



UNIVERSITÀ DEGLI STUDI DI
CASSINO E DEL LAZIO MERIDIONALE

Corso di Dottorato in
ISTITUZIONI, MERCATI E COMPORTAMENTI

INDIRIZZO *COMPORTAMENTI*
XXXII CICLO

*Empatia e libero arbitrio:
un colloquio aperto fra neuroscienze, etica e
psicologia*

SSD: M-PSI/01

Coordinatore del Corso
Chiar.ma Prof.ssa
Rosella Tomassoni

Dottoranda
Dott.ssa
Federica Madonna

Supervisore
Chiar.ma Prof.ssa
Rosella Tomassoni

Indice

Introduzione	p. 5
---------------------	------

Parte I

La neuroetica

Capitolo I

Il quadro di riferimento attuale della ricerca	p. 14
--	-------

Parte II

L'empatia cognitiva

Capitolo II

Che cos'è l'empatia

2.1 La difficoltà di una definizione unanime	p. 26
2.2 Differenza fra empatia, simpatia e contagio emotivo	p. 42
2.3 Empatia ed altruismo: alle origini della morale?	p. 52

Capitolo III

L'aspetto cognitivo dell'empatia

3.1 Come nasce il problema del <i>mentalizing</i>	p. 71
---	-------

3.2 L'empatia fra <i>Theory Theory</i> e <i>Simulation Theory</i>	p. 85
3.2.1 La <i>folk-empathy</i> nella <i>Theory Theory</i>	p. 86
3.2.2 La neuro-empatia e la <i>Simulation Theory</i>	p. 93

Capitolo IV

Fra *mindreading* e *metacognition*: un rapporto controverso

4.1 La posizione di Peter Carruthers: <i>Mindreading is prior</i>	p. 105
4.1.2. <i>Model 1: Two mechanism</i>	p. 107
4.1.3. <i>Model 2: One mechanism, two modes of access</i>	p. 109
4.1.4. <i>Model 3: Metacognition is prior</i>	p. 110
4.1.5. <i>Model 4: Mindreading is prior</i>	p. 111
4.2 <i>Pro</i> Carruthers	p. 121
4.3 <i>Contra</i> Carruthers	p. 125
4.4 La risposta di Peter Carruthers	p. 134

Parte III

Il libero arbitrio

Capitolo V

Il dibattito sul libero arbitrio

5.1 Il determinismo: punti di forza, carenze e difficoltà	p. 138
5.2 L'indeterminismo: un'alternativa, forse, non troppo convincente	p. 146

Capitolo VI

Libero arbitrio e neuroscienze

6.1 L'esperimento di Benjamin Libet: constatazioni e perplessità	p. 164
6.2 "Il nuovo Libet": vecchi intenti, medesime difficoltà	p. 172
6.3 La proposta di Michael Gazzaniga: <i>the interpreter</i>	p. 179
6.4 Il libero arbitrio fra emozione ed illusione	p. 183
6.5 Il libero arbitrio, la responsabilità morale e le neuroscienze	p. 193

Capitolo VII

Verso un nuovo approccio al libero arbitrio?

- | | | |
|-----|--|--------|
| 7.1 | Fra “certezze” ed “incertezze”: una soluzione impossibile? | p. 198 |
| 7.2 | Il <i>free will</i> e la <i>folk psychology</i> | p. 213 |
| 7.3 | Il libero arbitrio ed il ruolo della coscienza nella visione sperimentale del filosofo statunitense Eddy Nahmias | p. 221 |
| 7.4 | Una nuova fenomenologia per il libero arbitrio? | p. 228 |
| 7.5 | Prospettive future | p. 241 |

Conclusioni p. 247

Bibliografia p. 263

Introduzione

Come la parola stessa *introducĕre* suggerisce, il passaggio dal “non conoscere” al “conoscere” è “la finalità” delle prime pagine che l’ipotetico lettore incontra. Tra euforia, curiosità e, probabilmente, un po’ di scetticismo, l’emozione che ci avvolge quando apriamo un testo, sbirciando tra le pagine nomi, parole o frasi che possano colpirci, trova il suo *locus* finale nell’*Introduzione*.

Un passaggio difficile che vede protagonisti sia il lettore sia l’autore; un primo incontro in cui si confrontano menti, ambizioni e, soprattutto, aspettative: per l’uno di apprendere; per l’altro di “farsi apprendere”.

È con questo intento, seppur modesto per l’eventuale qualità del lavoro, che ci si accinge ad introdurre le pagine che seguiranno, tentando di render chiaro quel filo rosso che conduce dal “buio” alla “luce” o, per la sottoscritta, che l’ha condotta dalla “luce” al “buio”: un “buio” culturale che evidenzia (laddove non fosse già chiaro) quanto “l’amore per il sapere” sia quella spinta ad apprendere sempre e costantemente ciò che non si conosce. Socrate avrebbe detto *io so di non sapere* ed è con questa stessa umiltà che il “farsi apprendere” trova qui la sua *ratio*: nessuna presunzione e nessuna superbia, infatti, hanno accompagnato il percorso che si delineerà, ma solamente la *curiosità* è stata la compagna di viaggio. Curiosità di capire perché, in che modo e, soprattutto, quando.

Andiamo per ordine.

Il titolo *Empatia e libero arbitrio: un dialogo aperto fra neuroscienze, etica e psicologia* evidenzia, *in primis*, il contesto culturale entro cui gli studi che si leggeranno sono stati condotti; come sul finire dell’Ottocento il Filosofo era anche scienziato, così, oggi, la multidisciplinarietà del “sapere” è un requisito sempre più richiesto e necessario per gli studi che interessano ogni ambito della cultura.

Una settorializzazione esasperata, infatti, è causa di incomunicabilità: un *puzzle*, cioè, senza alcun aggancio. Se è vero che nel panorama culturale odierno si protende sempre più verso questa “globalità”, mantenendo ben salde le competenze di ognuno senza sopravanzare quelle dell’altro, è pur vero che per arrivare (di nuovo, oserei dire) a tale intento si sono dovute creare le condizioni che lo hanno reso possibile. In particolare, è un fatto noto ed oramai entrato nella memoria collettiva che negli ultimi decenni la tecnologia ha vissuto un

incremento delle sue potenzialità e ciò ha comportato uno sviluppo progressivo delle “scienze del cervello” o neuroscienze. Si hanno a disposizione mezzi in grado di constatare con una maggiore precisione, in alcuni casi, o di “vedere”, in altri, il cervello umano da una prospettiva diversa, essendo in grado di “intervenirvi” direttamente.

Sarebbe stato impensabile, ad esempio, qualche decennio fa, pensare che si sarebbe potuto osservare (con tutte le limitazioni del caso) “il cervello a lavoro” o ragionare su temi prettamente etici dai dati provenienti dalle neuroscienze; in questo senso è da intendersi la “rivoluzione concettuale” e culturale dell’epoca contemporanea.

Non si sta esprimendo un giudizio di valore, ma soltanto ponendo l’accento su un “fatto oggettivo” di ciò che sta succedendo nel nostro presente.

Una spinta in questa direzione è stata fornita, inconsapevolmente, dai casi di pazienti passati alla storia come *split brain*, ossia *cervello diviso*: sul finire degli anni Settanta, infatti, i neuroscienziati Michael Gazzaniga, Roger Sperry *et alii*, per bloccare gravi attacchi epilettici che rendevano la vita dei pazienti completamente disabilitante, recisero il corpo calloso fra i due emisferi cerebrali, convinti (come di fatto accadrà) che l’interruzione dei canali comunicativi fra questi ultimi avrebbe interrotto i relativi attacchi epilettici. Il lettore informato saprà certamente che i pazienti con *split brain*, divenuti vere e proprie “superstar” delle neuroscienze, hanno suscitato un intensissimo dibattito (ancora in corso) sullo *status* della coscienza: date le testimonianze pervenuteci, si lasciava intuire che in tali pazienti oltre ad esser divise le comunicazioni sinaptiche fra i due emisferi si fosse recisa anche la loro capacità di controllare loro stessi o di essere artefici dei propri atti. Sarebbe mancata, vale a dire, tutta quella sfera controversa (e per certi aspetti sfocata), che renderebbe l’essere vivente consapevole di se stesso.

Molteplici sono stati gli apporti teorici, le ipotesi, le discussioni che si sono succedute nell’ambito della *Filosofia della Mente* e non solo, dando origine ad una dicotomia, volendo semplificare al massimo il discorso, sulla natura della coscienza: è “materiale” o “immateriale”? Fondamentalmente, da come si sarà intuito, il problema di scelta fra le due *res* di cartesiana memoria è rimasto inalterato: cambiano i contesti, ma non i protagonisti.

Solo negli ultimissimi anni, nel 2017, lo psicologo statunitense Yair Pinto, operando una serie di *test* con i pochi pazienti *split brain* non decimati dal naturale corso della vita, avrebbe fornito dei dati la cui interpretazione andrebbe verso la smentita di quanto Gazzaniga *et alii*, da quasi quaranta anni, sostengono: l’unità della coscienza.

Come sarà facile intuire, tali studi stanno aprendo nuovi scenari e nuovi ordini di problemi; ma, al di là del dibattito che ne è scaturito, è bene notare che entro questo contesto generale i

protagonisti della scena restano sempre due: le neuroscienze, da una parte; la riflessione, dall'altra.

Ciò, allora, comincia a delineare il senso implicito del titolo in oggetto: essendo il presente un lavoro che pertiene alla riflessione morale (attorniata, come detto, dalla psicologia e dalle neuroscienze), la quale, per sua natura, non può essere avulsa dal contesto e dovendo noi trarre da questa un insegnamento valido ai fini della speculazione, non si poteva prescindere proprio da quel *background* culturale che ha in parte alimentato ed in parte contenuto il dibattito stesso. Un lavoro, cioè, che con una bibliografia recentissima ed in parte inedita in Italia (si consultino, a tal proposito, il paragrafo 3.3 sulla proposta, di Daniel Batson, i capitoli, rispettivamente, 5 e 8 sulle posizioni di Peter Carruthers ed Eddy Nahmias) tenta di inserirsi in un dibattito attuale, *in itinere*, per capirne le difficoltà, gli intenti, i problemi e le ragioni.

Porre l'accento sulle protagoniste dello "spettacolo", in questo caso ci si riferisce alle neuroscienze, non significa sposare *in toto* l'accezione più generale dell'eliminativismo o del riduzionismo in genere, secondo cui i fenomeni di un certo livello (la coscienza, la morale, la mente ecc.) sono spiegabili attraverso aspetti di base più semplici tali da far "scompare" i primi in vista di questi ultimi; il "riduzionismo cattivo", come è stato da molti definito, non è – ad avviso della scrivente – un approccio adeguato all'interrogazione del cervello umano, poiché è da considerare che il connubio fra Scienza e Filosofia non è mai stato scevro da diatribe; si dubita che proprio in questi anni, in questo "nuovo" contesto culturale esse possano "andare d'accordo".

Laddove, infatti, ciò dovesse accadere significherebbe l'appiattimento e la relativa scomparsa dell'una nell'altra e viceversa.

È pur vero, però, che la Scienza *tout court* ha il compito, per sua natura, di spiegare i fenomeni complessi, tentando di porre l'accento su quelli più semplici che li costituiscono e ciò non significa che questi ultimi, essendo più semplici, abbiano "meno realtà" di quelli più complessi: citando, ad esempio, gli oramai conosciutissimi *mirror neurons*, non è detto che il meccanismo specchio "sia" l'empatia o che, viceversa, l'empatia non abbia quel meccanismo che comporta la sua stessa esistenza.

Come è facile intuire, i tempi del *pensiero* sono diversi da quelli della *Scienza*, tanto che se il primo si interroga sulla legittimità dei cambiamenti straordinari che la seconda paventa, la seconda persegue il suo "ideale di scoperta" imperterrita: il problema diviene il prezzo da pagare.

Consapevoli del fatto che per natura stessa di un testo quest'ultimo debba rispettare dei crismi per la sua fruizione da parte di un ipotetico lettore e che, *ab origine*, qualsiasi autore opera delle scelte che reputa necessarie per il raggiungimento del fine propostosi, anche in questo caso il lavoro, per uno o per un altro verso, è frutto di cernite.

Veniamo alle scelte sull'empatia.

Da come si leggerà nelle pagine successive, parlare di *empathy* significa, per lo scenario qui brevemente esposto, porre in risalto un dialogo su un campo di indagine che pertenga tanto l'etica quanto le scienze del cervello. Si sta assistendo, però, dagli anni '90 in poi, ad una sorta di "euforia culturale" (se mi si consente l'espressione) che utilizza quale *jolly* l'importantissima scoperta, ad opera di Rizzolatti, Sinigaglia *et alii*, dei neuroni specchio per conferire una base neurale alla spiegazione di molteplici fenomeni: dai "meno tangibili" alle malattie meno conosciute.

Dall'empatia all'autismo i *mirror neurons* sono additati come *la* risposta a qualsiasi problema.

Ciò, per molteplici ragioni, non può essere credibile.

Quando si studia un fenomeno è dovere dello studioso accorto interrogare il fenomeno stesso sulla sua possibile identità: capirne la natura, le implicazioni, l'*idea* (per utilizzare un linguaggio platonico) diviene un passo necessario per poter sciogliere, un domani, il dubbio dal quale era partita l'analisi. Se ciò, quindi, pertiene "la buona norma" di qualsiasi ricerca, per la mia personale esperienza è stato complicatissimo capire la natura dell'empatia per *la soluzione*, avanzata da molti studiosi, sulla sua natura: "*empathy* è il sistema *mirror*", non mi convinceva. Non è stato, poi, un particolare testo anziché un altro a farmi trovare quel filo rosso che mi ha, successivamente, permesso di scrivere le pagine che seguono, ma una faticosa costruzione, tassello dopo tassello, di assemblaggio del "mosaico empatia" a farmi capire che, oggi, quando si parla di questo fenomeno così controverso si affrontano, parallelamente, due diversi livelli di indagine: da una parte, quello neuroscientifico; dall'altra, quello etico.

Il primo, come detto, considera l'apporto delle scienze del cervello quale mezzo primario per avviare la naturalizzazione cognitiva e morale dell'essere umano di cui, come detto, la neuroetica è la massima esemplificazione ed è in questo particolare aspetto che il *jolly specchio* è utilizzato in maniera non proprio consona: non sto, qui, negando che i *mirror neurons* non siano *parte* o tassello di un quadro ben più complesso di spiegazione neurale del funzionamento cerebrale, ma sto sostenendo che giudizi *tranchant* ed apodittici su un campo di indagine che si trova solamente agli esordi non costituiscono prova di onestà

intellettuale. Da un punto di vista neuroscientifico, infatti, i gravi limiti strutturali delle tecnologie a disposizione, oggi, non consentono un approccio consapevole ed olistico al “problema cervello” tale da poterlo *spiegare* in maniera evidente così come si farebbe se si mostrasse ad un viandante la cartina geografica del luogo in cui trova; è con tale prudenza, allora, che i neuroscienziati debbono fare i neuroscienziati ed i filosofi fare i filosofi.

Un dialogo proficuo si può costruire solo se si ha rispetto delle competenze di ognuno degli addetti ai lavori senza confonderle o miscelarle, poiché in gioco c'è un prezzo enorme da pagare: la conoscenza della natura umana.

È stato questo “scivolamento” di campi, questo mare increspato, che mi ha impedito, all'inizio, di “capire” ed è proprio questa “zona grigia” che (mi sento di sostenere) mischia le carte, dando a molti l'illusione di aver trovato l'enigma al problema.

L'empatia, quindi, *non* è il *sistema specchio*, ma è *anche* questo.

Il secondo livello di indagine è, come detto, quello prettamente etico.

È su questa bidimensionalità di piani che ho operato una prima scelta per la costruzione del già citato filo rosso.

Non essendo una neuroscienziata e non potendo fornire elementi nuovi alla ricerca in oggetto, ho preferito non cimentarmi nella riproposizione delle continue scoperte che si stanno succedendo, giorno dopo giorno, delle aree cerebrali che potrebbero interessare la spiegazione dell'empatia di base, poiché tale scelta avrebbe reso poco originale il lavoro. Ho preferito, invece, dare un taglio al problema prettamente etico.

Certa, infatti, che non avrei potuto dare un apporto originale alla spiegazione dell'empatia primaria ed altrettanto certa che ciò avrebbe significato annoiare il lettore con un'ennesima ricerca che trattasse del rapporto *neuroscienze/empatia*, ripetendo nozioni già note al grande pubblico, l'intenzione è stata, allora, concentrarsi su ciò che rende controversa, intricata e problematica la natura dell'empatia secondaria o cognitiva per mezzo della quale un ipotetico osservatore sarebbe in grado di *leggere la mente* di un'altrettanta ipotetica vittima. Questo aspetto è stato il *focus* del mio interesse, poiché, come si spiegherà successivamente, ha fornito l'*incipit* per avviare uno studio nuovo, forse, originale, sia sulla natura dell'empatia sia sul rapporto fra quest'ultima ed il libero arbitrio.

Un discorso diverso, invece, va fatto per il ricorso alla spiegazione dell'uso delle neuroscienze nella seconda parte del lavoro, riguardante il libero arbitrio.

Stante le precauzioni sopra descritte sul rispetto delle competenze, degli ambiti di ricerca e delle limitazioni strutturali delle medesime scienze del cervello, diverso è stato interfacciarsi con il problema del *free will*, partendo da un'ottica neuroscientifica.

La letteratura su questo particolare tema trova una minore popolarità fra il senso comune. Infatti, ad oggi sono pochissimi i tentativi che si sono succeduti per fornire delle linee guida sulla naturalizzazione della scelta: Benjamin Libet, agli inizi degli anni Ottanta del XX secolo, e Chun Siong Soon *et alii*, nei primissimi anni del XXI, sono, certamente, i due “pionieri simbolo” che hanno tentato di esplorare un campo ancora ben lontano dall’esser chiaro.

In questo caso, allora, è stato necessario confrontarsi con le teorie, i dubbi e le perplessità che hanno accompagnato (ed accompagnano tuttora) il dibattito sul libero arbitrio, poiché influenzato, ancora oggi, da quei dati ricavati dai due neuroscienziati con i quali, volente o nolente, è doveroso confrontarsi.

A questo punto, diviene necessaria un’ulteriore precisazione.

Al lettore non sarà passata inosservata la constatazione che nel titolo del testo sono presenti due fenomeni: empatia e libero arbitrio; che, sia nella tradizione classica sia in quella più contemporanea, sono sempre presentati distinti.

L’intento della scrivente si basa proprio su tale aspetto: convinta che l’empatia non sia ciò che, comunemente, la cultura odierna denota, ha tentato di darne una spiegazione, partendo dal concetto di libero arbitrio. L’ipotesi, cioè, è che ciò che si intende, erroneamente, per *empathy*, “altro non sarebbe” che una manifestazione particolare di un concetto diverso, il *free will*, appunto, con cui l’agente *sceglierebbe*, consapevolmente, di “calarsi nei panni altrui”, operando, quindi, uno sforzo cognitivo di ciò che viene definita *lettura della mente*. Se ciò fosse vero, si avrebbero, necessariamente due conseguenze:

1. I fenomeni conosciuti come “empatia” (*empathy accuracy, co-cognition* etc.) denoterebbero un denotato inesistente: parlerebbero, vale a dire, di ciò che non esiste, poiché non sarebbe la capacità emotiva la fonte della medesima empatia, ma tutt’altro;
2. Bisognerebbe ridefinire il concetto di empatia o eliminandolo *in toto* o essendo attenti a chiarire che quando si parla di *empathy* si indicano solamente quegli aspetti “di base”, neurobiologici, del funzionamento cerebrale che renderebbero abile l’uomo ad interfacciarsi con la realtà circostante.

Sono conscia, come già sostenuto altrove¹, che Roma non fu costruita in un solo giorno e che ciò che mi propongo di raggiungere è la costruzione di un “Impero” di cui, al momento, credo si siano delineate, aleatoriamente, neppur le fondamenta, ma ho creduto necessario, per *amor del sapere*, esporre ciò che ho appreso o, meglio, l’interpretazione che ne ho ricavato, affinché

¹ FEDERICA MADONNA, *Considerations about cognitive empathy*, in “American Journal of Humanities and Social Science Research”, Vol. 3, Issue 4, pp. 34-40.

possa io stessa confrontarmi e costruire un dialogo proficuo con chi avrà il desiderio, la pazienza di “leggermi”, *in primis*; di dialogare, *in secundis*.

Con tale lente di ingrandimento, allora, si dovrà vedere l'accostamento dell'empatia e del libero arbitrio in un contesto culturale, la neuroetica, che mi fornisce, funzionalmente, gli strumenti concettuali e neuroscientifici per avviare un discorso unanime; alla luce ulteriore della convinzione che, come detto, la *Conoscenza* è sempre e solo una.

Il percorso che si delineerà nelle pagine seguenti è stato il frutto di un “atto empatico”, volendo restare in tema, poiché, calandomi nei panni dell'ipotetico lettore, ho creduto opportuno illustrare ciò che verrà, seguendo un percorso che andasse dal generale al particolare. In quest'ottica, il primo capitolo *La neuroetica: il quadro di riferimento attuale della ricerca* indicherà (qui solamente accennate) le linee guida del nascente quadro concettuale che, nella sua faticosa lotta per l'affermazione tra le scienze universalmente riconosciute come tali, ha lo scopo di fornire gli strumenti entro cui l'analisi in oggetto si colloca. Non è un caso, infatti, che si sia posto al di fuori delle due sezioni successive (Parte II. *L'empatia cognitiva* e Parte III. *Il libero arbitrio*) quale sezione a sé, affinché, una volta delineata la mappa, il lettore sia, poi, libero di muoversi per la meta desiderata.

Nella seconda parte, proseguendo, si delineerà quel “mosaico empatia” che, con non poche difficoltà, ho assemblato a partire dai suoi innumerevoli tasselli: da un'analisi storica del concetto all'oggetto di discussione attualissimo di un eventuale rapporto fra *mindreading/metacognition*, non tralasciando le ripercussioni sul legame che legherebbe l'empatia stessa all'altruismo quale motivazione prosociale al comportamento morale. Si spera che alla fine della lettura risulteranno più chiari quei dubbi e quelle incredulità che nascono, inevitabilmente, quando ci si accinge a capire il fenomeno empatico.

Anche in questo aspetto, il terzo protagonista del presente lavoro ha costituito un asse centrale alla chiarificazione concettuale delle tematiche affrontate: la *Psicologia*. L'utilizzo del modello di Daniel Batson, ad esempio, per mostrare l'alternativa a quello proposto da Martin Hoffman sull'intricato rapporto empatia/altruismo è stato un mezzo necessario per il raggiungimento del fine.

Si tiene a precisare che le fonti usate nella stesura, sia del primo capitolo sia delle due sezioni ulteriori, sono state per la maggior parte di origine anglosassone/analitica e, soprattutto, ai più sconosciute in Italia.

È il caso, come detto, di Batson, del rapporto fra *metacognition/mindreading*, esemplificato in questa sede dalla figura del filosofo statunitense Peter Carruthers e della “filosofia

sperimentale” del suo collega Eddy Nahmias, il quale, occupandosi del problema *free will*, giunge alla “costruzione” di un metodo di indagine che lui stesso definisce “psico-filosofico”. Sarà così solo al termine di ciò che si è ritenuto necessario illustrare nell’ipotetica scala a pioli del “problema empatia” che si giungerà alle *Conclusioni*, che, forse, paradossalmente, hanno costituito il punto di partenza per reinterpretare quanto appreso sul problema, spiegandone motivazioni e ragioni.

Ciò ha consentito di avviare un discorso inter-disciplinare legato non soltanto al rispetto delle “barriere invisibili” che circondano, oramai, la Cultura dal Novecento in poi, ma un dialogo quanto più onnicomprensivo per la spiegazione di quello che sembra apparire un unico fenomeno, il *freempathy* (quale unione fra libero arbitrio ed empatia cognitiva) per cui, un giorno, credo, giungerò a sostenere che “l’empatia non esiste”.

Adesso mera convinzione nella mia testa e che mi auguro sarà oggetto di una trattazione futura nel mio continuo e, spero futuro, *iter* di crescita.

Parte I

La neuroetica

Capitolo I

Il quadro di riferimento attuale della ricerca

Quando nel pieno della notte, il giovane Kenneth Parks, negli Stati Uniti, guidò per circa ventitré chilometri, raggiungendo la casa dei suoceri e li accoltellò entrambi, non negò le accuse rivoltegli dalla Giustizia americana, ma si dichiarò innocente per “non aver commesso *consciamente* il fatto”, poiché in stato di sonnambulismo². Parks, subito dopo il processo, fu assolto: “incapace di intendere e di volere”, dichiarò la Giuria; ma, fatto sta, che per come andarono i fatti nella notte del 23 maggio 1987, il caso suscitò non poco interesse.

L'imputato confessò di non ricordare nulla di quando si mise in macchina, di come arrivò a casa dei suoceri, di come eseguì quegli omicidi efferati; sostenne solo che, dopo l'accadimento, si recò alla stazione di Polizia più vicina, confessando che aveva la sensazione di aver compiuto “qualcosa” di orribile. In quell'occasione, infatti, si rese conto di avere delle ferite profonde alle braccia (tendini recisi, gli diagnosticarono, successivamente, al Pronto Soccorso) e di esser completamente sporco di sangue.

La responsabilità di Parks, giuridicamente parlando, fu dichiarata nulla, tanto che, come detto, fu assolto; ma eticamente il caso non poteva passare inosservato.

Ciò contribuì ad accentuare dubbi e perplessità sul significato di “azione responsabile”; infatti, quali sono o potrebbero essere le condizioni in cui un individuo possa ritenersi tale? E quando un atto può esser definito “volontario”?

² Il sonnambulismo è classificato quale disturbo del sonno (classificato, a sua volta, quale *parasomnia del sonno MREM* – manifestazioni motorie complesse indesiderate – o disturbi dell'*arousal*) che può presentarsi, generalmente, entro due ore dall'addormentamento. I comportamenti messi in atto dal sonnambulo possono essere fra i più semplici (sedersi sul letto, compiere gesti inconsulti) o fra i più complessi, come nel caso di Parks, che diventano pericolosi per chi li compie. Il sonnambulo, infatti, è, in tali circostanze, completamente incosciente e, soprattutto, la sua particolarità è l'amnesia che accompagnerà, al risveglio, l'attività messa in essere durante la notte; di fatto egli sta dormendo e continua a dormire. A ciò si imputa l'amnesia al risveglio

Inutile dire che sul concetto di responsabilità morale esiste un'intera tradizione di pensiero che, per certi aspetti o per certi altri, pone l'accento o sull'"individuo" o sull'"ambiente circostante", ma ciò che accese l'interesse e la curiosità del caso furono le sue modalità; al centro del dibattito, infatti, non verteva più la storica diatriba fra necessitazione causale e contingenza casuale, ma il *quid* dell'uomo nella sua interezza: la coscienza.

Dopo che per millenni i filosofi hanno dibattuto sul rapporto *logos*-materia, prima, mente-cervello, poi, il caso di Parks sarebbe potuto esser interpretato quale "prova provata" della sovversione del classico problema dualista per essere indirizzato ad un nuovo ed innovativo approccio alla mente umana: neuroscientifica.

Un insieme favorevole di contingenze hanno permesso agli scienziati, sul finire degli anni Ottanta, di "affacciarsi" allo studio del cervello con tecniche e tecnologie in continuo progresso per tentare di svelare il "mistero cervello", fornendo alle neuroscienze un'ascesa costante e continua nel panorama culturale odierno. Si faceva strada, cioè, un modo nuovo di indagare i problemi dell'umano fino a quel momento competenza della sola speculazione scientifica, lasciando presagire – cosa che, di fatto, avverrà – una interazione fra etica e neuroscienze resa possibile, da una parte, dal miglioramento tecnologico; dall'altra, dalla nascita delle neuroscienze cognitive, in grado di mettere in relazione l'attività neuronale con il pensiero ed il comportamento³; come sarà facile intuire, ne sono scaturite nuove visioni del mentale e, soprattutto, della persona in quanto tale.

L'idea di indagare scientificamente l'etica, chiedendosi, cioè, se la morale posseda una base neurale, è da far risalire all'incidente che coinvolse l'operaio del Vermont, Phineas Gage, nel 1848, allorché una barra di ferro

[dal peso di, n.d.r.] *6 chilogrammi*; [...] *lung[a, n.d.r.] 110 centimetri* e [con, n.d.r.] *un diametro di poco più di 3 centimetri*⁴,

gli trapassò da parte a parte il cranio, entrando all'altezza della guancia destra e fuoriuscendone alla sua sommità. L'eccezionalità dell'incidente fu dovuta al fatto che non solo Gage sopravvisse, ma conservò intatte le sue funzioni cognitive: riusciva ad interloquire, a comprendere il linguaggio, a ricordare, a camminare...tuttavia, come scrisse il medico che lo ebbe in cura, era come se

³ ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Neuroetica*, Il Mulino, Bologna, 2011, p. 9.

⁴ JOHN HARLOW, *Recovery from the passage of an iron through the head*, in "Publications of the Massachusetts Medical Society", 2, 1868, pp. 327-347 in ANTONIO DAMASIO, *L'errore di Cartesio. Emozione, ragione e cervello umano*, Adelphi Edizioni, Milano, 1995, p. 36.

*l'equilibrio, per così dire, tra la sua facoltà intellettuale e le sue disposizioni animali era stato distrutto*⁵.

Gage, cioè, non era più Gage. Divenne

*bizzarro, insolente, capace a volte delle più grossolane imprecazioni, da cui in precedenza era stato completamente alieno; poco riguardoso nei confronti dei compagni; insofferente di vincoli o consigli che contrastassero con i suoi desideri; a volte tenacemente ostinato, e però capriccioso e oscillante; sempre pronto a elaborare molti programmi di attività future che abbandonava non appena li aveva delineati...Un bambino, nelle sue manifestazioni e capacità intellettuali, ma con le passioni animali di un adulto robusto*⁶,

in netto contrasto con la personalità moderata ed intellettivamente apprezzabile che aveva sempre esibito prima dell'incidente. Come scrive Antonio Damasio:

*Senza volerlo, l'esempio di Gage indicava che qualcosa nel cervello aveva a che fare specificamente con proprietà peculiarmente umane, tra cui la capacità di anticipare il futuro e di pianificare in accordo con tale anticipazione, all'interno di un ambiente sociale complesso; il senso di responsabilità verso se stessi e verso gli altri; la capacità di predisporre la propria sopravvivenza in modo deliberato, in ottemperanza al proprio libero volere*⁷.

Quando agli inizi degli anni Novanta del secolo scorso Damasio *et alii* studiarono il cranio di Gage, ne imputarono la dissociazione, ossia la discrepanza fra le credenze e le convinzioni prima dell'incidente ed il comportamento manifesto dopo, alla distruzione dell'area prefrontale nei suoi settori orbitari e mediani: ulteriori studi su pazienti che presentavano tumori localizzati nella stessa area cerebrale hanno fatto supporre, sia al *team* di Damasio sia al mondo scientifico in genere, che essa possa rappresentare, probabilmente, la “sede” della morale umana⁸.

⁵ *IDEM*, p. 37.

⁶ *IVI*.

⁷ ANTONIO DAMASIO, *L'errore di Cartesio*, op. cit., p. 40.

⁸ *IDEM*.

Si dovrà attendere il 2002, quando, in un convegno internazionale, la filosofa Adina Roskies proporrà una “etichetta” che raccoglierà, immediatamente, questo fervente clima culturale, interessato dalle nuove tecnologie sullo studio del cervello ed i classici problemi etici sull’uomo e sul loro utilizzo, sintetizzabile nel neologismo *neuroetica*. Il tutto lasciava presagire un dialogo aperto e sempre più frequente tra la classe dei problemi etici e le odierne neuroscienze; un connubio, vale a dire, fra etica e scienza che non celava, a sua volta, la difficoltà interpretativa del costituendo approccio, esplicabile, da una parte, nell’*etica delle neuroscienze*, che si sarebbe occupata, al pari della bioetica, della legittimità delle tecniche neuroscientifiche nella sperimentazione sull’uomo; dall’altra, nelle *neuroscienze dell’etica*, che, sulla scia dell’incidente di Gage, avrebbero indagato le basi neurali della morale. Come sarà facile immaginare, l’impatto dello studio diretto del cervello sta comportando la nascita e la rimodulazione dei classici problemi etici con prospettive e paradigmi a volte diversi, altre volte ribaltati:

La duale separatezza fra mente e corpo [ad esempio, n.d.r.] ha determinato l'emergenza di nuove riflessioni etiche. Lo svilupparsi delle conoscenze relative al sostrato biologico dei comportamenti dell'essere umano ha richiesto una metamorfosi nella percezione del corpo e della mente⁹.

Difficilmente si sarebbe pensato di porre, in etica, il problema della *privacy cerebrale*¹⁰: la riflessione secondo cui ogni individuo ha diritto di custodire o potenziare le proprie capacità intellettive o, più in generale, le facoltà cognitive. Sapere, infatti, come funziona il cervello per determinati compiti (seppur, si badi bene, la “conoscenza” è lungi dall’esser tale), ha portato nei *college* statunitensi all’utilizzo di un medicinale, il *Ritalin*, in grado di aumentare le prestazioni cognitive e, di conseguenza, raggiungere risultati migliori rispetto alla propria *baseline* “naturale”. Nello specifico, l’acido gamma-aminobutirrico (GABA), migliora la plasticità neuronale e, quindi, l’apprendimento. Alti livelli di GABA inibiscono al cervello la sua “fissità”, favorendo la nascita di nuove sinapsi e, in certi casi, di nuove abilità. Su tale scia, il neuroscienziato Allan Snyder, addirittura, ha ipotizzato (e nulla ci vieta di pensare che un giorno sarà praticabile) di rendere possibile il potenziamento neurobiologico attraverso la

⁹ FABRIZIO RUFO, ANDREA BORRI, *Il cervello non è più quello di una volta*, in ANDREA CERRONI, FABRIZIO RUFO (A CURA DI), *Neuroetica tra etica e società*, Utet, Torino, 2013, p. 125.

¹⁰ Cfr. JULIAN SAVULESCU E NICK BOSTROM (a cura di), *Human Enhancement*, Oxford University Press, NY, 2009.

stimolazione trans-cranica con la (probabile) conseguenza di sfruttare al massimo le nostre capacità cerebrali¹¹: $10^7 \times 100.000.000.000$ di interconnessioni intramodulari o intraneuroniche¹². Una manipolazione diretta del cervello che, sulla scia dell'accezione *etica delle neuroscienze*, pone riflessioni urgenti: è opportuno, è lecito, migliorare il proprio Sé? È legittimo ricorrere al potenziamento e non al trattamento delle facoltà cognitive? Che cosa comporterebbe, tutto ciò, a livello di autenticità del proprio Io?

Se per *potenziamento* si intende il rafforzamento, la valorizzazione o l'incremento delle proprie capacità finanziarie, economiche o, in questo caso, cognitive e per *trattamento*, di contro, si indica l'applicazione di un prodotto o corpo fisico a diversi processi per conseguire determinati effetti, allora si capiscono i motivi di tanta perplessità sull'argomento: che conseguenze si avrebbero per il soggetto sottoposto a potenziamento per la conoscenza "originale" che avrebbe su di sé laddove non fosse passibile di tali mezzi artificiali? I dubbi sorgono sui loro scopi, lontani dal voler tangere "la verità" dell'individuo che si troverebbe, innanzitutto, a vivere inautenticamente se stesso; lontano dalla conoscenza di Sé e dall'immagine consuetudinaria – costruitasi faticosamente nel lungo processo di formazione individuale – con cui l'Io ha da sempre convissuto.

Oltre ai controversi e delicati problemi posti, solo per citarne i più evidenti, non sarebbero da sottovalutare quelli relativi alle metodologie d'uso: chi dovrebbe essere responsabile della somministrazione dei mezzi per migliorare le capacità cognitive? Lo Stato, l'individuo stesso? Se si lasciasse ad arbitrio del singolo ci si troverebbe, *sine dubio*, di fronte ad una consistente disparità sociale: ipotizzando, infatti, che tali mezzi abbiano un costo elevato, solo i facoltosi o i figli di facoltosi potrebbero accedervi. La distribuzione delle risorse, vale a dire, non sarebbe mai equanime.

Si tornerebbe al classico problema di Rawls sulla costruzione di una società equa: sarebbe, stavolta, sufficiente, il *velo d'ignoranza*? Quale sarebbe, ad esempio, la posizione originaria del "giudicante"?

Non è insolito, infatti, leggere fra gli innumerevoli volumi editi dal 2002 ad oggi¹³ tesi che evidenziano con vigore preoccupazioni e limiti della nascente neuroetica, ma, a tal proposito, non si dovrebbero dimenticare le parole di Adina Roskies al riguardo, con particolare attenzione e cura al problema della libertà:

¹¹ NEIL LEVY, *Neuroetica. Le basi neurologiche del senso morale*, Apogeo Editore, Milano, 2013, p.97.

¹² FABRIZIO RUFO, ANDREA BORRI, *Il cervello non è più quello di una volta*, op. cit., p. 123.

¹³ Si rimanda il lettore al capitolo II.

Recent developments in neuroscience raise the worry that understanding how brains cause behavior will undermine our views about free will and, consequently, about moral responsibility. The potential ethical consequences of such a result are sweeping. I provide three reasons to think that these worries seemingly inspired by neuroscience are misplaced. First, problems for common-sense notions of freedom exist independently of neuroscientific advances. Second, neuroscience is not in a position to undermine our intuitive notions. Third, recent empirical studies suggest that even if people do misconstrue neuroscientific results as relevant to our notion of freedom, our judgments of moral responsibility will remain largely unaffected. These considerations suggest that neuroethical concerns about challenges to our conception of freedom are misguided¹⁴.

Si deve tener presente, infatti, che ci si trova di fronte alla *naturalizzazione della morale*: quel programma di ricerca che vuole ricondurre l'etica alle sue basi neurobiologiche, portando l'uomo ad essere lui stesso oggetto di studio delle scienze naturali. A tal proposito, si dovrà esser pronti a:

1. *Studiare tutti i fenomeni umani con gli stessi concetti della Scienza tout court;*
2. *Studiare tutti i fenomeni umani attraverso le leggi della fisica, dando per scontato che qualsiasi evento abbia una causa ed un effetto (ciò potrebbe essere problematico per alcuni ordini di problemi esaminati dalla stessa filosofia morale come, ad esempio, il concetto di scelta, di libero arbitrio o della stessa mente);*
3. *Saper ricostruire nel tempo, artificialmente, tutti i fenomeni emergenti dell'essere umano¹⁵,*

sostenendo, infine, che l'uomo è per sua natura portato all'agire morale.

L'utilizzo di speciali tecniche di indagine, quali la risonanza magnetica funzionale (fMRI), la *brain imaging* o la tomografia a emissione di positroni (PET), permettono, presumibilmente, diagnosi sempre più precise rispetto alle sole ipotesi concettuali, in grado di migliorare (ma non di stravolgere) i temi dell'etica classicamente dibattuti.

¹⁴ADINA ROSKIES, *Neuroscientific challenges to free will and responsibility*, in "Trends in Cognitive Sciences", Vol. 10, n. 9, pp. 419-423, p. 419.

¹⁵ Cfr. ANDREA LAVAZZA, *Neuroscienze e filosofia morale*, in "Rivista di Filosofia", Vol. XCVIII, n. 3, dicembre 2007, pp. 327-358.

In questo senso, allora, l'accezione delle *neuroscienze dell'etica* apre quegli spiragli conoscitivi, interdisciplinari, che prevedono, per certi aspetti, non tanto una "ristrutturazione del sapere", ma una rielaborazione epistemologica del sapere stesso:

Perciò il terreno specifico della neuroetica dovrebbe attenersi alla riflessione circa ciò che apprendiamo su noi stessi e il nostro 'funzionamento', grazie, principalmente, (ma non esclusivamente) alle neuroscienze. In altre parole, è la naturalizzazione forte dell'indagine sull'essere umano a rendere pertinente una metadisciplina che si occupi dell'ambito multidisciplinare descritto. A essere oggetto di studio per il suo carattere nuovo e controverso, quindi, non sarebbe ciò che pensiamo di fare, ma ciò che sappiamo o che crediamo attendibilmente di sapere¹⁶.

Il problema, allora, si pone proprio su ciò che "crediamo attendibilmente di sapere": in che modo le neuroscienze potranno o possono cambiare la conoscenza che abbiamo su noi stessi? In svariati modi, in realtà.

I pazienti prosopagnostici sono coloro i quali non sono in grado di riconoscere i volti, neppure quelli dei familiari, a meno che non li sentano parlare; al contrario di coloro affetti da *Il delirio di Capgras*, i quali si rendono immediatamente conto della familiarità del volto, ma, sebbene sappiano che assomiglia *esattamente* ad un familiare, restano convinti non solo che la persona di fronte non sia chi dica di essere, ma la considerano un impostore. Recenti studi sostengono che in entrambi i casi menzionati ad essere responsabile delle patologie sia un funzionamento anomalo del sistema autonomico; un insieme di meccanismi di controllo in grado di mantenere regolare l'omeostasi del corpo. La doppia dissociazione fra questo sistema ed il riconoscimento dei volti provoca, in tali pazienti, la consapevolezza di un'esperienza diversa fra la visione di un familiare e di un estraneo, ma, neurobiologicamente, essi sono incapaci di capire chi si trova loro di fronte.

È facile, allora, condividere coloro i quali sostengono che la neuroetica sta facendo nascere un'immagine del tutto diversa del Sé rispetto alla tradizione ed è con altrettanta semplicità che non ci si dovrebbe stupire di fronte ad una *riqualificazione* del concetto stesso di *persona*, anche se, contrariamente alle aspettative, tale aspetto risulta essere il grave assente nel dibattito.

¹⁶ ANDREA LAVAZZA, *Neuroetica*, op. cit., p. 26.

La distinzione fra persone¹⁷ e cose del mondo è una riflessione che attraversa l'intera tradizione del pensiero: non per questo, infatti, la neuroetica pone nuovi dilemmi, ma, in questo caso, contribuisce a vecchi intenti con nuove difficoltà: quelle delle scienze del cervello.

Il problema che si pone, allora, è scegliere se tentare di naturalizzare anche il concetto di “persona” oppure lasciare la riflessione al di fuori delle nascenti tendenze. A tal proposito, il dibattito acceso sull'*American Journal of Bioethics* dal saggio di due autrici, Martha J. Farah e Andrea Heberlein¹⁸, evidenzia proprio il *punctum dolens* del ragionamento: o si tenta di individuare caratteristiche fisiche o neurobiologiche che costituiscono la *persona* quale entità fisica che agisce ed interagisce nel mondo (naturalizzandola) oppure si tenta il suo definitivo abbandono:

Quando un termine ha troppi significati e quando i suoi significati tecnici si mescolano con quelli quotidiani diventa un problema usarlo. Ed è esattamente questo che sta accadendo con il termine 'persona', che viene utilizzato con significati non solo diversi ma addirittura opposti [...] Meglio, allora, abbandonare questo termine, essendo quello di 'individuo umano' del tutto sufficiente (oltre che biologicamente fondato): eliminiamo dai discorsi filosofici (e quindi anche etici e bioetici) il termine persona e così elimineremo anche i fastidiosi dibattiti e chiacchiericci intorno alla questione se qualcuno o qualcosa sia persona o meno¹⁹.

Secondo le due autrici, infatti, il contributo che le neuroscienze possono dare al dibattito è la consapevolezza che la categoria “persona” è un costrutto derivante dalla nostra neurobiologia; un costrutto del cervello, cioè, applicato al mondo, tramite il quale, istintivamente, “cabliamo” il nostro relazionarci all'altro: sia esso oggetto o essere vivente dotato di valore. Una sorta di *person network*, ovvero di un sistema neurale capace di distinguere un conspecifico, che “regolerebbe”, immediatamente, il comportamento da tenere verso l'altro; una risposta evolutiva in grado di indirizzare le cure e la sopravvivenza di un conspecifico all'interno di un

¹⁷Scelta arbitraria di ricondurre il discorso, sennò troppo lungo, al solo *Homo Sapiens Sapiens*, pur condividendo pienamente l'assunto della intraspecificità del concetto stesso: *persone* sono tutti quegli *animali umani e non umani* che possiedono capacità di scelta, volizioni, desideri e son capaci di gestire se stessi ed altri secondo propri fini e scopi.

¹⁸ MARTHA J. FARAH E ANDREA HEBERLEIN, *Personhood and Neuroscience: Naturalizing or Nihilating?*, in “American Journal of Bioethics, Neuroscience”, Vol. 7, 2007, pp. 37-48.

¹⁹ GIOVANNI BONIOLO, *Dalla 'persona' all' 'individuo': una soluzione filosofica a partire dal fondamento biologico*, in GIOVANNI BONIOLO, GABRIELE DE ANNA E UMBERTO VINCENTI (a cura di), *Individuo e persona. Tre saggi su chi siamo*, Bompiani, Milano, 2007, p. 15.

gruppo sociale, il quale, a sua volta ed in termini utilitaristici, contribuirebbe alla distinzione fra “persone” e “non persone” con ciò che, eticamente, ne conseguirebbe.

Non si tratta, a dire il vero, di un approccio riduzionista alla questione, ma di una reinterpretazione del dubbio cartesiano (*Che cosa sono, dunque, io?*²⁰) in chiave neuroscientifica o, per meglio dire, neuroetica.

È su questa scia che le neuroscienze hanno tentato di occuparsi, già sul finire degli anni Ottanta, di aspetti speculativi epistemologici o metafisici di temi dibattuti, classicamente, in ambito filosofico che sembravano lontani da qualsiasi “materialità” presunta o ipotizzata del *quid* umano; il libero arbitrio, ad esempio, è uno di questi.

Il dilemma di sempre, che si alternava fra posizioni fra loro contrastanti o alle volte complementari, tra la reale o fittizia esistenza della libertà nell’uomo, è stato affrontato alla luce dell’approccio *neuroscientifico all’etica* per poi ricondurre il discorso alla sola etica *tout court*: l’avvento delle “scienze della mente”, infatti, ha portato alla realizzazione di veri e propri esperimenti, curati in laboratorio, in cui la grande indagata è proprio la volizione umana. Si è assistito, collateralmente, ad un fenomeno di *follow up* (come oggi si usa dire): dalla coda di una tradizione del pensiero che si è interessata al problema della libertà, si parte, nuovamente, per tornare “indietro”; ma stavolta non ai grandi pensatori, ma all’uomo stesso, introiettando, se mi si concede l’espressione, il problema.

Sembrerebbero, oramai, marginali i dubbi e le perplessità relativi alla metafisica del mondo (determinato o indeterminato?) per porre l’accento sull’interrogativo se l’uomo, quale essere vivente, sia dotato di meccanismi cerebrali tali che la loro attivazione comportino una scelta anziché un’altra. Diviene, oramai, una “questione cerebrale” in cui la speculazione etica, apparentemente, sottostà alla scienza: non più, come nella tradizione medievale, *Philosophia ancilla Theologiae*, ma *Philosophia ancilla Scientiae*.

Ciò ha consentito, forse per un falso principio di autorità, a molti neuroscienziati di proclamare la loro “verità” su temi cogenti, pertinenti all’uomo in quanto tale, come ad esempio il libero arbitrio, la coscienza, l’intersoggettività. Non è insolito, infatti, leggere sulle testate scientifiche statunitensi titoli allarmisti del tipo: “I neuroscienziati hanno scoperto che non siamo liberi”; oppure “È tutta colpa dei geni”; e simili, dando a pensare al vasto pubblico eterogeneo che, effettivamente, non siamo liberi per cui responsabili delle nostre azioni; un’equazione che, per effetto domino, dovrebbe essere applicata anche ai *serial killer* ed ai malfattori più comuni. I colpevoli, seguendo tale ragionamento, non sarebbero più quegli individui che commettono un illecito giuridico, ma il loro patrimonio genetico. Tale approccio

²⁰ Cfr. RENÈ. DESCARTES, *Meditazioni Metafisiche*, Edizioni Laterza, Roma-Bari, 2016.

comporterebbe, di conseguenza, che una serie di altre questioni riguardanti il diritto, l'economia, la sociologia, l'antropologia etc. renderebbe urgentissima una loro rivisitazione completa tanto che sembrerebbe (quasi) che tali scienze dovrebbero “essere scritte” daccapo. In tutto ciò, usando un'espressione in una accezione negativa tipica di questi tempi, i *filosofi*, “colpevoli” di vivere i “tempi morti” della speculazione e non potendo (apparentemente) fornire un *utile* contributo alla società, *devono rimanere in poltrona*.

Tale atteggiamento, come si comprende, rende ancora più urgente una riflessione che non può essere taciuta.

Se il pregio delle *neuroscienze dell'etica* è stato quello di *suggerire* un'impostazione diversa, sicuramente nuova, dei classici problemi etici, dall'altra, però, ha portato ad una confusione di ruoli e competenze inaudita.

Socrate soleva dire: *kánete ó, ti boreíte na kánete* (“fai ciò che è di tua competenza fare”) ed è una lezione che molto spesso si dimentica.

Quando i neuroscienziati o gli scienziati, in genere, proclamano la “verità” sui temi cari al *quid* umano lo fanno travalicando le competenze che gli competono: essi hanno l'onore (d'altronde è il loro mestiere) di capire, neurobiologicamente, come funziona il cervello, ma per ciò che concerne l'interpretazione dei dati ed il modo in cui essi vengano “inglobati” all'interno del dibattito etico, epistemologico... è una competenza che non spetta loro, ma “ai cari e vecchi” filosofi.

Ecco allora spiegate le parole che, poco più su, si riportavano della Roskies: seppure dovesse uscire dalla scienza che il libero arbitrio non esiste, ciò non comporterebbe nulla per *l'essenza* dell'umano perché ciò che, evidentemente, potrebbero spiegare le neuroscienze sarebbe solo il modo in cui “funzioniamo”.

Ciò è ben diverso dal gridare a voce alta che le peculiarità umane siano pura illusione o che, addirittura, non esistano.

Con questo senso, allora, d'ora in poi si dovranno apprendere le “lezioni delle neuroscienze”: il loro dire fornirà, certamente, “materiale” su cui il filosofo dalla poltrona che gli compete (la poltrona), potrà capire o quantomeno interrogarsi sull'uomo in quanto tale.

Una confusione di competenze si è avvertita anche all'interno del grande tema della intersoggettività: dalla scoperta dei *neuroni specchio* all'inizio degli anni Novanta, è tornata all'attenzione degli studiosi e del vasto pubblico il tema della *simpatia* di *humeana* memoria, conosciuta come *empatia*. Si è assistito al fiorire di “soluzioni”, a volte anche un po' bizzarre, dei più disparati problemi: dall'eziologia di alcune malattie alla spiegazione antropologica del comportamento umano, commettendo lo stesso errore a proposito del *free will*.

L'empatia è un discorso molto complesso che, proprio a causa del suo declinarsi in molteplici modi, si presta ad incomprensioni ed incongruenze: già la sua definizione, ad esempio, è problematica, poiché esistono tante definizioni quanti sono i modi per spiegarla.

Certamente, anche in questo caso, l'approccio neuroscientifico al tema ci permette di comprendere le sfaccettature del fenomeno e di rifarci ad una chiave neurobiologica dei motivi dell'accadimento, donando di scientificità qualsiasi discorso affrontato: sapere, ad esempio, quali siano i meccanismi sottesi all'emozione o alla presa di decisione contribuisce, *sine dubio*, all'elaborazione di una teoria più consona alla realtà dei fatti più di quanto si potesse fare nei secoli scorsi.

Si assiste, però, simultaneamente, alla "empatia mania": qualsiasi fenomeno è riconducibile a *neuroni specchio* e qualsiasi "stortura" psicologica, medico-clinica, evolucionistica è riconducibile al meccanismo di Rizzolatti *et alii*. Con ciò non si sta sostenendo che la scoperta operata dal laboratorio di Parma sia da sottovalutare, anzi, tutt'altro! Si tenta solo di evidenziare che tale è il fervore culturale per questa interdisciplinarietà fra etica e neuroscienze che, a volte, si tralascia l'importanza gnoseologica del dubbio di cartesiana memoria (e non solo).

Anche per il tema dell'empatia si corre il rischio di "appiattare" il discorso ai soli dati neuroscientifici che possono o potrebbero spiegare il fenomeno: ciò che si omette di sottolineare è che questi ultimi non lo esauriscono *in toto*, ma solo una parte, ossia quella relativa alla sola visione/interazione con altri su compiti o situazioni "semplici". Non a caso, infatti, si parla propriamente di una "empatia di base o primaria" ed una "secondaria o cognitiva": la prima, spiegabile attraverso il ricorso ai meccanismi neuroscientifici di stimolo/risposta (i *neuroni specchio* e simili); la seconda, con l'aspetto più controverso del fenomeno relativo all'accesso nella mente altrui e, conseguentemente percepirne su di sé lo stato emotivo. In questo senso, allora, si parla di *mentalizing* o di *lettura della mente*: un accesso privilegiato alla mente dell'altro a base della intersoggettività, forse dell'altruismo e forse, ancora, della morale.

Come si vede, anche in questo caso il ruolo del filosofo resta certo, evidente, indubbio nelle competenze e nelle riflessioni. Non si tratta, ribadisco, di un riduzionismo fine a se stesso, da molti inteso quale depauperamento dell'essere umano, ma solo di un ripristino di competenze per la fuoriuscita di un risultato condiviso: la comprensione dell'uomo in quanto tale.

D'altronde, uno degli intenti della neuroetica è anche questo: comprensione a tutto tondo della complessità umana.

Parte II

L'empatia cognitiva

Capitolo II

Che cos'è l'empatia

2.1 Le difficoltà di una definizione unanime

Quando David Hume, nel lontano 1739, utilizzò il termine *simpatia* dubitiamo che avesse mai pensato che un giorno, sia pur lontano, avrebbe dato adito a dibattiti, controversie e posizioni contrapposte su ciò che tale termine, con il concetto denotato, avrebbe dovuto o potuto significare.

Siamo nel 2019: sono trascorsi più di tre secoli ed il “problema simpatia” è ancora in via di risoluzione.

Si sente, oramai, sempre più spesso parlare di *empatia* e si sente, altrettanto spesso, distinguere la “prima empatia” dalla “seconda”, con progressivi mutamenti di soggetti e complementi oggetto per chiarirne i tratti, il fenomeno nel suo complesso, le sue caratteristiche.

Il problema, povero Hume, consta proprio in ciò: come definire l'empatia? Che cos'è?

Molti autori partecipanti al dibattito direbbero che è un fenomeno complesso, intrigato, sofisticato e ciò lo si evince già dalla difficoltà di fornire una risposta quanto più unanime possibile. Eppure non è così!

Per quali ragioni? Perché, contrariamente ad altri fenomeni, dare una definizione di *empatia* risulta così difficile?

Gli stessi pensatori (antichi e contemporanei), autori del progressivo dibattito sul tema, inizierebbero, immediatamente, a sottolineare che, comunemente, il termine *empathy*

presuppone la presenza di almeno due individui per la constatazione del fenomeno per poi cominciare a spiegarne le peculiarità la cui *ratio* deve, necessariamente, rimandare al tema dell'*intersoggettività*. Forse, proprio quest'ultima potrebbe essere il collante universale del fenomeno, ma, ahimè, ho da dissentire.

Quando si parla, oggi, di *empatia* non si dovrebbe tralasciare l'analisi dei motivi per cui si abbia tanta difficoltà nel "catalogarla", "etichettarla", soprattutto quando, filosoficamente, si è appresa la lezione di Ockham (*pluralitas non est ponenda sine necessitate*, recitava a cavallo fra il XIII ed il XIV secolo) e, soprattutto, quando si entra nel mondo della filosofia della mente.

Procediamo con ordine.

Il *punctum dolens* della definizione del fenomeno, ad avviso della scrivente, risiede nel fatto che l'empatia, per come oggi molti la intendono (torneremo più avanti sull'argomento), ossia l'"essere nei panni dell'altro e provare le medesime emozioni dell'altro" sia una traslitterazione emotiva, forse epistemologica, di uno strumento gnoseologico utilizzato dai "padri teorizzatori" del fenomeno stesso per spiegare il mondo, l'universo, la conoscenza *tout court* ora secondo un empirismo di fondo, ora seguendo un'idea positivista o romantica della cornice culturale in cui si sono adoperati.

David Hume, nel *Trattato sulla Natura Umana*²¹, parlando a proposito delle passioni, scriveva:

All'inizio di questo Trattato si è osservato come tutte le idee derivino dalle impressioni, e come questi due tipi di percezione differiscano solo per il diverso grado di forza e di vivacità con cui colpiscono l'anima. Le parti che compongono le idee e le impressioni sono perfettamente simili. Il modo e l'ordine del loro apparire possono essere identici. Quindi, i diversi gradi della loro forza e vivacità sono il solo punto che le distingue: e dal momento che questa differenza può in qualche misura essere rimossa da una relazione fra le impressioni e le idee, non c'è da meravigliarsi che un'idea di sentimento o di una passione possa ravvivarsi tanto da diventare quello stesso sentimento o passione. La viva idea di un oggetto si avvicina sempre all'impressione che se ne ha; e non c'è dubbio che possiamo provare un malessere o un dolore per la sola forza dell'immaginazione, e far quindi diventar reale una malattia pensandoci spesso. Ma questo accade con ancor maggiore evidenza nelle opinioni e nelle affezioni; ed è principalmente qui che un'idea viva si converte in un'impressione. Le nostre affezioni dipendono da noi stessi e dalle operazioni interne della mente più di qualsiasi altra

²¹ Cfr. DAVID HUME, *Trattato sulla Natura Umana*, Editori Laterza, Roma-Bari, 2004.

*impressione; e per questa ragione esse sorgono nel modo più naturale dall'immaginazione e da ogni viva idea che ce ne formiamo. Questa è la natura e la causa della simpatia; ed è in questo modo che riusciamo a entrare tanto profondamente nelle opinioni e nelle affezioni altrui ogni volta che le scopriamo*²².

L'analisi del modo in cui si incorre nella conoscenza di un fatto o di un evento, partendo da idee o da impressioni per poi giungere all'intelletto, è il motivo per cui Hume analizza il problema delle passioni e, quindi, della *simpatia*. Con tale termine si suole, in genere, indicare *l'azione reciproca delle cose fra loro o la loro capacità di influenzarsi a vicenda*²³; un significato che anche lo stesso filosofo inglese conserva, poiché esclude qualsiasi fusione emotiva tra le persone fra cui intercorre. Non a caso, infatti, Hume stesso sottolinea che essa rappresenta il principio interno alle passioni grazie a cui si realizza il passaggio da idee ad impressioni: uno strumento epistemologico, quindi, che dal mondo dell'emotività o delle passioni è in grado di “saltare” a quello intellettuale, razionale, dando ordine al mondo. Non a caso, egli tende a sottolineare che è partendo dall'io (non quello psicologico o *essenziale* della tradizione novecentesca), inteso quale agente umano, la simpatia fa scoprire a se stessi un certo grado di affinità o dissimiglianza con gli oggetti o con gli altri io, usando una terminologia non consona a quella humeana. A tal fine, Hume scrive:

*È certo evidente che quando proviamo simpatia per le passioni e i sentimenti degli altri, questi moti dell'animo appaiono inizialmente nella nostra mente come mere idee, e vengono pensati come appartenenti a un'altra persona, non diversamente da come concepiamo qualsiasi altro dato di fatto. È altresì evidente che le idee delle affezioni altrui si convertono nelle impressioni stesse che rappresentano e che le passioni sorgono in conformità alle immagini che ce ne formiamo [...]*²⁴.

Si comprende chiaramente che la natura della simpatia risiede in una *compartecipazione immaginativa alla vita degli altri*²⁵ in base a cui, superando il naturale egoismo umano, genera una “primaria” conoscenza altrui; in opposizione o in assonanza con il proprio io quale metro di valutazione, la simpatia contribuisce a “spiegare” il mondo emotivo dell'altro, stante la natura stessa del suo principio. Nonostante essa riguardi l'area passionale, Hume sostiene

²² *IBIDEM*, pp. 334-335.

²³ Cfr. NICOLA ABBAGNANO, *Dizionario di Filosofia*, Utet Libreria, Milano, 2006, lemma “simpatia”.

²⁴ DAVID HUME, *Trattato sulla Natura Umana*, op. cit., p. 335.

²⁵ IDA CAPPIELLO, *La morale della simpatia di David Hume*, Liguori Editore, Napoli, 1983, p. 87.

che possiede una somiglianza con lo stesso intelletto e che, se equiparate, corrisponderebbe, perfettamente, alla natura dello stesso: principio di causa ed effetto, principio di somiglianza, principio di contiguità.

Tanto è sentita quale vera da Hume la somiglianza fra la simpatia e l'intelletto che il filosofo inglese, in più di un'occasione, non manca di ribadire che:

Tutto ciò è oggetto della più comune esperienza e non dipende da una qualche ipotesi propria della filosofia. Questa scienza ha il solo ruolo di spiegare i fenomeni [...]»²⁶.

Non serve a Hume analizzare o indagare “oltre” la natura simpatetica, poiché risulta *evidente di per sé* che con essa si penetri nel mondo delle passioni o dei sentimenti: con una certa forzatura, si potrebbe dire che con tale atteggiamento il filosofo inglese precederà, di circa due secoli, l'atteggiamento fenomenologico, poiché anche Hume, come Husserl o la Stein, a suo modo, cerca l'essenza nella forma. Ovviamente, con un metodo ed un approccio impensabile per allora.

Il senso della speculazione del Libro III del *Trattato* era quello di dare autorevolezza alla morale, indicando la simpatia quale principio a cui ricorrere per spiegare non solo la genesi dei sentimenti morali, ma anche del modo in cui gli uomini riescano ad approvare l'utilità e la piacevolezza di alcune virtù (ritenute tali) e disapprovarne l'insolenza o la ripugnanza di altre. Una strutturazione, forse più debole, si avrà nell'opera tarda nella *Ricerca sui principi della morale*²⁷ in cui Hume, dopo la sfortunata vicenda del *Trattato*, depaupererà l'impianto concettuale delegato, precedentemente, alla simpatia, giungendo alla definizione di una *simpatia sociale*, ossia

*come ciò che consente una partecipazione alle emozioni altrui quasi coincidente con la benevolenza e il sentimento di umanità*²⁸.

Essa, vale a dire, non si presenta più quale unico principio fondante la sfera passionale in cui assume le caratteristiche di percezione neutra di trasmissione della somiglianza o dissimiglianza in stretta relazione all'Io “interrogante”, ma assume una connotazione *intersoggettiva* in cui il *sensu interno* di ognuno rimanda alla propria percezione sull'eventuale approvazione del comportamento. L'oggettività della morale, dunque,

²⁶ *IVI*.

²⁷ DAVID HUME, *Ricerca sui principi della morale*, Editori Laterza, Roma-Bari, 2018.

²⁸ *IDEM*, p. XI.

poggerebbe anche su questo sentimento comune a tutti gli uomini (appunto, la simpatia), che, grazie ai principi della benevolenza, della qualità e dell'utilità, garantirebbe la sua universalità. Scrive, a tal proposito:

Tutto quello che può avere un valore di qualunque specie, si classifica da sé così naturalmente sotto la divisione di utile o piacevole, l'utile e il dolce, che non è facile immaginare perché si dovrebbero svolgere ulteriori ricerche, o considerare la questione come materia di sottile investigazione²⁹.

A partire dal “cambio di rotta” operato dallo Scozzese nell’opera più tarda, si evince, seppur con le dovute differenze rispetto al *Trattato*, un uso del termine *simpatia* di stampo ontologico, ossia quale propensione o apertura verso l’altro; su cui, con l’apparente semplicità della strutturazione morale, si giungerebbe non solo ad un’etica unanimemente condivisa, ma anche all’autorevolezza della morale stessa. Si badi bene, però, che il senso di *simpatia* verso l’altro non è una capacità personale di entrare in comunione con il Tu, inteso quale immedesimazione nella sfera emotiva altrui, ma una “categoria” del pensiero in grado di identificare i principi universali della morale.

Sostiene, infatti,

[...] c'è un po' di benevolenza, per quanto poca, infusa nel nostro cuore, qualche scintilla di amicizia per la specie umana, qualche particella della colomba impastata nella nostra struttura, insieme con gli elementi del lupo e del serpente. Supponiamo pure che questi sentimenti così generosi siano così deboli da risultare insufficienti a muovere anche una mano o un dito del nostro corpo; essi devono tuttavia dirigere le determinazioni della nostra mente, e, restando invariato il resto, debbono produrre una calma preferenza per quello che è utile e giovevole all'umanità rispetto a ciò che le è dannoso e pericoloso³⁰.

È vero altresì che, sia nel *Trattato* sia nella *Ricerca*, Hume scrive che la simpatia permetterebbe di entrare nel vissuto emotivo altrui, ma tali asserzioni, se avulse dalla filosofia che ne disciplina la *ratio*, tenderebbero a fuorviare il lettore dal concetto humaneo, ponendo una falsa premessa per un falso ragionamento.

²⁹ *IBIDEM*, p. 161.

³⁰ *IDEM*, p. 167.

Alla luce di considerazioni che si faranno nel prosieguo del lavoro, si evincerà che non è un errore metodologico tentare di capire che cosa sia l'empatia partendo direttamente dallo Scozzese, poiché, nella sua filosofia, sono diversi i punti di congiunzione con le odierne concezioni empatiche, sia nel suo rapporto con la morale sia nel suo rapporto fra due o più individui. In questo senso, allora, se negli anni in cui Hume era attivo si fosse parlato di *empathy*, si sarebbe certamente sostenuto che la futura *simpatia* avrebbe dovuto denotare quell'apparato psichico, suffragato dalla razionalità dell'intelletto, in grado di valutare, come fosse un giudice, gli aspetti normativi della morale: quel *sensu interno* citato nella *Ricerca* che si sarebbe distanziato dal greco *pathein* ("sentire dentro", "soffrire", "patire"), utilizzato più volte nella filosofia aristotelica, ad esempio, quale catarsi emotiva dello spettatore nei confronti del rappresentato. Una prima definizione moderna, quindi, che già per i contemporanei di Hume sarebbe sembrata inadatta a spiegare il fenomeno.

Adam Smith, ad esempio, è colui il quale, nella *Teoria dei sentimenti morali*³¹, è meno interessato agli aspetti logici o gnoseologici della sua teoria, poiché la sua ricerca, sulla strada già percorsa da Hutcheson, mira a dare una fondazione razionale all'etica, esplicitando ed indagando più approfonditamente la sfera emotiva, contrariamente da quanto fatto da Hume. La centralità della simpatia, quindi, è da collocare all'interno della fondatezza morale entro cui, a sua volta, collocare le spiegazioni dell'approvazione o della disapprovazione morale: una centralità emotiva, dunque, che non mirava a comprendere *in nuce* il problema dell'intersoggettività *tout court*, ma che lo toccava solo perché avente a che fare con la sfera della morale. Come giustificare un comportamento accettabile da uno ritenuto indegno se non attraverso l'ausilio della simpatia così intesa e dello *spettatore interno* quale "uno sopra le parti"? Non a caso, dato l'utilizzo strumentale della simpatia, un elemento di novità per il successivo dibattito sull'*empathy* è svolto dal ruolo che l'immaginazione ha ed assume all'interno del contesto sociale e morale di Smith. A tal proposito, egli scrive:

Dal momento che non abbiamo esperienza diretta di ciò che gli altri uomini provano, non possiamo formarci alcuna idea della maniera in cui essi vengono colpiti in altro modo che col concepire ciò che noi stessi proveremmo nella loro stessa situazione [...] i nostri sensi [...] non ci hanno mai condotto, e mai potranno condurci, al di là della nostra persona, ed è solo attraverso l'immaginazione che noi possiamo concepire quali siano le sue

³¹ Cfr. ADAM SMITH, *Teoria dei sentimenti morali*, Bur Rizzoli, Milano, 2016.

sensazioni. E tale facoltà non può aiutarci in questo, altro che col rappresentarci quali sarebbero le nostre sensazioni se fossimo noi al posto suo [...]³².

Il ruolo dell'immaginazione è, in realtà, un aspetto dell'odierno dibattito che un po' si è perso forse a causa o per merito delle odierne scoperte neuroscientifiche sul modo in cui dovrebbe funzionare il cervello umano. Oramai è noto al grande pubblico che la scoperta dei *neuroni specchio* e delle successive aree cerebrali in grado di "intervenire" nel rapporto con l'altro o con gli oggetti è utilizzato da molti per spiegare parte della natura empatica: di questo fenomeno, cioè, di cui ancora non si è saputa dare alcuna definizione unanime.

Tornando a Smith, il quale continua la sua dissertazione sulla simpatia e sul ruolo svolto dall'immaginazione con l'esempio del funambolo (erroneamente associato alla sola figura di Lipps, ma ben presente nella tradizione Settecentesca già da Du Bos) e ripreso, successivamente, da Edith Stein nel Novecento, ribadisce che:

la simpatia non nasce dalla semplice vista di una passione, quanto dalla vista della situazione che la suscita. Proviamo a volte, al posto di un altro, una passione della quale lui stesso sembra del tutto incapace, perché, quando ci mettiamo nei suoi panni, quella passione sorge in noi dall'immaginazione, nonostante non sorga in lui nella realtà [...]³³.

Si è lontani, quindi, dall'utilizzo dell'empatia quale strumento epistemologico in grado di "entrare" nelle altre menti come si possedesse un accesso privilegiato ad esse con cui non solo si capirebbe l'altro, ma si riuscirebbe in sé a provare – a riprodurre, "simulando" – le emozioni altrui: la generalizzazione odierna *empathy*, contrariamente a quanti, facendosi ingannare dal titolo dell'opera di Smith, ritengono quest'ultimo uno dei primi teorici del concetto di empatia insieme a Theodor Smith, in realtà, sbagliano.

