro l'ordine di discorso adottato nella vita di tutti i giorni, nei tribunali e spesso anche nelle scienze umane e sociali è una capacità “realmente” (nel senso di una esistenza relativa) posseduta dagli esseri umani; ossia, se si indica con il simbolo L_{fp} l'ordine di discorso della *folk-psychology*, il libero arbitrio in quanto L_{fp} -*fictum* gode di L_{fp} -esistenza relativa. Di conseguenza tutte le asserzioni che menzionano tale L_{fp} -*fictum* possono essere L_{fp} -vere o L_{fp} -false, sebbene non esistano processi cerebrali (e, più in generale, L_0 -processi fisico-chimici) che lo implementino. E ciò è epistemologicamente rilevante. Infatti da un lato tutte le asserzioni riguardanti il libero arbitrio sono a rigore L_0 -false (oppure L_0 -insensate) e devono essere espunte da ogni concezione scientifica del mondo, dato che il libero arbitrio non gode di L_0 -

esistenza assoluta, ma dall'altro la loro L_{fp} -verità (o L_{fp} -falsità) risulta in moltissime occasioni molto utile o addirittura indispensabile. Nella vita di tutti i giorni e spesso anche nei tribunali o nelle scienze umane e sociali è inevitabile prendere decisioni e interagire con gli altri partendo dal presupposto che noi esseri umani siamo capaci di agire di nostra propria libera volontà. Il concetto di esistenza relativa consente perciò di conciliare la L_{fp} -verità di certe L_{fp} -affermazioni del senso comune, verità alla quale non potremmo rinunciare senza indesiderabili conseguenze pratiche, con la loro L_0 -falsità (o insensatezza) dal punto di vista scientifico³⁷.

In secondo luogo Quine aveva in mente soprattutto la fisica quando, rifiutando il relativismo, sosteneva che le concezioni fondamentali della realtà prevalenti tra gli scienziati nel corso del tempo si succedono l'una all'altra non in modo arbitrario, bensì sulla base di criteri pragmatici alla ricerca di una sempre migliore adeguatezza empirica. Ma – quando, facendo uso della distinzione tra esistenza assoluta ed esistenza relativa, la validità dell'epistemologia di Quine venga estesa anche alle scienze cognitive, umane e sociali – anche il suo non-relativismo, limitato in Quine al modo nel quale si succedono nel tempo le teorie appartenenti ad un'unica scienza, deve essere esteso al modo nel quale entrano in relazione gli impegni ontologici delle varie scienze. Si verrà a sostenere in tal modo che ad esempio i rispettivi impegni ontologici della fisico-chimica e delle scienze naturali in generale, da un lato, e delle scienze cognitive, umane e sociali dall'altro, lungi dall'appartenere a ordini di realtà ontologicamente irriducibili e dall'essere dotati tutti di esistenza relativa sullo stesso piano, risultano essere ordinabili gerarchicamente mediante delle relazioni di implemen-

³⁷ Tuttavia la consapevolezza che molte affermazioni riguardanti la volontarietà o meno di questa o quell'azione, pur essendo L_{fp} -vere, sono L_0 -false (o sono L_0 -insensate), dovrebbe indurre i giudici, allorché consultino uno psicologo per sapere se un certo imputato era, come si suol dire, “capace di intendere e di volere” nel momento in cui ha commesso un certo crimine, a tenere in maggior conto di quanto ciò spesso non avvenga il fatto che il confine tra atti volontari e atti involontari è difficilmente tracciabile in molti casi *border-line* per la semplice ragione che il concetto di libero arbitrio proprio del senso comune è un concetto ambiguo e sfuocato, estraneo ad una visione scientifica del mondo.

tazione da reale a virtuale. E pertanto gli psicologi ad esempio, se vogliono formulare teorie empiricamente corrette, devono tener conto del fatto che gli stati mentali, pur essendo dei legittimi oggetti di studio della psicologia dotati di esistenza relativa entro quella scienza, possiedono però poteri causali solo nella misura in cui sono delle entità virtuali (e non fittizie!) implementate da processi fisico-chimici.

In conclusione con queste due correzioni, partendo dalla Quine-verità, è possibile giungere ad una concezione pragmatistica della “natural-verità” che possa fungere da fondamento ad una ontologia del mentale del tutto adeguata alle esigenze del naturalismo cognitivo. Secondo questa concezione una teoria scientifica o, più in generale, una rappresentazione culturale $RC(O_n^*)$ è “natural-vera” se e solo se risultano soddisfatte le seguenti condizioni:

1. $L_n-RC(O_n^*)$, appartenente all'ordine di discorso L_n , è L_n -Quine-vera quando il suo oggetti interno O_n^* , un oggetto virtuale dotato di L_n -esistenza relativa, sia sufficientemente simile sotto qualche riguardo rilevante all'oggetto virtuale interno (anch'esso L_n -esistente) di $L_n-RC(O_n^{**})$, cornice teorica generale che viene provvisoriamente accettata per *default* entro L_n come la migliore rappresentazione possibile, in termini pragmatici, della realtà.
2. O_n^* gode indirettamente di L_0 -esistenza assoluta (e pertanto è anche dotato di poteri causali) quando, non essendo un *fictum*, sia implementato in ultima analisi (eventualmente mediante uno o più passaggi intermedi) da dei processi fisico-chimici (inclusi i processi biologici) che siano l'oggetto interno di teorie del tipo $L_0-RC(O_0^*)$.
3. Gli oggetti interni O_0^* di tali L_0 -teorie, appartenenti all'ordine di discorso L_0 proprio della fisico-chimica e di tutte le scienze naturali, sono sufficientemente simili (o quanto meno si suppone per delle buone ragioni che siano sufficientemente simili) sotto qualche riguardo rilevante

a qualche aspetto di O_0^{**} , oggetto interno (anch'esso L_0 -esistente) di $L_0\text{-RC}(O_n^{**})$, cornice teorica generale che, sebbene non ancora pienamente conosciuta, viene ritenuta per *default* dalla maggioranza degli scienziati entro L_0 come L_0 -Quine-vera, in quanto essa sembra essere – sebbene solo in modo olistico, in termini pragmatici e salvo prova in senso contrario – come la migliore concezione generale possibile dell'intero mondo fisico, l'unico mondo che goda di esistenza assoluta in modo diretto.

4. Conclusione.

Se si torna al problema da cui la precedente analisi del concetto di verità ha preso le mosse, risulta ora più chiaro come si possa evitare l'apparente contrasto tra la concezione corrispondentistica della verità propria del senso comune e le concezioni pragmatistiche sopra menzionate (*in primis* la concezione della natural-verità) riguardo alla verità o meno anche della più semplice e banale delle percezioni. Queste due concezioni della verità delle percezioni (e, più in generale, di tutte le rappresentazioni mentali e culturali, comprese le teorie scientifiche) sembrano essere entrambe irrinunciabili, l'una soprattutto per l'utilità che ha nella vita di tutti i giorni e l'altra principalmente come idea-guida della ricerca scientifica. Ma al tempo stesso esse sembrano essere anche difficilmente conciliabili.

Tuttavia l'analisi precedente sulle varie teorie della verità può aiutare a comprendere come un naturalista possa uscire da questo dilemma. Esiste sicuramente un contrasto tra il "corrispondentismo mediante somiglianza e identità" e la natural-verità propria di un'epistemologia naturalizzata. Ma l'aver introdotto nel concetto di natural-verità una distinzione tra verità assoluta e verità relativa ad un certo ordine di discorso consente di evitare che quel contrasto divenga una contraddizione. Quando affermo che la mia percezione che il semaforo è rosso è vera perché esso è effettivamente rosso, sto dicendo che la mia percezione $L_n\text{-RM}(O_n)$, che identifica O con O_n , è L_n -simil-vera. Quando invece valuto quella mia affermazione nei termini della natural-verità, essa è a rigore L_0 -falsa

(o L_0 -insensata), perché O_n , (il *quale* “essere rosso”) in quanto oggetto che pretende di essere reale ma che in effetti è solo un oggetto fenomenico, in ultima analisi non gode di alcuna L_0 -implementazione fisico-chimica, ossia non gode di alcuna esistenza assoluta e pertanto è solo un L_n -*fictum* privo, in quanto tale, di autentici poteri causali³⁸.

Inoltre, riguardo alla possibilità di avere percezioni ingannevoli, chi noti che una percezione può essere vera ma inutile o, viceversa, falsa ma utile sta facendo implicitamente riferimento alla definizione della verità come “corrispondenza mediante somiglianza e identità”, mentre chi replica che l’identità tra verità e utilità vale non per le singole percezioni, bensì per il meccanismo biologico che le genera, si sta riferendo alla natural-verità di teorie scientifiche di tipo neo-darwiniano. In altre parole, nella vita di tutti i giorni una L_n -percezione di O da parte di un certo agente può essere sì L_n -simil-vera o L_n -simil-falsa indipendentemente dalle funzioni che questa percezione può svolgere nel guidare le azioni di quell’agente su O , ma il suo essere natural-vera o natural-falsa non è invece affatto indipendente da tali funzioni. In conclusione chi – per criticare le teorie naturalistiche dell’Intenzionalità e le concezioni causalistiche o pragmatistiche della verità a cui tali teorie si appoggiano – faccia appello alla possibilità dell’errore dice qualcosa che è sì intuitivamente ragionevole e convincente, ma solo finché si resti nell’ordine di discorso della *folk-psychology* e si intenda perciò la verità in un senso corrispondentistico ingenuo. Quando invece si abbandoni il punto di

³⁸ Si noti che non è corretto obiettare che in effetti le onde elettro-magnetiche che colpiscono il mio occhio hanno il potere di causare in me la percezione del colore rosso se sono di un certa lunghezza. È verissimo ovviamente che esse, in un certo senso, hanno quel potere; ma ciò significa unicamente che quelle onde elettromagnetiche sono una concausa essenziale della mia percezione-di-rosso, non significa affatto che il *quale* “essere rosso” sia ad esse identico e proprio per questo goda di esistenza assoluta e sia dotato di poteri causali. La mia percezione-di-rosso (che è in effetti un “percepire rossamente”, cfr. *supra* capitolo secondo) gode di esistenza assoluta in quanto è implementata da processi cerebrali, mentre il suo oggetto interno, il colore rosso in quanto contenuto di quella percezione e da essa separabile solo nel “formato coscienziale”, è solo un *fictum* privo di ogni potere causale.

vista della *folk-psychology*, allora appare chiaro che il concetto di verità non può più essere sganciato dalla funzione pratica che le percezioni svolgono nel guidare le interazioni tra gli esseri umani ed il mondo.

Capitolo settimo

Kant e le scienze cognitive sulla natura dell'Io

1. Coscienza e autocoscienza da Cartesio a Hume

Possono le celebri pagine della seconda edizione della *Kritik der reinen Vernunft* sull'appercezione trascendentale (Kant 1787, pp. 110-112) essere di un qualche aiuto oggi per quei filosofi della mente d'indirizzo naturalistico e per quegli scienziati cognitivi (siano essi neuroscienziati o psicologi) che si interrogano sulla natura della coscienza e dell'autocoscienza? La prima tentazione è di rispondere seccamente no! Che può avere a che fare infatti uno dei maggiori classici della "filosofia continentale" con la filosofia analitica o post-analitica e, più in particolare, con la filosofia della mente figlia di quell'incontro, a partire dagli anni Sessanta, tra pensiero filosofico e scienze cognitive che può essere chiamato (duplice) "svolta cognitiva" (cfr. *supra* capitolo primo)? Ma chi ragionasse così darebbe una risposta affrettata ed errata alla domanda summenzionata. Non solo perché, da un lato, gli studiosi di Kant più attenti e più aperti al dibattito filosofico contemporaneo dimostrano oggi grande interesse per la svolta cognitiva in filosofia della mente (ad es. Barale 2006 e Brook 2008), ma anche e soprattutto perché, in modo molto più specifico e puntuale rispetto al tema qui affrontato, una rilettura attenta delle pagine dedicate da Kant all'appercezione trascendentale rivela una insospettata attualità di questo concetto nel dibattito attuale sulla natura dell'Io (o come molti preferiscono dire in inglese, del *Self*)³⁹. Per comprenderlo, conviene fare prima un brevissimo *excursus* storico sul concetto di Io o *Self*.

Mettendo in parentesi i precedenti reperibili nella filosofia antica e medievale, prendiamo le mosse da Cartesio. Com'è noto, quest'ultimo sosteneva che io sono una *res cogitans*, una mente, ossia una sostanza separabile dopo la mia morte dal mio corpo (Cartesio 1641). Questa concezione della mente proposta da Cartesio fu criticata nel corso del XVII e XVIII secolo da più parti. In particolare venne rifiutata per un verso da T. Hobbes nelle *Objectiones tertiae* alle

³⁹ Normalmente "Io" viene inteso in filosofia della mente come la traduzione italiana di *Self*. Ma qui (come già anticipato nell'*Introduzione* e nel secondo capitolo) proporrò invece una distinzione tra questi due concetti.

Meditationes di Cartesio (Hobbes in Cartesio 1641, pp. 159-185), perché essa presuppone una soluzione dualistica del problema mente-corpo, e per un altro da Locke e Hume.

Ammettiamo – sosteneva Locke – che la teoria della metempsi-cosi sia vera e che in me riviva l'anima di Nestore o di Tersite. Potrei dire allora che, in quel caso, io sarei Nestore o Tersite? Sicuramente no, perché, anche se la mia anima fosse la stessa che animò il corpo di Nestore o di Tersite, io non ricorderei alcun episodio delle loro vite. Pertanto io non sono la mia anima, bensì la continuità dei miei ricordi (Locke 1690, pp. 512-515). Da ciò Hume trasse la conclusione che io non sono una sostanza, come voleva Cartesio, bensì solo un insieme di *perceptions*, ossia di stati mentali legati tra loro da associazioni operate dall'immaginazione:

ciò che chiamiamo *una* mente [*a mind*] non è altro che un fascio o collezione di percezioni differenti, unite da certe relazioni, e che si suppongono, sebbene erroneamente, dotate di una perfetta semplicità e identità (Hume 1739-1740, I p. 220; trad. modificata, corsivo mio).

Questa concezione della mente, che la declassa da sostanza a collezione di proprietà, permetteva inoltre a Hume di dubitare che essa sia una entità semplice e unica, della quale possiamo essere *self-conscious*. Ci sono dei metafisici – osservava ironicamente Hume – che dicono di avere coscienza di sé per via introspettiva. Beati loro! Io non sono capace di auto-percepirmi (Hume 1739-1740, pp. 263-265)!

La concezione lockiana e humiana della mente come collezione di stati mentali è divenuta nel Novecento la cosiddetta *bundle theory* del mentale dopo che le idee di Locke e Hume erano state riprese e rielaborate da W. James agli inizi del secolo XX per trasformarle nella concezione della coscienza e del *Self* non come cosa, bensì come «flusso», o meglio come «*stream of consciousness*» (James 1890 capitoli IX e X).