Costoro, infatti, tralascerebbero la non ovvia considerazione che fra il Settecento e l'Ottocento il fermento culturale europeo, soprattutto inglese e tedesco, non muoveva da una settorializzazione disciplinare comune alla cultura contemporanea, ma tentava di fornire risposte a domande sull'intera speculazione sull'uomo e sul mondo che non tralasciava, certamente, né la sfera passionale/sentimentale né quella razionale. Si pensi, ad esempio, alla divergenza di impostazione fra il *Positivismo* ed il *Romanticismo*: è all'interno di quel

³² *IDEM*, pp. 81-82.

³³ *IBIDEM*, p. 86.

paradigma³⁴, infatti, che si consuma, da una parte, l'esaltazione della Scienza e della Scienza evolucionistica e, dall'altra, il tentativo, poi brillantemente riuscito, di razionalizzare il sentimento (si pensi al geniale Hegel).

Con ciò si vuole solamente far notare che utilizzare la "teoria della simpatia" di quell'epoca per spiegare l'odierna è un errore epistemologico e gnoseologico: non si potrebbe, infatti, fornire lo stesso senso con cui quei filosofi hanno parlato di "sentimento" per supportare o teorizzare il modo contemporaneo con cui se ne parla. Con ciò non voglio sostenere che una genealogia del concetto di empatia o di qualsiasi altro concetto sia inutile o errata, ma solo precisare la cautela con cui far "riemergere" talune tematiche estrapolate dal contesto. Lo stesso errore, ad esempio, di chi, tornando al *focus*, sostiene che Theodor Lipps è stato il padre teorico dell'empatia. Analizzando le sue opere con questa ipotetica lente di ingrandimento, si dovrebbe immediatamente notare che il senso con cui Lipps parla effettivamente di *Einfühlung* non è la teorizzazione di un fenomeno prettamente psicologico da "etichettare" all'interno dell'odierna Psicologia, quindi un precursore dei suoi tempi, ma è mosso dalle medesime esigenze dei molti prima di lui: dare una fondatezza psicologica al mondo, partendo da se stessi. La rivoluzione copernicana operata da Kant nelle *Critiche* è, per traslitterazione, la stessa *ratio* che muove Lipps, prima nel campo estetico e dopo in quello epistemologico: l'empatia, in questo secondo contesto, diviene, ad avviso della scrivente, un atto epistemico con cui "dare ordine al mondo". Nel secondo libro dell'estetica, infatti, egli scrive che, dopo la teoria, ora ha necessità di dedicarsi all'analisi dei sentimenti, intesi quali

*attività di creazione del mondo e degli oggetti: una creazione in senso logico e contemporaneamente psicologico, che finisce ancora una volta nel relativizzare le basi della conoscenza [...]*³⁵.

L'empatia, quindi, diviene un'attività teoretica che, per certi aspetti a volte sembra oscillare fra la natura dell'*Io Penso* kantiano e l'*Io Infinito* di Fichte, tanto che Lipps stesso sente, all'interno della sua opera, di dover chiarire che l'*Einfühlung* è, innanzitutto, un'attività. Un'attività che sembra impostare un discorso relativo ad una "Psicologia Trascendentale" in cui i tre ambiti conoscitivi, individuati da Lipps, devono la loro esistenza solo ed esclusivamente alla presenza di un Io che, rapportandosi a se stesso, configura il mondo dei fenomeni e dei soggetti.

³⁴ THOMAS KHUN, *La struttura delle rivoluzioni scientifiche*, Piccola Biblioteca Einaudi, Torino, 2009.

³⁵ MARIA ROSARIA DE ROSA, *Theodor Lipps. Estetica e critica delle arti*, Guida Editori, Napoli, 1990, p. 40n.

A tal proposito sostiene:

Vi sono tre ambiti conoscitivi. So delle cose, di me stesso e di altri Io. La prima conoscenza ha per fonte la percezione sensibile. La seconda la percezione interna, cioè il coglimento dell'Io, che avviene immediatamente oppure retrospettivamente nel ricordo, insieme con le sue determinazioni, con le esperienze vissute di esigenze varie, con le sue attività, i suoi atti e sentimento, e al contempo con le sue relazioni riguardo agli oggetti. La fonte del terzo tipo di conoscenza, infine, è l'empatia³⁶.

Tanto che Lipps sente di dover chiarire o quantomeno specificare il senso dell'*Einfühlung* quale

'oggettivazione di me stesso in un oggetto diverso da me' a prescindere dal fatto che ciò che è obiettivato meriti o meno il nome di sentimento in senso vero e proprio³⁷.

Prescindendo dall'analisi minuziosa che il filosofo tedesco conduce sull'empatia, il punto di svolta della considerazione secondo cui sarebbe alquanto dubbio considerare Lipps il padre teorico dell'odierna *empathy* risiede proprio nella citazione appena riportata. Sostenere, infatti, che si debba parlare di una oggettivazione di sé in un oggetto distinto da sé dà adito all'interpretazione che, da una parte, si stia parlando di un processo costruttivo del reale in cui si debba conferire un senso ed un significato ad un'operazione psicologica tipica del "demiurgo empirico" e, dall'altra, non si pone l'accento sulle possibili conseguenze emotive del presunto atto psicologico (cioè sulla ipotetica teoria dell'empatia), poiché Lipps stesso tende a precisare che tale aspetto diviene completamente secondario. La natura di ciò che è obiettivato passa in secondo piano rispetto al processo che l'ha posto in essere.

Proprio per tali considerazioni, sarei incline a pensare che ciò di cui Lipps sta parlando non è *empathy* propriamente detta, ma un evento distinto: si passa da un processo (in parte creativo) della psicologia trascendentale, prima, ad un processo di *noesis*, dopo. Che in tale sviluppo subentri anche la sfera sentimentale sembrerebbe pressoché ovvio, poiché la complessità dell'uomo *tout court* implica, necessariamente, anche questa parte della sua natura.

La definizione che si addice all'empatia di Lipps, allora, non può non essere diversa da quella fornita da chi lo ha preceduto, proprio perché gli aspetti considerati ed il senso della sua

³⁶ THEODOR LIPPS, *Fonti della conoscenza. Empatia*. Articolo reperito dalla Biblioteca di Ateneo dell'Università di Bologna, *Alma Mater Studiorum*.

³⁷ *IDEM*.

speculazione si muovono da esigenze distinte e da contesti culturali completamente differenti da Hume o da Smith, tanto per citare Filosofi già incontrati. Non a caso, infatti, l'esigenza di Lipps è quella di dare vita al mondo attraverso una forza persuasiva ed onnicomprensiva, muovendosi in contesto romantico e sostenendo che:

*Qui 'empatia' significa allora: nel momento in cui appercepisco un oggetto, esperisco una determinata modalità del mio comportamento interiore come se appartenesse a tale oggetto appercepito, o come se vi si trovasse quale sua componente. Questa modalità appare come se mi fosse data o comunicata in tale oggetto [...]*³⁸.

Anche in questo caso, come del resto nei precedenti, si trova una definizione di *empathy* che pone l'accento su alcuni aspetti legati non solo all'atto epistemico, ma all'intersoggettività: come più tardi sosterrà Edmund Husserl, sarà tramite questa "propensione" verso il mondo che renderà consapevole l'Io della sua mancata monadologia, essendo un Io fra Io; tematica che, nella fattispecie, sarà poi analizzata minuziosamente da Edith Stein, allieva di Husserl.

In un quadro concettuale in cui la stessa definizione del fenomeno è effimera, il lavoro svolto da Martin Hoffman alla fine degli anni Settanta del secolo scorso fornisce delle linee di indagine e di spiegazione di notevole importanza.

Dal greco *enpatheia* (*en*, dentro; *patheia*, passione), al tedesco *Einfühlung* (*ein*, dentro; *Fühlung*, emozione) all'inglesismo *empathy*, Hoffman nota che spiegare il fenomeno come "essere all'interno del sentimento altrui" potrebbe essere equivoco e poco esplicativo per cui, mantenendo più o meno stabile il concetto, sostiene che sarebbe:

*di gran lunga preferibile definire l'empatia non nei termini del suo risultato (coincidenza affettiva) ma dei processi alla base della relazione tra il sentimento dell'osservatore e quello del modello*³⁹,

tanto che la definisce

*la scintilla da cui nasce l'interesse umano per gli altri, il collante che rende possibile la vita sociale*⁴⁰.

³⁸ IDEM, pp. 47-48.

³⁹ MARTIN HOFFMAN, *Empatia e sviluppo morale*, Il Mulino, Bologna, 2008, p. 54.

⁴⁰ IBIDEM, p. 23.

Affinché possa esservi empatia, la condizione essenziale è, secondo Hoffman, il coinvolgimento dei processi psico-biologici dell'osservatore

*che facciano sì che i sentimenti di una persona (l'osservatore, n.d.r.) siano più consoni alla situazione di un altro che non alla propria*⁴¹.

Il problema di Hoffman, motivo per cui ha dedicato le sue ricerche al problema del “prendersi cura dell'altro” per oltre trentanni, è capire le ragioni del comportamento prosociale; ancor prima, cioè, che si arrivi alla teorizzazione o alla condivisione di un paradigma etico condiviso: in questo quadro concettuale, il *focus* sull'empatia è, ancora una volta, diverso; non tanto, come detto, sul risultato, ma sulle sue cause.

Prescindendo dall'analisi precisa e puntuale che si svolge nelle pagine dell'opera del 2000, l'aspetto interessante per la seguente trattazione è la consapevolezza di Hoffman, poi dimostrata anche a livello neuroscientifico, che il fenomeno empatico può essere suddiviso sotto due grandi aspetti:

- l'aspetto primario;
- l'aspetto secondario.

Nel primo rientrano tutti quei processi di attivazione empatica pre-razionali, pre-verbali ed istintuali che, così come si evince, non sono, cioè, mediati da una consapevolezza razionale o coscienziale della sofferenza o della gioia dell'osservatore nei confronti di una ipotetica vittima.

Un esempio di quanto detto è fornito proprio dalla *mimesi* o *mimicry*, ossia una risposta involontaria ad uno stato emotivo osservato che, attraverso l'*imitazione* prima e la *retroazione* poi, giunge ad una concomitanza emotiva fra l'osservatore e l'osservato.

*Dapprima, e in modo automatico, le variazioni dell'espressione facciale, della voce e della postura dell'osservatore riproducono, in sincronia, le più lievi variazioni nell'espressione dei sentimenti da parte di un'altra persona [...]. Dopo di che, i cambiamenti della muscolatura facciale, vocale e posturale dell'osservatore suscitano una retroazione afferente che causa nell'osservatore sentimenti che coincidono con quelli della vittima*⁴².

⁴¹ *IDEM*, p. 54.

⁴² *IBIDEM*, p. 62.

A riprova del fatto che le espressioni facciali tenderebbero ad influenzare le esperienze emozionali, Charles Darwin, negli scritti del 1862, sosteneva che il fenomeno onnicomprensivo dell'*imitazione* poteva essere la causa prosociale che spingerebbe un individuo di una specie ad aiutarne un altro; e recenti esperimenti del 1994, condotti dal neuroscienziato Johnson D. Laird *et alii*⁴³, hanno constatato che la *retroazione* potrebbe essere sia una spinta emotiva forte all'aiuto reciproco sia una spinta all'azione pro-sociale a fondamento della moralità: non si spiegherebbe in modo diverso, infatti, il motivo per cui ai soggetti cui era stato chiesto di osservare facce sorridenti, mantenendo il volto accigliato, in un caso; il volto sorridente, nel secondo, alla fine dell'esperimento riferivano di sentirsi più arrabbiati, nel primo caso; più felici, nel secondo, senza una motivazione apparentemente valida.

Sulla natura della retroazione, però, sono stati sollevati diversi dubbi, poiché se è vero che l'espressione facciale modifica l'umore del soggetto, chi non dice che una forma debole di autopercezione non sia la causa del cambio dello stato emotivo? Ciò inficerebbe la legittimità di tutti quegli studi in sviluppo progressivo sulla presenza dell'empatia già nella prima infanzia in cui, dando per assodato che la *mimicry* operi già nelle prime fasi dello sviluppo del bambino, gli permetterebbe di rispondere "empaticamente" al sentimento altrui senza averlo prima mai esperito; al contrario, laddove la *mimicry* si sviluppasse solo in un secondo momento, l'empatia nella prima infanzia sarebbe il risultato di un processo educativo e di un condizionamento classico di pavloviana memoria anche se gli studi di Paul Ekman, prevalentemente, indicano l'universalità delle emozioni e delle espressioni facciali, causate dall'integrazione neurale⁴⁴.

A differenza della mimesi, il condizionamento classico è un altro meccanismo preverbale dell'attivazione empatica che risponde solo alla situazione emotiva della vittima; i cambiamenti dell'espressione facciale condizionano la risposta dell'osservatore che non si caratterizza, come nella mimesi, per la diretta corrispondenza fra sentimento osservato e manifestato dalla stessa, ma resta ad uno stadio "più superficiale" di corrispondenza, stavolta, fra un'espressione che rimanda ad un sentimento ed uno stimolo neutro in grado di provocare lo stesso sentimento nella vittima.

⁴³ JASON. D. LAIRD, JAY J. WAGENER, MICHAEL. HALAL E MICHELLE SZEDGA, *Remembering what you feel: Effects of emotion and memory*, in "Journal of Personality and Social Psychology", 42, pp. 646-675.

⁴⁴ PAUL EKMAN, WALLACE V. FRIESEN, MAUREEN O'SULLIVAN ED ANTHONY CHAN, *Universals and cultural differences in the judgments of facial expressions of emotion*, in "Journal of Personality and Social Psychology", 53, 1987, pp. 712-717.

A differenza di quanto si credeva, il condizionamento classico può avvenire anche nel neonato e, successivamente, la sofferenza o la gioia di un bambino può essere condizionata da qualsiasi persona, non solo dalla madre⁴⁵.

Altro meccanismo automatico di preattivazione empatica è rappresentato dall'associazione diretta; la quale si differenzia dal condizionamento classico, poiché non presuppone che la propria sofferenza sia accompagnata o suscitata da quella della vittima. L'unico requisito, infatti, è la presenza di uno stimolo o di un segnale che rimandino l'osservatore al sentimento di sofferenza provato in una data situazione in presenza di quello o di un quell'altro segnale, apparentemente neutro.

La caratteristica principe dei primi meccanismi di attivazione empatica è, come detto, la loro automaticità, rapidità ed inconsapevolezza, fondamentali, seguendo Martin Hoffman, per varie ragioni:

- a. *Sono automatici, involontari e agiscono molto rapidamente;*
- b. *Permettono ai bambini nella prima infanzia e in età preverbale, come pure agli adulti, di rispondere empaticamente alla sofferenza altrui;*
- c. *Fanno sì che il bambino già in tenera età metta in relazione la sofferenza empatica propria con quella di altre persone, ciò che contribuisce alla sua aspettativa di sofferenza quando è esposto alla sofferenza altrui;*
- d. *Hanno una certa capacità di autorinforzarsi, perché alimentano comportamenti di aiuto che possono produrre sollievo empatico;*
- e. *Conferiscono una dimensione involontaria alle future esperienze empatiche*⁴⁶.

I tre meccanismi pre-verbali descritti, nonostante attivino immediatamente una risposta di tipo empatico nell'osservatore, continuano ad essere oggetto di numerose teorie grazie all'intenso sviluppo neuroscientifico⁴⁷ e donano pragmaticità al fenomeno, per merito delle continue scoperte sul funzionamento cerebrale⁴⁸; presentano, però, gravi limiti strutturali affinché possano rappresentare la *ratio* di un comportamento pro-sociale. La sola lontananza della vittima, ad esempio, può ingenerare una totale indifferenza nel presunto spettatore. Non a

⁴⁵ ELLIOTT M. BLASS, JUDITH R. GANCHROW E JACOB E. STEINER, *Classical conditioning in newborn humans 2-48 hours of age*, in "Infant Behavior and Development", 7, 1984, pp. 223-235.

⁴⁶ MARTIN HOFFMAN, *Empatia e sviluppo morale*, op. cit., p. 73.

⁴⁷ Cfr. GIACOMO RIZZOLATTI E CORRADO SINIGAGLIA, *So quel che fai. Il cervello che agisce e i neuroni specchio*, Raffaello Cortina Editore, Milano, 2005.

⁴⁸ Cfr. SIMON BARON-COHEN, *La scienza del male. L'empatia e le origini della crudeltà*, Raffaello Cortina Editore, Milano, 2012.

caso, lo stesso Hoffman sostiene che tali meccanismi non contribuiscono allo sviluppo dell'aspetto secondario o metacognitivo dell'empatia, che, garantito dalla presenza del linguaggio e dallo sviluppo cognitivo, permetterà, successivamente, di focalizzarsi sul problema attuale di fondatezza epistemologica dell'empatia quale strumento di comprensione della mente altrui.

Andando per gradi, Hoffman indica tre meccanismi secondari:

- l'*associazione mediata*, tramite il linguaggio;
- l'*assunzione di ruolo* o *role taking*;
- la *combinazione*, sintesi dei due meccanismi precedenti.

Nella prima, la mediazione verbale è in grado di trasmettere la sofferenza dell'ipotetica vittima e la decodifica della semantica consente al lettore di relazionarsi, dapprima, con la propria situazione personale, laddove esperita quale simile a passate esperienze; dopo, con l'eventuale evocazione visiva dell'ipotetica vittima bisognosa d'aiuto alla quale può rispondere provando, effettivamente, sofferenza empatica per i meccanismi di *mimesi* o *associazione diretta*.

Le caratteristiche fondamentali dell'*associazione mediata*, le ritroveremo anche nelle successive; esse sono:

- il tempo necessario per la decodifica del messaggio;
- lo sforzo cognitivo per comprenderlo;
- l'elaborazione semantica, attraverso cui il lettore crea una distanza psicologica *volontaria* fra sé e la vittima presunta.

L'*assunzione di ruolo* o *role taking* rappresenta il mettersi al posto dell'altro, ma, contrariamente al senso comune sull'empatia, in questa particolare forma di attivazione cognitiva il ruolo centrale lo svolge l'immaginazione, poiché è solo tramite quest'ultima che l'osservatore può avviare lo sforzo cognitivo di "cambiare la propria mente", come dicono gli inglesi.

Non importa, per la *ratio* del seguente lavoro, se l'assunzione di ruolo sia centrata su di sé o su un altro⁴⁹, poiché ciò che si vuole far notare è un aspetto che sarà, poi, spiegato nelle *Conclusioni*: il grave limite cognitivo del meccanismo stesso, definito da Hoffman *deriva egoistica*.

Ad ulteriore esplicitazione della posizione della scrivente, è utile riportare la posizione dello psicologo statunitense, il quale sostiene che:

⁴⁹ MARTIN HOFFMAN, *Empatia e sviluppo morale*, op. cit., pp. 78-83.

*Quando ci mettiamo al posto della vittima e rievochiamo ricordi personali carichi di risonanze affettive, questi ricordi, a volte, possono assumere il controllo della nostra risposta e spostare la nostra attenzione dalla vittima a noi stessi [...] la figura della vittima passa in secondo piano. In altri termini, la relazione empatica con la vittima si impadronisce dell'osservatore per poi, ironicamente, spezzarsi, poiché l'affetto empatico entra in profonda risonanza con i bisogni dell'osservatore stesso, il quale sposta la propria attenzione, inizialmente centrata sulla vittima, su di sé [...]*⁵⁰.

Si assiste, cioè, ad un *dissolvimento* del fenomeno empatico che evidenzia, chiaramente, che sebbene si sia “solidali” con la vittima, non si è, in fin dei conti, la vittima stessa tanto che, di fronte a questo forte limite del fenomeno, Hoffman ritiene di chiarire che:

*Qualunque sia la spiegazione, la risposta affettiva dell'osservatore, sebbene inizialmente suscitata dallo stato affettivo della vittima, non va considerata, a mio giudizio, una forma di empatia, a meno che l'osservatore non rivolga di nuovo l'attenzione sulla vittima*⁵¹.

Anche qui, riprendendo il filo rosso del discorso, si può dare una definizione di empatia a seconda delle caratteristiche peculiari del fenomeno che si considerano: ora nei suoi aspetti primari o psico-biologici; ora nel suo aspetto cognitivo.

Indipendentemente dal merito delle ricerche condotte da Hoffman e dalle conclusioni cui giunge, è importante riconoscergli un positivo approccio metodologico perché per porre un capo alla “matassa *empathy*” opera un “percorso a ritroso”: sostiene, cioè, di dover partire dal risultato del fenomeno, dall'*explanandum*, per giungere, così, all'*explanans*, rendendo evidente, in tal modo, la difficoltà accennata poco più su nel considerare l'empatia quale accesso privilegiato alle altre menti.

Si potrebbe pensare, allora, che proprio in quello sforzo cognitivo operato dall'osservatore nei confronti della vittima risiederebbe “la soluzione” per il dubbio posto, ossia la spiegazione di che cosa sia l'empatia.

Ahimè, la risposta è ben più complessa di quanto si possa pensare, poiché proprio su tale interrogativo, evidenziato da Edith Stein (almeno nella sua accezione più contemporanea), risiede il senso di tutto un dibattito intercontinentale (di stampo analitico, soprattutto) che

⁵⁰ *IBIDEM*, p. 82.

⁵¹ *IVI*.

intreccia epistemologia, filosofia della mente, ermeneutica, neuroscienze...che ha condotto all'aggravarsi di idee e definizioni sull'empatia, annebbiando la natura effettiva del fenomeno e della nitidezza concettuale.

Ovviamente, non tutto è nato per caso: ci sono state delle ragioni storico-concettuali per cui si è giunti ad ampliare il ventaglio di possibilità dell'etichetta "empatia" dallo sviluppo del Positivismo Logico al progressivo allontanamento del problema dell'*action theory* della Filosofia della Mente; all'attribuzione intenzionale alla *folk psychology* di una razionalità presunta o immotivata. Fatto sta che, prescindendone l'analisi, uno degli aspetti che nel primo Novecento è emerso quale caratteristica empatica, più di quanto avessero già fatto Hume o Smith, è stata l'inter-relazione.

Anche sotto questo punto di vista, infatti, sono state avanzate diverse definizioni che ponessero l'accento su quest'altro aspetto poliedrico del fenomeno. In Edmund Husserl, ad esempio, è

da osservare che il tema dell'intersoggettività si articola lungo due diversi ordini di pensiero. Innanzitutto si tratta di comprendere quali siano le condizioni di possibilità dell'esistenza dell'altro [...] e si pone il problema della sua accessibilità per me, della possibilità della nostra comunicazione, dei mezzi della nostra intesa⁵²,

ossia dell'intersoggettività *tout court*.

Si assiste, in questo caso, al capovolgimento dell'interrogativo originario: non più come sia possibile interagire con l'altro da sé, ma come sia possibile essere e rimanere una *monade* fra le altre.

È facile intuire come il filone fenomenologico abbia dato vita a tutta una serie di considerazioni sul fenomeno che hanno portato alla concettualizzazione di particolari declinazioni dello stesso, che se per certi versi possono aiutare a capirne la natura per certi altri hanno contribuito ad intorbidire ulteriormente le acque.

Non a caso, come si scriveva più su, dare una definizione unanimemente condivisa dell'*empathy* trova non poche difficoltà: l'*Einfühlung* è un fenomeno poliedrico e trasversale, difficilmente circoscrivibile a dati "oggettivi"; tanto che, facilmente, si presta a confusioni o fraintendimenti con altri fenomeni ad essa apparentemente simili.

⁵² ALICE PUGLIESE, *La dimensione dell'intersoggettività. Fenomenologia dell'estraneo nella filosofia di Edmund Husserl*, Mimesis Editore, Milano, 2004, p. 72.

Sui motivi per cui nel dibattito odierno si è ricominciato ad asserire l'importanza del tema in oggetto sarà motivo di analisi nell'ultimo capitolo di questa prima parte dello scritto; poiché solo a quel punto saranno esplicitati i temi attuali, contemporanei, sul "problema empatia" nel suo aspetto cognitivo, oggetto di analisi e di critica del presente.

Per adesso, il lettore mi perdonerà, si dovrà passare all'illustrazione delle differenze fra *empatia*, *simpatia* e *contagio emotivo*.

2.2 La differenza fra empatia, simpatia e contagio emotivo

L'interesse per il concetto di *empatia* nelle Scienze Umane si deve alla traduzione del termine ad opera di Edward Titchener all'inizio del Novecento; il quale, compresa la portata rivoluzionaria che avrebbe potuto comportare l'estensione del termine in ambiti diversi rispetto a quello prettamente estetico, lo applicò ai rapporti umani. Probabilmente, lo psicologo Carl Rogers, usando *empathy*, iniziò a riferirsi a *live the attitudes of the other*⁵³, a vivere, cioè, quelle credenze e desideri dell'altro che, nel gergo di una oramai consolidata tradizione analitica, vengono definiti *atteggiamenti proposizionali* o *propositional attitudes*: tutti quegli stati mentali esprimibili linguisticamente da asserzioni del tipo 'A desidera che, crede che, pensa che...', le quali manifestano, quindi, lo *stato intenzionale* dell'agente⁵⁴.

In questo contesto, il concetto stesso di *intenzionalità* è coniugato nel senso attribuitole dal filosofo e psicologo tedesco Franz Brentano, a proposito dell'esplicitazione della differenza tra *fenomeno fisico* e *fenomeno psichico*: se il primo, infatti, trae la sua origine dagli oggetti esistenti nel mondo, studiati dalle scienze naturali e lontani da qualsiasi rapporto con la coscienza; il secondo nasce da sensazioni o da fantasie relative ad esperienze uditive, visive o tattili dell'agente da cui deriveranno le *rappresentazioni*, oggetto di giudizio e dei moti dell'animo. Caratteristica peculiare di questo tipo di fenomeno è l'*in-esistenza intenzionale*, ossia il riferimento sì ad un oggetto, ma di ordine mentale, psichico, esistente, cioè, solo nella sfera cognitiva dell'agente⁵⁵. In questo senso, quindi, la nozione di *atteggiamento proposizionale* si riferisce all'intenzione dell'agente di "riferirsi a, essere relativi a, dirigersi verso" stati di cose esistenti materialmente (*fenomeni fisici*) o esistenti "mentalmente" (*fenomeni psichici*); ossia, coniugando Rogers, a quegli stati intenzionali

⁵³ CARL ROGERS, *Client-centered therapy*, Houston-Mifflin, Boston, 1951, p. 29.

⁵⁴ ANTONIO RAINONE, *La riscoperta dell'empatia. Attribuzioni intenzionali e comprensione nella filosofia analitica*, Bibliopolis, Napoli, 2005, p. 49.

⁵⁵ Cfr. FRANZ BRENTANO, *La psicologia da un punto di vista empirico*, Biblioteca Universale Laterza, Roma-Bari, 1997.

dell'agente/interpretato, avente come oggetto percezioni, emozioni, verso cui l'interprete dirige la sua attenzione e predisposizione alla comprensione. Sorvolando, per il momento, sulle ragioni per cui l'empatia ha avuto nel corso degli ultimi cinquanta anni una diffusione notevole nel campo psicologico, si noti che, ad oggi, la sua indefinibilità in termini categorici ha contribuito ad incrementare diverse linee guida al suo studio; ognuno declinabile secondo caratteristiche e modelli arbitrari, generalizzabili in cinque grandi approcci:

1. *Temperamental or personality variables related to empathy –related responding*, in cui si presta attenzione al grado di intensità empatica provata dall'agente posto di fronte all'angoscia di una vittima, avendo quali variabili il suo temperamento e la sua personalità;
2. *The development of empathy-related responding*; da alcuni anni a questa parte, infatti, sono nati studi mirati sullo studio dello sviluppo emotivo-empatico dei minori già dalle loro prime fasi dello sviluppo cognitivo;
3. *The relation of empathy-related reactions to social behaviour, including prosocial behaviour, aggression and social competence*; riprendendo e riproducendo l'esperimento di Milgram, si è prestato attenzione al modo in cui l'empatia del carnefice si annichilisce di fronte alla figura dell'autorità o come, di contro, l'empatia possa essere la motivazione per l'agente ad attuare un comportamento prosociale, a fondamento, forse, dell'altruismo⁵⁶;
4. *Gender differences in empathy-related reactions*, nelle cui ricerche si tenta di fornire una spiegazione al perché un genere sarebbe più incline a provare empatia rispetto all'altro;
5. *Socialization correlates*, in cui si evidenziano i limiti a cui l'empatia è soggetta, riferendosi, in particolar modo, alla diversa risposta dell'agente di fronte ad una situazione di sofferenza della vittima: da un familiare o parente stretto ad un perfetto sconosciuto.

I cinque diversi indirizzi di ricerca, stigmatizzati (qui in poche righe) da Nancy Eisenberg⁵⁷, riassumono *the crux of the problem*⁵⁸, la quale, nonostante renda quasi impossibile definire

⁵⁶ DANIEL L. KREBS, *Empathy and Altruism*, in "Journal of personality and social psychology", 1975, XXXII, pp. 1134-1146.

⁵⁷ MICHAEL LEWIS E JEANNETTE M. HAVILAND-JONES, *Handbook of Emotions*, Guilford Press, New York, 2018, pp. 677-691.

⁵⁸ LAUREN WISPÈ, *The distinction between sympathy and empathy*, in "Journal of personality and social psychology", Vol. 50, n. 2, 1986, pp. 314-321, cit. p. 317.

l'empatia unanimemente, renderebbe possibile, allo stesso tempo, la validità della sua costruzione teorica: dotare di pragmaticità un concetto teorico significa attribuirgli attendibilità e credibilità⁵⁹.

I dubbi e le perplessità, però, non sono superati neppure da questo punto di vista, poiché diverse sono le domande senza risposta sui vari esperimenti che si sono condotti e si continuano a condurre sulla risposta “emotiva” dell'agente posto in una “situazione empatica”, in quanto, ad esempio, le diverse istruzioni forniscono già diverse risposte; risulta probabile, a questo punto, che gli stessi esperimenti possano indagare altri fenomeni che esulino dall'empatia *tout court*.

Un serio limite del costrutto di Cronbach e Meehl, infatti, risulta essere la mancanza di una “inconfutabile” evidenza che ne giustifichi la stessa validità. Non a caso una confusione simile si ripercuote anche sull'uso di due termini che potrebbero denotare o due concetti differenti o il medesimo che riguardano l'empatia, da un lato, e la simpatia, dall'altro. Anche se, storicamente, si è avuto modo di notare la traslitterazione del primo nel secondo, è pur vero che un effervescente clima culturale a carattere analitico si sta convincendo che se l'empatia è definibile per mezzo dei fenomeni del *role taking*, *perspective taking* o, in generale, del sentire in sé l'emozione altrui, la simpatia dovrebbe denotare

[...] *to the heightened awareness of the suffering of another person as something to be alleviated*⁶⁰,

ossia la consapevolezza del dolore altrui non come emozione vissuta in prima persona dall'osservatore, ma quale spinta motivazionale all'aiuto per il benessere stesso della vittima. A differenza dell'empatia, la simpatia sembrerebbe essere molto di più di una generica compassione: si parlerebbe, infatti, dell'attivazione di un processo coscienziale, di consapevolezza della sofferenza della vittima, appunto, che spingerebbe l'osservatore ad intervenire per alleviare lo stato di difficoltà della stessa. L'io, in questo caso, non diverrebbe né il mezzo epistemologico per conoscere la “mente altrui”, né l'oggetto dell'azione; perché lascerebbe spazio all'altro, sia come mezzo sia come fine dell'azione stessa.

Non si tratta di capire, ma di agire;

*empathy is a way of knowing. Sympathy is a way of relating*⁶¹.

⁵⁹ LEE J. CRONBACH E PAUL E. MEEHL, *Construct validity in psychological tests*, in “Psychological Bulletin”, n. 52, 1955, pp. 281-387.

⁶⁰ LAUREN WISPÈ, *The distinction between sympathy and empathy*, op. cit., p. 318.

L'empatia, per di più, non solo è un'esperienza che può sorgere inintenzionalmente, ossia indipendentemente dalle proprie predisposizioni ad empatizzare, ma alcuni studi sembrerebbero circoscrivere la predisposizione empatica al proprio gruppo di appartenenza, familiare o sociale che sia, riscontrando una certa difficoltà ad estendere il fenomeno in soggetti lontani culturalmente e fisicamente dall'osservatore; un aspetto che sembrerebbe essere assente nella simpatia, la quale, oltre ad essere durevole piuttosto che momentanea, avrebbe come unico scopo il benessere dell'altro, indipendentemente dalla provenienza o dal grado di familiarità.

La simpatia o *preoccupazione simpatetica* (*sympathetic concern*), quindi, sembrerebbe rispondere a tre particolari caratteristiche:

- a. *responds to some apparent threat or obstacle to an individual's good or well-being;*
- b. *has that individual himself as object;*
- c. *involves concern for him and thus for his well-being*⁶².

La preoccupazione per l'altro apre, inevitabilmente, la simpatia al campo della "morale pratica", identificando valori e disvalori ad essa pertinenti; *in primis*, quindi, una considerazione quasi "obbligatoria" consiste nella natura stessa della *preoccupazione simpatetica* o *sympathetic concern*, per poi considerare, *in secundis*, la relazione, alquanto complicata, fra il "valore neutrale della persona" e la persona stessa.

Se la preoccupazione simpatetica è tale perché ha come oggetto la preoccupazione stessa per qualcuno o qualcosa, ciò significa che l'oggetto cui essa è rivolta ha un valore in sé; d'altro canto, se per *valore* si intende, genericamente, *qualsiasi oggetto di preferenza o di scelta*⁶³ (termine esteso dagli Stoici all'oggetto di scelta o di preferenza tipica del dominio dell'etica - per cui avente una natura morale) allora ciò significa che il benessere dell'altro diviene *valore normativo* per la stessa. Dunque, la simpatia è una preoccupazione durevole per il benessere dell'altro, che implica, per sua stessa natura, un "codice" di valori e disvalori, arbitrariamente riferito ad ogni singolo.

La ricerca dell'universalità tipica dei costrutti morali, cioè, non pertiene la sua natura, poiché, essendo diretta da ogni agente a "oggetti di preoccupazione simpatetica", distinti a seconda

⁶¹ *IVI*.

⁶² STEPHEN DARWALL, *Empathy, Sympathy, Care*, in "Philosophical Studies", 89, 1998, pp. 261-282, p. 261.

⁶³ NICOLA ABBAGNANO, *Dizionario di Filosofia*, Utet Libreria, Torino, 2006, lemma *valore*.

delle proprie aspettative, inclinazioni, predisposizioni e temperamento, non può avere un fine teleologico, ma solo un fine in se stessa per se stessi. Ciò comporta, inevitabilmente, che nella “scala di priorità” della simpatia risultino esserci dei “valori” riconosciuti tali e dei “non valori”, lontani dalla spinta motivazionale all’agire; i primi sono stati definiti, genericamente, con l’espressione *value for a person*, i secondi *person-neutral value*⁶⁴. La differenza fra l’uno e l’altro è sostanziale: sostenere, infatti, il *value for a person* significa, letteralmente, “dare valore per la persona”, *a favore di* un individuo verso cui l’agente prova preoccupazione, quindi, motivazione ad agire. Si riconoscono importanti per il benessere dell’altro quelle azioni che, certamente, porteranno alla sua salute, al suo vigore, attribuendo, arbitrariamente, valori normativi a talune di esse rispetto ad altre; in questo senso, il fine della simpatia per l’agente che la esperisce ha come oggetto stesso di attenzione l’altro come tale ed è proprio prendendo spunto da quest’attitudine particolare del sentimento che sono nati degli studi sul rapporto fra simpatia ed altruismo; anziché empatia/altruismo.

Tutto ciò, simultaneamente, comporta il “valore neutrale della persona”: la cernita arbitraria dell’oggetto avente *valore* per il fine da perseguire (il benessere dell’altro) comporta, *ad escludendum*, il “non valore” di altro a cui la “morale” *tout court* potrebbe attribuire un valore intrinseco. Non si tratta, cioè, di ammettere o di giustificare la liceità di un atto immorale per il benessere dell’altro, ma di considerare, dal proprio punto di vista per il fine da perseguire, un’azione *indifferente* o *neutrale*: il *personal-neutral value*, infatti, denota proprio tale caratteristica intrinseca dell’azione simpatetica che, andando oltre, potrebbe non esaurirsi nella semplice preoccupazione per il benessere dell’altro in sé, ma potrebbe comportare il desiderio di agire solo ed esclusivamente per “amore” dell’altro. Esulando da qualsiasi prospettiva egoistica ed egocentrica, la spinta ad agire per lenire la sofferenza altrui non sarebbe, cioè, il risultato di una necessità propria di alleviare quella provata su di sé a causa della consapevolezza del dolore che la vittima sta vivendo, ma, al contrario, potrebbe rappresentare, secondo alcune interpretazioni, la spinta all’altruismo o alla *cura*.

Con ciò non si sta tentando di reindirizzare il discorso alla fondazione delle “etiche della simpatia”, ricordi del XVIII secolo in cui la simpatia entrava nel dibattito sulla formazione del giudizio morale, ma di evidenziare quegli elementi sostanziali e differenti comuni tanto alla simpatia quanto all’empatia, per cercare di capirne, effettivamente, la natura.

Il *personal-neutral value*, però, non esaurisce la sua fonte problematica nell’estensione più o meno arbitraria della preoccupazione empatica, ma la estende anche al suo intricato rapporto con il “problema persona”, con cui la morale, inevitabilmente, si scontra. Forse, non si tratta

⁶⁴ STEPHEN DARWALL, *Empathy, Sympathy, Care*, op. cit., p. 275.

neppure di applicare la simpatia alla persona, ma di riflettere se si possa provare simpatia anche per le “non-persone”. Il termine, quasi pleonastico ribadirlo, è oggetto di complicate e controverse discussioni fra intrecci di problemi, di definizioni, di posizioni che, però, in maniera assoluta, si potrebbero generalizzare con l’ausilio di due espressioni: *l’essere una persona* o *l’agire come una persona*.

Il concetto stesso di “indifferenza o di valore neutrale”, infatti, chiama in causa il delicato compito di identificare i soggetti a cui attribuire l’“etichetta persona” per poi domandarsi se e come sia possibile provare simpatia per “entità” distinte da sé.

Per quanto la matassa ritorta della speculazione sia ancora ben lontana dal suo capo, in linee generali, le due qualità sopra citate traggono la loro origine dall’etica della *sacralità della vita* e dall’etica della *qualità della vita*; secondo le quali, rispettivamente, il concetto di persona o attiene a proprietà intrinseche nella natura dell’uomo oppure è condizionato dall’acquisizione progressiva delle capacità cognitive; laddove, infatti, queste ultime dovessero mancare o venir meno, lo *status* di persona muterebbe, irreversibilmente, in quello di *non persona*, rendendo urgente un’ulteriore riflessione sul rapporto simpatia/non persona.

Molto sommariamente, secondo il primo approccio *l’essere una persona* è, come detto, una qualità ontologica intrinseca all’essere umano e qualsiasi individuo vivente, sia esso consapevole di se stesso o sia manchevole di capacità cognitive, è da ritenersi tale⁶⁵; non esisterebbero, quindi, esseri viventi appartenenti alla specie *Homo Sapiens Sapiens* che siano delle *non persone* perché manchevoli di determinati requisiti; per cui *l’agire come una persona* diviene un aspetto collaterale: non si potrebbe inferire da ciò che un individuo compie concretamente nel mondo il suo essere sostanziale.

Non si è, cioè, ciò che si fa o non si riesce a fare.

Di etica completamente contrapposta – e da molti ritenuta radicale – è quella proposta dal filosofo australiano Peter Singer, il quale, portavoce di una bioetica laica, sostiene che se è vero che ogni *persona* ha diritto alla vita, è pur vero che non a tutti gli esseri umani è applicabile “l’etichetta” stessa, poiché manchevoli delle caratteristiche ontologiche che rendono l’individuo *persona* in quanto tale. Un essere, cioè, capace di pensare al futuro e di avere bisogni e desideri, essendo consapevole della propria esistenza nel tempo, è la definizione di *persona* che Singer sostiene: di contro all’idea di una sacralità della mera vita vegetativa, l’aspetto *qualitativo* della stessa deve essere il valore da difendere, connaturato al progetto stesso che ci si è preposti di compiere quali esseri senzienti, coscienti e consapevoli

⁶⁵ Cfr. JOSEF SEIFERT, *La morte cerebrale non è più la morte di fatto. Argomentazioni filosofiche* in ROSANNA BARCARO E PAOLO BECCHI (A CURA DI), *Questioni mortali. L’attuale dibattito sulla morte cerebrale e il problema dei trapianti*, Edizioni Scientifiche Italiane, Napoli, 2004, pp. 77-98.

in linea con le proprie *scelte*. Di conseguenza, rifiutando qualsiasi forma di specismo e condividendo la versione utilitaristica del *principio di uguaglianza*, teorizzato da Henry Sidgwick, secondo cui *il bene di ciascun individuo non è di maggiore importanza, da un punto di vista dell'universo, del bene di ogni individuo*, tale approccio porta all'estensione dello *status* di *persona* tanto agli animali non umani quanto agli umani con quelle particolari accezioni⁶⁶.

È bene chiarire che, seppure Singer riconosce all'interno della società umana due generi di esseri viventi, le persone e le non persone, ciò non significa che queste ultime debbano essere escluse dalla sfera dell'interesse morale, poiché il primo principio della sua etica sostiene che arrecare dolore o lasciar soffrire sia un male: un atteggiamento moralmente disonesto da condannare.

Ciò porta, come sarà facile intuire, ad una problematizzazione del rapporto simpatia/empatia *v/s* altro, poiché non c'è alcuna chiarezza sul fatto che la propensione all'alterità, in qualsiasi forma la si voglia definire, abbia la natura del *prendersi cura* oppure dell'*altruismo*⁶⁷.

Fatto sta che, nonostante si sia ribadita la predisposizione “monodiretta” della simpatia nei confronti di un singolo rispetto “all'eterodirettività” dell'empatia per la comprensione delle condizioni altrui, ad un'accurata analisi sembrerebbe che anche la prima sia condizionata dalle caratteristiche della situazione a cui, apparentemente, risponde il singolo: se, infatti, da una prospettiva in prima persona si crede che il sentimento vissuto dalla vittima non sia giustificabile a causa della situazione che sta vivendo, la preoccupazione simpatetica si “spegne”. Ciò ha fatto ipotizzare a molti che la simpatia non sia altro che una

*projective empathy, as when we share another's feeling as a way of caring (sympathetically) for her [...] When we projectively mirror other's feelings, we not only show them how they feel, we also show them that we agree with them about how to feel. We show we understand their feelings and signal our willingness to participate with them in a common emotional life. This makes projective empathy central to the formation of normative communities – like-minded groups who can agree on norms of feeling*⁶⁸.

Quando l'agente riflette *proiettivamente* i sentimenti dell'altro o degli altri, “subendoli” in prima persona, quest'ultimo sta comunicando non solamente la propria predisposizione a “compartecipare alle sofferenze altrui – quindi, di “empatia” in senso lato –, ma (aspetto

⁶⁶ Cfr. PETER SINGER, *Ripensare la vita. La vecchia morale non serve più*, Il Saggiatore, Milano, 1994.

⁶⁷ Per il rapporto fra *empatia* ed *altruismo* si rimanda al paragrafo successivo.

⁶⁸ STEPHEN DARWALL, *Empathy, Sympathy, Care*, op. cit., pp. 269-270.

importante per il discorso successivo) la propria inclinazione “ad essere d’accordo” con i sentimenti provati: l’individuo, in talune situazioni, si mostra, cioè, di capire il modo in cui sente il Tu e, secondo un’interpretazione del tutto soggettiva della scrivente, *sceglierà, coscientemente*, di voler partecipare ad una vita emotiva comune ad entrambi. In questo senso, allora, la simpatia si mostra come una proiezione del punto di vista altrui, focalizzato su quest’ultimo quale oggetto della predisposizione e, allo stesso tempo, concentrato sul riproporre/sul simulare, consapevolmente, il suo stato emotivo; un fenomeno che, miscelando le caratteristiche dell’empatia e della simpatia, è stato definito *proto-sympathetic empathy*:

*it brings the other’s relation to his situation into view in way that can engage sympathy on his behalf [...] Proto-sympathetic empathy is informed by projective empathy, but goes beyond it in not being felt entirely as from the other’s standpoint (or, at least, not without projected self-consciousness)*⁶⁹.

Come la simpatia, l’empatia proto-simpatetica pone al centro della sua azione l’altro nella difficile situazione emotiva che sta vivendo e, grazie alla proiezione verso l’oggetto della sua azione (il dolore del Tu) non solo la riproduce in sé (empatia), ma va al di là di essa non sentendosi relegata o “intrappolata” nella situazione stessa (simpatia). Il modo in cui, però, l’agente possa definire il tipo di emozione provata in prima persona durante la preoccupazione simpatetica, per come si voglia definire, l’empatia, l’empatia proiettiva o l’empatia proto-simpatetica è ancora fonte di dubbi e perplessità.

Una proposta avanzata da David A. Lishner, Daniel Batson *et alii* sembrerebbe porre al centro della discussione proprio tale interrogativo, identificando, in generale, due tipi di emozioni in grado di sciogliere il dilemma: la tenerezza, da una parte, la simpatia *tout court*, dall’altra⁷⁰. Ad un’analisi attenta, sostengono gli autori, si possono distinguere due forme differenti di bisogno, di cui il primo è definibile quale il “bisogno corrente”, che coinvolge la percezione dell’esistenza di una discrepanza fra una o più dimensioni del benessere attuale di un soggetto rispetto alla sua norma e la “vulnerabilità”. Si ha la constatazione, cioè, di una notevole differenza fra lo *status quo* della vittima e ciò che sarebbe per lei desiderabile, auspicabile e, data la natura semipermanente dello stato di bisogno, si proverebbe *simpatia*. Il secondo bisogno, invece, data la natura casuale dell’incontro con la vittima, l’osservatore proverebbe

⁶⁹ *IDEM*, p. 271.

⁷⁰ DAVID A. LISHNER, DANIEL C. BATSON, ELISABETH HUSS, *Tenderness and Sympathy: Distinct empathic Emotions Elicited by Different Forms of Need*, in “Personality and Social Psychology Bulletin”, 37(5), pp. 614-625.

tenerenza; pertinente, cioè, al sentimento empatico. Un sentimento che spingerebbe l'osservatore ad agire per alleviare non solo la sofferenza della vittima, ma anche la fonte di preoccupazione e di angoscia personale. In questo secondo caso, l'altruismo spiegherebbe i motivi dell'etichetta *moral empathy*, poiché spingerebbe l'agente a compiere atti pro-sociali e/o altruistici, sovrapponendo il sentimento provato in prima persona con quello della *compassione*; un "soffrire con" che, come l'empatia, è oggetto di confuse, numerose e varie definizioni.

Il primo stato di bisogno (con la relativa emozione provata) apre le porte alle ulteriori differenziazioni sulle dimensioni proprie dell'empatia, suddivise in quattro grandi aree di ricerca riconducibili all'*affective (emotional) empathy*, *cognitive empathy*, *behavioural empathy* e, come già detto, alla *moral empathy*⁷¹.

La prima dimensione, conosciuta anche con il nome di *affective matching*, ripropone il concetto di *Einfühlung*, teorizzato da Lipps, in cui si identifica la *capacità di sentire su di sé* i sentimenti o le emozioni dell'altro, coinvolgendo anche l'uso dell'immaginazione; differente, in questo senso, dalla seconda, nella quale il *focus* è la *comprensione* dell'emotività altrui, da una prospettiva non più in prima persona, ma "distaccata". Non a caso, infatti, quest'ultima è definita anche come *detached concern*, ossia "preoccupazione distaccata", obiettiva, senza che l'ipotetico osservatore rievochi in sé il dolore o la sofferenza altrui.

È bene distinguere, a tal punto, la *self-oriented perspective taking* dall'*other-oriented perspective taking*: se in quest'ultima, la prospettiva cognitiva e pragmatica è orientata sull'altro quasi come fosse una sorta (forse lo sarebbe) di simpatia, nel primo caso la prospettiva da cui guardare l'altro è, inizialmente, se stessi; subentra, cioè, una categoria "come se", una forma di identificazione con l'altro che fa capire e provare lo stato d'animo osservato all'osservatore.

I rischi della *cognitive empathy*, stante la *self-oriented perspective taking*, sono il fallimento della comprensione altrui (perché troppo presi da come ci si sentirebbe in quella situazione – il punto di vista, vale a dire, rimarrebbe l'osservatore e non la vittima); errori di predizioni, false assunzioni e sentimento di angoscia personale. Per l'empatia cognitiva, infatti, è necessaria una flessibilità e plasticità mentale tipica del "cambio di prospettiva".

Infine, la *behavioural empathy*, utile, soprattutto, in contesti clinici/terapeutici, è relativa alla capacità di comunicare la comprensione della precedente dimensione empatica al paziente,

⁷¹ DAVID JEFFREY, *Empathy, sympathy and compassion in healthcare: is there a problem? Is there a difference? Does it matter?*, in "Journal of Royal Society of Medicine", Vol. 109 (12), pp. 446-452.

appoggiandosi, inoltre, all'*affective empathy*, affinché si senta su di sé la sofferenza altrui, senza tralasciare la visione del mondo del paziente stesso (competenza empatica).

Stando, quindi, a quanto fin qui detto, sembrerebbe da molti condivisibile sostenere che

*sympathy is an emotion in its own right, whereas empathy is a way of experiencing emotions*⁷²:

un modo, vale a dire, di percepire le emozioni che ha portato, come si vedrà nel prosieguo, ad attribuirle un compito prettamente epistemologico. Allontanandosi, però, dalla complessità di questo tipo di *empathy*, è necessario chiarire che non sempre l'esperienza emotiva vissuta in prima persona equivale o a quest'ultima o alla preoccupazione empatica; poiché, frequentemente, ci si trova a vivere l'esperienza del *contagio emotivo*: una sorta di "infezione emotiva" che si percepisce immediatamente, non mediata da alcuna consapevolezza coscienziale. È come se l'individuo, ad esempio, entrando in una stanza in cui tutti i suoi colleghi ridono, catturasse *in medias res* l'ilarità e si adegasse al contesto; la conseguenza sarebbe una sentita risata, probabilmente senza neppure conoscere il motivo della stessa.

Questa peculiarità, forse prettamente umana, è stata spesso accostata o alla stessa simpatia o all'empatia *tout court* come uno strumento istintivo, di base, tramite cui l'individuo esperirà, poi, lo sviluppo cognitivo del suo sentire. Autore della "confusione" *post hoc ergo propter hoc* fu proprio Hume, il quale più volte ne *I principi della morale* sembra utilizzare il concetto di contagio emotivo dapprima come percezione immediata di comprensione dello stato altrui, poi nei termini attuali di simpatia o predisposizione/preoccupazione verso l'altro, infine, quale benevolenza universale.

Coloro i quali sostengono che il contagio emotivo sia una forma d'empatia trovano supporto in alcuni recenti studi sui suoi meccanismi neurobiologici: uno degli "imputati" sarebbe proprio il processo di *mimicry*; ma, a ben vedere, anche in questo caso si dovrebbe essere cauti nelle generalizzazioni.

Infatti, bisognerebbe capire, innanzitutto, di quale empatia si stia parlando o di quale tipo di approccio. Ciò evidenzia, ancora una volta, che, come Aristotele sosteneva già qualche millennio fa, l'uomo non può non essere un *animale sociale*.

Uno degli interrogativi più importanti e controversi sulla natura, rispettivamente, dell'empatia e della simpatia resta comprendere *in che modo* gli individui riescano a padroneggiare,

⁷² HEIDI L. MAIBOM, *Feeling for Others: Empathy, Sympathy and Morality*, in Special Issue "Other Minds and Moral Agency", "Inquiry", 52, 2009, pp. 483-499, p.489.

inconsapevolmente e senza alcun tipo di educazione in merito, la relazione con l'altro da sé, comprendendone le emozioni provate:

*One of the most important questions about sympathy and empathy is how the person does it, how the other person – the experimenter, client or sufferer – knows when someone else is sympathizing or empathizing, and what the sympathizer or empathizer feels when he or she is doing it. Perhaps this ability to transpose the ego is a unitary process common to both sympathy and empathy? Or perhaps two different processes are involved one peculiar to sympathy and another for empathy?*⁷³

Si torna, cioè, al punto di partenza.

2.3 Empatia ed altruismo: alle origini della morale?

Contrariamente a quanti sostengono che il pensiero occidentale abbia, nella sua storia, cambiato modi e metodi della ricerca filosofica a partire da Nietzsche, la scrivente si accoda a quanti, di contro, ritengono che quel cambiamento sia avvenuto *post* Hegel; con il Filosofo romantico, infatti, i problemi classici del pensiero hanno trovato una loro soluzione: la dicotomia tra spirito e materia è stata *risolta* nel movimento dialettico e tutto ciò ha donato alla tradizione successiva un modo nuovo, diverso, di approcciarsi alla filosofia stessa.

Seppur da molti criticato per non aver conferito all'uomo uno *status* privilegiato, il concetto di *individuo* in Hegel è rimesso nelle mani di una *ratio* globale che spieghi nella sua *realtà* il continuo ed infinito movimento dello Spirito a discapito di qualsiasi determinazione che possa far pensare ad *una notte in cui tutte le vacche sono nere*⁷⁴. Tolti gli *individui cosmico-storici* di cui lo Spirito stesso si serve per la concretizzazione e la realizzazione di sé, l'uomo diviene mero mezzo nel sistema razionale (nonostante un'attenzione particolare alla dicotomia corpo/"anima" sia, comunque, affrontata dallo stesso Hegel) per un "fine superiore", quale, appunto, la risultante dell'*Idea in sé e per sé*. Quando si pensa, infatti, al "problema dell'intersoggettività" nella *Fenomenologia dello Spirito* si riporta la "mente", immediatamente, alla dialettica servo/padrone da cui si potrebbe evincere che la libertà sia "solo" un "saper fare" e che i rapporti umani, riprendendo Hobbes, siano la realizzazione

⁷³ LAUREN WISPÈ, *The distinction between sympathy and empathy*, op. cit., pp. 317- 318.

⁷⁴ GEORGE WILHELM FRIEDRICH HEGEL, *Fenomenologia dello Spirito*, Bompiani, Milano, 2000, p. 67.

pratica dell'*homo homini lupus*, successivamente “risolti” in una “convinzione” di libertà “immateriale” che ricorderebbe, seppur con particolari determinazioni, quella precedente di Tommaso Campanella. Si sarebbe propensi a ritenere che, stante quanto riportato, il problema della inter-relazione fra individui nella dialettica hegeliana comporti o una sudditanza (seppur solo fisica) dell'altro oppure una falsa libertà del singolo che, scopertosi quale mera determinazione, sia in continua ricerca della sua unicità. Interpretando in tal modo il pensiero di Hegel, si sarebbe inclini a sostenere che, contrariamente a Leibnitz (in cui la monade risulta essere un microcosmo sufficiente di per sé) e per la natura stessa dello Spirito del Mondo, l'individuo hegeliano sia manchevole e che tenti di realizzare se stesso solo attraverso l'incontro-scontro con le determinatezze (apparentemente solo) materiali, prima, e con l'altro, dopo, venendo meno al concetto stesso di *individuo* quale essere “non divisibile”, un *unicum* fra *unici*. Se si dovesse accettare questa parziale ermeneutica, allora, si converrà, conseguentemente, che parlare di *altruismo* nell'opera del 1807 sarebbe sostanzialmente difficile, poiché “vivere per gli altri” (come dirà Comte nel 1852, ideatore del neologismo), nel senso odierno del termine, era al di fuori dell'intenzione di Hegel o, quantomeno, al di fuori di un'attenzione così “analitica” che nascerà negli ultimi anni del XX secolo e nei primi del XXI.

Volendo, però, continuare ad osare e sperando di non far saltare l'ipotetico lettore sulla sedia, si potrebbe sostenere che per quanto Hegel e Nietzsche siano due filosofi sostanzialmente contrapposti l'uno all'altro, in realtà, con una certa forzatura, si potrebbero definire simili. Identificando, infatti, l'amore per il prossimo con l'altruismo, Nietzsche, lo condanna attraverso la figura di Zarathustra, il quale sostiene che:

*Voi andate al prossimo sfuggendo a voi stessi e vorreste far di ciò una virtù; ma io lo leggo bene attraverso il vostro altruismo [...] Voi non sapete sopportare voi stessi e non vi amate abbastanza: ed ecco che volete sedurre il vostro prossimo inducendolo all'amore e farvi belli del suo amore*⁷⁵.

Il punto di congiunzione (e di differenza, allo stesso tempo) si concentra proprio sul concetto di individuo. Tolte, infatti, le contrapposizioni alla stragrande maggioranza note, si dovrebbe altresì notare che in entrambi quest'ultimo è un essere incompleto, manchevole e parziale che tenta di raggiungere la sua interezza attraverso l'incontro/scontro con l'altro, in Hegel, attraverso il solo “incontro”, in Nietzsche. Volendo andare oltre, in quest'ultimo la ricerca

⁷⁵ FRIEDRICH NIETZSCHE, *Così parlò Zarathustra*, Bompiani, Milano, 2010, pp. 56-57.

della sua completezza diviene non solo utilitaristica, ma anche subdola: solo attraverso il mezzo dell'altro che si manifesta con il suo "prendersi cura", l'individuo, consciamente, tende a cercare l'amore verso di sé, poiché solo ciò che gli sarà donato dall'altro sarà utile al suo fine: la ricerca della sua unicità.

In Hegel, invece, la *ratio* del processo è sempre e solo dialettica: è una libertà necessitata che l'individuo si incontri e si scontri con l'altro per la ricerca non tanto di sé, quanto della concretizzazione della vita dello Spirito; per cui si è ben lontani da qualsiasi mezzo utilitaristico conscio della singolarità per ricercare, dall'occhio onniscente dello Spirito stesso, la sua unicità. Esiste, infatti, solo ciò che avverrà: necessariamente in Hegel, "al di là del bene e del male" in Nietzsche.

Ed è proprio l'interrogativo se l'interesse umano per l'altro vada "al di là del bene e del male" che ha spinto i molti ad interrogarsi su un fenomeno, l'altruismo, che sembrerebbe andare contro l'idea che l'uomo sia di per sé egoista: che cosa motiva un individuo a prestare aiuto ad un suo simile⁷⁶, rimettendoci risorse e possibilità di avvantaggiare la trasmissione dei propri geni alle successive generazioni?

Richard Dawkins, filosofo statunitense, risponderebbe chiamando in causa "il gene egoista", riportando il discorso ad un duplice piano d'indagine: evolucionistico/biologico, da una parte; culturale (il concetto di "meme"), dall'altro; ciò, però, condurrebbe l'analisi ad una presa d'atto di un fenomeno di cui si vuole capire la *ratio*, l'altruismo, ovviando, in tal modo, ai progressivi interrogativi sulle affinità o sulle dissonanze con l'empatia⁷⁷.

Oggetto di dibattito dalla filosofia moderna in poi con questa medesima "etichetta", l'altruismo denota, in genere, quell'atteggiamento disinteressato verso il Tu, che pone in primo piano il benessere dell'altro anziché del proprio, evidenziando che il tratto distintivo di qualsiasi comportamento altruistico *versus* quello egoistico è la sua motivazione: se nel primo l'oggetto di interesse per chi aiuta è il raggiungimento del benessere altrui a discapito del proprio, nel secondo il comportamento verso l'altro diviene il mezzo, e non il fine, del suo fare.

Anche se, sostanzialmente, ci si riferisce a due concetti distinti, altruismo ed egoismo possono presentarsi, fenomenologicamente, come il medesimo fenomeno: entrambi, infatti, sono mossi da una motivazione propria; hanno uno scopo ultimo da raggiungere e, infine, sono interessati

⁷⁶ In questa sede, per motivi di scelta legata alla complessità del fenomeno ed alla ristrettezza a cui un testo deve necessariamente rispondere, si limiterà la ricerca ai soli animali umani.

⁷⁷ Cfr. RICHARD DAWKINS, *Il gene egoista. La parte immortale di ogni essere vivente*, Edizioni Mondadori, Milano, 1995.

all'incremento del benessere; il problema morale, a quel punto, diviene stabilire verso *chi* sia rivolto quest'ultimo.

Stante l'accento sull'importanza che la motivazione o le *intenzioni* di chi agisce *per* l'altro pongono sull'identificazione di un fenomeno anziché di un altro, è importante notare che esiste una differenza sostanziale fra il mero "comportamento dedito all'aiuto" ed il "comportamento altruistico" propriamente detto.

Tale distinzione tende a definire un "limite" fra i due, basandosi sull'assenza di una motivazione psicologica, nel primo, di contro alla sua presenza, nel secondo, che interesserebbe la capacità cognitiva dedita al *mindreading* dell'agente propenso all'aiuto.

L'esempio tipico dei pipistrelli, i quali rinunciano ad un pasto di sangue per favorire il compagno in difficoltà, potrebbe rientrare nel primo caso; poiché, attenendosi al "vocabolario evolucionistico", questi ultimi adopererebbero un comportamento che ridurrebbe la propria *fitness* riproduttiva nella futura generazione a favore di un conspecifico all'interno dello stesso gruppo di appartenenza.

In tal modo, i biologi sono in grado di spiegare molti dei comportamenti presenti in natura dagli organismi più semplici a quelli più complessi (fra cui l'uomo) sotto l'etichetta di "altruistici" nell'accezione di "comportamento dedito all'aiuto". E le abilità di *lettura delle altre menti*?

Non nasce dal nulla, quindi, il dubbio che sorge a chi scrive sull'eventuale importanza da attribuire alla presenza o all'assenza della "riflessione cognitiva" in chi si prodiga per un proprio conspecifico, poiché, per ciò che si avrà modo di leggere più avanti, una parte del dibattito in corso sul rapporto altruismo/empatia (nell'uomo) darebbe per presupposto o sarebbe incline a condividere che quando si parla di altruismo *tout court* si dovrebbe constatare la presenza di alcuni elementi fondanti: la capacità del *mindreading* o *lettura della mente* sarebbe uno fra questi. Uno sforzo ed un'abilità cognitiva che lascerebbe presupporre che il "comportamento dedito all'aiuto" sia un fenomeno distinto dal "comportamento altruistico"; non migliore o peggiore, ma semplicemente differente per cui non si dovrebbero far rientrare in un unico grande insieme eventi distinti, denotati con una medesima etichetta.

Laddove, ad esempio, si presupponesse la presenza di un "pensiero" in specie biologicamente semplici (ad esempio le amebe) in grado di porre l'accento su quelle che, comunemente, sono identificate "facoltà intellettive" tipiche di un cervello in grado di elaborare un comportamento motivato da ragioni, allora, probabilmente, si dovrebbe cominciare a pensare che anche specie meno complesse siano dotate di una "protocoscienza" o "consapevolezza primaria" o "coscienza" in grado di operare una riflessione (non importerebbe, a quel punto,

se fosse in grado di giungere o eguagliare o superare i livelli di quella umana). Ciò sarebbe sufficiente a definirle “esseri consci”, estendendo non soltanto il “dominio della coscienza” anche a tali organismi, ma inserendoli, di fatto, nella riflessione morale con tutto ciò che ne conseguirebbe.

Si dovrebbe operare, a quel punto, una ridefinizione del fenomeno *altruismo*, poiché, stante le nuove premesse, il “comportamento dedito all’aiuto” rimarrebbe un’espressione vuota: qualsiasi comportamento verso il prossimo dovrebbe essere identificato come “altruistico”; non più, cioè, spiegazioni di stampo bio-evoluzionistiche (l’aumento delle possibilità per un conspecifico anziché per sé di trasmettere il proprio corredo genetico alle prossime generazioni), ma morali, in cui l’agente (sia esso un’ostrica o un essere umano) sia motivato ad aprirsi verso il prossimo per l’incremento del benessere altrui.

Ci si dovrebbe trovare di fronte a ciò che viene definito *altruismo psicologico* di contro al mero *altruismo evoluzionistico*: nel primo, la motivazione all’aiuto avrebbe origini cognitive; nel secondo, la motivazione si declinerebbe come “preoccupazione istintiva ed istintuale” di matrice organica.

Riassumendo, in contrapposizione a ciò che sostengono, attualmente, biologi ed etologi, si dovrebbe parlare di *altruismo* in tutti gli organismi: da quelli più semplici a quelli più complessi, poiché ognuno sarebbe dotato di “coscienza”, di “mente”, di “pensiero” e di “consapevolezza”, eliminando, in tal modo, spiegazioni di fenomeni che non rientrino nel cosiddetto “altruismo morale”.

Una seconda ipotesi prevede che se tale riflessione dovesse dare esiti negativi (con certezza, vale a dire, si stabilisse l’assenza di coscienza negli organismi più semplici⁷⁸) allora la definizione di “altruismo quale comportamento dedito all’aiuto dell’altro in mancanza di motivazione” dovrebbe essere cambiata: se *altruismo* significa “incrementare il benessere dell’altro”, quale conseguenza di specifiche intenzioni dell’agente di giungere a questo particolare fine, non si potrebbe definire con la medesima “etichetta” un comportamento privo di una caratteristica essenziale del fenomeno stesso. In questo secondo caso, cioè, l’accostamento di un concetto morale ad uno biologico (per esemplificare il discorso) non troverebbe il *trait d’union*: a quel punto, in termini disgiuntivi e non inclusivi, o si cambierebbe il concetto di altruismo nel primo ambito per sposare quello del secondo o viceversa; *tertium non datur*.

⁷⁸ L’uso di un linguaggio deciso e perentorio è stata una scelta ponderata per marcare, ancora di più, le difficoltà di una eventuale risposta *tranchant*.

Tale discorso fornisce il pretesto per poter distinguere il secondo approccio al discorso altruistico che, come visto, è di stampo prettamente etico: con l'etichetta *moral altruism*, infatti, ci si riferisce proprio alla presenza di una motivazione ad agire per (e solo *per*) l'altro; la quale, come detto, fornisce "la linea di demarcazione" fra altruismo *tout court* ed egoismo. Un egoismo che potrebbe delinearsi in tre modi diversi:

1. un *egoismo psicologico*, in cui la motivazione ad agire per l'altro deriverebbe dalla ricompensa successiva all'azione in sé: uno stato edonistico o di egocentrismo smisurato farebbe sì, come sostenne Zarathustra, da poter "ricevere" il plauso emotivo o concreto dell'ipotetica vittima o della comunità purché l'Ego dell'apparente altruista possa essere messo in risalto;
2. un *egoismo etico*, a differenza del primo, in tale contesto cambiano, leggermente, le ragioni che muovono all'azione: l'egocentrico etico, vale a dire, è convinto che aiutare sia di aiuto a se stesso non per *vana gloria*, ma perché conscio che il ritorno emotivo della sua azione possa fornirgli strumenti utili alla sua "sopravvivenza psicologica". In questo senso, seppur ci si trova di fronte ad una mancanza di altruismo propriamente detto, forse, individui appartenenti alla categoria "egocentrici etici" potrebbero essere più individui psicologicamente fragili che meri calcolatori edonistici. Ciò non significa giustificare o alleviare le loro colpe morali, ma semplicemente evidenziare una differenza (oserei dire sostanziale) fra un tipo di egoismo ed un altro;
3. un *egoismo avente a che fare con la lettura della mente*, ossia un agente che, pur riconoscendo che l'altro si trovi in uno stato di necessità, decide di non intervenire.

Non stupisce, allora, il ricorso all'equazione

Egoismo = Immoralità,

quando nel senso comune si sente parlare di egocentrismo, ma, come già nel 1971 John Rawls ha avuto modo di illustrare, il sillogismo secondo cui si arriverebbe al risultato sopra indicato è erroneo: contrariamente a quanto si pensi, anche da un comportamento egoistico può derivare un comportamento virtuoso⁷⁹.

Trasformando il fine in mezzo, infatti, l'egoista potrebbe essere incline ad aiutare un proprio conspecifico in difficoltà non per *amor suo*, ma perché da ciò (mezzo) otterrà un bene (egoistico) maggiore: che sia gloria, ammirazione altrui o la consapevolezza di non incorrere

⁷⁹ Per ulteriori approfondimenti, si rimanda alle riflessioni di John Rawls sull'inconsistenza del sillogismo che legherebbe l'egoismo all'immoralità. Cfr. JOHN RAWLS, *A theory of justice*, Harvard University Press, Harvard, 1971.

in problemi giudiziari, questa forma di altruismo di per sé è di difficile catalogazione. È un ibrido fra ciò che, volendo riportare il discorso ad una esemplificazione funzionale per l'oggetto di indagine, le categorie nette del "buono" e del "cattivo" definiscono come tale: da una parte, si presenta l'aiuto oggettivo che l'agente fornisce all'altro il quale, effettivamente, incrementa il suo benessere; dall'altra, il Tu è ridotto a mero mezzo funzionale da chi, avendo le potenzialità, ha *scelto* di adoperarsi per l'altro per un fine diverso rispetto al benessere del Tu: costui, cioè, ha aiutato con la speranza di una ricompensa che gli donerà maggior piacere o che gli sarà maggiormente utile.

Se si dovesse porre l'accento solo sull'azione in sé, sarebbe difficile "giudicarla" non altruistica, poiché *de facto* l'aiuto c'è stato e le attese dell'agente di una ricompensa maggiore che non il benessere della vittima potrebbero essere disattese per motivi contingenti: potrebbe anche avvenire che il *dopo* (la speranza della ricompensa egocentrica dell'agente) di quell'azione non si verifichi mai. In tal caso, all'agente stesso rimarrebbe il mezzo senza il fine – dalla sua prospettiva; il fine in sé – da quella della vittima.

Si potrebbe, allora, definire non morale un'azione in cui il fine sia stato raggiunto (incremento benessere dell'altro), indipendentemente dalle motivazioni morali sottese a quella medesima azione da parte di chi si è prestato ad aiutare? In altri termini, ci si chiede se sia corretto impostare una riflessione sulla considerazione di un'azione quale altruistica o non altruistica, prescindendo dalle reali intenzioni che hanno spinto l'ipotetico osservatore a prodigarsi verso il prossimo.

Per un approccio meramente fenomenologico, crediamo di sì, poiché, stante la sua ottica, lo scopo è stato raggiunto: l'incremento del benessere della vittima.

Ad avviso di chi scrive il problema ed i limiti di una risposta siffatta si trovano in ciò che "non è scritto": il fine raggiunto non è tale di per sé, ma è solo *apparente*. Manca la constatazione o la presa d'atto di ciò che "sta dietro" il mero fenomeno "oggettivo", ossia le intenzioni, le motivazioni, la scelta o la "consapevolezza" che rendono l'uomo⁸⁰ un essere morale. L'intera tradizione teoretica e filosofica del pensiero ci insegnano che "ciò che è presente nel mondo" non è "il mondo in sé", ma è la percezione soggettiva che l'uomo in quanto tale se ne dà, partendo dai limiti strutturali e cognitivi che presenta biologicamente, costituzionalmente, fisiologicamente, prescindendo da qualsiasi falsa pretesa di obiettività. Per rimanere sull'oggetto di indagine, ancora una volta si evince che le *intenzioni* dell'agente

⁸⁰ Si ricorda che è stata condotta una scelta a monte nell'attenersi solo alla specie umana non per convinzioni specieiste, ma per restringere il dibattito all'oggetto d'indagine del presente lavoro.

e non il suo mero comportamento oggettivo fanno la differenza fra ciò che è altruistico e ciò che non lo è.

Riprendendo, se si vuole, l'imperativo categorico kantiano, secondo cui

*Agisci in modo da trattare l'umanità, sia nella tua persona sia in quella di ogni altro, sempre come fine e mai semplicemente come mezzo*⁸¹,

dovrebbe divenir ovvio, a questo punto, che sia legittimo l'interrogativo sull'eventuale relazione fra empatia ed altruismo, poiché entrambi hanno la loro ragion d'essere nell'intersoggettività.

L'una sarebbe causa del secondo o il secondo sarebbe causa della prima?

Il motivo per cui si tende a sottolineare una eventuale "progressione" tra un prima (altruismo) ed un dopo (empatia) trova la sua giustificazione in una serie di studi molto recenti; i quali, tra filosofia e psicologia sperimentale, stanno tentando (il condizionale e la prudenza sono d'obbligo!) di spiegare la natura del fenomeno empatico a partire proprio dall'altruismo, ponendolo quale motivazione pro-sociale all'agire morale. Dall'interesse gratuito verso il prossimo si genererebbe nell'individuo la spinta ad aiutare l'altro da sé (altruismo) che, concretizzandosi in un aiuto materiale, farebbe nascere in chi aiuta l'empatia.

L'altruismo e l'empatia, quindi, sembrerebbero essere candidati, da una parte, ad una *indifferenza morale*, poiché "al di fuori" di qualsiasi valutazione dell'agire morale; dall'altra, ad una parte sostanziale della morale stessa, poiché parti integranti del *comportamento*: non si rientrerebbe, cioè, in un ambito a-morale o "oltre" la morale, ma nel cuore della stessa, tentando di capirne, per certi aspetti, l'origine.

Complice del fiorente clima culturale è lo sviluppo delle scienze del cervello, il quale, oltre ad aggiornarsi con dati sempre più recenti e all'avanguardia sull'ipotetico funzionamento cerebrale, fornisce un valido aiuto al dibattito stesso, partendo *ab origine*: si pensi, ad esempio, al crescente interesse per la psicologia evolutiva del bambino, vista da molti quale possibilità "concreta" di "interrogare e di rispondere" ai dilemmi sulle capacità cognitivo-morali dell'essere umano.

L'altruismo e l'empatia rientrano proprio in questo discorso.

Se è vero che alcuni recenti studi pongono un prima ed un dopo al binomio oggetto di interesse in questa sede

(altruismo → empatia → comportamento morale)

⁸¹ IMMANUEL KANT, *Fondazione metafisica dei costumi*, UTET, Torino, 1995, p. 88.

ce ne sono altrettanti che, di contro, sostengono che la causa dell'altruismo sia la sofferenza o l'angoscia empatica; la quale, attraverso una motivazione ad agire, disinteressatamente, verso il prossimo, condurrebbe, concretamente, ad un comportamento morale

(sofferenza empatica → altruismo → comportamento morale).

L'esemplificazione di un dibattito tanto complesso e trasversale può essere fornito da due distinti modelli ad opera – cronologicamente – di due psicologi statunitensi: Martin Hoffman, da una parte; Daniel Batson, dall'altra.

Partendo dal presupposto che l'empatia sia il frutto della selezione naturale, Hoffman sostiene che quest'ultima sia (parafrasando) il mezzo o l'autore attraverso il quale l'uomo si interessi al benessere dell'altro, motivandolo ad un aiuto disinteressato; l'empatia, infatti, coerentemente a quanto sostenuto, è definita dallo psicologo statunitense

*come una risposta affettiva più consona alla situazione di un altro che non alla propria*⁸²,

in cui l'aspetto caratterizzante consta della presenza di sofferenza empatica nella ipotetica relazione che si instaurerebbe tra l'osservatore (che proverebbe, di riflesso, la medesima angoscia) e la vittima. Se è vero, inoltre, che essa ha un'origine naturale, deve essere altrettanto vero che sia

*una risposta multideterminata, che può essere suscitata da segni o indizi di sofferenza provenienti dalla vittima o dalla situazione in cui essa si trova*⁸³.

Stante questi presupposti, l'analisi di Hoffman prende in esame cinque diversi dilemmi morali nati da particolari situazioni morali (qui la ripetizione del termine è necessaria) in cui l'autore si interroga sull'eventuale risposta comportamentale dell'ipotetico osservatore, poiché tali situazioni forniranno, secondo il quadro generale della sua teoria, il pretesto per approfondire la maggior parte delle situazioni morali antesignane di ciò che sarà la motivazione ad agire.

A maggiore esplicitazione si riporta quanto, a tal proposito, sostiene Hoffman:

⁸² MARTIN HOFFMAN, *Empatia e Sviluppo Morale*, op. cit., p. 24.

⁸³ *IBIDEM*, p. 25.

1. *Nel primo tipo, il più semplice, vi è uno spettatore innocente del dolore o della sofferenza altrui (di tipo fisico, emotivo, economico). Il problema morale è: la persona darà aiuto? E, se non lo fa, come si sentirà?*
2. *Il secondo tipo implica un trasgressore che nuoce o è in procinto di nuocere a qualcuno (in modo accidentale, in una lotta, in una disputa). Qui il problema morale è: la persona eviterà di nuocere all'altro? E, in caso contrario, si sentirà poi colpevole?*
3. *Nel terzo tipo di dilemma, che combina elementi dei primi due, vi è un trasgressore virtuale che, pur essendo innocente, crede di aver fatto del male a qualcuno.*
4. *Il quarto tipo è più complesso: implica più parti morali (multiple moral claimants) tra le quali la persona è costretta a scegliere. Il problema morale è: a chi dare aiuto? E la persona si sentirà colpevole per aver trascurato gli altri?*
5. *Il quinto tipo, nel quale il prendersi cura si contrappone alla giustizia, implica non solo più parti morali, ma anche un conflitto fra la considerazione per il prossimo e temi più astratti quali i diritti, il dovere, la reciprocità. Il problema morale, in questo caso è: quale principio prevarrà, la cura o la giustizia? E ci sentiremo colpevoli per aver violato l'altro principio?⁸⁴.*

Il *punctum dolens* concerne l'origine stessa dell'empatia: stante la sua natura evuzionistica, il problema di Hoffman è individuarne le ragioni o le motivazioni che inducano un individuo a farsi carico del peso emotivo esperito prima dalla vittima e poi da se stesso, tenendo presenti, però, tutti i limiti, i particolarismi e le debolezze cui il fenomeno incorre.

Non è un caso che, in linea con i dibattiti recenti, consolidati, soprattutto, dagli anni '90 in poi – data simbolo della scoperta dei neuroni specchio – Hoffman differenzi i modi di attivazione empatica: da una parte, preverbale, automatica ed involontaria; dall'altra, consapevole e conscia.

Nell'ipotetico modello di spiegazione più semplice dell'attivazione empatica, che porterà, successivamente, l'individuo ad intervenire (altruismo) per l'altro in difficoltà, Hoffman presenta due attori principali: l'osservatore e la vittima in cui l'empatia risulta essere una risposta passiva ed involontaria, dipendente da stimoli superficiali, che interessa la presenza fisica, nello stesso luogo, di entrambi i protagonisti. Infatti, sostiene Hoffman:

⁸⁴ *IDEM*, p. 24.

*questa semplice forma di sofferenza empatica [che non richiede nessun particolare sforzo cognitivo dell'osservatore nei confronti della vittima, n.d.r.] è tuttavia importante, proprio perché mostra che gli esseri umani sono fatti in modo tale da poter provare, involontariamente e intensamente, le emozioni di un'altra persona – che spesso la loro sofferenza dipende da un'esperienza dolorosa altrui, piuttosto che propria*⁸⁵.

Ad escludendum, se ne deduce che l'altra modalità di attivazione empatica in un modello ideale che, stavolta, non prevede né la sola presenza *in loco* dei due attori, né il solo rapporto bidirezionale fra i due, consisterà in un contesto più complesso (quarto e quinto tipo, sopra citato dallo stesso Hoffman) in cui la presenza di sofferenza empatica nell'osservatore richiederà uno sforzo cognitivo dello stesso per riproporre in sé, consciamente, l'ipotetica angoscia provata da una vittima ideale: sono questi i casi, precisa lo psicologo statunitense, in cui la sofferenza empatica incontra i gravi limiti strutturali che la costituiscono, dando luogo, alle volte, ad uno “spegnimento” dell'empatia provata dall'osservatore. Ciò favorisce la nascita dell'interrogativo su come sia possibile, arrivati ad un certo punto della relazione ipotetica osservatore/vittima, la “fine” della *lettura della mente* dell'altro.

Quando si sente parlare, infatti, di *mindreading* legato all'empatia, ci si riferisce proprio alle ipotetiche situazioni, che Hoffman come tanti altri hanno delineato, in cui l'“entrare nei panni altrui” diventa, sommariamente, “operare uno sforzo cognitivo tale da leggere la mente dell'altro”: a quel punto, come si avrà modo di leggere nei capitoli che seguiranno, il problema diventa identificare il rapporto sussistente fra *metacognition* e *mindreading*; poiché l'aspetto più controverso del problema empatico non riguarda tanto quell'attivazione preconsa, preverbale, involontaria (rimanendo nel linguaggio di Hoffman) quanto quella conscia e “razionale”.

Nell'ipotetico modello stilistico del rapporto osservatore-vittima, lo psicologo statunitense sostiene che tre sono le forme automatiche di attivazione empatica nell'osservatore che porteranno alla comparsa dell'altruismo:

1. La *mimesi*;
2. Il *condizionamento classico*;
3. L'*associazione diretta*,

le quali, senza tornarvi sopra⁸⁶, descrivono “neuroscientificamente” il processo attraverso il quale nell'osservatore si genera la sofferenza empatica, spinta motivazionale all'agire morale.

⁸⁵ *IBIDEM*, p. 25.

⁸⁶ Si veda a tal proposito, il paragrafo 1.1.

La motivazione pro-sociale o morale viene definita da Hoffman come ciò che:

a) ha un carattere irresistibile ed obbligatorio; b) è qualcosa che sentiamo provenire dall'interno; c) ci fa sentire colpevoli quando compiamo o consideriamo la possibilità di compiere delle azioni che danneggino gli altri; d) fa sì che teniamo conto dei bisogni degli altri anche quando sono in conflitto con i nostri. Quando questo conflitto ha luogo, i processi attivatori dell'empatia che agiscono nella situazione dello spettatore possono non avere abbastanza forza da motivare una persona ad agire in modo prosociale⁸⁷.

Per quanto riguarda, invece, le due forme di attivazione empatica cognitivamente rilevanti (la *mediazione verbale* e *l'assunzione di ruolo*), Hoffman sostiene che hanno il pregio di estendere l'interesse empatico dell'ipotetico osservatore a persone lontane dalla sua vista e dallo stazionamento nello stesso luogo; interessando, come detto, quello sforzo cognitivo oggetto di un crescente e fiorente dibattito in merito.

Il problema "altruismo", perciò, sembrerebbe essere, per le prime forme di attivazione empatica, il risultato immediato di un processo naturale ("neuroscientifico") quasi come fosse la "conseguenza logica" di un'equazione i cui termini fossero due variabili ben definite: "funzionando" in un certo modo, giungeremmo, attraverso quella spinta motivazionale all'aiuto disinteressato verso l'altro, alla comparsa di un comportamento morale; nella seconda situazione, invece, l'altruismo diverrebbe non una spinta "istintiva", ma un prodigarsi verso il prossimo, risultato di una riflessione conscia, ponderata e razionale, facendo nascere il problema non più sul binomio empatia/altruismo, ma sulla natura stessa dell'altruismo *tout court*.

La differenza sostanziale fra il modello appena esposto di Hoffman e quello proposto dal suo collega Batson si basa sul ruolo affidato alla sofferenza empatica: se nel primo, come visto, essa diviene la causa che ha come effetto l'altruismo, nel secondo, l'altruismo diviene la causa e non l'effetto,

La sofferenza empatica, cioè, rientra nel modello di spiegazione pro-sociale non come protagonista indiscussa, ma come corollario ad una natura umana *sostanzialmente* altruistica.

Per giungere a sostenere questa tesi teoreticamente impegnativa, Batson *et alii* hanno messo alla prova la "naturale empatia" di ogni essere umano attraverso l'ideazione di alcuni esperimenti.

⁸⁷ MARTIN HOFFMAN, *Empatia e Sviluppo Morale*, op. cit., p. 31.

Il *team*, manipolando il livello di empatia e la percezione degli spettatori in una situazione anomala per la vittima in cui suscitasse, effettivamente, sofferenza empatica, ha evidenziato che non è l'empatia stessa a motivare all'altruismo, ma è la naturale tendenza dell'essere umano a prodigarsi per chi è in difficoltà a far nascere, successivamente, l'empatia. Non ci si muove all'azione per lenire una propria angoscia emotiva, nata dall'osservazione di uno stato di necessità del prossimo (come sostiene Hoffman quando parla dei rischi in cui incorre il fenomeno empatico), ma l'"attivazione emotiva" diviene il risultato di un "dirigersi verso l'altro", suscitato dalla percezione di un'assenza o di una diminuzione di benessere del Tu.

Tale "sensibilità", se mi si consente il termine, non dovrebbe, a sua volta, essere confusa con ciò che molti definiscono *empathic accuracy*, ossia quell'abilità di individuare emozioni o sentimenti esperiti dall'altro per poi riprodurli/sentirli come propri che l'odierno clima culturale definisce, talvolta, con l'etichetta di "simpatia" o contagio emotivo e, talaltre, con empatia (qualsiasi definizione essa abbia).

Per capire, effettivamente, ciò che lo psicologo statunitense intenda nella delucidazione del rapporto empatia/altruismo e, di contro, "egoismo", diviene necessario porre l'accento sulle definizioni dei fenomeni di cui si occupa, poiché un'interpretazione sbagliata dell'uno o dell'altro o una semplice presupposizione di una sfumatura anziché un'altra porterebbe, certamente, ad una incomprensione di fondo, stante la complessità stessa dell'argomento in oggetto. Direbbe Wittgenstein che la filosofia (ridotta e depauperata dell'essenza stessa del suo metodo d'indagine, che le conferisce lo *status* di *Scienza* e di *Ricerca del Sapere* che, legittimamente, le spetta – aggiungerebbe la scrivente), essendo *chiarificazione linguistica*, non assolverebbe al suo compito se non spiegasse in modo semplice e comprensibile ciò di cui si occupa. Stante ciò, si consideri che per "empatia" Batson intende

*One particular set of congruent vicarious emotions, those that are more other-focused than self-focused, including feelings of sympathy, compassion, tenderness, and the like...empathy is distinct from feelings of personal distress*⁸⁸

mentre

*Altruism is a motivational state with the ultimate goal of increasing another's welfare. Egoism is a motivational state with the ultimate goal of increasing one's own welfare*⁸⁹.

⁸⁸ DANIEL BATSON, *The altruism question: toward a social psychological answer*, Lawrence Erlbaum, Hillsdale, 1991, p. 86.

Come volevasi dimostrare, lo stesso Batson riconosce l'esigenza di specificare tre "espressioni chiave" che, consciamente, utilizza nel definire l'altruismo e l'egoismo: lo "stato motivazionale", ad esempio, è una fra queste.

Dovendo denotare, infatti, le ragioni psicologiche per cui un individuo dirige il proprio fare verso l'altro, lo scopo ultimo della sua azione dovrà rispettare alcune caratteristiche fondanti per esser definito, appunto, *goal-directed*, ossia la volontà precisa e puntuale dell'agente di ottenere quel fine e non altri. Tali caratteristiche, come scrive lo stesso Batson, sono:

- a) *The organism desires some change in his or her experienced world (this is what is meant by a goal);*
- b) *A force of some magnitude exists, drawing the organism toward the goal;*
- c) *If a barrier prevents direct access to the goal, alternative routes will be sought;*
- d) *The force disappears when the goal is reached.*

[...] *To set and to seek goals require high-level perceptual and cognitive processes generally associated with a developed neocortex of the sort found in higher mammals, especially humans*⁹⁰.

La seconda espressione passibile di un'eventuale incomprensione è *with the ultimate goal* che, alla luce della distinzione della *ratio* a fondamento dell'aiuto offerto (l'altro inteso come fine o come mezzo), Batson è attento a sottolineare che le sue posizioni vertono per l'imperativo categorico kantiano (volendo riprendere il discorso condotto poco più su) in cui l'altro diviene il significato ultimo, il senso in sé del prodigarsi verso il Tu. Infine, coerentemente con quanto appena sostenuto, l'ultima delle tre espressioni è rappresentata da *of increasing another's welfare* in antitesi a *of increasing one's own welfare*, su cui ci si soffermerà poco più avanti.

⁸⁹ DANIEL BATSON, LAURA L. SHAW, *Evidence for altruism: toward a pluralism of Prosocial Motives*, in "Psychological Inquiry", Vol. 2, n. 2, 1991, p. 108.

⁹⁰ *Ivi*. Essendo tale passaggio fondamentale per un'adeguata comprensione della teoria di Batson, si è reputato necessario aggiungere *hic* la traduzione concettuale (ad opera della scrivente) dall'inglese in italiano del passaggio indicato.

Si legge:

- a) *L'organismo desidera qualche cambiamento nella percezione che ha del (n.d.r.) mondo esperienziale (questo è ciò che si intende per obiettivo);*
- b) *Esiste una forza di una certa entità, che attira l'organismo verso lo scopo prefissatosi (n.d.r.);*
- c) *Se una barriera impedisce l'accesso diretto all'obiettivo, saranno cercati percorsi alternativi; e*
- d) *La forza scompare quando lo scopo viene raggiunto*

[...] *Per impostare e per ricercare obiettivi (congrui alle proprie motivazioni, n. d. r.) si richiedono processi percettivi e cognitivi di alto livello, generalmente associati con lo sviluppo della neocorteccia, tipica (n.d.r.) dei mammiferi superiori, specialmente dell'uomo.*

Tornando al nocciolo del problema, secondo lo psicologo statunitense i presupposti essenziali di qualsiasi teoria sulla relazione fra altruismo ed empatia debbono essere costituiti dalla percezione dell'altro in stato di bisogno e dal valore intrinseco del benessere altrui: capisaldi, se si vuole, dell'altruismo *tout court*.

La *perception of another as in need*⁹¹ si baserebbe sulla percezione di una discrepanza naturale fra lo stato attuale di un'ipotetica vittima e ciò che per essa sarebbe desiderabile in una dimensione emotiva in cui fossero assenti ansia, panico, preoccupazioni, pericoli; ciò, specifica Batson, non significa che tale percezione (che potremmo, in questo contesto, tradurre con “constatazione”) possa essere declinata quale sentimento empatico (riprodurre in sé lo stato emotivo dell'altro) dell'osservatore nei confronti dell'ipotetica vittima, ma, semplicemente, che il primo sia “naturalmente” in grado di rendersi conto di uno stato di necessità in cui versa l'altro. In tal senso, la traduzione del termine “percezione” con “constatazione” non è sbagliata, poiché Batson non ricerca una *substantia*, ma “semplicemente” uno *status quo* delle capacità umane tanto che sia la percezione dello stato di bisogno dell'altro sia la discrepanza fra il benessere reale o fittizio dell'ipotetica vittima sembrerebbero essere, simultaneamente, i componenti essenziali delle premesse della teoria dell'empatia-altruismo da lui proposta.

Altro presupposto fondamentale è il concetto denotato dall'espressione *intrinsic valuing of the other's welfare*, definito anche come *intrinsic or terminal valuing*, in cui l'altro è considerato dall'agente che si prodiga per il Tu come “persona” avente un fine in sé e non per ciò che quest'ultima potrebbe dargli in cambio ad aiuto avvenuto come fosse un *do ut des* di stampo (falsamente) altruistico.

Lo spartiacque tra un “altruismo genuino” ed un “mero utilitarismo” resta, come già accennato, lo scopo ultimo del comportamento che conduce, inevitabilmente, Batson *et alii* a tre osservazioni generali:

- a) *Helping, as a behavior, can be either egoistically or altruistically motivated; it is the end-state goal, not the behavior, that distinguishes an act as altruistic;*
- b) *Motivation for helping may be a mixture of altruism and egoism; it need not be solely or even primarily altruistic to have an altruistic component;*

⁹¹ DANIEL BATSON, *The empathy-Altruism Hypothesis*, Oxford Handbook Online, Oxford, 2014, p. 4.

- c) *Increasing the other's welfare is both necessary and sufficient to attain an altruistic end-state goal*⁹².

Affinché si sia in grado di discernere un comportamento “genuinamente altruistico” da uno “falsamente disinteressato”, lo psicologo statunitense sostiene che ci si dovrebbe far guidare da quattro principi generali che, se correttamente applicati, sarebbero in grado di individuare lo scopo ultimo reale e non fittizio dell’agente disposto al sacrificio di sé per gli altri o per l’altro. In tal senso, allora, sarebbe buona norma non lasciarsi mai convincere dai racconti di chi aiuta, poiché i *self-reports* non corrispondono quasi mai alle motivazioni reali di chi interviene per il Tu; bisognerebbe inferire il fine dell’azione non da ciò che si è detto o si è ascoltato, ma dall’azione stessa.

Laddove vi fossero situazioni ambigue in cui i fini ultimi del comportamento siano non chiari, allora, quasi certamente, lo scopo ultimo dell’azione non sarà fra quelli manifesti e, infine, a circostanze mutate, se si dovesse mantenere lo stesso *iter* procedurale per il comportamento altruistico in favore dell’ipotetica vittima, allora, probabilmente, la motivazione ad agire dell’osservatore potrebbe non essere genuinamente tale.

Tenendo, quindi, in considerazione gli aspetti generali sopra delineati dell’altruismo e dell’empatia, Batson, come detto, è giunto alla constatazione che, dopo più di trenta esperimenti condotti con studenti e civili sulla contrapposizione fra le teorie “classiche” sulla naturale propensione egoistica della natura umana e la sua proposta teorica (l’uomo è più altruista di quanto si pensi)⁹³, l’empatia non sarebbe la spinta motivazionale pro-sociale al comportamento morale, ma, contrariamente a quanto sostenuto da Hoffman, la naturale propensione verso l’altro (altruismo) “genererebbe” la “preoccupazione empatica” da cui scaturirà il comportamento morale.

Sono invertiti, cioè, i termini dell’equazione.

Da come si evince dai due modelli sopra descritti, un elemento comune ad entrambi è la capacità da parte dell’agente di “entrare nella mente altrui”; quindi, una capacità di *lettura*

⁹² DANIEL BATSON ET ALII, *Is empathic emotion a source of altruistic motivation?*, in “Journal of Personality and Social Psychology”, Vol. 40, n. 2, 1981, p. 291.

⁹³ Fra i molteplici lavori di Batson, per avere un’idea della metodologia esperienziale adoperata, si consigliano: C. D. BATSON, *These things called empathy: Eight related but distinct phenomena*, in JEAN DECETY E WILLIAM J. ICKES (A CURA DI), *The social neuroscience of empathy*, MIT Press, Massachusetts, 2009; DANIEL BATSON, *Two forms of perspective taking: Imagining how another feels and imagining how you would feel*, in KEITH D. MARKMAN, WILLIAM M. KLEIN E JULIE A. SUHR (A CURA DI), *Handbook of imagination and mental simulation*, Psychology Press, New York, 2009; DANIEL BATSON, *The Naked Emperor: Seeking a more plausible genetic basis for psychological altruism*, in “Economics and Philosophy”, 26, 2010, pp. 49–164; DANIEL BATSON, *Altruism in humans*, Oxford University Press, New York, 2011.

della mente in grado di oltrepassare le barriere verbali e non verbali tra soggetti per penetrare, direttamente, nei pensieri o nel mondo “virtuale” dell’altro.

Un modello alternativo ad entrambi è stato presentato dal filosofo statunitense Shaun Nichols in cui, contrariamente a Hoffman ed a Batson, l’altruismo non è presentato come (volente o nolente) una “facoltà superiore”, ma come una risposta, direi “istintiva” o “più semplice”, di abilità cognitive in grado di percepire, immediatamente, lo stato di bisogno dell’altro. In questo senso chi scrive ha osato spiegare il *Concern Mechanism* di Nichols con la parola, virgolettata, “istinto”, poiché lo stesso autore scrive che il meccanismo neuro-biologico in grado di dotare l’uomo della capacità di capire la mancanza di benessere dell’altro è in grado di “funzionare” anche al di sotto dei canoni presupposti del livello di coscienza, responsabili (reinterpretando Hoffman e Batson) di tutti i discorsi successivi all’abilità del *mindreading*:

altruismo → empatia oppure empatia → altruismo?

È ovvio che, come detto, in questa sede non si è potuto dar conto delle molteplici sfumature che rendono il dibattito tanto complesso quanto affascinante ed è altrettanto ovvio che i modelli di cui ci siamo serviti per spiegare il problema non sono in assoluto gli unici o gli “immutabili”, ma sono stati presi come spunto di riflessione per tentare di evidenziare la complessità tanto del “fenomeno empatia” quanto i molteplici ambiti pratici in cui esso viene “declinato”.

È interessante, a proposito di sfera emotiva intesa nella sua accezione più ampia e variegata, leggere nell’introduzione al testo di Antonio Damasio non soltanto ciò che un neuroscienziato intenda per *coscienza*, ossia

[...] *la chiave di una vita osservata nel bene o nel male, la nostra prima autorizzazione a conoscere tutto della fame, della sete, del sesso, delle lacrime, delle risa, dei calci, dei pugni, del flusso di immagini che chiamiamo pensiero, dei sentimenti, delle parole, delle storie, delle credenze, della musica e della poesia, della felicità e dell’estasi. Al livello più semplice e fondamentale, la coscienza ci fa riconoscere la spinta irresistibile a rimanere in vita e a prendersi cura di se stessi. Al livello più complesso ed elaborato, la coscienza aiuta a sviluppare un interesse per altri sé e a perfezionare l’arte della vita* [...] ⁹⁴,

quanto la definizione che egli fornisce dell’*emozione*,

⁹⁴ ANTONIO DAMASIO, *Emozione e Coscienza*, Biblioteca Scientifica Adelphi, Milano, 2007, pp. 17-18.

collezione di risposte [corporali n.d.r.], in gran parte osservabili pubblicamente [...] ⁹⁵,

poiché ciò apre la strada a spunti di riflessione che non sono, poi, così lontani dalle nostre vite.

In palio, vale a dire, c'è il modo stesso in cui l'uomo "vede", considera se stesso e da tale prospettiva non si può prescindere dal dibattito sulla naturalizzazione cognitiva.

L'interesse per Damasio (preso a modello di un approccio distinto da quello morale al problema empatia/altruismo), ad esempio, si trova nel fatto che il neuroscienziato, attenendosi a casi di pazienti affetti da gravi o gravissime lesioni neuronali, giunge a sostenere che esiste una differenza fra "emozione provata" e "sentimento provato". Differenza non solo di natura semantica, ma neurofisiologica; il "sentimento provato", definito come

l'esperienza mentale, privata di un'emozione [...] ⁹⁶

baserebbe la sua "ragion d'essere" nella "base neurale della consapevolezza" (qualsiasi "oggetto" essa denoti ed ammesso che possa avere una base neurale).

In che modo, allora, da una prospettiva neuroetica, si potrebbe "spiegare" il rapporto empatia/altruismo? È, forse, difficile pensare che l'intuizione "chiara e distinta" di essere un Io pensante, *in primis*, possa essere "sostituita" da un *Io cerebrale*, considerando, *in secundis*, che l'intento delle odierne scienze cognitive è quello di avviare una scienza globale dell'uomo, spiegando, neurobiologicamente, anche il rapporto che la mente o le stesse abilità di *mindreading* possano intrattenere con altre menti. Partendo da tale assunto, non si deve dimenticare che l'odierno naturalismo cognitivo si basa su quattro principi generali – declinati in maniera diversa a seconda dell'interpretazione dell'uno o dell'altro fenomeno considerato peculiare per spiegare l'ontologia del mentale – attraverso i quali si confuta, *de facto*, non soltanto il dualismo di stampo cartesiano, ma anche una particolare spiegazione dell'empatia cognitiva. Quest'ultima, infatti, utilizzando la mente per riprodurre in sé gli stati emotivi altrui, si troverebbe, in parte, denaturalizzata e ricondotta allo stesso livello dell'empatia di base o "neuroscientificamente spiegabile".

A tal proposito, Sandro Nannini sostiene:

⁹⁵ *IDEM*, p. 59.

⁹⁶ *IVI*.

1. La mente (così come la coscienza, lo spirito o la soggettività) fa parte del mondo reale;
2. La natura costituisce l'intero mondo reale [...].
3. La natura può essere conosciuta solo dalle scienze empiriche [...].
4. L'universo fisico è un sistema 'chiuso' [...] ⁹⁷.

La realtà, compresa quella mentale, quindi, si “tradurrebbe” in eventi prettamente fisici: gli stati psicologici (desideri, intenzioni, volontà...) sarebbero spiegabili ed identificabili in interazioni neurobiologiche, possibili solo grazie alla complessità cerebrale. Ciò, come fa notare Tim Crane, implica un atto di “riduzione” che, stando alla semantica del termine, non significa “eliminazione”, ma identificazione delle “entità di un dominio con una sottoclasse delle entità di un altro” ⁹⁸.

Una corrispondenza biunivoca tra uno stato mentale ed il suo substrato cerebrale è stata avanzata, agli inizi degli anni '50 del secolo scorso, dalla *Teoria dell'Identità di tipo* o *Tipo-Tipo* (si ricordino, ad esempio, Herbert Feigl, Ullin Place, Jack Smart), secondo cui lo stesso rapporto che intercorre fra la mente ed il cervello esisterebbe fra la temperatura e l'energia cinetica: la medesima natura esplicantesi in modi diversi. Tale teoria, come è facile intuire, si è trovata di fronte ad insormontabili difficoltà, sia di ordine neuroscientifico (come riuscire ad identificare quale sinapsi o quali processi cerebrali diano origine ad uno specifico stato mentale?) sia di ordine concettuale. Riprendendo, ad esempio, l'argomento della *realizzabilità multipla* ⁹⁹ di Hilary Putnam, si capisce bene come l'assunzione vincolante tra eventi mentali e stati cerebrali porterebbe *ex absurdo* alla possibilità che uomini e animali abbiano, per un dato stato mentale, un *identico* stato cerebrale.

Al di là del dibattito sui limiti o sulle possibilità offerte dalla Teoria dei tipi o della sua contrapposta teoria della *Simulation Theory* (delle quali si parlerà nel prossimo capitolo) è bene notare che fra ali moderate ed ali estremiste del pensiero sulle possibili spiegazioni di che cosa sia il mentale, si dovrebbe essere consci che viviamo in un periodo di passaggio, di “traghetamento” da un'epoca ad un'altra in cui ciò che verrà non significa “esser migliore del precedente”, ma semplicemente differente in cui, però, un'attenta valutazione del *continuum* delle idee ed il loro mutamento non dovrà essere abbandonata.

⁹⁷ Cfr. SANDRO NANNINI, *Naturalismo Cognitivo. Per una teoria materialistica della mente*, Quodlibet Studio, Macerata, 2007.

⁹⁸ Cfr. TIM CRANE, *Elements of Mind*, Oxford University Press, Oxford, 2001.

⁹⁹ Cfr. HILARY PUTNAM, *Mind, Language and Reality*, Philosophical Paper, Vol. 2, Cambridge University Press, Cambridge (Mass.), traduzione italiana di ROBERTO CORDESCI, *Mente, Linguaggio e Realtà*, Adelphi, Milano, 1987.

Capitolo III

L'aspetto cognitivo dell'empatia

3.1 Come nasce il problema del *mentalizing*

Si è spesso accennato nelle pagine precedenti che l'empatia è, oramai, associata allo strumento epistemologico dell'intersoggettività, ossia alla capacità di entrare “nella” mente dell'altro per comprenderne gli stati mentali o le emozioni: una forma conoscitiva, cioè, non mediata da forme logico-razionali, ma un “sentire intenzionale”, che, usando un'espressione di Scheler, rivela l'esistenza di beni e valori morali *nella gerarchia oggettiva tra valori*¹⁰⁰ dell'altro. In tale ottica, allora, l'azione morale di dirigersi consapevolmente verso il Tu, poggiandosi sulla naturale “spiritualità” di ogni persona, si traduce in *immedesimazione l'uno con l'altro*; il percepire in sé il Sé altrui, un *cosentire* in grado di “aprire” l'Io all'intersoggettività ed all' “interiorità” del Tu.

Nello schema tipo “osservatore-vittima”, ipotizzato da Hoffman, si parla di *empatia cognitiva* quando la comprensione dello stato emotivo della vittima risulta mediata da tre fattori distinti che anticipano quello sforzo mentale (cognitivo, appunto) di comprensione altrui: il tempo per l'elaborazione del messaggio; lo sforzo intellettuale per comprenderlo; la distanza psichica fra l'osservatore e la vittima. Questi renderebbero, perciò, tale tipo di empatia o tale sua caratteristica in grado di fornire “conoscenza” dello stato emotivo della vittima: uno sforzo concettuale che permetterebbe di “entrare nei panni altrui” o, detta in altro modo, di capirne la mente.

¹⁰⁰ Cfr. MAX SCHELER, *Il formalismo nell'etica e l'etica materiale dei valori*, Bompiani, Torino, 2013.

Il *mentalizing*, perciò, rappresenta l'aspetto più controverso del fenomeno empatico, poiché se è vero che esiste una certa corrente di pensiero che propone una correlazione significativa fra *Leib e Korper* è altrettanto vero che lo stesso concetto di mente è assai dibattuto¹⁰¹. L'aspetto centrale del problema, però, resta capire come sia avvenuta questa traslitterazione semantica dal campo emotivo a quello epistemologico e la chiave di lettura deve essere ricercata in colui il quale, volente o nolente, ha influenzato e continua ad influenzare il paradigma concettuale del dualismo mentale dal XVII secolo ad oggi: Renè Descartes.

Identificare nel *cogito* la *summa* dell'intera tradizione del pensiero filosofico, prima, e psicologico, poi, per cercare di capire *che cosa* rappresenti e *che cosa sia* l'uomo è una sintesi teoretica ed ontologica più che condivisibile: la domanda più antica del mondo, che dominava i pensieri, le riflessioni e le speculazioni dei filosofi greci sulla natura dell'uomo trova, infatti, in Cartesio una sua risposta ancora oggi al centro del dibattito nella Filosofia della mente. Nel 1641, Descartes, nelle *Meditazioni Metafisiche*¹⁰², dopo aver sottoposto a dubbio iperbolico tutte le sue certezze, conviene, in ultimo, che ciò di cui può effettivamente non dubitare è l'esistenza di se stesso: parlare, agire, mangiare, interrogarsi... ossia, capire che è, essenzialmente, una *cosa* che *pensa*, che *cogita*, e che, contemporaneamente, esperisce, percepisce ed opera. Un pensare ed un agire: una materia *pensante* in diretta connessione ad una che "si estende" nello spazio.

L'interazione fra il pensare ed il pensarsi quale essere senziente (la *res cogitans*) ed il convenire di essere una "sostanza estesa" (la *res extensa*) hanno dato il via, nella storia del pensiero, all'idea che l'uomo sia pensiero e materia o, nell'accezione contemporanea, un dualismo di mente e di corpo¹⁰³.

Quale fosse, all'epoca di Cartesio, l'inter-relazione fra le due *res* è dato a sapersi grazie alle stesse parole del filosofo francese, il quale sostiene che, grazie alla ghiandola pineale (oggi riconosciuta nell'epifisi), le due sostanze, aventi caratteristiche e peculiarità proprie, riescono ad interagire: una interazione che, se da un lato pone di fronte ad una evidenza empirica ("ho un corpo"), dall'altro questa non può percepire, con altrettanta immediatezza, l'attributo essenziale della natura umana: il pensiero, il *cogito*.

Cartesio, dunque, partendo dalla volontà di giungere ad una verità certa ed indubitabile, conclude, paradossalmente, la sua riflessione *duplicando* il problema originario; se l'intento fosse stato quello di trovare una verità tale da spiegare il mondo nella sua totalità finisce con il

¹⁰¹ FEDERICA MADONNA, *Mentalizing: the cognitive aspect of empathy* in "Journal of Advance Research in Social Science & Humanities", Vol. 5, Issue 2, 2019, pp. 69-78. Articolo reperibile online al seguente indirizzo: <https://jiaats.com/ojs31/index.php/ssh>

¹⁰² Cfr. RENATO CARTESIO, *Meditazioni Metafisiche*, Bompiani, Firenze, 2001.

¹⁰³ Cfr. ALFREDO PATERNOSTER, *Introduzione alla filosofia della mente*, Laterza, Roma-Bari, 2014.

sostenere che esistono *due mondi*, l'uno distinto dall'altro, l'uno *parallelo* all'altro, ognuno con proprie regole – o *attributi* – e con proprie specificità: il mondo del *cogito*, rispondente alla sostanza inestesa, non spaziale; quello della materia, la *res extensa*, consono agli attributi dello spazio e del tempo.

Una duplicazione del problema iniziale in cui l'uomo, sintesi simultanea dei due mondi, paradossalmente, riesce a pensare e ad agire seguendo regole o sistemi di regole completamente differenti ed *indifferenti* l'uno dall'altro: determinismo e rigido meccanicismo – principi della nuova fisica postgalileiana – ordinano la materia, i corpi degli animali e dell'uomo stesso; coscienza, creatività e libertà – peculiarità proprie della *res cogitans* – si esplicano nel mondo del pensiero di cui l'uomo, e solamente l'uomo, è degno ed unico rappresentante.

Fondamentalmente, sul problema di *come* funzioni il *cogito*, Cartesio non ci fornisce una spiegazione sistematica e dettagliata, tanto che, proprio su questo ed altre mancanze, secondo i suoi successori, il filosofo francese peccerebbe di imprecisione. In realtà, il *focus* principale dell'intera argomentazione cartesiana, utile al presente scritto, è incentrato sul fatto che Descartes, da una parte, tenta – indipendentemente dai giudizi valoriali – di fornire una spiegazione al perché esista in noi una “facoltà superiore” – la coscienza, il pensiero, la mente –; dall'altra, però, compie una *fallacia* epistemica nel suo ragionamento. Infatti, da una premessa di tipo ontologico (esistono due diverse sostanze) giunge ad una *conclusio* epistemologica (possiedo un corpo; mentre di questo posso dubitare, non potrei mai tentennare nell'asserire di possedere una mente); tanto che sembrerebbe voglia superare la sua disavvertenza, chiamando in causa un elemento prettamente *estes*, la ghiandola pineale, per uscire dall'*impasse* dell'incomunicabilità dei due piani sostanziali.

Un tentativo di chiarire *come* ed *in che modo* l'uomo possa interagire o relazionarsi agli altri che alcuni pensatori, da una parte, potrebbero interpretare con molte forzature quale debole, debolissimo sforzo di Cartesio di non ricorrere a Dio – come Malebranche, suo contemporaneo – per spiegare la causalità degli effetti del mentale sul fisico: dal pensiero, cioè, al corpo; dall'altra, per gli stessi sarebbe difficile, alla fine del '600, annoverare Descartes fra coloro che si sarebbero avvicinati, timidamente, al programma di *naturalizzazione cognitiva* (ossia, come già detto, quel filone di studi che vuole spiegare gli aspetti cognitivi umani su base prettamente biologica, sostenendo e tentando di mostrare che l'uomo sia per sua costituzione fisica portato all'agire cognitivo), addirittura interpretando il suo meccanicismo quale prodromo di ciò che sarà, poi, l'avvento delle neuroscienze.

Nonostante la filosofia moderna rispondesse a domande e problemi tipici di quell'epoca, gli interrogativi sollevati da Cartesio, però, dopo quattrocento anni, sembrerebbero provenire proprio da quell'intento, anche se, ovviamente, ciò rappresenta una interpretazione *post hoc* con paradigmi e conoscenze decisamente diversi da quelli posseduti dal filosofo francese; sebbene le difformità evidenti, due dei suoi interrogativi rappresentano, tutt'oggi, il perno attorno a cui ruotano le domande sul funzionamento della mente, ossia:

1. *come e perché i fenomeni mentali sono differenti dagli altri? Si deve, cioè, presupporre o inferire una loro specificità sostanziale, ontologica, rispetto a quelli fisici o materiali?*
2. *in che modo si potrebbe dar prova della loro natura biochimica di contro all'ipotesi di una natura prettamente metafisica?*¹⁰⁴

Lo scontro-incontro, hegelianamente parlando, delle domande sopra riportate sono state il filo conduttore che, fra riduzionismo ed antiriduzionismo, fra la metafisica classica ed il progredire della scienza odierna, hanno condotto (e conducono), inevitabilmente, ad una forte discrasia fra l'immagine, forse un po' "romantica", che abbiamo di noi stessi e quella "più vera", più pertinente all'immagine *naturale*, fornitaci dalla scienza *tout court*. È proprio a partire da questa, infatti, che negli ultimi sessant'anni si è ricorso, sempre più spesso, ai metodi, alle scoperte o alle stesse ipotesi di lavoro di cui la scienza è costituita per tentare di rispondere al dualismo sollevato, in modo preponderante, da Cartesio: cambiano i contesti, ma non i problemi. Infatti, è per molti fonte di dubbio e perplessità capire come la materia di cui sia costituito il nostro sistema nervoso possa, addirittura, *pensare* o essere *consapevole*¹⁰⁵; dal semplice potenziamento/trattamento delle facoltà cognitive umane – quale, ad esempio, la farmacologia – si è giunti a parlare, addirittura, di *lettura della mente* o *mindreading*: la *res cogitans*, cioè, spiegata attraverso la *res extensa* grazie, proprio, alle neuroscienze.

Se è vero che sarebbe per noi difficile immaginare come potrebbe cambiare, se mai potesse, la visione di Cartesio, oggi, se apprendesse le lezioni delle neuroscienze cognitive¹⁰⁶ è pur vero che dal punto di vista epistemologico, in generale, approcciarsi al "problema mente" significa, innanzitutto, porsi l'interrogativo primario di capire:

¹⁰⁴ MICHELE DI FRANCESCO, *Introduzione alla filosofia della mente*, Carocci Editore, Roma, 2002, p. 20.

¹⁰⁵ Cfr. ASTRO CALISI, *Oltre gli orizzonti del conosciuto. La sfida della mente alla scienza del XXI secolo*, Editrice Uni Service, Trento, 2011.

¹⁰⁶ Cfr. FEDERICA MADONNA, *Il problema aperto della mente in relazione alle neuroscienze. Che cosa ne direbbe, oggi, Cartesio?*, Aracne Editrice, Roma, 2018.

1. che cosa sia questa capacità, forse non prettamente umana, che consenta all'uomo di conoscere se stesso¹⁰⁷, *in primis*, e di relazionarsi agli altri, *in secundis*¹⁰⁸;
2. presupporre di capire le intenzioni, le volizioni ed i pensieri di altre menti;
3. non trascurare, in ultimo, la capacità di ogni Sé di conoscersi attraverso quel continuo processo maieutico di interrogazione incessante fra sé e sé con gli altri.

Ciò, metafisicamente, porterebbe all'elaborazione dei contenuti concettuali delle altre menti, interrogandosi non più sugli "stati mentali" altrui, ma sul modo in cui gli altri capiscano gli altri da Sé.

In questa direzione, la tradizione continentale ha perseverato (e, forse, lo fa tuttora) a comprendere il "problema mente" come iscritto in un circolo ermeneutico, legato al problema del senso e del significato delle espressioni o dei comportamenti del singolo; sulla scia del nuovo slancio dato da Heidegger, Husserl, Gadamer e molti altri. Infatti, quest'ultima adottava una particolare pratica interpretativa che, occupandosi di una certa "matrice testuale" (si pensi, ad esempio a Collingwood o allo stesso Gadamer) veniva considerata l'unica in grado di fornire gli strumenti necessari alla comprensione di tutte le scienze. Tale tendenza, in parte, cambiò nei primi anni del Novecento, poiché, legandosi soprattutto alla tradizione della *simpatia*, prese ad oggetto l'arte in quanto tale: artefice di tale influenza e, se si vuole, di questo "cambio di rotta", fu (ad avviso della scrivente) Friedrich Schleiermacher, il quale sostenne che nella "scienza artistica" l'autore non seguiva le regole logico-razionali, ma "ricreava" se stesso. Volendo interpretare il pensiero dell'autore, l'artista trasponeva su tela parte di sé oggettivandosi, affinché il fruitore dell'opera potesse "entrare" in contatto con l'intimo Io dell'artista stesso, per comprenderne emozioni e sentimenti: un messaggio "universale" in grado, cioè, di travalicare le barriere della conoscenza ordinaria per essere, in tal modo, compreso "immediatamente".

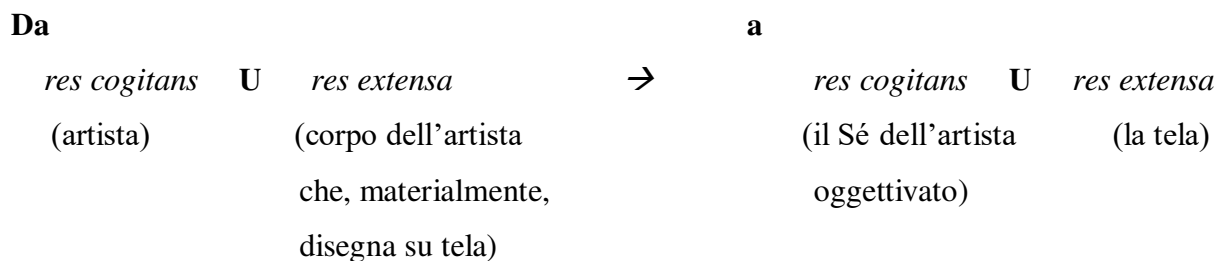
Fu da questo momento in poi che concetti come "empatia, risentire in sé, mettersi nei panni dell'altro..." e simili tornarono, prepotentemente, all'attenzione speculativa, i quali riproponevano, continuativamente, il dualismo cartesiano. Si pensi, ad esempio, allo stesso Schleiermacher. Nell'operazione di "ricreazione di Sé" la *res cogitans* dell'artista si sarebbe trasposta su tela, rimanendo (in parte) *cogito* e (in parte) materia; per il tramite di due *res* estese (il corpo dell'artista ed il quadro) si sarebbe assistito, vale a dire, ad un continuo e progressivo passaggio fra le due sostanze tale da sopraelevare, in ultimo, lo *status* di una

¹⁰⁷ Cfr. GIOVANNI REALE, *Socrate. Alla scoperta della sapienza umana*, Rizzoli, Milano, 2001.

¹⁰⁸ Cfr. TIM CRANE, *Fenomeni mentali. Un'introduzione alla filosofia della mente*, Raffaello Cortina Editore, Milano, 2003.

“semplice tela” in quello di “opera d’arte”, ossia si sarebbe passati ad una *res extensa*, portatrice di un valore (morale?) universale.

Seguendo la schematizzazione presente:



Risulta evidente che, dal dualismo sostanziale dell’artista, lo stesso trasporta su tela la sua dualità, “immettendo” su quest’ultima parte di sé (quindi, *cogito*), rendendola “viva” attraverso un’unione inscindibile con il medesimo supporto fisico che ha consentito l’*impressione* di Sé al di fuori di sé: la tela stessa.

Non è un caso, infatti, che il medesimo fruitore per giungere *in medias res* alla fruibilità di tale dualismo debba abbandonare i canoni consuetudinari dell’apprendimento per giungere ad una “immedesimazione” con il Sé cogitante oggettivato dell’artista per il tramite di un’altra strada.

La *simpatia* settecentesca forniva, così, lo strumento essenziale per poter spiegare questo processo, divenendo, allora, ciò che diverrà il “mettersi nei panni dell’altro” o empatia.

Un’interpretazione fra mondi diversi e paralleli (*res cogitans* e *res extensa*) che poneva, urgentemente, all’attenzione il problema di come sarebbe stato possibile, allora, fare dell’empatia un mezzo conoscitivo che esulasse dal campo artistico per proporsi in quello della vita vissuta, delle relazioni sociali da cui la vita dell’individuo è costituita: un modo in cui fosse possibile, cioè, esser certi dell’esistenza delle altre menti che diverrà un problema di marginale importanza per i filosofi della tradizione ermeneutica; così come lo è l’attuale uso che si fa del concetto di *empathy*. Centrale tornerà ad essere l’interpretazione del rapporto fra il testo ed il suo scritto, concentrandosi, soprattutto, sulla struttura del significato ed il significato stesso, che porterà, ad esempio, al concetto di “significato onnicomprensivo del fenomeno” e simili.

L’esempio celebre nella filosofia del linguaggio del caso della traduzione del termine *gavagai* in “coniglio” costituisce, anche sotto il punto di vista del percorso concettuale dell’empatia (ad avviso della scrivente) un ulteriore punto di svolta per lo stesso. Quando, infatti, il linguista indica all’indigeno un coniglio, pronunciando l’interrogazione: “Gavagai?” è ben

consapevole che per avviare una comprensione di ciò che costituirà la risposta del suo interlocutore dovrà, solo se sarà in grado, riuscire ad “andare oltre”, “superare” (se mi si consente l’espressione) ciò che Quine stesso chiamò “il mito del museo”; l’ipotesi (da lui rifiutata) secondo cui le parole non siano altro che “etichette” affianco alle parole “esposte” su uno scaffale di un museo che le definiscono.

Quello che importa sottolineare a riguardo non è la ripercussione semantica nell’opera di Quine dell’esperimento in questione, ma il fatto che Quine stesso sia consapevole che per “capire l’altro” occorra un mezzo ulteriore rispetto alle sterili “leggi di copertura” di hempeliana memoria; tale strumento diviene, appunto, l’empatia, la quale, però, è connaturata sostanzialmente alle teorie comportamentistiche del parlante. Quine sostiene che:

Il linguista rileva il proferimento di ‘Gavagai’ da parte del nativo nei casi in cui egli stesso, nella posizione del nativo, potrebbe aver detto ‘Coniglio’. Così prova a dire ‘Gavagai’ nelle occasioni in cui egli sarebbe spinto a dire ‘Coniglio’, e attende l’approvazione dei nativi. Incoraggiato, adotta provvisoriamente ‘Coniglio’ come traduzione. L’empatia domina l’apprendimento del linguaggio, sia nel caso del bambino sia nel caso del linguista sul campo. Nel caso del bambino si tratta dell’empatia da parte del genitore. Questo valuta l’appropriatezza dell’enunciato osservativo del bambino considerando il suo orientamento e come la scena apparirebbe dal suo punto di vista. Nell’altro caso si tratta dell’empatia da parte del linguista che opera sul campo allorché avanza la sua prima congettura su ‘Gavagai’ sulla base del proferimento e dell’orientamento del nativo, e ancora quando chiede ‘Gavagai’, in una successiva occasione che gli sembra favorevole, per ottenere l’assenso del nativo. Tutti noi abbiamo una straordinaria capacità di empatizzare la situazione percettiva di un’altra persona, ancorché ignari dei meccanismi fisiologici o ottici della sua percezione¹⁰⁹.

Oggetto di un’attenzione speculativa specifica che, contrariamente a Quine, non lasciasse “semplicemente” intuire una “compagna di viaggio” nella sua opera più influente, l’empatia è stata ripresa da Hilary Putnam e Robert Nozick tra gli anni Settanta ed Ottanta del secolo scorso.

¹⁰⁹ WILLARD VAN ORMAN QUINE, *Three Indeterminacies*, IN R. BARRETT E R. GIBSON (A CURA DI), “Perspective on Quine”, Basil Blackwell, Oxford, 1990, pp. 1-16, cit. p. 4

Nel panorama analitico contemporaneo, infatti, i due filosofi statunitensi hanno sostenuto la comprensione empatica, donandole plausibilità e semplicità, il primo; lo statuto epistemologico, il secondo. In quest'ultimo, la spiegazione del *Verstehen* diviene:

*l'altro agirebbe come agirei io e che io agirei come penso (sulla base di una proiezione immaginativa) che agirei*¹¹⁰,

tracciando le fila di ciò che, negli anni successivi, sarà definita *lettura della mente*.

La ripresa del concetto di *empathy* non si lega solo ed esclusivamente ai filosofi sopra citati, ma, certamente, essi sono stati fra coloro che hanno contribuito al riaccenderne i riflettori, a maggior ragione che, proprio in quegli stessi anni, le “ricostruzioni razionali” fornite dal modello di Hempel e dal Neopositivismo Logico cominciarono ad “andar strette” per la spiegazione dei fenomeni emotivi e razionali dell'uomo comune. Sempre in Nozick, si legge:

*limitarsi a inserire un'azione in uno schema razionale, mostrandone l'intelligibilità come perseguimento intenzionale di certi scopi e desideri alla luce della situazione in cui l'agente credeva di trovarsi, non costituisce di per sé una spiegazione dell'azione; sarebbe necessario anche aggiungere che la persona è generalmente disposta a perseguire razionalmente i suoi scopi, o, almeno, che lo era nell'occasione data. Con alcune persone un particolare problema esplicativo è semmai quello di spiegare perché si comportano razionalmente quando lo fanno*¹¹¹.

L'intelligibilità di un comportamento, quindi, non deriverebbe da un'assunzione di razionalità così come nelle Scienze naturali, ma dalle conoscenze empiriche che si hanno sull'agente e sulla situazione che sta vivendo, onde evitare di sbagliare nello spiegare il suo agire: ritorna, seppur in termini non congrui all'originale, la situazione tipo dell'interpretazione radicale di quineana memoria negli aspetti centrali dell'intersoggettività e dell'attribuzione intenzionale. Seguendo un certo tipo di interpretazione, si potrebbe sostenere che, progressivamente, nella Filosofia analitica, proprio per la sua *ratio*, si è cercata una spiegazione alternativa all'ontologia dualista, ricorrendo, progressivamente, ad altre forme di spiegazione della sfera emotiva ed agentiva dell'individuo, ricercando quelle stesse “leggi di copertura” a cui si sarebbero attribuite razionalità e “scientificità”. Un esempio potrebbe essere quello dello

¹¹⁰ ROBERT NOZICK, *Philosophical Explanation*, Clarendon Press, Oxford, 1981, p. 638.

¹¹¹ *IDEM*, p. 639.

stesso Quine, il quale sosteneva, in un contesto prettamente linguistico, che non vi è conflitto fra empatia e comportamento manifesto, dando legittimità allo strumento psicologico di comprensione dell'altro per evitare, appunto, di ricorrere a qualsiasi ontologia mentalistica: non a caso, Quine stesso si affida non solo al principio di carità, ma anche alla *folk psychology*, il sapere comune, che diverrà centrale per la spiegazione del comportamento empatico.

La concezione teoretica della *folk psychology* fu teorizzata da Sellars¹¹² negli anni Sessanta con il *Mito di Jones*, attraverso il quale il filosofo statunitense, rifiutando qualsiasi ordine dualistico del mentale, volle tradurre il comportamento manifesto in un sapere empiricamente fondato:

*il frutto di un lungo processo di risposte pubblicamente rinforzate a oggetti pubblici (incluse le esecuzioni verbali) in situazioni pubbliche*¹¹³.

La “storiella fantastica” di Jones poneva l’accento sulla situazione dello stesso Jones, uno scienziato terrestre, il quale, rifiutando l’idea di possedere stati mentali quali credenze e desideri, era convinto, nondimeno, che fosse ugualmente in grado di articolare pensieri su altri individui a cui attribuire quegli stessi stati mentali attraverso l’utilizzo di un *corpus* di conoscenze comunemente condivise definito, appunto, *folk psychology*. Jones, infatti, rinnegando qualsiasi “linguaggio mentalistico”, è talmente certo, secondo Sellars, del rifiuto dell’ontologia dualista che ipotizza che, laddove un marziano dovesse entrare in contatto con un terrestre (quindi, due soggetti per ambiente, cultura, linguaggi, usanze... diversi) ed apprendere la psicologia comune di quell’ambiente in cui vive il suo nuovo amico terrestre, il “nuovo ET” sarebbe in grado, analogamente, di capire gli stati mentali (credenze o desideri) dell’abitante umano.

Il *Mito di Jones* è stato visto da molti come alternativa al cartesianesimo attraverso il quale l’intersoggettività non si sarebbe più dovuta spiegare attraverso i canoni problematici degli stati mentali e dei problemi concettuali cui, necessariamente, incorreva (come, infatti, associare ad un “linguaggio mentalese” esperito in prima persona un’obiettività fruibile in terza persona?), bensì attraverso il ricorso ad un insieme costituito di credenze condivise che avrebbero conferito legittimità al comportamento manifesto.

¹¹² Cfr. WILFRID SELLARS, *Science, Perception and Reality*, Humanities Press, New York, 1963.

¹¹³ *IBIDEM*, p. 62.

Si tratta, per Sellars, di un *escamotage* per evitare quella trappola filosofica secondo cui, eliminando il dualismo, si potrebbe incorrere nell'incomunicabilità dei propri stati emotivi o "mentali"; o perché ritenuti errori categoriali (si ricordi, ad esempio, Gilbert Ryle) o, di contro, perché non categorizzabili. Rimarrebbe esclusa, in linea di principio, qualsiasi relazione intersoggettiva.

Era chiaro l'intento di Sellars di ricondurre la definizione di "stato mentale" alla risultante di un *output* in risposta ad un *input*, eliminando qualsiasi uso linguistico del mentale, ma ciò avrebbe comportato, come detto, l'isolazionismo monadico che già Husserl, parlando di intersoggettività, aveva scongiurato o aveva tentato di scongiurare tramite l'ausilio dell'empatia con tutte le critiche cui incorse.

L'ascesa del funzionalismo, in realtà, "risolse" il *gap*, donando alla *folk psychology* la struttura di una vera e propria teoria condivisa che, proprio come una teoria scientifica, è dotata di costrutti, concetti, premesse e principi: il *central action principle of folk psychology* (CAP), ad esempio, sostiene che se una persona vuole x e la desidera e pensa che una particolare azione y possa farla giungere ad avere x, allora a parità di condizioni, l'azione y sarà preferibile per raggiungere il proprio scopo¹¹⁴.

Si comprende, quindi, che si cominciò ad attribuire alla *folk psychology* un carattere primariamente epistemico, il cui strumento d'azione era affidato all'empatia, tramite la quale l'attribuzione di uno stato intenzionale di un agente sarebbe derivato dal suo comportamento manifesto; cambia, inoltre, la natura stessa dell'*inferenza teoretica* che, a differenza dell'inferenza per analogia, non si basa più sull'assunzione di un accesso privilegiato ai propri stati mentali per estenderli (per analogia, appunto) sull'altro, ma, in maniera "più distaccata", assume il *corpus* di conoscenze "già date" per spiegare lo stato emotivo-intenzionale che si trova a vivere. In questo senso, quindi, Karsten Stueber definisce questa concezione della *folk psychology* quale "concezione distaccata", la quale, però, incoraggia, per Stueber, non solo l'impovertimento della stessa agentività umana, definibile attraverso "leggi di copertura" già date, ma corre seriamente il rischio di falsare, completamente, le spiegazioni pseudo-scientifiche che la *folk psychology* vuole fornire.

L'incapacità, ad esempio, di riconoscere i propri stati mentali quali ragioni di un'azione rappresenta una seria mancanza: se l'individuo non avesse un accesso privilegiato in prima persona alle ragioni delle proprie azioni non sarebbe in grado, poi, di mantenere l'immagine sociale che egli proietta nel mondo; non saper attribuire a se stessi un dato comportamento,

¹¹⁴ Cfr. KARSTEN STUEBER, *Rediscovering Empathy. Agency, Folk Psychology and the Human Sciences*, The MIT Press, Massachusetts, 2010.

riconoscendolo come proprio, significherebbe essere un mistero per se stessi, rendendo di fatto impossibile qualsiasi intersoggettività. La “concezione distaccata della folk psychology”, quindi

*has distorted our understanding of the primary domain of folk psychology, that is, its use for the interpretation of conspecifics. In its primary domain, folk psychology cannot be understood as a theory that is adopted from the view from nowhere [...] but rather has to be minimally understood as being adopted from the engaged perspective. It is adopted from the point of view of agents who in experiencing their agenthood have the view themselves in folk-psychological terms and who use the same terminology to explain and predict the behavior of other human beings whom they conceive of as agents like themselves [...]*¹¹⁵.

Comprendere se stessi quali autori di un'azione significa capirne le ragioni per cui se ciò è ipotizzabile per sé lo sarà anche per tutti coloro che agiscono all'interno di un contesto comune, rendendo manifesto, non solo il comportamento adottato, ma anche le motivazioni sottese. In questo senso, subentra il carattere epistemico dell'empatia: quell'accesso privilegiato alle menti altrui diviene, nel paradigma di una concezione “propriamente vissuta” della *folk psychology*, lo strumento conoscitivo che consente al singolo di non essere una monade fra monadi, ma di interloquire e di rapportarsi all'altro da sé. La “lettura della mente” diventa, esattamente, la comprensione delle ragioni o delle motivazioni altrui, donando all'empatia sì un carattere epistemico, ma anche “universale”.

Diversi autori hanno voluto attribuire all'*empathy* il principio di razionalità, secondo cui qualsiasi azione un agente compia la compierebbe utilizzando norme di razionalità condivise; così come diversi altri non credono che la *folk psychology* possa essere considerata un paradigma razionale di spiegazione morale.

L'ambiguità stessa della nozione di razionalità, assumibile, a seconda del contesto e delle situazioni, in termini “soggettivi” o “oggettivi” (facendo appello, cioè, all'abilità dell'individuo di raggiungere lo scopo prefissato) è oggetto di ampi dibattiti e di numerose posizioni contraddistinte fra loro che, sommariamente, si possono suddividere fra i *proponents of the rationality view* e *proponents of the nonrationality view*. Essi, risultato di alcune ricerche accumulate in ambito psicologico negli ultimi trentanni, dimostrerebbero che l'essere umano, per sua natura, è portato ad allontanarsi dal “modello standard di razionalità”

¹¹⁵ *IDEM*, p. 47.

(la logica o la matematica), facendosi influenzare da stati emotivi o euristici, così come gli esperimenti di Daniel Kahneman ed Amos Tversky evidenzerebbero¹¹⁶.

Alcuni filosofi, come Stephen Stich, hanno sostenuto che, alla luce di tali evidenze empiriche, si dovrebbe abbandonare l'idea aristotelica di *animali razionali* per avviarsi verso un paradigma più adeguato della propensione umana di tradire gli *standard* della razionalità per sposarne uno più "emotivo" o "euristico"; a posizioni così drastiche se ne sono succedute diverse che, con una certa prudenza, raccomandano solo di tener presente che, *a volte*, gli esseri umani agiscono con poca razionalità; non che ne siano completamente assenti. Fatto sta che l'interpretazione degli esperimenti di Kahneman e Tversky, secondo altri, è stata compromessa da alcune assunzioni preliminari che ne avrebbero inficiato i "risultati", giungendo ad una enunciazione del modello di non razionalità dell'umano del tipo che segue:

1. *Our best theories of logic, probability, and decision making articulate the appropriate standards for the evaluation of behavior as rational or irrational;*
2. *Humans are rational animals if and only if their minds are rationally organized; that is, they are rational if and only if they are programmed according to the principles of our best normative theories of rationality;*
3. *If our minds are rationally programmed, then our behavior (including our reasoning) would conform sufficiently to the principles of our best normative theories;*
4. *The psychological experiments cited above show that our behavior does not sufficiently conform to the principles of our best normative theories.*

*It thus follows that humans are not rational animals*¹¹⁷.

Esulando dalle confutazioni che si sono succedute al modello così schematizzato, è importante notare che il dibattito sull'assunzione di razionalità o di non razionalità nella scelta euristica di un comportamento piuttosto che di un altro sembrerebbe essere troppo sbrigativo e troppo allarmista per sconvolgere il filosofo *in poltrona* che tenta di capire la natura umana. Il punto, cioè, è che, partendo da un problema prettamente ontologico (la natura della mente, risalente, simbolicamente a Cartesio), si è passati ad una denaturalizzazione del significato comune di *folk psychology* (tramite Sellars) per conferire a quest'ultima uno statuto epistemico e gnoseologico; attraverso l'utilizzo dello *strumento empatia* e per il mezzo del

¹¹⁶ Cfr. DANIEL KAHNEMAN E AMOS TVERSKY, *On the Reality of Cognitive Illusions*, in "Psychological Review", 103, 1996, pp. 582-591.

¹¹⁷ KARSTEN STUEBER, *Rediscovering Empathy*, op. cit. p. 55.

dibattito sulla filosofia del linguaggio, si è giunti, nuovamente, al problema da cui si era partiti: che natura ha l'uomo: dualista o "monista"?

In questo giro lungo e contorto del pensiero, è nell'ultimo passaggio che si è iniziato a parlare di *mindreading* o *mentalizing* quale strumento diretto di comprensione dell'altro, poiché dotando la *folk psychology* di una legittimità al pari della scienza, ne è conseguito che anche il suo utilizzo conferisce legittimità alle conoscenze da essa prodotte. Ecco, allora, la nascita di tutti i problemi sullo *status* ontologico dell'empatia, sull'*episteme* da essa prodotta e sulla sua natura confusa: che tipo di conoscenza è (che plausibilità o validità possa avere) una conoscenza che, passando per il tramite di un "sentire comune", si arroga il diritto di *spiegare* il Tu quale alterità distaccata dal Sé?

Non a caso, in un contesto già abbastanza torbido, coloro che hanno contribuito ad agitare ancor di più le acque, introducendo l'espressione *mindreading*, sono stati proprio due etologi, David Premack e Guy Woodruff, che hanno contribuito, *sine dubio*, a riportare all'attenzione questo particolare uso dell'empatia, esulandola, però, dal contesto della *folk psychology*, ed aggiungendo, forse, falsi problemi ad altri falsi problemi.

I due etologi statunitensi, nel saggio *Does the chimpanzee have a Theory of Mind?*¹¹⁸, pubblicato nel 1978, ipotizzarono che anche gli scimpanzé possiedono quella che si potrebbe definire *intelligenza machiavellica*¹¹⁹ propria di una teoria della mente: l'abilità, vale a dire, di capire le intenzioni altrui, adottando contromosse o completandone l'azione.

Sintomo ed espressione di una forte intelligenza sociale e di una strategia interpretativa propria a cui il filosofo statunitense Daniel Dennett ha dato il nome di *atteggiamento intenzionale*¹²⁰:

L'atteggiamento intenzionale è la strategia per interpretare il comportamento di un'entità (non importa se persona, animale o artefatto), trattandola come se fosse un agente razionale che orienta la propria 'scelta' d'azione' prendendo in 'considerazione' le proprie 'credenze' e i propri 'desideri' [...] ¹²¹.

[Una sorta di, n.d.r.] 'modulo' con una 'teoria della mente', progettato per generare credenze del secondo ordine (credenze riguardo a credenze e altri stati mentali altrui) [...] ¹²².

¹¹⁸ DANIEL PREMACK E GUY WOODRUFF, *Does the chimpanzee have a theory of mind?*, in "Behavioral and Brain Sciences", 4, 1978, pp. 515-629.

¹¹⁹ Cfr. RICHARD W. BYRNE E ANDREW WHITEN, *Machiavellian Intelligence*, Oxford University Press, Oxford, 1997.

¹²⁰ DANIEL DENNETT, *L'Atteggiamento Intenzionale*, op. cit.

¹²¹ *IBIDEM*, pp. 38-39.

¹²² *IBIDEM*, p. 480.