2. Il Self oggi tra filosofia e scienze cognitive

Queste due concezioni del *Self* si ritrovano anche oggi tra i filosofi della mente e gli scienziati cognitivi. Per quanto concerne la concezione cartesiana, soprattutto tra i filosofi non mancano infatti coloro che – si dichiarino essi apertamente antinaturalisti o preferi-

scano invece aderire al grande alveo del naturalismo liberalizzato e dell'emergentismo – sostengono comunque la irriducibilità ontologica degli stati mentali in generale e del *Self* in particolare a processi cerebrali o a loro ridescrizioni in termini funzionali. Sul versante opposto si trovano invece quei neuroscienziati e quei filosofi naturalisti che, sebbene divisi su molti altri argomenti, concordano sul rifiuto di ogni forma di dualismo ontologico tra mente e corpo e pensano che ciascuno di noi sia semplicemente identico all'attività del proprio cervello (cfr. *supra* capitolo primo).

Ora, non c'è dubbio che coloro che oggi sostengono questo secondo orientamento, con tutti i tentativi di riduzione della coscienza e dell'autocoscienza a processi cerebrali che esso comporta, sono eredi, sebbene in modo non sempre pienamente consapevole, di quella concezione materialistica del mentale che in età moderna trova la propria origine nel pensiero di Hobbes.

Tuttavia molto evidente è in questo genere di studi anche l'eredità di Locke, Hume e James. Ciò risulta particolarmente chiaro quando dagli scritti dei neuroscienziati ci si sposti sulle interpretazioni che delle loro teorie hanno dato molti filosofi di orientamento naturalistico. Patricia Churchland ad esempio nella sua introduzione alla «neurofilosofia» – disciplina a cavallo tra filosofia e neuroscienze da lei stessa fondata – ha chiarito, appoggiandosi in particolare a varie patologie del sistema nervoso, che il *Self* deve essere inteso non come una cosa, bensì come un insieme di funzioni espletate da processi cerebrali distinti (Churchland 2002, p. 59 sgg.):

Il Self non è un singolo schema rappresentazionale perfettamente coerente e unificato riguardo al quale noi avremmo credenze altrettanto pienamente coerenti e *unificate*. Il Self è piuttosto qualcosa che assomiglia ad una squadriglia di capacità che volano in ordine sparso. [...] La capacità *fondamentale* consiste però probabilmente nella coordinazione di bisogni, scopi, percezioni e ricordi con il controllo motorio (Churchland 2002, p. 63; trad. mia, corsivi dell'autrice).

Quest'ultimo accenno al carattere fondamentale che ricopre il senso di sé nel coordinamento senso-motorio mette tuttavia in evidenza che Patricia Churchland da un lato segue Hume e James nel ritenere che, quando siamo coscienti di noi stessi, «non sembra che ci sia alcuna

cosa che noi percepiamo come il nostro self [there does not seem to be any self *thing* there to perceive]. Ciò che si può cogliere introspettivamente è un *flusso* di percezioni visive, suoni, odori, emozioni, ricordi, pensieri e così via» (Churchland 2002, p. 59; trad. mia, corsivi dell'autrice). Ma dall'altro lato ella non vuole neppure eliminare il *self* dal novero di ciò che è reale e causalmente efficace:

Il cervello fa sì che noi pensiamo di avere un Self. Questo vuol dire che il Self che io penso di essere non è reale? No, esso è tanto reale quanto lo è qualsiasi attività del cervello. Certo, vuol dire però che il proprio Self non è un etereo pezzettino di "roba spirituale" [an ethereal bit of "soul stuff"]. Ma è tanto reale quanto lo è, ad esempio, la coerente attività neuronale che vi rende capaci di camminare o di pensare al riscaldamento globale oppure di ritrovare la strada di casa dopo una passeggiata nel bosco. L'attività cerebrale è una cosa del tutto reale (Churchland 2002, p. 124; trad. mia).

Molto più vicino allo spirito dissacratore di Hume e della *bundle theory* della mente è invece da venti anni a questa parte ciò che sostiene D.C. Dennett. Egli ammette da un lato che ovviamente esistiamo, se intendiamo dire con ciò che persino quando dubitiamo che il nostro cervello possa ospitare un *Self* noi siamo comunque un soggetto che si sta interrogando sulla propria esistenza. Ma dall'altro Dennett aggiunge subito dopo che non esistiamo, se la nostra esistenza viene scambiata per l'esistenza di qualche entità che controlli i nostri corpi, i nostri pensieri e le nostre decisioni, ossia per l'esistenza di una *res cogitans* (Dennett 1991, p. 459). Dennett sfugge (o almeno ritiene di sfuggire) alla contraddizione per cui in un senso esistiamo ed in un altro no sostenendo che certamente non siamo delle entità spirituali installate nel nostro corpo, ma esistiamo nondimeno nel senso che il nostro *Self* è il modello mentale che il cervello si costruisce dell'insieme delle proprie attività in quanto organo preposto al controllo delle interazioni dell'organismo con il mondo esterno (compresi non solo gli oggetti materiali, ma anche gli altri esseri umani e l'insieme della società e della cultura in cui siamo immersi):

Secondo la mia teoria – scrive Dennett – un sé [*Self*] non è un punto matematico, ma un'astrazione definita dalle miriadi di attribuzioni e interpretazioni (incluse le auto-attribuzioni e le

auto-interpretazioni) che hanno composto la biografia del corpo vivente di cui è il Centro di Gravità Narrativa. Come tale, svolge un ruolo singolarmente importante nell'incessante economia cognitiva di quel corpo vivente, perché tra tutte le cose dell'ambiente di cui un corpo attivo deve farsi un modello mentale, nessuna è più cruciale di se stesso (Dennett 1991, p. 474).

Per comprendere questo passo occorre tener presente che Dennett, dicendo che il *Self* è un'astrazione e più precisamente un «Centro di Gravità Narrativa», intende dire che esso non è un'entità reale distinta dal corpo; è piuttosto l'insieme dei processi cerebrali che implementano la rappresentazione che diamo di noi stessi in quanto *focus* virtuale verso il quale facciamo convergere l'insieme delle informazioni che abbiamo, per via propriocettiva e introspettiva, su ciò che sentiamo, immaginiamo, pensiamo e facciamo. E, poiché l'immagine che noi esseri umani abbiamo di noi stessi è largamente affidata al linguaggio e alla nostra autobiografia, si comprende in quale senso tale *focus* sia, appunto, narrativo.

Dennett non si nasconde che questa concezione del *Self* va incontro a varie obiezioni. In primo luogo si dirà – egli ammette – che «il guaio dei centri di gravità è che non sono reali; sono delle finzioni teoriche». Ma – replica Dennett – «questo non è il guaio dei centri di gravità; è il loro vanto. Essi sono *magnifiche* finzioni» (Dennett 1991, p. 477; corsivo dell'autore]. In secondo luogo – prosegue Dennett, introducendo la seconda e principale possibile obiezione alla sua teoria – che ne sarà della mia responsabilità morale, se io sono una finzione? Come può una finzione agire liberamente ed essere responsabile di ciò che fa? Lo può – replica Dennett – se intendiamo la libertà del volere come quella capacità di autocontrollo sulle proprie azioni che un essere vivente acquisisce quando sviluppi un cervello che, nel rispetto di tutte le leggi di natura, sia in grado, in quanto sistema complesso, di auto-rappresentarsi⁴⁰.

Dennett colloca inoltre questa concezione del *Self* entro un'articolata critica di gran parte delle scienze cognitive, alle quali rimprovera, anche quando credano di rifiutare il dualismo cartesiano, di restare in effetti prigionieri del mito che esista in noi un omuncolo,

⁴⁰ Sui sistemi complessi cfr. ad es. Kaplan e Glass (1995).

non importa se materiale invece che spirituale, che dall'interno del nostro corpo ne guiderebbe le azioni. Dennett chiama questo mito «Teatro Cartesiano», intendendo dire con ciò che molti scienziati cognitivi, ancora prigionieri della “analogia mente-computer”, vanno in cerca dell'esistenza nel cervello di un processore centrale (una C.P.U.) capace di elaborare in modo seriale, secondo un programma, l'*input* sensoriale per trasformarlo in *output* motorio. In altre parole per chi resti prigioniero del mito del Teatro Cartesiano le nostre rappresentazioni mentali divengono coscienti e acquisiscono per ciò stesso il controllo motorio del corpo una alla volta, allorché salgono su un Palcoscenico interno alla nostra mente e si presentano una dopo all'altra al Pubblico di un tale Teatro Interiore, ossia al *Self*, al Soggetto Autocosciente, all'Agente che guida il corpo nella sua interazione con l'esterno (Dennett 1991, pp. 119-129). A questa concezione della mente, che Dennett con un ardito ossimoro taccia di essere una forma di «materialismo cartesiano» (Dennett 1991, p. 125), egli contrappone il «Modello delle Molteplici Versioni» (Dennett 1991, pp. 129-146). Secondo Dennett, come nessuno possiede in un certo momento La Versione Autentica di un proprio *preprint* che egli faccia girare tra i propri amici e che corregga in continuazione sulla base dei loro suggerimenti via via che essi li inviano, allo stesso modo non esiste, in maniera univoca, ciò che costituisce “Il Contenuto” della nostra coscienza in un certo istante. In ogni momento circolano nel nostro cervello vari pacchetti di informazioni (implementati dall'attività di circuiti cerebrali funzionanti in parallelo) sullo stato del mondo esterno e del nostro corpo (incluso il cervello stesso) e conseguenti possibili progetti d'azione; pacchetti d'informazioni e progetti d'azione che, quasi fossero folletti in lotta tra loro, sono in competizione per accedere al controllo del sistema motorio (incluso l'apparato fonatorio). I folletti che per un certo tempo riescono a vincere la competizione e perciò riescono a guidare le nostre azioni (inclusi i nostri *speech acts*) sono anche in grado di costruire retrospettivamente una storia del loro successo (la storia è sempre scritta *ex-post* dai vincitori!) e con ciò a delineare anche la rappresentazione di un *Self* fittizio che ritiene (ora) di essere stato (allora) L'Agente (al singolare) che ha compiuto tali azioni: è così che io mi sento l'autore di ciò che ho fatto. E naturalmente questo sentimento di *agency* è al tempo stesso prodotto e parte del meccanismo che mi consente di

agire come una persona libera e responsabile. Come si vede, con Dennett la concezione umana della mente come collezione di stati mentali giunge alle sue estreme conseguenze. Il *Self* secondo Dennett è una finzione mediante la quale il cervello, auto-rappresentandosi, riesce ad acquisire un maggior controllo e un maggior potere di coordinamento dei movimenti del corpo in risposta alle informazioni fornite dai sensi sull'ambiente esterno e sul corpo stesso.

Questa dissoluzione del *Self* proposta da Dennett sta avendo oggi un notevole successo⁴¹ e sembra inevitabile concludere che, tra tutte le teorie del *Self* oggi prevalenti, è, insieme alla concezione proposta da Patricia Churchland, quella che più di ogni altra è in accordo con il naturalismo cognitivo. Basta però fare questi due esempi, Dennett e la Churchland, per comprendere che, se essi concordano nel rifiuto senza appello di ogni forma di dualismo, tuttavia riguardo alla realtà del *Self* propongono due tesi tutt'altro che coincidenti. Per Patricia Churchland il *Self*, in quanto identico a un insieme di processi cerebrali, è – se mi si passa l'espressione – realmente reale, mentre per Dennett è un Agente solo virtuale o addirittura fittizio.

Ma se è così, allora la concezione di Patricia Churchland è più vicina al senso comune, dato che ciascuno di noi ritiene spontaneamente di essere qualcosa di realmente reale! E tuttavia ciò entra in contraddizione con la tesi, anch'essa presente peraltro nel suo pensiero, che, se il *Self* è reale, lo è solo perché coincide con dei processi cerebrali che, in quanto «squadriglia di capacità che volano in ordine sparso» (cfr. *supra*), sono privi di quell'unità e coerenza che ciascuno di noi presta a sé stesso come agente libero e capace di compiere azioni volontarie.

La concezione del *Self* proposta da Dennett incappa invece nell'obiezione opposta: da un lato il suo concetto di "*Self* fittizio" (e quindi non reale, nel senso comune del termine) è pienamente in accordo con l'immagine che le neuroscienze ci offrono dell'attività

⁴¹ Cfr. in particolare, oltre al filosofo Metzinger (2009), lo psicologo Wegner (2002). Quest'ultimo ha però radicalizzato la teoria di Dennett, giungendo contro Dennett stesso alla conclusione che non solo il *Self* non è una *res cogitans*, ma che la proprietà essenziale tradizionalmente attribuita a quest'ultima, ossia la capacità di agire di propria libera volontà, è in realtà un'illusione. Un'illusione utilissima e inevitabile; e tuttavia un'illusione (ho sostanzialmente accettato questa concezione di Wegner sia in Nannini (2007a, pp. 135-158) sia *supra* capitolo secondo).

cerebrale, poiché da un punto di vista neurologico il *Self* è riconducibile ad un insieme piuttosto frammentario e solo parzialmente coerente di processi cerebrali che si sviluppano in parallelo e sono privi di un vero e proprio centro di coordinamento; ma, dall'altro, ciò non corrisponde all'immagine che ciascuno di noi ha di sé. Nessuno di noi sarebbe capace di vivere e agire normalmente se non si sentisse un individuo. Perché l'immagine che ho di me funziona così bene ed è anzi indispensabile alla mia salute mentale, se essa è errata ed illusoria? Dennett stesso avverte il problema quando sostiene che il nostro essere dei *Self* fittizi deve essere nondimeno sufficiente per renderci liberi e responsabili delle nostre azioni. Ma la soluzione che egli dà di tale problema soffre della debolezza di cui soffrono tutte le concezioni compatibilistiche del libero arbitrio: è proprio vero che, per non dubitare della veridicità di quel sentimento di libertà che accompagna (e rende possibili, sembra) le mie azioni volontarie, mi basta sapere che le loro cause sono interne al mio cervello, anche se di fatto, dato che ogni effetto segue necessariamente alla propria causa e le cause interne delle mie azioni sono a loro volta determinate da cause esterne completamente fuori dal mio controllo, so bene che non avrei potuto fare nulla di diverso da ciò che ho fatto? Se il *Self* è fittizio, anche la capacità di agire liberamente è illusoria! Più in generale: come faccio ad essere l'individuo libero e capace di azioni volontarie che mi sento di essere e che sono anche per gli altri e con gli altri, se i processi cerebrali che mi implementano sono invece, da un lato, causalmente determinati e, dall'altro, frammentari e privi di quell'unità che io mi sento di avere e dimostro di avere come autore di un comportamento al tempo stesso coerente ed imprevedibile?