I due etologi, dopo aver osservato il comportamento dei primati non umani, hanno constatato che, “fenomenologicamente”, i loro atteggiamenti erano guidati da qualcosa di più sofisticato di una semplice risposta ad uno stimolo di pavloviana memoria; come se:

[essi dessero, n.d.r.] *senso a ciò che ved[ono, n.d.r.], assumendo che l'agente umano desider[i, n.d.r.] la banana e fa di tutto per raggiungerla [...]*¹²³,

entrando in *empatia* con gli sperimentatori. Ciò ha portato, inevitabilmente, i due studiosi a cambiare il punto di vista interpretativo con cui, precedentemente, avevano osservato il comportamento degli animali non umani, ponendo al centro delle proprie riflessioni non più l'interrogativo specifico su *che cosa sa del mondo uno scimpanzé?*¹²⁴, ma, allargando il proprio orizzonte di mondo,

*che cosa sa il soggetto circa ciò che qualcun altro conosce (ipotizza, crede, pensa ecc..) del mondo?*¹²⁵,

dove con il termine “soggetto” non ci si riferiva più esclusivamente all'uomo, bensì a qualsiasi essere senziente. Mutando il proprio paradigma di riferimento¹²⁶ e spostando il *focus* dell'osservazione dal mondo fisico a quello “metarappresentazionale“, “metaconcettuale” degli scimpanzé, gli autori sono giunti alla conclusione che i primati non umani possiedono una “teoria della mente”, giustificandone la presenza grazie all'uso dell'empatia, seppur declinata in un modo decisamente *sui generis*;

*The empathy view starts by assuming that the animal imputes a purpose to the human actor, indeed understands the actor's predicament by imputing a purpose to him. The animal takes over the actor's purpose, as it were, and make a choice in keeping with that assumed purpose. The empathy view diverges only in that it does not grant the animal any inferences about another's knowledge; it is a theory of mind restricted to purpose [...]*¹²⁷.

¹²³ *IDEM.*

¹²⁴ *IDEM*, p. 522.

¹²⁵ *IVI.*

¹²⁶ Cfr. THOMAS KHUN, *La struttura delle rivoluzioni scientifiche*, op. cit., 2009.

¹²⁷ ALVIN GOLDMAN, *Simulation Mind. The Philosophy, Psychology and Neuroscience of Mindreading*, Oxford University Press, Oxford, 2006, p. 11.

La restrizione del concetto stesso di empatia a degli “stati intenzionali” non è stata chiarita in modo preciso e puntuale dai due etologi; tanto che tale aspetto e talaltri riguardanti l’assunzione da parte dei primati non umani di una capacità di *mentalizing* sono stati oggetto di forti critiche e confutazioni, che, però, hanno determinato una spinta propulsiva agli studi dell’attività del *mindreading* o, semplicemente, della “lettura delle altre menti”¹²⁸.

3.2 L’empatia fra *Theory Theory* e *Simulation Theory*

L’insieme delle conoscenze delineatesi negli anni dal Neopositivismo Logico ad oggi ha subito, come è evidente, una progressiva mutazione che si è consolidata all’interno di due grandi paradigmi di riferimento: la *Theory Theory*, da una parte; la *Simulation Theory*, dall’altra.

La *folk psychology*, vale a dire, ha assunto “sembianze” e “vesti” distinte a seconda del modo in cui è stato interpretato il relazionarsi all’altrui soggettività, cercando, sempre e comunque, di ovviare o di superare il problema cartesiano di conoscenza fra *res cogitans*.

Oggi parlare di due concezioni nettamente distinte è fuorviante, poiché il discorso sulla *folk psychology* è divenuto la risultante ibrida di entrambi i paradigmi. Infatti, porre dei “paletti virtuali” sui rispettivi confini è un compito arduo ed alle volte quasi impossibile: il punto di congiunzione, indipendentemente dagli sviluppi di entrambi, si ritrova nella capacità, forse non del tutto umana, di inter-relazionarsi con l’altro e di capirne, *in qualche modo*, desideri, intenzioni, volizioni attraverso quel *mindreading* o “lettura della mente”, basantesi, appunto, sulle capacità cognitive dell’essere umano; oggetto, queste ultime, di una certa discordia interpretativa che caratterizza, come un filo rosso, l’una e l’altra concezione.

Il problema, vale a dire, si gioca sulla natura delle molteplici e svariate attribuzioni che l’agente può concedere agli stati mentali dell’altro da sé, interrogandosi, infatti, se queste ultime seguano una “teoria scientifica” (la *folk psychology*) oppure delle specifiche capacità cognitive, frutto dell’evoluzione.

Nel primo caso, la *folk psychology* sarebbe equiparata, come già detto, ad un quadro teoretico utilizzato dall’agente attraverso cui inferire quello dell’interlocutore: Daniel Dennett, ad esempio, pur condividendo in parte questa impostazione, sostiene che equiparare la “psicologia di senso comune” a teoria empirica è fuorviante, poiché essa è sì un insieme di

¹²⁸ Cfr. SUZANNE CUNNINGHAM, *What is a mind? An introduction to the Philosophy of Mind*, Hackett Publishing Company, Inc., Cambridge, 2000.

generalizzazioni condivise, ma la loro natura è solo normativa per l'uso razionale dell'agente nel vivere sociale.

Nel secondo caso, invece, la *folk psychology* sarebbe una “teoria” falsa, in quanto, come Fodor sosteneva, si riferirebbe ad oggetti (gli stati mentali) completamente inesistenti, che qualsiasi teoria aspiri ad esser tale non dovrebbe considerare nella sua ontologia strutturale.

In altre parole,

*Ciò significa, dunque, che [la folk psychology n.d.r.] debba essere letta secondo un'accezione psicologica realistica, come se tali principi venissero in qualche modo rappresentati interiormente dalle persone competenti nella psicologia del senso comune? Oppure significa semplicemente che la struttura delle nostre pratiche interpretative di psicologia del senso comune vengono rese trasparenti in modo sistematico solo se si concretizzano in una teoria psicologica?*¹²⁹.

3.2.1. La folk-empathy nella Theory Theory

La “Teoria della Teoria” o *Theory Theory* (d'ora in poi “TT”) è quella spiegazione del comportamento umano che si propone di comprenderlo non attraverso il ricorso a particolari o speciali capacità cognitive, ma utilizzando un insieme di leggi “comuni” (*folk*) attraverso le quali attribuire credenze, desideri e volizioni all'agente altro da sé. Gli psicologi della TT

*[...] tendono a considerare le strutture concettuali sottostanti alle nostre capacità di psicologia del senso comune come identiche o assai simili alla struttura concettuale del discorso ordinario della psicologia del senso comune. Inoltre e di pari importanza, il paradigma della teoria della teoria attrae in quanto fornisce un paradigma che ci permette di osservare lo sviluppo cognitivo e l'accrescimento della complessità di concettualizzazione del bambino in modo unificato in diversi campi di indagine. Ciò ci permette di considerare la maturazione della nostra comprensione degli aspetti fisici, biologici e psicologici del mondo come dipendente dall'acquisizione di strutture conoscitive sempre più complesse riguardanti le diverse aree di interesse dell'uomo [...]*¹³⁰.

¹²⁹ KARSTEN R. STUEBER, *L'empatia*, Il Mulino, Bologna, 2010, p. 69.

¹³⁰ *IDEM*, p. 160.

La prima significativa distinzione da fare, all'interno di questo quadro teorico, è quella relativa al modo in cui intendere la comprensione degli stati mentali altrui. Secondo David Lewis, ad esempio, la psicologia di senso comune sarebbe niente più che una raccolta di

[...] *tutte le ovvietà che puoi pensare riguardo la relazione causale degli stati mentali, gli stimoli sensoriali e le risposte motorie* [...].

Aggiungi anche le ovvietà all'effetto che uno stato mentale cade sotto un altro [...]. *Forse ci sono anche banalità di altre forme. Includi solo le ovvietà che sono comuni conoscenza tra noi: ciascuno le conosce e ognuno sa che ciascun altro le conosce, e così via* [...] ¹³¹.

Secondo Lewis, infatti, la psicologia di senso comune altro non sarebbe che la sommatoria di un insieme di ovvietà, organizzate concettualmente per l'attribuzione intenzionale di credenze e volizioni, che farebbero parte di un insieme integrato di concetti generali per la "lettura della mente" altrui. In questo senso, quindi, il quadro teorico della *TT esterna*, cioè, sarebbe fornito solo da quell'insieme di ovvietà del contesto di riferimento cui appellarsi per la comprensione del comportamento altrui; la spiegazione dell'*empatia* – si potrebbe sostenere, forse, con una certa forzatura – si troverebbe "al di fuori di sé", quale strumento concettuale, quindi cognitivo, cui appellarsi per rendere possibile l'intersoggettività propriamente detta.

Se per ogni teoria scientifica si accetta unanimemente che l'uso di un termine derivi dall'insieme di credenze che costituiscono il paradigma dominante allora anche nel caso della psicologia di senso comune, in accordo a quanto sostenuto da Sellars sul *Mito di Jones*, si dovrà concedere che l'uso funzionale dei termini intenzionali derivino il loro significato dall'insieme delle ovvietà concettualmente organizzate per capire *la mente* altrui. In questo modo qualsiasi *mindreading* sarà supportata da altrettante "credenze teoricamente organizzate" sul concetto di *mente*, costituito, a sua volta, da un corpo di conoscenze, "leggi comuni", in grado di connettere stati mentali con risposte corporee, comportamentali con altri stati mentali e così via.

L'ulteriore differenziazione all'interno della *TT esterna* concerne il *modo* in cui la *Theory of Mind* sarebbe acquisita e, di contro, anche *da che cosa* o *da chi* discenderebbe: da innate capacità cognitive o dalla stessa *folk psychology*?

Diversi autori, non condividendo quanto Lewis sostiene sulla natura della psicologia di senso comune, propongono una versione *innatista* della *folk psychology* sulla scia delle

¹³¹ DAVID LEWIS, *Psychological and theoretical identification*, in "Australian Journal of Philosophy", n. 50, 1972, p. 256.

considerazioni di Fodor, potendo, addirittura, parlare di una “psicologia di senso comune funzionalista”: essendo i concetti integrati in una rete integrata, l’uso dell’attribuzione intenzionale all’altro dipenderebbe dal mero uso *funzionale* del ruolo causale assegnatoli per il raggiungimento dello scopo finale, ossia capire l’altro.

A tal proposito, Davies Stone sostiene:

[...] *Si può dire che i concetti mentali che comprendono la nostra psicologia quotidiana o del senso comune – come credenze, desideri, speranze, provare dolore ecc – sono parte di una rete integrata di concetti, così che comprendere uno di questi concetti richiede la comprensione di qualcuno o tutti gli altri [...]*¹³².

I *Child-Scientist Theory-Theory*¹³³, ad esempio, sostengono che gli studi sullo sviluppo cognitivo dei bambini può essere illuminante per gettar luce sul modo in cui spiegare il *mentalizing*: basandosi sull’analogia fra un bambino stesso ed uno scienziato, essi sostengono che la psicologia di senso comune, così come qualsiasi teoria scientifica non permette solo un mero immagazzinamento di informazioni, ma mette nella condizione, chi ne fruisce, di elaborare, a sua volta, spiegazioni generali dei macrofenomeni osservati. Tale passaggio, però, sarebbe possibile solo grazie alla postulazione di regolarità non direttamente osservabili, ma ipotizzabili proprio grazie alle generalizzazioni teoriche condivise dalla teoria scientifica o, in questo caso, dalla *folk psychology* adottata; tenendo ben presente che le stesse teorie scientifiche sono sottoposte a cambiamenti per integrare quelle *anomalie* altrimenti ignorate, gli stessi paradigmi concettuali vanno incontro a parziali mutamenti, integrazioni o, come ci dice Kuhn, ad abbandoni progressivi.

Per descrivere, allora, lo sviluppo psico-motorio osservabile nei bambini, secondo i *Child-Scientist* della TT, lo scienziato dovrà, necessariamente, ricorrere ad un incremento di osservazioni via via sempre più complesso per il fine propostosi; in tal caso, nello studio dello sviluppo cognitivo dei bambini dai tre ai cinque anni, egli, necessariamente, ricorrerà a fasi prestabilite, postulate come tali dall’insieme di conoscenze acquisite dalla teoria e dall’osservazione. Succederà, allora, ad una fase di “psicologia del desiderio”, una “psicologia intermedia del desiderio-credenza” ed una “psicologia

¹³² DAVIES STONE, *The mental simulation debate: a progress report*, in PETER CARRUTHERS E PETER K. SMITH (A CURA DI), “Theories of Theories of Mind”, Cambridge University Press, Cambridge, 1996, p. 121.

¹³³ Cfr. ALISON GOPNIK E ANDREW N. MELTZOFF, *Words, Thoughts and Theories*, Mit Press, Massachusetts, 1997.

del desiderio-credenza”¹³⁴, in cui la forza motrice del cambiamento in questo modello sarà data dalla forza dell’apprendimento progressivo che il bambino compie durante gli anni considerati. Secondo i teorici modulari della TT¹³⁵, i cambiamenti concettuali protagonisti di queste tappe non deriverebbero da un continuo sviluppo delle capacità cognitive, ma dall’acquisizione sempre più ricca del mondo che circonda il minore; essi mettono in evidenza l’esistenza di un progressivo sviluppo del bambino più o meno sempre uguale, poiché *standard* sarebbero le modalità di accrescimento delle conoscenze teoriche della *folk psychology* in cui si relazionerebbe il bambino stesso. In questo contesto, perciò, il passaggio ad una fase successiva deriverebbe, secondo la TT *esterna*, da una rivisitazione del *corpus* acquisito in maniera più profonda e più sofisticata, mentre per i teorici modulari alla base del cambiamento vi sarebbe un’interazione fra diversi moduli cognitivi, coinvolti nel processo di *mentalizing*. Per usare l’impianto teorico di Baron-Cohen, ad esempio, un bambino riuscirebbe ad attribuire desideri e volizioni ai suoi simili attraverso la funzionalità del modulo ID e SAM, responsabili degli stadi epistemici della mente. Stando alla sua posizione, infatti, le capacità di “lettura della mente” dipenderebbero da quattro distinti sottosistemi modulari: il riconoscimento dell’intenzionalità (ID, *Intentionality Detector*); il riconoscimento della direzione dello sguardo (EDD, *Eye Detection Detector*); il meccanismo dell’attenzione condivisa (SAM, *Shared Attention Mechanism*); il modulo della teoria della mente (ToMM, *Theory of Mind Model*)¹³⁶.

Secondo gli ultimi scritti di Leslie, il modulo ToMM avrebbe, principalmente, la funzione di accumulare una molteplicità di concetti sugli stati mentali che contribuirebbero, poi, alla formazione delle “metarappresentazioni sull’agente”, ossia su possibili spiegazioni del “mondo vissuto” dall’altro¹³⁷. Sebbene Leslie sia riconosciuto come un teorico della TT, sembrerebbe, a detta di molti, che negli ultimi anni abbia in parte rimodulato le sue convinzioni; poiché, ad esempio, la specificazione che il modulo ToMM non abbia ruoli causali nella formazione dell’attribuzione intenzionale all’altro significherebbe già un distacco dalla concezione modulare funzionalista della TT. Al di là delle differenze interpretative all’interno dello *status* della natura *esterna* della TT, con le parole stesse di Ian M. Ravenscroft se ne potrebbe dare una generica definizione come segue:

¹³⁴ Cfr. KEIN BARTSCH E HENRY M. WELLMAN, *Children Talk about the Mind*, Oxford University Press, Oxford, 1995.

¹³⁵ Cfr. SIMON BARON-COHEN, *Mindblindness: An Assay on Autism and Theory on Mind*, MIT Press, Massachusetts, 1995.

¹³⁶ *IDEM*.

¹³⁷ ALAN M. LESLIE, *How to acquire a representational theory of mind*, in DAN SPERBER (A CURA DI), “Metarepresentations: A Multidisciplinary Perspective”, Oxford university Press, Oxford, 2000, pp. 197-224.

[The “external” concept of folk psychology is, n.d.r.] a theory of mind implicit in our everyday talk about mental states. In the every day traffic of our lives we make remarks linking sensory experiences to mental states; mental states to other mental states; and mental states to behavior. Thus we remark that the smell of freshly baked bread made Sally feel hungry; that Sally wanted to go on diet because she thought that she was overweight; and that Sally went to the fridge because she desired a piece of chocolate cake [...] ¹³⁸.

Se, quindi, quanto appena detto si riferisce alla versione “esternalista” della TT, altro discorso lo si deve ai teorici *interni* della TT, secondo i quali la capacità di “mentalizzazione” deriverebbe da capacità cognitive proprie dell’agente, che, coadiuvate dal *corpus* conoscitivo della *folk psychology*, conferirebbe uno stato volitivo all’altro da sé.

Le linee di ricerca di questa particolare versione della TT si potrebbero schematizzare in quattro generici punti che forniscono il senso della differenza interpretativa tra i due approcci; essi, seppur con delle differenze, condividono il nucleo centrale dell’impianto teorico; ossia l’uso di concetti della *folk psychology* per riuscire nella risoluzione del problema dell’attribuzione.

Per una maggiore esplicitazione del problema si è ritenuto di riassumerli come segue:

- a. I teorici della TT *interni*, a differenza della variante *esterna*, non sono interessati a fornire una particolare rappresentazione mentale (si ricordi, ad esempio, lo stesso Baron-Cohen), poiché centrale diviene il modo in cui la *folk psychology* è “rappresentata” nel linguaggio del pensiero o da una rete di connessioni concettuali che, attraverso l’ausilio delle capacità cognitive dell’agente, conferisce un senso ed un significato all’altro da sé ¹³⁹.

Come è facile intuire, tale posizione solleva il classico problema di come sia possibile una teoria “interna” di tali rappresentazioni mentali: si torna, cioè, all’interrogativo cartesiano su come sia possibile, per una materia, *pensare*, considerando, inoltre, che i recenti sviluppi neuroscientifici pongono ulteriori problemi sul modo in cui un neurone (per semplificare) possa rispondere *concettualmente* ad uno stimolo esterno attraverso una rete sinaptica.

¹³⁸ IAN M. RAVENSCROFT, *Folk Psychology as a Theory*, in “Stanford Encyclopedia of Philosophy”, consultabile al seguente indirizzo internet:

<https://stanford.library.sydney.edu.au/archives/spr2010/entries/folkpsych-theory/> .

¹³⁹ STEPHEN STITCH E SHAUN NICHOLS, *Folk Psychology: simulation or tacit theory?*, in “Mind and Language”, n. 7, pp. 35-71;

Gareth Evans, a tal proposito, ha suggerito che posti di fronte ad un sistema cognitivo S che emette X ogni qualvolta ha un certo *input* “X&Y”, il cui problema è quello di stabilire in quali condizioni il sistema risponda come *output* “&-eliminazione” per cui rimarrebbe solo X; una possibilità sarebbe quella di ipotizzare che il sistema S possieda una *rappresentazione* dell’*output*; in tal modo, tutte le volte che fosse sollecitato da “X&Y”, applicando la regola “&-eliminazione” si avrebbe X¹⁴⁰. Tale proposta, però, genera non poche perplessità quando si parla di “concetti mentali.

Ad oggi, è ancora inimmaginabile riuscire a catalogare con determinati *input* e relativi *output* la materia cerebrale: dove si troverebbe, infatti, il “ricordo” o la “rappresentazione” di cui Evans parla? È vero sì che la moderna farmaceutica è in grado, grazie a dei beta-bloccanti, di inibire la formazione del ricordo se ingeriti entro sei ore dall’evento (pratica tuttora in uso negli ospedali statunitensi per pazienti ricoverati per gravi, gravissimi traumi vissuti o subiti¹⁴¹), ma è pur vero che progressi sulla “localizzazione cerebrale” per ciò che concerne il pensiero astratto, è ancora *in itinere* (confidando che, un giorno, il cervello riuscirà a capire se stesso).

- b. I teorici della TT *interna* non rifuggono all’idea che la *folk psychology* possa essere appresa quale contenuto concettuale che solo da quel momento in poi potrà coadiuvare le capacità cognitive per la relativa “lettura della mente”.

Il problema centrale per ciò che concerne questo aspetto è proprio la natura stessa della *folk psychology*: alcuni autori, infatti, sostengono che essa sia in gran parte innata e che, seguendo Fodor, l’uomo abbia un particolare modulo cerebrale predisposto alla sua acquisizione. Data la particolare assonanza del tema con quello relativo al nativismo della psicolinguistica, non sono mancati confronti e parallelismi sul tema: tanto i primi quanto i secondi, infatti, ricorrendo alla teoria darwiniana, sostengono che sia la “mentalizzazione” sia il “linguaggio” siano predisposizioni evolutive frutto della selezione naturale. Esulando dalle critiche avanzate alla possibile eredità ambientale e genetica della psicolinguistica, un dubbio sollevato ai teorici della TT *interna* è stato quello della “povertà di argomentazione”: come è possibile, si chiede, che un bambino nella prima infanzia (0-3 anni) possa essere esposto ad un ambiente talmente ricco di varietà comportamentali tale che possa acquisire parte dei concetti della *folk psychology*? Sembrerebbe, cioè, poco plausibile sostenere che, data la giovanissima età, un minore possa avere già una teoria *morale* nel senso di *comportamentale* completa

¹⁴⁰ Cfr. GARETH EVANS, *Semantic Theory and Tacit Knowledge*, in SAM HOLTZMAN E CHRISTOPHER M. LEICH (a cura di), “Wittgenstein: To Follow a Rule”, Routledge & Kegan Paul, Londra, 1981.

¹⁴¹ Si rimanda il lettore al capitolo I.

dell'uomo. Alvin Gopnik, Henry Wellman, Andrew Meltzoff e molti altri hanno tentato di fornire una risposta, ipotizzando il modello in termini di “sviluppo della teoria” nello stesso modo in cui Fodor ha sostenuto che il modulo predisposto all'acquisizione della *folk psychology* rifletterebbe le classiche fasi di sviluppo delle capacità cognitive¹⁴².

- c. Un aspetto problematico cui incorrono i teorici della TT *interna* è come spiegare l'autismo.

Se fosse vero che la *folk psychology* avesse una natura innata, si dovrebbe spiegare il motivo per cui bambini autistici, sottoposti a *test* per la capacità di mentalizzazione, non risponderebbero con dei risultati più o meno soddisfacenti in linea con le predizioni (viste le difficoltà stesse che l'autismo comporta). Stando ai dati raccolti, secondo i teorici della TT *interna*, a tali bambini mancherebbe la capacità stessa di utilizzare la *folk psychology*¹⁴³ o, di contro, potrebbero essere l'evidenza che il costrutto teorico su cui si basano le medesime ricerche sia sbagliato. In un caso o in un altro, gli studi in questo filone stanno proseguendo anche grazie alle continue scoperte neuroscientifiche che, ricordiamo, sono ancora ben lontano dal “presentare al mondo” l'organo più complesso senza più segreti.

- d. I nuovi sviluppi della TT *interna* sembra stiano giungendo dall'incontro di due discipline: la filosofia e la psicologia sociale.

Barbara von Eckard, psicologa sociale, sostiene, ad esempio, che l'ambito di questo particolare indirizzo psicologico non si esaurisca al solo studio “dell'uomo in società”, ma inerisca anche particolari aspetti della stessa filosofia nelle sue tematiche della *folk psychology*, TT *interna*. Secondo la studiosa, il giudicare gli altri, quindi farne oggetto prima di mentalizzazione e dopo di oggetto di valutazioni proprie, potrebbe essere il sintomo di poca astuzia intellettuale da parte del giudicante e di relativa poca stima da attribuire a se stesso. Ciò comporterebbe, per i teorici della TT *interna*, che esempi del genere (con sofisticazioni e peculiarità che lasceremo scoprire al lettore) potrebbero essere un segno che *qualcosa* nella *folk psychology* innata “non abbia funzionato”¹⁴⁴.

Il filone intrapreso è solo all'inizio, ma promette per il futuro allettanti dati di indagine.

¹⁴² Se tali risposte siano convincenti, persuasive o manchevoli non spetta a noi, in questa sede, dirlo, ma ciò che, invece, ci è doveroso segnalare è la loro stessa esistenza per l'ipotetico lettore.

¹⁴³ SIMON BARON-COHEN, ALAN LESLIE ED UTA FRITH, *Does the Autistic Child have a 'Theory of Mind*, in “Cognition”, n. 21, 1985, pp. 37-46.

¹⁴⁴ Cfr. BARBARA VON ECKARD, (1997) *The Empirical Naivete of the Current Philosophical Conception of Folk Psychology*, in MARTIN CARRIER E PETER MACHAMER (A CURA DI), “Mindscapes”, University of Pittsburgh Press). Pittsburgh, 1997.

Da come si evince e da quanto era stato predetto, il quadro teorico della TT è decisamente controverso e multifaccettato, ma non lo è di meno quello opposto: la *Simulation Theory* (d'ora in poi "ST"); il quale, riportando il discorso ad un piano prettamente storico, si potrebbe far risalire a Lipps ed al suo famosissimo esempio della visione del funambolo.

Una considerazione, però, prima di passare ad illustrare l'altro complesso mondo della ST, sembrerebbe d'obbligo. Si è aperta questa sezione, cercando di illustrare, brevemente, ma (speriamo) con efficacia argomentativa, il modo in cui il concetto di *empatia* si è "trasformato" negli ultimi cinquanta anni per spiegare i motivi, le ragioni per cui, oggi, nel dibattito attuale si parla di *empatia cognitiva*. È stata definita *co-cognition*, *empatia ricostruttiva*...ma il punto su cui vorremmo soffermarci è il seguente: nel discorso *in itinere*, dalla TT alla ST, dalla *folk psychology* al *mindreading*, che fine ha fatto l'empatia?

Nella TT in generale essa diviene uno "strumento" prima esterno all'agente per attribuire stati intenzionali all'altro, poi, in qualche modo, interno all'osservatore che, coadiuvato dall'insieme di conoscenze o di ovvietà della *folk psychology*, torna all'interno del soggetto per svolgere il suo ruolo canonico. È, certamente, una esemplificazione, ma è stato un passaggio esplicativo necessario per non perdere di vista quel filo rosso che si sta cercando di tracciare con questo studio. Considerazioni, che si promette di approfondire nelle *Conclusioni*, in cui si cercherà di sollevare un dubbio sulla natura stessa dell'empatia nella sua accezione cognitiva.

Se, allora, il passaggio dell'*empathy* nella TT è stato dall'esterno all'interno, seppur, in quest'ultimo caso, mediata *in primis* da concetti e rappresentazioni, il quadro teorico che sostiene che essa sia solo ed esclusivamente derivante da specifici meccanismi cerebrali è quello relativo alla ST.

3.2.2. La neuro-empatia e la *Simulation Theory*

Il secondo paradigma concettuale che prende avvio dall'utilizzo delle *folk psychology* per spiegare l'apertura o la comprensione dell'altro a partire da "stati mentali" che ripropongono in sé quelli del Tu è, come detto, quello della ST.

A livello generale, esso sostiene che il *mindreading* non sia più guidato dalla teoria (*theory driven*), ma dai processi cerebrali dell'interprete stesso (*process driven*): la psicologia di senso

comune, quindi, non si baserebbe su capacità teoriche di apprendimento o di comprensione, ma sull'uso

[...] *delle motivazioni e delle risorse emotive proprie di una persona e dalla sua particolare capacità di ragionamento pratico* [...] ¹⁴⁵.

Parafrasando le parole di Gordon, si comprende come il *focus* della ST sia quello di spiegare l'intersoggettività attraverso un modello prettamente *naturale* o *biologico* ed *egocentrico*. Le capacità cerebrali tipiche dell'uomo e, probabilmente, anche di altri animali non umani, sarebbero in grado di avviare "internamente" la "comprensione" degli stati mentali altrui grazie a particolari meccanismi cognitivi, che, dagli anni '90 in poi, si ipotizzano siano i neuroni specchio insieme ad altre aree dell'intero encefalo. Ciò, quindi, consentirebbe all'agente di "mettersi nei panni altrui", trasponendo il punto di vista proprio sul mondo dal personale a quello dell'altro, comprendendo/"sentendo su di sé" emozioni, volizioni o credenze e desideri del Tu. Compiendo *naturalmente* questo tipo di operazione, l'osservatore, di contro, non si alienerebbe nell'osservato, ma utilizzerebbe se stesso quale prototipo di comprensione; divenendo, in tal senso, il *modello egocentrico* dell'empatia stessa. Quest'ultima diverrebbe lo strumento biologico attraverso il quale "risolvere" il dualismo cartesiano, poiché fornirebbe una spiegazione trascendentale della comunicazione fra potenziali monadi, entrando *in medias res* nella mente altrui.

A questo proposito, non è insolito trovare autori che, condividendo la ST, spieghino l'*empathy* con il suo risultato "naturale", ossia con il *mindreading* o *mentalizing*; tali espressioni sarebbero intese come l'attribuzione di stati intenzionali che non escluderebbero *a priori* l'ausilio di concetti, ma subentrerebbero solo a simulazione avvenuta (per cui in un secondo momento) ¹⁴⁶ anche se alcuni recenti studi sembrerebbero porre in evidenza che l'utilizzo della competenza psicologica potrebbe essere piuttosto marginale ¹⁴⁷.

Il termine "simulazione" consente la comprensione del modello egocentrico, sia in termini neuroscientifici sia cognitivi: se, infatti, nel primo caso l'espressione è usata per indicare l'attivazione (solitamente automatica, istintiva preconsa) di una risposta neurale a seguito di un comportamento osservato, nel secondo la psicologia non nega l'ausilio di ragionamenti euristici nel vivere quotidiano e di ragionamenti simulazionistici del proprio o dell'altrui

¹⁴⁵ ROBERT M. GORDON, 'Radical' Simulationism, in PETER CARRUTHERS E PETER K. SMITH (A CURA DI), "Theories of Theories of Mind", op. cit., p. 11.

¹⁴⁶ Cfr. ALVIN GOLDMAN, *Simulation Mind*, op. cit.

¹⁴⁷ VITTORIO GALLESE, *Before and Below 'Theory of Mind': Embodied Simulation and the Neural Correlates of Social Cognition*, in "Phil. Trans R. Soc. B.", 362, pp. 659-669.

comportamento alla luce delle parziali conoscenze a disposizione in quel momento. Il termine “simulazione”, quindi, applicato alla teoria della *folk psychology*, vorrebbe evidenziare proprio questo particolare aspetto: un’attenzione biologica, di neuroscienza cognitiva e di psicologia sociale quale possibile spiegazione dell’intersoggettività. Non a caso,

Ciò che i teorici della simulazione definiscono simulazione è essenzialmente la procedura con cui una persona usa i propri processi cognitivi in una modalità imitativa allo scopo di acquisire informazioni sulle altre menti per predire e spiegare le azioni di un’altra persona. Simulando un’altra persona, io cerco di capire le sue reazioni, chiedendomi come reagirei alla situazione in cui questa si trova, pensando al suo problema sia ‘nella realtà’ – come nel caso in cui il problema si presenta sulla scacchiera – sia ‘nell’immaginazione’, immaginando, cioè, di trovarmi in una situazione simile alla sua¹⁴⁸.

La distinzione operata da Stueber tra “la realtà” e “l’immaginazione” non nasce casualmente: in un articolo di qualche anno fa, i filosofi Martin Davies e Tony Stone¹⁴⁹ ponevano l’accento sulla distinzione fra una *simulazione di situazione effettiva* ed una *simulazione di finzione off-line*.

Nel primo caso, come Stueber ha scritto, i due autori si riferivano ad un problema reale, di vita vissuta, in cui entravano in gioco i processi di simulazione e le euristiche del ragionamento normativo per superare e risolvere una difficoltà: dal dubbio se muovere o meno l’alfiere su una scacchiera a quale *personal computer* poter comprare. Nel secondo, invece, ci si troverebbe di fronte ad un problema “fittizio” o “virtuale”: ad esempio, la risoluzione di un’equazione aritmetica. In questi casi *off line*, cioè, noi alimenteremmo il nostro sistema cognitivo con credenze, desideri e volizioni senza alcun tipo di riscontro immediato della loro “esattezza” nella vita concreta. Se queste sono le differenze, ci si dovrebbe aspettare che anche i processi di “simulazione di base” siano differenti, per cui, come hanno suggerito Stich e Nichols¹⁵⁰, non si potrebbe più parlare di simulazione *tout court*, ma di altri processi ancora da spiegare, poiché secondo “i dettami” (per usare un’espressione forte) della ST i meccanismi sottesi al *decision-making* ed al *mentalizing* devono essere gli stessi, altrimenti non si può più parlare di *simulazione*.

¹⁴⁸ KARSTEN STUEBER, *L’Empatia*, op. cit. p. 167.

¹⁴⁹ MARTIN DAVIES E TONY STONE, *Mental Simulation, tacit theory and the threat of collapse*, in “Philosophical Topics”, 29, nn1/2, 2001, pp. 127-173.

¹⁵⁰ STEPHEN STITCH E SHAUN NICHOLS, *Cognition penetrability, rationality and restricted simulation*, in “Mind and Language”, 12, nn3/4, 1997, pp. 297-326.

Non a caso, Stueber, non condividendo quanto sostenuto da Stich e Nichols nello scritto “incriminato”, aggiunge a conclusione della spiegazione del termine “simulazione” quanto segue:

[...] *Utilizzando così la mia stessa attrezzatura cognitiva, potrei servirmi di ulteriori conoscenze teoriche riguardanti il mondo – conoscenze che presumo di condividere con l'altra persona [...]*¹⁵¹;

“segno tangibile” di quanto sia altrettanto multifaccettato e complesso il quadro teorico della ST.

Sull’ambiguità del termine *simulazione* si sono distinte all’interno della ST due filoni concettuali in cui si è posto l’accento su un aspetto interpretativo o su un altro: nel primo caso, il significato è stato spiegato quale “somiglianza”, la *Mental Simulation as Resemblance* (MSR1), nel secondo come “riutilizzo”, la *Mental Simulation as Reuse*.

Il problema, come si sarà intuito, oltre che semantico è soprattutto interpretativo.

A fondamento delle due macrodistinzioni, infatti, si celano le diverse spiegazioni date all’interrogativo di che cosa mai significhi che uno stato mentale sia la *simulazione* mentale di quello altrui; per quanto le recenti scoperte neuroscientifiche stiano tentando di fornire una “risposta”, il problema resta etico.

Nell’esaminare sostanzialmente il primo approccio, il MSR1, come suggerisce la parola stessa, interpreta gli “stati mentali” in termini di “somiglianza”, per cui un agente avrà il medesimo stato mentale M di un altro M* se rispetta la schematizzazione che segue:

1. *Both M and M* are mental states*
2. *M* resembles M in some significant respects*¹⁵².

Da come si evince, il concetto di “somiglianza” richiede:

- uno stesso ruolo funzionale dello stato mentale M;
- una base neurale simile;
- un contenuto simile;
- una stessa fenomenologia¹⁵³;

¹⁵¹ KARSTEN STUEBER, *L’Empatia*, op. cit. p. 167.

¹⁵² STANFORD ENCYCLOPEDIA OF PHILOSOPHY, lemma *Folk Psychology as Mental Simulation*, consultabile online al seguente indirizzo

<https://stanford.library.sydney.edu.au/entries/folkpsych-simulation/> .

per cui M sta simulando M* se e solo se si trova nella condizione di *essere la simulazione mentale di*, ossia ricorda (*resemblance*) quello di cui è la simulazione diretta. Un problema in cui questa spiegazione incorre è il modo in cui spiegare, effettivamente, tale somiglianza: come si potrebbe, infatti, verificare *de re* ciò che si assume come premessa del discorso stesso? A tal proposito, sempre Goldman ha risposto a tali difficoltà sostenendo che la relazione di somiglianza deve essere simmetrica, per cui:

Token state M is a mental simulation of token state M if and only if:*

1. *Both M and M* are mental states*
2. *M* resembles M in some significant respects*
3. *In resembling M, M* fulfils at least one of its functions*¹⁵⁴.

Ciò evidenzia come l'immaginazione sarebbe equiparata a qualsiasi simulazione (visiva, credenziale o di "scelta") per cui non si finirebbe nell'annoso problema di spiegare come classificare la simulazione mentale M* quale simile alla propria M.

Su tale aspetto, nonostante dissidi e nuove proposte, questa accezione della MSR quale *resemble* sembrerebbe fornire un discreto accordo tra i simulazionisti, anche se la stessa interpretazione della ST quale *reuse* sembrerebbe riscuotere un certo consenso.

Secondo quest'ultima, infatti, la spiegazione della ST quale "somiglianza" non marcherebbe abbastanza la distinzione fra l'approccio della ST e quello della TT, poiché la spiegazione data al processo di *resemble* non porrebbe l'accento sulla distinzione che di fatto vuole esistere tra spiegare un *meccanismo* cerebrale a fondamento della comprensione dell'agentività (ST) e spiegare un *processo* cognitivo che può riguardare, data l'ambiguità dell'uso del termine, la comparsa di particolari attività cognitive che, facendo uso di metarappresentazioni, spiegherebbero il comportamento manifesto (TT).

Susan Hurley, a tal proposito, mette in evidenza che la ST *reuse* dovrebbe presentarsi come segue:

Token state M is a mental simulation of token state M if and only if:*

1. *Both M and M* are mental states*
2. *M is generated by token cognitive process P*
3. *M* is generated by token cognitive process P**

¹⁵³ JUSTIN C. FISHER, *Does Simulation Theory Really Involve Simulation?*, in "Philosophical Psychology", 19(4), 2006, pp. 417–432.

¹⁵⁴ Cfr. ALVIN GOLDMAN, *Simulation Mind*, op. cit.

4. *P* is implemented by the use of a token cognitive mechanism of type *C*
5. *P** is implemented by the reuse of a token cognitive mechanism of type *C*¹⁵⁵.

Da tale esemplificazione si evince, secondo Hurley, che se il processo cognitivo può essere realizzato in strutture diverse, ciò che consente alla cognizione di potersi esprimere non è la capacità stessa di esistere, ma il meccanismo (cerebrale) che la rende possibile; non a caso, Hurley parla di *implementation* proprio per configurare quelle strutture fisiche organizzate che consentono il processo risultante: la cognizione, appunto.

Per ovviare alle difficoltà che entrambe le proposte hanno riscontrato, i simulazionisti hanno pensato di unirle così da configurare la ST quale risultante sia della *resemble* sia del *reuse*, in modo da superare, almeno ipoteticamente, i problemi sottesi; ne è nato il PROC, secondo cui

Token process P is a (mental) simulation process if and only if:*

1. *P** generates token state *M**
2. *M** resembles another token state, *M*, in some significant respects
3. Both *M* and *M** are mental states
4. *M* is generated by token process *P*
5. Both *P* and *P** are cognitive processes
6. *P* is implemented by the use of a token cognitive mechanism of type *C*
7. *P** is implemented by the reuse of a token cognitive mechanism of type *C*¹⁵⁶.

In realtà, seguendo la letteratura sul tema, sembrerebbe quasi che da questa fusione i simulazionisti siano incorsi in ulteriori gravi problemi con conseguenze, forse, poco prevedibili all'atto della "fondazione" del PROC.

Una premessa necessaria al discorso riguarda la consuetudine, oramai divenuta tale, tra i simulazionisti di argomentare sull'esistenza di due distinti processi di simulazione: l'*high-level* ed il *low-level*¹⁵⁷, anche tenendo conto degli studi di Goldman¹⁵⁸.

Per quanto riguarda "l'alto livello di simulazione", un processo può essere identificato come tale se soddisfa almeno tre caratteristiche fondanti:

¹⁵⁵ SUSAN HURLEY, *Understanding Simulation*, in "Philosophy and Phenomenological Research", 77(3), 2008, pp. 755–774.

¹⁵⁶ STANFORD ENCYCLOPEDIA OF PHILOSOPHY, lemma *Folk Psychology as Mental Simulation*, consultabile online al seguente indirizzo
<https://stanford.library.sydney.edu.au/entries/folkpsych-simulation/>

¹⁵⁷ FRÉDÉRIQUE DE VIGNEMONT, *Drawing the Boundary Between Low-Level and High-Level Mindreading*, in "Philosophical Studies", 144(3), 2009, pp. 457–466.

¹⁵⁸ Cfr. ALVIN GOLDMAN, *Simulation Mind*, op. cit.

- a. Devono essere processi coscienti, quindi, sotto il diretto controllo della volontà dell'agente;
- b. Devono soddisfare le peculiarità della scala PROC, ossia i meccanismi di simulazione devono essere implementati dal riutilizzo di particolari meccanismi cognitivi;
- c. Gli *output* risultanti di un dato meccanismo D dovranno essere uguali a quelli del medesimo meccanismo M; identificato, ad esempio, al processo di visualizzazione, di immaginazione visiva e motoria, il *decision-making* e simili¹⁵⁹.

Sulla scia dell'enunciazione di quanto appena detto è nato un intenso ed acceso dibattito sul modo in cui "catalogare" gli stati percettivi. Alcuni studi su pazienti affetti da NEGLET, ad esempio, hanno teso a dimostrare che l'immaginazione visiva ed il processo di visione siano generati dagli stessi meccanismi, per cui, soddisfacendo b., si potrebbe inserire come simulazione ad alto livello anche la visione; il problema sarebbe quello di riuscire a rispondere, affermativamente, alla domanda se esistono processi di simulazione ad alto livello i cui *output* siano simulazioni mentali di stati non sensoriali¹⁶⁰. In realtà, al di là del dibattito suscitato da questi studi, nella letteratura simulazionista tre sono i processi cognitivi oggetto di particolare attenzione: "immaginare un desiderio"; "immaginare una decisione"; "immaginare una credenza". Senza entrare troppo nello specifico, in questa sede ci si limita a segnalare che dal dibattito emergerebbe che tutti e tre i processi sopra menzionati, quindi i meccanismi a fondamento dell'empatia, sarebbero sotto l'unico meccanismo cognitivo dell'immaginazione: alcuni hanno voluto nominarla *re-creative imagination*¹⁶¹ ed altri ancora *enacemement imagination*. Scrive, a tal proposito, Goldman:

What is distinctive to high-level simulation is the psychological mechanism [...] that produces it, the mechanism of imagination. This psychological system is capable of producing a wide variety of simulational events: simulated seeings (i.e., visual imagery) [...], simulated motor actions (motor imagery), simulated beliefs, [...] and so forth [...]. In producing simulational outputs, imagination does not operate all by itself [...]. For example, it recruits

¹⁵⁹ *IBIDEM*.

¹⁶⁰ STANFORD ENCYCLOPEDIA OF PHILOSOPHY, lemma *Folk Psychology as Mental Simulation*, consultabile online al seguente indirizzo

<https://stanford.library.sydney.edu.au/entries/folkpsych-simulation/> .

¹⁶¹ Cfr. GREGORY CURRIE E IAN RAVENSCROFT, *Recreative Minds: Imagination in Philosophy and Psychology*, Oxford University Press, Oxford, 2002.

*parts of the visual system to produce visual imagery [...]. Nonetheless, imagination 'takes the lead' in directing or controlling the other systems it enlists for its project*¹⁶².

Basandosi sugli studi neuroscientifici, secondo i quali sarebbe comune a tutti il meccanismo definito *default-network*, partecipe alle molteplici attività dell'immaginare, si spiegherebbe perché la sua attività potrebbe essere definita quale simulazione ad alto livello: consapevolezza, controllo volontario e "riuso", secondo il protocollo PROC, conferirebbero a quest'ultima la possibilità di spiegare le modalità per cui sarebbe congruo "immedesimarsi nei panni altrui" o, quantomeno, provare empatia da un punto di vista cognitivo.

Se, allora, questi sono i meccanismi di simulazione *high-level*, diversi sono quelli *low-level*, secondo cui l'azione, come nel primo caso, deve soddisfare due requisiti fondanti:

- a. Devono essere automatici, inconsci e pre-verbali;
- d. Devono soddisfare le peculiarità della scala PROC;

si richiamano, cioè, alle caratteristiche del *mirroring process*, secondo cui la simulazione può esser definita tale se e solo se l'azione segue dall'attivazione di un meccanismo specchio, o nella diretta visione (sistema specchio) o nella sua percezione dell'intenzione (sistema canonico).

Data questa premessa necessaria, l'incontro tra livelli di simulazione e "scala PROC" ha generato, come accennato, un vero e proprio sconvolgimento teorico all'interno della ST che potrebbe portare, secondo alcuni autori, al collasso della ST in favore della TT.

La prima preoccupazione nata dalla PROC la espone Heal, in uno scritto del 1994, sul modo in cui spiegare i processi cognitivi, partendo dai meccanismi di base che li sottintendono.

Si supponga, ad esempio, che la mente umana sia dotata di due meccanismi specifici:

1. Mec 1, che prende in considerazione credenze e desideri (*folk psychology*) per generare, a sua volta, delle decisioni (*output*);
2. Mec 2: che, partendo da Mec 1, genera decisioni *immaginate* quale *output*;

il primo meccanismo, cioè, produce decisioni concrete che si attueranno nella vita pratica; il secondo immaginerà soltanto le decisioni da prendere.

¹⁶² ALVIN GOLDMAN, *Replies to Perner and Brandl, Saxe, Vignemont, and Carruthers*, in "Philosophical Studies". 144(3): 2009, pp. 477-491.

Si consideri ora l'amico Frank, il quale, in un pomeriggio assolato, decide di entrare in un bar per bere una bibita fresca: Frank, cioè, desidera bere una birra (Mec 1) ed immagina di poterlo fare in un bar (Mec 2). Il problema che pone Heal è il seguente: la decisione immaginata di Mec 2 conta come decisione reale di compiere l'azione di Mec 1? È, cioè, la decisione immaginata una simulazione mentale di compiere l'azione stessa? Se la risposta sarà affermativa, allora, scrive Heal, la ST PROC ha seri problemi interni, poiché i due processi di simulazione, stando alla ST, dovrebbero essere distinti in base al principio che i *processi cognitivi* abbiano una natura distinta dai *meccanismi cognitivi*, in quanto, per problemi terminologici, l'espressione "simulazione mentale di" dovrebbe significare: "avere il medesimo meccanismo cerebrale che, nelle medesime condizioni, funziona nell'osservatore e nell'osservato nello stesso modo, a prescindere dalla cognizione"¹⁶³.

Ciò comporta, però, un'ulteriore seria difficoltà: come si può sostenere, infatti, che la simulazione mentale di una mela rossa di un agente, che vive nella metropoli New York nel 2019, possa essere la medesima di uno stesso agente che, o vissuto nel suburbio romano all'epoca di Augusto o che vivrà su Marte fra cento anni, possa essere sempre la stessa simulazione visiva ottenuta nel presente, ottenuta nel passato e ottenuta nel futuro? Perché ciò comporterebbe che nel momento in cui si abbia o l'immagine di una mela rossa o la visualizzazione della stessa si avrebbe, simultaneamente, la lettura della mente di tutta l'umanità passata, presente e futura della mela rossa. Una conseguenza che, come sarà facile intuire, solleva una certa perplessità in merito. Proprio su tale aspetto si concentra l'ultima perplessità che desta preoccupazione: come potrebbero le diverse versioni della ST aiutare a costruire una teoria del *mindreading* efficace in termini simulazionistici?

Scriva Gallagher:

[...] *given the large diversity of motives, beliefs, desires, and behaviours in the world, it is not clear how a simulation process ... can give me a reliable sense of what is going on in the other person's mind [...]*¹⁶⁴.

È complicato, infatti, capire come poter apportare una giustificazione epistemica all'attribuzione intenzionale di stati mentali altrui, anche perché la ST non si interroga, per sua natura, se tali attribuzioni possano valere quale forma di conoscenza, ma solo sul modo in

¹⁶³ JANE HEAL, *Simulation vs Theory-Theory: What is at Issue?*, in CHRISTOPHER PEACOCKE (A CURA DI), "Objectivity, Simulation, and the Unity of Consciousness: Current Issues in the Philosophy of Mind", Oxford University Press, Oxford, 1994, pp. 129–144. Articolo reperibile online al seguente indirizzo <https://www.thebritishacademy.ac.uk/pubs/proc/volumes/pba83.html>.

¹⁶⁴ SHAUN GALLAGHER, *Simulation Trouble*, in "Social Neuroscience", 2(3–4), 2007, p. 355.

cui queste ultime siano elaborate “meccanicamente” a “livello centrale”; d’altro canto, constatano alcuni autori, se è vero che capire le intenzioni altrui è un’operazione di una qualche natura che funziona, date le differenze fra le menti, il *mindreading* potrebbe non basarsi su meccanismi dati, ma su altro genere di spiegazione.

Sembrerebbe che le direzioni prese dai ST stiano andando verso una forma computazionale della teoria stessa, trasformando, parzialmente, il costrutto base da cui la ST era partita: meccanismi cognitivi, cioè, distinti dai processi cognitivi. Scrive, a tal proposito, Dennett:

[...] *If I make believe I am a suspension bridge and wonder what I will do when the wind blows, what “comes to my mind” in my make-believe state depends on my knowledge of physics. Why should my making believe I have your beliefs be any different? In both cases, knowledge of the imitated object is needed to drive the “simulation”, and the knowledge must be... something like a theory [...]*¹⁶⁵.

Anche laddove i simulazionisti dovessero rispondere all’accusa del collasso della propria teoria, sostenendo che i meccanismi cerebrali negli esseri umani sono tutti simili (per cui varrebbe la generalizzazione secondo cui “gli esseri umani sono psicologicamente simili”), tornerebbe, ancora una volta, il *focus* problematico della ST: si baserebbe, vale a dire, su un concetto tipico assunto della TT.

Nonostante l’ampio dibattito in merito, anche nella ST, così come nella TT, si possono individuare quattro aree di ricerca prioritarie che caratterizzano ed indirizzano le riflessioni in merito:

1. La prima è relativa alla *falsa credenza*.

Riprendendo in parte Gallagher, se è vero che l’attribuzione intenzionale a fondamento del *mindreading* funziona nella maggior parte delle volte, è altrettanto vero che per una stessa stragrande maggioranza l’osservatore ha necessità di “aggiustare” le proprie credenze sull’altro. Ciò si spiegherebbe proprio per l’utilizzo di quel modello *egocentrico* o, detto altrimenti, *in prima persona*, che i teorici della ST sposano: saremmo, cioè, sempre inclini ad errori egocentrici, ossia ad attribuire all’altro ciò che noi stessi, in quel momento ed in quella situazione, proveremmo, cadendo nel cosiddetto “pregiudizio di conoscenza”; daremmo per scontato, cioè, che l’altro, diverso da sé, sappia, provi ciò che noi sappiamo e proviamo. Non a caso, come lo stesso Goldman sostiene, l’accuratezza dell’attribuzione mentale ha successo

¹⁶⁵ DANIEL DENNETT, *The Intentional Stance*, MIT Press, Massachusetts, 1987, pp. 100-101.

solo se “mettiamo da parte noi stessi”, mettendo “in quarantena” il nostro Io, sostituendo i propri stati mentali con quelli altrui ed “aggiustando” proprie attribuzioni e credenze secondo credenze, desideri e volizioni dell’altro. Un cambio di prospettiva che, come lo stesso Goldman afferma, non è immune da “fatica”, poiché i propri “stati mentali genuini”, come li definisce, tendono a ripresentarsi nel processo di *mindreading* con una certa esigenza. Si comprendono, allora, le parole di Gallagher a riguardo:

[...] *if I project the results of my own simulation onto the other, I understand only myself in that other's situation, but I don't understand the other [...]*¹⁶⁶.

2. La seconda area di interesse è quella relativa allo sviluppo psico-cognitivo dei bambini nell’infanzia (la *Priority of self or other ascription*) nei suoi aspetti intersoggettivi.

Recenti indirizzi di studio stanno facendo emergere che in questa particolare età evolutiva le cosiddette capacità empatiche si presenterebbero man mano che il bambino cresce. Ciò ha portato molti teorici della ST a coadiuvare tali studi, poiché potrebbero, un giorno, “provare” che, effettivamente, l’attribuzione intenzionale è un meccanismo neuro-biologico che, una volta “maturato”, consentirebbe, di conseguenza, lo sviluppo dei successivi processi cognitivi. Alcuni esperimenti, però, metterebbero in evidenza che, in realtà, il riconoscimento degli stati mentali altrui non si baserebbe su un meccanismo o su un’inferenza analogica, incrementando, ulteriormente, dibattiti e nuove proposte o rivisitazioni della stessa ST¹⁶⁷.

3. La terza area di interesse concerne l’incessante incremento degli sviluppi neuroscientifici in merito al “funzionamento cerebrale” (*Neural Simulation*).

È a tutti noto che il laboratorio di Parma, negli anni ’90, abbia rappresentato una spinta propulsiva a livello mondiale per i progressivi sviluppi; se tali scoperte dovessero, un giorno, essere “assemblate” in una “teoria della mente” a carattere prettamente neuroscientifico, comporterà che i teorici della ST avrebbero una convincente e solida base empirica per poter sostenere la loro posizione.

4. La quarta ed ultima area di interesse concerne la “Impenetrabilità Cognitiva” (*Cognitive impenetrability*).

¹⁶⁶ SHAUN GALLAGHER, *Simulation Trouble*, op. cit., p. 355.

¹⁶⁷ ALISON GOPNIK ED HENRY WELLMAN, *Why the Child's Theory of Mind Really Is a Theory*, in “Mind and Language”, n. 7, 1992, pp. 145-71.

Come accennato nel punto 3, risulta ad oggi difficile condividere che tutti i processi mentali siano di stampo trascendentale: le molteplici teorie sulla mente, vale a dire, pongono seri dubbi, non tanto sulla naturalizzazione cognitiva (secondo un certo filone), ma sulla “frenesia intellettiva”, scatenatasi negli ultimi decenni, che il *mindreading* sia spiegabile, attualmente, con le poche scoperte scientifiche operate sul cervello.

Da come si evince dall’intero discorso, il “problema empatia” resta, tuttora, lontano da una sua possibile “risoluzione”.

Capitolo IV

Fra mindreading e metacognition: un rapporto controverso

4.1 La posizione di Peter Carruthers: *Mindreading is prior*

Alla luce della consapevolezza raggiunta finora grazie al percorso delineato, dovrebbe esser divenuto evidente che il problema principe dell'*empatia cognitiva* si basa proprio sulla rete di processi a cascata che essa metterebbe in moto affinché siano rispettati i passaggi relativi a quel “rivivere in sé gli stati emotivi altrui”; i quali renderebbero possibile l'intersoggettività *tout court*: “entrare nella mente dell'altro” per acquisirne una conoscenza “epistemica”; “riproporla in sé”; esperire, cognitivamente e sensorialmente, la medesima emozione dell'altrui soggettività.

In tale contesto, si è anche avuto modo di capire come nel corso degli anni, le due principali risposte al dilemma sul modo in cui possa essere possibile siano state fornite dai paradigmi della *Theory Theory* e della *Simulation Theory*. Ciò ha comportato, se si vuole, lo slittamento del *focus* del problema (il rapporto fra la lettura della mente del Tu – *third-person mindreading* – e lo sforzo cognitivo dell'osservatore per avviare il medesimo processo del *mindreading* – *first-person metacognition* –), tralasciando quello sull'eventuale rapporto fra il *mindreading* e *metacognition* e sulla stessa natura del *mindreading*.

Gli interrogativi che si stanno succedendo, infatti, sono inerenti al tipo di relazione sussistente fra il primo ed il secondo processo, ponendo in essere il dubbio sull'eventuale esistenza di una differenza (nel caso, di che tipo?) fra l'accesso alle altre menti e l'accesso rivolto verso se stessi; se non si comprende il *quid* della lettura della mente sarà altrettanto complicato capirne le modalità.

A tal proposito, il filosofo statunitense Peter Carruthers ha pubblicato, nel 2009, un articolo che resta, tuttoggi, oggetto di dibattito: *How we know our minds: The relationship between mindreading and metacognition*¹⁶⁸ in cui, oltre a fornire un sunto delle principali posizioni che si sono succedute sulla relazione in questione, ribadisce la sua posizione personale.

È da far notare che secondo alcuni lettori delle tesi di Carruthers (Daniel Dennett è uno fra questi) il merito del filosofo statunitense si trova, soprattutto, nelle argomentazioni apportate a sostegno di un particolare modello di spiegazione della relazione in questione in grado, forse, di rendere più chiaro il tipo di relazione fra lettura della mente e cognizione sulla cognizione.

A tal proposito, nel rapporto *mindreading/metacognition* diversi protagonisti al dibattito hanno imputato a Carruthers di non aver eliminato, progressivamente, il problema dell'accesso dell'Io a se stesso, poiché ciò comporterebbe non più una relazione fra lettura degli stati mentali altrui e le capacità cognitive in merito, ma allo scivolamento del problema del modo in cui l'Io sia in grado di porsi, a se stesso, quale oggetto di studio e di interrogazione; implicando, inevitabilmente, un'alienazione sostanziale di stampo dualistico del problema in oggetto. In che modo, infatti, dar conto di un Io che studi se stesso, esulando dall'eventuale scissione "scientifica" che dovrebbe conseguirne, coinvolgendo, così, il problema della natura del mentale? Non è un caso, allora, che molti fautori della posizione avanzata da Carruthers abbiano suggerito allo stesso di concentrarsi soltanto sul rapporto Io-altro. Infatti, il modo in cui il sistema *mindreading is prior* è stato strutturato dallo stesso filosofo statunitense lascerebbe trasparire una distinzione dei meccanismi del Sé (uno per gli stati di lettura della mente, l'altro per l'interpretazione degli stati metacognitivi) quali *res* distinte di cui una già inserita nel processo e l'altra "da inserirsi" in un secondo momento¹⁶⁹.

Gli esseri umani, esordisce Carruthers, sono degli "accaniti lettori della mente" (*inveterate mindreaders*¹⁷⁰), poiché, quasi inconsapevolmente, tentano di attribuire all'altro credenze, pensieri, stati o volizioni esperite in quel momento o condivise come parte integrante del proprio *modus agendi*, avviando un processo di *metarepresentations* (ossia *representations of representational states*) nel quale l'oggetto di studio diviene l'altro da sé (terza persona). Laddove, al contrario, tale sforzo cognitivo (riflessivo) dovesse partire da sé (prima persona) e, quindi, la capacità di "lettura della mente" si rivolgesse non più verso l'esterno, ma verso se stessi, l'individuo sarebbe il protagonista del processo di *metacognition*, ossia di una

¹⁶⁸ Cfr. PETER CARRUTHERS, *How we know our minds: The relationship between mindreading and metacognition*, in "Behavioral and Brain Sciences", 32, 2009, pp. 121-138.

¹⁶⁹ Cfr. DANIEL DENNETT E BRYCE HUEBNER, *Banishing 'I' and 'we' from accounts of metacognition*, in "Behavioral and Brain Sciences", 32, 2009, pp. 148-149.

¹⁷⁰ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 121.

cognizione (sforzo o processo di pensiero) sullo stesso processo di cognizione che porterebbe, inevitabilmente, alla cognizione stessa (il *meta* della parola “meta-cognizione”):

But human mental life, too, is richly metarepresentational, containing frequent attributions of mental states to ourselves. This sort of first person metarepresentation is generally referred to as ‘metacognition’¹⁷¹.

All’interno del quadro concettuale odierno, diverse sono state le risposte fornite sull’eventuale rapporto che, come detto, legherebbe l’empatia alla comprensione dell’altro ed in questo effervescente clima culturale, che promette nuovi sviluppi del binomio in questione, Carruthers individua quattro modelli generali (ognuno con proprie peculiarità e limiti), utili ai fini della sua trattazione, per giungere, poi, all’esplicitazione dell’uso epistemico tipico dell’empatia non solo per la “conoscenza” dell’altro, ma anche della propria. A tal fine, seguendo la strada tracciata dalle riflessioni del filosofo statunitense, ci si accinge ad illustrare quei modelli di spiegazione del rapporto in oggetto (*Model 1: Two mechanism, two modes of access; Model 2: One mechanism, two modes of access; Model 3: Metacognition is prior; Model 4: Mindreading is prior*) che gli consentiranno una enunciazione *chiara e distinta* del modo in cui lui stesso concepisce il binomio *mindreading-metacognition*.

4.1.2 *Model 1: Two mechanism, two modes of access*

La prima possibilità di spiegazione, che il filosofo statunitense “etichetta” come *Model 1: Two mechanisms*, è fornita dall’ipotesi di Nichols e Stich¹⁷² secondo cui l’*architettura della mente* (il cui modello è, oramai, ampiamente adottato sia dalle scienze cognitive sia dalla filosofia) “conterrebbe” due tipi di stati (credenze e desideri), i quali non solo differirebbero dal punto di vista funzionale e causale, ma, soprattutto, dalle modalità di interazione fra loro stessi ed altri “modelli” della mente.

Alcuni desideri, ad esempio, saranno causati da percezioni corporali derivante dai sensi (ad esempio, il desiderio di bere una bevanda fresca); altri da “desideri strumentali” o “sotto-obiettivi” che, tramite un ragionamento deduttivo, forniranno un ragionamento pratico o di

¹⁷¹ *IVI*.

¹⁷² Cfr. SHAUN NICHOLS E STEPHEN P. STICH, *Mindreading. An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*, Oxford University Press, Oxford, 2003.

pianificazione futura del comportamento da adottare. La proposta di Nichols e Stich (quindi la prima delle due assunzioni oggetto di critica da parte di Carruthers), sta nel fatto che, condividendo loro – come si è già intuito – la *teoria modulare della mente* di Jerry Fodor – i sistemi di ragionamento e pianificazione pratica non potranno avere solo “il compito” di generare meri “sotto-obiettivi”, ma dovranno essere in grado di determinare su quale struttura di “obiettivi” e di “sotto-obiettivi” debba essere “indirizzato” l’*input* oltre a *pianificare, progettare* l’azione da eseguire. Solo in questo secondo momento lo stimolo cognitivo sarà trasmesso al sistema di monitoraggio del corpo, che avvierà, in seguito, il controllo sulla sua realizzazione motoria.

Se tale spiegazione concerne il *mindreading*, altro varrà per la *metacognition* la quale (ed ecco la seconda assunzione criticata da Carruthers) consta, comunque, di due o più meccanismi di “monitoraggio di sé”, distinti gli uni dagli altri: un primo, ad esempio, fornirà conoscenza sui propri stati esperenziali, partendo da percezioni corporee; un secondo fornirà conoscenza sui propri atteggiamenti proposizionali. Un meccanismo, quindi, sarà dedito alla lettura della mente; un altro alla cognizione di sé. Rifacendosi ai recenti studi su pazienti schizofrenici ed autistici, che hanno contribuito ad una chiarificazione concettuale del modello proposto da Nichols e Stich, essi sostengono che

[...] *should be cases of people who can attribute mental states to others successfully but who have difficulty in attributing mental states to themselves, as well as instances of people who maintain reliable access to their own mental states while losing their capacity to attribute such states to other people*¹⁷³.

Ciò suggerisce ai due filosofi statunitensi una sorta di “lista” dei punti essenziali che una qualsiasi teoria sulla lettura della mente dovrebbe trattare; in particolar modo:

- *Normal adults readily attribute large numbers of beliefs to other people, even when they have no apparent evidence. Moreover, most of these attributions are correct;*
- *Normal adults are also very good at predicting the inferences of other people;*
- *Though adults are good at attributing some perceptual states, their ability is limited; they are quite bad at attributing perceptual states that result from perceptual illusions;*

¹⁷³ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 122.

- *Though adults are good at predicting some decisions, their ability to predict decisions is restricted; in a wide range of contexts, adults are quite bad at predicting the decisions that people will make.*
- *For normal children, though not for children with autism, there is a significant leap in mindreading abilities after the third birthday;*
- *Desire attribution precedes belief attribution in young children;*
- *Autistic children apparently have some capacity for desire attribution despite their difficulties with belief attribution;*
- *Children can predict behaviour effectively before they exhibit understanding of belief;*
- *Both pretend play and mindreading are deficient in autistic children¹⁷⁴.*

4.1.3. Model 2: One mechanism, two modes of access

In aperta opposizione all'impianto di Nichols e Stich, il secondo modello di interazione fra *mindreading* e *metacognition*, denominato da Carruthers *Model 2: One mechanism, two modes of access*, scientificamente avrebbe un vantaggio significativo per ciò che concerne la "legge della parsimonia".

Per spiegare, infatti, la conoscenza degli altri e la conoscenza di sé, tale modello si farebbe ricorso ad una sola facoltà mentale e non a due distinte, poiché la conoscenza che si avrebbe degli altri o di sé "poggerebbe" su una sola capacità mentale anche se il tipo di "apprendimento" conseguente sarebbe sempre derivato dalle percezioni sensoriali (in senso di *corporee*) e da se stessi (autopercezione). Esemplificandone la *ratio*, quest'ultima renderebbe evidente l'esistenza di una sola facoltà metarappresentazionale (quindi, *one mechanism*), che si distinguerebbe, al suo interno, in due diverse modalità di accesso agli stati mentali (*two modes of access*): usando due diversi canali di trasmissione degli *input*, il meccanismo metarappresentazionale si occuperebbe sia della modalità basata sulla percezione sia sulla percezione introspettiva. Anche qui, come in Nichols e Stich, si spiegherebbero casi di evidente dissociazione fra le due modalità di accesso all'unico meccanismo (e non ai due, ipotizzato dai primi), con i sintomi che i pazienti autistici e schizofrenici presenterebbero.

¹⁷⁴ SHAUN NICHOLS E STEPHEN P. STICH, *Mindreading*, op. cit., p. 78. Sulle posizioni contrarie a questo primo modello non è il caso di soffermarsi in questa sede, poiché si rischierebbe di uscire fuori tema, ma, certamente, le alternative a questo modello illustrate di seguito forniranno al lettore una panoramica degli argomenti in discussione.

Dalle parole di Carruthers si coglie una certa ironia sul modello riportato, incapace, a suo avviso, di dar conto, sostanzialmente, dell'apporto epistemico (eventuale) dell'empatia al soggetto che la esperisce:

[...] *there should be cases in which self-knowledge is compromised (because the introspective inputs to the metarepresentational faculty have been disrupted), whereas other-knowledge is intact (because the faculty itself remains undamaged and still has access to perceptual input). And it predicts that there should be cases where both self-knowledge and other-knowledge are compromised, by virtue of damage to the metarepresentational faculty itself [...] But there should be no cases where other-knowledge is damaged while self-knowledge is left intact, except by virtue of massive multi-modal perceptual failure*¹⁷⁵.

4.1.4 Model 3: Metacognition is prior

Il terzo modello, il *Model 3: Metacognition is prior*, mira a sottolineare che l'aspetto della metacognizione è precedente, prioritario, rispetto a quello del *mindreading* per cui ciò significherebbe che l'attribuzione degli stati mentali all'altro dipenderebbe da un atto introspettivo sui propri stati mentali, reso possibile attraverso i meccanismi di simulazione. Non è un caso, infatti, che il maggior fautore di tale modello sia proprio Goldman, il quale, dagli anni '90 ad oggi, ha in parte rivisto o abbandonato la sua posizione iniziale. Ciò lo ha portato ad essere oggetto di forti critiche per un articolo del 1993 in cui, Goldman stesso, sostenne che l'accesso ai nostri atteggiamenti proposizionali sarebbe stato "mediato" dalla consapevolezza degli stessi sentimenti fenomenici che contraddistinguerebbero gli atteggiamenti proposizionali¹⁷⁶.

Recentemente, mantenendo, comunque, la posizione che l'introspezione utilizzerebbe un codice innato del linguaggio del pensiero, secondo le caratteristiche neurali che lo renderebbero possibile, Goldman continua a sostenere che il processo di *mindreading* sia possibile grazie ad un atto immaginativo, introspettivo, che, proiettando sull'altro credenze e volizioni, simuli in se stesso ciò che penserebbe, crederebbe, desidererebbe l'altro da sé. In questo senso, quindi, la metacognizione diviene primaria rispetto alla lettura della mente, poiché quest'ultima non sarebbe possibile senza l'introspezione.

¹⁷⁵ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 122.

¹⁷⁶ Cfr. ALVIN GOLDMAN, *The psychology of folk psychology*, in "Behavioral and Brain Sciences", Vol. 16, Issue 1, 1993, pp. 15-28.

Sia il *Modello 2* sia il *Modello 3* concordano nel sostenere che esistono casi in cui le due capacità mentali sono danneggiate (con la precisazione che ciò in Goldman si verifica ogni qualvolta la capacità introspettiva è distrutta indipendentemente dalla funzionalità di quella del *mindreading*, proprio perché – come rende evidente il modello stesso – quest’ultima è subordinata al funzionamento della prima), ma discordano, di contro, sulle spiegazioni fornite a casi di presenza di *metacognition* ed assenza di *mindreading*. Ciò, per il filosofo statunitense, sarebbe possibile a causa della relativa distruzione dei processi di simulazione che renderebbero possibile la lettura della mente.

4.1.5 *Model 4: Mindreading is prior*

L’ultimo modello, quello condiviso e rivisitato da Carruthers, è il *Model 4: Mindreading is prior*, che si contrappone a quanto sostenuto da Goldman e da tanti come lui, che, sposando la *Simulation Theory* e, quindi, in generale il concetto di *simulazione*, condividono la priorità della *metacognition* sui meccanismi che renderebbero operativa la “lettura della mente”.

Da come si evince dalla stessa “etichetta”, nel *Modello 4* la prima è il risultato della trasformazione delle proprie capacità intellettuali rivolte su di sé:

*Instead of mindreading being grounded in metacognition, it maintains that metacognition is merely the result of us turning our mindreading capacities upon ourselves*¹⁷⁷.

Sono diversi gli autori che hanno contribuito o stanno contribuendo alla continua precisazione di aspetti o problemi rimasti irrisolti in questo modello (si pensi, ad esempio, a quanto sostenuto da Daniel Wegner, Michael Gazzaniga, Alison Gopnik e molti altri) e l’intento di Carruthers si situa su questa scia, avviando anche una distinzione fra *introspezione* e *discorso interiore* (*inner speech*) che porterà, nel prosieguo, all’introduzione del concetto di *confabulation*.

Come per il *Modello 3*, anche per il quarto esiste una sola facoltà coinvolta in entrambe le forme di attività della lettura della mente e della cognizione su se stessi; tanto che, usando, essenzialmente, gli stessi *input* ed attraverso una manipolazione di quelli percettivi e comportamentali, si giungerebbe, secondo Carruthers, a poter indurre i soggetti a

¹⁷⁷ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 123.

“confabulare” sulle attribuzioni degli stati mentali a se stessi. In tal modo, si avvierebbe un falso processo interpretativo:

However, the account also predicts that it should be possible to induce subjects to confabulate attributions of mental states to themselves by manipulating perceptual and behavioral cues in such a way as to provide misleading input to the self-interpretation process (just as subjects can be misled in their interpretation of others)¹⁷⁸.

Stando al modello *Mindreading is prior*, infatti, un processo di attribuzione a se stessi di stati mentali quali credenze o volizioni non può, per la medesima natura del modello, consentire che esista un’abilità distinta dalla lettura della mente, chiamata *introspezione*, perché ciò significherebbe smentire il modello stesso che prevede, come detto, una sola facoltà in cui l’applicazione pratica su se stessi comporta una comprensione dei molteplici *input* sensoriali cui l’individuo è soggetto.

Per questo motivo, sostiene Carruthers, non si può accettare la versione proposta da Alison Gopnik, la quale, nonostante sposti le medesime premesse, finisce per dare un’interpretazione del *mindreading* “dualista”: nel momento in cui l’attribuzione di tutti gli stati mentali a sé sia basata o su una teoria di fondo o su alcune assunzioni di base, le quali hanno tutte le caratteristiche per esser definite “interpretative”, si finisce, inevitabilmente, in una spiegazione duale del *mindreading*.

Ci si baserebbe su *input* sensoriali e comportamentali oggetto di fortissime critiche.

Come sostengono, ad esempio Nichols e Stich¹⁷⁹ o lo stesso Goldman¹⁸⁰, non è necessario per alcuno di noi essere in movimento per capire che cosa si stia pensando o che cosa si stia pianificando, poiché sono “consapevolezze” che si possiedono pur seduti, comodamente, in poltrona; in che modo sarebbe possibile, allora, parlare di interpretazione anziché introspezione?

Il *mindreading*, infatti, dovrebbe interpretare.

Altre varianti del *Modello 4* sono quelle di Wegner¹⁸¹ o di Dan Spencer e Deirdre Wilson¹⁸² alle quali Carruthers imputa l’incapacità di stabilire un confine chiaro e nitido tra l’accesso introspettivo e quello interpretativo per gli atteggiamenti proposizionali da attribuire a se

¹⁷⁸ *IVI*.

¹⁷⁹ Cfr. SHAUN NICHOLS E STEPHEN P. STICH, *Mindreading*, op. cit.

¹⁸⁰ Cfr. ALVIN GOLDMAN, *Simulating Mind*, op.cit.

¹⁸¹ Cfr. DANIEL WEGNER, *The Illusion of Conscious Will*, The MIT PRESS, Cambridge, 2002.

¹⁸² Cfr. DAN SPENCER E DEIRDRE WILSON, *Pragmatics, Modularity and Mindreading*, in “Mind and Language”, Wiley, 17 (1), 2002, pp. 3-33.

stessi, lasciando quasi intuire che la metacognizione abbia o un ruolo primario rispetto alla lettura della mente o che essa sussista indipendentemente dalla seconda.

In entrambi i casi, la peculiarità del *Modello 4* sarebbe compromessa.

Una premessa metodologica importante è data dalla precisazione che ciò che Carruthers stesso sta delineando non è un profilo sul modo in cui si diventi *consapevole* degli stati percettivi o di come si “acceda alla coscienza”, bensì sul modo in cui se ne acquisisca conoscenza. In tale ottica, qualsiasi ricorso alla “coscienza” si dovrà sempre intendere quale consapevolezza dei fenomeni coscienziali e non un loro tentativo di spiegazione. Il problema della coscienza, vale a dire, resta fuori dall'intento dell'autore.

Se queste ed altre varianti incorrono in critiche o confutazioni, l'intento di Carruthers è quello, come già detto, di fornirne una in grado di “risolvere” tali perplessità, sostenendo che:

There is just a single metarepresentational faculty, which probably evolved in the first instance for purposes of mindreading (or so I shall argue in sect. 5). In order to do its work, it needs to have access to perceptions of the environment. For if it is to interpret the actions of others, it plainly requires access to perceptual representations of those actions. Indeed, I suggest that, like most other conceptual systems, the mindreading system can receive as input any sensory or quasi-sensory (e.g., imagistic or somatosensory) state that gets “globally broadcast” to all judgment-forming, memory-forming, desireforming, and decision-making systems¹⁸³.

Ai fini della lettura della mente, tale capacità deve poter avere accesso, in un primo momento, agli stimoli sensoriali propri ed all'ambiente circostante per fornire a se stessa il “materiale” sensorio da “trasmettersi globalmente” per poi, in un secondo momento, rendere tale “materiale” fruibile a tutti gli altri sistemi di giudizio, di formazione della memoria, di pianificazione e simili. Il sistema di *mindreading*, dovrebbe, successivamente, essere in grado di attribuire a se stesso tali percezioni *in un modo incapsulato (in an “encapsulated” way)* senza ausilio di ulteriori *input*:

Receiving as input a visual representation of a man bending over, for example, it should be capable of forming the judgment, “I am seeing a man bending over.” (At least, this should be possible provided the visual state in question has been partially conceptualized by other mental faculties, coming to the mindreading system with the concepts man and bending

¹⁸³ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 123.

*over already attached. I return to discuss the significance of this point shortly.) This is the way in which introspection of perceptual, imagistic, and somatosensory mental events is achieved, I suggest*¹⁸⁴.

Benché il sistema della lettura della mente abbia accesso agli stati percettivi (non nel senso che questi ultimi siano oggetto di analisi da parte della *coscienza* tale che quest'ultima possa partire dai dati sensori per poi giungere ad un livello di analisi *cosciente* da parte dell'Io), ciò non significa, precisa Carruthers, che esso abbia, *in primis*, l'accesso successivo e diretto ai meccanismi della formazione dei giudizi (il *decision making*), poiché l'attribuzione a se stessi delle varie attitudini proposizionali sarebbe il risultato di un processo veloce ed inconsapevole di interpretazione di sé.

Ciò non significa, *in secundis*, che siano solo il comportamento manifesto e le circostanze fisiche del soggetto a fornire le basi per l'interpretazione, perché anche percezioni, percezioni sonore, immagini visive, frasi utilizzate dal “linguaggio interiore”, stati emotivi possono contribuire ad avviare il processo interpretativo.

Tali precisazioni potrebbero ovviare alle difficoltà incontrate nella posizione di Gopnik (un'interpretazione “dualista” del meccanismo unico del *mindreading*) evidenziate, come detto, da Goldman, Nischols e Stich; potendo accedere a qualsiasi *input* percettivo visibile o “invisibile” (stati mentali, emotivi, “linguaggio del pensiero” ecc), il sistema *mindreading* utilizzerebbe questi ultimi per interpretare le proprie intenzioni, le quali saranno, poi, messe in atto concretamente oppure no. Un uomo comodamente seduto in salotto che pensi di andare verso lo scaffale della libreria per leggere quel particolare libro avrà, in tal modo, attivo il sistema di lettura della mente grazie agli *input* “interni”, senza ricorrere all'attesa che quel comportamento si concretizzi (alzarsi e prendere il libro), evitando qualsiasi interpretazione dualista della mente e del meccanismo stesso. Non vi è, cioè, alcuna ragione per pensare che la formulazione di un giudizio o di una credenza necessitino di un intervento introspettivo attraverso la metacognizione. Infatti, in qualsiasi caso, la formulazione di quel giudizio o di quella credenza sarà il risultato di un atto interpretativo dato dal funzionamento del sistema *mindreading* che, come detto, accede a qualsiasi tipo di *input* esterno ed “interno”; per cui, continua Carruthers, non ci si deve lasciare indurre in errore quando si risponde ad una domanda, recuperando nella memoria il concetto ricercato: questo processo non ha la natura di un giudizio metacognitivo (introspettivo), ma solo interpretativo, poiché è lo stesso meccanismo di funzionamento del *mindreading*.

¹⁸⁴ *IVI*.

Risulta controintuitivo, sostiene Carruthers, pensare che “capire se stessi” non sia un atto introspettivo, ma interpretativo, poiché intere tradizioni filosofiche si sono succedute nella determinazione della comprensione del soggetto, ponendo l’accento sugli aspetti che, oggi, si definirebbero della *metacognition*. C’è da sottolineare, però, che ciò evidenzia che l’uomo è pre-teoreticamente incline in modo deciso a pensare che nell’intersoggettività si vada verso l’attribuzione di credenze e volizioni all’altro attraverso interpretazioni; di contro ad un accesso immediato, “consapevole”, di comprensione di sé che avverrebbe con un atto introspettivo.

Al di là della posizione personale della scrivente, Carruthers sostiene che, neuroscientificamente, esistono casi riportati dalla letteratura dei “pazienti con il cervello diviso” che possano confermare quanto sostenuto. Difatti, come Gazzaniga evidenzia, i pazienti sottoposti a commisurotomia hanno consapevolezza dell’intervento subito e dei suoi effetti, ma, di fatto, non giungono mai a spiegare il loro comportamento come conseguenza dell’operazione. Frasi del tipo: “sto scegliendo, probabilmente, in tal modo perché ho gli emisferi cerebrali divisi chirurgicamente per cui se scelgo *a* sarà perché solo l’emisfero dx ha percepito l’*input* di contro a quello sinistro” sono assenti dalle esperienze degli *split brain*, quasi si volesse ricercare in tale mancanza la giustificazione “razionale” o scientifica del modello proposto da Carruthers stesso.

Al contrario, tali pazienti riferiscono sulle loro scelte senza intoppi o “interruzioni”, “dimostrando” che il loro senso di Sé (emisfero sinistro) sembra intatto dopo l’intervento. Come riporta lo stesso Gazzaniga:

*All the evidence indicates that separation of the hemispheres creates two independent spheres of consciousness within a single cranium, that is to say, within a single organism. This conclusion is disturbing to some people who view consciousness as an indivisible property of the human brain. It seems premature to others, who insist that the capacities revealed thus far for the right hemisphere are at the level of an automaton. There is, to be sure, hemispheric inequality in the present cases, but it may well be a characteristic of the individuals we have studied. It is entirely possible that if a human brain were divided in a very young person, both hemispheres could as a result separately and independently develop mental functions of a high order at the level attained only in the left hemisphere of normal individuals*¹⁸⁵.

Ciò, secondo Carruthers, a riprova del fatto che se i pazienti avessero consapevolezza o avessero un accesso diretto al sistema di introspezione degli *input* dovrebbero aver

¹⁸⁵ MICHAEL GAZZANIGA, *The Split Brain in Man*, in “Scientific American Inc”, 1967, p. 29. Articolo reperibile al link <https://pdfs.semanticscholar.org/d8dd/bd1721eb8b98fcf5119b5b39fcb91044a8d2.pdf>

consapevolezza anche della situazione anomala che vivono, riportando, casomai, pensieri o riflessioni “manchevoli” così come è “manchevole” la comunicazione diretta fra un emisfero e l’altro. Non essendo, però, mai accaduto, stando a quanto riportato sia negli studi pubblicati di Gazzaniga sia nei suoi personali resoconti, si evidenzia ancora una volta che l’accesso ai nostri giudizi o alle nostre decisioni è sempre introspettivo nonostante la propria personale incapacità a distinguere la natura dell’atto stesso; quindi

[...] *our inability to discriminate shows that we have no subjectively accessible reason to believe in the existence of introspection. So anyone who is wondering whether or not introspection is real should realize that they have no reason they can offer for thinking that it is, in advance of examining the evidence*¹⁸⁶.

Un’accusa che è stata rivolta a Carruthers è stata quella di fornire una spiegazione a metà, poiché sostenere la fattività di un modello anziché di un altro significa sposare, innanzitutto, una teoria della mente tale da inquadrare all’interno di un dato contesto (la teoria) il funzionamento dell’oggetto preso in esame (la mente) e di conseguenza anche il *mindreading is prior* non può sottrarsi a ciò. Innanzitutto, esso deve fornire un profilo dell’accesso agli stati mentali che gli agenti avrebbero in quanto dotati di mente e già le scelte potrebbero essere due: la prima, secondo cui lo stesso meccanismo di lettura della mente sia in grado di interpretare se stessa; la seconda, gli agenti avrebbero un accesso diretto (quindi introspettivo) ai loro stati mentali.

Entrambe le soluzioni, però, complicherebbero di molto il dibattito, poiché conseguenza della prima opzione sarebbe, necessariamente, l’analisi critica non solo dei casi in letteratura con dati neuroscientifici inclini a quel tipo di interpretazione di cui il modello si funge portavoce, ma anche dell’analisi di svariate prove (prese dai diversi ambiti disciplinari) che suffragherebbero l’ipotesi di partenza. La seconda conseguenza, non meno complicata della prima, comporterebbe che gli agenti siano *naturalmente* inclini ad attribuire a sé e agli altri, avendo l’onere di specificare, quindi, che cosa significhi quel “naturalmente” e quali teorie lo spieghino.

In un modo o in un altro, è indubbio, secondo Carruthers, che qualsiasi elemento si aggiunga alla “semplice e concisa” complessità computazionale della mente indurrà in errore. È, infatti, un aspetto familiare nella “Scienza” *tout court* condividere l’euristica di una spiegazione o di una teoria, andando verso una maggiore semplificazione dei fenomeni da spiegare per cui,

¹⁸⁶ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 127.

ceteris paribus, una volta fornita una spiegazione convincente dell'errata attribuzione della natura introspettiva delle proprie attitudini proposizionali, non si potrà non convenire sulla semplice spiegazione della mente secondo il modello *Mindreading is prior*, in cui un solo modulo (quello della lettura della mente) è postulato ed utilizzato per spiegare fenomeni che altri modelli utilizzano con un altro grande numero di complessità.

Volendo trarre una considerazione generale sulla differenza tra i modelli proposti, la diatriba – come visto – si gioca sulla relazione fra introspezione ed interpretazione; se con la prima espressione ci si riferisce a

*l'auto-osservazione interiore cioè l'osservazione che l'io fa dei propri stati interni*¹⁸⁷

con la seconda si intende

il risalire da un segno al suo significato [in cui, riprendendo Platone, si colgono, n.d.r.] *oggetti che, invece, pongono un problema perché 'colpiscono il senso con impressioni fra loro opposte'; 'interpretazioni' sono dette per l'appunto queste impressioni opposte* (Repubblica, 523b sgg)¹⁸⁸.

Da ciò ne consegue, logicamente, non solo l'ulteriore contrapposizione fra i dati disponibili all'una o all'altra capacità, ma se tali dati siano o meno accessibili anche alla coscienza. Non si può tacere, infatti, sull'interrogativo del ruolo che svolge o svolgerebbe quest'ultima nell'accesso introspettivo o interpretativo delle proprie attitudini proposizionali e quelle da attribuire all'altro da sé; poiché la condivisione dell'una o dell'altra posizione potrebbe esaltare, far collassare o, addirittura, fondere in un'unica teoria le posizioni di partenza che, originariamente, si fossero presentate come contrapposte ed inconciliabili.

A tal proposito, Carruthers sostiene:

Now, it does not follow, logically, from all this (and much more) data that there is no such thing as introspection for propositional attitudes. For there might be one set of such events to which we have introspective access while there is another set that we can't introspect; and hence, whenever our behavior is caused by attitudes drawn from the latter set, we are forced to self-interpret (and often to confabulate). What might be proposed, in effect,

¹⁸⁷ NICOLA ABBAGNANO, *Dizionario di Filosofia*, Utet, Torino, 2006, lemma "introspezione".

¹⁸⁸ IDEM, lemma "interpretazione".

*is that there is both a conscious and an unconscious mind. Judgments and decisions within the conscious mind are introspectable, whereas judgments and decisions within the unconscious mind can only be known (if at all) by turning our mindreading capacities upon ourselves. And just such a view seems to be endorsed by some of those who have been most prolific in demonstrating the reality of metacognitive attitude attribution via processes of interpretation and confabulation. Thus, both Wegner (2002) and Wilson (2002) allow that we do sometimes have introspective access to our (conscious) thoughts, even if much of the time our access to our own propositional attitudes is interpretative, and often confabulatory*¹⁸⁹.

Le riflessioni sul problema accesso cosciente/accesso non cosciente, quindi, non può essere ignorato, poiché, come detto, non è una conseguenza logica del ragionamento fin qui esposto che, necessariamente, il funzionamento del modulo *mindreading* interpelli l'uso della coscienza; potrebbe anche darsi, come già proposto da alcuni autori, che “la funzione coscienza” subentri per certi aspetti mentre per altri ne rimanga fuori. Esisterebbe, quindi, una “mente cosciente” che potrebbe interpretare decisioni e giudizi ed esisterebbe una “mente incosciente” che, introspettivamente, formulerebbe altre decisioni ed altri giudizi. Il problema nel problema, allora, diverrebbe capire in quali circostanze subentrerebbe l'una ed in quali l'altra, chiedendosi, di conseguenza, quale sia quest'architettura mentale che renderebbe fattibile tutto ciò, quale la relazione fra le due “menti” (consapevole ed inconsapevole) e quale il ruolo svolto dalla coscienza¹⁹⁰.

L'ipotesi che Carruthers insieme ad altri¹⁹¹ si sente di sostenere, poiché derivante da ipotesi speculative suffragate dall'interpretazione dei dati ricavati da pazienti schizofrenici, è quella secondo cui il modulo *mindreading* sarebbe costituito da due sub-meccanismi con funzioni diversi: il Sistema 1 (S1) ed il Sistema 2 (S2). Infatti, sostiene che:

This is that human reasoning processes may be divided into two very different types, often now referred to as “System 1” and “System 2. System 1 (which is really a set of systems, arranged in parallel) is fast, unconscious, hard to alter, universal to all thinkers, and evolutionarily ancient. System 2, in contrast, is slow and serial, characteristically conscious,

¹⁸⁹ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 132.

¹⁹⁰ Non sembra opportuno al fine del discorso *in itinere* aprire una lunga parentesi su che cosa si intenda per *coscienza* e, soprattutto, quali siano le posizioni odierne che, all'interno del contesto neuroetico, abbiano fornito dati a favore dell'una o dell'altra interpretazione. Ci si limita, in questa sede, a ricordare che riflettere su che cosa sia la coscienza non è mero esercizio speculativo, ma un problema pratico che urgerebbe una risposta che, probabilmente, non giungerà tanto velocemente.

¹⁹¹ Cfr. KEITH FRANKISH, *Mind and Supermind*, Cambridge University Press, Cambridge, 2007; Cfr. JONATHAN EVANS E DAVID OVER, *Rationality and Reasoning*, Psychology Press Publishers, East Sussex (UK), 1996.

malleable in its principles of operation, admits of significant variations between individuals, and is evolutionarily novel. And a number of authors have emphasized the important constitutive role played by imagery (especially inner speech) in the operations of System 2 (Carruthers 2009; Evans & Over 1996; Frankish 2004). Likewise, others have demonstrated the crucial role played by inner speech in the performance of tests of executive functioning (which are likely to implicate System 2), such as the Wisconsin Card Sorting Task (Baddeley et al. 2001). For when inner speech is suppressed by the need to shadow an irrelevant speech stream while performing the task, performance collapses¹⁹².

Affinché tale ipotesi abbia successo, lo stesso Carruthers tende a sottolineare che quel *inner speech* o gli eventi immaginari coscienti di ognuno giochino, praticamente, il giusto ruolo causale per l'attivazione sia del S1 sia, successivamente, del S2: una decisione, ad esempio, non solo non dovrebbe scaturire da alcun ragionamento pratico, ma dovrebbe azionare, direttamente, il funzionamento del S1 che, attivando il sistema motorio conseguente, porterà l'agente ad agire praticamente. Solo in un secondo momento, tramite l' "entrata in gioco" del meccanismo S2, l'agente stesso sarà in grado di *interpretare* il suo fare e darsi, quindi, una spiegazione *cosciente* del suo agire. Qualcosa di simile, sostiene Carruthers, avverrebbe anche per una pianificazione futura, poiché le intenzioni, che, apparentemente, sembrano dei limiti al ragionamento *tout court*, costituiscono la struttura base su cui si fonderà, poi, l'agire stesso. Non si deve pensare che, per quanto appena esposto, l'individuo "sia agito" anziché "agire"; lontano, quindi, da qualsiasi consapevolezza cosciente di sé e dalle percezioni in prima persona, poiché la *ratio*, sia dell'introduzione dei due sistemi sia della posizione stessa di Carruthers, si ritrova in una certa corrente di pensiero sul "problema coscienza" oggi fortemente condivisa e discussa, contemporaneamente.

Almeno per quanto la scrivente pensa di aver compreso, la tesi di fondo che sembrerebbe condivisa dal filosofo statunitense è quella secondo cui la coscienza "altro non sarebbe" che la risultante di una funzione olistica dei due emisferi cerebrali che, per "apparire", necessiterebbe di tempo. La spiegazione, cioè, di ciò che Gazzaniga definisce *the interpreter*¹⁹³, sarebbe la diretta conseguenza da una parte del S1 (rimanendo nella tesi di Carruthers), che (inconsapevole, veloce, universalmente condiviso, poiché di origini evolutive) avvierebbe la decisione di agire in un dato modo e solo successivamente (quando, cioè, la "funzione coscienza" "appaia" quale risultato di intensi processi cerebrali scambiatisi

¹⁹² PETER CARRUTHERS, *How we know our minds*, op. cit., p. 133.

¹⁹³ Per una spiegazione approfondita si rimanda al capitolo VI.

fra i due emisferi) subentrerebbe “a giochi fatti” per fornire una spiegazione – quindi, un’interpretazione – di ciò che, apparentemente, si sarebbe compiuto *prima* della coscienza stessa. L’illusione di agire consapevolmente sarebbe, a sua volta, la risultante di un posticipo nel futuro dell’azione intrapresa così da garantire al soggetto stesso il pieno controllo di sé e dell’azione medesima.

In tale contesto si colloca, allora, la spiegazione di Carruthers sul funzionamento dei due sistemi considerati e del ruolo svolto dal *mindreading*; a questo punto risulta anche più chiaro il perché il modulo della lettura della mente sia prioritario rispetto a quello della metacognizione: sostenere, infatti, o la precedenza del secondo o l’esistenza di uno parallelo significherebbe ribaltare, completamente, la spiegazione di stampo *naturale* della coscienza.

La *metacognition*, cioè, anticiperebbe, temporalmente, l’ingresso in gioco della funzione coscienza per poi dare a quei meccanismi inconsci ed involontari una spiegazione successiva; in tal modo, sarebbe inspiegabile la coscienza come risultante di una specifica e complessa attività cerebrale, dovendole donare, poi, una natura di stampo trascendente (?), cadendo, inevitabilmente, nel problema cartesiano delle due *res*.

Il modello *mindreading is prior*, secondo Carruthers, trova conferme non solo nei dati ricavati (come già detto) dai malati schizofrenici e dagli autistici, ma anche da ricerche sperimentali condotte da vari autori (Hurlburt ed Akhter¹⁹⁴) nelle quali si evidenzerebbe che l’attribuzione cosciente a sé ed agli altri di specifiche attitudini derivi, effettivamente, da quei “discorsi interiori” o confabulatori che il soggetto compie quando si trova in modalità cosciente; a riprova del fatto che il *mindreading* agisce inconsapevolmente, catturando tutti gli *input* sensori da trasmettere, poi, al sistema S2 e, quindi, all’interpretazione cosciente.

La posizione di Carruthers, come è facile immaginare, sta facendo discutere moltissimo sia quali *pro* sia quali *contro* per ciò che significherebbe per il discorso legato all’intersoggettività ed all’architettura mentale di ogni singolo. Quest’ultima si collocherebbe, laddove l’autore statunitense avesse ragione, in una spiegazione di stampo prettamente neuroscientifica sul funzionamento cerebrale, sposando la naturalizzazione del mentale con ciò che ne seguirebbe; se, di contro, si sbagliasse, bisognerebbe capire in che direzione spingerebbero tali spiegazioni ed in che modo si avvicinerrebbero, si allontanerebbero o rifiuterebbero l’ipotesi di una “mente fatta di soli neuroni e scambi biocellulari”.

Il discorso, vale a dire, resta aperto.

¹⁹⁴ Cfr. RUSSELL HURLBURT E SARAH AKHTER, *Consciousness and Cognition*, in “Elsevier Journal”, 17, 2008, pp. 1364-1374.

4.2 Pro Peter Carruthers

Il filosofo statunitense, nella sua intensa disamina sulla spiegazione della validità del suo modello, sostiene che l'esistenza del "modulo lettura della mente" non può che avere origini evolutive; il meccanismo di adattamento e di sopravvivenza che la causalità della *fitness* genetica ha trasmesso ai posteri ha fatto sì da sviluppare nell'uomo ed in altre specie viventi la capacità di "capire", di "afferrare" le intenzioni e volizioni altrui.

Gli studi sui primati non umani, a tal proposito, sono in progressivo aumento e si sta delineando una vera e propria corrente di pensiero che raccoglie dati, esperimenti, speculazioni e teorie sulla continuità cognitiva fra le specie viventi. Per ciò che concerne esplicitamente l'uomo, Carruthers, come detto, sostiene che la capacità metacognitiva non può che derivare da una "svolta evolutiva" che, sotto la pressione dell'evoluzione stessa, ha fatto sì o da incrementare una capacità latente del *mindreading* o da spingere verso una sua specifica settorializzazione settoriale di alcune capacità in grado di costruire e mantenere un'unitaria immagine di sé (come già diversi altri autori hanno ipotizzato). Infatti, egli sostiene che:

From the perspective of a "mindreading is prior" account, no separate story needs to be told. Since metacognition, on this view, results from turning one's mindreading capacities upon oneself, its emergence will be a by-product of the evolution of mindreading. (This isn't to say that metacognition might not have come under secondary selection thereafter, perhaps by virtue of helping to build and maintain a positive selfimage, as Wilson [2002] suggests)¹⁹⁵.

Si evince, quindi, che ciò che finora è stata definita *empatia cognitiva* potrebbe esser stata la risultante di un'emergenza evolutiva, così come potrebbe aver avuto il medesimo "destino" la stessa *metacognition*; tale strategia argomentativa è stata ripresa da alcuni autori¹⁹⁶, i quali, contrariamente agli intenti di Carruthers, sostengono che per le medesime ragioni si evincerebbe che il modello che spiegherebbe la lettura della mente non sarebbe quello sostenuto dal filosofo statunitense, ma quello del *metacognition is prior*.

È evidente che ciò significa ribaltare e stravolgere gli intenti originali: lo scopo di Carruthers, infatti, sembrerebbe essere una "naturalizzazione dell'empatia cognitiva" quale strumento di adattamento sociale di un membro di una specie in una comunità, onde evitare che qualsiasi

¹⁹⁵ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 128.

¹⁹⁶ Cfr. JUSTIN J. COUCHMAN, MARIANA V. C. COUNTINHO, MICHAEL J. BERAN, J. DAVID SMITH, *Metacognition is prior*, in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009.

spiegazione di stampo dualista possa essere interpellata per spiegare l'attribuzione di intenti e volizioni all'altro da sé. Sostenere, infatti, che, evolutivamente, la capacità del *mindreading* si sia potuta, progressivamente, "specializzare" per mantenere una percezione di sé quanto più stabile e duratura significa delegare all'intersoggettività una funzione evolutiva senza ricorrere alla *res cogitans* o ai suoi derivati.

Tale intento è tangibile nella medesima spiegazione che Carruthers opera nella precisazione che non esiste introspezione, ma interpretazione: solo in questo modo, come sostenuto, si può ricorrere ad un uso della "funzione coscienza" quale attività risultante dalla complessa interazione dei due emisferi cerebrali e non denotarla come chissà quale entità misteriosa o di altra natura; con ciò né Carruthers, credo, né la scrivente stanno prendendo posizione sul "problema coscienza".

A proposito di questo aspetto non secondario del *Modello 4*, alcuni autori, fra cui Jonathan Evans¹⁹⁷, sostengono che la posizione di Carruthers aiuterebbe a risolvere il problema dualista dei processi cognitivi: aspetti coscienti ed aspetti incoscienti della cognizione stessa. Evans, ad esempio, si sente di corroborare la tesi secondo cui il soggetto né abbia un accesso introspettivo ai propri stati coscienti e né sia consapevole del principio dell'ipotetica catena causale che "determinerà" il suo fare, ma che, solo in un secondo momento, ne diventerà cosciente; l'illusione, quindi, della priorità della metacognizione "risolverebbe", come detto, il relativo problema dualista del mentale. Benché l'intento operato da Carruthers sia pienamente condiviso da Evans (preso a simbolo di molti altri autori che condividono ciò che si sta per scrivere), Evans stesso (e con lui tutti coloro che si riconoscono nelle sue posizioni) non si spiega i motivi per cui il filosofo statunitense sia rimasto ancorato ai propositi di quegli stessi dualisti che tanto ha cercato di allontanare; si riferisce, cioè, all'idea secondo cui debbano esistere due sistemi cognitivi distinti (che Carruthers definisce "Sistema 1" e "Sistema 2"); i quali, genericamente, sono "etichettati" come "coscienti" e "non coscienti". Con i primi, si delinerebbero le forme di conoscenza "esplicite"; con le seconde quelle "implicite" e, stando all'interpretazione che Carruthers adduce di entrambi i sistemi, si potrebbe evincere, secondo Evans, un errore nel ruolo svolto dalla consapevolezza:

The problem as I see it (Evans 2008;2009) is that it is a mistake to use consciousness in the definition of the distinction between Systems 1 and 2. A more satisfactory definition is that the latter requires access to central working memory, whereas the former does not. For

¹⁹⁷ Cfr. JONATHAN ST. B. T. EVANS, *Introspection, confabulation and dual-process theory*, in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009.

*sure, this implies that something about a Type 2 process is conscious, as the contents of working memory tend to reflect in conscious experiences (Andrade 2001)*¹⁹⁸.

Onde evitare, infatti, inutili accuse sostanziali per una superficialità linguistica, sarebbe opportuno evitare qualsiasi ricorso ai termini “cosciente/non cosciente”, poiché ciò implicherebbe spiegare la natura delle abilità cognitive, la motivazione, il ragionamento non tralasciando, comunque, la natura stessa della coscienza. È a tal proposito, infatti, che la posizione di Carruthers potrebbe tornare utile per ciò che concerne, appunto, quelle stesse attitudini proposizionali in gioco quando si parla di intersoggettività: se fosse vero che l’abilità del *mindreading* fosse “implementata” naturalmente nel cervello, ciò spiegherebbe, di conseguenza, il modo in cui la mente interpreterebbe non solo l’altro in quanto tale, ma anche se stessa. Rimarrebbe, però, l’interrogativo su che cosa significhi, effettivamente, “confabulare” e come l’introspezione (ammesso che esista quale fenomeno) sia possibile.

Carruthers risponderebbe, come visto, riportando l’interpretazione che ha dato dell’esperimento di Hurlburt in cui i pensieri o le considerazioni riportate dai soggetti nei momenti più disparati della giornata al suono del *beep* del dispositivo avrebbero evidenziato proprio l’uso *inconsapevole* del *mindreading* nelle modalità note. In effetti, se è vero che diversi autori convengono con il filosofo statunitense che “il modulo lettura della mente” possa funzionare in modo *inconsapevole* e che solo in un secondo momento possa *far divenire cosciente* l’agente, è altrettanto vero che la spiegazione “furtiva” dello “spostamento all’indietro” della consapevolezza è vista quale espediente frettoloso di spiegazione. Dominic Murphy¹⁹⁹, ad esempio, è dell’avviso che, stando proprio a quest’aspetto, la posizione di Carruthers sembrerebbe un “disperato espediente” sul modo in cui un agente diventi consapevole di sé, tralasciando, di contro, che il sol fatto di “essere nel mondo ed interagire con esso” apra le porte a spiegazioni più “veridiche” dell’auto-consapevolezza. Murphy, cioè, sostiene che non solo si diviene consapevoli in una molteplicità di modi che Carruthers non ha considerato, ma che tale consapevolezza non significa, in ogni caso, *interpretare* se stessi:

When Hurlburt (1997) discusses the case of Donald, for instance, he does not show that Donald was interpreting himself at all when he was beeped. Donald noted, for instance, that his son had left the record player on again (Hurlburt 1997, p. 944). He reported attending to a fact about his environment. He was later brought to see the facts he noticed as evidence that

¹⁹⁸ *IDEM*, p. 143.

¹⁹⁹ Cfr. DOMINIC MURPHY, *Varieties of self-explanation*, in “Behavioral and Brain Sciences”, Vol. 32, Issue 2, 2009.

*he harbored unacknowledged anger toward his son. Donald went over his own transcript after beeping and interpreted his behaviour as he might anyone else's. This case helps Carruthers in one way, but it also shows something he misses about the beep cases. Donald did not report anything about his own mind or engage in any self-interpretation when he was beeped. Many of Hurlburt's subjects do not, when beeped, report anything that looks like introspection or any other self-examination. They report thinking about the world, not about themselves*²⁰⁰.

L'asserzione, cioè, che la deliberazione ordinaria debba essere accompagnata da un'attività coscienziale dell'agente sembrerebbe superflua e non necessaria, poiché spesso la credenza di un'auto-attribuzione è solo un ricordo e non un processo interpretativo. A tal proposito, sostiene:

*I can assert a belief that p via the same procedure that would I go through in order to assert that p. This might be the result of working out what I believe via self-mindreading. But in other cases, when I state my belief that p, I am just remembering that p is true. Interpretative evidence is not needed. We do not have to assume that any interpretation is going on at all. In other cases, I put myself in a position to assert a belief by wondering if it is true. This is the way to handle the case of the woman who wonders what car she will go home in; we do not need to think of her as interpreting herself at all. Rather, the beeper leads her to say where her attention is focused, and it is focused on the world, so that is what she talks about. And attending to the world is not introspection or self-interpretation, even if it lets you say what you are thinking about*²⁰¹.

Carruthers sembrerebbe rimanere schiavo dell'idea che la *metacognition* necessiti di un qualsiasi “dispositivo neurale” per poter funzionare, senza considerare che spesso lo “sguardo interiore” dell'agente non nasconde chissà quali realtà mistiche, ma, semplicemente, una considerevole capacità di deliberare sul mondo; con credenze o conoscenze che non sono isolate dal mondo stesso, ma concernono ciò che sappiamo o crediamo di sapere sulla realtà circostante.

²⁰⁰ *IDEM*, p. 156.

²⁰¹ *IVI*.

4.3 *Contra Carruthers*

Il lavoro di Carruthers è stato interpretato da molti quale operazione di “demistificazione” o di “ridimensionamento” dei risultati di alcuni esperimenti sulle capacità metacognitive degli animali non umani a favore di una lettura più “oggettiva” dei dati stessi. L’intento di Carruthers, così come di altri prima di lui, è stato, innanzitutto, tentare di dimostrare che la distinzione fra “metacognizione debole” e “metacognizione genuina” è, spesso, utilizzata con un vizio interpretativo a favore della tesi che si intende sostenere. Partendo, infatti, dagli studi *in itinere* sugli animali non umani, passando per l’esame dei dati prodotti da pazienti autistici e schizofrenici, il filosofo statunitense giunge a sostenere che la distinzione fra *mindreading* e *metacognition* non solo è dubbia, ma ridondante ai fini della spiegazione evolutiva delle capacità umane di lettura della mente.

Se questo ed altri propositi costituiscono dei “punti a favore” della sua opera è pur vero che diversi altri autori, come è facile immaginare, o si sono mostrati scettici sulla sua posizione o fortemente contrari non solamente sull’uso idoneo o non idoneo degli studi *in progress* delle abilità nei non umani, ma anche su tutta l’impostazione del discorso condotto. Sotto accusa, vale a dire, si trova l’interpretazione carruthersiana di ciò che, originariamente, sarebbe dovuta essere l’intersoggettività.

Per coloro che rientrano nella prima categoria di scettici, è interessante notare che le conclusioni evoluzionistiche cui giunge Carruthers sono messe in dubbio per ciò che concerne l’enfasi che l’autore statunitense impiega per spiegare la singola capacità metacognitiva, ponendola, però, in relazione con la ricchezza attitudinale che, proprio evolutivamente, sostiene essere in relazione o in continuità con gli animali non umani²⁰². Sarebbe, cioè, che lo stesso vizio interpretativo che quest’ultimo ha imputato nel suo articolo ad etologi, filosofi, neuroscienziati o psicologi che si siano occupati dell’eventuale esistenza della *metacognition* e del *mindreading* nelle specie non umane lo abbia commesso lui stesso nel momento in cui ha lasciato trasparire la capacità di lettura della mente quale “epifenomeno evolutivo”, senza apportare ulteriori prove a riguardo.

Senza smentire, ovviamente, l’evoluzione *tout court*, gli autori di questo disappunto a Carruthers sostengono che un approccio alternativo potrebbe essere abbandonare un vago “pregiudizio di specie”, per concentrarsi, come già suggerito da Gallaghers, su tutte quelle specie viventi dotate di “meccanismi integrati” che potrebbero portare ad una strutturazione

²⁰² Cfr. CAMERON BUCKNER, ADAM SHRIVER, STEPHEN CROWLEY E COLIN ALLEND, *How “weak” mindreaders inherited the earth*, in “Behavioral and Brain Sciences”, Vol. 32, Issue 2, 2009.

funzionale, da un punto di vista sinaptico, di una elaborazione delle capacità della metacognizione e della lettura della mente; seppur meno complesse di quelle umane. Abbandonando l'assunzione che il *mindreading* sia dipendente o in parte congiunto alla *metacognition*, si sarebbe “più liberi” di poter investigare se l'implementazione di meccanismi neurobiologici in maniera progressiva da un'ottica evolutivista possa spiegare la natura delle due capacità “superiori” dell'*Homo Sapiens Sapiens*.

A tal proposito, alcuni autori sostengono:

This perspective focuses us on the flexibility and adaptive significance of the evolved mechanisms which constitute the swarms, for a wide range of organisms in a variety of social environments (including humans in theirs). These suggestions are in the spirit of Dennett (1983), who advocated the usefulness of metarepresentational hypotheses in devising new experiments, accepting from the beginning that animals and humans will “pass some higher-order tests and fail others” (p. 349)²⁰³.

In tale impostazione, però, se si è intrapresa una diatriba fra i molteplici aspetti che concernono il *mentalizing*, giungendo, progressivamente, ad un livello di sofisticazione sempre più raffinata (o, forse, inutile!) la grande assente nel dibattito rimane proprio ciò che, paradossalmente, ha dato l'*incipit* al dibattito stesso: l'empatia cognitiva.

Non è un caso, infatti, che in un articolo di Philippe Rochat²⁰⁴, ciò che potrebbe sembrare un'annotazione *soft* a Carruthers ed ai più sul problema in questione, in realtà, si rivela, forse, l'accusa più devastante e più forte all'intero sistema: riproponendo l'attenzione su William James, che già nel 1872 si chiedeva se fossimo uomini o automi, lo psicologo statunitense rivendica che le abilità di lettura della mente e della metacognizione dovrebbero essere concettualizzate come “qualcosa di più” di meri processi razionali di attribuzioni di stati proposizionali ad altri.

Esistono, cioè, *sentimenti ed emozioni che ci rendono esseri senzienti ed incommensurabili alle macchine* [trad. mia]

²⁰³ *IBIDEM*, pp. 140-141.

²⁰⁴ Cfr. PHILIPPE ROCHAT, *Social-affective origins of mindreading and metacognition*, in “Behavioral and Brain Sciences”, Vol. 32, Issue 2, 2009.

(At the core of these issues, there is more than the rational attribution of mental states to the self or to others (theories of mind). There are also feelings and emotions, what make us sentient and incommensurate to machines²⁰⁵),

tanto che nel dominio sociale la “trascendenza di sé” si spiegherebbe come l’intera sfera emotiva messa in gioco quando si entra in contatto con l’altro, quando si intraprende con il Tu un qualsiasi tipo di relazione umana:

In the social domain, “self transcendency” corresponds first and foremost to the feeling of our relation to others, not just to a propositional attitude turned toward others or onto the self (metacognition)²⁰⁶.

Di questo “fenomeno spontaneo”, Rochat sostiene che le sue radici si radicano nella propensione “istintiva”, “naturale” a “sapere dove ci si trova nella relazione con l’altro”, quale consapevolezza, vale a dire, dello “scambio emotivo” fra sé ed il Tu; infatti, afferma:

The origins of a tendency to attribute attitudes (i.e., adopt a propositional attitude stance) toward others as well as toward the self are rooted in the basic propensity to know where we stand emotionally in relation to others: whether we feel affective distance or closeness, a sense of affiliation and approval, or on the contrary a sense of disapproval and rejection from others²⁰⁷.

In quest’ottica, prosegue, dovrebbero essere letti gli studi sugli infanti o sugli animali non umani che “mostrerebbero” che la natura delle capacità del *mindreading* o della *metacognition* derivano dall’urgenza di uno scambio intenzionale e comunicativo fra soggetti di stampo non razionale, ma ontologicamente connaturato all’emotività:

These facts demonstrate that at the roots of mindreading and metacognition, there is a complex pragmatics of intentional communicative exchanges that is not just rational, but ontologically affective and emotional²⁰⁸.

²⁰⁵ *IBIDEM*, p. 160.

²⁰⁶ *IVI*.

²⁰⁷ *IVI*.

²⁰⁸ *IVI*.

Uno scambio di difficile accertamento.

Il *mindreading*, infatti, dovrebbe fungere da sistema di attribuzione all'altro di credenze, desideri e volizioni e, simultaneamente, essere in grado di attribuire quelle stesse credenze, desideri e volizioni anche a se stessi; il problema nel problema sull'eventuale distinzione fra *mindreading/metacognition* mette in luce non solo che il sistema di lettura della mente non può avere un valore epistemico, ma, soprattutto, che l'eliminazione di un processo distinto della *metacognition* aggraverebbe solamente la spiegazione mentalistica.

Secondo Friedman e Petrashek²⁰⁹, ad esempio, che condividono tali appunti rivolti a Carruthers, le eventuali attribuzioni di stati intenzionali all'altro sono coerenti al contesto entro cui l'osservatore e l'osservato agiscono, per cui nessun valore conoscitivo si può abduire dal processo di *mindreading* e, quindi, di "lettura della mente".

A tal fine, sostengono:

One's intention to drink coffee says little about the world except perhaps that people sometimes drink coffee; and it says little about other people because relatively few share this intention at any time, meaning that it will seldom be useful to quickly extend this intention to others. So mindreading may not require access to such mental states. If the mindreading system lacks this access, it will also be lacking for metacognition²¹⁰.

Non è un caso, comunque, che, nella manchevolezza del discorso presentato da Carruthers, alcuni autori abbiano mostrato delle perplessità sul significato che il filosofo statunitense attribuisce alla parola "metacognizione"; soprattutto nel significato dello stesso suffisso "meta" a cui scuole di pensiero diverse attribuiscono altri intenti ed altri denotati; ciò significherebbe inficiare sia la stessa posizione di Carruthers sia la stessa riflessione sulle attitudini proposizionali.

Cominciando dal principio, il filosofo statunitense esordisce nel suo articolo che, probabilmente, le scienze cognitive sarebbero loro stesse incerte sul significato da attribuire all'espressione "metacognizione", da cui:

²⁰⁹ Cfr. ORI FRIEDMAN E ADAM R. PETRASHEK, *Non interpretative metacognition for true beliefs*, in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009, pp. 146-147.

²¹⁰ *IBIDEM*, p. 147.

*Generally the term is used, as it has been throughout this article, to mean cognition about one's own cognition. Metacognition, in this sense, is inherently higher-order, involving metarepresentations of one's own first-order cognitive processes as such*²¹¹

si potrebbe dedurre che la metacognizione implicherebbe, forse, necessariamente, un'abilità di ordine superiore, cioè intellettivamente impegnativa.

Ciò che al filosofo statunitense alcuni autori²¹² imputano, innanzitutto, è l'accusa di arbitrarietà della medesima definizione di *metacognition* nel dibattito delle scienze cognitive, poiché la definizione unanimemente riconosciuta come *standard*

*Imagine two components X and Y (where X and Y could be the same), related in such a way that state information flows from Y to X, and control information flows from X to Y. Component X is in a monitoring and control relationship with Y, and when Y is a cognitive component, we call this relationship metacognitive monitoring and control)*²¹³

è, oramai, entrata anche nel dibattito sulla relazione *mindreading/metacognition* e, soprattutto, da un punto di vista sostanziale, non sembra chiaro perché per parlare di *metacognition* si debba, necessariamente, invocare un *higher order level* di complessità cognitiva; ignorando, completamente, qualsiasi attenzione al processo anche nei suoi aspetti "base". In questo senso, abbandonando completamente la formale distinzione di Tarsky sulla differenza tra i meta-linguaggi per rappresentare il pensiero proposizionale quale semplificazione funzionale alla logica formale, Anderson e Perlis ribadiscono che le informazioni (gli *input* cui si riferisce Carruthers) sono usate e manipolate per l'attribuzione proposizionale non soltanto ad un livello cognitivo di impegno intellettuale (l'*higher order level*), bensì anche come mera guida al comportamento rispetto ai diversi *target* dell'ambiente in cui si agisce:

*Such monitoring and control is indeed (we maintain) cognition about cognition – is thus metacognition – without having to be/utilize higher-order representations of cognition as such*²¹⁴.

²¹¹ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 129.

²¹² Cfr. MICHAEL L. ANDERSON E DON PERLIS, *What puts the "meta" in metacognition?*, in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009, pp. 138-139.

²¹³ MICHAEL ANDERSON E TIM OATES, *A Review of Recent Research in Metareasoning and Metalearning*, in "AI Magazine", Vol. 28, n°1, Spring, 2007, p. 7.

Articolo reperibile online al seguente indirizzo

<https://www.aaai.org/ojs/index.php/aimagazine/article/view/2025>

²¹⁴ MICHAEL L. ANDERSON E DON PERLIS, *What puts the "meta" in metacognition?*, op. cit., p. 139.

Non è un caso, infatti, che a Carruthers sia imputato di non aver addotto prove di stampo evoluzionistico sulla diretta dipendenza della metacognizione dalla lettura della mente; impossibile, la prima, da ricondurre esclusivamente alla sola esperienza soggettiva.

Seppur si volesse sorvolare sulla definizione di *metacognition* da lui fornita e non rispondente al dibattito attuale nelle scienze cognitive, non si potrebbe tacere sulla semplificazione adottata a proposito della stessa metacognizione. Gli studi di Koriat *et alii*, infatti, mettono in luce l'intricato rapporto fra la "monitorazione" metacognitiva ed il controllo metacognitivo nel soggetto stesso.

Da una parte, quando il controllo di sé è guidato dagli obiettivi, il monitoraggio, l'osservazione degli effetti prodotti dall'agente in vista di quegli obiettivi controlla il processo metacognitivo così da migliorare, eventualmente, gli sforzi soggettivi per giungere al proprio fine attraverso un incremento sensorio dell'autopercezione di sentimenti e sensazioni di conoscenza. Dall'altra, quando l'azione viene eterodiretta dalla constatazione di dati, la medesima cognizione con l'influenza della sfera emotiva del soggetto tendono a rallentare i processi di controllo attraverso la percezione cosciente (da parte dell'individuo stesso) di *feedback* negativi per l'azione che sta svolgendo²¹⁵. Il *focus*, cioè, evidenzia l'intricata relazione di causa ed effetto tra esperienza soggettiva e comportamento manifesto, rimanendo, in Carruthers, ad un livello o completamente silente o completamente generico, che potrebbe, forse, ricordare il problema (al lettore già erudito su tali temi) solo nel momento in cui sostiene che la metacognizione potrebbe operare in un formato rappresentazionale differente.

La constatazione che la sfera emotiva non resta esclusa dal processo di metacognizione non può non far sorgere il sospetto che, forse, potrebbe rappresentare l'incertezza soggettiva nella guida al processo empirico di "azione nel mondo", guidando in un modo o in un altro il processo decisionale noetico senza coinvolgere, come sostiene Carruthers, il processo interpretativo concettuale²¹⁶.

Se da una parte, infatti, si rimanesse ad un livello generalissimo di adozione dell'ipotesi di Carruthers (secondo cui nell'architettura mentale la mente rimane non solo il *focus* di qualsivoglia attività, ma, ribadendo che azioni ed eventi siano, genericamente, un suo risultato) molti non potrebbero non convenire con Carruthers stesso; dall'altra, andando ad

²¹⁵ Cfr. ASHER KORIAT, HILIT MA'AYAN E RAVIT NUSSINSON, *The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior*, in "Journal of Experimental Psychology: General", 135(1), 2006, pp. 36-69.

²¹⁶ Cfr. JOËLLE PROUST, *Overlooking metacognitive experience*, in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009.

approfondire, strutturalmente, quell'architettura proposta, molti noterebbero che le assonanze e le similitudini con quanto già proposto da Skinner sono notevoli e sostanziali²¹⁷.

Un ulteriore problema riscontrato nell'elaborazione del *Modello 4* risulta essere una interpretazione dualistica dei meccanismi di controllo del *mindreading*: sembrerebbe, cioè, che ne esista uno per le azioni fisiche attraverso un *feedback* percettivo ed uno per le azioni mentali attraverso un *feedback* metacognitivo, quindi, interpretativo. Al di là delle constatazioni più ovvie che si possano esprimere contro la mancata unidirezionalità dell'architettura mentale, Proust fa notare a Carruthers che la meta-azione è funzionalmente simile alla meta-cognizione, nonostante un profilo evolutivo che potrebbe esser stato diverso. Tali mancanze metodologiche potrebbero, allora, evidenziare che si potrebbe giungere a considerazioni differenti rispetto a quelle proposte dal filosofo statunitense anche a proposito dell'utilizzo dei dati rinvenuti da pazienti schizofrenici o autistici²¹⁸.

Volendo addurre un parallelismo che potrebbe sembrare *ex abrupto* non pertinente all'oggetto di indagine, secondo alcuni autori l'apparente "rivoluzione" concettuale apportata da Carruthers nel sostenere che dall'abilità del *mindreading* si potrebbe "dedurre" quella della *metacognition* sembrerebbe riproporre la "rivoluzione concettuale" che già mezzo secolo fa ed oltre avanzò Skinner, rigettando le teorie e la metodologia di Watson.

Nello scritto *The Operational Analysis of Psychological Terms*²¹⁹, per l'appunto, il padre del Comportamentismo sosteneva, così come Carruthers, che i resoconti privati (*inner speech*) fossero il risultato di un uso linguistico formalizzato in un vocabolario unanimemente riconosciuto: l'atto metacognitivo, cioè, altro non sarebbe che una pratica pubblica della comunità verbale che, utilizzando una convenzione secondo cui l'uomo avrebbe un accesso privilegiato di prima persona a se stesso, il "racconto interiore" si spiegherebbe solo ed esclusivamente tramite un vocabolario di stampo comportamentista. In linea con Skinner, ritornando a Carruthers, l'*inner speech* si spiegherebbe come risposta a quell'insieme di *input* "interni" ed "esterni" derivanti dal funzionamento del sistema del *mindreading*, rifiutando, completamente, l'introspezione. Inoltre, la stessa suddivisione operata da Carruthers in *Sistema 1* e *Sistema 2* ricorda, esplicitamente, la suddivisione elaborata da Skinner come ulteriore testimonianza che nel panorama contemporaneo si sta sempre più convergendo verso

²¹⁷ Cfr. CHARLES CATANIA, *Cognitive science at fifty*, in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009, pp. 141-142.

²¹⁸ Cfr. PHILIP ROBBINS, *Guilt by dissociation: Why mindreading may not be prior to metacognition after all*, in "Behavioral and Brain Sciences" Vol. 32, Issue 2, 2009, pp. 159-160.

²¹⁹ Cfr. BORING SKINNER, *The Operational Analysis of Psychological Terms*, in "Behavioral and Brain Sciences", Vol. 7, Issue 4, 1984, pp. 547-553.

un felice spozalizio fra la tradizione delle scienze cognitive e della scuola psicologica a testimonianza che la Cultura è una.

Un altro aspetto fortemente criticato al sistema di Carruthers concerne la negazione o il rifiuto dell'atto introspettivo a favore della sola interpretazione, tanto che una delle critiche avanzate ha posto l'accento sulla non sempre chiara distinzione che il filosofo statunitense opera fra i concetti denotati dalle espressioni "confabulazione", "confidenza" ed "introspezione". Lo stesso, posto di fronte all'interrogativo di fornire una risposta a che cosa rappresenti la *confabulation* rimane generico, vago, facendo dipendere il significato dal tipo di approccio o dal tipo di interpretazione sottesa, lasciando nel vago il suo interlocutore o lettore²²⁰. A tale mancanza si aggiungerebbe quella riguardante la "confabulazione" sulla quale lascerebbe intendere che si potrebbe definire in tal modo quella classe di attribuzioni attitudinali rivolte verso sé attraverso il meccanismo del *mindreading*, generando, così, un processo "interpretativo" e non "introspettivo".

Della prima, Carruthers sostiene che è quel processo attraverso cui

*that accesses information about the subject's current circumstances, or the subject's current or recent behavior, as well as any other information about the subject's current or recent mental life*²²¹,

di conseguenza, come già detto, qualsiasi accesso di prima persona alle attitudini proposizionali rivolte a se stessi non possono che avere un'origine interpretativa. Ciò, però, apre un dilemma su che tipo di correlazione o di differenza possa esistere fra la confabulazione e l'interpretazione stessa, avendo fornito di quest'ultima una definizione troppo generica ed onnicomprensiva.

Tre sarebbero, a parere di Fiala e Nichols, le possibilità che si aprirebbero innanzi a Carruthers per uscire dall'*empasse* in cui si sarebbe arenato:

1. *Confabulations include all self-attributions that result from interpretation;*
2. *Confabulations include all false self-attributions that result from interpretation, and accurate interpretative self-ascriptions do not count as confabulatory;*
3. *Confabulations include only a proper subset of false selfattributions resulting from interpretation*²²².

²²⁰ BRIAN FIALA E SHAUN NICHOLS, *Confabulation, confidence and introspection* in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009.

²²¹ PETER CARRUTHERS, *How we know our minds*, op. cit., p. 123.

Ipotizzando che l'opzione 2 sintetizzi la posizione del filosofo statunitense, i due autori non condividono la rinuncia all'atto introspettivo, poiché le prove apportate a sostegno del rifiuto dell'*introspection* a favore dell'*interpretation* sembrerebbero aneddotiche, poco convincenti e pertinenti soltanto a pochi o rari casi della letteratura. L'immediatezza ed alta consapevolezza con cui si sarebbe inclini a pensare che qualsiasi processo di attribuzione a sé o agli altri possa essere interpretativo grazie alla strutturazione stessa dell'abilità del *mindreading*, si spiegherebbero, secondo Fiala e Nichols, solo attraverso l'ausilio di una confabulazione personale o atto di introspezione che ben si allontanerebbe da una mera interpretazione degli *input* ricevuti.

Attraverso lo studio di un'ampia letteratura a riguardo (si pensi, ad esempio, ai lavori di Nisbett e Wilson del 1977 a quelli più recenti di Thalia Wheatley del 2007, passando per Gazzaniga sui casi di "cervello diviso"), i due autori giungono a sostenere che in ognuna di queste ed altre ricerche quel senso di immediatezza ed alta consapevolezza (manifestazioni tipiche, secondo Carruthers) del processo interpretativo è completamente assente a riprova del fatto che la metacognizione è un processo di gran lunga più complicato che la spiegazione di un modello ad una o più uscite.

Non è un caso, infatti, che la stessa posizione avanzata a proposito dell'esistenza dei due sistemi (S1 e S2) in cui, si ricorda, non si avrebbe un accesso cosciente ai giudizi o alle decisioni se non "a dado tratto" dalla coscienza, non convince. Essendo due i sistemi di "organizzazione funzionale" della mente dovrebbero essere due i tipi di accesso consentiti: l'uno con accesso interpretativo a pensieri inconsci e l'altro ad accesso introspettivo a pensieri consci²²³.

Il rifiuto di una posizione "mista" in virtù della sola spiegazione *mindreading is prior* raccoglie la posizione di tanti, i quali sostengono che, seppur prezioso il contributo di Carruthers, in realtà sia troppo restrittivo per dar conto (come si diceva poco più su) sia della cognizione sia dell'intersoggettività, ovviando, in tal modo, a prender posizione in merito al dibattito sulla differenziazione e particolarizzazione sia della metacognizione sia della stessa interpretazione.

²²² BRIAN FIALA E SHAUN NICHOLS, *Confabulation, confidence and introspection* in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009, p. 144.

²²³ Cfr. KEITH FRANKISH, *How we know our conscious minds: Introspective access to conscious thoughts*, in "Behavioral and Brain Science", Vol. 32, Issue 2, 2009.

4.4 La risposta di Peter Carruthers

Essendo stato chiamato in causa più volte da diversi e variegati punti di vista, Peter Carruthers non avrebbe mai potuto, per deontologia professionale, tacere. D'altronde, come il bistrattato Socrate ci insegna dal 399 a. C. solo il *dialeghenstain* è fonte di conoscenza e di arricchimento.

Per quanto riguarda le obiezioni rivoltegli sulla natura della facoltà del *mindreading*, con la prudenza tipica di chi sa che non possono esistere giudizi *tranchant*, il filosofo statunitense conviene con tutti coloro che hanno assunto o hanno ribadito che non possa esservi una sola facoltà nell'architettura mentale per cui, contro coloro che sostengono il contrario, Carruthers ribadisce che il *mindreading system* non può essere il "solo blocco monolitico" di risoluzione dei problemi posti.

A tal proposito, afferma che:

I can accept that the mindreading faculty is not monolithic, but is actually a cluster of more specialized mechanisms working together in concert, somewhat as Buckner, Shriver, Crowley, & Allen (Buckner et al.) suggest. Indeed, this is what I actually believe, following Baron-Cohen (1995) and Nichols and Stich (2003)²²⁴.

L'appunto postogli, si rammenti, riguarda il fatto che il sistema di lettura della mente per funzionare debba aver accesso anche alle credenze dell'individuo per avere un materiale su cui poter lavorare, perciò l'attribuzione a se stessi di quelle credenze non può che derivare proprio da un atto interpretativo; il problema, cioè, di giustificare o di apportare ulteriori spiegazioni al riguardo sembrerebbe superfluo. Su tale assunto Carruthers non può esimersi dal rispondere, perché parlare della capacità di lettura della mente significa congegnare una spiegazione dell'architettura mentale sottostante e non, semplicemente, soffermarsi su uno o su un altro aspetto del singolo meccanismo esposto: non è un caso che, alla fine dell'articolo in questione, il filosofo statunitense sostenga che saranno le stesse scoperte neuroscientifiche a dirimere, un giorno, le controversie a testimonianza che, come ben il Socrate "platonizzato" mostrava e come più volte all'interno del seguente lavoro è stato ribadito, la Cultura è una.

Gran parte della cognizione ruota attorno "ad uno spazio di lavoro", che può risultare dal funzionamento di diversi meccanismi in parallelo e, perciò, soffermarsi sull'idea che possa

²²⁴ PETER CARRUTHERS, *Mindreading underlies metacognition*, in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009, p. 165.

esservi un solo processo computazionale significherebbe che il dialogo si sarebbe spostato su un automa o su una macchina e non di certo sul funzionamento della *mente* umana.

In tale ottica è da leggersi il ricorso di Carruthers all'introduzione dei due sistemi (S1 e S2), i quali, con caratteristiche e compiti distinti, potrebbero fornire una spiegazione quanto più pertinente "al vero". Ciò non significa, come Hernik, Fearon, & Fonagy e Rochat hanno evidenziato, un abbandono della sfera emotiva, "colpevole" del *mindreading*: il discorso condotto da Carruthers non vuole essere una difesa o una condivisione delle "risposte" date al problema del *mindreading* come spiegazione del modo in cui un individuo attribuisce stati mentali, credenze e volizioni all'altro attraverso o la *folk psychology* o forme innate (*Theory-Theory* o *Simulation Theory*), poiché ciò significherebbe, ontologicamente, soffermarsi su categorie o dibattiti che ben si allontanano dall'intento dell'Autore. Non ha obiezioni o difficoltà a sostenere che esistono prove sia a favore di un meccanismo evolutivo del *mindreading* in specie maggiormente complesse come l'uomo sia a favore della tesi contraria; *TT* o *ST* sono solo un approccio particolarizzato al problema più generale dell'architettura mentale.

In questo senso, allora, l'etichetta assegnata al suo modello (*mindreading is prior*) non deve essere fonte di ambiguità sul significato che Carruthers stesso vuole attribuire a quel "prior": non è la predizione di una facoltà emersa prima di altre nel processo evolutivo, bensì la denotazione di uno stesso sistema (la lettura della mente) che sottende o da cui nasce l'abilità della metacognizione.

A maggiore esplicitazione della sua posizione, si riporta, di seguito, quanto da lui stesso ribadito:

In response to Mills & Danovitch, I should also stress that although a finding that mindreading competence is developmentally prior to metacognition would support a "mindreading is prior" account (because it would be inconsistent with the other three alternatives), this is not actually a prediction of the account. For the latter claims only that it is the same system that underlies our mindreading capacity that gets turned upon ourselves to issue in metacognition. It does not claim that the first occurs in development before the latter. (In this respect, the label "mindreading is prior" may be somewhat misleading. I intend it only to refer to a functional and/or evolutionary priority)²²⁵.

²²⁵ *IBIDEM*, p. 167.

La stessa ambiguità potrebbe essere stata riscontrata nell'uso che il filosofo statunitense impiega del termine "introspezione"; Carruthers, infatti, non nega l'esistenza dell'introspezione *tout court*, ma le nega uno statuto speciale nella spiegazione dell'architettura mentale proposta. Essa, infatti, è stata oggetto di un doppio fraintendimento legato proprio all'utilizzo del termine: da una parte, si denota la percezione corporale di un'emozione o di un sentimento provato dall'individuo, il quale sarebbe in grado di capire il suo stato psico-fisico non da una riflessione su se stesso, bensì da una percezione sensoria del proprio corpo; dall'altra, nell'accezione che Carruthers sposa, si intende la riflessione sulle percezioni che interessano il corpo stesso. In questo senso, lo sforzo cognitivo su sé darebbe origine sì ad un atto "introspettivo", la cui origine, però, è da ricercarsi nella facoltà del *mindreading* (il quale, come detto, catturerebbe ogni *input* "interno" ed "esterno" per fornirne una trasmissione generale ai sistemi interconnessi al *mindreading* per il processo decisionale e simili). Cambia, vale a dire, il tipo di accesso alla propria mente e, di conseguenza, a quelle altrui; proprio per questo motivo,

I defined introspection negatively for the purposes of the target article. For I wanted to leave open "inner sense" accounts as well as "application of a tacit theory" views of introspection, according to each of which the attribution of mental states to oneself is inferential (but still quite different from the attribution of mental states to other people)²²⁶.

La differenza fra un significato ed un altro del concetto di "introspezione" non si gioca sulla diversità dei "dati" utilizzati per il processo, bensì sull'utilizzo dei canali a disposizione: non due capacità distinte con distinti "risultati", bensì un'unica capacità (il *mindreading*) con due diversi canali informativi, aventi, ognuno, risultati differenti. In questo senso "mindreading is prior".

A conferma di quanto sostenuto, Carruthers esplicita che i dati derivanti dai pazienti con il "cervello diviso", sostenendo che l'interpretazione più adeguata alla condizione di tali pazienti non può che essere quella che lui stesso propone.

Un dubbio, però, sembra alla scrivente non sia stato sollevato: se è vero che il cervello possiede una certa plasticità per cui è in grado di "trovare strade alternative" nel momento in cui determinate sinapsi siano compromesse, come si potrebbe esser certi che il cervello di tali pazienti non abbia già operato in tal modo quando "interrogati"? Chi ci assicura, cioè, che le evidenze riportate siano verifiche sui reali effetti dell'operazione subita?

²²⁶ *IDEM*, p. 168.

Se a tale dubbio fosse riconosciuta una certa autorevolezza, allora le prove che Carruthers in questo senso riporta a sostegno del suo modello potrebbero delegittimarsi e l'aspetto interessante, a quel punto, sarebbe capire in che modo il *mindreading is prior* potrebbe portarne delle ulteriori.

Quale ultimo appunto degno di nota al fine della trattazione del tema in oggetto nel seguente lavoro, è importante far rilevare al lettore la risposta del filosofo statunitense all'accusa di non aver usato il termine "metacognition" con accuratezza scientifica, fonte, a questo punto, di ambiguità e relativo scetticismo.

In realtà, in quella che sarebbe potuta esser letta quale superficialità dell'autore, si mostra la consapevolezza di un intenso dibattito a monte dell'uso del termine imputato che richiama, genericamente, la relazione fra *metacognition/metarepresentation* o come controllo di un atto cognitivo su se stessi o come processo inferenziale dall'atto metacognitivo con caratteristiche e peculiarità proprie o simili che farebbero, quindi, pensare all'uso epistemico del pensiero. Quest'ultimo sarebbe meramente concettuale o proposizionale? Che tipo di contributo possono fornire gli studi *in itinere* sulle capacità metacognitive degli animali e dei pazienti autistici o schizofrenici? Resta, ancora una volta, il dilemma di capire in che modo l'uomo possa accedere a se stesso. Allora l'intento di Carruthers, così come di tutti coloro che sposano "l'amore per il sapere", resta l'augurio che i progressi del pensiero (riflessivo e scientifico) possano, un giorno, redimere controversie e perplessità sul modo in cui il cervello funzioni.

Al di là di tale auspicio, le risposte di Carruthers agli appunti o ai disappunti postigli restano fonte di ulteriore approfondimento quale punto nodale da cui si diramano i progressi apportati al problema "empatia cognitiva": non soltanto per i temi trattati in sé, ma per avere un'idea quanto più chiara di che cosa sia il "mondo celato" della stessa empatia e di che cosa, effettivamente, essa denoti.

Capitolo V

Il dibattito sul libero arbitrio

5.1 Il determinismo: punti di forza, carenze e difficoltà

Quando si parla di “determinismo” si suole indicare un insieme di idee o di dottrine che spaziano dalla fisica alla filosofia²²⁷: determinata può essere la teoria delle particelle elementari attraverso cui la fisica classica comprende il reale oppure quell’insieme di idee comuni sull’uomo, prive di un certo argomento critico, secondo cui quest’ultimo non sarebbe né libero né responsabile. Un terzo modo in cui sovente si intende la parola “determinismo” è quello attraverso il quale si denota una particolare tesi filosofica, secondo cui il mondo non solo è rigidamente necessitato – per cui ogni evento B non può non manifestarsi – ma, contestualmente, il suo apparire segue causalmente dal fatto A che lo ha preceduto; seguendo, quindi, la rigida causazione necessaria, esso è regolato secondo un certo ordine che si estenderebbe tanto agli oggetti inanimati, quanto a quelli animati.

Per i deterministi il problema di come dar conto del libero arbitrio nell’*animale razionale* o di come conciliare la libertà con la rigida determinazione è del tutto secondario, poiché il *focus* della tesi è individuare il principio cardine che regola il mondo e l’universo *tout court*. Sembrerebbe, cioè, che si consideri (secondo una certa corrente all’interno del variegato mondo determinista) l’uomo simile ad una macchina di cartesiana memoria con la differenza

²²⁷ Per i diversi sensi in cui si può intendere la parola “determinismo”, si veda TIM HONDERICH, *Sei davvero libero? Il problema del determinismo*, Il Saggiatore, Milano, 1996.

che la *res cogitans* altro non sarebbe che la stessa *res extensa*, la quale obbedirebbe alle medesime leggi fisiche del cosmo²²⁸.

Il *continuum* fra mente e materia ha fatto riscuotere al determinismo un notevole consenso in ambito scientifico, ma, contemporaneamente, ha anche alimentato un articolato dibattito intorno al libero arbitrio, facendo sorgere il dubbio sul modo in cui l'uomo, in un contesto rigidamente necessitato, possa "realmente" avere una volontà libera.

Secondo la visione di una corrente del determinismo – il compatibilismo – il principio causale e la libertà umana non devono considerarsi mutuamente esclusivi, ma, al contrario, mutuamente inclusivi: a differenza della tesi classica, infatti, per il compatibilista si definisce libero colui che, non essendo impossibilitato ad agire per impedimenti o costrizioni esterne, compirà qualsiasi azione dettatagli dalla sua stessa volontà.

È essenziale notare che, definita in tal modo, la libertà non è affatto in contraddizione con la causalità: l'azione, discendendo direttamente dalla volontà dell'agente e derivante, a sua volta, dalle infinite cause necessarie che la determinano, si manifesta nelle azioni materialmente compiute dallo stesso; infatti, la volontà, non essendo in grado di desiderare altro rispetto a ciò che desidera, è completamente determinata da tutti quei fattori al di fuori del suo controllo.

Se questa è la tesi di partenza che, *prima facie*, sembrerebbe essere una pacifica soluzione all'interrogativo determinista sul libero arbitrio in relazione al principio causale, ad un'analisi attenta si evince la grande difficoltà in cui lo stesso compatibilismo incorre: affermare che la libertà si estende solo alle azioni dell'agente ed alla volontà che, però, non è in grado di desiderare altro rispetto a ciò che desidera, perché necessitata nel suo stesso desiderare, significa che non può esservi libertà d'azione, in quanto essa non discende direttamente dalle medesime intenzioni dell'agente.

Seppur la volontà, in teoria, potrebbe autodeterminarsi, di fatto essa è rigidamente sottoposta al determinismo causale che ne impedisce la possibilità di scelta, essendo tanto la volontà stessa quanto l'azione sottoposte alla rigida necessitazione causale.

A tale obiezione i compatibilisti hanno ribadito che il ricorso alla causalità non significa impedire l'azione dell'agente; la spiegazione di tale posizione è fornita dal filosofo statunitense Harry Frankfurt²²⁹ che i compatibilisti odierni reputano di fondamentale importanza.