3. L'«Io-penso» di Kant

Come uscire da questa *impasse*? È qui che un ritorno a Kant può essere oggi utile al filosofo naturalista per trovare una mediazione tra il considerare il *Self* come reale o come fittizio. Questo problema assomiglia infatti per certi versi a quello che Kant affrontava nella «Deduzione trascendentale delle Categorie», quando mediante il concetto di «Io penso»⁴² (o «appercezione trascendentale» che dir si

⁴² Ovviamente il kantiano *Das: ich denke* è traduzione tedesca del cartesiano *Cogito (ergo sum)*.

voglia) (Kant 1787, p. 110) cercava di aprirsi una terza via tra le concezioni della mente come cosa (*res*) pensante o come mucchio (*heap*) di *perceptions* (ossia di stati di coscienza) proposte rispettivamente da Cartesio e da Hume⁴³. L'appercezione trascendentale non era per Kant, infatti, né un'unica sostanza né una dispersa collezione di stati di coscienza; era piuttosto quell'attività di unificazione dell'intelletto che, agendo in base ad un principio, dà unità alla collezione stessa dei miei stati di coscienza rendendoli costitutivi di un unico mondo fenomenico. Più precisamente l'appercezione trascendentale è l'attività del mio intelletto in quanto capacità di costruire un'immagine coerente del mondo fenomenico riportando ad unità il molteplice dell'intuizione sensibile. Essa non deve essere perciò confusa con la variegata coscienza empirica che accompagna le singole intuizioni e percezioni e che mediante il «senso interno» mi offre anche un'immagine di me stesso sotto la forma della «unità *soggettiva* della coscienza» (Kant 1787, p. 114; corsivo di Kant].

Sebbene secondo Kant – come secondo Cartesio del resto – non possa esservi coscienza senza autocoscienza e quindi io non possa mai essere cosciente di qualcosa se non sono cosciente anche del fatto che sono io ad essere cosciente di quel qualcosa, tuttavia vi è una grande differenza in Kant tra l'appercezione trascendentale e la coscienza empirica dei singoli stati di coscienza. Mentre ogni rappresentazione cosciente empirica è diversa dalle altre in virtù del suo contenuto, la consapevolezza che tutto ciò che provo o penso è un mio stato di coscienza – una consapevolezza che accompagna tutte le

⁴³ Occorre qui fare una precisazione: la deduzione trascendentale delle categorie, soprattutto nella versione della seconda edizione della *Kritik der reinen Vernunft* che sarà qui esaminata, non è – com'è noto – l'esposizione di una teoria descrittiva ed empirica (diremmo oggi una teoria psicologico-cognitiva) sulla natura dell'autocoscienza; è piuttosto il punto focale di una teoria normativa e trascendentale della conoscenza, mediante la quale Kant intende fondare e giustificare (non descrivere o spiegare!) l'applicabilità dei concetti puri dell'intelletto ai fenomeni. Non c'è dubbio pertanto che l'uso che qui farò della distinzione tra unità analitica e unità sintetica dell'appercezione per formulare una teoria naturalistico-fisicalistica dell'Io è sicuramente contrario allo spirito di fondo della filosofia di Kant. Tuttavia, una volta fatta questa precisazione, sarebbe stupido non utilizzare gli spunti preziosi che un filosofo della mente di oggi può trovare nei testi kantiani solo perché essi appartengono ad un orizzonte di pensiero diverso da quello a cui sono nondimeno adattabili.

intuizioni e percezioni sia del senso interno sia del senso esterno – è sempre uguale a se stessa e proprio perciò, in quanto «rappresentazione [*Vorstellung*]»⁴⁴ *a priori*, riesce a riportare all'unità di una singola autocoscienza tutti gli stati di coscienza da essa accompagnati. In altre parole, mentre ogni rappresentazione empirica è diversa dalle altre in virtù del suo contenuto, l'Io-penso – in quanto «atto della spontaneità» del mio intelletto, «appercezione pura» e «originaria» (Kant 1787, p. 111) che rende a me coscienti tutte le altre mie rappresentazioni empiriche riportandole entro un'unica autocoscienza – è una rappresentazione sempre identica a sé e costituisce ciò che tutte le altre mie rappresentazioni hanno in comune, o meglio è ciò che fa sì che quelle altre rappresentazioni, essendo ricondotte dall'Io-penso all'unità di una autocoscienza, divengano le mie rappresentazioni.

Ora, sebbene le densissime pagine in cui Kant definisce l'appercezione trascendentale non siano di chiarezza cristallina, tuttavia è noto che egli individua due lati dell'appercezione trascendentale stessa, due lati che egli distingue come «unità sintetica» e «unità analitica dell'appercezione» (Kant 1787, p. 111). In altre parole l'Io-penso è, per un verso, la capacità di unificare secondo regole *a priori* le rappresentazioni empiriche fino a riportarle all'unità di un'autocoscienza, e, per un altro, è un marchio di fabbrica che, presente nel contenuto di ognuna delle rappresentazioni ricondotte ad unità, attesta per così dire l'identità dell'autore del processo di unificazione:

Infatti la coscienza empirica, che accompagna diverse rappresentazioni, è in sé dispersa e senza relazione con l'identità del soggetto. Questa relazione dunque non ha luogo ancora per ciò che io accompagno colla coscienza ciascuna delle rappresentazioni, ma perché le *compongo* tutte l'una con l'altra, e sono consapevole della loro sintesi. Solo perciò, in quanto posso legare in *una coscienza* una molteplicità di rappresentazioni date, è possibile che io mi rappresenti *l'identità della coscienza in queste rappresentazioni stesse*; cioè, l'unità *analitica* dell'appercezione è possibile solo a patto che si presupponga una unità *sintetica* (Kant 1787, p. 111; corsivi di Kant).

⁴⁴ *Vorstellung* è in Kant un'espressione generica applicabile a qualsiasi stato di coscienza (Kant 1787, p. 250) ed è equivalente perciò al termine *idea* in Locke o *perception* in Hume.

Nella nota successiva Kant chiarisce – ma di nuovo secondo il suo stile severo ed accademico, non molto *friendly* nei confronti del lettore – in che senso ogni unità analitica presupponga sempre una unità sintetica. Rendiamo appena un po' più esplicita l'argomentazione fornitaci da Kant in questa nota. Vedo ad esempio una mela e percepisco che è rossa, succosa e di forma sferica. Combino queste «note» (*Merkmale*) puramente percettive – rosso-qui-ora, sferico-qui-ora ecc., ossia proprietà particolari, “tropi” come si dice oggi – secondo la regola di unificazione fornitami dal concetto puro dell'intelletto (o categoria che dir si voglia) di “sostanza e accidente”. Ottengo così il concetto empirico di mela. Ogni mela, in quanto sostanza individuale, è ridotta ad essere una collezione di alcune sue note. Compio operazioni analoghe nei confronti di altri concetti empirici di altri corpi rossi: il cinabro, le stoffe purpuree ecc. Poi, confrontando tali concetti empirici, rilevo che contengono tutti alcune note simili. È a questo punto che, mediante un procedimento di astrazione e generalizzazione guidato di nuovo dal concetto puro di “sostanza e accidente”, costruisco a partire da note percettive particolari del tipo ‘rosso-qui-ora’ il concetto empirico della proprietà “rosso in generale”. Kant sottolinea con chiarezza che l'analisi dei concetti empirici delle cose rosse, analisi mediante la quale si riesce a formare il concetto della proprietà generale “essere rosso”, non sarebbe possibile se l'intelletto non avesse prima costruito, mediante processi di sintesi che richiedono la categoria di sostanza e accidente, alcuni concetti empirici di cose rosse.

In modo analogo – se si proietta l'argomentazione della nota sul testo principale – io riconosco ognuno degli stati mentali di cui sono cosciente come un mio stato di coscienza, e quindi ritrovo nel suo contenuto una rappresentazione di me (unità analitica dell'appercezione), solo perché sono stato io (in quanto unità sintetica) a unificare quelle rappresentazioni in un'unica autocoscienza. L'unità analitica è il marchio di fabbrica che rinvia ad una unità sintetica in quanto processo di produzione. Quindi l'Io-penso è sia la rappresentazione di una certa attività dell'intelletto (unità analitica) sia l'attività stessa che produce quella rappresentazione e ne giustifica la veridicità (unità sintetica).

Ora, com'è noto, questa concezione kantiana dell'Io come attività – come un fare, un agire, non un essere – ha conosciuto una fortuna enorme nell'idealismo tedesco, o meglio per certi aspetti ne è stata

addirittura all'origine. Tuttavia, se si pone per un momento in parentesi la storia della fortuna del pensiero di Kant e si guarda solo, in termini teoretici e non storici, alla distinzione tra unità analitica e unità sintetica dell'appercezione, si vede bene che tale distinzione offre una terza via non solo tra dei filosofi precedenti a Kant come Cartesio e Hume, ma anche tra dei filosofi contemporanei come Patricia Churchland e Daniel Dennett. Se infatti paragoniamo l'unità sintetica dell'appercezione all'insieme dei processi cerebrali che, come sostengono oggi i "costruttivisti" in psicologia e neuroscienze (ad es. Roth 2001), costruiscono le mappe cerebrali del mondo esterno e del proprio corpo, vediamo subito che l'unità analitica può essere paragonata, viceversa, a quel sentimento di identità e appartenenza che, essendo un prodotto di quei processi cerebrali stessi, accompagna le nostre percezioni coscienti e le nostre azioni volontarie e consente di combinarle in un'unica scena coerente.

Pertanto, se si applica la lezione di Kant al dibattito odierno sulla realtà o meno del *Self* e la si combina con l'introduzione di una distinzione concettuale e terminologica tra il *Self* stesso e l'Io (o *Ego* in inglese)⁴⁵, si può paragonare il *Self* all'unità sintetica dell'appercezione e l'Io all'unità analitica. In altre parole si può intendere per *Self*, sulla scia di Edelman (2004), una descrizione in termini funzionali di quella *higher order property* della dinamica cerebrale che, conferendo unità e coerenza all'attività del cervello (così come l'unità sintetica dell'appercezione la conferisce alla nostra rappresentazione del mondo fenomenico), permette al cervello stesso di guidare in modo flessibile ed efficace l'interazione di tutto il corpo con il mondo esterno a partire dal coordinamento senso-motorio. L'Io può essere inteso invece, paragonandolo all'unità analitica dell'appercezione, come il contenuto della rappresentazione idealizzata che il cervello si dà, nel "formato" della coscienza, del suo stesso *Self* (*supra* capitolo secondo).

Di conseguenza, mentre il *Self*, essendo una ridecrizione in termini funzionali di una proprietà della dinamica cerebrale, gode di realtà virtuale, l'Io invece, essendo in quanto *Io** solo il contenuto di una rappresentazione idealizzata di quella stessa attività nel formato della coscienza, è un *fictum*; e quindi è una mera finzione priva di

⁴⁵ Riprendo da Michele Di Francesco (1998) la distinzione tra Io e *Self*, ma la utilizzo con un intento "eliminativistico" molto lontano dal suo.

effettivi poteri causali (*supra* capitolo terzo). L'Io insomma è un oggetto fittizio al pari, ad esempio, di Pegaso. Il mio pensare a Pegaso è qualcosa di reale ed è identico ad una certa attività del mio cervello; ma ciò non significa che Pegaso stesso sia reale! In modo analogo il *Self*, in quanto funzione dell'attività cerebrale, è reale, o meglio gode di una realtà virtuale implementata da una proprietà sistemica della dinamica cerebrale. l'Io invece, in quanto *Io**, ossia in quanto contenuto della rappresentazione idealizzata che il cervello si dà del suo stesso *Self*, è solo una – pur utilissima – finzione.

In conclusione mediante la distinzione tra *Self* ed Io si può uscire dall'*impasse* che sembra scaturire dalla apparente incapacità del naturalismo cognitivo di spiegare come si possa conferire unità e realtà all'Io, in accordo con il senso comune, e al tempo stesso identificarlo, come suggeriscono le neuroscienze, con un *Self* che consiste in un insieme di processi cerebrali privi di altrettanta unità e coerenza.

4. *L'Io come Direttore Assente*

Volendo precisare ulteriormente la distinzione tra *Self* ed Io si può dire anche che il cervello, con la sua attività parallela e distribuita, è paragonabile ad un'orchestra che suoni a podio vuoto, quando si immagini inoltre che ogni orchestrale, osservando chi il primo violino chi il compagno di fila o qualche altro orchestrale, si costruisca la finzione di un Direttore in realtà Assente e che solo grazie a tali finzioni gli orchestrali riescano a suonare all'unisono. In tal caso il Direttore Assente sarebbe sì solo una finzione implementata da processi di coordinazione tra gli orchestrali in effetti frammentari e distribuiti, ma l'usare tale finzione da parte degli orchestrali svolgerebbe nondimeno una funzione essenziale, senza la quale l'orchestra non potrebbe suonare correttamente. È troppo azzardato pensare che qualcosa di analogo avvenga nel mio cervello e che l'Io del quale sono cosciente e con il quale mi identifico in quanto autore dei miei atti liberi e volontari sia il contenuto (in larga misura idealizzato e in parte illusorio) di una rappresentazione che il mio cervello dà a se stesso di quella attività di coordinamento dei suoi vari processi che costituisce il suo *Self*? Una rappresentazione che è sì ingannevole – visto che tale autore in effetti non esiste (dunque io, a rigore, non esisto!) e che la reale attività di auto-coordinazione del cervello è frammentaria e distribuita, come

pensa Dennett – ma che nondimeno, dato che il contenuto di tale rappresentazione è parte del contenuto di tutti i miei stati di coscienza, costituisce un attrattore, definibile in termini matematico-vettoriali, che dà alla dinamica complessiva dei correlati neuronali di tali stati di coscienza unità e coerenza. La teoria matematica e fisica dei sistemi complessi potrebbe darci qualche indicazione preziosa per ricostruire la dinamica complessiva dei processi cerebrali: forse in essa potremmo ritrovare anche l'implementazione neurologica della rappresentazione dell'Io in quanto Direttore Assente⁴⁶.

Questa ipotesi, che radicalizza la tesi dennettiana secondo la quale ciò che Dennett chiama il *Self* (e che qui invece viene inteso più precisamente come l'Io) è una finzione, presenta certo enormi differenze rispetto alla concezione kantiana dell'appercezione trascendentale. In primo luogo quest'ultima, in quanto unità sintetica, è sì per Kant un'attività, ma dell'intelletto, non del cervello! In secondo luogo l'Io-penso, in quanto unità analitica dell'appercezione, ossia in quanto rappresentazione di me che ritrovo in ogni mia rappresentazione empirica, è secondo Kant una rappresentazione per nulla ingannevole e anzi assolutamente veritiera dell'attività di unificazione operata dall'intelletto. Tuttavia, concesse queste grandi differenze, resta vero che Kant fornisce alle odierne scienze cognitive un duplice spunto: in primo luogo il *Self*, in quanto attività di sintesi che costruisce un'immagine coerente del mondo, costruisce anche un'immagine di se stesso (ossia costruisce l'Io) e in questa immagine si rappresenta come un soggetto agente; in secondo luogo occorre distinguere il *Self* (= unità sintetica dell'appercezione) dall'Io (= unità analitica dell'appercezione).