²²⁸ Cfr. FEDERICA MADONNA, *Il problema aperto della mente in relazione alle neuroscienze. Che cosa ne direbbe, oggi, Cartesio?*, op. cit.

²²⁹ Per ulteriori informazioni, si consulti DERK PEREBOOM, *Living without Free Will (Cambridge Studies in Philosophy)*, Cambridge University Press, Cambridge, 2001.

L'argomento, denominato delle "analisi gerarchiche", sostiene che gli agenti, oltre alle normali volizioni dettate dalla volontà, sono in grado di formare "desideri di secondo ordine", aventi come oggetto altri desideri, garantendo alla strutturazione gerarchica di desideri e metadesideri, che la volontà e l'azione compiuta dall'agente discendano direttamente da se stessi²³⁰; ciò dimostrerebbe che la volontà, sia di fatto sia in teoria, non sia rigidamente determinata e condizionata dagli eventi esterni.

Per quanto l'argomento di Frankfurt possa sembrare risolutivo, non chiarisce, però, il problema prospettato dagli indeterministi: le scelte dell'agente in che modo derivano, *genuinamente*, dalla sua volontà²³¹. Se, come sostiene De Caro, concediamo che l'analisi delle intenzioni di Frankfurt funzioni, Frankfurt stesso non ha spiegato come i desideri di secondo ordine "nascono"; se essi sono determinati da fattori che sfuggono al controllo dell'agente, il problema prospettato dagli indeterministi ritorna invariato: per quale motivo considerare un agente libero quando compie azioni dettate dai suoi metadesideri su cui egli non esercita nessun controllo diretto? Se anche si volesse supporre che tali metadesideri siano a loro volta il risultato di altri metametadesideri, il problema rimarrebbe immutato, poiché si innescherebbe un regresso all'infinito che non risolverebbe la *quaestio*.

Per rispondere, infatti, a coloro che sostengono che la vera natura della libertà non risiede nel non essere impediti a compiere ciò che la volontà detta, i compatibilisti sono ricorsi a "l'analisi condizionale dell'espressione della nozione del poter fare altrimenti", secondo cui la libertà deve essere interpretata in termini ipotetici: se, cioè, la volontà dell'agente fosse stata diversa nel momento in cui egli si fosse trovato a dover compiere una scelta allora il percorso d'azione intrapreso sarebbe stato diverso proprio perché l'azione sarebbe discesa direttamente dalla volontà del singolo.

Scrivo, a tal proposito, George Edward Moore:

- a. "Egli avrebbe potuto fare altrimenti";

significa che:

- b. Se egli avesse scelto/tentato/desiderato/voluto/enunciato/intrapreso di fare altrimenti, egli avrebbe agito altrimenti²³².

²³⁰ Cfr. MARIO DE CARO, *Il libero arbitrio, Una introduzione*, Editori Laterza, Milano, 2011.

²³¹ *IDEM*, p. 66.

²³² Cfr. GEORGE E. MOORE, *Ethics*, Oxford University press, London, 1912; traduzione italiana di PREDAVAL M.V. MAGRINI (A CURA DI), *Etica*, Franco Angeli, Milano, 1982.

La verità di b. – come sottolinea Moore – è in perfetto accordo con la visione compatibilistica secondo cui un agente, nonostante sia necessitato in ciò che fa²³³, avrebbe sempre, in linea di principio, un’alternativa valida con cui si manifesterebbero il suo libero arbitrio e la sua libertà di scelta.

Se a., quindi, coincide con la visione compatibilistica e b. ne è la sua diretta spiegazione, allora entrambe le proposizioni sostengono la medesima verità: compatibilismo e libero arbitrio sono possibili insieme.

Nonostante tale analisi potrebbe salvare *free will* e responsabilità morale dalla deriva nichilista prospettata dagli incompatibilisti, resta legittimo il dubbio sulla genuinità del ragionamento di Moore: stando ad un’analisi formale, infatti, l’esito cui il ragionamento ipotetico perviene è che la *conclusio* discende direttamente dalle premesse, le quali, a loro volta, non forniscono informazioni ulteriori rispetto a quelle già contenute nella tesi ancora tutta da dimostrare. Resta, perciò, un ragionamento apodittico.

Altra conseguenza diretta è lo spostamento dell’analisi non più sul rapporto necessitato e causale che lega l’uomo al mondo, ma solo ed esclusivamente l’uomo a se stesso: dalla verità di b. (*Se egli avesse scelto/tentato/desiderato/voluto/enunciato/intrapreso di fare altrimenti, egli avrebbe agito altrimenti*) non si può abduire la verità di a. (secondo cui, si ricorda, *Egli avrebbe potuto fare altrimenti* – la tesi da dimostrare –) ma, piuttosto, l’asserzione

c. Egli avrebbe potuto fare altrimenti.

Giungere alla *conclusio* c. significa sostenere non più la verità di a., ma quella di non a.: se un uomo non avesse potuto scegliere di fare altrimenti, allora non avrebbe potuto fare diversamente anche se lui era tale che se avesse scelto di fare altrimenti, lo avrebbe potuto fare. Tale ragionamento, per quanto corretto da un punto di vista logico, non nasconde, tuttavia, qualche perplessità; esso, infatti, non mette in luce il motivo per cui un agente potrebbe esser ritenuto responsabile per ciò che fa: dove si trova e da dove si inferisce la sua attribuzione di responsabilità?

Tralasciando, al momento, il problema aperto sul rapporto reale o virtuale fra la metafisica della libertà e della responsabilità, il compatibilista Robert Chisholm, per ovviare alla deriva scettica o, addirittura, nichilista cui potrebbe incorrere il compatibilismo – laddove si sostenesse che non vi sarebbe responsabilità morale per la scelta operata dall’agente,

²³³ Cfr. ROBERT CHISHOLM, *Human Freedom and self*, in GARY WATSON (A CURA DI), *Free Will*, Oxford University Press, USA, 2003.

rigidamente necessitato nella volontà e nell'azione –, sostiene che quando un evento o stato di cose (o un insieme di eventi o stati di cose) causa qualche altro evento o stato di cose, allora si ha un esempio di *causalità transeunte*²³⁴; espediente trovato dal filosofo per tentare, in tal modo, di salvaguardare sia la tesi compatibilista sia il rapporto fra *free will* e responsabilità. Tale causalità implicherebbe ciò che Chisholm chiama *causalità immanente*, espressione chiaramente ripresa dalla *Fisica* di Aristotele:

*And I shall say that when an agent, as distinguished from an event, causes an event or state of affairs, then we have an instant of immanent causation. The nature of what is intended by the expression 'immanent causation', may be illustrated by this sentence from Aristotle's Physics: 'Thus, a state moves a stone, and is moved by a hand, which is moved by a man'*²³⁵.

Nella frase: “un bastone muove una pietra, la quale è mossa dalla mano, che è mossa dall'uomo” per Chisholm risiede il senso della sua posizione; ossia l'esplicazione dei concetti della *causalità transeunte* e della *causazione immanente*. Il bastone, essendo mosso dalla mano, mossa a sua volta dai suoi muscoli, azionati a loro volta dal cervello, è stato mosso dall'uomo (*causazione immanente*) che si esplica nel mondo (*causazione transeunte*) ogniqualvolta un uomo “fa qualcosa” (*doing A*), “facendo sì che accada” (*making something happen*)²³⁶.

Una perplessità sulla posizione di Chisholm riguarda il medesimo concetto di *causazione immanente*: che cosa significa ed in che cosa consiste la causalità propria dell'agente? Non sembra, in altre parole, esserci alcuna differenza fra “l'accadere di un evento” e “l'essere la causa di un evento”

*(What, for example: the difference between A's just happening, and the agents' causing A happen?*²³⁷).

Il punto, per Chisholm, non risiede nel fatto di dover spiegare la differenza o la natura del nesso causale, ma di porre l'accento – come riprende da Reid –

²³⁴ Si legge: “I shall say that when one event or state of affairs (or set of events or states of affairs) causes some other event or state of affairs, then we have an Instance of *transeunt* causation”, *IDEM*, p. 22.

²³⁵ Si legge “E sosterrò che quando un agente, distinto da un evento, causa un evento o uno stato di cose, abbiamo in quell'istante una causazione immanente. La natura di ciò che intendo con l'espressione causazione immanente può essere illustrato da questa frase tratta dalla Fisica aristotelica: ‘Dunque, un bastone muove una pietra che è mosso da una mano che è mossa da un uomo’”, *IVI*.

²³⁶ *IVI*.

²³⁷ *IDEM*, p. 23.

*The conception of an efficient cause may very probably be derived from the experience we have had of our own power to produce certain effects*²³⁸:

l'uomo, in altre parole, per Chisholm, sarebbe equiparabile al primo motore immobile aristotelico, capace, cioè, di essere all'origine di infinite cause ed effetti che avranno luogo da lui stesso.

Tale affermazione non significa chiedersi se l'uomo libero possieda o meno il libero arbitrio, ma solo se si possa ritenere responsabile delle sue azioni: se si è, cioè, liberi e responsabili di volere o di fare ciò che vorremmo o che desidereremmo fare²³⁹.

Sarebbe da chiedersi in che modo, "autenticamente", si possa esser certi di volere effettivamente ciò che si desidera: se, in altre parole, la volontà fosse rigidamente necessitata sarebbe logica la conseguenza che l'individuo voglia ciò che, poi, vorrà. In tal modo, non conoscerebbe l'alternativa. Anche ammettendo l'ipotesi di Chisholm di equiparare l'uomo ad un primo motore immobile, fonte di infinite catene causali aventi origine da lui stesso, non risolverebbe il problema del vincolo *free will*/determinazione necessaria; inoltre, vi sarebbe un ulteriore problema: in che modo dar conto del rapporto fra molteplici primi motori immobili che, interagendo, comporterebbero, di conseguenza, altrettante infinite catene causali, originatesi da loro stessi? Si tornerebbe, *ab origine*, al problema di Husserl sulle inter-relazioni fra monadi, stavolta aggravata, da una parte, dal modo in cui dar conto del libero arbitrio in un contesto così ipotizzato; dall'altra, sarebbe urgente dar conto di un'ulteriore problema nel problema: quale sarebbe il rapporto fra la necessitazione del mondo e le necessitazioni dei primi motori immobili? Ci sarebbe una priorità di una su un'altra? E secondo quale principio e secondo quale natura delle stesse?

Ritornando a Chisholm, la sua posizione sembrerebbe, come ribadito precedentemente in un breve passaggio, un espediente per ovviare alla *crux* del compatibilismo: sostenere che la causalità transeunte e quella immanente non abbiano l'intento di fornire una risposta all'esistenza del libero arbitrio, servendo "solo" da salvagente per la responsabilità morale, significa non voler fornire – consapevolmente – una risposta al dubbio postogli. In tal senso, Chisholm commetterebbe un'ulteriore mancanza, dimostrando di non considerare un'intera

²³⁸ Si legge: "Sulla concezione di una causa efficiente che potrebbe, molto probabilmente, esser derivata dall'esperienza che noi possiamo avere un potere nel produrre certi effetti", *IBIDEM*, p. 24.

²³⁸ *IDEM*, p. 25.

²³⁹ Cfr. BRUCE AUNE, *Hypotheticals and Can: Another Look*, Vol. 27, n. 6, Oxford University Press, Oxford, 1968, pp. 191-195.

corrente di studi sulla *responsibility* che tenta di distaccarla dall'esistenza reale o ipotetica del *free will*, onde evitare di far risultare la seguente equazione veritiera:

non esiste il libero arbitrio → non esiste la responsabilità morale.

A proposito di *responsibility*, la posizione di Chisholm al riguardo è comunemente condivisa dai compatibilisti anche se, ultimamente, le sue argomentazioni sono entrate in crisi in quanto si sostiene che, se anche le dichiarazioni del filosofo compatibilista siano nel complesso condivisibili, alcune sue assunzioni sull'analisi del "poter fare altrimenti"²⁴⁰ non sembrerebbero scevere da gravi difficoltà.

Il ragionamento: *S avrebbe potuto fare altrimenti* significa che *Se S avesse scelto/tentato/desiderato/voluto/enunciato/intrapreso di fare altrimenti piuttosto che A, S avrebbe agito altrimenti* appare sospetto. Infatti, se da una parte, si può concedere che possa esservi un senso di "potrebbe" per cui valga l'inferenza secondo cui S non avrebbe potuto fare B – ossia qualcosa di diverso rispetto a quello che *de facto* ha fatto –, dall'altra, ci si potrebbe trovare di fronte ad un significato di "potere" che stia ad indicare che l'individuo stesso "fosse impossibilitato" a fare altro.

In questo secondo caso la tesi di partenza potrebbe garantire la deduzione che se S sceglie B, per S fare B è, in qualche modo, necessario e, d'altro canto, se S "non può" scegliere di fare B, allora è necessario che B non venga scelto.

Il punto centrale è che nelle affermazioni di Chisholm sembra che la nozione del "poter fare altrimenti" coinvolga necessariamente il libero arbitrio senza valutare che non ogni senso del verbo "potere" implica in maniera egualitaria il discorso sulla libertà: se un uomo è disposto o sceglie di compiere un atto (ad esempio, trattenere il respiro) che sia causalmente sufficiente per il suo scopo, e lo compie, allora sarà per lui logicamente impossibile compiere qualsiasi altra azione nel bel mentre trattiene il respiro. In questo caso, la sua impossibilità a compiere azioni diverse rispetto a quelle che, di fatto, sta compiendo; non significa che egli sia limitato nella sua libertà, ma, al contrario è proprio la sua libertà che esclude la possibilità di fare altrimenti rispetto a ciò che effettivamente fa.

I diversi tentativi di conciliare in termini modali la nozione del "poter fare altrimenti" con l'esercizio della propria volontà tramite la nozione della *causal impossibility*²⁴¹ (la causazione "impossibilitante" nel far sì che un agente possa agire diversamente rispetto a ciò che *de facto*

²⁴⁰ *IBIDEM.*

²⁴¹ *IDEM.*

ha scelto di compiere) cadono in errore: si giunge sempre alla conclusione che “poter fare altrimenti” implichi, necessariamente, un senso di libertà che potrebbe benissimo non implicare (si ricordi l’esempio dell’uomo che decide di trattenere il respiro).

Discostandosi dalle analisi delle possibilità alternative e da quanto finora emerso, non v’è dubbio che parlare di libero arbitrio significa, inevitabilmente, parlare del suo rapporto problematico con il determinismo; ossia di quello che Peter van Inwagen ha sostenuto essere la prima delle due parti in cui la questione *free will* si divide: da una parte il *Traditional Problem*, cioè se l’uomo possieda o meno il libero arbitrio; dall’altro il *Compatibility Problem*, ovvero se libero volere e determinismo siano compatibili²⁴².

In tal senso, con la rivisitazione del compatibilismo operata da Chisholm sulla possibilità di fare altrimenti, è sembrato a molti e per diverso tempo che il *free will* avesse trovato, apparentemente, una sua risoluzione; una parvenza legata al fatto che alcuni aspetti del *Compatibility Problem* non sembrano siano stati affrontati o abbiano trovato una soluzione nella tesi avanzata dallo stesso Chisholm.

Peter van Inwagen sostiene, infatti, che per dichiarare plausibile una tesi volta a dimostrare la presenza della libertà in un qualsiasi contesto non serve solamente la possibilità di fare altrimenti e l’autodeterminazione della propria volontà, bensì anche il controllo diretto dell’azione da parte del medesimo soggetto agente; mancando quest’ultimo aspetto, decisamente non secondario, sia nel compatibilismo sia nel determinismo, mancherebbe, di conseguenza, la stessa possibilità della libertà. Riconducendo tali riflessioni su quest’ultima, van Inwagen ne deduce l’impossibilità della libertà proprio a causa della mancanza di controllo da parte dell’agente di *tutti* i fattori che lo condurrebbero all’azione. Se ciò fosse vero, sotto il suo controllo dovrebbero essere incluse tanto le leggi di natura quanto tutti gli stati passati dell’universo per soddisfare, pienamente, l’aspetto secondo cui il medesimo agente abbia il pieno controllo delle infinite cause e concause che lo necessitano nella sua azione. Una condizione, precisa van Inwagen, che non si è mai in grado di soddisfare, tanto che nessuno, all’interno del quadro disegnato dal determinismo, può agire mai liberamente²⁴³. Siamo, ad esempio, nel 49 a. C. e Cesare ha varcato il Rubicone, dando inizio alla guerra civile; si può legittimamente supporre che prima di tale azione lo stesso Cesare si fosse interrogato sulle possibili conseguenze, ponderando i vantaggi e gli svantaggi che ciò gli avrebbe procurato.

²⁴² Cfr. PETER VAN INWAGEN, *An essay on free will*, Oxford University Press, USA, 1986.

²⁴³ *IBIDEM*.

Se questo è il contesto, il nocciolo del *Consequence Argument* – l’argomento di van Inwagen – è che, se il determinismo fosse vero, Cesare non avrebbe agito liberamente in quanto egli non sarebbe stato in grado di avere sotto il suo controllo i fattori che lo avrebbero portato in tale situazione: non era in grado, cioè, né di cambiare le leggi di natura né tanto meno di modificare gli eventi passati dell’universo per averne il diretto controllo²⁴⁴.

Da ciò si dimostrerebbe che il determinismo rende impossibile il “concretizzarsi” della libertà: né il controllo diretto dell’azione né la scelta di questa fra diversi corsi d’azione possibili sono accessibili all’agente determinato e, di conseguenza, le condizioni della libertà positiva non sono garantite²⁴⁵.

Sembrerebbe proprio che le difficoltà interne alla stessa tesi deterministica, sia nella sua versione *hard* sia nelle sue versioni compatibilistiche, le rendano impossibile il compito di spiegare e garantire la libertà in un contesto fortemente necessitato: mancando, infatti, la possibilità di fare altrimenti ed il diretto controllo delle cause che condurrebbero ad un’azione necessaria, in che modo un agente potrebbe mai dirsi libero? Dove si trova lo spazio logico e pratico della libertà?

La plausibilità del compatibilismo sembrerebbe gravemente compromessa, tanto che molti guardano ed hanno guardato alla tesi contrapposta – l’indeterminismo – come *la* possibilità di dirimere le controversie e fornire una spiegazione plausibile, sia del libero arbitrio sia della metafisica della libertà.

5.2 L’indeterminismo: un’alternativa, forse, non troppo convincente

L’indeterminismo è quella tesi ontologica che si contrappone, come detto, al determinismo e sostiene che il mondo non è regolato da nessun tipo di causalità necessaria, ma dalla contingenza; quel principio secondo cui un evento potrebbe tanto accadere quanto non accadere, senza comportare alcun cambiamento nel mondo²⁴⁶.

Così come nel determinismo, anche nell’indeterminismo esistono diverse correnti di pensiero – sempre definibili come indeterministe – che negano *in toto* o accettano, secondo certe condizioni, il nesso causale che, comunque, non è mai necessario.

²⁴⁴ Cfr. MARIO DE CARO, *Il libero arbitrio. Una introduzione*, op. cit.

²⁴⁵ Per *libertà positiva* si intende quella posizione secondo cui ogni agente, per mezzo del proprio libero arbitrio, sia in grado di autodeterminarsi, essendo pienamente responsabile delle proprie scelte; per *libertà negativa*, al contrario, si intende la possibilità di agire senza costrizioni o impedimenti esterni (il tipo di libero arbitrio sostenuto dai compatibilisti).

²⁴⁶ Cfr. LUDWIG WITTGENSTEIN, *Tractatus logico-philosophicus*, Einaudi Editore, Torino, 1996.

Nell'*indeterminismo radicale*, definito anche *indeterminismo acausale*, qualsiasi compromesso con il principio causale, tipico del determinismo e non solo, è completamente abbandonato: ogni azione, vale a dire, accade *senza* una causa che la generi.

A differenza, infatti, del campo delle scienze naturali, le azioni *indeterministiche radicali* hanno un carattere puramente *intenzionale*²⁴⁷: alla base di un'azione, sostiene Carl Ginet – il maggiore fautore di tale corrente – vi è sempre un motivo da ricercare nell'intenzione dell'agente, priva di una qualsiasi struttura causale. Il problema metafisico della connessione fra la volizione e la ragione di un'azione è risolto, secondo Ginet, dalla convinzione che all'interno delle medesime volizioni sono da ricercarsi sempre le ragioni di un'azione che costituiscono la *ratio* o “la causa” della stessa. In tal senso, allora, il limite concettuale della difficoltà di controllare la propria azione quale risultato di un atto volitivo sarebbe superata dalla constatazione che ad accompagnare qualsiasi azione mentale semplice vi sarebbe l'*actish phenomenal quality*, la consapevolezza fenomenologica di essere l'autore indiscusso della propria azione²⁴⁸:

*Se un evento è una propria azione, allora il mio determinarla richiede semplicemente che io la compia; e il mio compiere l'azione, che è per l'appunto la sua esecuzione, è compatibile con il fatto che quell'evento sia indeterminato, non causalmente necessitato dai suoi antecedenti*²⁴⁹.

A sua volta l'agente sarebbe consapevole che l'intenzione, ossia la diretta spiegazione dell'esecuzione dell'azione, derivi dalla sua stessa intenzione:

*la tesi che un'azione libera non può essere determinata nomicamente dai suoi antecedenti non implica che un agente non possa determinare quale azione libera compie o che egli non possa compiere un'azione libera per date ragioni*²⁵⁰.

Coniugando il problema dell'induzione (di *humana memoria*) con l'acausalità, gli indeterministi radicali sostengono che la connessione, *a posteriori*, fra un'azione e la sua esecuzione – il cui risultato consiste in un approccio prettamente fenomenologico all'evento stesso – è spiegato attraverso il ricorso al ragionamento logico-deduttivo, il quale esclude *in*

²⁴⁷ Cfr. JASON R. LUCAS, *The Freedom of the Will*, Clarendon Press, Oxford, 1970.

²⁴⁸ Cfr. CARL GINET, *On Action*, Cambridge University Press, Cambridge, 1990.

²⁴⁹ *IBIDEM*, p. 127.

²⁵⁰ CARL GINET, *Reasons Explanation of Action: An Incompatibilist Account*, in “Philosophical Perspective”, Vol. 3, 1989, p. 43.

toto qualsiasi spiegazione fisica dell'evento stesso, facendo appello alla sola natura intenzionale dell'azione stessa:

*La sola cosa necessaria per la verità di una determinata spiegazione basata su ragioni, oltre al darsi dell'azione spiegata, è che l'azione sia stata accompagnata da un'intenzione con il giusto tipo di contenuto*²⁵¹.

Una spiegazione che, però, lascia al lettore non pochi dubbi, poiché

*se un'azione è un evento incausato – anche se ‘accompagnato’ da un'intenzione che lo riguarda, come scrive Ginet – essa è ipso facto indipendente da tutti gli eventi precedenti, inclusi quelli che riguardano l'agente [...]: ‘un evento totalmente incausato è un evento che scaturisce dal nulla, che ha la sua origine nel nulla’ [...]. La rinuncia al nesso causale tra agente e azione sembra dunque compromettere non solo la possibilità dell'azione libera, ma dell'azione tout court*²⁵².

Sembra evidente che il punto focale della tesi indeterminista, indipendentemente dalle sue correnti, sia l'assunzione di un concetto di libertà umana più vicino al senso comune piuttosto che ad una tesi filosoficamente così impegnativa. A differenza del determinismo, infatti, sembrerebbe, *prima facie*, ragionevole ed ontologicamente plausibile pensare la libertà quale scelta fra possibilità alternative e fra percorsi d'azione resi attuabili non in linea di principio (metafisicamente), ma *categoricamente*, empiricamente, salvando, di conseguenza, anche la responsabilità morale dell'agente. È ovvio, altresì, che i fautori della tesi contrapposta agli indeterministi radicali, fra cui il filosofo statunitense John Martin Fisher, non condividano tale posizione, poiché:

*There is simply no good reason to suppose that causal determinism in itself (and apart from considerations pertaining to alternative possibilities) vitiates our moral responsibility*²⁵³.

In accordo al *source compatibilism* – e, quindi, alla difesa del compatibilismo operata da Frankfurt – per Fisher un agente è da ritenersi moralmente responsabile non per la

²⁵¹ CARL GINET, *On Action*, op. cit., p. 138.

²⁵² MARIO DE CARO, *Il libero arbitrio. Una introduzione*, op. cit., pp. 43-44.

²⁵³ JOHN FISHER, *My Way*, Oxford University Press, Oxford, 2006, pp. 201-202.

disponibilità di un percorso d'azione alternativo, bensì per il fatto che quella data azione abbia una storia che la porta ad essere, simultaneamente, l'effetto e la causa della medesima azione di cui l'agente è tanto la causa quanto l'effetto. Una posizione che, fra punti di convergenza e di discontinuità, è vicina, se ci si sofferma un attimo a riflettere, a quella del già citato Chisholm.

Tra i fautori dell'incompatibilismo radicale, poi mutato in *source incompatibilism*, il filosofo statunitense Derk Pereboom sostiene che in qualsiasi scenario compatibilista, sia esso fantascientifico o realistico, l'agente non potrà mai ritenersi libero e responsabile, poiché non ci sarà mai, da parte dell'agente stesso, una scelta "genuina" derivante direttamente da sé e non dalla rigida concatenazione causale cui è sottoposto. Pereboom, con il *Multiple-Case Manipulation Argument*²⁵⁴, sostiene, *contra Fischer*, che

[...] la prospettiva ottimistica secondo cui concepire l'esistenza senza questo genere di libero arbitrio non sarebbe distruttivo per la moralità o per il significato che attribuiamo all'esistenza stessa; per certi aspetti, anzi, potrebbe essere perfino benefica.

Rifiuto inoltre un incompatibilismo in base al quale sia cruciale la disponibilità di possibilità alternative al fine di spiegare la responsabilità morale e accetto, invece, un incompatibilismo che attribuisce il ruolo più rilevante alla storia causale dell'azione. Sostengo che la responsabilità morale di un agente per un suo atto sia spiegata in primo luogo dal fatto che l'azione ha una storia causale nella quale l'agente è la sorgente del suo atto in modo specifico. Opto quindi per un incompatibilismo dell'agente come causa (source incompatibilism), in opposizione all'incompatibilismo delle possibilità alternative (leeway incompatibilism)²⁵⁵.

Il senso principe dell'argomento delle manipolazioni risiede nel fatto che per difendere il libero arbitrio nei modi e nei termini dell'incompatibilismo duro si debba, inevitabilmente, affrontare il compatibilismo e mostrare come la manipolazione inconsapevole della volontà di un agente nell'esecuzione di un'azione generi la convinzione della sua assenza di responsabilità morale per quella e per tutte le azioni conseguenti da quella incriminata.

L'argomento di Pereboom si articola, ipotizzando quattro diversi scenari, che analizzeremo nelle prossime pagine; ad uno di essi in particolare, l'autore concederà la verità del

²⁵⁴ DERK PEREBOOM, *Lo scetticismo ottimistico su libertà e responsabilità* (traduzione di Andrea Lavazza), in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Quanto siamo responsabili? Filosofia, Neuroscienze e Società*, Codice Edizioni, Torino, 2013, pp. 127-155.

²⁵⁵ *IDEM*, p. 129.

compatibilismo; la difficoltà a cui i compatibilisti saranno chiamati a rispondere sarà fornire una spiegazione plausibile dei motivi per cui nei primi tre casi l'agente non possa essere ritenuto responsabile (nonostante il contesto risponda ai criteri compatibilistici) e perché nell'ultimo (il quarto), a parità di condizioni, quello stesso agente degli scenari fantascientifici se calato in un contesto ordinario di vita vissuta non debba più essere considerato responsabile. Per Pereboom, una parzialità così evidente non è giustificabile, poiché non esistono differenze in grado di garantire la mancanza di responsabilità in un mondo e non in un altro:

*A additional challenge for the compatibilist is point out a relevant and principled difference between any two adjacent cases that would show why the agent might be morally responsible in the later example but not in the earlier one. I argue that this can't be done, and that the agent's non-responsibility therefore generalizes from the first of the manipulation examples to the ordinary case*²⁵⁶.

La *ratio* dei *Manipulation Cases* è, come si è potuto intuire, avviare la falsificazione del principio causale rigidamente necessitato in favore della tesi incompatibilista. Tornando agli scenari di Pereboom, in ognuno dei quattro casi, il Professor Plum decide di uccidere la Sig.ra White (atto che di fatto compirà), poiché consapevole che la sua morte gli arrecherà dei notevoli vantaggi personali.

L'azione su cui il filosofo statunitense incentra la sua riflessione per dimostrare l'incompatibilità fra determinismo e responsabilità morale è proprio la decisione di Plum di uccidere White (*a basic mental action*)²⁵⁷, la quale soddisfa tutte le effettive e possibili condizioni compatibiliste evocate:

- da Hume²⁵⁸; l'azione non è estranea al carattere del Sig. Plum, poiché per lui le sue personali motivazioni egoistiche hanno un peso notevole nella presa di decisione cosciente di qualsiasi azione per cui, da un punto di vista "personale", nonostante la consapevolezza delle ragioni morali che lo indurrebbero a non compiere l'azione, Plum non risulta costretto dalle catene deterministiche ad uccidere la Sig.ra White;

²⁵⁶ DERK PEREBOOM, *Free will, Agency and meaning of Life*, Oxford University Press, Oxford, 2014, p. 75.

²⁵⁷ *IDEM*.

²⁵⁸ Cfr. DAVID HUME, *Trattato sulla Natura Umana*, op. cit.

- da Frankfurt²⁵⁹; la volontà di uccidere la Sig.ra White (desiderio di primo ordine) è conforme ai desideri di secondo ordine (la considerazione personale che uccidere White, nonostante sia sbagliata, persista in Plum per giungere ad un vantaggio personale);
- da John Fisher e Mark Ravizza²⁶⁰; Plum sarebbe in grado di modificare la sua azione se, razionalmente, ponderasse che le conseguenze dell'eventuale uccisione della Sig.ra White gli procurassero più danni che benefici;
- da Jay Wallace²⁶¹; secondo cui Plum ha l'abilità di analizzare, cambiare e comprendere le sue azioni in base alla morale di riferimento.

Pertanto, resta da chiedersi se, in un mondo compatibilista, Plum sia davvero responsabile della sua azione in concomitanza alla variabile della manipolazione introdotta da Pereboom.

Nel Caso I, si legge:

Case I: A team of neuroscientists is able to manipulate Professor Plum's mental state at any moment through the use of radio-like technology. In this case, they do so by pressing a button just before he begins to reason about his situation. This causes Plum's reasoning process to be strongly egoistic, which the neuroscientists know will deterministically result in his decision to kill White. Plum does not think and act contrary to character since his reasoning processes are frequently egoistic and sometimes strongly so. His effective first-order desire to kill White conforms to his second-order desires. The process of deliberation from which his action results is reasons-responsive; in particular, this type of process would have resulted in his refraining from deciding to kill White in some situations in which the reasons were different. Still, his reasoning is not in general exclusively egoistic, since he often regulates his behavior by moral reasons, especially when the egoistic reasons are relatively weak. He is also not constrained to act as he does, in the sense that he does not act because of an irresistible desire – the neuroscientists do not induce a desire of this kind²⁶².

²⁵⁹ Cfr. HARRY FRANKFURT, *Freedom of the Will and The Concept of Person*, in "The Journal of Philosophy", Vol. 68, 1971, pp. 5-20.

²⁶⁰ Cfr. JOHN FISHER E MARK RAVIZZA, *Responsability and Control: A Theory of Moral Responsibility*, Cambridge University Press, Cambridge, 1998.

²⁶¹ Cfr. JAY WALLACE, *Responsability and the Moral Sentiments*, Harvard University Press, Cambridge (MA), 1994.

²⁶² DERK PEREBOOM, *Free Will, Agency and Meaning in Life*, op. cit., p. 77.

Si legge: "Un'equipe di neuroscienziati è in grado di manipolare in qualsiasi momento gli stati mentali del professor Plum attraverso l'uso di una sorta di radiocomando. In questa circostanza, essi lo fanno schiacciando

Da quanto si evince dal testo, un *team* di neuroscienziati, in grado di manipolare gli stati mentali del Professor Plum (credenze, volizioni e desideri) attraverso una tecnologia a radiofrequenza, sarebbe pronto a modificare il suo processo deliberativo laddove quest'ultimo si astenesse, alla fine, dalla decisione e dall'atto di uccidere la Sig.ra White. Sono specificate, inoltre, due condizioni degne di nota: il forte egoismo di Plum, coerente con la sua personalità, non lo rende inabile a capire la differenza fra quello che sarebbe moralmente giusto e quello che sarebbe moralmente sbagliato (tanto che Plum stesso sarebbe in grado di modificare il suo agire, seguendo non i suoi desideri o bisogni egoistici, ma ciò che la morale dominante ritiene "moralmente giusto"); i neuroscienziati, nonostante siano in grado di determinarne *a priori* il processo deliberativo, inducendolo a compiere un'azione piuttosto che un'altra, non possono intervenire sul desiderio stesso se è soddisfatta la condizione secondo cui quest'ultimo gli derivi direttamente dalla confluenza dei desideri di primo e di secondo ordine.

La domanda che pone Pereboom è se, intuitivamente, si consideri Plum responsabile delle sue azioni, nonostante la presenza di un'eventuale manipolazione sulla sua "psiche"; almeno nel Caso I, il filosofo statunitense vuole evidenziare la personalità di Plum ed il suo processo deliberativo, facendo rimanere in secondo piano la presenza di una eventuale manipolazione da parte dei neuroscienziati rispetto all'intuizione di un'altrettanta presunta o veritiera responsabilità dell'agente.

L'intervento *neutrale* dei "burattinai scientifici" accrescerebbe, da una parte, le disposizioni egoistiche di Plum; dall'altra non indebolirebbe l'agentività del protagonista che, però, intuitivamente verrebbe percepito meno responsabile della sua decisione.

Il secondo caso, a differenza del primo, è costruito in modo da essere più coerente con uno scenario reale:

un tasto appena prima che egli cominci a ragionare sulla propria situazione. Ciò fa sì che il ragionamento di Plum sia egoistico, tale che i neuroscienziati sanno lo porterà deterministicamente alla decisione di uccidere la signora White. Plum non decide e non agisce in modo contrario al suo carattere, dato che i suoi ragionamenti non sono infrequentemente egoistici. Il suo desiderio efficace di primo ordine di uccidere la signora White si conforma ai suoi desideri di secondo ordine. Il processo deliberativo dal quale discende l'azione è sensibile alle ragioni; in particolare, quel tipo di ragionamento avrebbe portato alla decisione di non uccidere la signora White in alcune situazioni in cui le ragioni fossero state differenti. Ancora, il suo ragionamento in generale non è completamente egoistico, dato che egli spesso accorda il suo comportamento alle ragioni morali, specialmente quando le ragioni egoistiche risultano relativamente deboli. Egli non è nemmeno costretto, nel senso che non agisce in base a un desiderio irresistibile: i neuroscienziati non gli inducono un simile desiderio", in *Quanto siamo responsabili?* traduzione di ANDREA LAVAZZA, *Quanto siamo responsabili?*, op. cit., p. 132.

Case II: *Plum is like an ordinary human being, except that neuroscientists programmed him at the beginning of his life so that his reasoning is frequently but not always egoistic (as in Case 1), and sometimes strongly so, with the consequence that in the particular circumstances in which he now finds himself, he is causally determined to engage in the egoistic reasons-responsive process of deliberation and to have the set of first and second-order desires that result in his decision to kill White. Plum has the general ability to regulate his behavior by moral reasons, but in his circumstances, due to the strongly egoistic character of his reasoning he is causally determined to make his decision. The neural realization of his reasoning process and of the resulting decision is exactly the same as it is in Case 1 (although the external causes are different). At the same time, he does not decide as he does because of an irresistible desire*²⁶³.

A differenza del Caso I, stavolta in Plum una manipolazione da parte del *team* di neuroscienziati è fattivamente presente nella storia: alla nascita, infatti, essi lo hanno “programmato” ad un processo deliberativo il più delle volte egoistico, che segua, cioè, le sue inclinazioni personali nell’accrescere il proprio benessere a discapito degli altri in quasi tutte le circostanze che si troverà a vivere.

Sotto tale aspetto la somiglianza con il Caso I è evidente: Plum ragiona seguendo proprie motivazioni egoistiche; è in grado di cambiare idea, seguendo la morale a discapito dei propri desideri ed infine (coerentemente con la sua personalità) non agisce per un desiderio irrefrenabile che potrebbe inficiare la sua razionalità – come nel Caso I –, ma sempre coerente con un processo “razionale”.

Intuitivamente, sostiene Pereboom, Plum non risulta essere moralmente responsabile delle sue azioni e delle sue decisioni, poiché l’intromissione alla nascita della manipolazione operata dal *team* di neuroscienziati lo deresponsabilizza dalle decisioni assunte: quanto, vale a dire, c’è di Plum “reale” nella sua decisione di uccidere la Sig.ra White?

Il terzo scenario, a differenza dei primi due, è più simile ad una situazione ordinaria:

²⁶³ DERK PEREBOOM, *Free Will, Agency and Meaning*, op. cit., p. 77.

Si legge: “Plum è in tutto simile a un normale essere umano tranne per il fatto che alcuni neuroscienziati lo hanno programmato fin dall’inizio della sua esistenza in modo che, frequentemente ma non sempre, ragioni in modo egoistico (come nel Caso 1), con la conseguenza che in circostanze particolari nelle quali ora si trova è causalmente determinato a compiere un processo decisionale sensibile alle ragioni egoistiche e ad avere l’insieme dei desideri di primo e di secondo ordine che sfociano nella sua determinazione di uccidere la signora White. Plum ha la capacità generale di regolare il proprio comportamento in base a ragioni morali, ma in quelle circostanze, a motivo del carattere egoistico del suo ragionare, è causalmente determinato a prendere quella decisione. La realizzazione cerebrale del suo processo di ragionamento e della conseguente decisione è esattamente quella del Caso 1 (benché siano diverse le cause esterne). Nello stesso tempo, Plum non agisce in base a un desiderio irresistibile”, in *Quanto siamo responsabili?*, traduzione di ANDREA LAVAZZA, *Quanto siamo responsabili?*, op. cit., pp. 133-134.

Case III: *Plum is an ordinary human being, except that he was causally determined by the rigorous training practices of his family and community in such a way that his reasoning processes are often but not exclusively rationally egoistic (in this respect he is just like he is in Cases 1 and 2). This training took place when he was too young to have the ability to prevent or alter the practices that determined this aspect of his character. This training, together with his particular current circumstances, causally determines him to engage in the strongly egoistic reasons-responsive process of deliberation and to have the first and second-order desires that result in his decision to kill White. Plum has the general ability to regulate his behavior by moral reasons, but in his circumstances, due to the strongly egoistic nature of his reasoning processes, he is causally determined to make his decision. The neural realization of this reasoning process and of his decision is the same as it is in Cases 1 and 2. Here again he does not decide as he does due to an irresistible desire*²⁶⁴.

A differenza dei casi precedenti, stavolta il condizionamento egoistico di Plum (il quale lo porta a ragionare in termini di convenienza personale in quasi tutte le circostanze che si trova a vivere) gli deriva dalla formazione della famiglia e della comunità in cui è cresciuto; per cui, sostiene Pereboom, il condizionamento è avvenuto in un'età in cui era impossibile per Plum interrogarsi sul biasimo o sulla lode di tali atteggiamenti. Coerentemente con Frankfurt, inoltre, i suoi desideri di primo ordine di voler uccidere la Sig.ra White trovano un corrispettivo in quelli di secondo ordine, in cui, razionalmente, Plum, per vantaggio personale, vuole uccidere l'ostacolo ai suoi profitti. Come nei casi I e II, il Professore, però, non è totalmente condizionato dalla formazione ricevuta: a volte, infatti, è in grado di cambiare il suo agire, rispettando la "doppia morale" di *humeana memoria*, ma il più delle volte il processo deliberativo è determinato dal suo egoismo. Anche qui, infine, la decisione di Plum non è spinta da un desiderio irrazionale, ma da un ragionamento freddamente razionale.

²⁶⁴ DERK PEREBOOM, *Free Will, Agency and Meaning*, op. cit., p. 78.

Si legge: "Plum è un normale essere umano tranne per il fatto che è stato determinato causalmente da un rigoroso programma di addestramento condotto in famiglia e nella sua comunità di modo che, frequentemente ma non sempre, ragiona in modo egoistico (come nel Caso 1 e nel Caso 2). Tale addestramento è avvenuto quando egli era troppo piccolo per avere la capacità di impedire o modificare il programma che ha determinato questo aspetto del suo carattere. L'addestramento, insieme con le sue particolari circostanze attuali, lo determina causalmente a compiere un processo decisionale sensibile alle ragioni egoistiche e ad avere l'insieme dei desideri di primo e secondo ordine che sfociano nella sua determinazione di uccidere la signora White. Plum ha la capacità generale di regolare il proprio comportamento in base a ragioni morali, ma nelle circostanze attuali, a motivo del carattere egoistico del suo ragionare, è causalmente determinato a prendere quella decisione. La realizzazione cerebrale del suo processo di ragionamento e della conseguente decisione è esattamente quella dei Casi 1 e 2. E, ancora, Plum non agisce in base a un desiderio irresistibile", in *Quanto siamo responsabili?*, traduzione di ANDREA LAVAZZA, *Quanto siamo responsabili?*, op. cit., pp. 134-135.

Intuitivamente, sostiene Pereboom, il compatibilista, anche in questo caso, è portato a ritenere Plum moralmente responsabile, adducendo quale spiegazione che il controllo – seppur neutrale – del *team* di neuroscienziati rende plausibile l’assenza del controllo e della reale percezione dell’azione di Plum nei casi I e II; nel III, la determinazione causale dietro l’azione di Plum, lo rende inevitabilmente non responsabile per le medesime ragioni dei casi I e II. Il quarto ed ultimo scenario è studiato da Pereboom in modo che Plum sia calato in un contesto quotidiano di vita vissuta:

Caso IV: *Physicalist determinism is true – everything in the universe is physical, and everything that happens is causally determined by virtue of the past states of the universe in conjunction with the laws of nature. Plum is an ordinary human being, raised in normal circumstances, and again his reasoning processes are frequently but not exclusively egoistic, and sometimes strongly so (as in Cases 1-3). His decision to kill White results from his strongly egoistic but reasons-responsive process of deliberation, and he has the specified first and second-order desires. The neural realization of his reasoning process and decision is just as it is in Cases 1-3. Again, he has the general ability to grasp, apply, and regulate his behavior by moral reasons, and it is not due to an irresistible desire that he kills White*²⁶⁵.

In accordo alla più generale definizione di mondo determinista, Pereboom precisa che, anche in questa situazione, il processo deliberativo di Plum di voler uccidere la Sig.ra White è dettato dal suo egoismo e dalla sua deliberazione razionale, in accordo all’analisi gerarchica di Frankfurt. Pure in questo caso, la decisione non è scaturita né da un desiderio irresistibile né dall’incapacità di Plum di rendersi conto di che cosa sia giusto o sbagliato secondo la morale dominante.

Ad un’analisi delle medesime condizioni sarebbe difficile stabilire la deresponsabilizzazione morale di Plum nel Caso IV, poiché, come in quelli precedenti, il processo deliberativo e l’influenza neutrale del contesto non giocano un ruolo differente nella constatazione della responsabilità morale del Professore in un caso piuttosto che in un altro.

²⁶⁵ DERK PEREBOOM, *Free Will, Agency and Meaning*, op. cit., p. 79.

Si legge: “Il determinismo fisicalistico è vero, vale a dire ogni cosa nell’universo è fisica e ogni cosa che accade è causalmente determinata in virtù degli stati passati dell’universo, in congiunzione con le leggi di natura. Plum è un normale essere umano, cresciuto in normali circostanze, e i suoi processi di ragionamento, frequentemente ma non sempre, sono egoistici (come nei Casi 1-3). La sua decisione di uccidere la signora White discende dai suoi processi deliberativi, che rispondono a ragioni, ed egli ha gli specifici desideri di primo e secondo ordine. La realizzazione cerebrale del suo processo di ragionamento e della sua decisione è quella di cogliere le ragioni morali, di applicarle e di regolare in base ad esse il proprio comportamento, e non è per un desiderio irresistibile che egli uccide la signora White”, in *Quanto siamo responsabili?*, traduzione di ANDREA LAVAZZA, *Quanto siamo responsabili?*, op. cit., p. 135.

Da ciò Pereboom vuole evidenziare che il determinismo causale, legato alle circostanze o ad altri agenti, non è una condizione sufficiente e necessaria per stabilire la responsabilità morale dell'agente, il quale, *de facto*, sarebbe giustificato nelle sue scelte per quelle infinite catene causali “dietro” l'agentività che minano, alla fonte, la dicotomia determinismo/responsabilità come i compatibilisti, al contrario, vorrebbero dimostrare. In questo modo:

[...] *the manipulation argument aims to persuade the natural compatibilist and the agnostic that their resistance to incompatibilism is best given up*²⁶⁶,

evidenziando, come sostiene Pereboom, che non c'è differenza fra i *manipulation cases* e gli agenti in ordinarie situazioni deterministiche.

L'argomento del filosofo statunitense, come è facile intuire, ha suscitato non pochi dibattiti sulla plausibilità del paragone fra la manipolazione (reale o presunta) ed il compatibilismo, adducendo, *contra Pereboom*, il fatto che in tutti i casi proposti la migliore spiegazione al comportamento di Plum derivi proprio dalla manipolazione cui – direttamente o indirettamente – è soggetto, indipendentemente che si cali l'agente in un contesto ordinario o fantasioso.

Pereboom, però, fa notare che la spiegazione della mancanza di responsabilità del Professore, derivante solo dalla presenza di terzi che potrebbero influenzare il suo processo deliberativo, è, di per sé, una spiegazione che non pone a fuoco il motivo centrale per cui intuitivamente il compatibilista è portato a credere alla deresponsabilizzazione di Plum nella decisione di uccidere la Sig.ra White.

Se, infatti, lo scenario dei quattro esempi riportasse, quale premessa, la verità dell'incompatibilismo duro (assenza, quindi, di qualsiasi nesso causale), probabilmente la percezione su Plum cambierebbe.

Alfred Mele, in uno scritto del 2007²⁶⁷, risponde a Pereboom, sostenendo che:

[...] *these imagined data do help in the case of compatibilists who have “the intuition” that Plum is not morally responsible for the killing in one or more of Pereboom’s stories. Other things being equal, given that that intuition is generated in them both by a deterministic case of manipulation and by an indeterministic analogue of that case, but not by a comparable deterministic case involving no manipulation, the manipulation featured in the*

²⁶⁶ DERK PEREBOOM, *Free Will, Agency and Meaning*, op. cit., p. 81.

²⁶⁷ Cfr. ALFRED MELE, *Free Will and Luck: Replies to My Critics*, in “Philosophical Explorations”, n. 10, 2007, pp. 195-210.

*relevant cases is a better candidate for an explanation of their “nonresponsibility” intuitions about these cases than determinism is. The judgment that the determinism in a deterministic manipulation case provides the best explanation of these compatibilists’ “nonresponsibility” intuitions about it is silent on the analogous indeterministic case, and it yields the prediction that these compatibilists will have “the intuition” that Plum is not morally responsible for the killing in any straightforward deterministic story I might tell that involves no manipulation and no monkey business of any kind. Obviously, the imagined data do not warrant that prediction*²⁶⁸.

Da come si evince dalle parole di Mele, prese a “simbolo” di una intera corrente di pensiero che sposa una linea *hard* del compatibilismo, l’intuizione della mancanza di responsabilità nelle azioni di Plum è indotta dal contesto: dall’idea, cioè, che sussista una qualche forma di manipolazione che, inevitabilmente, conduce il lettore *prima facie* a sgravarlo da qualsiasi partecipazione immorale, derivante da Plum stesso, sia per le decisioni sia per le azioni. Non a caso, nel passaggio sopra riportato, Mele sostiene che se si proponessero gli stessi quattro casi con la variante indeterministica, l’impressione finale sarebbe la medesima: Plum non sarebbe ritenuto moralmente responsabile.

In risposta a Mele, Pereboom scrive:

*[...] manipulation per se cannot explain Plum’s non-responsibility in Cases 1 and 2, and neither can bypassing. Causal determination remains a candidate for the best explanation of his non-responsibility in these examples, and this does not conflict with manipulation absent causal determinism yielding non-responsibility in indeterministic cases. Moreover, if the incompatibilist’s concern is to be engaged, in reflective assessment of the manipulation examples the initial intuition that agents can be morally responsible in ordinary deterministic situations should not be granted whatever weight we would unreflectively assign to it. While the objections Mele raises need to be taken into consideration when assessing the strength of the four-case argument, my sense is that they do not undercut its force*²⁶⁹.

Al di là delle confutazioni reciproche che continuano a susseguirsi fra l’una e l’altra parte, un aspetto non trascurabile è il senso che si applica ai concetti di “azione” e di “libero arbitrio”. Se, ad esempio, il già citato Chisholm considera l’azione e la sua origine alla stregua del

²⁶⁸ *IDEM*, pp. 204-205.

²⁶⁹ DERK PEREBOOM, *Free Will, Agency and Meaning in Life*, op. cit., pp. 88-89.

primo motore immobile di aristotelica memoria (nelle modalità e nelle caratteristiche precedentemente riportate), Pereboom, convinto sostenitore dell'incompatibilismo, ribadisce che:

*I understand 'action' to denote not only intentional bodily movements, but also purely mental items such as intention-formations and decisions [...]*²⁷⁰.

Si chiarisce, allora, il motivo per cui il centro della discordia del dibattito fra compatibilisti ed indeterministi sia legato al problema del controllo dell'azione: i primi, come evidenzia Chisholm, hanno bisogno di spiegare, metafisicamente e fenomenologicamente, come avviene un'azione "libera" all'interno delle rigide catene causali; i secondi, scevri dal principio di causalità, hanno bisogno, al contrario, di spiegare come l'agente abbia il controllo della propria azione, adducendo come spiegazione le proprie intenzioni o volizioni quali "cause non causate" dell'azione stessa.

Il problema del libero arbitrio nell'*hard incompatibilism* è legato, quindi, alla contingenza cui il mondo è sottoposto: escludendo *in toto* qualsiasi forma di principio di causa o di nesso causale, gli indeterministi radicali non sono in grado di spiegare il modo in cui l'agente possa avere il diretto controllo dell'azione. Proprio per tali difficoltà, all'interno della grande famiglia indeterministica, sono nate delle correnti che hanno introdotto un principio causale *sui generis*, denominate *indeterminismo causale* o *incompatibilismo libertario*.

A differenza dei compatibilisti, i libertari sostengono un concetto di libertà molto vicino a quello del senso comune, in quanto si considera libero colui che, essendo in grado di autodeterminarsi, non ha soltanto il controllo diretto della sua azione, ma ha *de facto* una possibilità di scelta fra corsi d'azione alternativi. Ciò sembrerebbe maggiormente in grado di garantire lo spazio logico e pratico della libertà umana, grazie alla quale all'agente sarebbe garantita la possibilità genuina della scelta, discendente direttamente dalla sua stessa volontà; l'importanza centrale delle possibili scelte alternative da parte dell'agente non vengono intese come una mera possibilità teorica della nozione del "poter fare altrimenti", bensì come la capacità pratica e *categorica* dell'agente di operare una scelta – secondo le proprie intenzioni – del corso d'azione da seguire.

Tutto ciò è possibile nel mondo incompatibilista proprio perché non vige alcun nesso necessario che conduce e pilota inevitabilmente il suo fare e la volontà dell'agente in un dato corso d'azione: mancando, vale a dire, il rigido determinismo causale, nel mondo libertario

²⁷⁰ *IDEM*, p. 89.

sembra che il libero arbitrio possa trovare la sua massima esplicazione. *Hic et nunc*, in maniera *categorica*, l'agente è in grado di compiere una scelta attraverso la quale percorrere uno fra gli infiniti corsi d'azione possibili. La scelta non diviene una possibilità in linea di principio, ma un fatto accreditato che certifica la sua esistenza.

La tesi incompatibilistica a differenza di quella compatibilista, non sembrerebbe, allora, incorrere in gravi problemi concettuali che impediscano una spiegazione coerente e scevra da difficoltà del binomio libero arbitrio/mondo, ma ad un'analisi attenta si comprende che proprio il rifiuto del nesso causale costituisce un grave problema; in che modo, infatti, l'agente avrebbe la capacità del controllo della sua azione? In un mondo in cui vige l'indeterminato, in che modo si può esser certi che la conseguenza dell'azione operata dall'agente sia opera sua e che non sia, invece, la conseguenza di un altro corso d'azione intrapreso da altri?

Una possibile soluzione secondo gli indeterministi libertari è quella di introdurre il concetto di *causalità indeterministica*: un nesso causale che non necessiterebbe gli effetti, ma ne accrescerebbe solo la probabilità che si verificano. Essa, a differenza del nesso causale comunemente conosciuto, avrebbe una natura diversa ed agirebbe nel momento preciso in cui l'agente si trovi a dover prendere una decisione: in quell'attimo, stando ai libertari, il nesso causale ordinario cesserebbe di esistere, dando spazio alla causalità indeterministica, che garantirebbe non solo la possibilità di scegliere fra diversi corsi d'azione, ma anche il controllo dell'azione da parte dell'agente. Per quanto di natura diversa, infatti, la causalità indeterministica rimane pur sempre causalità, di conseguenza l'agire dell'agente causerà *probabilisticamente* come effetto la sua azione.

C'è da dire che tale causalità, oltre ad essere decisamente fantasiosa, sembrerebbe essere anche poco credibile.

Uno dei problemi nella sua introduzione nel dibattito sull'assenza eventuale di una rigida necessitazione concerne, innanzitutto, la sua natura: che tipo di causalità potrà mai essere? Come può, infatti, una causalità non determinarne l'effetto? Per definizione essa è il legame relazionale che intercorre fra un prima ed un dopo (causa-effetto); porre un velo probabilistico fra essi potrebbe, in qualche modo, determinare la salvaguardia della libertà nell'indeterminismo, garantendo il controllo dell'azione, ma non salvaguarderebbe la risoluzione di quegli stessi problemi sostanziali legati proprio all'utilizzo di una causalità la cui *ratio* non sia quella della causalità stessa.

Un altro problema in cui incorre l'indeterminismo è quello relativo alla localizzazione indeterministica, ossia

*non è affatto chiaro dove debba avvenire la rottura della catena causale deterministica che porta all'azione*²⁷¹.

Il problema è l'individuazione di una falla indeterministica nella catena deterministica che sorregge il mondo tale da garantire all'agente la scelta dell'azione da compiere in completa libertà: dove, cioè, collocare l'indeterminismo?

La difficoltà a spiegare in maniera convincente il modo in cui l'agente indeterminista possa effettivamente avere il pieno controllo della propria azione, sia attraverso la causalità probabilistica sia attraverso la faglia indeterministica, è stata, forse, "risolta" all'interno di una posizione che promette di superare le difficoltà proprie della tesi di partenza, mostrando la perfetta congiunzione fra libertà e causalità: l'*Agent Causation*.

Riprendendo in parte la tesi di Thomas Reid, l'*Agent Causation* o il *source incompatibilism* sostengono che il fattore di controllo delle azioni è rappresentato dallo stesso agente, il quale, oltre ad essere in grado di autodeterminarsi, origina, grazie alla propria volontà, nuove catene causali: in tal modo non solo si garantirebbe il pieno controllo dell'azione all'agente, ma si lascerebbe anche lo spazio ai possibili corsi d'azione alternativi.

La difficoltà centrale di tale posizione, però, rimane quella già incontrata a proposito della causalità indeterministica: l'inconcludente spiegazione di che cosa significhi l'origine di una nuova causazione da parte degli agenti. Riprendendo in parte la tesi sull'uomo come *primo motore immobile* aristotelico, i libertari sostengono che quando si procede nell'analisi a ritroso delle cause degli eventi si giungerà, inevitabilmente, ad un punto in cui si dovrà sostenere che solo l'uomo è l'origine di una infinita catena di eventi: proprio per questo l'azione da esso compiuta può dirsi libera. È come se le azioni dell'agente avessero una natura intenzionale e non causale.

Tale posizione non è, però, immune da critiche e perplessità: come spiegare, ad esempio, l'idea che l'uomo – e solamente lui – sia un'eccezione nell'ordine del mondo, rappresentando un motore immobile capace di dar avvio ad una serie infinita di effetti? Ciò, ad onore del vero, sembra non solo una mossa poco felice per ovviare alle difficoltà dell'indeterminismo, ma anche una posizione che nega *in toto* le stesse premesse da cui parte: se, ancora una volta, si sta parlando di un mondo governato dalla contingenza, perché postulare un mondo sorretto dal principio di causa tipico del determinismo? Inoltre, tale posizione cade nel dualismo cartesiano: in che modo questo motore immobile (o coscienza o mente) riuscirebbe ad

²⁷¹ MARIO DE CARO, *Il libero arbitrio. Una introduzione*, op. cit., p. 51.

autodeterminarsi? Significherebbe presupporre che esso sia costituito da una materia/sostanza diversa rispetto al mondo.

Così come quanto avvenuto per il compatibilismo, si è recentemente sostenuto che lo stallo teorico dell'*Agent Causation* è causato dalle diverse interpretazioni date alle parole “poter fare”, “abilità” “avrebbe potuto fare altrimenti” e “potere dell’agente” – nel senso di *power* – che non hanno fatto altro che allontanare l’indeterminismo (in qualsiasi accezione lo si consideri) dalla risoluzione del grave problema interno alla sua stessa tesi. Si dovrebbe ripartire da tale consapevolezza per poi concentrarsi, successivamente, sulla ridefinizione del libero arbitrio. Parlare di *free will* significa considerarlo come:

*the power to be the ultimate creator and sustainer of one’s own ends or purpose*²⁷².

Affinché, dunque, un agente possa esser ritenuto in possesso del libero arbitrio, egli deve essere in grado di mostrarsi come autore e matrice dei propri desideri e possedere, in termini *categorici*, la possibilità di fare altrimenti. La scelta, però, da sola, non è in grado di mostrare la responsabilità morale di un agente, poiché

*to be ultimately responsible for an action, an agent must be responsible for anything that is a sufficient reason (condition, cause or motive) for the action’s occurring*²⁷³.

Come sostiene Robert Kane, nell’argomento *Ultimate Responsibility*, l’agente per dichiararsi “essenzialmente responsabile” dovrà avere non solo una o più ragioni sufficienti che giustificano l’azione che sta svolgendo, ma anche tutte quelle esperienze della sua vita passata che lo hanno condotto ad essere ciò che è:

*For UR does require that we could have done otherwise with respect to some acts in our past life histories by which we formed our present characters. I call these ‘self-forming actions’ or SFAs*²⁷⁴.

²⁷² Si legge: “Il potere di essere il creatore ultimo e sostenitore dei propri intenti o dei propri propositi”, ROBERT KANE, *Free Will: new direction for an ancient problem*, in “Journal of Philosophy”, 96, pp. 222-223.

²⁷³ Si legge: “Per essere essenzialmente responsabili di un’azione, un agente deve essere responsabile di qualsiasi motivo che sia una ragione sufficiente (condizioni, cause o motivi) per le azioni che si verificano”, *IVI*.

²⁷⁴ Si legge: “Per essere essenzialmente responsabili si richiede che potremmo fare altrimenti nel rispetto di alcuni avvenimenti della nostra vita passata che hanno formato la nostra personalità. Chiamerò queste pietre fondanti il nostro Io come ‘azioni autoformanti’ o SFAs”, *IDEM*, p. 223.

Le azioni “autoformanti” sono ciò che consente all’agente, non solo di essere dichiarato responsabile per le proprie scelte, anche di possedere l’alternativa alla scelta da compiere, rimanendo sempre in accordo a quella che è la personalità ed il carattere dell’agente stesso. Esse potrebbero rappresentare l’alternativa al concetto della causalità indeterministica per superare e risolvere il problema del controllo dell’azione da parte dell’agente nel mondo incompatibilista. Si consideri, a tal proposito, l’esempio di Martin Lutero: quando si distaccò definitivamente dalla Chiesa di Roma è ipotizzabile che egli sostenesse che “non era in grado di fare altrimenti” e, in effetti, Lutero per Kane diceva il vero. Dato il suo carattere, la sua personalità e le sue ragioni, egli realmente non avrebbe potuto agire in maniera diversa da come di fatto agì proprio perché nessuno compie atti che contrastino il proprio Io. L’agire secondo la propria volontà significa, allora, agire *con e tramite* il libero arbitrio, in quanto l’uomo è frutto delle scelte “autoformanti” che lo hanno “costruito” e che lo rendono responsabile per ciò che fa. Scrive Kane:

*If none of our acts were self-forming in this way, we would not be ultimately responsible for anything we did*²⁷⁵.

A tal proposito, riprendendo la *Compatibility Question*, Kane sostiene che, introducendo l’argomento dell’*Ultimate Responsibility*, si mostrerebbe non solo il senso in cui le azioni dell’agente siano determinate, ma, soprattutto, in che modo possa esser ritenuto responsabile, ovviando alle gravi difficoltà cui la famiglia incompatibilista incorre.

Le SFAs accadono e si rivelano tali in quei momenti di scelta sofferta fra due alternative, nel tentativo di dominare la tentazione di fare un qualcosa anche se si desidera, fortemente, fare altro. Kane suggerisce che tale struggimento riflette in regioni cerebrali appropriate un “caos indeterministico”, un *stirring up of chaos*²⁷⁶, percepite dall’agente come “anima in movimento” quando, invero, si tratta di un’esperienza formativa (SFAs) che rispecchia un preciso percorso neurale. Tale esperienza l’agente la percepisce come:

*opening of a window of opportunity that temporarily screen off complete determination by influence of the past*²⁷⁷.

²⁷⁵ Si legge: “Se nessuno dei nostri atti è autoformante nel modo in cui si è detto, allora non potremmo essere essenzialmente responsabili per qualcosa che abbiamo compiuto”, *IBIDEM*, p. 225.

²⁷⁶ *IDEM*, p. 228.

²⁷⁷ Si legge: “L’apertura di una finestra di possibilità che schermo temporalmente la determinazione ad agire in un certo modo secondo l’influenza del passato”, *IVI*.

Certamente decidere in questa situazione di scelta struggente fra due possibili alternative non significa che la scelta definitiva sia in qualche misura *determinata* proprio perché il “salto neurale indeterministico” – a fondamento del processo neurale – esclude qualsiasi forma di causazione possibile dalla volontà futura dell’agente. Tale “finestra di possibilità”, che si aprirebbe all’agente nell’atto della scelta, rappresenta la falla indeterministica in cui “collocare” l’indeterminismo: quello spazio logico entro cui la causalità diverrebbe mera probabilità.

Il processo indeterministico a fondamento delle SFAs, inoltre, rappresenta non solo la difficoltà della scelta, ma ciò che renderà l’agente prettamente responsabile delle proprie azioni: contrariamente a quanti sarebbero portati a pensare che parlare di “scelta indeterministica” non sarebbe sinonimo di “scelta casuale, capricciosa o accidentale”, ciò significherebbe compiere scelte che sono state *determinate* da quel “rumore indeterministico” (*the indeterministic noise*²⁷⁸) che renderebbe gli agenti moralmente responsabili. Appagate tali condizioni, le SFAs sono soddisfatte e gli agenti sono in tal modo in grado di esercitare il pieno controllo sulle loro vite *hic et nunc*. Kane chiama tale condizione *plural voluntary control*²⁷⁹, secondo cui l’agente può esser ritenuto moralmente responsabile solo se nei fatti egli riesce a compiere ciò che era nelle sue intenzioni compiere: le intenzioni, in altre parole, non rendono alcuno responsabile degli atti *potenziali che avrebbe voluto compiere*.

Infine, contro coloro che hanno ipotizzato l’esistenza di *extra factor strategies*, Kane sostiene che, stando alla sua posizione, non vi è alcun motivo di ricorrere a fattori metafisici che giustifichino la coesistenza di libero arbitrio e indeterminismo, in quanto il primo è spiegato in termini di sforzi e di processi neurali che spetterà alle ricerche future individuare.

Il richiamo di Kane alla “scienza del cervello” nello studio del libero arbitrio non è casuale: negli ultimi anni, dato lo stallo teorico fra le due grandi posizioni in gioco, si è fatta strada l’idea di poter spiegare il *free will* tramite l’utilizzo delle nuove strumentazioni scientifiche in grado di vagliare, captare e studiare (anche se ancora in maniera del tutto limitata) l’attività cerebrale. L’idea che sta prendendo corpo è che il libero arbitrio sia il risultato di una particolare attivazione cerebrale tale anche da arginare, per pochi millisecondi (mms), la presenza della coscienza di cui il lavoro pionieristico del neurofisiologo Benjamin Libet costituisce un esempio.

²⁷⁸ *IBIDEM*, p. 229.

²⁷⁹ *IDEM*, p. 10.

Capitolo VI

Libero arbitrio e neuroscienze

6.1 L'esperimento di Benjamin Libet: constatazioni e perplessità

Che il dibattito sul libero arbitrio stia subendo il fascino delle neuroscienze è un dato incontrovertibile; ma che si possa parlare di vera e propria *rivoluzione* in ambito metafisico è tutto da verificare.

Certo è che l'idea di indagare fisiologicamente il *free will* non è nata dal nulla: agli inizi degli anni '80, infatti, contravvenendo agli scettici più convinti ed anticipando di trenta anni il cambiamento storico-culturale di queste ultime decadi, Benjamin Libet, sulla scia degli studi condotti da Hans Kornhuber²⁸⁰, tentò la verifica ipotetica del concetto di libero arbitrio attraverso i metodi consoni alla scienza galileiana, sperando di capire se la scelta fosse “opera” del cervello, prima, della coscienza, poi, o il contrario. Con “la tecnologia” a

²⁸⁰ Ad onor del vero, colui che per primo intentò lo studio del libero arbitrio tramite l'utilizzo delle neuroscienze fu Hans H. Kornhuber agli inizi degli anni Sessanta del Novecento. L'esperimento era pressoché simile a quello conosciuto di Libet, tanto che anche lo stesso Kornhuber giunse all'identificazione del *potenziale di preparazione*, successivamente battezzato quale potenziale di prontezza. Di solito l'esperimento di Kornhuber non viene nominato nel dibattito sul libero arbitrio in quanto risulta del tutto inconcludente: lo scienziato, infatti, non riuscì a calcolare l'eventuale scarto temporale che intercorreva fra l'attivazione dell'area supplementare motoria (dedita al movimento) ed il presentarsi della volontà nel soggetto di flettere il dito. Solo dopo una ventina d'anni Libet riuscì a sopperire a tale mancanza ed è proprio per questo motivo che il lavoro di Kornhuber è caduto nel dimenticatoio.

Per ulteriori approfondimenti, cfr. FRANCESCA TEMPIA, *Decisioni libere*, in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, op. cit.

disposizione in quegli anni, Libet, partendo direttamente dalla (neuro) fisiologia dell'intero encefalo, ingegnò l'esperimento sui cui dati si sta ancora dibattendo.

Procedette in tal modo. Applicò ad un volontario degli elettrodi nelle aree supplementari motorie (responsabili del movimento) ed un elettroencefalogramma (EEG), capace di visualizzare l'attività elettrica del cervello. Fu chiesto, poi, al soggetto di flettere il dito quando ne avesse avvertito il desiderio, nonché di ricordarsi il momento esatto in cui l'avesse desiderato tanto che per una maggiore precisione della raccolta dati gli fu posto innanzi un orologio luminoso: la posizione indicata dalle lancette sarebbe stata in grado di individuare l'istante preciso nel quale si sarebbe formata l'intenzione consapevole dell'agente.

Grazie alla ripetizione costante dell'esperimento, Libet, a differenza di Kornhuber, riuscì a quantizzare ed a misurare i dati ricavati in termini di tempo, sostenendo, in ultimo, che era alquanto probabile l'esistenza di un minimo scarto temporale fra l'esecuzione del movimento richiesto e la consapevolezza del soggetto di compierlo, quantizzabile intorno ai 250mms; tempo dedotto dalla constatazione di due "eventi" fortemente significativi:

- gli agenti avvertivano il "desiderio" di flettere "liberamente" la falange circa 200mms prima del compimento dell'azione;
- intorno ai 500mms prima dello stesso, si registrava un'intensa attività cerebrale, chiamata da Libet stesso *readiness potential*²⁸¹.

Tutto ciò suscitò non poche riflessioni; una fra queste fu l'interrogativo di capire che cosa mai significasse la primarietà dell'attività cerebrale su quella coscienziale: *punctum dolens* dell'attuale neuroetica e non solo.

La difficoltà dell'esperimento di Libet non consistette tanto nella sua esecuzione, ma nell'interpretazione stessa dei dati ricavati o, meglio, nella loro "indeterminatezza", poiché se l'intento fosse stato, come detto, quello di prescindere dalle classiche e *sempiternae* diatribe sul *free will*, fallì miseramente. I dati, infatti, potevano essere indirizzati e, quindi, interpretati, o *contra* la tesi deterministica o *contra* quella indeterministica, contemporaneamente.

Nonostante per la prima il problema della volontà costituisca un aspetto marginale²⁸², una particolare interpretazione dell'esperimento avrebbe potuto evidenziare che l'assenza di qualsiasi controllo da parte dell'agente, teoricamente, dovrebbe avere quale certezza, tanto

²⁸¹ *IDEM*, p. 71.

²⁸² Si ricordi che per il determinismo *tout court* l'azione è strettamente necessitata dalle infinite catene di cause e di concause che l'hanno generate, per cui, come già spiegato al capitolo VI, la volontarietà dell'agente diviene un fatto marginale o, per lo meno, acquisisce importanza solo all'interno delle correnti deterministiche che si interrogano sul libero arbitrio.

teoretica quanto fenomenologica, di essere la causa di un effetto e ciò avrebbe invalidato l'intera *ratio* della stessa posizione deterministica.

Tradotto in termini "sostanziali", i dati del neurofisiologo mostrarono la mancanza "sempre e comunque" del soggetto in se stesso: egli, cioè, sembrava essere coscientemente "in ritardo" rispetto all'esecuzione materiale del movimento a cui precedeva, fisiologicamente, un'intensa attività cerebrale (il *potenziale di prontezza*), che preparava il dito alla flessione. Se, cioè, il rigido nesso causale è il solo che specifica un dato corso di azione, scoprire che l'azione dell'agente sia determinata non da questo, bensì da specifiche strutture sinaptiche, significherebbe scoprire che la causa di un'azione potrebbe non essere il soggetto stesso; che fine avrebbe fatto, allora, il principio causale? Un dato che, appunto, sembrava esser suggerito dall'esperimento di Libet.

Una seconda interpretazione fu fornita dalla tesi indeterministica in cui, addirittura, i dati del neurofisiologo sembravano invalidare completamente i presupposti cardine della libertà positiva: la mancanza di controllo diretto sull'effetto (probabilistico o reale) e l'impossibilità, da parte dell'agente, di scegliere percorsi d'azione alternativi; ossia, compiere scelte diverse rispetto a quelle dettate dal proprio cervello.

Contrariamente a quanto molti avrebbero potuto pensare, si tornava, con gli esperimenti di Libet, al classico problema affrontato dalla filosofia moderna sul rapporto filosofia/scienza: il metodo d'indagine sulla natura sfidava, forse, le intuizioni più radicate dell'essere umano e della sua presunta "libertà", rivolgendosi, stavolta come allora, all'antico problema di comprensione di se stessi. L'imbarazzo del neurofisiologo significò anche che la sola Scienza, per quanto "metodologicamente oggettiva", non avrebbe, per sua natura, saputo spiegare le difficoltà cui il concetto di *free will* incorre tanto che ciò che stupisce ancora oggi è la spiegazione fornita da Libet stesso sul modo in cui interpretare i dati ricavati.

Secondo il neurofisiologo, infatti, i veri determinanti causali non potevano risiedere nella "libera scelta"²⁸³ dell'agente, ma negli eventi neuronali precedenti che, generando il "potenziale di prontezza" e la successiva attività motoria, rendevano il soggetto inconsapevole, persino a se stesso, della sua stessa volontà. In questi termini, i 250mms rappresentavano, agli occhi di Libet, l'incapacità del controllo conscio dell'agente del proprio agire e, contestualmente, tentava di salvare dalla deriva nichilista il libero volere, che, a quel punto, diveniva mera illusione: i determinanti causali dell'azione che, da soli, avrebbero reso il soggetto incapace di essere parte attiva al "processo della scelta", sarebbero diventati parte

²⁸³ Per ulteriori informazioni per l'identificazione della *scelta* a livello fenomenologico, cfr. ROBERTA DE MONTICELLI, *Che cos'è una scelta?*, in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, op. cit.

passiva dello stesso attraverso ciò che Libet chiamò *libertà di veto*; che avrebbe permesso all'agente la possibilità di interrompere quella serie di cause e concause a livello sinaptico (che porterebbero, inevitabilmente, all'azione), essendo divenuto conscio della sua stessa volontà.

La consapevolezza, in tal modo, sembrava seguire un dato evento cerebrale per poi “apparire” solo nel mentre l'azione si compiva (il significato dello scarto temporale dei 250mms), ipotizzando non solo che la coscienza non fosse la causa delle scelte compiute, ma che essa non avesse un ruolo determinante nella libertà d'azione. Libet, così, salvava il concetto di libero arbitrio, il controllo diretto sull'azione e forniva, seppur dualisticamente, una spiegazione del fenomeno coscienza/libertà/mondo.

Come era facilmente prevedibile, l'esperimento di Libet è stato fortemente criticato, sia metodologicamente sia concettualmente.

Per quanto riguarda il primo aspetto, fra le diverse confutazioni avanzate negli anni, è rilevante notare che, per quanto un EEC possa misurare in “maniera certa” l'attività cerebrale, è altrettanto vero che esso non è sufficiente a dirimere controversie neuroscientificamente più rilevanti sull'attivazione di un'area o di un gruppo di neuroni prima di un altro. Sarebbe complicato anche oggi, a distanza di circa quaranta anni, rispondere a tali controversie pur facendo rimanere *in auge* il problema sul loro significato.

Degna di nota è anche l'obiezione sulla scelta “libera” di muovere una falange ad arbitrio del candidato: in che modo, infatti, può essere annoverata quale “scelta libera” un'azione “programmata” scientificamente all'interno di un contesto già prestabilito? Il problema, e qui ci si sposta nell'ambito concettuale o interpretativo dell'esperimento libetiano, è il significato stesso di libertà/libero arbitrio.

Il neurofisiologo avanzò la “libertà di veto” per salvare la casualità sostanziale che investirebbe tanto l'uomo quanto il mondo; ciò significherebbe, di contro, che l'agente non solo non avrebbe “libertà d'azione”(in qualsiasi modo la si interpreti), ma gli mancherebbe anche il potere decisionale “in positivo”: non avviare *ex novo* un'azione, ma solo la scelta di interromperla oppure no. Ad aggravare ulteriormente il quadro catastrofico che una certa lettura di Libet suggerisce, si aggiungerebbe il dubbio sulla natura dei determinanti causali: se essi sono inconsci, perché non si dovrebbe ipotizzare che quelli della volontà “in negativo” siano consapevoli? Che cosa, cioè, distinguerebbe, sostanzialmente, le prime dalle seconde? È esplicativa, a tal proposito, la riflessione di De Caro:

[...] è semplicemente arbitrario da parte di Libet e dei suoi continuatori sezionare la catena causale fermandosi al momento dell'incremento del potenziale di prontezza senza risalire al momento precedente in cui i soggetti (cosciamente, per quanto ne sappiamo) decidono di aderire alla richiesta degli sperimentatori. La tesi che i determinanti causali dell'azione di flettere il dito abbiano soltanto carattere inconscio appare dunque mal fondata²⁸⁴.

Ciò che, in generale, emerge dai disappunti e dalle obiezioni avanzate è il ruolo svolto – seppur con qualche millisecondo di ritardo – dalla coscienza. Resta quest'ultima, infatti, al centro del dibattito, lasciando trasparire sullo sfondo che, in definitiva, la coscienza sia il frutto di una complessa ed intricata struttura cerebrale che per “apparire” necessita di tempo:

*Does this negate freedom of choice, reducing it to the status of an illusion? Does mere fact that brain activity goes on outside of our awareness mean that we are automatons that live with the pleasant delusion that we are free?*²⁸⁵.

Il problema nel problema, allora, è come spiegare il dato ricavato da Libet: che ruolo svolge la coscienza nei processi decisionali? E come mai un approccio naturalizzato al “fenomeno coscienza” comporterebbe il diniego della libertà di scelta?

Interrogativi che si pone lo stesso Donald:

*Questa negazione della libertà di scelta, conduce all'illusione del libero arbitrio? La mera attività cerebrale al di fuori della consapevolezza significa [forse, n.d.r.] che siamo automatismi che vivono con la costante delusione che siamo liberi?*²⁸⁶.

I 250ms lascerebbero suggerire una separazione fra il fisico ed il mentale, secondo cui il libero arbitrio altro non sarebbe che l'insieme degli eventi neuronali che lo stesso soggetto subirebbe²⁸⁷; l'idea emergente è quella ricalcata dalla distinzione cartesiana fra *res cogitans* e *res extensa* in cui lo scarto temporale quantizzato dal neurofisiologo sarebbe “la prova” di una

²⁸⁴ MARIO DE CARO, *Libero arbitrio e neuroscienze*, in ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Neuroetica*, Il Mulino, Bologna, 2011, p. 75.

²⁸⁵ MERLIN DONALD, *Consciousness and the Freedom of Act*, in ROY F. BAUMEISTER, ALFRED R. MELE & KATHLEEN D. VOHS (A CURA DI), *Free Will and Consciousness: How Might They Work?*, Oxford University Press, Oxford, 2010, p. 11-12.

²⁸⁶ *IDEM*, p. 13.

²⁸⁷ Cfr. JOHN SEARLE, *Libertà e neurobiologia*, Bruno Mondadori, Milano, 2005.

netta separazione fra l'attività neurale avviata dal cervello e la consapevolezza dell'agente, sempre "in ritardo" rispetto alla prima.

Etichettare con "monismo" o "dualismo" le diverse posizioni che si sono succedute da Cartesio ad oggi nella tradizione della filosofia della mente sarebbe, certamente, riduttivo²⁸⁸; ma ciò non toglie che "in una selva di *ismi*"²⁸⁹ essi sintetizzino, immediatamente, il cuore del problema. Che cosa significherebbe e che conseguenze ci sarebbero per la libertà umana discendere da una coscienza "fatta di neuroni" o da una "sostanza pensante" distinta dal "mero" cervello neurale?

Dietro una domanda apparentemente semplice, ne consegue l'esistenza o l'inesistenza della stessa all'interno del contesto entro cui si dibatte, in definitiva, l'esistenza o l'inesistenza dello stesso libero arbitrio.

Uno fra gli attori protagonisti del palcoscenico dialettico suscitato dalla natura della coscienza ed aspro critico nei confronti della posizione dualistica sposata da Libet è Daniel Dennett, filosofo americano. Egli, infatti, sostiene che Libet stesso ha indotto a credere che il "vuoto morale" (da molti così definito) dei 250mms indichi il tempo necessario all'uomo per venire a conoscenza di un'azione che il cervello ha intrapreso senza "consultare prima" la consapevolezza cosciente che abita il medesimo corpo. Il problema, vale a dire, si trova nell'interpretazione dei dati forniti dallo stesso neurofisiologo:

[...] *quando pensate di stare decidendo, voi in realtà state solo guardando passivamente una sorta di registrazione interna trasmessa in ritardo (i 250mms di ritardo) della vera decisione, che è avvenuta in modo inconscio nel vostro cervello molto prima che 'a voi venisse voglia' di muovere il dito*²⁹⁰.

In *Coscienza*²⁹¹, Dennett sostiene che laddove si sposasse il dualismo di sostanza cartesiano, la mente potrebbe essere paragonata ad un grande Quartier Generale, in cui diversi "omini" verdi eseguano le decisioni (coscienti) da far poi eseguire alla macchina corpo. In tal senso, la coscienza diverrebbe mera spettatrice dell'attività cerebrale come se il lavoro di preparazione motoria del soggetto per compiere un'azione non fosse una faccenda che possa interessare la stessa consapevolezza.

²⁸⁸ Cfr. MICHELE DI FRANCESCO, *Una introduzione alla filosofia della mente*, op. cit.

²⁸⁹ *IBIDEM*, p. 17.

²⁹⁰ DANIEL DENNETT, *L'evoluzione della libertà*, Raffaello Cortina Editore, Milano, 2004, p. 94.

²⁹¹ Cfr. DANIEL DENNETT, *Coscienza. Che cosa è*, Editore Laterza, Roma-Bari, 2009.

Dennett chiama questa metafora “il Teatro Cartesiano”, in cui lo spettatore è personificato dalla mente che assisterebbe come in un cinema alla visione del flusso di comandi che l’attività cerebrale avvierebbe indipendentemente sia dalla volontà cosciente sia dalla mente stessa. È come se l’attività cerebrale fosse intrapresa da diversi “omini” verdi di cui sarebbe impossibile intravedere lo specifico referente: ciò, allora, comporterebbe la vanificazione del concetto di responsabilità morale, in quanto essendo il soggetto “sempre in ritardo” rispetto alle decisioni prese da quegli stessi omini verdi, in che modo potrebbe dirsi responsabile delle proprie azioni?

Questo comporterebbe, allora, considerare mera illusione il libero arbitrio?

Secondo Dennett assolutamente no. Perché mai, infatti, la “funzione coscienza” dovrebbe attendere l’arrivo dell’informazione per fare la sua comparsa nel Quartier Generale? Perché non considerare che la medesima coscienza possa aver dato inizio *inconsiamente* all’avvio dell’azione? Il punto, cioè, che Dennett vuole sottolineare è il cercare di capire i motivi per cui si debba avere così tanta riluttanza a concepire la mente come un’intensa attività cerebrale: mente e cervello come la medesima “sostanza”.

Egli sostiene che l’interpretazione “corretta” dei dati forniti da Libet è che i processi decisionali coscienti richiedono del tempo: il tempo necessario affinché il cervello sia in grado di dar vita alla complessa “funzione coscienza”. Di conseguenza, essendo *mind* e *brain* fatti della medesima sostanza, il libero arbitrio sarebbe, per il filosofo americano, l’insieme di tutti quei meccanismi cerebrali che sottostanno alla formazione della consapevolezza così che una volta accettata l’idea che vi sia solo la strabiliante complessità cerebrale, si abbandonerà anche l’illusione molto attraente del “Teatro Cartesiano”. L’intento di Dennett, quindi, è quello di avviare un’interpretazione naturalizzata, come detto, dell’esperimento degli anni ’80, escludendo, di conseguenza, qualsiasi ricorso a posizioni “mentalistiche” che si allontanino dalla neurofisiologia dei processi cognitivi.

Certo è che, come sarà facile intuire, dall’esperimento di Libet ad oggi si sono susseguite una molteplicità di posizioni che, in sintesi, tendono o a salvare l’esistenza del libero arbitrio (come in Dennett) o a negarlo completamente: gli scettici, ad esempio, sono coloro che non si sbilanciano né dall’una né dall’altra parte. All’interno di tale quadro, i dualisti, contrariamente a Dennett, sostengono che il libero arbitrio non può essere un’illusione, poiché, altrimenti, l’agente non sarebbe in grado di accedere ai processi neurali che determineranno il suo pensare o il suo agire²⁹².

²⁹² Cfr. SAM HARRIS, *Free will*, New York University Press, New York, 2012.

Se i “compatibilisti classici” (come Dennett) sostengono che la mancanza di un controllo diretto sui processi cerebrali non sia sinonimo di *alterità*, ma, piuttosto, che l’agente sia l’autore dei propri progetti, poiché sempre “fisicamente e cerebralmente istanziato in se stesso”²⁹³, di altra natura è il *free will* per la posizione contraria.

Secondo gli incompatibilisti tale ragionamento è la prova di un errore interpretativo di fondo: il libero arbitrio, infatti, avrebbe una natura prettamente psicologica. Riprendendo le tesi di Daniel Wegner e di coloro i quali condividono la medesima linea interpretativa, l’esperienza soggettiva di essere artefici della propria azione sarebbe “niente più” che un’illusione funzionale della mente per giustificare a se stessa la sua “azione tardiva” rispetto a quella prettamente neurobiologica (i 250mms di ritardo individuati, già, dallo stesso Libet).

Ne consegue che, naturalizzando la volontà e sostenendo che scelte, sforzi e intenzioni non siano che parti di una catena di cause che precedono la consapevolezza cosciente sulle quali l’agente stesso non può esercitare alcun controllo, sia la volontà sia il libero arbitrio non sarebbero altro che un’illusione.

In effetti, l’accusa di dualismo cartesiano fu più volte rivolta a Libet, il quale non la smentì mai, lasciando presupporre che condividesse questa visione del mentale.

Ciò che emerge dall’esperimento è l’assenza di una teoria adeguata che possa fornire una spiegazione plausibile ai dati ricavati da strumenti scientifici piuttosto semplici²⁹⁴, tanto che non sono mancati i tentativi di conciliare forzatamente i dati ricavati dall’esperimento (l’assenza del controllo dell’azione da parte dell’agente e della volontà cosciente nel compiere un movimento, ossia la presenza dei 250mms) a favore o della tesi deterministica o di quella indeterministica²⁹⁵.

Per quanto riguarda la prima tesi, molti hanno sostenuto che la decisione di flettere il dito può derivare dall’intensa attività del potenziale di prontezza il cui esito (l’effetto) condurrebbe alla consapevolezza del soggetto di compiere l’azione; mentre per quanto riguarda la seconda, molti hanno sostenuto che l’esistenza dei 250mms può essere la prova tanto agognata dell’esistenza di leggi a-causali che risolverebbero il problema della localizzazione indeterministica²⁹⁶.

²⁹³ *IBIDEM.*

²⁹⁴ Cfr. MARIO DE CARO, *Il libero arbitrio*, in ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Neuroetica*, Il Mulino, Bologna, 2011.

²⁹⁵ Cfr. J. DAVID HAYNES, *Posso prevedere quel che farai*, in ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Neuroetica*, op. cit.

²⁹⁶ *IDEM.*

Proprio per tali motivi, alcuni anni fa il gruppo di lavoro del neuroscienziato Soon ha deciso di ripetere il famoso esperimento, seguendo la stessa linea metodologica, ma con strumenti decisamente più sofisticati di quelli utilizzati da Libet.

6.2 “Il nuovo Libet”: vecchi intenti e medesime difficoltà

Nel 2008, il gruppo di lavoro del neuroscienziato Soon ha messo a punto una nuova versione del famoso esperimento libetiano, apportando delle modifiche specifiche per quanto riguarda l'utilizzo di speciali neurotecnologie: il *team* statunitense è ricorso, infatti, all'ausilio della risonanza magnetica funzionale (fMRI) in grado di misurare i livelli di ossigenazione del sangue e, decisamente, più precisa nella rilevazione dell'attività cerebrale rispetto al semplice elettroencefalogramma, usato anni prima da Libet²⁹⁷. Il compito assegnato al soggetto, in fase sperimentale, consisteva sempre nel muovere la falange di un dito per premere un pulsante quando ne avesse avvertito il desiderio, con alcune modifiche di poco conto nella scenografia dell'esperimento²⁹⁸ (ad esempio, la sostituzione dell'orologio con un quadrante luminoso).

Si noti che l'utilizzo della fMRI, se ha il pregio di fornire maggiori dettagli visivi sull'incremento dell'attività encefalica, ha anche dei notevoli limiti strutturali: questa, infatti, non permette né la visualizzazione a cascata dei processi neuronali che portano all'azione né la visione in tempi reali di ciò che accade. I dati da essa forniti, in altre parole, sono sempre in ritardo (di alcuni secondi) rispetto al verificarsi dell'evento cerebrale che essa sta monitorando; avendo, inoltre, una griglia di risoluzione molto bassa, non permette di avere una panoramica precisa dell'intera attività cerebrale durante lo svolgimento di un compito. Nonostante tali limiti, i dati forniti dalla macchina mettono, comunque, in luce la presenza effettiva di uno specifico scarto temporale fra la prima e la seconda. Uno scarto quantizzabile di circa 10 secondi dalla comparsa dell'una e dell'altra.

Gli studi libetiani hanno avuto il merito indiscusso di aprire le porte al tentativo di naturalizzare aspetti metafisici dell'azione apparentemente lontana dalla cognizione: come, infatti, si sarebbe mai potuto pensare, al tempo di Cartesio, che la mente potesse essere indagata neuroscientificamente?²⁹⁹.

²⁹⁷ Cfr. JOHN D. HAYNES, *Posso prevedere quel che farai*, in ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Neuroetica*, op. cit.

²⁹⁸ *IBIDEM*.

²⁹⁹ Cfr. PATRICK HAGGARD E MARTIN EIMER, *On the relation between brain potentials and the awareness of voluntary movements*, in “Exp. Brain Res.” n. 126, 1999, pp. 128-133; ITZHAK FRIED, ROY MUKAMEL, GABRIEL KREIMEN, *Internally generated preactivation of single neurons in human medial frontal cortex predicts volition*,

Al di là del dibattito che si è incrementato negli anni sul *cogito* o su taluni aspetti del libero arbitrio, è bene ricordare che nel 2014, sulla scia del famoso esperimento libetiano ed il tentativo avvenuto nel 2008, il *team* di neuroscienziati di Soon *et alii* ha riprodotto, nuovamente, lo studio con constatazioni conclusive “inaspettate”.

L’ipotesi di partenza, stavolta, è stata dettata dalla necessità di indagare il *free will problem* in maniera tendenzialmente diversa rispetto a quanto già presentato nella tradizione, poiché fondamentale è stato impostare l’esperimento con l’intento di rispondere a tre questioni principali emerse dagli studi precedenti:

- a. che cosa può insegnarci l’attività cerebrale sul libero arbitrio? O, meglio, può quest’ultima *anticipare*, significativamente [n.d.r.], il risultato del [l’eventuale, n.d.r.] decisione libera?
- b. che cosa può spiegare questa attività cerebrale “precoce”?
- c. gli esperimenti che si sono succeduti da Libet ai giorni nostri affrontano realmente il problema del libero arbitrio? Se sì, da quale angolazione e da quale punto di vista?³⁰⁰

A tal riguardo è esemplificativa la seguente posizione:

Three main questions arise from this early research and will be addressed here. First, what can be learned from brain activity about free will decision-making; or more specifically, can brain activity predict the outcome of free decisions in advance? We will review some key papers demonstrating convincingly that this is indeed possible. Second, what might explain this early decision-related brain activity? We will argue that many current studies fail to link their results to cognitive decision models. However, this does not mean that there is a lack of appropriate theoretical models for free decision tasks. We will review other related studies, which suggest that sequential-sampling models, or evidence-accumulation models, as used for perceptual and cognitive decision problems [...] can explain results from free decision studies. We will provide evidence that, as would be expected in the framework of these models, the decision-makers’ use of implicit contextual information can explain the most puzzling phenomena of recent free decision studies. Finally, we will discuss whether experimental studies actually tackle the philosophical question of free will. We will argue that this is not

in “Neuron” n. 69, 2011, pp. 548-562; LOGAN SCHNEIDER, ELISE HOUDAYER, OU BAI E MARK HALLETT, *What we think before a voluntary movement*, in “J. Cognitive Neuroscience” n. 25, 2013, pp. 822-829.

³⁰⁰ Cfr. STEFAN BODE, CARSTEN MURAWSKI, CHUN SION-SOON, PHILIPP BODE, JUTTA STAHL, PHILIP L. SMITH, *Demystifying ‘free will’: The role of contextual information and evidence accumulation for predictive brain activity*, in “Neuroscience and Biobehavioral Review”, 47, 2014, pp. 636-645.

*necessarily the case. While some papers might imply that they do, the focus of the actual research is usually directed at different questions [...]*³⁰¹.

Un'operazione preliminare, necessaria al fine di evitare incomprensioni, sia concettuali sia linguistiche, è stata la scelta di adottare un vocabolario scientifico quanto più "neutrale" possibile.

Decidere, infatti, di avviare un esperimento sul libero arbitrio apre, come è facile intuire, "infinite" premesse concettuali da delucidare, che implicano, di conseguenza, prendere posizione e dirigere, fin dal principio, l'andamento dell'esperimento e le sue constatazioni conclusive.

In quest'ottica, infatti, deve esser valutata la scelta di alcuni autori di adottare le espressioni "azione volontaria, movimento volontario, decisione interna"³⁰² piuttosto che la classica espressione *free will*. Quest'ultima, come ricordano Soon *et alii*, lascia presagire un significato di "azione libera" ambiguo: da una parte, l'idea che la libertà, pur declinandosi metafisicamente in molteplici modi, significhi "assenza di costrizioni esterne"; dall'altra, sulla scia del cartesianesimo ancora vigente, l'idea che artefice della suddetta azione debba essere la coscienza e non il cervello. Non a caso, per tentare di essere "più neutrali possibili", diversi altri scienziati diversificano la fonte dell'azione: "attiva" o "passiva"³⁰³, dando seguito – forse inconsapevolmente – alla differenza *in nuce* dell'azione libera: generata "dall'interno" o subita "dall'esterno"? Il problema, quindi, sembrerebbe, innanzitutto delineare una definizione universalmente accettata di "azione libera" anche nell'area scientifica, poiché, contrariamente a quanti avanzano proposte di definizione, attualmente sembra impossibile parlare anche solo di "azione", prescindendo dall'ambiente circostante: esistono, infatti, delle aree cerebrali che si attivano, sia in un caso sia in un altro, manifestando delle sovrapposizioni evidenti tali da non garantire, in maniera *clara et distincta*, la differenza netta fra i due casi. *Ad escludendum*, il medesimo problema si presenta laddove si volesse indicare, neuroscientificamente, quando un'azione non è libera: stesse ambiguità e stessi limiti che, come in Libet, si presentano, sotto (non) mutate spoglie, oggi.

A tal fine è interessante notare che Soon *et alii*, nonostante approccino il problema da un punto di vista neuroscientifico, si pongono il problema del libero arbitrio da una prospettiva

³⁰¹ *IDEM*, p. 637.

³⁰² Cfr. PATRICK HAGGARD, *Human volition: towards a neuroscience will*, in "Nat. Rev. Neurosci", 9, 2008, pp. 934-946.

³⁰³ Cfr. ALEXANDER BATTHYANY, *Mental causation and free will after Libet and Soon: reclaiming conscious agency*, in ALEXANDER BATTHYANY, AVSHALOM C. ELITZUR (A CURA DI), *Irreducibly Conscious. Selected Papers on Consciousness*, Winter, Heidelberg, 2009.

metafisica quale premessa al loro lavoro; ad ulteriore testimonianza non soltanto del connubio inscindibile fra etica e neuroscienze, ma che *de facto* è l'interpretazione dell'oggetto d'indagine che fa protendere per una spiegazione o per un'altra. Importante diviene la metafisica della scelta a fondamento dell'agentività, esperibile in laboratorio, affinché si possa tentare di naturalizzare l'uomo in quanto tale; ciò non significa, si ribadisce, che si "svaluti" la natura umana, anzi! Trovare le basi neurali del *free will* o della coscienza etc significherebbe comprendere la straordinaria complessità *naturale* del mondo: un risultato "ben riuscito" di un insieme innumericamente contabile di particelle e forze che, interagendo fra loro, avrebbero dato a vita tanto alle amebe quanto al cervello umano (per circoscrivere il discorso solo ed esclusivamente al pianeta Terra). Un "miracolo trascendentale" che conferisce fascino e mistero alla natura umana. Ed il libero arbitrio?

A tal proposito, riprendendo il filo rosso, il *team* di Soon sostiene:

The problem with finding a clear distinction between free and non-free decisions already demonstrates an important point: a truly free decision in the intuitive sense might not exist

*Moreover, there might not be any decisions that are either fully internally generated, or fully externally determined. Free decisions might simply be the decisions for which we least understand the origin and underlying mechanisms- the question then arises which processes generate these decisions*³⁰⁴.

L'esperimento del 2014 ha riprodotto, a grandi linee, quello del 2008: ad un agente è stato chiesto di pigiare un bottone quando ne avesse avvertito il desiderio; ciò che è mutato è stato il supporto neuroscientifico per lo studio dell'attività cerebrale. Attraverso l'analisi di classificazione del modello multivariato (MVPA) insieme alla *fMRI* e la *brain imaging*, l'*equipe* di Soon ha individuato una rete di regioni cerebrali coinvolte nella presa di decisione volontaria e nella successiva azione: la corteccia prefrontale mediale (PFC), le aree sensorie (SMA e pre-SMA), l'area cingolata rostrale, i ganglia basali e la corteccia parietale posteriore (PPC). L'aspetto che più ha destato sorpresa è stato che le decisioni (in termini di "azione") possono essere predette dall'attività della corteccia prefrontale mediale anteriore e dalla prefrontale mediale posteriore sette secondi prima (e non dieci come nell'esperimento del 2008): il tempo segnalato dalla "comparsa" della consapevolezza cosciente. La PFC e la

³⁰⁴ STEFAN BODE, CARSTEN MURAWSKI, CHUN SION-SOON, PHILIPP BODE, JUTTA STAHL, PHILIP L. SMITH, *Demystifying free will: The role of contextual information and evidence accumulation for predictive brain activity*, in "Neurosci Biobehav Rev.", 2014, p. 638.

PPC sarebbero, quindi, responsabili tanto dell'esecuzione materiale dell'azione quanto della sua predizione a livello inconscio, così come lascia pensare lo stesso esperimento di Soon *et alii* del 2013³⁰⁵.

Considerando che i dati ricavati sono suscettibili di tutte le limitazioni strutturali delle tecnologie usate, resta il fatto che, psicologicamente e per la prospettiva della *folk psychology*, tutto ciò fa sembrare la consapevolezza poco più di una fievole percezione delle decisioni inconscie, rendendo la vita molto simile alla visione di un film.

Come già chiarito in premessa, gli autori non intendono sostituirsi al “mestiere del filosofo” per “approntare” o “indirizzare” il dibattito metafisico sul libero arbitrio a favore di una o di un'altra posizione, ma, per ciò che pertiene le loro competenze, evidenziare ai molti che “funzioniamo in un certo modo”:

*The validity of this assumption is pertinent to whether neural activity preceding a trial is influenced by memory of the prior trial(s) or preparation for the upcoming decision [...]. We argue that the assumption of truly independent decisions is both unlikely and unproductive for understanding the underlying cognitive processes, and for clarifying the relationship of these results to the free will problem. However, sequential dependencies of choices by no means contradict the idea that these choices are prepared unconsciously, as argued before [...]*³⁰⁶.

Al di là delle critiche rivolte all'esperimento³⁰⁷, Soon *et alii* suggeriscono che, benché le scelte non siano consciamente preparate, ognuna di esse è improbabile che “nasca dal nulla”, nel senso di essere completamente indipendente da quelle che l'hanno preceduta. Il contesto così creato, influisce ed influenza, in tal modo, il processo di decisione, non necessariamente in un modo conscio, ma in forme di “linea guida” per le scelte che verranno:

[...] the states of these decision networks do not exist in temporal isolation, but result from a causal flow of events, which may also be involved in prior choices. This not

³⁰⁵ CHUN SION SOON, AH HE, STEPHAN BODE, JOHN-DYLAN HAYNES, *Predicting free choices for abstract intentions*, in “Proc. Natl. Acad. Sci. USA”, 110 (15), 2013, pp. 5733-5734.

³⁰⁶ STEFAN BODE, CARSTEN MURAWSKI, CHUN SION-SOON, PHILIPP BODE, JUTTA STAHL, PHILIP L. SMITH, *Demystifying “free will”: The role of contextual information and evidence accumulation for predictive brain activity*, op. cit., p. 639.

³⁰⁷ Cfr. REI AKAISHI, KIM UMEDA, K. SAKAI, *Autonomous mechanism of internal choice estimate underlies decision inertia*, in “Neuron”, 81 (1), 2014, pp. 195-206; JESSE J. BENGSON, TODD A. KELLEY, XIAOKE ZHANG, JANE-LING WANG, GEORGE R. MANGUN, *Spontaneous neural fluctuations predict decisions to attend*, in “J. Cogn. Neurosci”, 26, 204, pp. 2578-2584.

*mean, however, that we consciously and intentionally make new decisions based on our decision history. In other words, the task context might have had a crucial role in creating the observed decision patterns, as well as contributed to the early time point of decision patterns. Thus, acknowledging the existence of sequential effects and choice biases would contribute to advancing our understanding of the origin of free choices*³⁰⁸.

Se così fosse, si aprirebbe un nuovo campo d'indagine per la psicologia cognitiva.

Il ruolo cruciale svolto dall'attività neurale inconscia porterebbe alla predizione della scelta "futura" che il soggetto eseguirebbe, consapevolmente, dopo un certo lasso temporale dalla sua origine sinaptica, facendo nascere, di conseguenza, notevoli perplessità sul concetto di "libera scelta" e di consapevolezza cosciente. L'assunzione che le scelte potrebbero essere frutto inconsapevole di una serie di preferenze operate negli anni o nel contesto di riferimento, significherebbe che l'agente sarebbe fortemente determinato, lasciando presagire un ruolo del tutto marginale alla coscienza:

*[...] Finally, although it does not resolve the philosophical question, the demonstration of unconscious brain processes for free decision-making inspires more theoretically and methodologically sophisticated research into the mechanisms of human decision-making than have hitherto been undertaken*³⁰⁹.

Dai numerosi esperimenti condotti neuroscientificamente sul libero arbitrio, emerge la negazione di un'intuizione fortemente radicata nella psicologia comune, secondo cui il compimento di un'azione discenderebbe direttamente dalla *libera* volontà del soggetto e non dalla precisa attivazione di chissà quale area cerebrale. La constatazione, invece, che esistono specifiche aree cerebrali che si attivano prima della comparsa della consapevolezza dell'individuo provoca non poche perplessità: quale relazione causale esisterebbe fra l'attività neurale e la volontà?

Il problema di Libet torna invariato.

Se si volesse sposare la tesi deterministica, si dovrebbe dimostrare l'esistenza di una connessione causale costante fra una specifica attività cerebrale e la conseguente volontà che ne deriverebbe, grazie all'ausilio di una misurazione tecnica attendibile.

³⁰⁸ STEFAN BODE, CARSTEN MURAWSKI, CHUN SION-SOON, PHILIPP BODE, JUTTA STAHL, PHILIP L. SMITH, *Demystifying free will: The role of contextual information and evidence accumulation for predictive brain activity*, op. cit., p. 640.

³⁰⁹ *IDEM*, p. 642.

Ciò che emerge dall'esperimento di Soon *et alii*³¹⁰ è l'impossibilità attuale (come in Libet), di fornire un'interpretazione adeguata dei dati raccolti così da corroborare o la tesi deterministica o quella indeterministica sulla presa di decisione cosciente: una impossibilità che potrebbe non rivelarsi più tale in un futuro prossimo grazie al progredire delle scoperte scientifiche e tecnologiche³¹¹.

La scienza e la filosofia sono ancora limitate nell'identificazione di una strada pratica e concettuale volta alla spiegazione di che cosa effettivamente sia la libertà ed in che modo essa "funzioni".

Proprio a tal proposito, il filosofo Joshua Knobe sostiene che, prima ancora di cimentarsi nuovamente nello studio del libero arbitrio a livello etico e pratico, sarebbe opportuno condurre uno studio psicologico preliminare sull'uomo stesso: solo attraverso un approccio analitico della psicologia umana si sarà successivamente in grado di rispondere alle questioni portanti dell'etica.

Scrive a tal proposito:

*L'idea chiave è che se arriviamo a capire meglio la psicologia che c'è dietro le intuizioni filosofiche possiamo distinguere meglio quali intuizioni siano degne di fiducia e quali vadano messe da parte come inaffidabili o fuorvianti. La nostra speranza è che questo lavoro ci porti a una conoscenza più profonda riguardo alle convinzioni delle persone sulle grandi questioni filosofiche*³¹².

Come lo stesso Knobe sostiene, studi preliminari sulle convinzioni e sulla psicologia umana potrebbero riuscire a dirimere numerose controversie sugli interrogativi ancora aperti della filosofia come, appunto, il libero arbitrio. L'intuizione di base della filosofia sperimentale – il nuovo approccio psicologico alla scienza – risiede nel fatto che, come già ribadito, a volte si ha l'impressione che posti di fronte ad un problema la mente

*[...] sembra tirarci allo stesso tempo in due direzioni diverse, come se dentro di noi ci fossero due voci separate che danno due risposte opposte alle stesse domande*³¹³.

³¹⁰ Sulle procedure di controllo utilizzate e sulle migliorie che gli stessi Soon e colleghi hanno previsto per i futuri esperimenti, cfr. JOHN D. HAYNES, *Posso prevedere quel che farai*, in ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Neuroetica*, op. cit.

³¹¹ *IDEM*.

³¹² JOSHUA KNOBE, *Esperimenti sul pensiero*, articolo reperibile on line al seguente link <http://download.kataweb.it/mediaweb/pdf/espresso/scienze/2012/01/31/115530199-3cebd776-d977-417c-bf9c-dc7a1a1ee84e.pdf>

³¹³ *IDEM*, p. 50.

Ciò ha fatto nascere in Knobe l'idea che a fondamento di tutte le intuizioni contrastanti sul *free will* si nasconda una tensione fra due diverse forme della cognizione umana: la riflessione teorica astratta e quella emotiva che, simultaneamente, spingono il pensiero in direzioni opposte.

Alcuni esperimenti hanno mostrato che un'ulteriore dato che influenza il giudizio nei soggetti analizzati è la componente temporale: a seconda che si immagini un evento collocato nel tempo presente o nel futuro prossimo, si è portati ad impiegare processi cognitivi diversi. Nel primo caso si sarà indotti a ragionare con la cognizione emotiva, nel secondo con quella teorica astratta e, dunque, con l'uso di un ragionamento maggiormente riflessivo. Tali studi – ancora in via di sviluppo – potrebbero, in futuro, essere in grado di spiegare i motivi per cui vige ancora tanto imbarazzo teorico sul problema del libero arbitrio.

Anche laddove l'ipotesi di lavoro di Knobe un giorno dovesse rivelarsi veritiera, da un punto di vista neuroscientifico, si dovrà comunque spiegare il significato dello scarto temporale fra l'azione ed il pensiero individuato anche da Soon e dal suo *team*: se fosse vera l'idea di Dennett – secondo cui i 250ms individuati da Libet corrispondono al tempo necessario al cervello per avviare il complesso sistema *coscienza* – in che modo si riuscirebbe a conciliarla con l'idea di Knobe?

Tutto ciò rispecchia l'enorme difficoltà di conciliare la sfera filosofica con quella neuroscientifica tanto che molti autori si sono persuasi dell'illusorietà del libero arbitrio, convinti che sarà una sfida difficile, forse impossibile, spiegare *free will* o volontà cosciente dal “solo” funzionamento cerebrale.

6.3 La proposta di Michael Gazzaniga: *the interpreter*

Il neuroscienziato Michael Gazzaniga, in un libro relativamente recente, ma, sostanzialmente, ancora molto attuale, ha tentato di avviare un discorso di comprensione sul libero arbitrio e sulla responsabilità morale che non considerasse i dati forniti dalla *fMRI* sulle attività di specifiche aree cerebrali o sul funzionamento degli emisferi cerebrali³¹⁴, ma tentasse di spiegarlo, innanzitutto, da una prospettiva teorica. Un percorso “all'inverso”, in cui si è delineata dapprima una riflessione sui possibili dati e poi la loro raccolta pratica.

L'idea di Gazzaniga è, innanzitutto, quella di chiedersi da che cosa effettivamente l'uomo debba ritenersi o considerarsi libero e perché comprendersi in termini fisici/meccanicistici

³¹⁴ Cfr. MICHAEL GAZZANIGA, *Chi comanda? Scienza, mente e libero arbitrio*, Codice Edizioni, Milano, 2013.

allontanerebbe l'idea dell'agente responsabile: per quanto il determinismo – inteso come tesi generale sul mondo – possa effettivamente riscontrare notevoli simpatie fra scienziati e neuroscienziati, la cui *ratio* sia di essere in grado di spiegare il comportamento umano, è evanescente così come lo è il tentativo indeterminista. Secondo il neuroscienziato, infatti, i limiti delle due posizioni sono individuabili nell'impostazione del problema solo su un piano strettamente metafisico, lontano dal reale, senza partire dall'unica base quantomeno certa che governa il mondo, ossia l'uomo ed il suo cervello.

Su questa scia, egli sostiene che l'approccio neuroscientifico (iniziato, simbolicamente, da Libet) è completamente errato: il punto non è quello di comprendere in quali aree o luoghi specifici del cervello si svolga l'azione, ma in che modo la complessa interazione dei diversi moduli cerebrali interagisca per dar vita all'idea di un Sé unitario. La tesi di Gazzaniga sul libero arbitrio, allora, è che esso costituisca nulla più che un'illusione, diretta conseguenza del complesso funzionamento cerebrale: una sorta di *emergenza* derivante dalla comparsa della coscienza, che consisterebbe in una funzione cognitivamente impegnativa: da un livello di organizzazione ad un altro.

Gazzaniga introduce, così, il concetto di *the interpreter*: la capacità dell'emisfero sinistro di *interpretare* il reale in pochi millisecondi a seconda delle informazioni sensoriali, percettive ed emotive che il soggetto ha a disposizione in quel tempo limitato. La chiave per comprendere un'emergenza è capire, a sua volta, che esistono diversi e molteplici livelli di organizzazione cerebrale complementari l'uno all'altro e che solo attraverso una visione d'insieme dell'intero funzionamento cerebrale si può comprendere il senso di un particolare fenomeno apparentemente unitario. In ciò, allora, cadrebbe l'approccio di Libet: sezionando una particolare area cerebrale per la comprensione del libero arbitrio, egli ha perso di vista il quadro d'insieme decisamente più complesso ed omogeneo che lo ha generato.

Da che cosa nasce, allora, l'illusione del libero arbitrio?

Dalla creatività dell'emisfero sinistro: quest'ultimo, rispetto a quello destro che ha compiti decisamente più analitici e schematici, è quello dedito alla creatività e, grazie a degli studi fatti su persone con la sindrome del cervello diviso³¹⁵, il *team* del neuroscienziato ritiene che quell'immagine di un Sé unitario, che pilota la nostra coscienza psichica, altro non sarebbe che il risultato di una complessa interazione di moduli a livelli diversi che l'emisfero sinistro integrerebbe con i dati forniti dall'esterno. L'emisfero destro, infatti, tende a massimizzare i

³¹⁵ Denominata anche *split-brain*, la sindrome del cervello diviso si verifica in quei pazienti nei quali, per motivi medici, viene interrotta, tramite la recisione sinaptica del corpo calloso, qualsiasi comunicazione fra l'emisfero destro e quello sinistro. Ciò comporta una vera e propria separazione cognitiva e coscienziale all'interno del medesimo paziente, per cui la parte destra e quella sinistra vedono, percepiscono...il mondo e interagiscono con questo in modi decisamente diversi simultaneamente.

dati percepiti e lo stesso Gazzaniga, a seguito di ulteriori studi condotti dall'*equipe* del neuroscienziato Paul Corballis su pazienti con il *cervello diviso*, non esclude che un modulo dell'interprete possa trovarsi anche in questo secondo emisfero. Il "processo" è così descritto da Gazzaniga:

*L'interprete lavora con le informazioni che ha a disposizione. Riceve i risultati delle elaborazioni operate da moltissimi moduli diversi, ma non riceve l'informazione che vi sono moltissimi moduli diversi, e nemmeno di come questi moduli funzionino; tantomeno riceve l'informazione sull'esistenza di un sistema di riconoscimento di pattern nell'emisfero destro. L'interprete è un modulo che spiega gli eventi in base alle informazioni che riceve*³¹⁶.

L'idea dell'esistenza dell'interprete o, semplicemente, di questa specifica capacità dell'emisfero sinistro, è nata in Gazzaniga – come precedentemente si diceva – grazie a degli studi effettuati su pazienti *split brain*; in quanto, così come già ricordato precedentemente, essi non possiedono sinapsi di collegamento fra i due emisferi, per cui ognuno dei due fornisce informazioni specifiche e discordanti alla coscienza, che sembra anch'essa divisa in due.

In pazienti sani l'interprete è ciò che sintetizza a livello interpretativo qualsiasi *input* "sparso" nel "caos sinaptico" per dar senso a ciò che accade, facendo nascere la sensazione illusoria di essere in grado di "agire liberamente"; *the interpreter*, per sua natura, è, allora, quel meccanismo cerebrale che interpreta non solo le ragioni o le motivazioni dei comportamenti, ma anche ciò che accade nel cervello a livello microscopico.

Il cervello, quindi, non può essere considerato come un sistema statico ed immutabile in cui *tutto* debba essere determinato come, eventualmente, il mondo fisico in cui interagisce. Sfatando quest'assunzione erronea (e, quindi, sfiduciando la tesi deterministica), Gazzaniga, facendo appello alla teoria dei sistemi complessi, sostiene l'*esistenza* del *free will* e della responsabilità delle proprie azioni seppur in modo distinto rispetto al senso di *esistenza* conferito, usualmente, ad un oggetto fisico inserito in un contesto. L'*esistenza* di cui parla Gazzaniga, infatti, è da intendersi quale risultante di un'*emergenza* cerebrale in cui dalla complessa interazione sinaptica conseguirebbe tanto il libero arbitrio quanto la responsabilità morale. In questo senso, allora, si può parlare di "illusione" del *free will*, poiché non può essere un "luogo" fisico, ma una conseguenza di processi *down-up* del complesso sistema

³¹⁶ MICHAEL GAZZANIGA, *Chi Comanda?*, op. cit., p. 100.

cerebrale. In ciò, quindi, falliscono – come detto – tanto gli intenti di Libet quanto quelli di Soon *et alii* che si sono succeduti negli anni.

Introducendo, inoltre, il concetto di *complementarietà*, il neuroscienziato ribadisce che non è importante individuare il sistema o i sistemi responsabili dell'azione, ma esser consapevoli (dal punto di vista dell'interrogante) che *free will* o responsabilità morale siano il frutto di elementi *complementari* l'uno agli altri, ossia il risultato di “componenti biologiche” che, lavorando “insieme” (in questo senso, *complementari*), costituiscano i propri stati mentali da cui *emergerà* il “sistema libero arbitrio” o il “sistema responsabilità morale”.

Il rovescio della medaglia si ritrova nel fatto che, nonostante al *free will* Gazzaniga abbia conferito uno *status* particolare di *esistenza*, che potrebbe esser scambiata per una tesi che sostenga l'illusorietà dello stesso, il neuroscienziato propone di continuare a sostenere l'esistenza del *free will* per due ragioni centrali: la prima è che, per quanto illusorio, il sentimento di unitarietà dell'agente è funzionale al suo vivere sociale ed al suo senso di responsabilità; la seconda ragione risiede nelle degenerate conseguenze comportamentali a cui si assisterebbe laddove si diffondesse la convinzione che il libero arbitrio effettivamente non esista. Per quanto riguarda questa seconda ragione, Gazzaniga fa riferimento ai tanti esperimenti psicologici che si stanno susseguendo in questi anni per cercare di capire che cosa potrebbe mai accadere se non si credesse più nel libero volere. Diversi dati hanno mostrato che, data tale premessa, il comportamento sociale cambierebbe in peggio, poiché, non ritenendosi più responsabili, gli uomini si lascerebbero andare a comportamenti al limite della legalità.

Per quanto riguarda la prima ragione, Gazzaniga, rifacendosi agli esperimenti di Lau *et alii*³¹⁷, sostiene che il sentimento di unitarietà del Sé, che l'agente ha nel momento in cui compie qualsiasi atto, deriva da un'elaborazione dell'interprete *post hoc*, che sposta all'indietro nel tempo il sentimento cosciente di voler compiere una data azione qualche millisecondo dopo che il soggetto la sta già compiendo. Un'interpretazione *post hoc*, ossia dopo che il fatto sia accaduto, che l'interprete compie per dare una spiegazione razionale, ad esempio, all'atto istintivo di togliere un dito da vicino una fiamma, che contribuisce alla formazione di un sentimento cosciente di essere padroni della propria volontà.

Se, allora, illusorio è il sentimento del libero arbitrio, lo sarà anche il senso di responsabilità?

Per Gazzaniga, assolutamente no:

³¹⁷ *IDEM*, p. 120-122.

I cervelli sono macchine automatiche che seguono percorsi decisionali, ma analizzando i singoli cervelli in isolamento non si riesce a mettere a fuoco la capacità di essere responsabili: la responsabilità è una dimensione di vita che deriva da uno scambio sociale e lo scambio sociale richiede più di un cervello. Quando più cervelli interagiscono, cominciano ad emergere elementi nuovi altrimenti imprevedibili, stabilendo un nuovo insieme di regole³¹⁸.

Un insieme di regole che, socialmente, vengono qualificate come appartenenti al senso di responsabilità morale e di libertà che non risiederebbero nel cervello, ma

[...] nello spazio fra cervelli, nell'interazione fra persone³¹⁹.

In tal modo, tanto il libero arbitrio quanto la responsabilità morale sarebbero salvati nonostante un approccio neuroscientifico. Che cosa dire, allora, a proposito dell'esistenza della volontà e della coscienza? Sarebbero anch'esse frutto di un'interpretazione *post hoc* che il cervello farebbe per donare all'agente un sentimento del Sé unitario? Come si è già detto, per Gazzaniga è questa la natura del libero arbitrio e, stando al panorama contemporaneo, tale approccio non è il solo: anche alcuni psicologi, fra cui Daniel Wegner, sono giunti alle medesime conclusioni tramite strade concettuali decisamente differenti.

6.4 Il libero arbitrio fra emozione ed illusione

Lo psicologo Daniel Wegner è colui il quale, nel panorama delle scienze cognitive contemporanee, si avvicina considerevolmente alla tesi di Gazzaniga, tanto che, per molti, le due tesi sarebbero l'una il complemento dell'altra: da una parte lo studio neuroscientifico del libero arbitrio, dall'altra quello psicologico/scientifico della volontà cosciente. Anche Wegner, così come Gazzaniga, comincia con l'evidenziare che parlare di libertà d'azione per la mentalità comune significa sostenere che l'atto compiuto da un agente deve essere accompagnato dalla sensazione di controllare consapevolmente il proprio agire come facente parte del proprio vissuto soggettivo: la volontarietà dell'atto – ossia l'esperienza fenomenologica di governare la propria libertà – è ciò che contribuisce a livello psicologico a

³¹⁸ *IBIDEM*, p. 148.

³¹⁹ *IDEM*.

conservare nell'uomo il senso di responsabilità morale, dovuto alla sensazione di essere la causa che produce l'effetto:

*l'esperienza di voler compiere [consapevolmente, n.d.r.] un'azione con la causazione dell'azione da parte della mente [cosciente, n.d.r.]*³²⁰

è vivere la propria esperienza, la quale contribuisce ad alimentare l'immagine delle intenzioni coscienti come cause del proprio agire³²¹.

Il rapporto che si instaura fra la propria volontà e l'effetto che essa provoca, infatti, è dovuto proprio all'esperienza della volontà cosciente che aiuta ad identificare le azioni come proprie, distinguendole da ciò che accade intorno al soggetto operante. È come se, nel linguaggio metaforico, l'esperienza della volontà “marchiasse a fuoco” l'atto compiuto non facendolo percepire al soggetto soltanto come “proprio” – così come potrebbe essere un qualsiasi oggetto materiale – ma come profondamente ancorato e appartenente al proprio Sé psicologico: “quest'azione è *mia*” diviene simile a dire “quest'azione è il prolungamento fisico del mio Sé psicologico che si dà a vedere nel mondo reale”.

Contro coloro che sostengono che la volontà altro non sarebbe che un cambiamento cerebrale visibile nella struttura neurale, tanto da identificarsi con la comparsa della consapevolezza, Wegner, riprendendo le parole di T. H. Huxley, secondo cui:

*La volizione [...] è un'emozione indicativa di cambiamenti fisici, non una causa di tali cambiamenti*³²²,

concorda nel sostenere che l'esperienza della volontà cosciente sembra più simile ad un'emozione provata dall'agente piuttosto che ad una reale modificazione chimica a livello cerebrale. Alcuni studi scientifici, infatti, mettono in seria difficoltà la suddetta corrispondenza, mostrando che non sempre la volontà accompagna azioni desiderate o compiute dall'agente. Un esempio di ciò può essere fornito dalla sindrome della mano aliena, una malattia neuropsicologica dovuta a particolari lesioni cerebrali, in cui il paziente percepisce le azioni della propria mano come se essa agisse con una sua propria autonomia. Chi ne è affetto, difatti, sostiene di non volere e di non sentire come proprie le azioni

³²⁰ Cfr. DANIEL WEGNER, *L'illusione della volontà cosciente*, in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI, (A CURA DI), *Siamo davvero liberi?*, op. cit., p. 103.

³²¹ *IDEM*.

³²² *IDEM*, p. 105.

compiute dalla “mano aliena”: se non si ha la sensazione di compiere l’azione, come poter sostenere che vi sia volontà?

La maggior parte degli esseri umani percepisce le altre entità come viventi, poiché le riconosce dirette verso uno scopo, essendo in grado di pensare ai possibili obiettivi che vogliono raggiungere; applicando tale capacità anche a se stessi, l’uomo risulta in grado di pensarsi come un agente causale, essendo in grado di comprendere sé e gli altri tramite credenze, desideri e progetti, che contribuiscono alla convinzione di possederne una. Ciò, a sua volta, determina l’esperienza in prima persona di possedere una volontà cosciente, proprio in quanto si è in grado di immaginare le azioni da compiere prim’ancora che si compiano, contribuendo, da una parte, ad alimentare la causazione mentale sul mondo fisico; dall’altra, mascherandone il suo carattere prettamente illusorio.

I casi della sindrome della mano aliena e dell’ipnosi (in quest’ultimo, sembra proprio che il soggetto venga privato completamente della consapevolezza dell’azione, essendo ridotto a mero automa) suggeriscono che non sempre esiste una connessione fra l’azione compiuta e la sensazione di agire: se realmente l’agente è portato a compiere azioni che non desidera in che modo, allora, potrebbe dirsi realmente libero? Il punto, secondo Wegner, non è chiedersi “dove” mai si trovi il libero arbitrio, ma sentirne in prima persona l’esperienza cosciente che si esplica nella medesima esperienza della volontà³²³: interpretandola come reale causa dell’azione, l’uomo sperimenta un senso di libertà indipendente da qualsiasi reale connessione fisica fra attività cerebrale e azione; il pensare se stessi come artefici di un’azione è il motivo secondo cui il sentimento e l’esperienza della volontà legano in maniera indissolubile l’agente al suo fare.

In quest’ottica, perciò, è da interpretare la posizione di Wegner, definita come *Teoria della causazione mentale apparente*, secondo cui la volontà cosciente, impossibilitata nel cogliere la complessità degli stessi meccanismi cerebrali che la generano, inferisce il suo ruolo causale fra il pensiero e l’atto osservato. L’illusione della volontà come movente dell’azione suggerirebbe, così, al soggetto la sua *agentività causale*, ossia la capacità di agire nel mondo per differenziarsi dagli esseri inanimati e percepirsi come un Sé agente; ciò sarebbe sufficiente a mostrare che l’intenzione cosciente (la volontà) sarebbe un inganno funzionale, un’illusione, che la mente farebbe a se stessa. Proprio come un gioco di prestigio risulta all’occhio dello spettatore il risultato misterioso di particolari capacità del mago – ignorando completamente i meccanismi fisici capaci di creare un’illusione – nel medesimo modo la sequenza causale reale di stampo neurale, che soggiace al comportamento umano, coinvolge

³²³ *IBIDEM.*

un insieme enormemente complicato di meccanismi, contribuendo, in tal modo, all'idea di una volontà cosciente che preceda le azioni fisiche del corpo di cui, oltretutto, il soggetto agente ne fa esperienza.

La mente, afferma Wegner, crea in tal modo un'illusione cosciente proprio per sopperire alla mancanza di informazioni sulle cause che danno il via alle azioni corporali, producendo una sensazione continua di autospiegazione secondo cui solo a lei è imputabile il principio di qualsiasi movimento.

Il problema, tuttavia, consta proprio in ciò, ossia nell'impossibilità teorica da parte della mente di conoscere i meccanismi neurali che la governano, facendo risultare l'uomo un vero e proprio enigma già a se stesso. Ecco allora spiegata l'illusorietà della volontà: essendo incapace di conoscere le reali cause della coscienziosità del movimento, la mente crea – proprio come un mago – la percezione fallace che sia solo ed esclusivamente lei la causa prima di sé, facendo sperimentare all'uomo l'esperienza in prima persona della causalità del mentale sul fisico.

Risulta, a tal punto, ancora più facile spiegare l'intuitività del dualismo cartesiano e del perché, in effetti, quest'ultimo risulti essere ancora oggi la posizione più sostenuta, sia nella percezione comune sia per alcune teorie scientifiche.

Nella formazione cosciente dell'illusorietà di sé come causa prima, l'uomo, secondo Wegner, inferisce una relazione causale fra il proprio pensiero e l'azione, tendendo a vedersi come l'autore dell'atto quando ha, con un dato anticipo, pensieri coscienti sull'azione medesima da compiere; affinché la volontà sia esperita come cosciente, è necessario che il pensiero o l'intenzione di agire siano percepite dal soggetto prima che proceda all'azione, essendo coerenti con questa e con la certezza che non ci siano state altre cause potenziali a produrre l'effetto. L'unicità, la coerenza e la priorità risultano essere, allora, le tre fonti chiave per l'esperienza autocosciente della volontà pratica, arrivando a credere che, effettivamente, le intenzioni abbiano un potere causale sul mondo fisico.

Stando a questa logica, i reali motivi che spingono all'azione non possono mai essere coscienti e, di conseguenza, non stupisce che movimenti prettamente automatici compiuti dall'agente possano non essere l'espressione diretta della volontà. È probabile, inoltre, che i movimenti automatici siano il risultato di processi cerebrali inconsci proprio perché fondamentali al benessere fisico dell'organismo al contrario di quelli sotto la rigida esperienza soggettiva, che possono godere di un tempo di formazione cosciente maggiore (così da avere il tempo di far produrre alla mente l'illusione poi cosciente della causazione mentale sul mondo fisico).

Risulta chiaro, a tal punto, come la volontà in quanto tale non sia il risultato di un meccanismo neurale ancora da individuare, ma quello di un processo operato dal cervello stesso che interpreta le interconnessioni neurali dei pensieri come qualcosa che abbia una natura distinta rispetto a sé.

L'ipotesi di lavoro di Wegner, contrariamente a quella di Libet o di Soon, non è tanto quella di indagare, a livello cerebrale, l'intenzionalità cosciente quale attività preconsocia che determina qualsiasi azione volontaria, ma quell'emozione – cosciente – che accompagna qualsiasi azione compiuta dall'agente: il *free will*, quindi, come emozione.

In tal senso, sposando la tesi di Gazzaniga e di molti altri filosofi (ad es. Daniel Dennett), la coscienza subentrerebbe sempre *dopo* l'inizio di un'attività, tanto da non potersi mai dire la matrice di un'azione volontaria e consapevole. In quest'ottica Wegner sostiene che il libero arbitrio sia un'illusione e che la medesima illusione della volontà sia un sistema di adattamento dell'uomo all'ambiente di fondamentale importanza: essa costituirebbe “la bussola della mente”, la spia di essere agenti causali nel mondo capaci di autopercepirsi quali esseri senzienti in grado di formare e organizzare le risposte adeguate agli stimoli esterni; che sia il risultato di una forte emozione o la prontezza di agire in una gara sportiva.

L'illusione della volontà cosciente, allora, è particolarmente utile come guida all'uomo stesso, essendo in grado non solo di far percepire al soggetto la paternità dell'atto da compiere, ma anche di conservare il senso di responsabilità delle azioni che serve quale presupposto alla morale. Sarebbe un errore, a tal proposito, concludere che la volontà sia un'illusione irrilevante: oltre ad essere la causa del proprio Sé o della propria identità personale, essa ha pur sempre una base sinaptica che spetterà alla neuroetica scoprire, affinché risulti più chiara la sua natura fisica e psicologica.

Se, allora, la volontarietà dell'agente sembra essere solamente un gioco di prestigio che la mente farebbe a se stessa, il problema di come definirsi moralmente responsabili resta insoluto e l'unica alternativa possibile sarebbe quella di separare la libertà dal concetto stesso di responsabilità morale. In che modo ciò è possibile?

Al di là delle discrepanze esistenti fra compatibilisti e libertari, il *focus* centrale dei diversi argomenti riportati è sempre quello secondo cui la ricerca psicologica (nonostante sia ben lontana dall'esser risolutiva) sul libero arbitrio potrebbe notevolmente aiutare il dibattito ancora in corso perché fornirebbe al filosofo un dato “oggettivo” sulla sua esistenza o meno. Tale apporto verrebbe proprio dal sentire comune, dalle *folk intuitions*, e proprio da tale

assunto ha preso piede il lavoro dello psicologo Roy F. Baumeister³²⁴, il quale vuole coniugare i dati empirici derivanti dalle neuroscienze con le intuizioni comuni sul libero arbitrio.

Contro la teoria di Wegner e contro quelle che dichiarano l'illusorietà o lo scetticismo dell'esistenza del *free will*, si dibattono due diverse linee argomentative, che, seppur con alcune limitazioni, sostengono entrambe l'esistenza del libero arbitrio: la prima, prende in esame gli aspetti empirici del problema; la seconda, si concentra sull'interpretazione della nozione del Sé³²⁵.

Stante alla prima posizione, si evidenzia come la coscienza giochi un ruolo decisivo nell'attuazione del comportamento, ma, al contempo, stante la limitatezza strutturale delle attuali tecnologie, sembrerebbe pressoché impossibile fornire dati certi ed incontrovertibili sul ruolo da essa svolto e sulla eventuale identificazione dell'area specifica in cui risiede.

Tale strategia argomentativa, infatti, si articola nell'enunciazione di tre argomenti principali volti a dimostrare l'assoluta compatibilità fra neuroscienze e libero arbitrio:

1. la volontà cosciente è in grado di esercitare il *potere di veto* teorizzato da Libet;
2. la volontà cosciente è in grado di generare piani di azione in grado di influenzare *de facto* il comportamento pratico;
3. la volontà cosciente esercita un ruolo causale in qualsiasi pratica quotidiana non osservabile da un laboratorio.

Già dalla prima premessa emergono decisive contraddizioni *in nuce*, poiché si deve tener presente che:

*There are several ways that conscious representations of future actions seem essential to the way we perform them. One involves inhibition*³²⁶.

Se si dovesse scoprire che il *potere di veto* teorizzato da Libet non sia una capacità che la volontà effettivamente possieda (ossia la capacità del soggetto – o della sua coscienza – di interrompere in quei 250mms che separano la consapevolezza cosciente dall'esecuzione

³²⁴ Cfr. ROY F. BAUMEISTER, *Understanding Free Will and Consciousness on the Basis of Current Research Findings in Psychology*, in ROY F. BAUMEISTER, ALFRED R. MELE E KATHLEEN D. VOHS (A CURA DI), *Free will and consciousness: How might they work?*, Oxford University Press, New York, 2010.

³²⁵ Cfr. STEPHEN MORRIS, *The impact of the Neuroscience on the Free Will Debate*, in "Florida Philosophical Review", Vol. 9, Issue 2, 2009, pp. 56-78.

³²⁶ Si legge: "[...] ci sono diversi modi nei quali la coscienza rappresenta le azioni future che sembrano essenziali al modo in cui le eseguiamo. Uno coinvolge l'inibizione", *IDEM*, p. 58.

dell'azione il processo causale che condurrebbe, inevitabilmente, all'azione) allora si sarebbe in grado di stabilire che la coscienza non sia in grado di fermare l'esecuzione dell'azione, dato che la sua natura sinaptica glielo impedirebbe. Si contesta, cioè, la natura *spirituale* della volontà, fornita dall'interpretazione di Libet, per rivendicare la sua natura prettamente cerebrale non preoccupandosi di fornire dati pratici a sostegno di tali dichiarazioni:

*At this stage, however, this hypothesis lacks robust empirical support... Nahmias has offered no compelling empirical evidence that such adjusting or inhibiting actually occurs. At best then, we are left at an impasse regarding the notion that our conscious will possesses veto power*³²⁷.

Il *focus* dell'argomento empirico sembrerebbe, dunque, quello di sostenere che il modo *più ovvio* di constatare che la volontà cosciente abbia un significativo ruolo causale nel pilotare il comportamento sia di rendersi conto di quanto i suoi fini o i suoi scopi siano pilotati interamente dalla consapevolezza cosciente:

*Indeed, the most obvious way that conscious will play a causal role in behavior involves developing general goals or plans for behaviour*³²⁸.

Ciò che non può passare inosservato è l'*ovvietà* secondo cui si dovrebbe constatare il ruolo causale che la volontà cosciente giochi nell'indirizzare “verso scopi o fini il comportamento”: se anche tale ovvietà fosse vera, in che modo tale asserto potrebbe contrastare – se non proprio confutare – le tesi di Wegner? E perché la coscienza sarebbe necessaria per dirigere il comportamento quando i semplici automatismi incoscienti dimostrerebbero esattamente il contrario?

Supponiamo, per amor di discussione, che la volontà cosciente abbia un ruolo significativo nell'elaborazione dei fini o degli scopi verso cui il comportamento dell'agente è diretto; supponiamo, ancora, che si conosca, realisticamente, il modo in cui tale influenza agisca (un dato che al momento è ancora ignoto): sarebbe ciò sufficiente a garantire l'esistenza del libero arbitrio? C'è ragione di dubitarne in quanto non si può credere che la sola volontà cosciente

³²⁷ Morris scrive: “In conclusione, tale tesi perde il suo supporto empirico... Nahmias non ha offerto nessuna convincente evidenza empirica che le sostenga effettivamente. Nel migliore dei casi poi, ci ritroviamo in una situazione di stallo per quanto riguarda l'idea che la nostra volontà cosciente possieda il potere di veto”, *IBIDEM*, pp. 58-59.

³²⁸ Scrive Nahmias: “Infatti, il modo più ovvio in cui la coscienza gioca un ruolo causale nel comportamento coinvolge lo sviluppo di scopi o piani generali del comportamento”, *IDEM*, p. 59.

abbia un potere causale così forte da consentire l'esercizio del libero arbitrio. È difficile vedere come ogni piano comportamentale sia sotto lo specifico controllo del soggetto: alcuni di essi, infatti, potrebbero essere causati da processi cerebrali inconsci, che determinerebbero la specifica natura di questi movimenti. Il punto focale si concentra su una domanda antecedente a quella da cui Wegner ed altri come lui sono partiti: non tanto se la volontà possa – e in che misura – influenzare e dirigere “i piani del comportamento” (*general plans of behavior*³²⁹) dell'agente, quanto indagare se quest'ultima sia un qualcosa che si oppone profondamente ai processi cerebrali inconsci.

La maggiore plausibilità della teoria di Wegner si basa sul fatto che lo psicologo statunitense ha analizzato “casi al limite” in cui, cioè, volontà cosciente e inconsapevolezza delle azioni compiute si trovavano a convivere nel medesimo agente, simultaneamente. Ciò fornisce alla teoria di Libet-Wegner una maggiore evidenza empirica che, anche laddove dovesse essere errata l'interpretazione che ne è stata data, la pone su un piano decisamente più plausibile delle semplici – seppur convincenti – tesi della prima linea argomentativa, a cui, però, si contesta la mancanza di dati pratici.

La seconda linea argomentativa a difesa del libero arbitrio sostiene che, indipendentemente dalla vero-falsità delle tesi di Libet-Wegner, quest'ultima non potrebbe mai costituire una minaccia per l'esistenza del libero arbitrio.

In effetti, ad un primo sguardo, la teoria di Libet-Wegner sembra realmente la negazione *in toto* della libertà decisionale dell'agente, in quanto in essa è racchiuso il timore diffuso dalla mentalità comune di non essere liberi di agire secondo le proprie volontà:

*In order for it to be correct to say that a person causes or controls a particular action of hers (a necessary condition for free will), it must be the case that the act of conscious willing that she takes to be causing the action in question (the moment of perception of conscious willing is denoted by Libet as time t) truly does play a causal role in producing the action*³³⁰.

Tale approccio assume che il potere dell'individuo, il suo *Sé*, è causa delle proprie azioni, ma se da una parte rende possibile l'esistenza del libero volere, dall'altra implicherebbe,

³²⁹ *IBIDEM*, p. 62.

³³⁰ Scrive Dennett: “Per esser corretti, dire che una persona causi o controlli una sua particolare azione (una condizione necessaria per il libero arbitrio), significa sostenere che l'atto della volontà cosciente che l'agente intraprende causa l'azione in questione (il momento in cui il soggetto percepisce la volontà cosciente è denotato da Libet come il momento t) e che realmente essa giochi un ruolo causale nella produzione dell'azione”, cit. in STEPHEN MORRIS, *The impact of neuroscience on the free will debate*, op. cit., p. 65.

necessariamente, una visione dualistica per cui la volontà sarebbe responsabile di una modificazione cerebrale che corrisponderebbe all'azione da compiere.

Seguendo l'analisi di Dennett³³¹, si sostiene che tale interpretazione cade vittima del già citato "Teatro Cartesiano": il dualismo intrinseco fa sì da generare l'idea erronea secondo cui "ridurre" l'attività mentale al funzionamento dei processi neurali significati, inevitabilmente, eliminare o minacciare l'esistenza del libero arbitrio, poiché, messo in discussione il Sé cosciente, si eliminerebbe anche la volontà deliberatrice.

Solo quando si riuscirà a comprendere che la propria *ipseità* non comporta alcun allontanamento dall'attività cerebrale (in quanto il Sé psicologico "altro non sarebbe" che il funzionamento sinaptico) allora si comprenderà che, seppur le tesi di Libet e Wegner sostengono il vero, il libero arbitrio rimarrà indenne.

Tale posizione, però, per quanto possa sembrare convincente, incorre in alcuni problemi; uno dei quali è quello relativo all'assenza di una salda teoria pratica sull'esistenza del *free will* e della responsabilità morale.

Il punto debole di tali argomentazioni risiede nel fatto che un agente possa esser sempre ritenuto responsabile in qualsiasi caso di coscienza si trovi; in quanto il proprio Sé comprende tanto quei momenti "isolati" della consapevolezza cosciente quanto i processi e le azioni involontarie o inconscie. Si consideri, ad esempio, un sonnambulo che uccide durante il sonno così come è stato per Ken Parks: è o non è da ritenersi responsabile delle sue azioni?

Stando a tale posizione, egli lo sarebbe: uno dei suoi scopi principali è quello di stabilire una tesi filosoficamente difendibile della responsabilità morale, credendo che essa si basi sulle interazioni politiche e sociali dell'individuo in un qualsiasi contesto in cui essa sia implicita nelle azioni che si compiono³³²; perché mai, infatti, il Sé non dovrebbe includere i sofisticati processi inconsci che causano l'uccisione di un uomo mentre dorme? Nel caso del sonnambulo che uccide durante il sonno, allora, egli non dovrebbe esser giudicato colpevole in quanto l'agente ha compiuto un'azione al di fuori del suo controllo cosciente.

Tale approccio, però, anziché giovare *in toto* alla discussione sul tema in questione, per un verso elimina le sostanziali differenze che intercorrono fra le diverse posizioni che si sono da sempre occupate del tema del libero arbitrio; per un altro fa giungere ad una regressione del pensiero anziché ad una sua evoluzione³³³, nonostante si sia concorde sull'effettiva esistenza del *free will*.