A differenza di Kant possiamo ritenere tuttavia che tale auto-rappresentazione possa essere ingannevole e che non necessariamente un pensiero debba essere il pensiero di qualcuno, perché in primo luogo – come dice Edelman citando James - «a pensare sono i pensieri stessi» (Edelman 2004, p. 111) e in secondo luogo perché, come sosteneva anche B. Russell, il verbo “pensare” può essere considerato un verbo impersonale al pari di “piovere” (Russell 1921, p. 23). Se si accettano queste due premesse, allora si può concludere,

⁴⁶ Un modello della dinamica non lineare del cervello è stato sviluppato da Izhikevich (2007).

contro Cartesio, che “Penso, dunque sono” va trasformato in “Circola pensiero nel mio cervello (ad esempio mediante la comparsa di processi di sincronizzazione tra circuiti cerebrali oscillanti); e ciò crea nel cervello, ad opera del cervello e per il cervello (ossia per me) un ulteriore pensiero (ossia un altro fenomeno di sincronizzazione cerebrale) il cui ingannevole contenuto (quando tale fenomeno cerebrale assuma mediante la sincronizzazione il formato della coscienza fenomenica) è che esista un autore unico di tutti i miei pensieri; ossia il mio cervello stesso crea l'illusione che io esista in quanto entità distinguibile dal complesso dell'attività cerebrale” (cfr. *supra* capitolo quarto).

Se è così, allora esiste davvero il Genio Maligno immaginato da Cartesio: è il cervello stesso il Genio Maligno che si auto-inganna e quindi, auto-ingannandosi, inganna anche me, che sono parte del suo auto-inganno! Solo che questo Genio per un verso è molto più potente di quanto Cartesio riteneva possibile che fosse, perché il suo inganno più grande è il *Cogito* cartesiano stesso, ossia è il far credere che nessun inganno sia possibile a meno che non esista qualcuno che è ingannato. In altre parole il cervello costruisce mediante la sua dinamica un modello idealizzato e semplificato della sua stessa attività. In questo modello tale attività viene rappresentata, nel formato della coscienza fenomenica, come l'opera di un Soggetto Agente. Ma in effetti tale Soggetto non è né l'Autore della rappresentazione messa in scena dal cervello né il Pubblico che ad essa assisterebbe (se la rappresentazione avesse luogo in un Teatro Cartesiano). In realtà l'Autore, il Pubblico e il Teatro stesso non esistono. La messa in scena di questa commedia pirandelliana senza autore è un prodotto del funzionamento automatico del cervello. Ed il Soggetto che sembra esserne l'Autore è solo, al pari del Pubblico con il quale pretende di coincidere, un personaggio immaginario interno alla rappresentazione stessa; un personaggio le cui proprietà essenziali – soprattutto la sua pretesa di essere un agente libero ed un'unica mente cosciente – non trovano alcun correlato nei processi cerebrali che implementano la rappresentazione di cui fa parte.

Ma, per altro verso, questo potentissimo Genio che è il cervello non è affatto Maligno, perché l'illusione che esso crea, sebbene sia ingannevole, è probabilmente utilissima per la sopravvivenza del corpo di cui fa parte (conoscere più nel dettaglio perché questo

inganno sia così utile è un compito però che solo la teoria dell'evoluzione biologica potrà forse assolvere). Comunque sia io sono solo una finzione del mio cervello, ma – come dice Dennett – si tratta di una finzione magnifica!

Capitolo ottavo

Freud e il naturalismo cognitivo

1. *L'inconscio freudiano e l'Intenzionalità brentaniana*

La psicoanalisi ha sempre goduto di cattiva stampa presso i filosofi della mente di orientamento analitico e naturalistico-cognitivo (Grünbaum 1984). Il concetto di inconscio in particolare è sempre apparso molto sospetto. Ha giocato contro di esso, da un lato, il retaggio della identificazione cartesiana tra pensiero e coscienza fatta propria da Hume e, attraverso di lui, da buona parte degli psicologi cognitivi e dei filosofi della mente anglofoni. Oggi però da alcuni anni a questa parte le cose sono cambiate. Da un lato l'inconscio, sia pure sotto la nuova distinzione tra “inconscio psicoanalitico” e “inconscio cognitivo” (Migone 2007), è divenuto oggetto di studio da parte degli scienziati cognitivi, in generale, e dei neuroscienziati in particolare (Hassin, Uleman & Bargh 2006). Dall'altro si sono moltiplicati i tentativi di connettere la psicoanalisi alle neuroscienze (Mancia 2006), come del resto S. Freud stesso aveva cercato di fare – entro i limiti, ovviamente, delle conoscenze del suo tempo sull'attività del cervello – nel suo *Entwurf der Psychologie* (Freud 1895).

Freud infatti – com'è noto – nel definire gli stati mentali inconsci come quegli stati che, pur non essendo coscienti, sono propriamente mentali (e non puramente fisiologici), spezzò l'identificazione cartesiana tra mente e coscienza. Poté farlo in primo luogo perché individuò l'essenza degli stati mentali non nell'essere stati di coscienza, bensì nell'essere stati dotati di Intenzionalità nel senso brentaniano del termine, e in secondo luogo perché, distaccandosi invece in ciò dal pensiero di F. Brentano, collocò il concetto di Intenzionalità entro una concezione complessiva del mentale almeno tendenzialmente naturalistica e materialistica.

Ora è proprio in questa duplice chiave che il concetto brentaniano di Intenzionalità è stato utilizzato nella seconda metà del XX secolo da alcuni dei più noti filosofi della mente per naturalizzare il mentale (cfr. *supra* capitolo primo). In tal modo essi, tramite il ricorso al concetto di Intenzionalità, hanno coniugato il rifiuto del dualismo mente-corpo con il rigetto, altrettanto netto sebbene implicito, di ogni riduzione cartesiana della mente alla sola coscienza. In una concezione naturalistica del mentale, infatti, l'Intenzionalità trova la

propria origine nel coordinamento senso-motorio ed è riducibile ad una proprietà di primo ordine della dinamica cerebrale, mentre la coscienza fenomenica può essere a sua volta naturalizzata solo se viene interpretata come una ulteriore proprietà di secondo ordine della medesima dinamica cerebrale, una proprietà che dà ad essa unità e coerenza (cfr. *supra* capitolo secondo).

Resta qualcosa della lezione di Freud nella odierna concezione naturalistica del mentale? In un certo senso sì. Ovviamente il concetto freudiano di inconscio come prodotto di processi di rimozione va perduto. E tuttavia resta l'idea che, per giungere ad una concezione scientifica della mente, si debba distinguere in essa la coscienza dall'Intenzionalità e considerare quest'ultima, non la prima, come il fondamento del mentale, aprendo con ciò la strada alla sua naturalizzazione. Vediamo più nel dettaglio come possa essere articolato questo confronto teorico tra Freud e l'odierno naturalismo cognitivo.

2. *Coscienza e Intenzionalità in Cartesio e F. Brentano*

Per chiarire la genesi del concetto di inconscio in Freud conviene partire da una brevissima ricostruzione del concetto di mente nella storia del pensiero filosofico. Già nell'antichità Democrito, Platone e Aristotele si ponevano la domanda se ciascuno di noi sia solo un corpo o abbia anche un'anima. E, se abbiamo un'anima, essa che cosa è? Di che cosa è fatta? Da allora queste domande non hanno mai cessato di riproporsi. Fra le risposte date ad esse in epoca moderna la più comunemente accettata – anche da parte di pensatori che avevano visioni del mondo diverse e sostenevano teorie epistemologiche contrapposte come l'innatismo e l'empirismo – è quella che risale a Cartesio, il quale identificava la mente con la coscienza: gli stati mentali sono stati di coscienza (Cartesio 1644, p. 25). Reagendo all'aristotelismo della scolastica medievale, Cartesio riteneva che fosse necessario sostituire il concetto di anima con quello di mente (Cartesio 1641, p. 343). Mentre l'anima secondo Aristotele serviva fra l'altro a tenere in vita il corpo del quale costituiva la forma, Cartesio invece pensava che il corpo vivente fosse un meccanismo autosufficiente, cioè che la vita fosse spiegabile interamente in termini meccanici e che gli animali fossero solo delle sofisticatissime macchine costruite da quell'ingegnere e matematico onnisciente che è Dio (Cartesio 1637, pp. 327-329). La mente, posseduta unicamente

dagli uomini e non dagli animali, era invece secondo Cartesio una sostanza immateriale diversa dal corpo capace, per sua natura, di pensare. O meglio, la mente era coincidente con la coscienza e l'autocoscienza, ossia con quella consapevolezza di sé che accompagna, oltre a ogni nostro pensiero, anche la più insignificante delle nostre sensazioni di dolore o di piacere e che è invece incompatibile con la proprietà essenziale di ogni corpo: l'estensione (l'occupare cioè un certo spazio). Avere uno stato mentale voleva dire per Cartesio trovarsi in un certo stato di coscienza e autocoscienza. Di conseguenza gli stati mentali accompagnati da inconsapevolezza e involontarietà (la cui esistenza certo nessuno ignorava anche prima di Freud) venivano assimilati tutti sia da Cartesio sia da moltissimi filosofi e psicologi a lui successivi a stati fisiologici. In questa prospettiva cartesiana ogni essere umano è una macchina (un corpo esteso) più una coscienza (una mente immateriale). Pertanto tutto ciò che non appartiene alla coscienza appartiene necessariamente alla macchina. Neppure Cartesio dubitava del fatto che spesso noi esseri umani facciamo molte cose in modo automatico, involontario, senza pensarci. Ma, in questo caso – diceva Cartesio – noi agiamo come se fossimo un mero meccanismo; non siamo noi che agiamo, è il nostro corpo che si muove secondo leggi di natura. Le nostre braccia, ad esempio, si muovono in occasione di un atto involontario come possono muoversi le lancette di un orologio. In conclusione non c'è posto all'interno della concezione cartesiana della mente per degli stati inconsci che siano propriamente mentali, cioè non meramente fisiologici e quindi non spiegabili in termini puramente meccanici.

L'identificazione cartesiana tra mente e coscienza ha avuto un successo enorme tra tutti i filosofi e gli psicologi successivi ed è ancora largamente accettata, a quasi quattro secoli dalla sua formulazione, anche da molti di coloro che, riguardo al rapporto mente-corpo, difendono invece teorie contrapposte al dualismo cartesiano tra *res cogitans* e *res extensa*. È qui, nella vitalità della tradizione cartesiana – filtrata in molti casi dall'empirismo di Locke e di Hume –, che possiamo trovare una delle ragioni dell'estraneità e talvolta ostilità dei filosofi analitici contemporanei verso il pensiero di Freud, in generale, e verso il suo concetto d'inconscio in particolare⁴⁷.

⁴⁷ Ciò non significa ovviamente che il concetto di inconscio non si sia storicamente

Esiste tuttavia nella storia del pensiero filosofico e psicologico moderno un'altra concezione del mentale, secondo la quale – come già è stato detto nei capitoli precedenti – gli stati mentali sono stati Intenzionali, ossia sono stati che hanno un significato (o contenuto che dir si voglia) e si riferiscono intrinsecamente ad altro da sé. La grande differenza tra quest'ultima concezione del mentale e quella cartesiana non sempre è stata vista con chiarezza. E anzi molti filosofi (incluso ad esempio B. Russell 1921, p. 288) hanno ritenuto che le due concezioni fossero coincidenti, poiché tutti gli stati di coscienza sarebbero anche Intenzionali e viceversa. Ma, in realtà, non ci sono buone ragioni per credere in questa coincidenza della coscienza con l'Intenzionalità. Vediamo perché.

La teoria che identifica gli stati mentali con gli stati dotati d'Intenzionalità risale al filosofo e psicologo austriaco della seconda metà dell'Ottocento Franz Brentano (1874). Questi riprendeva, per caratterizzare il mentale e distinguerlo dal fisico, un concetto della scolastica medievale, il concetto di *intentio* (che aveva un significato molto più largo di quello che ha oggi la parola “intenzione” nel linguaggio ordinario). La *intentio* di uno stato mentale era secondo Brentano l'oggetto (o stato di cose o evento o proprietà ecc.) a cui tale stato si riferisce così come il significato di una parola è ciò a cui mi riferisco quando la pronuncio. Più precisamente Brentano era in cerca, mediante il recupero del concetto scolastico di *intentio*, di un criterio capace di distinguere il mentale dal fisico che fosse indipendente dalla distinzione cartesiana tra *res cogitans* e *res extensa* (anche se non necessariamente in contrasto con essa).

Gli stati mentali – pensava Brentano – non sono riducibili a stati fisici; ma che cosa è che li contraddistingue? Quale proprietà essi possiedono, sempre e necessariamente, che nessuno stato o oggetto fisico può possedere? Questa proprietà – rispondeva Brentano – è appunto l'Intenzionalità, ossia il fatto che ogni stato mentale è necessariamente diretto verso un oggetto (concreto o astratto che sia). Gli stati fisici sono invece sempre e necessariamente non Intenzionali.

Inoltre il fatto che nessuno stato fisico possa essere Intenzionale provava – secondo Brentano – che nessuno stato mentale è riducibile

formato ben prima di Freud. Cfr. ad esempio i classici studi di Whyte (1960) e Ellenberger (1970), nonché Rand (2004).

ad un qualsiasi stato fisico (ad esempio, ad uno stato cerebrale). E in effetti ancor oggi uno dei problemi fondamentali che i fisicalisti devono affrontare è provare che le proprietà Intenzionali degli stati mentali sono traducibili in termini fisici (in primo luogo in termini cerebrali). Il fisicalista, in altre parole, è impegnato a mostrare come dei sistemi fisici, purché sufficientemente complessi come lo è il cervello umano nella sua interazione con il resto del corpo e con il mondo esterno, possano acquisire Intenzionalità brentaniana (o meglio come possano simulare, mediante processi puramente fisici, tale acquisizione). Se questo non fosse possibile, Brentano, insieme a tutti i dualisti, avrebbe ragione e il mentale non sarebbe riducibile al fisico.

3. Intenzionalità e naturalismo in Freud

Se ci chiediamo ora quale fra queste due concezioni del mentale – quella cartesiana secondo la quale mente e coscienza coincidono e quella di Brentano secondo la quale gli stati mentali sono stati Intenzionali – sia stata fatta propria da Freud, è facile rispondere che egli accettava senz'altro la seconda. La prima non poteva accettarla di sicuro, perché in essa, come abbiamo visto, non c'è spazio per distinguere tra ciò che è propriamente mentale e ciò che è conscio. Pertanto in essa non c'è spazio per un inconscio che non sia semplicemente qualcosa di puramente fisiologico: vale a dire non c'è spazio per l'inconscio nel senso propriamente freudiano.

Del resto Brentano fu professore di Freud all'Università di Vienna ed ebbe una qualche influenza su di lui, tant'è vero che ad un certo punto perfino le convinzioni atee che Freud aveva abbracciate fin dall'adolescenza e che poi lo accompagneranno con rinnovato vigore per tutta la vita conobbero un momento di incertezza. Per alcuni anni il giovane Sigmund pensò che anche il deismo potesse essere preso in seria considerazione, ma poi, com'è noto, Freud dimenticò in fretta la lezione del suo vecchio professore di filosofia e tornò ben presto ad essere ateo (Gay 1988, p. 27).

La concezione che del mentale aveva Brentano come regno dell'Intenzionalità subì tuttavia una sorte ben diversa dal teismo nella testa del suo giovane allievo. Essa rimase infatti sempre presente in Freud, scavò a fondo nella sua mente e finì per divenire parte integrante delle sue idee anche in età matura, come prova ad esempio questo passo tratto dalle *Vorlesungen zur Einführung in die Psychoa-*

nalyse pubblicate nel 1917; un passo appartenente ad un gruppo di pagine nel quale Freud, appoggiandosi su esempi forniti dai lapsus verbali, sta riassumendo l'essenziale delle tesi sostenute in *Zur Psychopathologie des Alltagslebens* (Freud 1901):

Abbiamo esaminato le condizioni più generali nelle quali un lapsus verbale [*ein Versprechen*] ha luogo, quindi le influenze che determinano il tipo di deformazione provocata dal lapsus, ma finora non abbiamo preso in considerazione l'effetto del lapsus considerato in sé stesso, indipendentemente dalla sua origine. Se ci decidiamo a far questo, dobbiamo finalmente trovare il coraggio di dire che in alcuni degli esempi anche il risultato del lapsus ha un senso [*Sinn*]. Che cosa vuol dire "ha un senso"? Ebbene, vuol dire che l'effetto del lapsus ha diritto a essere considerato un atto psichico [*psychischer Akt*] pienamente valido, perseguente un proprio fine [*Ziel*], espressione di un contenuto e di un significato [*eine Äusserung von Inhalt und Bedeutung*] (Freud 1917, p. 35).

Ora, che il lapsus sia un atto sì inconscio, ma propriamente psichico (non meramente fisiologico, come ad esempio la digestione) e che sia tale giustappunto perché ha un fine, un contenuto e un significato è comprensibile solo se si dà una definizione del mentale di tipo brentaniano e non cartesiano: l'atto inconscio è psichico, in senso proprio e non meramente metaforico, perché è un atto Intenzionale. Se Freud avesse accettato l'identificazione cartesiana tra mente e coscienza, le affermazioni sopra citate non avrebbero senso: come potrebbe un atto inconscio (ossia, in termini cartesiani, non mentale e puramente fisiologico) avere un fine, un contenuto e un significato? Ma, se non li avesse, un lapsus non sarebbe più ciò che Freud pretendeva che fosse, ossia la manifestazione di un desiderio inconscio. Un lapsus privo di significato sarebbe semplicemente un intoppo nel funzionamento puramente fisico del cervello (come un lampo sullo schermo del televisore quando c'è uno sbalzo di tensione). Senza la concezione brentaniana del mentale, addio psicoanalisi! Mi sembra evidente perciò che, se dobbiamo collocare Freud rispetto alla grande distinzione filosofica fra il mentale inteso come coscienza e il mentale inteso come Intenzionalità, Freud appartiene sicuramente a questa seconda corrente.

Freud dunque era un Intenzionalista al pari di Brentano. Ma era perciò anche un dualista come il suo vecchio professore di filosofia riguardo alla relazione tra mente e corpo? Direi proprio di no, almeno

se è lecito confrontare la meta-psicologia esposta da Freud nel già menzionato *Entwurf einer Psychologie* (un testo giovanile scritto nel 1895 ma pubblicato postumo solo nel 1950) con le soluzioni date al problema mente-corpo che, come già è stato detto, dominano l'intero pensiero filosofico occidentale e sono ancor oggi largamente diffuse: dualismo, fisicalismo (o materialismo) e funzionalismo (cfr. *supra* capitolo secondo). Se questo confronto è lecito – e non si vede perché non dovrebbe esserlo –, allora è sensato collocare Freud (quanto meno il Freud autore dell'*Entwurf*) tra i materialisti. Freud sosteneva infatti in quello scritto che si dovesse costruire una psicologia scientifica, intesa come una scienza naturale accanto alla fisica e alla biologia, e che si dovesse cercare una traduzione dei concetti psicologici in termini neurologicamente plausibili. Si legga al riguardo l'inizio di quel testo:

L'intento di questo progetto è di dare una psicologia che sia una *scienza naturale*, ossia di rappresentare i processi psichici come stati quantitativamente determinati di *particelle materiali* identificabili, al fine di renderli chiari e incontestabili (Freud 1895, p. 11, corsivi miei).

In conclusione, se Freud avesse potuto conoscere, quando scriveva l'*Entwurf*, la teoria dell'identità di U. Place o l'eliminativismo dei Churchland vi si sarebbe sicuramente in gran parte riconosciuto. Inoltre, se si guarda non solo all'*Entwurf* del 1895 ma anche a molti altri scritti successivi, è facile notare che Freud si è sempre opposto (anche se più o meno esplicitamente) al dualismo mente-corpo. E non poteva non opporvisi perché, tra l'altro, era un determinista convinto e quindi non credeva, a differenza di quanto pensano in genere i dualisti, nell'esistenza del libero arbitrio. Ad esempio, a proposito di coloro che pensano che i lapsus siano eventi accidentali e che quindi sia errato spiegarli rintracciandone la causa nell'inconscio, Freud notava nelle già menzionate *Vorlesungen zur Einführung in die Psychoanalyse* che «chi spezza così il determinismo naturale in un singolo punto manda all'aria l'intera concezione scientifica del mondo» (Freud 1917, p. 29).

Tuttavia occorre riconoscere che, partito in gioventù da una evidente accettazione del materialismo, Freud approdò nei suoi scritti più maturi ad un naturalismo che oggi giudicheremmo più

vicino al funzionalismo che al materialismo in senso stretto. Freud infatti si convinse progressivamente nel corso dello sviluppo del suo pensiero, forse perché scoraggiato dal basso grado di sviluppo delle neuroscienze del suo tempo, che la meta-psicologia dell'*Entwurf* almeno provvisoriamente doveva essere lasciata nel cassetto e che la psicoanalisi doveva andare avanti per proprio conto indipendentemente dagli studi di neurobiologia, venendo a formare un proprio repertorio di concetti chiaramente ed esclusivamente psicologici come ad esempio l'inconscio, il preconscious, la rimozione, l'Io, l'Es, il Super-Io ecc. Se risultasse che tutti gli atti mancati hanno un senso – scriveva Freud ad esempio di nuovo nelle *Vorlesungen* – potremmo «allora lasciar da parte tutti i fattori fisiologici e psicofisiologici e dedicarci a indagini puramente psicologiche sul senso, cioè sul significato, sull'intenzione, dell'atto mancato» (Freud 1917, p. 36).

Pertanto, se combiniamo quest'ultima affermazione con il rifiuto, mai messo in dubbio da Freud, del dualismo mente-corpo o di altre ipotesi idealistiche o comunque spiritualistiche, si può plausibilmente trarre la conclusione che anch'egli, al pari di Aristotele e dei funzionalisti contemporanei, era in cerca negli scritti della maturità di una terza via tra dualismo e materialismo; una terza via che consentisse lo studio scientifico della mente anche in assenza di progressi significativi nella conoscenza del funzionamento del cervello. E tuttavia non è neppure implausibile sostenere che, in realtà, Freud si sia adattato ad una scelta imposta dallo scarso sviluppo, ai suoi tempi, degli studi sul cervello, ma non abbia mai ripudiato, come concezione ontologica di base, il materialismo della giovinezza.

4. Freud e l'odierna filosofia della mente

L'odierna filosofia della mente diffida – come abbiamo detto – della scientificità della psicoanalisi; quindi non lascia molto spazio al concetto freudiano di inconscio e stenta a liberarsi della identificazione cartesiana tra mente e coscienza. Una spia di tale diffidenza può essere individuata nel fatto che persino filosofi che, pur in forme diverse, si dichiarano naturalisti difficilmente usano la parola “inconscio” quando vogliano distinguere gli stati di coscienza da altri stati mentali o da altri stati funzionali del cervello. Si veda ad esempio la distinzione introdotta da D. Chalmers (1996 e 2010) tra *consciousness* e *awareness* o quella, altrettanto fortunata, proposta da N. Block (1995 e 2007) tra *pheno-*

menal consciousness e access consciousness.

E tuttavia la somiglianza tra Freud e l'odierna filosofia della mente riemerge implicitamente su un altro piano quando si noti che molti filosofi contemporanei fondano, al pari di Freud, la loro concezione della mente e dell'Io su una naturalizzazione dell'Intenzionalità, ossia su una interpretazione dell'Intenzionalità che la renda compatibile con il naturalismo. Alcuni di loro anzi, come abbiamo già visto, pur evitando l'uso del concetto freudiano di inconscio, condividono con Freud una concezione del mentale chiaramente anticartesiana: si pensi a D.C. Dennett o ai coniugi Churchland.

Questi ultimi infatti si sono scontrati con coloro che, pur dichiarandosi naturalisti e utilizzando anch'essi il concetto brentaniano di Intenzionalità, concepiscono il mentale in termini sostanzialmente cartesiani. J. Searle ad esempio ha sostenuto in consonanza con il pensiero di Brentano e di Husserl che le rappresentazioni mentali, lungi dal godere di una Intenzionalità semplicemente derivata come quella delle parole che traggono il loro significato da convenzioni linguistiche, sono intrinsecamente Intenzionali. Ad esempio la parola "cane" si riferisce agli animali pelosi che abbaiano in virtù di una semplice convenzione. Ma il pensare ai cani, ossia il concetto di cane, si riferisce a quei certi animali in modo intrinseco ed essenziale. In linea di principio noi parlanti italiani potremmo decidere di chiamare "cani" gli animali che miagolano e "gatti" quelli che abbaiano. Ma, se pensiamo ai cani, non possiamo decidere che il nostro pensiero si riferisca ai gatti! Inoltre – ed è qui che Searle torna a riavvicinarsi a Cartesio e, come Brentano, pensa che tutti gli stati Intenzionali coincidano con degli stati di coscienza – uno stato mentale può godere di Intenzionalità intrinseca proprio perché o è direttamente uno stato di coscienza o quanto meno è uno stato fisiologico suscettibile di divenire cosciente (Searle 1992 e 1997).

Dennett nella sua replica a Searle ha invece rifiutato la distinzione tra Intenzionalità intrinseca e Intenzionalità derivata: come sono coloro che parlano italiano a stabilire una relazione d'Intenzionalità per convenzione tra la parola "cane" ed i cani, allo stesso modo è stata l'evoluzione biologica a far sì che, per ragioni di convenienza pratica, un certo stato del nostro cervello (da noi vissuto come un certo stato di coscienza) si riferisca convenzionalmente ad un certo stato di cose nell'ambiente circostante, pur non avendo nulla in comune con esso da un punto di vista fisico. L'Intenzionalità del mentale è derivata e riposa

implicitamente su convenzioni al pari dell'Intenzionalità del linguaggio (Dennett in Searle 1997, pp. 77-107).

È pertanto evidente che almeno la concezione del mentale di Dennett e in larga misura anche quella di molti altri naturalisti contemporanei, a partire dai Churchland, di fatto – indipendentemente dalle loro intenzioni – si incrocia con ciò che pensava Freud, allorché da un lato definiscono il mentale non mediante la coscienza bensì mediante l'Intenzionalità, ma dall'altro rifiutano il dualismo mente-corpo che Brentano invece faceva proprio.

5. Psicoanalisi e naturalismo cognitivo oggi

In conclusione, se è vero che i filosofi della mente di orientamento naturalistico farebbero bene a rivedere il loro atteggiamento di diffidenza verso la meta-psicologia di Freud, non è forse vero che anche gli psicoanalisti freudiani⁴⁸, per converso, dovrebbero scendere dalla loro torre d'avorio e sporcarsi le mani con le scienze cognitive (in particolare con la neuroscienza cognitiva) ed il connesso naturalismo filosofico? Io penso che dovrebbero (e ritengo anche – come ho già detto all'inizio – che fortunatamente alcuni di loro comincino a farlo davvero).

Agli psicoanalisti infatti, almeno a quelli che si sentono in larga parte eredi del pensiero di Freud, si aprono oggi strade alternative riguardo al rapporto mente-corpo. La prima è quella di coloro che, a mio avviso un po' testardamente, rimangono fedeli alla lettera delle opere di Freud, ma, in sostanza, vanno contro il loro spirito, perché hanno nei confronti degli sviluppi recenti delle scienze cognitive (e in particolare della neuroscienza cognitiva) un atteggiamento di diffidenza e rifiuto pregiudiziale esattamente opposto a quello che Freud aveva nei confronti della neurobiologia del suo tempo.

Il secondo atteggiamento è quello di chi sviluppa, sì, il pensiero di Freud in modo più libero e creativo, ma lo piega in una direzione opposta a quella concezione scientifica del mondo che Freud stesso ha sempre difesa e che lo avrebbe portato ad aborreire quel matrimonio della psicoanalisi con il misticismo e la religione che oggi molti vorrebbero invece celebrare (ad es. Fattori e Vandi 2016).

⁴⁸ Mi limito qui a menzionare gli psicoanalisti freudiani perché la distanza tra altre forme di psicoanalisi (ad esempio quella junghiana) ed il naturalismo cognitivo mi appare veramente difficile da colmare.

La terza prospettiva è quella di chi ritiene sì che sia possibile sviluppare la psicoanalisi andando oltre il punto al quale Freud era giunto, ma poi continua a tenere separata la psicologia dalle neuroscienze, com'era forse inevitabile fare un secolo fa, ma non è più giustificato oggi alla luce degli enormi recenti progressi di queste ultime. Questo atteggiamento rinunciatario è, almeno in parte, il retaggio duro a morire del clima culturale affermatosi circa trent'anni fa in Italia (e in Francia) con il movimento culturale e politico dell'antipsichiatria, quando sembrò a molti psichiatri di dover necessariamente scegliere tra due approcci alla malattia mentale ed al modo di curarla incompatibili tra loro: uno basato principalmente sulla logoterapia ed il reinserimento sociale ed uno fondato, invece, sugli psicofarmaci e la concezione della malattia mentale come vera e propria malattia organica. Nella pratica questa distinzione non ha mai retto fino in fondo ed oggi l'antipsichiatria è rivista criticamente anche da alcuni di coloro che quarant'anni fa l'avevano promossa (Corbellini e Jervis 2008).

Si profila quindi la possibilità di un quarto punto di vista riguardo all'atteggiamento da tenere verso l'eredità freudiana: quello di chi, pur attingendo ancora a piene mani alla lezione di Freud, non si rifiuta di rivedere, anche profondamente, i suoi concetti e le sue teorie alla luce dei progressi e delle recenti scoperte di tutte le scienze cognitive (comprese *in primis* le neuroscienze). Quello psicoanalista infatti che, lavorando oggi gomito a gomito con i neuroscienziati, ricercasse le basi neurologiche, ad esempio, della coscienza o del lavoro onirico, pronto, se necessario, a "commettere parricidio" e a rinnegare Freud nella misura in cui i risultati sperimentali lo richiedano, non solo non lo tradirebbe, ma realizzerebbe anzi quella che nell'*Entwurf* egli confessava essere la sua più segreta aspirazione.

Riferimenti

- Aglioti S., Berlucchi G. 2013, *Neurofobia: chi ha paura del cervello?*, Milano, Cortina.
- Ammaniti M., Gallese V. 2013, *The birth of intersubjectivity: psychodynamics, neurobiology, and the self*, New York, Norton & Company.
- Armstrong D.M. (1968), *A materialist theory of mind*, revised edition, London and New York, Routledge 1993.
- Baars B. 1988, *A cognitive theory of consciousness*, Cambridge, Cambridge University Press.
- Baars B. 1997, *In the theater of consciousness*, Oxford, Oxford University Press.
- Barale M. 2006, *Sfondi e confini ontologici delle contemporanee filosofie della mente*, «Nuova civiltà delle macchine», XXIV, 1.
- Beckermann A. 1999, *Analytische Einführung in die Philosophie des Geistes*, Berlin, de Gruyter.
- Beckermann A., Flohr H., Kim J. (eds) 1992, *Emergence or reduction? Essays on the prospects of nonreductive physicalism*, Berlin, de Gruyter.
- Benini A. 2009, *Che cosa sono io? Il cervello alla ricerca di se stesso*, Garzanti, Milano.
- Benini A. 2012, *La coscienza imperfetta: le neuroscienze e il significato della vita*, Milano, Garzanti.
- Benini A. 2017, *Neurobiologia del tempo*, Milano, Cortina.
- Bennett M.R., Dennett D.C., Hacker P.M.S., Searle J.R. (introduction and conclusion by D. Robinson) 2007, *Neuroscience and philosophy: brain, mind, and language*, New York, Columbia University Press.
- Bennett M.R., Hacker P.M.S. 2003, *Philosophical foundations of neuroscience*, Malden MA, Blackwell.
- Bergson H. (1922), *Durée et simultanéité : à propos de la théorie d'Einstein*, septième édition, Paris, P.U.F. 1968.
- Bieri P. 1995, *Why is consciousness puzzling?*, in Metzinger T. (ed), *Conscious experience*, Paderborn, Schöning.
- Biran I., Chatterjee A. 2004, *Alien hand syndrome*, «Archives of Neurology», 61, 292-294.
- Biran I., Giovannetti T., Buxbaum L., Chatterjee A. 2006, *The alien*

- hand syndrome: what makes the alien hand alien?* «Cognitive Neuropsychology» 23(4), 563–582.
- Biuso A.G. 2008, *Dispositivi semantici: introduzione fenomenologica alla filosofia della mente*, Catania, Villaggio Maori Ed.
- Block N. (1995), *On a confusion about a function of consciousness*; ora in N. Block, O. Flanagan e G. Güzelde (eds), *The nature of consciousness: philosophical debates*, Cambridge, MA, The MIT Press, 1997, pp. 375-415.
- Block N. 2007, *Consciousness, function, and representation: collected papers*, vol. 1, Cambridge, MA, The MIT Press.
- Boden M.A. (ed.) 1990, *The philosophy of artificial intelligence*, Oxford, Oxford University Press.
- Boniolo G., Dorato M. (1997), *Dalla relatività galileiana alla relatività generale*, in Boniolo G. (a cura di), *Filosofia della fisica*, Milano, Bruno Mondadori 2000, pp. 5-167.
- Bourguine P., Nadal J.P. 2004, *Cognitive economics*, Berlin, Springer.
- Bourguine P., Walliser B. (eds.) (1992), *Economics and cognitive science*, Oxford, Pergamon Press.
- Brentano F. (1874), *Psychologie vom empirischen Standpunkt*; trad. it. *Psicologia dal punto di vista empirico*, Roma-Bari, Laterza 1997.
- Brook A. 2008, *Kant's View of the Mind and Consciousness of Self*, «Stanford Encyclopedia of Philosophy», <http://plato.stanford.edu/archives/win2008/entries/kant-mind/>.
- Brooks R. A. 1990, *Elephants don't play chess*, «Robotics and Autonomous Systems», 6, 3-15.
- Brooks R. A. 1991, *Intelligence without representations*, in «Artificial Intelligence Journal», 47, 139-159.
- Cartesio (1637), *Discours de la méthode*; trad. it. *Discorso sul metodo*, in Cartesio, *Opere filosofiche*, vol. 1, Roma-Bari, Laterza 1994.
- Cartesio (1641), *Meditationes de prima philosophia*, trad.it *Meditazioni metafisiche, Obbiezioni e risposte* in Cartesio, *Opere filosofiche*, vol. 2, Roma-Bari, Laterza 1994.
- Cartesio (1644), *Principia philosophiae*; trad. it. *I principi della filosofia*, in Cartesio, *Opere filosofiche*, vol. 3, Roma-Bari, Laterza 1994.
- Carvajal J.J.R. et alii (2012), *Visual Anosognosia (Anton-Babinski*

- Syndrome): Report of Two Cases Associated with Ischemic Cerebrovascular Disease*, «Journal of Behavioral and Brain Science» 02(03), 394-398.
- Cassirer E. (1923-1929), *Philosophie der symbolischen Formen*; trad. it. *Filosofia delle forme simboliche*, Firenze, La Nuova Italia 1961-1966.
- Cerroni A., Rufo F. (a cura di) (2009), *Neuroetica: tra neuroscienze, etica e società*, Torino, UTET.
- Chakravartty A. (2011), *Scientific realism*, in “Stanford Encyclopedia of Philosophy”, <<http://plato.stanford.edu/entries/scientific-realism/>>
- Chalmers D. (1996), *The conscious mind: in search of a fundamental theory*; trad. it. *La mente cosciente*, Milano, McGraw-Hill Libri Italia, 1999.
- Chalmers D. 2010, *The character of consciousness*, Oxford, Oxford University Press.
- Chomsky N. (1957), *Syntactic structures*; trad. it. *Le strutture della sintassi*, Bari, Laterza 1970.
- Chomsky N. (1975), *Reflections on language*; trad. it. *Riflessioni sul linguaggio*, Torino, Einaudi 1981.
- Chomsky N. 1992, *A minimalist program for linguistic theory*, Cambridge MA., MIT Press.
- Churchland P.M. 1988, *Matter and consciousness: a contemporary introduction to the philosophy of mind*, Cambridge MA, The MIT Press, revised edition.
- Churchland P.M. (1989), *A neurocomputational perspective: the nature of mind and the structure of science*; trad. it. *La natura della mente e la struttura della scienza: una prospettiva neuro-computazionale*, Bologna, Il Mulino 1992.
- Churchland P.M. (1995), *The engine of reason, the seat of the soul*; trad. it. *Il motore della ragione, la sede dell'anima*, Milano, Il Saggiatore 1999.
- Churchland P.M. 2007, *Neurophilosophy at work*, Cambridge, Cambridge University Press.
- Churchland P.M., Churchland P.S. 1996, *Do we propose to eliminate consciousness?*, in R.N. McCauley (ed.), *The Churchlands and their critics*, Oxford, Blackwell, pp. 297-301.
- Churchland P.S. 1986, *Neurophilosophy: toward a unified science of*

- the mind-brain*, Cambridge MA, The MIT Press.
- Churchland P.S. 2002, *Brain-wise: studies in neurophilosophy*, Cambridge MA, The MIT Press.
- Churchland P.S. (2011), *Braintrust: what neuroscience tells us about morality*; trad. it. *Neurobiologia della morale*, Milano, Cortina 2012.
- Churchland P.S. (2013), *Touching a nerve: the self as brain*; trad. it. *L'io come cervello*, Milano, Cortina 2014.
- Corbellini G., Jervis G. 2008, *La razionalità negata: psichiatria e antipsichiatria in Italia*, Torino, Bollati-Boringhieri.
- Corbellini G., Sirgiovanni E. 2013, *Tutta colpa del cervello: un'introduzione alla neuroetica*, Milano, Mondadori.
- Crick F. (1994), *The astonishing hypothesis: the scientific search for the soul*; trad. it. *La scienza e l'anima*, Milano, Rizzoli 1994.
- Crick F., Koch C. 1990, *Towards a neurobiological theory of consciousness*, «Seminars in the Neurosciences», 2, pp. 263-275.
- Damasio A.R. (1994), *Descartes'error: emotion, reason and the human brain*; trad. it. *L'errore di Cartesio: emozione, ragione e cervello umano*, Milano, Adelphi 1995.
- Damasio A.R. 1999, *The Feeling of what happens: body, emotion and the making of consciousness*, London, Heinemann.
- Damasio A.R. (2003), *Looking for Spinoza: joy, sorrow and the feeling brain*; trad. it. *Alla ricerca di Spinoza: emozioni, sentimenti e cervello*, Milano, Adelphi 2003.
- Damasio A.R. (2010), *Self comes to mind: constructing the conscious brain*; trad. it. *Il sé viene alla mente: la costruzione del cervello cosciente*, Milano, Adelphi 2012.
- Dawkins R. (1976), *The selfish gene*; trad. it. *Il gene egoista*, Bologna, Zanichelli 1983.
- Dawkins R. (1982), *Extended phenotype: the gene as the unit of selection*; trad. it. *Il fenotipo esteso: il gene come unità di selezione*, Bologna, Zanichelli 1986.
- Dawkins R. (1986), *Blind watchmaker*; trad. it. *L'orologiaio cieco: creazione o evoluzione?* Milano, Fabbri 2009.
- Dawkins R. (2009), *The greatest show on earth: the evidence for evolution*; trad. it. *Il più grande spettacolo della Terra: perché Darwin aveva ragione*, Milano, Mondadori 2010.
- De Caro M., Macarthur D. (eds) (2004), *Naturalism in question*;

- trad. it. *La mente e la natura. Per un naturalismo liberalizzato*, Fazi, Roma, 2005.
- Dehaene S. (2014), *Consciousness and the brain: deciphering how the brain codes our thoughts*; trad. it. *Coscienza e cervello: come i neuroni codificano il pensiero*, Milano, Cortina 2014
- Dellantonio S. 2007, *La dimensione interna del significato: esternismo, internismo e competenza semantica*, Pisa, ETS.
- Dennett D.C. (1969), *Content and consciousness*; trad. it. *Contenuto e coscienza*, Bologna, Il Mulino, 1992.
- Dennett D.C. (1978), *Brainstorms: philosophical essays on mind and psychology*; trad. it. *Brainstorms: saggi filosofici sulla mente e la psicologia*, Milano, Adelphi 1991.
- Dennett D.C. 1984, *Elbow room: the varieties of free will worth wanting*, Cambridge, MA.
- Dennett D.C. (1987), *The Intentional stance*; trad. it. *L'atteggiamento intenzionale*, Bologna, Il Mulino 1993.
- Dennett D.C. 1988, *Quining qualia*, in Marcel A.J., Bisiach E. (eds), *Consciousness in contemporary science*, Oxford, Clarendon Press, pp. 42-77.
- Dennett D.C. (1991), *Consciousness explained*; trad. it. *Coscienza, che cosa è*, Roma-Bari, Laterza 2009.
- Dennett D.C. (1995), *Darwin's dangerous idea: evolution and the meanings of life*; trad. it. *L'idea pericolosa di Darwin*, Torino, Bollati-Boringhieri 1997.
- Dennett D.C. 1998, *Brainchildren: essays on designing minds*, Cambridge, MA, The MIT Press.
- Dennett D.C. (1999), *The zombic hunch: extinction of an intuition?*; ora in Dennett (2013a, cap. 54)
- Dennett D.C. (2003), *Freedom evolves*; trad. it. *L'evoluzione della libertà*, Milano, Cortina 2004.
- Dennett D.C. (2005), *Sweet dreams: philosophical obstacles to a science of consciousness*; trad. it. *Sweet dreams: illusioni filosofiche sulla coscienza*, Milano, Cortina 2006.
- Dennett D.C. (2013a), *Intuition pumps and other tools for thinking*; trad. it. *Strumenti per pensare*, Milano, Cortina 2014.
- Dennett D.C. (2013b), *Zombies and zimboes*; ora in Dennett (2013a, cap. 55).
- Dennett D.C. 2016, *Illusionism as the Obvious Default Theory of*

- Consciousness*, «Journal of Consciousness Studies», 23, 11-12, pp. 65-72.
- Dennett D.C. (2017), *From bacteria to Bach and back: The evolution of minds*; trad. it. *Dai batteri a Bach: come evolve la mente*, Milano, Cortina 2018.
- Di Francesco M. 1998, *L'io e i suoi sé*, Milano, Cortina.
- Di Francesco M. 2000, *La coscienza*, Roma-Bari, Laterza.
- Dretske F.I. 1981, *Knowledge and the flow of information*, Cambridge MA, MIT Press.
- Dretske F.I. 1995, *Naturalizing the Mind*, Cambridge MA, The MIT Press.
- Dreyfus H.L. 1999, *What Computers Still Can't Do: A critique of artificial reason*, Cambridge, MA.
- Dummett M. (1988), *Origins of analytical philosophy*; trad. it. *Origini della filosofia analitica*, Torino, Einaudi 1990.
- Dummett M. (1991), *The logical basis of metaphysics*; trad. it. *La base logica della metafisica*, Bologna, Il Mulino 1996.
- Durkheim É. (1894), *Les règles de la méthode sociologique*; trad. it. *Le regole del metodo sociologico*, Milano, Comunità 1963.
- Edelman G.M. (1987), *Neural darwinism: the theory of neuronal group selection*; trad. it. *Darwinismo neurale: la teoria della selezione dei gruppi neuronali*, Torino, Einaudi 1996.
- Edelman G.M. (1989), *The remembered present: a biological theory of consciousness*; trad. it. *Il presente ricordato: una teoria biologica della coscienza*, Milano, Rizzoli 1991.
- Edelman G.M. (1992), *Bright air, brilliant fire: on the matter of the mind*; trad. it. *Sulla materia della mente*, Adelphi, Milano, 1993.
- Edelman G.M. (2004), *Wider than the sky: the phenomenal gift of consciousness*; trad. it. *Più grande del cielo: lo straordinario dono fenomenico della coscienza*, Torino, Einaudi 2004.
- Edelman G.M. (2006), *Second nature: brain science and human knowledge*; trad. it. *Seconda natura: scienza del cervello e conoscenza umana*, Milano, Cortina 2007.
- Edelman G.M., Tononi G. (2000), *A universe of consciousness: how matter becomes imagination*, New York, Basic Books.
- Einstein A. (1917), *Über die spezielle und allgemeine Relativitätstheorie (gemeinverständlich)*; trad. it. *Relatività: esposizione divulgativa*, Torino, Boringhieri 1960.

- Einstein A. (1949), *Autobiographisches – Autobiographical notes*; trad. it. *Note Autobiografiche*, in P.A. Schilpp (a cura di), *Albert Einstein, scienziato e filosofo: autobiografia di Einstein e saggi di vari autori*, Torino, Einaudi 1958, pp. 3-49.
- Ellenberger H.F. (1970), *The discovery of the unconscious: the history and evolution of dynamic psychiatry*; trad. it. *La scoperta dell'inconscio: storia della psichiatria dinamica*, 2 voll., Torino, Boringhieri, 1976.
- Ellis G.F.R. 2016, *How can physics underlie the mind?*, Berlin, Springer.
- Engel A.K., Singer W. 2001, *Temporal binding and the neural correlates of sensory awareness*, «Trends in Cognitive Sciences», 5 (1), http://bernardbaars.pbworks.com/f/***+CH+ATTEN+&+CSNS+Engel+&+Singer+Time+binding+TICS+2001.pdf
- Engel A.K. 2003, *Temporal Binding and the Neural Correlates of Consciousness*, in A. Cleeremans (ed.), *The Units of Consciousness: Binding, Integration and Dissociation*, Oxford, Oxford University Press.
- Engel A.K. 2005, *Neuronale Synchronisation und Wahrnehmungsbewusstsein*, in C.S. Hermann, M. Pauen, J.W. Rieger, S. Schicktanz (Hg.), *Bewusstsein: Philosophie, Neurowissenschaften, Ethik*, München, Fink, pp. 216-241.
- Fattori L., Vandi G. 2016, *Psicoanalisi e fede: un discorso aperto*, Milano, Angeli.
- Feigl H. 1967, *The “Mental” and the “Physical”: essay and a post-script*, Minneapolis MN, University of Minnesota Press.
- Fodor J.A. 1975, *The Language of Thought*, New York, Thomas Y. Crowell Co.
- Fodor J.A. 1981, *Representations: philosophical essays on the foundations of cognitive science*, Cambridge MA, The MIT Press.
- Fodor J.A. (1983), *The modularity of the mind: an essay on faculty psychology*; trad. it. *La mente modulare: saggio di psicologia delle facoltà*, trad. it. Bologna, Il Mulino 1988.
- Fodor J.A. (1987), *Psychosemantics: the problem of meaning in the philosophy of mind*; trad. it. *Psicosemantica: il problema del significato nella filosofia della mente*, Bologna, Il Mulino 1990.
- Fodor J.A. (1998), *Concepts: where cognitive science went wrong*; trad. it. *Concetti: dove sbaglia la scienza cognitiva*, Milano,

McGraw-Hill 1999.

- Fodor J.A. (2000), *The mind doesn't work that way: the Scope and limits of computational psychology*; trad. it. *La mente non funziona così: la portata e i limiti della psicologia computazionale*, Roma-Bari, Laterza 2004.
- Fodor J.A., Piattelli-Palmarini M. (2010), *What Darwin got wrong*; trad. it. *Gli errori di Darwin*, Milano, Feltrinelli 2010.
- Frege G. (1892), *Über Sinn und Bedeutung*, trad. it. in Frege G., *Senso, funzione, concetto*, Roma-Bari, Laterza 2000.
- Freud S. (1895), *Entwurf einer Psychologie*; trad. it. *Progetto di una psicologia*, Torino, Boringhieri 1976.
- Freud S. (1901), *Zur Psychopathologie des Alltagslebens*; trad. it. *Psicopatologia della vita quotidiana*, Torino, Boringhieri 1971.
- Freud S. (1917), *Vorlesungen zur Einführung in die Psychoanalyse*; trad. it. *Introduzione alla psicoanalisi*, Torino, Boringhieri 1969.
- Gallagher S., Zahavi D. (2008), *The phenomenological mind: an introduction to philosophy of mind and cognitive science*; trad. it. *La mente fenomenologica: filosofia della mente e scienze cognitive*, Milano, Cortina 2009.
- Gallese V., Goldman A. (1998), *Mirror neurons and the simulation theory of mind-reading*, «Trends in Cognitive Sciences», 12, pp. 493-501.
- Gardner H. (1985), *The mind's new science: a history of the cognitive revolution*; trad. it. *La nuova scienza della mente: storia della rivoluzione cognitiva*, Milano, Feltrinelli 1988.
- Gay P. (1988), *Freud: a life for our time*; trad. it. *Freud: una vita per i nostri tempi*, Milano, Bompiani 1988.
- Glanzberg, M. (2016), *Truth*, «The Stanford Encyclopedia of Philosophy», <https://plato.stanford.edu/archives/win2016/entries/truth/>
- Gould S.J. (2002), *The structure of evolutionary theory*; trad. it. *La struttura della teoria dell'evoluzione*, Torino, Codice Edizioni 2003.
- Gozzano S. 2007, *Pensieri materiali*, Torino, UTET 2007.
- Grünbaum A. 1984, *The Foundations of Psychoanalysis: A Philosophical Critique*, Berkeley, University of California Press.
- Haggard P., Libet B. 2001, *Conscious intention and brain activity*, «Journal of Consciousness Studies» 8, 11, pp. 47-63.
- Hassin R.R., Uleman J.S., Bargh J.A. (eds) 2006, *The new uncon-*

- scious, Oxford, Oxford University Press.
- Hempel C.G. 1965, *Aspects of scientific explanation and other essays in the philosophy of science*, New York-London, The Free Press-MacMillan.
- Hinton, G. E., Sejnowski, T. J., 1986, *Learning and Relearning in Boltzmann Machines*, in Rumelhart D.E., McClelland J.L. (eds), *Parallel distributed processing: explorations in the micro-structure of cognition*, Volume 1: *Foundations*. Cambridge MA, The MIT Press, pp. 282–317.
- Hook S. (ed.) 1960, *Dimensions of mind*, New York, New York University Press.
- Horwich P. 1998, *Truth*, Oxford, Clarendon Press.
- Hume D. (1739-1740), *A treatise of human nature: being an attempt to introduce the experimental method of reasoning into moral subjects*; trad. it. *Trattato sulla natura umana*, in *Opere filosofiche I*, Roma-Bari, Laterza 1992.
- Husserl E. (1913), *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*; trad. it. *Idee per una fenomenologia pura e per una filosofia fenomenologica*, Torino, Einaudi 1965.
- Innocenti A. 2009, *L'economia cognitiva*, Roma, Carocci.
- Izhikevich E.M. 2007, *Dynamical systems in neuroscience: the geometry of excitability and bursting*, MIT Press, Cambridge, MA.
- Jackson F. (1982), *Epiphenomenal qualia*, ora in Lycan W.G. (ed.), *Mind and cognition: a reader*, Blackwell, Oxford 1990, pp. 469-477.
- Jackson F. 1986, *What Mary didn't know*, «Journal of Philosophy», 83, pp. 291-295.
- Jackson F. 1998, *Minds, methods, and conditionals: selected essays*, London, Routledge.
- James W. (1890), *The principles of psychology*; trad. it. *Principi di psicologia*, Milano, Principato 1965.
- Johnson-Laird P.N. (1983), *Mental models*; trad. it. *Modelli mentali*, Bologna, Il Mulino 1988.
- Kandel E.R. et alii (1995), *Essentials of neuroscience and behavior*; trad. it. *Fondamenti delle neuroscienze e del comportamento*, Milano, Casa Editrice Ambrosiana 1999.

- Kant I. (1787), *Kritik der reinen Vernunft*, 2. Auflage; trad. it. *Critica della ragion pura*, Roma-Bari, Laterza 1993.
- Kaplan D. and Glass L. (eds) 1995, *Understanding nonlinear dynamics*, New York, Springer.
- Kim J. 1993, *Supervenience and mind: selected philosophical essays*, Cambridge, Cambridge University Press.
- Kim J. (1998), *Mind in a physical world: an essay on the mind-body problem and mental causation*; trad. it. *La mente e il mondo fisico*, Milano, McGraw-Hill, 2000.
- Kim J. (2005), *Physicalism, or something near enough*; trad. it. *La mente ai confini del fisicalismo*, Milano, Melquìades, 2012.
- Koppelberg D. 1999, *What is naturalism today – And what should it be tomorrow?*, in D. Marconi (a cura di), *Naturalismo e naturalizzazione*, Vercelli, Mercurio, pp. 23-45.
- Kripke S.A. 1971, *Identity and necessity*, in Munitz M. (ed.), *Identity and individuation*, New York, New York University Press, pp. 160-164.
- Kuhn T.S. (1962), *The structure of scientific revolutions*; trad. it. *La struttura delle rivoluzioni scientifiche*, Torino, Einaudi 1969.
- Lanfredini R. 2010, *Fenomenologia naturalizzata e fenomenologia naturale: Il problema degli stati qualitativi*, in M.C. Amoretti (a cura di), *Natura umana e natura artificiale*, Milano, Angeli, 117-134.
- Lavazza A. (a cura di) 2008, *L'uomo a due dimensioni: il dualismo mente-corpo oggi*, Milano, Mondadori.
- Legrenzi P. (a cura di) 1980, *Storia della psicologia*, Il Mulino, Bologna.
- Legrenzi P., Umiltà C. 2009, *Neuro-mania: il cervello non spiega chi siamo*, Bologna, Il Mulino.
- Leibniz G.W. (1710), *Essais de théodicée sur la bonté de Dieu, la liberté de l'homme et l'origine du mal*; trad. it. *Saggi di teodicea sulla bontà di Dio, sulla libertà dell'uomo, sull'origine del male*, Milano, Rizzoli 1993.
- Levin J. (2013), *Functionalism*, in «Stanford Encyclopedia of Philosophy», <https://plato.stanford.edu/entries/functionalism/>
- Levine J. 1983, *Materialism and qualia: the explanatory gap*, in «Pacific Philosophical Quarterly» 64, pp. 354-61.
- Levine J. 2001, *Purple haze : the puzzle of consciousness*, Oxford,

- Oxford University Press.
- Libet B. (2004), *Mind time: the temporal factor in consciousness*; trad. it. *Mind time: il fattore temporale nella coscienza*, Milano, Cortina 2007.
- Lindemann B. (2014), *Mechanisms in world and mind: perspective dualism, system theory, neuroscience, reductive physicalism*, Exeter, Imprint Academic. <https://books.google.it/books?id=-Ly7BAAAQBAJ&printsec=frontcover&dq=bernd+lindemann+top+%22Mechanisms+in+world%22&hl=it&sa=X&ved=0ahUKEwilwvf-n5LdAhUD2KQKHRdRDfwQ6AEIJzAA#v=onepage&q=top%20down&f=false>
- Locke J. (1690), *An Essay Concerning Human Understanding*; trad. it. *Saggio sull'intelletto umano*, Milano, Bompiani 2004.
- Lowe E.J. 2000, *An Introduction to the Philosophy of Mind*, Oxford, Oxford University Press.
- Makeig S. 2001, *Swartz Center for computational neuroscience: vision overview*, <https://scn.ucsd.edu/VisionOverview.html>
- Malcolm N. 1966, *L. Wittgenstein: a memoir*, Oxford, Oxford University Press.
- Mancia M. (2006) (a cura di), *Psychoanalysis and neurosciences*; trad. it. *Psicoanalisi e neuroscienze*, Milano, Springer 2007.
- Matsushashi M., Hallet M. 2008, *The timing of the conscious intention to move*, «European Journal of Neuroscience», 28 (11), pp. 2344-51. PMC 4747633. PMID 19046374. doi:10.1111/j.1460-9568.2008.06525.x.
- McGinn C. 1982, *The character of mind*, Oxford, Oxford University Press.
- McGinn C. 1991, *The problem of consciousness*, Oxford, Blackwell.
- Metzinger T. 2003, *Being no one: the self-model theory of subjectivity*, Cambridge MA, The MIT Press.
- Metzinger T. (2009), *The ego tunnel: the science of mind and the myth of the self*; trad. it. *Il tunnel dell'io: scienza della mente e mito del soggetto*, Milano, Cortina 2010.
- Migone P. 2007, *L'inconscio psicoanalitico e l'inconscio cognitivo*, «Il Ruolo Terapeutico», 105, pp. 51-61.
- Minelli A. 2007, *Forme del divenire*, Torino, Einaudi.
- Nagel T. (1974), *What is it like to be a bat?*, trad. it. *Che cosa si*

- prova a essere un pipistrello?*, in Nagel T., *Questioni mortali*, Milano, Il Saggiatore 1986, pp. 162-175.
- Nannini S. 1992, *Cause e ragioni: modelli di spiegazione delle azioni umane nella filosofia analitica*, Roma, Editori Riuniti.
- Nannini S. 2004, *Mental causation and intentionality in a mind naturalizing theory*, in A. Peruzzi (ed.), *Mind and Causality*, Amsterdam-Philadelphia, John Benjamins, pp. 69–95.
- Nannini S. 2006, *Theorien mentaler Repräsentation*, in Sandkühler H.J. (Hg.), *Theorien und Begriffe der Repräsentation: Beiträge zu einem Workshop*, Bremen, Schriftenreihe der von der VolkswagenStiftung geförderten Forschergruppe Repräsentation, Bd. 1, pp. 91-98.
- Nannini S. 2007a, *Naturalismo cognitivo: per una teoria materialistica della mente*, Macerata, Quodlibet.
- Nannini S. 2007b, *Intentionality Naturalised*, in Beany M., Penco C., Vignolo M. (eds), *Explaining the mental: naturalist and non-naturalist approaches to mental acts and processes*, Angerton Gardens UK, Cambridge Scholars Publ., pp. 124-134.
- Nannini S. 2010, *Cognitive naturalism and cognitive neuroscience: a defence of eliminativism and a discussion with G. Roth*, in De Caro M., Egidi R. (eds), *The architecture of knowledge: epistemology, agency, and science*, Roma, Carocci, pp. 35-56.
- Nannini S. 2011, *L'anima e il corpo: un'introduzione storica alla filosofia della mente*, nuova ed. Roma-Bari, Laterza.
- Nannini S. 2013, *La coscienza può essere naturalizzata?*, in Nannini S. e Zeppi A. (a cura di), *La mente può essere naturalizzata?*, Colle Val d'Elsa (Siena), Salvietti e Barabuffi, pp. 117-142.
- Nannini S. 2016, *Consciousness naturalised*, in «Idee» n.s. 6 (11), pp. 21-40.
- Nannini S., Sandkühler H.J. (eds) 2000, *Naturalism in the cognitive sciences and the philosophy of Mind*, Frankfurt a. M., Peter Lang.
- Neisser U. (1967), *Cognitive psychology*; trad. it. *Psicologia cognitivista*, Firenze, Giunti-Barbera 1975.
- Noë A. 2004, *Action in perception*, Cambridge MA, The MIT Press.
- Noë A. (2009), *You are not your brain*; trad. it. *Perché non siamo il nostro cervello: una teoria radicale della coscienza*, Milano, Cortina 2010.

- O'Connor T., Wong H.Y. (2015), *Emergent properties*, «Stanford Encyclopedia of Philosophy», <https://plato.stanford.edu/archives/sum2015/entries/properties-emergent/>
- Oliverio A. 1995, *Biologia e filosofia della mente*, Roma-Bari, Laterza.
- Oliverio A. 2009, *Introduzione alle neuroscienze*, Roma-Bari, Laterza.
- Papineau D. 2016, *Naturalism*, in «Stanford Encyclopedia of Philosophy», <https://plato.stanford.edu/archives/win2016/entries/naturalism/>.
- Parrini P. 1995, *Conoscenza e realtà: saggio di filosofia positiva*, Roma-Bari, Laterza.
- Parrini P. 2011, *Il valore della verità*, Milano, Guerini e Associati.
- Parrini P. 2012, *Realismi a prescindere. A proposito di realtà ed esperienza*, «Iride», XXV, 67, pp. 495-520.
- Pastore L., Dellantonio S. 2017, *Internal perception: the role of bodily information in concepts and word mastery*, Berlin, Springer.
- Pastore L., Dellantonio S. (a cura di) (2009), *Percezione, rappresentazione, coscienza: questioni filosofiche e problemi epistemologici*, Pisa, ETS.
- Pauen M., Roth G. 2001, *Neurowissenschaften und Philosophie*, München, Fink.
- Pauen M., Roth G. 2008, *Freiheit, Schuld und Verantwortung: Grundzüge einer naturalistischen Theorie der Willensfreiheit*, Frankfurt a. M., Suhrkamp.
- Perconti P. 2008, *L'autocoscienza: cosa è, come funziona, a cosa serve*, Roma-Bari, Laterza.
- Perconti P. 2011, *Coscienza*, Bologna, Il Mulino.
- Petitot J. et alii (eds.) (1999), *Naturalizing phenomenology: issues in contemporary phenomenology and cognitive science*, Stanford, Stanford University Press.
- Piattelli-Palmarini M. (ed.) 1980, *Language and learning: the debate between Jean Piaget and Noam Chomsky (the Royaumont debate)*, part 1 and part 2, Cambridge MA, Harvard University Press.
- Pievani T. 2006, *La teoria dell'evoluzione: attualità di una rivoluzione scientifica*, Bologna, Il Mulino.

- Pitt D. 2017, *Mental Representation*, «Stanford Encyclopedia of Philosophy», <https://plato.stanford.edu/archives/spr2017/entries/mental-representation/>
- Place U.T. (1956), *Is consciousness a brain process?*, ora in Place U.T., *Identifying the mind: selected papers of U.T. Place*, Oxford, Oxford University Press, 2004, pp. 45-69.
- Pöppel E. (1997a), *A hierarchical model of temporal perception*, «Trends in Cognitive Sciences», vol. I, n. 2, pp. 56-61.
- Pöppel E. 1997b, *The brain's way to create "nowness"*, in Atmanspacher H., Ruhnau E.(eds), *Time, temporality, now: experiencing time and concepts of time in an interdisciplinary perspective*, Berlin, Springer, pp. 107-120.
- Pöppel E., Logothetis N. 1986, *Neuronal oscillations in the human brain: discontinuous initiations of pursuit eye movements indicate a 30-Hz temporal framework for visual information processing*, «Naturwissenschaften», vol. LXXIII, n. 5, pp. 267-268.
- Popper K.R. (1994), *Knowledge and the body-mind problem: in defence of interaction*; trad. it. *La conoscenza e il problema mente-corpo*, Bologna, Il Mulino 1994.
- Prigatano G.P., Schacter D.L 1991, *Awareness of deficit after brain injury: clinical and theoretical issues*, Oxford, Oxford University Press.
- Putnam H. (1960), *Minds and machines*; trad. it. *Menti e macchine*, in Putnam H., *Mente, linguaggio e realtà*, Milano, Adelphi, 1987, pp. 392-416.
- Putnam H. (1963), *Brains and behavior*; trad. it. *Cervelli e comportamenti*, in Putnam H., *Mente, linguaggio e realtà*, Milano, Adelphi, 1987, pp. 334-371.
- Putnam H. 2016, *Naturalism, realism, and normativity*, Cambridge MA, Harvard University Press.
- Quine W.v.O. (1951), *Two dogmas of empiricism*; trad. it. *Due dogmi dell'empirismo*, in Quine W.v.O., *Il problema del significato*, Roma, Ubaldini 1966, pp. 20-44.
- Quine W.v.O. (1953), *From a logical point of view*; trad. it. *Il problema del significato*, Roma, Ubaldini 1966.
- Quine W.v.O. (1960), *Word and object*; trad. it. *Parola e oggetto*, Milano, Il Saggiatore 1970.
- Quine W.v.O. 1969, *Ontological relativity and other essays*, New

- York, Columbia U.P.
- Racine E. 2010, *Pragmatic neuroethics*, Cambridge MA, MIT Press.
- Ramachandran V.S, Hirstein W. 1999, *The science of art: a neurological theory of aesthetic Experience*, «Journal of Consciousness Studies», 6 (6-7), pp. 15-51.
- Rand N. (2004), *The hidden soul: the growth of the unconscious in philosophy, psychology, medicine, and literature, 1750-1900*; trad. it. *L'anima nascosta: nascita ed evoluzione del concetto di inconscio in filosofia, psicologia, medicina e letteratura*, «Psicoterapia e Scienze Umane», 42, 1, 2008, pp. 13-40.
- Rizzolatti G., Sinigaglia C. 2006, *So quel che fai: il cervello che agisce e i neuroni specchio*, Milano, Cortina.
- Robinson H. 2016, *Dualism*, «Stanford Encyclopedia of Philosophy», <https://plato.stanford.edu/entries/dualism/>
- Robinson W. 2004, *Understanding phenomenal consciousness*, Cambridge, Cambridge University Press
- Romano G. 2002, *La mente mimetica: riflessioni e prospettive sulla teoria della simulazione mentale*, Siena, Protagon.
- Romano G. 2006, *Essere per: il concetto di "funzione" tra scienze, filosofia e senso comune*, Macerata, Quodlibet.
- Romano G. 2009, *Thoughtful Things: an investigation in the descriptive epistemology of artifacts*, Eindhoven, Technische Universiteit.
- Rorty R. (1967), *The linguistic turn*; trad. it. *La svolta linguistica*, Milano, Garzanti 1994.
- Roth G. 1994, *Das Gehirn und seine Wirklichkeit: kognitive Neurobiologie und ihre philosophischen Konsequenzen*, Frankfurt a. M., Suhrkamp.
- Roth G. 2001, *Fühlen, Denken, Handeln: wie das Gehirn unser Verhalten steuert*, Frankfurt a. M., Suhrkamp.
- Roth G., Grün K.J. (Hg.) 2006, *Das Gehirn und seine Freiheit: Beiträge zur neurowissenschaftlichen Grundlegung der Philosophie*, Göttingen, Vandenhoeck und Ruprecht.
- Roth G., Strüder N. 2014, *Wie das Gehirn die Seele macht*, Stuttgart, Cotta.
- Rumelhart D.E., McClelland J.L. (eds) (1986), *Parallel distributed processing: explorations in the microstructure of cognition*; trad. it. parziale *PDP: microstruttura dei processi cognitivi*, Bologna,

- Il Mulino 1991.
- Russell B. (1905), *On Denoting*; trad. it. in A. Bonomi (a cura di), *La struttura logica del linguaggio*, Milano, Bompiani, 1973, pp. 179-195.
- Russell B. (1912), *The problems of philosophy*; trad. it. *I problemi della filosofia*, Milano, Feltrinelli 1973.
- Russell B. (1921), *The analysis of Mind*; trad. it. *L'analisi della mente*, Roma, Newton Compton 1970.
- Ryle G. (1949), *The concept of mind*; trad. it. *Il concetto di mente*, Roma, Laterza 2007.
- Searle J.R. (1980), *Minds, brains, and programs*; trad. it. *Menti, cervelli e programmi*, in Hofstadter D.R. e Dennett D.C. (a cura di), *L'io della mente*, Milano, Adelphi 1981, pp. 341-360.
- Searle J.R. (1992), *The rediscovery of the mind*; trad. it. *La riscoperta della mente*, Torino, Bollati-Boringhieri 1994.
- Searle J.R. (1997), *The mystery of consciousness*; trad. it. *Il mistero della coscienza*, Milano, Cortina 1998.
- Sellars W. (1956), *Empiricism and the philosophy of mind*; trad. it. *Empirismo e filosofia della mente*, Torino, Einaudi 2004.
- Sellars W. 1975, *The adverbial theory of the objects of sensation*, in «Metaphilosophy», 6, pp. 144–160.
- Singer W. 2001, *Consciousness and the binding problem*, in P.C. Marijuàn (a cura di), *Cajal and consciousness: scientific approaches to consciousness on the centennial of Ramón y Cajal's Textura*, New York, New York Academy of Sciences, pp. 123-146.
- Singer W. 2003, *Ein neues Menschenbild? Gespräche über Hirnforschung*, Frankfurt a. M., Suhrkamp.
- Singer W. 2004, *Synchrony, oscillations, and relational codes*, in Chalupa L.M., Werner J.S. (eds), *The visual neurosciences*, Cambridge MA, The MIT Press, 2nd ed., pp. 1665-1681.
- Smart J.J.C. (1963), *Materialism*; ora in Borst C.V. (ed.), *The mind-brain identity theory*, New York, Macmillan 1970, 159-170.
- Smith D.W. et alii (eds) 2005, *Phenomenology and philosophy of mind*, Oxford, Oxford University Press.
- Soon C.S. et alii 2008, *Unconscious determinants of free decisions in the human brain*, in «Nature Neuroscience», 11, pp. 543-545, doi: 10.1038/nn.2112
- Spinoza B. (1677), *Ethica ordine geometrico demonstrata*; trad. it.

- Etica*, Milano, Bompiani 2007.
- Stamenov M.I., Gallese V. (eds.) 2002, *Mirror neurons and the evolution of brain and language*, Amsterdam-Philadelphia, John Benjamins.
- Stephan A. (1999), *Emergenz: von der Unvorhersagbarkeit zur Selbstorganisation*, Dresden-München, Dresden University Press.
- Stich S., Warfield T.A. (eds) 1994, *Mental Representation*, Oxford, Blackwell.
- Stoljar D. 2015, *Physicalism*, «Stanford Encyclopedia of Philosophy», <https://plato.stanford.edu/entries/physicalism/>
- Strawson P.F. (1950) *On Referring*; trad. it. in A. Bonomi (a cura di), *La struttura logica del linguaggio*, Milano, Bompiani 1973, pp. 197-224.
- Swaab D. (2010), *We are our brains: a neurobiography of the brain, from the Womb to Alzheimer's*; trad. it. *Noi siamo il nostro cervello: come pensiamo, soffriamo e amiamo*, Roma, Elliot 2011.
- Swinburne R. 1986, *The evolution of the soul*, Oxford, Oxford University Press.
- Swinburne R. 2013, *Mind, brain, and free will*, Oxford, Oxford University Press.
- Tacca M.C. 2010, *Seeing objects: the structure of visual representation*, Paderborn, Mentis Verlag.
- Tarski A. 1956, *Logic, semantics, metamathematics: papers from 1923 to 1938*, Oxford, Clarendon Press.
- Tye M. 1984, *The adverbial theory of visual experience*, «Philosophical Review», 93, pp. 195-225.
- Wegner D.M. 2002, *The illusion of conscious will*, Cambridge MA, The MIT Press.
- Weiskrantz L. 2009, *Blindsight: a case study spanning 35 years and new developments*, Oxford, Oxford University Press.
- Whyte L.L. (1960), *The unconscious before Freud*; trad. it. *L'inconscio prima di Freud: una storia dell'evoluzione della coscienza umana*, Roma, Astrolabio-Ubaldini 1970.
- Yalowitz S. 2014, *Causal closure of the physical in the argument for monism*, «Stanford Encyclopedia of Philosophy». supplementary document file (causal-closure.html) associated with *Anomalous Monism*, <https://plato.stanford.edu/archives/win2014/entries/anomalous-monism/>

Zeki S. 1999, *Inner Vision: An Exploration of Art and the Brain*, Oxford, Oxford University Press.

Zeki S. (2008), *Splendours and Miseries of the Brain*, Oxford, Blackwell.

Zeppi A., Plebe A., Perconti P. 2016, *Looking at complexity the other way around*, Conference Paper: MIUCS 2016 Warsaw
https://www.researchgate.net/publication/304323964_Looking_at_complexity_the_other_way_around



Sandro Nannini

Mente e cervello

La seconda svolta cognitiva della mente e altri saggi

Gli stati mentali da un punto di vista ontologico sono riducibili a vari aspetti e proprietà dell'attività cerebrale. Questo il filo rosso che lega gli otto saggi qui riproposti. L'incontro della filosofia della mente con le scienze cognitive consente di formulare una soluzione naturalistico-fisicalistica del problema mente-corpo (capitolo primo). Ad esempio l'Intenzionalità del mentale, l'Io e la coscienza sono presumibilmente aspetti degli stati mentali implementati da tre diverse *higher order properties* della dinamica cerebrale (capitolo secondo). Inoltre la soluzione che deve essere data al problema mente-corpo è simile a quella che in genere viene data al problema individuo-società (capitolo terzo). Il cambiamento di paradigma dal dualismo cartesiano al naturalismo cognitivo presenta alcuni punti di somiglianza con la rivoluzione che investì la fisica agli inizi del Novecento (capitolo quarto). Il naturalismo cognitivo qui proposto implica il realismo scientifico (capitolo quinto) ed una concezione pragmatistica della verità (capitolo sesto). Infine il concetto kantiano di appercezione trascendentale può aiutare a chiarire la distinzione tra il Self e l'Io (capitolo settimo), così come è possibile un incontro tra naturalismo fisicalistico e psicoanalisi freudiana (capitolo ottavo).

Sandro Nannini è un ex-professore di Filosofia Teoretica e di *Philosophy of Mind* dell'Università di Siena. Dopo essersi interessato allo strutturalismo francese (*Il pensiero simbolico: saggio su Lévi-Strauss* 1981), alla teoria dell'azione (*Cause e ragioni* 1992) e alla meta-etica (*Il Fanatico e l'Arcangelo* 1998), si è dedicato interamente alla filosofia della mente (cfr. in particolare *L'anima e il corpo* 2002 e 2011, *Naturalismo cognitivo* 2007 e *La notte di Minerva* 2008). Su Corisco è comparsa una raccolta di saggi a lui dedicata (C. Lumer e G. Romano (a cura di), *Dalla filosofia dell'azione alla filosofia della mente* 2018).