³³¹ Cfr. DANIEL DENNETT, *Coscienza*, op. cit., pp. 160-195.

³³² Si rimanda il lettore alla lettura dell'ultimo paragrafo di questo capitolo.

³³³ Per informazioni più dettagliate, si consiglia la lettura dell'articolo già citato di STEPHEN MORRIS, *The impact of the Neuroscience on the Free Will Debate*, op. cit., pp. 67-73.

Se tale ipotesi di lavoro continua ad essere seguita da diversi neuroscienziati e filosofi, un altro filone di ricerca si sta concentrando sull'idea che l'intenzione cosciente di un'azione sia fortemente influenzata dagli eventi che accadono subito dopo che l'azione sia stata eseguita; è come se, vale a dire, le intenzioni si *ricostruissero* parzialmente a seconda delle conseguenze delle proprie azioni³³⁴.

Alcuni esperimenti hanno suggerito che posto il soggetto nelle medesime condizioni sperimentali di Libet – con la sola differenza di un segnale sonoro alla pressione del tasto – quando il *feedback* uditivo veniva ritardato rispetto alla simultanea pressione del tasto, i soggetti riportavano che l'intenzione cosciente di flettere il dito si presentava in concomitanza alla comparsa del segnale sonoro: nonostante, cioè, il compito fosse stato svolto prima della comparsa del *feedback* uditivo, i soggetti riportavano che le loro intenzioni coscienti si erano presentate solo *dopo* che si fosse presentato lo stesso. Le intenzioni, vale a dire, avevano subito una parziale *ricostruzione*, basandosi sulla conseguenza di un'azione compiuta. Differenti esperimenti, seppur con strumenti diversi, hanno riportato i medesimi dati, indipendentemente dallo specifico compito assegnato al soggetto; ciò ha suggerito l'idea che quando l'agente giudica introspettivamente il momento esatto in cui ha "l'intenzione di", il suo giudizio viene parzialmente ricostruito a seconda delle conseguenze delle proprie azioni. In conclusione, si può sostenere che la tesi di Wegner (la volontà è un'illusione) più che dimostrare l'assenza del ruolo causale della coscienza nelle azioni comportamentali, suggerisce, invero, che la coscienza ha un ruolo meno significativo di ciò che si crede. I casi da lui esaminati (sindrome della mano aliena e l'ipnosi) mostrano che l'esperienza della coscienza non è mai veritiera:

*The unusual cases from the psychological literature demonstrate only that the experience of will is not always veridical, not that it is never veridical*³³⁵.

In mancanza di dati neuroanatomici, che dimostrino che i processi cerebrali siano casualmente interconnessi, manca, in ogni caso, una prova convincente sull'illusorietà del libero arbitrio. Dovunque stia la verità, tali studi, in fase germinale, se da una parte promettono di contribuire

³³⁴ Cfr. DAVIDE RIGONI E MARCEL BRASS, *La libertà: da illusione a necessità*, paragrafo *L'intenzione cosciente come processo ricostruttivo* in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, op. cit.

³³⁵ Scrive Wegner: "I casi inusuali ripresi dalla letteratura psicologica dimostrano solo che l'esperienza della volontà non è sempre veritiera, ma che non lo è mai", cit. in EDDY NAHMIA, *Scientific Challenge to Free Will* in CONSTANTINE SANDIS E TIMOTHY O'CONNOR (A CURA DI), *A Companion to the Philosophy of Action*, Wiley-Blackwell, 2010, Hoboken, New Jersey, p. 351.

ad una conoscenza maggiormente adeguata dell'uomo stesso (uno fra gli scopi principali della neuroetica), dall'altra alimentano le già forti perplessità esistenti sulla realtà effettiva del *free will*.

Se la volontà (che si è detta essere la componente fondante del libero arbitrio) si rivela essere un gioco di prestigio prodotto dalla mente, di cui anche la sua percezione cosciente possa essere frutto di un adeguamento psicologico del soggetto alla situazione vissuta (come Wegner sostiene) allora – molti autori deducono – anche la libertà altro non sarebbe che un'illusione.

6.5 Il libero arbitrio, la responsabilità morale e le neuroscienze

Il problema a cui molti autori stanno tentando di rispondere è proprio quello di salvare il principio della responsabilità morale dalla deriva scettica, affinché non risulti controversa la sua assegnazione; è alquanto problematico, però, separare la responsabilità dalla libertà – per molti considerata il suo presupposto fondamentale – nonostante ci siano diversi tentativi in corso. Per riuscire in tale intento si dovrebbe dimostrare o l'indipendenza concettuale che lega la responsabilità alla libertà oppure l'effettiva esistenza della seconda affinché sia salva anche la prima.

La forte preoccupazione che angoschia i sostenitori della responsabilità morale indipendente dal concetto della libertà è quella relativa al rapporto fra il principio delle possibilità alternative ed il principio di causalità, declinato o in termini deterministici o indeterministici.

Il principio delle possibilità alternative, si ricordi, è quello che sostiene la capacità da parte dell'agente di scegliere fra diversi percorsi d'azione, affinché, insieme al controllo esercitato sulle medesime, siano rispettate le condizioni necessarie per parlare di atto libero in senso positivo. Il problema che sorge, invero, è come conciliare tale principio con la causazione necessaria tipica del determinismo o con l'assenza di causalità tipica dell'indeterminismo. La questione, allora, concerne il modo in cui conciliare il problema della libertà/responsabilità con i due schieramenti sopra nominati e, dato che risulta problematico spiegare il libero arbitrio con entrambi, diversi autori hanno provato a teorizzare il principio di responsabilità indipendentemente dal concetto di libertà, affinché risulti salvo da derive scettiche.

La prima ipotesi in questione sostiene che il vero punto dolente del principio di responsabilità è il suo rapporto con la causazione necessaria tipica del determinismo piuttosto che con la sua

assenza nel mondo indeterministico; il problema, infatti, riguarda la conciliazione fra il principio delle possibilità alternative e la causazione necessaria cui l'agente è sottoposto.

Se ad una prima lettura potrebbe sembrare assente il concetto di responsabilità morale nel contesto determinista, perché mancherebbe (per le difficoltà precedentemente esaminate) il libero arbitrio, l'individuo, essendo "impossibilitato" a scegliere fra diversi corsi d'azione e manifestando la sua scelta, subirebbe solamente le causazioni deterministiche del mondo; in realtà non è così, poiché nefaste sarebbero le ripercussioni pratiche: come condannare, ad esempio, un malfattore se non lo si considera responsabile delle sue azioni? La responsabilità, in altre parole, secondo una certa linea argomentativa, non potrebbe scindersi dalla libertà, ma deriverebbe da una rigida catena necessaria di cause e conseguenze che renderebbero vana qualsiasi metafisica della libertà.

Di contro, molti sostengono che il principio delle possibilità alternative (ritenuto quasi una sorta di verità *a priori* quasi come fosse – se non lo è davvero – il principio caratterizzante il mondo indeterminista) sia falso e protenderebbero, di conseguenza, per un mondo deterministico; almeno in questo la responsabilità morale sarebbe salva. Un agente, allora, deve essere ritenuto responsabile anche quando non sia nelle sue capacità fare altrimenti, poiché nessuna costrizione esterna può essere considerata motivo sufficiente a discolparlo dalle sue azioni; qualsiasi atto compia sarà sempre il risultato di una scelta volontaria operata dal medesimo soggetto morale per fini utilitaristici propri. La coercizione non escluderebbe, così, la responsabilità.

Per mostrare tale assunzione e la falsità del principio delle possibilità alternative, il filosofo Harry Frankfurt ricorre ad un esempio: si immagini che qualcuno minacci Jones con la condizione che se non compirà l'azione indicata dal malfattore subirà una pena da Jones stesso considerata inaccettabile. Le reazioni e le decisioni di Jones, posto di fronte a questo difficile *aut aut*, potrebbero essere diverse: nel primo caso ci si può immaginare Jones come una persona poco ragionevole, poco attenta alle conseguenze delle azioni e che abbia deciso di compiere l'azione impostagli dalla minaccia ben prima che questa facesse la sua comparsa. In tal caso, la coercizione esterna non avrebbe esercitato alcuna influenza sull'agire di Jones, invalidando la tesi secondo cui essa spoglierebbe di qualsiasi responsabilità morale l'agente sottoposto ad una sua coercizione.

Il secondo caso è che Jones, rimasto atterrito dalla minaccia, abbia deciso di obbedire al malfattore, dimenticandosi completamente le intenzioni che aveva prima del fatto traumatico: in tal caso, siccome a Jones si è prospettato solo il corso d'azione presentato dal malfattore, la

responsabilità morale è esclusa dall'atto di Jones, poiché non è nelle sue capacità pensare altrimenti.

Un'ultima possibilità è data dal caso in cui Jones abbia già precedentemente deciso di compiere l'azione che lo condurrà in modo inevitabile alla conseguenza da lui stesso ritenuta inaccettabile e che, in tale frangente, abbia ricevuto la minaccia dal malfattore: in tal modo, indipendentemente dal fatto che Jones sia o meno rimasto atterrito da tale minaccia, egli ha agito seguendo le sue motivazioni.

Il punto interessante è che, ipotizzando che egli abbia agito seguendo le sue ragioni e non quelle imposte dal malfattore, la coercizione esterna non ha influito sulle sue azioni, tanto da ritenerlo pienamente responsabile delle sue scelte. In questo caso Jones non è stato impossibilitato a compiere un'azione alternativa: avrebbe potuto sfidare la minaccia con le relative conseguenze, laddove avesse voluto.

Ciò, secondo Frankfurt, invalida il principio delle possibilità alternative, in quanto sia teoricamente sia praticamente. Jones avrebbe potuto fare altrimenti ed è stata solo una sua responsabilità accettare di piegarsi alla coercizione esterna, proponendo una rivisitazione del suddetto principio:

*una persona non è moralmente responsabile per ciò che ha fatto, se l'ha fatto solo perché non avrebbe potuto fare altrimenti*³³⁶.

Nelle tre ipotesi comportamentali di Jones, cioè, solamente il secondo caso lo salverebbe dalla responsabilità morale di compiere l'azione in quanto non esistono, né teoricamente né praticamente, possibilità alternative alla sua condotta. Stando, allora, così le cose, la versione del principio delle possibilità alternative modificata è, secondo Frankfurt, del tutto irrilevante per la corretta attribuzione della responsabilità morale, mostrando che responsabilità e libertà non sono concetti mutuamente inclusivi.

Indipendentemente, allora, dal rapporto fra la libertà, il determinismo e l'indeterminismo, risulta evidente che l'essenza della responsabilità rimane sempre la stessa: un agente è ritenuto responsabile delle proprie azioni anche nel caso in cui non possa fare altrimenti purché l'azione dipenda da una sua autonoma decisione: un agente (nel mondo determinista), che si trovi ad agire secondo la rigida causazione che lo necessita, sarà, allora, sempre ritenuto

³³⁶ HARRY FRANKFURT, *Possibilità alternative e libertà morale*, in MARIO DE CARO (A CURA DI), *La logica della libertà*, Meltemi Editore, Bari, 2002, pp. 117-132.

responsabile della sua condotta anche se teoricamente e praticamente non possa fare altrimenti.

Il principio delle possibilità alternative, modificato da Frankfurt, non è immune da confutazioni che minano la stessa posizione del filosofo statunitense, come De Caro evidenzia³³⁷.

La prima critica mossagli è che la sua argomentazione sembra viziata da una *petitio principii*, ossia da un ragionamento fallace in cui la proposizione da provare è presupposta nelle sue stesse premesse: nel caso in questione, la proposizione da dimostrare è la presenza della responsabilità morale di Jones.

Nel modo in cui Frankfurt ha condotto la sua argomentazione, la responsabilità è sempre inferita dalla premessa del ragionamento e mai provata effettivamente: Jones, ad esempio, potrebbe esser stato oggetto di malattie psichiatriche o di traumi cerebrali per cui sarebbe stato incapace di scegliere ragionevolmente l'azione da compiere. In tal caso, le considerazioni da fare sulla sua condotta sarebbero cambiate. Inoltre, Frankfurt non definisce mai il tipo di scenario in cui Jones si trova ad agire: se si trovasse nel mondo determinista, allora, per definizione, egli avrebbe la piena responsabilità delle sue azioni (e, in tal caso, l'ipotesi di Frankfurt non avrebbe aggiunto nulla di nuovo); se si trovasse nel mondo indeterminista sarebbe da spiegare come e dove avvenga la causazione indeterministica che spieghino l'effettivo controllo dell'azione da parte di Jones e la sua responsabilità morale.

In tal modo, fallisce un primo tentativo di dividere la responsabilità morale dal concetto della libertà, in quanto quest'ultima sembra essere il presupposto essenziale alla prima.

La seconda ipotesi sulla possibilità di salvare la responsabilità dalla deriva scettica nel caso in cui dovesse figurare inesistente il libero arbitrio è quella di Peter Strawson³³⁸.

Il filosofo inglese sostiene che la responsabilità è indipendente dal contesto deterministico o indeterministico, poiché trova la sua ragion d'essere negli "atteggiamenti reattivi" di cui l'uomo è capace. Solo dai rapporti interpersonali dipende la comparsa della responsabilità morale che ancora quest'ultima alla costituzione naturale dell'uomo stesso; un presupposto ben più saldo del concetto di libertà. Indipendentemente, allora, che si consideri il mondo determinista o indeterminista, l'uomo sarà sempre responsabile, perché il concetto stesso di responsabilità dipende dalla natura umana che interagisce con l'altro in qualsiasi circostanza: da un punto di vista biologico ed etologico è decisamente più vantaggioso, ai fini della sopravvivenza e della riproduzione, l'aiuto reciproco piuttosto che il solo egoismo.

³³⁷ Cfr. MARIO DE CARO, *Il libero arbitrio. Un'introduzione*, op. cit.

³³⁸ Cfr. PETER F. STRAWSON, *Libertà e risentimento*, in MARIO DE CARO (A CURA DI), *La logica della libertà*, pp. 77-116.

Il senso di responsabilità, discendente dalla convinzione di possedere il libero arbitrio, potrebbe essere, in effetti, un meccanismo adattativo, evolutosi per un essere sociale quale appunto l'uomo: anche laddove dovesse scoprirsi l'illusorietà del *free will*, ciò non cambierebbe e non intaccherebbe il senso di responsabilità delle azioni dell'agente che si ritroverebbe a rispondere in qualsiasi circostanza – deterministica o indeterministica – delle proprie azioni.

Qualcuno, però, potrebbe far notare che i dati provenienti dagli studi di Knobe e Nicholas sostengono esattamente il contrario: in un mondo privo di libero arbitrio, la gente comune sarebbe più incline ad atteggiamenti delinquenziali ed egoistici; come rispondere, allora, alle accuse rivolte a tali studi da Nahmias, secondo cui i due filosofi avrebbero indotto gli agenti a pensarsi scevri di responsabilità in un mondo completamente determinato? La risposta si ritrova nel fatto che per ovviare alle difficoltà incontrate (come spiegare la responsabilità, la volontarietà e la libertà?) negli ultimi anni sta nascendo un approccio a tali questioni che non si fossilizzi sui problemi apparentemente irrisolvibili del determinismo e dell'indeterminismo, ma che riparta dalla psicologia del senso comune³³⁹.

In che modo tale indirizzo di ricerca possa redimere le controversie e giungere ad un'ipotesi comunemente condivisa è ancora in fase di elaborazione.

Per orientare il discorso sul libero arbitrio, in modo diverso rispetto a quello classico, i fautori del nuovo approccio, Nahmias – *in primis* – insieme a Morris, Nadelhoffer e Turner, suggeriscono che il primo passo da compiere è partire da un nuovo e diverso studio fenomenologico del *free will*; tale convinzione prende avvio dalla constatazione che ogni autore ha sempre prodotto una teoria del libero volere partendo solo ed esclusivamente dalla sua fenomenologia dell'evento. Uno studio intimo ed introspettivo, utilizzabile per spiegare il comune senso con cui la gente intendeva e sentiva il proprio senso di agentività causale; ignorando, completamente, la fenomenologia del suo modo privato e personale di *vivere* il libero arbitrio. Proprio per tale motivo, Nahmias *et alii* suggeriscono una nuova (?) fenomenologia dello stesso.

³³⁹ Cfr. EDDY NAHMIAS, STEPHEN MORRIS, THOMAS NADELHOFFER AND JASON TURNER, *The Phenomenology of Free Will*, in "Journal of Consciousness Studies", 1 (7-8), 2004, pp. 162-179.

Capitolo VII

Verso un nuovo approccio al libero arbitrio?

7.1 Fra “certezze” ed “incertezze”: una soluzione impossibile?

Si è più volte avuto modo di ricordare nel percorso fin qui tracciato, che l'utilizzo e l'incremento delle nuove tecnologie sul cervello sono divenute complementari al discorso etico (la *neuroetica*, ne costituisce un esempio), ma, come in ogni nuovo “giocattolo” a disposizione dello scienziato, in queste tappe iniziali le sue regole di utilizzo sono ancora in fase di messa a punto; su questa scia si colloca la critica di Nahmias e dei tanti come lui che sostengono la medesima posizione, richiamando all'attenzione non tanto la legittimità dei mezzi utilizzati sull'uomo, ma, in sede di dibattito sull'effettiva o sull'illusoria esistenza del libero arbitrio, sul modo in cui i dati forniti dalla stessa Scienza vengono *interpretati*.

Il problema (e qui si ritorna agli esordi) è la traslitterazione di competenze che ha investito ogni figura professionale della società contemporanea in cui, ahimè, capita (come fosse un diritto) che un neuroscienziato *si convinca che possa, anche, essere* filosofo, relegando quest'ultimo, come fosse obsoleto, a rimanere “seduto in poltrona”, alla stregua di un vecchio ed arcaico strumento oramai *non più utile*. Il discorso si aprirebbe a riflessioni diverse e discordanti se si potesse, in questa sede, porre l'accento sulla legittimità del binomio

utilità = valore morale

non, però, sulle riflessioni già condotte da Mill o da Bentham, ma sulla scia di quelle tracciate dalla Scuola di Francoforte (Adorno, Benjamin, Marcuse...) in cui, sostanzialmente, la società

del falso bisogno ingloba in un vortice senza fine i tempi del pensiero, perché non a passo con la velocità del finto “progresso”. Il quale corre; corre senza sapere *dove* voglia giungere.

È su questa scia, allora, che oggi – forse più di qualsiasi altra era vissuta dall’uomo – si pone la reale *urgenza* che la Filosofia *debba* “restare in poltrona”, interpretando tale espressione non come una *diminutio*, ma come l’unica ricchezza, unica speranza per l’umanità; affinché ritrovi *il lume della ragione*.

Non deve stupire che tale discorso investa anche aspetti, forse, più specifici della cultura, quale la presunta o veritiera esistenza del *free will*, poiché ciò che accade è che in tale confusione di competenze e di ruoli la “Scienza” si arroghi il diritto di *capire* e non di *constatare*; in quest’ottica si debbono leggere i numerosi tentativi di fornire una risposta neuroscientifica al problema in oggetto, alimentando posizioni come quella di psicologi come Jonathan Bargh, il quale sostiene che:

*The phenomenological feeling of free will is very real [...] but this strong feeling is an illusion*³⁴⁰;

oppure del neuroscienziato John-Dylan Haynes, il quale scrive che dai dati ricavati dalla *fMRI*,

*[t]here’s not very much space for operation of free will. The outcome of a decisions shaped very strongly by brain activity much earlier than the point in time when you feel to be making a decision*³⁴¹.

Uno dei problemi che si pone sulla divulgazione di posizioni come questa è che tali dichiarazioni possono avere un impatto significativo, sia nell’ambito giuridico sia sul singolo aspetto cognitivo –identitario di qualsiasi agente –, poiché il modo in cui i *mass-media* o la stessa *Scienza* presentano le loro convinzioni non è adeguato alla comprensione del problema. È, perciò, importante, sostiene Nahmias, preso a simbolo di una corrente di pensiero che rivendica, legittimamente, il ritorno ed il ripristino del rispetto delle proprie competenze, considerare le nuove scoperte in un’ottica mediata dal buon senso, dalla ragionevolezza e, soprattutto, da attente riflessioni. Titoli di giornale del tipo:

³⁴⁰ JONATHAN BARGH, *Free will is un-natural* in JOHN BAER, JAMES KAUFMANN AND ROY BAUMEINSTER (A CURA DI), *Are we free? Psychology and Free Will*, in “International Journal of Philosophical Studies”, 20 (5), pp. 725-734.

³⁴¹ JOHN DYLAN HAYNES, *Case closed for free will?* In *Science NOW Daily News*, aprile 2008.

*Free will is not the defining features of humanness, modern neuroscience implies, buti s rather an illusion that endures only because biochhemical complexity conceals the mechanisms of decion making*³⁴²;

oppure

*And since all behavior is caused by our brains, wouldn't this mean all behaviorcould potentially be excused?*³⁴³

non fanno che incentivare false convinzioni sull'effettiva illusorietà del libero arbitrio, alimentando scetticismi e discorsi viziati dalle premesse, che non hanno alcuna corrispondenza con ciò che, attualmente, si è appreso, dando luogo (aspetto più pericoloso), oltre che a derive nichiliste, a comportamenti ai limiti della legalità; poiché gli agenti vengono indotti a credere all'inesistenza della libertà e, di conseguenza – per molti – alla propria responsabilità morale.

Da un punto di vista etico, il *punctum dolens* si attornia sul dilemma se sia possibile compensare e preservare la metafisica della scelta con il determinismo; non a caso sono nate diverse correnti all'interno del *determinism*, che, come visto, tentano di conciliare la rigida necessitazione causale con il “libero” volere dell'uomo. La critica/constatazione che Nahmias rivolge ad un certo panorama culturale attuale riguardo la Scienza e l'interpretazione che essa offre dei dati da lei stessa ricavati, si concentra su un errore concettuale di fondo; non si tratta, come detto, di una superficialità sistemica nell'affrontare riflessioni “care alla filosofia”, ma dell'impostazione metodologica.

In genere, il problema sull'esistenza del *free will* è la presunta verità del determinismo per cui gli scienziati (sulla via di Libet, Soon *et alii*) stanno tentando di “dimostrare al mondo” che se sussiste una causazione sinaptica non può derivarne la “libertà di volere” dell'agente stesso, risultato di cause ed effetti al di fuori del suo controllo; l'io, in altri termini, si troverebbe a “subire” ciò che “il cervello” (qualsiasi ente o insieme di funzioni denoti tale termine) compie.

È difficile, infatti, trovare un verbo che possa spiegare un'attività indipendente del *brain* senza rischiare di cadere in posizioni o troppo radicali o troppo “ingenua”; perché è altrettanto

³⁴² Cfr. la Rivista SCIENCENOW, edizione del 12 giugno 2007.

³⁴³ Cfr. JEFFREY ROSEN, *The brain on the stand*, in “The New York Times”, 11 marzo 2007.

complicato capire e condividere una teoria sulla mente che non tradisca i dati ricavati dalle neuroscienze ed il vizio di molti di volerli interpretare adattandoli alla propria visione del problema.

Indipendentemente dai dubbi scettici sulla metodologia utilizzata, riprendendo l'oramai già tracciato filo rosso, Nahmias sostiene che il ragionamento della "Scienza" è viziato già dalla premessa; ossia

[D1] *Free will requires that determinism is false.*

[D2] *Science is showing that determinism is true.*

[D3] *So, science is showing that we do not have free will*³⁴⁴.

Ciò renderebbe evidente che:

*one problem with the argument is that the scientists simply assume that premise is true, ignoring the substantial philosophical literature supporting compatibilism, the view that free will does not require the falsity of determinism*³⁴⁵.

L'ingenuità con cui la Scienza *assume*, quasi dogmaticamente, la verità della prima premessa è emblematica: di quale determinismo, infatti, si sta parlando?

Peter van Inwagen, ad esempio, ha evidenziato che con esso si debba intendere una tesi sul mondo che fornisca, momento per momento, la descrizione completa di un singolo sistema o dell'universo tutto, connaturato alle medesime leggi fisico-chimiche che lo rendono possibile; in aggiunta, van Inwagen non dimentica di fornire l'ultima *conditio sine qua non*: la possibilità, anzi, l'assoluta certezza, che con quei dati a disposizione si sia in grado di prevedere l'evoluzione di quel sistema con tutti i suoi singoli componenti.

Non è lontano, vale a dire, dall'arte della lettura degli auspici con la sostanziale differenza che nel mondo determinista la previsione del futuro non sarà opera di magie o stregonerie varie, ma del risultato *certo* proveniente da premesse altrettanto *certe*: lo stato dei fatti al tempo dell'interrogazione determinista. Un sistema chiuso, cioè, in grado di prevedere, con accuratezza e precisione, ciò che accadrà *certamente* in un futuro istante *t*, poiché

³⁴⁴ EDDY NAHMIA, *Scientific Challenge to Free Will*, in CONSTANTINE SANDIS E TIMOTHY O' CONNOR (A CURA DI), *A companion to the Philosophy of Action*, Wiley-Blackwell, Hoboken (New Jersey), 2010, pp. 345-356.

³⁴⁵ *IDEM*, p. 346.

*In a deterministic system, given identical circumstances, the same causal antecedent will always have the exact same effects*³⁴⁶.

Ne consegue, allora, che gli scienziati non possono essere in grado di porre un veto o di fornire una garanzia alla veridicità o meno della tesi esaminata, poiché l'atteggiamento speculativo assunto non consente loro di valutare o di scindere le inclinazioni incompatibiliste attraverso le quali potrebbero leggere i correlati statistici forniti dalle tecnologie utilizzate; per cui, sostiene Nahmias, ciò che *realmente* lo scetticismo degli scienziati dimostra è l'assunzione di una tesi sul mondo che nulla ha a che fare con determinismo/indeterminismo. Ad un'accurata analisi, allora, si evince che essi sposano il *Mechanism*, cioè:

*[...] is the view that mental phenomena can be fully explained in terms of their component neurobiological parts and in terms of the organization of, and interactions between, these parts*³⁴⁷,

comportando, di conseguenza, nessuna differenza fra la veridicità eventuale del *Mechanism* e l'esistenza del libero arbitrio o, viceversa, la veridicità eventuale del *Mechanism* e la veridicità eventuale del determinismo.

Un primo errore concettuale operato dagli scienziati è quello di esser convinti, quindi, di adottare il determinismo, quando, in realtà, tale assunzione dogmatica si mostra, appunto, non ammissibile e non veritiera. Essi, infatti, si confondono con il *meccanicismo*, ossia quella visione secondo cui i fenomeni possono essere spiegati attraverso il ricorso alle loro componenti neurobiologiche e di organizzazione totale del sistema entro cui sono inseriti, interagendo costantemente fra loro³⁴⁸. Data la sua natura "microcellulare", il meccanicismo potrebbe esser vero se il determinismo fosse falso (ad esempio se le parti di un sistema *meccanico* interagissero in modo indeterministico) o, al contrario, esser falso se un sistema fosse deterministico (si pensi al caso in cui le leggi della natura fossero deterministiche come quelle psicologiche con la differenza che queste non potessero essere spiegate in termini di processi fisici). In tali e talaltri casi simili, la confusione fra meccanicismo e determinismo non può essere presentata come la certezza incontrovertibile dell'inesistenza al *free will*, inficiando, in tal modo, gli studi, le riflessioni e, cosa più grave, il comportamento di tutti

³⁴⁶ *IVI*.

³⁴⁷ Cfr. CARL CRAVER, *Explaining the Brain: Mechanism and the Mosaic Unity of Neuroscience*, Oxford University Press, New York, 2007.

³⁴⁸ *IVI*.

quegli agenti che si vedrebbero privati sia della loro *agentività causale* sia della personale *responsabilità* legata alle proprie azioni. Sostenere che un sistema più o meno complesso mantenga il suo equilibrio e perseveri nella sua esistenza attraverso particolari componenti (in questo caso di natura neurobiologica) non significa escludere che lo stesso possa interrelazionarsi all'ambiente circostante in un modo "libero", indipendentemente dal modo in cui, cioè, il principio causale sia o rigidamente presente o probabilisticamente ammesso; la *constatazione* dei dati non implica la loro diretta *interpretazione*.

I problemi sorgono sia quando questa tesi viene reinterpreta alla luce della fisica dei quanti, che dimostrerebbe la sua falsità epistemica e pragmatica (la premessa D2) sia quando l'incompatibilismo quantico viene condiviso dagli stessi filosofi nella speranza che fornisca una tesi coerente e senza minacce all'esistenza del libero arbitrio. A ciò, precisa Nahmias, si aggiunge un problema ancora più pregnante: coloro i quali sostengono l'inesistenza del *free will* sono neuroscienziati e psicologi, che per impostazione professionale e per il percorso formativo intrapreso, ignorano in parte o completamente l'intera letteratura filosofica a riguardo.

Scrive Nahmias:

While, in practice, these scientists may aspire to offer deterministic explanations and laws for the systems they study, these sciences are not in a position to establish determinism. Most of their discoveries involve statistical correlations that are compatible with indeterminism. Moreover, it is implausible to think that indeterministic interactions at the microphysical level never have an effect on the way things happen at, say, the neurobiological level. And 'near-enough determinism' is simply not sufficient to make incompatibilist arguments go through (i.e. to justify premise D2)³⁴⁹.

In tal senso, il ragionamento fin qui condotto dalla Scienza, in genere, non può non mutarsi come segue:

- [M1] *Free will requires that mechanism is false.*
- [M2] *Science is showing that mechanism is true.*
- [M3] *So, science is showing that we do not have free will³⁵⁰.*

³⁴⁹ *IVI.*

³⁵⁰ EDDY NAHMIAS, *Scientific Challenges to Free Will*, op. cit., p. 347.

Così come accade con quello riguardante il determinismo, anche in questo caso, sostiene Nahmias, vige una consapevolezza errata delle scoperte che la Scienza porta avanti: mancando una visione “oggettiva” del funzionamento olistico del cervello è più probabile che i dati ricavati non mostrino la verità del *mechanism* (premessa M2), ma, piuttosto, *assumano* un ragionamento euristico su probabili spiegazioni del *decision making* o del comportamento; che possano *spiegare* l’inconsistenza di qualsiasi teoria ribadisca l’esistenza del libero arbitrio. Ciò, allora, conduce, inevitabilmente, all’inesattezza concettuale della prima premessa; quella, cioè, secondo cui il *free will richiede* che il *Mechanism* sia falso: perché mai, si chiede Nahmias, dotare il libero arbitrio di componenti neurobiologiche – spiegare, cioè, come “funzioni” meccanicamente la presa di decisione cosciente – significherebbe privare l’uomo “di speciali poteri” connaturati alla sua natura? In che modo, cioè, tale spiegazione renderebbe vana l’esistenza della libertà?

È con quest’ottica e non con l’intento con cui, effettivamente, le ha scritte che si dovrebbero leggere le parole del neuroscienziato Read Montague; il quale, però, parte da una prospettiva differente: quella secondo cui la naturalizzazione (ammettendo che sia possibile) delle capacità decisionali significherebbe, immediatamente, la negazione tanto della libertà quanto, più specificatamente, del libero volere:

*Free will is the say idea that we make choices and have thoughts independent of anything remotely cousin to the idea of the soul – the concept that ‘you’, your thoughts and feelings, derive from an entity that is separate and distinct from the physical mechanism that is separate and distinct from the physical mechanisms that make up your body... Consequently, the idea of free will is not even in principle within reach of scientific description*³⁵¹.

Non è difficile scorgere dietro le parole del neuroscienziato (esemplificative dei tanti che condividono le medesime posizioni) un timore diffuso che, purtroppo, identifica nel binomio

Naturalizzazione = soppressione

l’idea che tentare di capire come “il cervello funzioni” significhi eliminare, completamente, “speciali poteri” coscienziali all’agente. Ciò ingloberebbe, ovviamente, l’intera natura umana, giungendo ad una deriva nichilista su temi quali coscienza, libero arbitrio, volontà, responsabilità; poiché, forse per l’influenza ininterrotta che Cartesio continua ad esercitare dal XVII secolo, risulta difficile ai più non condividere o non “cadere” in spiegazioni dei

³⁵¹ READ MONTAGUE, *Free will*, in “Current Biology”, 18, 2008, p. 584.

fenomeni di stampo dualistici. Esempio emblematico, si ricordi, è lo stesso Libert, il quale giunse ad ipotizzare la *libertà di veto* pur di non mandare alla deriva l'importanza che il controllo sulle proprie azioni esercita sull'agente quale aspetto identitario di se stessi.

Secondo Nahmias, prima di sostenere che i dati neuroscientifici possano minare l'esistenza del *free will*, si dovrebbe primariamente essere in grado di assumere che la volontà cosciente sia riconducibile ad un *naturalistic account*³⁵², ossia ad un "profilo naturalizzato" (si potrebbe tradurre) che funga da principio metodologico piuttosto che da una tesi propriamente detta:

*If one stipulates that free will requires a nonphysical soul, then free will would face metaphysical objections, dating back to Descartes, which do not require scientific discoveries to illustrate – namely, explaining how non physical minds can causally influence physical bodies. If the challenge to free will is really supposed to derive from naturalism, then no specific discoveries from cognitive science do much to advance this challenge. General conclusions about the metaphysics of mind are unlikely to be illuminated with specific scientific findings, including the oft-discussed ones about where and when various events happen in the brain*³⁵³.

Bisognerebbe mutare, come già ribadito, la convinzione di fondo che la naturalizzazione della natura umana non comporterà un annichilimento delle sue peculiarità, ma un suo "disvelamento" adeguato. In questo senso, allora, le scoperte scientifiche potrebbero *aiutare a spiegare* il libero arbitrio, piuttosto che *spiegarlo*, dimostrando, ancora una volta, che la sola Scienza (per via delle limitazioni cui il suo metodo la pone) non può essere in grado di sostenere o di negare *de facto* aspetti della natura umana complicati perfino per gli stessi filosofi. L'ultima parola, cioè, non spetta a lei.

Si giunge, allora, all'ennesimo mutamento di ragionamento a fondamento del paradigma scientifico, utile a spiegare i dati ricavati secondo quel "profilo naturalizzato" che dovrebbe aiutare la Scienza stessa ad ovviare ai limiti interpretativi dei ragionamenti precedenti per avviare una *interpretazione* dei dati quanto più *oggettiva* possibile:

[E1] *Free will requires that one's actions properly derive from one's conscious reasoning, decisions, and intentions.*

³⁵² EDDY NAHMIA, *Scientific Challenges to Free Will*, op. cit., p. 348.

³⁵³ EDDY NAHMIA, *Is free will an illusion? Confronting Challenges from the Modern Mind Science*, in WALTER SINNOTT-ARMSTRONG (A CURA DI), *Moral Psychology: Free Will and Moral Responsibility*, MIT Press, Massachusetts, 2014, p. 10.

[E2] *Science is showing that our actions do not properly derive from our conscious reasoning, decisions and intentions.*

[E3] *So, science is showing that we do not have free will*³⁵⁴.

Secondo Nahmias, ad un'analisi approfondita, si evince che, se le scoperte scientifiche sono conciliabili nelle espressioni [E2] [E3], dovrebbe essere chiaro che la prima premessa (a differenza delle precedenti) rimanda ad una posizione ben precisa all'interno del dibattito sul *free will* e non ad una assunzione derivante dai dati forniti dalle stesse tecnologie utilizzate; la teoria sottesa alla premessa [E1] viene definita da Nahmias stesso *modular epiphenomenalism*³⁵⁵, secondo cui

*The neural correlates of our conscious experience of deciding or intending an action (the 'C module') are not causally responsible for producing that action; instead, distinct, non-conscious processes or modules (the 'NC modules') cause the action, while NC modules also cause the activity in the C modules. Hence 'conscious will' is an illusion (Wegner 2002) in that the processes underlying our conscious choices and intentions do not in fact cause our action – they are epiphenomenal*³⁵⁶.

Reinterpretando le principali posizioni che si sono succedute dagli anni Ottanta ad oggi, Nahmias è convinto che un esempio di spiegazione *epifenomenistica* del libero arbitrio sia proprio quella di Libet; poiché, grazie al concetto di *potenziale di prontezza*, il neurofisiologo ha fornito una descrizione fisiologica del *potere di veto* (che – come si ricorderà – l'agente eserciterebbe 250mms prima della presa di decisione di muovere o meno il dito); l'assunzione metafisica implicita nel discorso di Libet è che esiste uno scarto temporale fra la coscienza ed i correlati neurali: più tardiva ad “apparire” la prima, più “immediati” e più “veloci” i secondi. Il problema, ancora una volta, si incentra sul fatto che i dati a disposizione sono interpretabili in molteplici modi e si sposano con molteplici ipotesi di lavoro differenti, tanto che da soli non sono sufficienti a mostrare che l'attività neurale non conscia causi *deterministicamente* una particolare azione. Lo stesso Libet, infatti, non condusse delle ricerche sugli agenti che avessero, a pari condizioni, il desiderio di flettere o di non flettere il dito: cosa vieta di pensare che i dati sul “potenziale di prontezza” potessero essere diversi o, addirittura, inesistenti?

³⁵⁴ EDDY NAHMIA, *Scientific Challenges to Free Will*, op. cit., p. 348.

³⁵⁵ Cfr. EDDY NAHMIA, *When consciousness matters: a critical review of Daniel Wegner's The Illusion of Conscious Will*, in “Philosophical Psychology”, 15, 2002, pp. 527-541.

³⁵⁶ IDEM, p. 348.

Ciò sembrerebbe ridurre anche il ruolo della coscienza a mera spettatrice di quel “Teatro Cartesiano” di cui parla Dennett³⁵⁷, “bollando” il discorso su un dualismo di fondo *assunto*, quasi dogmaticamente, come veridico.

Il medesimo errore interpretativo è stato condotto dallo stesso *team* di Soon nel 2008³⁵⁸, il quale asserisce che:

*two specific regions in the frontal and parietal cortex of the human brain had considerable information taht predicted the outcome of a motor decision the subject had not yet consciously made*³⁵⁹,

sostenendo, quindi, che

*subjective experience of freedom is no more than an illusion and than an illusion and that our actions are initiated by unconscious mental processes long before we become aware of our intention to act*³⁶⁰.

Lo stesso esperimento di Soon *et alii*, evidenzia Nahmias, potrebbe benissimo non mostrare la distanza temporale tra l’attività cerebrale e la consapevolezza di muovere il dito, proprio a causa dei gravi *deficit* tecnologici in cui l’esperimento incorre.

In tali circostanze, allora, l’interpretazione possibile sarebbe, da una parte, la mancanza effettiva di spontaneità dell’atto – e, quindi, si spiegherebbe il perché dell’assenza completa della coscienza – da parte degli agenti per via della ripetitività ed automaticità del gesto da compiere; dall’altra, che questi ultimi non avessero l’*intenzione* di agire, ma, piuttosto, la consapevolezza di compiere un atto.

Stessa posizione riscontrata in Daniel Wegner (il quale sostiene che ai fini del libero arbitrio parlare di *volontà cosciente* è irrilevante, poiché la stessa coscienza sarebbe il risultato di un’attività cerebrale prim’ancora che ne diventiamo consapevoli)³⁶¹ e nei tanti che condividono l’idea secondo cui parlare di processi coscienziali in termini neurali equivalga a *constatare* l’illusorietà del libero arbitrio per quelle motivazioni che più volte si sono ribadite. Secondo Nahmias, le conclusioni di tali tesi dipenderebbero da tre discutibili assunzioni:

³⁵⁷ Cfr. DANIEL DENNETT, *La coscienza. Che cos’è*, op. cit.

³⁵⁸ Cfr. CHUN SIONG SOON, MARCEL BRASS, HANS-JOCHEN HEINZE E DYLAN HAYES, *Unconscious determinants of free decisions in the human brain* in *Nature Neuroscience*, 11, 2008, pp. 543-545.

³⁵⁹ *IBIDEM*, p. 545.

³⁶⁰ EDDY NAHMIA, *Is free will an illusion?*, op. cit., p. 10.

³⁶¹ Cfr. DANIEL WEGNER, *The illusion of Conscious Will*, op. cit.

1. L'identificazione dell'attività neurale inconscia che precede la consapevolezza sarebbe interpretabile come "decisione motrice" dell'intenzione ad agire;
2. La conclusione che i 250ms siano il simbolo di un'attività che *bypassi* i processi coinvolti nella formazione dell'intenzione cosciente piuttosto che rappresentare il lavoro che avviene *attraverso* tali processi;
3. L'identificazione di *esperienza cosciente* con *decisione cosciente* (e, di conseguenza, l'identificazione del tempo delle loro risposte con il tempo delle loro decisioni coscienti)³⁶².

Ci si troverebbe di fronte a due alternative significative: se le prime due premesse sono false, allora l'attività cosciente potrebbe rappresentare il precursore causale dell'attività che darà luogo alle decisioni o alle intenzioni coscienti; in tal modo, la prima non potrebbe essere interpretata quale rappresentazione di una *decisione* attuale. Se, al contrario, la terza premessa fosse falsa, allora gli agenti potrebbero, semplicemente, riferire la consapevolezza di un *impulso* a muoversi, piuttosto che una decisione cosciente, conseguente alle istruzioni dell'esperimento che non prevedeva in Libet e Soon una particolare mobilità o spontaneità d'azione da parte degli stessi soggetti coinvolti. Non vi sarebbe stata, da parte di questi ultimi, la constatazione di una differenza significativa fra un mero *impulso* a flettere il dito e la *coscienza* di volerlo effettivamente.

In un caso o in un altro, la rilevazione dei dati non fornì alcuna *prova* della veridicità di una o di tutte e tre le ipotesi conclusive.

Contro la prima, ad esempio, si potrebbe sostenere che l'attività registrata nella SMA potrebbe rappresentare quella solita che precede la consapevolezza o, semplicemente, il normale flusso elettrico delle sinapsi; si deve esser cauti, infatti, nell'etichettare un "fenomeno constatato" come "la prova" del correlato neurale della decisione cosciente o dell'intenzionalità, poiché, come è risaputo, l'interpretazione dei dati non è mai univoca. Tale considerazione vale ancor più se si considerano la decisione o l'intenzionalità (contro la seconda): dopo tutto, se assumiamo che la coscienza abbia uno o più correlati neurali, allora dovremmo aspettarci che l'esperienza cosciente non sorga dal nulla! Infatti, la stessa strutturazione dell'esperimento libetiano potrebbe, come hanno constatato Pockett e Purdy³⁶³,

³⁶² EDDY NAHMIAS, *Is free will an illusion?*, op. cit., p. 14.

³⁶³ Cfr. SUSAN POCKET E SUZANNE PURDY, *Are voluntary movements initiated preconsciously? The relationship between readiness potential, urges and decisions*, in WALTER SINNOTT-ARMSTRONG E LYNN NADEL (A CURA DI), *Conscious Will and Responsibility: A Tribute to Benjamin Libet*, Oxford University Press, New York, 2010.

infiare i dati ricavati, destrutturizzando tutte le successive interpretazioni che Libet stesso ed altri dopo di lui ne hanno dato.

Da ciò, secondo Nahmias, si ricava che:

*There is simply no evidence yet to show that such conscious deliberation, reasoning and planning lack these causal effects on what we do or that our conscious monitoring of our behavior is not critically involved in how we carry out and adjust our actions. On the contrary, there is evidence that conscious ‘implementation intentions’ influence actions*³⁶⁴.

Dopo tutto, se si assume che il libero arbitrio richieda fattori metafisici per il suo esercizio, allora qualsiasi spiegazione meccanicistica della presa di decisione cosciente suggerisce l'illusorietà dello stesso (come già ribadito a proposito del naturalismo); i problemi relativi a posizioni simili a quelle di Wegner o Gazzaniga (esemplificativi, ognuno per il proprio ambito, del modo in cui intendere il libero arbitrio) sarebbero, in linea generale, due:

1. L'implicita ed immotivata assunzione che la coscienza non possa essere naturalizzata;
2. Anche ammettendo che essa sia spiegabile grazie a processi neurali, la psicologia quanto le neuroscienze peccano, secondo Nahmias, di parzialità: a seconda, cioè, dell'aspetto del dibattito che meglio definisce o spiega ciascuna posizione, privilegiano uno o un altro aspetto³⁶⁵.

Il *focus* di queste e diverse altre posizioni simili, sostiene Nahmias, si incentra sul fatto che l'intera discussione si attornia sul dato neuroscientifico, scambiando quelle che sono teorie sulla mente (la *sopravvenienza* del mentale sul fisico, ad esempio, ne è una fra queste) per mere discussioni scientifiche; travalicando non solo i confini conoscitivi delle competenze che sono proprie agli “addetti al lavoro”, ma minando alla base un ragionamento con le medesime premesse sbagliate.

È in quest'ottica che dovrebbero essere considerati i casi della letteratura psico-scientifica (il fenomeno dei cervelli divisi, la “sindrome della mano aliena” e simili), i quali dimostrerebbero solo che l'esperienza della volontà non è sempre attendibile; non che non lo sia mai.

³⁶⁴ EDDY NAHMIA, *Is Free Will an Illusion?*, in WALTER SINNOTT-ARMSTRONG (A CURA DI), *Moral Psychology: Free Will and Moral Responsibility*, op. cit., p. 17.

³⁶⁵ *IDEM*, p. 12.

Senza i dati neuro-anatomici oggi a disposizione si potrebbe “dimostrare” l’illusorietà della coscienza solo attraverso la guida delle “illusioni visive”, le quali, a loro volta, mostrerebbero che la rilevanza dei processi cerebrali sottesi potrebbero non essere causalmente interconnessi; l’ambiguità e la facilità con cui l’occhio umano cade in errore è una prova più che attendibile della cautela con cui ci si dovrebbe avvicinare a temi tanto delicati quanto controversi.

Scrivo, a tal proposito, Nahmias:

[...] *indeed, as with visual illusions, explanations for illusions of will may be offered in terms of a generally reliable system, which sometimes produces inaccurate output because of some unusual feature of the situation (all the cases discussed by Wegner are unusual in important ways). The fact that we sometimes perform complex behaviours with conscious intentions (for example, under hypnosis or with subliminal priming) does not show that, on the many occasions when we do perform complex behaviours with conscious intentions, the conscious mental processes are causally irrelevant*³⁶⁶.

Sicuramente tali interpretazioni potrebbero essere o tutte giuste o tutte errate, ma anche laddove fosse “oggettivamente vero” che in determinate circostanze si agisse senza coscienza per poi divenir consapevoli dell’azione eseguita *post hoc* ciò non minaccerebbe l’esistenza del libero arbitrio; rivelerebbe solo parte del meccanismo neurobiologico della presa di decisione; sistema ben diverso dal *free will*. Alcune ricerche sul giudizio morale, infatti, sembrerebbero mostrare proprio questo: solo dopo aver eseguito un’azione, ci si ferma ad analizzare profondamente le cause e le ragioni che hanno spinto l’agente a compierla³⁶⁷.

Uno degli aspetti significativi su cui riflettere dal dialogo *Scienza/Filosofia* potrebbe essere il seguente: ammettendo che la deliberazione cosciente abbia una base neuro-biologica, ciò potrebbe “dimostrare” che, forse, il ruolo cosciente della stessa volontà possa essere meno influente di quanto si possa pensare.

Ciò suggerirebbe, allora, una sfida ancora più pregnante all’esistenza eventuale o reale del libero arbitrio, mutando il ragionamento di fondo del paradigma scientifico come segue:

³⁶⁶ EDDY NAHMIA, *Scientific challenges to Free Will*, in TIMOTHY O’CONNOR E CONSTANTINE SANDIS (A CURA DI), *A companion to the Philosophy of Action*, op. cit., p. 351.

³⁶⁷ Cfr. JOSHUA GREEN, *The secret joke of Kant’s soul*, in WALTER SINNOTT-ARMSTRONG (A CURA DI), *Moral Psychology. The neuroscience of Morality: Emotion, Disease and Development*, MIT Press, Cambridge, 2007, pp. 35-79; JONATHAN HAIDT, *The emotional dog and its rational tail: a social intuitionist approach to moral judgment*, in “Psychology Review”, 108, 2001, pp. 814-834.

[R1] *Free will requires that one's actions properly derive from decisions or intentions that one has at some point consciously considered, or at least that one would accept, as one's reasons for acting;*

[R2] *Science is showing that our actions do not properly derive from decisions or intentions that we have consciously considered or would accept as our reasons for acting. Rather, our decisions are produced by other factors and we rationalize them after the fact;*

[R3] *So, science is showing that we do not have free will*³⁶⁸.

Sulla premessa [R1], Nahmias sostiene che sarebbe plausibile sostenere che molte teorie filosofiche sul *free will* sarebbero incline a condividerla, benché le ricerche neuroscientifiche non aggiungano molto né al dibattito né alla verità della premessa [R2] da cui, semplicemente, si potrebbe constatare che, come detto, si dovrebbe prender atto che, in talune circostanze, la coscienza potrebbe avere un ruolo meno significativo di quanto si creda. Tale evidenza, però, aprirebbe un ulteriore problema nel problema: da un punto di vista psicologico, in che modo gli agenti reagirebbero “scoprendo” che, probabilmente, sono meno consapevoli di quanto credono?

Diverse ricerche hanno dimostrato che quando le persone apprendono dalla Scienza che libero arbitrio, responsabilità morale, coscienza sono solamente parole senza alcun denotato il loro comportamento muta in peggio: è come se le remore morali del vivere sociale venissero completamente abbandonate, per lasciarsi andare ad atteggiamenti o medesimi comportamenti delinquenziali, poiché convinti che non ne siano loro stessi gli artefici ed i responsabili; gli studi condotti nel 2008 da Vohs e Schooler sono, a tal proposito, esemplificativi.

L'idea condivisa da molti è che (semplificando il discorso), ognuno sia dotato di “poteri particolari”, quasi “magici”, che lo rendano un agente “libero” e consapevole per cui ciò porrebbe, ancor di più, l'urgenza di fornire una *risposta* a che cosa sia il libero arbitrio.

Non è un caso, infatti, che quando la Scienza lascia intendere nei suoi (falsi) ragionamenti che *free will* e responsabilità morale siano delle illusioni, il perno su cui dovrebbe porre chiarezza è la spiegazione di che cosa possa mai essere questo “speciale potere dell'agente” di fornire allo stesso la sua unità psicologica quale *libero* e *consapevole*. Ancora una volta, sostiene Nahmias, il discorso diviene viziato da premesse e ragionamenti erronei; stavolta, però, in grado di confondere ancor più le acque sul *fattore X* in grado di *spiegare* l'agentività causale del soggetto.

³⁶⁸ EDDY NAHMIAS, *Scientific challenges to Free Will*, op. cit., p. 354.

Per rendere esemplificativo il discorso, si potrebbe usare il seguente sillogismo:

1. *Free will requires that X is not the case.*
2. *Science is showing that X is the case (for humans).*
3. *Thus science is showing that humans lack free will*³⁶⁹.

Assumendo che il libero arbitrio coinvolga le capacità che permetta all'agente di essere moralmente responsabile per le sue scelte ed azioni, la conclusione di questo ragionamento dovrebbe essere che:

4. *Science is showing that humans are not morally responsible agents*³⁷⁰.

Questo schema mette in luce che qualsiasi ragionamento si compia sul libero arbitrio dipende da ciò che è individuabile nel “fattore X” che le premesse 1 e 2 pongono quale perno centrale al sillogismo. Ulteriore problema, allora, sarà quello di identificarne un significato, *in primis*, per fornirne una spiegazione (scientifica), *in secundis*; poiché un errore interpretativo di quel sistema o di quell'insieme di sistemi che potrebbero essere la causa del *free will* comporterebbero conseguenze “disastrose”: scoperte scientifiche falsamente giustificabili e la convinzione per la gente comune di non possedere l'agentività causale ne sarebbero solo alcune fra quelle ipotizzabili.

Come capita spesso, infatti, gli scienziati su tale aspetto continuano ad essere molto ambigui su ciò che dovrebbe denotare il *quid* dell'agentività causale, lasciando aperte diverse interpretazioni su che cosa possa mai essere. Per quanto si possa riflettere sulla natura di attribuzione del “fattore X” in un modo o in un altro, resta centrale il fatto che qualsiasi costruzione paradigmatica del ragionamento scientifico volta a *spiegare* fenomeni “metafisici” della cognizione umana errerà sempre per le ragioni già esplicitate nel corso del percorso fin qui condotto.

In questo esercizio di pensiero, che lungi dall'essere un'apologia di sé, pone all'attenzione un messaggio forte e, come più volte ribadito, urgente: forse più che in altre epoche, le scoperte neuroscientifiche pongono il *Filosofo* di fronte alla necessità di capire, ponderando e riflettendo, non tanto (per come la vede la scrivente) su nuovi problemi e nuove difficoltà, ma

³⁶⁹ EDDY NAHMIA, *Is Free Will an Illusion? Confronting Challenges from the Modern Mind Science*, in WALTER SINNOTT-ARMSTRONG (A CURA DI), *Moral Psychology: Free Will and Moral Responsibility*, MIT Press, Massachusetts, 2014, p. 5.

³⁷⁰ *IVI*.

su un modo più “adeguato alla realtà” di quegli stessi problemi e di quelle stesse difficoltà da sempre incontrate.

7.2 Il *free will* e la *folk psychology*

Immaginiamo, sostiene Nahmias, un futuro non troppo lontano dalla nostra contemporaneità in cui alcuni neuroscienziati possano prevedere, con un’accuratezza pari al 100%, le volizioni ed i desideri di un agente che indossi, come cappuccio, una speciale *brain imaging* in grado di anticipare prim’ancora che al soggetto stesso, la decisione che prenderà nella cabina elettorale di votare per l’uno o per l’altro schieramento³⁷¹.

Gli interrogativi che ne scaturirebbero sarebbero, immediatamente, due:

- a. se ritenessimo responsabile l’agente;
- b. ammesso che una tale tecnologia possa esistere, che cosa dimostrerebbe sull’esistenza del libero arbitrio.

Lo scenario potrebbe anche suggerire che spiegare il *free will* in termini neurobiologici infici, considerevolmente, la percezione della sua esistenza, certamente per molti più vicina ad una visione metafisica: il *transcendence self* in contrapposizione al *transcendent self*. In tale ottica, infatti, molti sarebbero portati a credere che, al di là delle diatribe filosofiche fra deterministi ed indeterministi, lo stesso aspetto identitario dell’uomo sarebbe messo in crisi, poiché, eliminando il libero arbitrio, si eliminerebbe, di conseguenza, anche la responsabilità morale dell’azione.

In quest’ottica, negli ultimi anni, diversi sono stati i lavori sulla psicologia di senso comune riguardante la nozione del *free will* che stanno rivelando dei dati inaspettati sul modo in cui, indipendentemente dalla “speculazione” *tout court*, venga percepito dalla *folk psychology*, che, a differenza di Sam Harris³⁷² o di Joshua Knobe³⁷³ – inclini ad una visione metafisica o del *transcendence self* in grado di garantire l’esistenza del libero arbitrio e, quindi, della responsabilità morale – si avvicina molto più ad una condivisione del *transcendent self* o “visione scientifica” di quanto si sia portati ad immaginare.

Per quanto riguarda il *free will*, il naturalismo, secondo Nahmias, offre la teoria più adeguata a capire e descrivere la mente e l’agentività, basantesi su due considerazioni principali:

³⁷¹ Cfr. ELIZABETH O’NEILL, EDOUARD MACHERY (A CURA DI), *Current Controversies in Experimental Philosophy*, Routledge, New York, 2014.

³⁷² Cfr. SAM HARRIS, *Free Will*, Free Press, New York, 2012.

³⁷³ Cfr. JOSHUA KNOBE E SHAUN NICHOLS, *Experimental Philosophy. Volume 2*, Oxford University Press, Oxford, 2014.

1. la scelta non è separata – metafisicamente – dai processi neurofisiologici che la sottendono;
2. una visione naturalistica del *free will* risulterebbe compatibile anche con la teoria compatibilistica del libero arbitrio;

poiché:

*naturalism does not commit one to a reductionistic ontological thesis that says the only things that really exist are whatever entities physics determines compose everything (such as subatomic particles or strings), not to a reductionistic epistemological thesis that says the best explanations will be those offered by lower-level science (e.g., physics or neuroscience)*³⁷⁴.

Secondo Nahmias *et alii*, al di là di quelle che possono essere le tesi personali sul concetto di libertà o responsabilità morale, uno spiraglio di un'eventuale comprensione delle variabili in gioco sul *free will* potrebbero essere fornite proprio dall'interrogazione rivolta al senso comune sul modo in cui essa affronti tematiche così complesse. Bisognerebbe, cioè, puntare la propria bussola non tanto solo ed esclusivamente sull'Accademia o su chi "costruisce" il "sapere", ma sull'altro attore protagonista del processo conoscitivo: la *folk psychology*. Ciò nella convinzione che percezioni "più genuine" possano dirimere controversie oramai "avvitate" su loro stesse.

In quest'ottica è importante notare che la gente comune – in questo senso *folk psychology* – non è impegnata nella ricerca di una visione trascendente che confligga con il naturalismo o con l'incompatibilismo, poiché il *sentire comune* – e su questo Nahmias è d'accordo con Knobe – associa alla *mente* quei particolari poteri causali che, nello scenario dualistico, si tradurrebbero in una distinzione fra "il Sé trascendente" o "pensiero" e "stati cerebrali" o "cervello".

Questa "teoria" immediata, priva di intercessioni teoretiche, Nahmias la definisce *Theory-lite* o, mantenendo il significato in italiano, "teoria leggera, poco impegnativa".

Da quest'ultima si comprende che la gente comune:

- tende a spiegare il comportamento personale e quello altrui in base ai desideri o volizioni presenti;

³⁷⁴ EDDY NAHMIAS, *A naturalistic vision of Free Will*, in ELIZABETH O'NEILL, EDOUARD MACHERY (A CURA DI), *Current Controversies in Experimental Philosophy*, Routledge, New York, 2014, p. 104.

- non tende a cambiare la *Theory-lite* sul libero arbitrio, nonostante scoperte o nuovi modi di guardare la questione possano sembrare più ragionevoli ad un'analisi approfondita;
- adotta l'equazione
desiderio → comportamento → libero arbitrio = responsabilità morale.

Sostiene, a tal proposito, Sellars:

*People may be surprised to find out that solid objects are composed of mostly empty space, but that does not change most of their beliefs about, or talk of, solid objects, much less the way they interact with solid objects. It does not lead people to believe or say that solidity is an illusion*³⁷⁵.

Nonostante ricerche successive abbiano tentato di mostrare che il senso comune tenda, maggiormente, verso una *Theory-laden* (concettualmente “ricca”) del libero arbitrio, condividendo una visione trascendente del Sé e del dualismo di sostanza mente/corpo³⁷⁶ piuttosto che una *Theory-lite*, Nahmias sostiene che ciò sarebbe lontano dal vero: gli esperimenti condotti su un campione molto diversificato per cultura, età, etnia o religione hanno evidenziato esattamente il contrario.

È abbastanza frequente che le persone, indipendentemente dalla propria tradizione teologica o convinzione dualistica, non interiorizzano i dogmi metafisici, “etichettando” i concetti di anima, mente e pensiero a “qualcosa” che possa spiegare velocemente le capacità cognitive umane³⁷⁷ perché non lasciano condizionare le loro credenze sul libero arbitrio o la responsabilità morale dai concetti metafisici di “anima” o da quelli naturalistici di stati cerebrali o processi neurobiologici³⁷⁸.

La *folk psychology*, vale a dire, non si sentirebbe “minacciata” da una visione naturalistica del libero arbitrio.

Su tali convinzioni, l'ipotesi avanzata dal *team* di Nahmias è stata quella di capire quanto la *Theory-lite* sulla mente e sull'agentività possa essere messa in discussione dall'apprensione di

³⁷⁵ WILFRIED SELLARS, *Philosophy and the scientific image of man*, in ROBERT COLODNY (A CURA DI), *Science, perception and Reality*, Humanities Press, Ridgeview, p. 43.

³⁷⁶ Cfr. JESSE BERING, *The folk psychology of soul in Behavioral and Brain Sciences*, vol. 29 (5), 2006, pp. 453-462; PETER BLOOM, *Descarte's Baby: How the Science of Child Development Explains What Makes Us Human*, Basic Books, New York, 2004.

³⁷⁷ Cfr. EDDY NAHMIA, STEPHEN MORRIS, THOMAS NADELHOFFER, JASON TURNER, *Is incompatibilism intuitive?* In “Philosophy and Phenomenological Research”, 72 (1), 2006, pp. 28-53.

³⁷⁸ Cfr. ANDREW E. MONROE E BERTRAM F. MALLE, *From uncased will to conscious choice: the need to study, not speculate about people's folk concept of free will* in *Review of Philosophy and Psychology*, 1(2), 2010, pp. 211-224; ALFRED MELE, *Another scientific threat to free will?*, “The Monist”, 95, 2012, pp. 422-440.

una visione naturalistica sul libero arbitrio, suffragata dalle neuroscienze. L'esperienza, in breve, è consistita nel sottoporre al campione considerato la spiegazione di uno scenario rigidamente deterministico per valutare, poi, le risposte fornite. Data la seguente premessa:

*Recent brain scanning studies have shown that specific patterns of brain activity can be used to predict simple decisions several seconds before people are consciously aware of those decisions. Imagine that in the future brain scanning technology becomes much more advanced. Neuroscientists can use brain scanners to detect all the activity in a person's brain and use that information to predict with 100% accuracy every single decision a person will make before the person is consciously aware of their decision. The neuroscientists cannot, however, do anything to change brain activity and hence they cannot directly influence thoughts and actions*³⁷⁹.

Si consideri il caso di Jill:

Suppose that in the future a woman named Jill agrees, as part of a neuroscience experiment, to wear this brain scanner for a month (it is a lightweight cap). The neuroscientists are able to use real-time information about her brain activity to predict everything that Jill will think or decide, even before she is aware of these thoughts or decisions. However, they cannot alter her brain activity to change what she thinks and does
380

Tornando a Jill:

On election day, Jill is considering how she will vote for President and for Governor. Before she is aware of making any decisions, the neuroscientists can see, based on her brain

³⁷⁹ Si legge: "Recenti studi di *brain scanning* hanno mostrato che specifici *pattern* cerebrali possono essere usati per predire semplici decisioni pochi secondi prima che l'agente ne sia divenuto consapevole. Immagina che in un futuro prossimo questa tecnologia si affini ulteriormente. I neuroscienziati potrebbero usare il *brain scanner* per rilevare l'attività cerebrale delle persone ed usare queste informazioni per predire con un'accuratezza pari al 100% qualsiasi decisione che prenderanno prima che gli agenti stessi ne diventino consapevoli. Essi, però, non potrebbero [in alcun modo, n.d.r.] cambiare l'attività cerebrale del soggetto onde influire sulle sue scelte immediate e future" in EDDY NAHMIA, *A naturalistic vision of Free Will*, in ELIZABETH O'NEILL, EDOUARD MACHERY (A CURA DI), *Current Controversies in Experimental Philosophy*, op. cit., p. 5.

³⁸⁰ Si legge: "Supponiamo che in un futuro prossimo una donna di nome Jill sia d'accordo con i neuroscienziati di indossare (come fosse un cappuccio) il *brain scanner* per un mese. I neuroscienziati sono in grado di usare in tempo reale le informazioni provenienti dall'attività cerebrale per predire tutto ciò che Jill penserà o deciderà anche prima che lei ne diventi consapevole. L'unico limite che i neuroscienziati hanno è l'impossibilità di alterare la sua attività cerebrale", in *IVI*.

activity, that she is about to decide to vote for Smith for President and Green for Governor. Just as the neuroscientists predicted, Jill votes for Smith for President and Green for Governor. As with her decisions to vote for Smith for President and Green for Governor, the neuroscientists are able to predict every decision Jill ends up making with 100% accuracy while she is wearing the scanner.

Occasionally, Jill tries to trick the neuroscientists by changing her mind at the last second or by stopping herself from doing something that she just decided to do, but the neuroscientists predict these events as well. Indeed, these experiments confirm that all human mental activity is entirely based on brain activity such that everything that any human thinks or does could be predicted ahead of time based on their earlier brain activity³⁸¹.

Il *test*, sottoposto ad un campione di 122 partecipanti, richiedeva agli stessi di esprimere in questo contesto delineato dalle premesse e dal caso specifico di Jill quanto fossero d'accordo con alcune asserzioni sul libero arbitrio, sulla responsabilità morale e sulla possibilità di scelta, dividendole in gruppi da tre: il primo, riguardante la scelta di Jill sul voto; il secondo, l'atteggiamento generale di Jill nel mese in cui ha indossato il *brain scanner*; il terzo, l'atteggiamento dello stesso candidato che, ipotizzandosi al posto di Jill, avrebbe dovuto spiegare quale atteggiamento avrebbe assunto laddove questa tecnologia esistesse già.

Il responso, sostiene Nahmias, non ha invalidato la predizione di partenza: i partecipanti, nonostante la descrizione di uno scenario rigidamente deterministico, non hanno considerato la rigida necessitazione causale una minaccia all'esistenza del libero arbitrio, mostrando che, in termini generali, il senso comune pensa che concetti quale il Sé o la mente siano qualcosa che trascendano l'ordine causale. In tale contesto, solo una piccola minoranza ha considerato l'attività cerebrale come causa del libero arbitrio una evidenza dell'incompatibilità dello stesso con lo scenario deterministico. Infatti, dall'analisi del *test* si evince che:

³⁸¹ Si legge: "Il giorno delle elezioni, Jill sta pensando a come voterà per il Presidente e per il Governo. Prima che lei sappia quale sarà la sua decisione a riguardo, i neuroscienziati, analizzando i dati a disposizione, sanno già quale sarà la decisione di Jill: votare per il candidato Smith alle presidenziali; votare per il candidato Green al Governo. Proprio come previsto, Jill, nella cabina elettorale, voterà per Smith e Green. Così come i neuroscienziati sono stati in grado di predire la decisione di Jill, essi sono in grado di predire ogni decisione che lei prenderà almeno fino a quando indosserà lo *scanner* cerebrale. Occasionalmente, infatti, Jill tenta di ingannare i neuroscienziati cambiando idea all'ultimo minuto o fermandosi un attimo prima di compiere l'azione, ma, nonostante i suoi tentativi, i neuroscienziati predicono questi eventi e ne sono a conoscenza ben prima di Jill stessa. Questi esperimenti, infatti, confermano che tutta l'attività mentale umana è interamente basata sull'attività cerebrale così che la prima possa essere predetta con una pari accuratezza alle decisioni di Jill, basandosi, completamente sulla seconda", in *IVI*.

91% agreed that 'Jill voted for Governor of her own free will' (with 50.8% strongly agreeing, mean = 6.09), and 90.2% agreed that 'Jill was responsible for how she voted for Governor' (mean = 6.05). Similarly, 88.5% agreed that 'During the experiment, when the neuroscientists predicted what Jill did, she had free will' (mean = 6.04), 94.3% agreed that she was responsible for what she did (mean = 6.14), 87.7% agreed that she made choices about what to do (mean = 6.00), and 86.9% agreed that what she did was 'up to her' (mean = 5.95).

Only 12.3% agreed that 'If this technology existed, then people would not have free will' (mean = 2.39), only 18.9% agreed that 'If this technology existed, then people would not be morally responsible for their actions' (mean = 2.53), only 9.0% agreed that 'If this technology existed, then people would not really make choices' (mean = 2.53), and only 11.5% agreed that 'If this technology existed, then people would not truly deserve blame for their bad actions' (mean = 2.23)³⁸².

L'interpretazione che Nahmias *et alii* hanno fornito sull'analisi dei dati è stata l'ipotesi che la maggior parte delle persone considerino che le decisioni di Jill dipendano (solo, n.d.r.) dalla sua consapevolezza di compierle per delle ragioni, dei desideri o delle volizioni. Ciò suggerirebbe, allora, che, implicitamente, la *folk psychology* resti ad un livello di spiegazione *lite* dei fenomeni mentali (come già Platone aveva evidenziato, suddividendo la *doxa* dalla *noesis!*), sia sposando qualche forma di fisicismo non riduttivo sia eludendo, dall'interpretazione dello scenario, l'attività cerebrale quale causa diretta degli stati mentali di Jill.

Ad ulteriore convalida dei dati ottenuti, Nahmias *et alii*³⁸³, nel 2014, ispirandosi allo scenario riportato da Sam Harris³⁸⁴ nel libro *Free Will*, constatarono che le persone, contrariamente a quanto creduto, non ritenevano la spiegazione del libero arbitrio in termini di correlati sinaptici una minaccia alla loro identità se si fossero spiegati cerebralmente i fenomeni cognitivi.

Lo scenario era il seguente:

Imagine a perfect neuroimaging device that would allow us to detect and interpret the subtlest changes in brain function.... the experimenters knew what you would think and do

³⁸² IDEM, p. 7.

³⁸³ Cfr. EDDY NAHMIAS, JASON SHEPARD, E JANE REUTER, *It's OK if 'My Brain Made Me Do It': People's Intuitions about Free Will and Neuroscientific Prediction*, in "Cognition", 133(2), 2014, pp. 502-513.

³⁸⁴ Cfr. SAM HARRIS, *Free Will*, op. cit.

*just before you did it. You would, of course, continue to feel free in every present moment, but the fact that someone else could report what you were about to think and do would expose this feeling for what it is: an illusion*³⁸⁵.

Prima della descrizione, si aggiungeva la specificazione di una nuova tecnologia, il *brain scanner*, in grado di tradurre in azioni le connessioni sinaptiche avvenenti nel soggetto che lo avrebbe indossato:

*Neuroscientists can use brain scanners to detect all the activity in a person's brain and use that information to detect the activity that causes decisions and predict with 100% accuracy every single decision a person will make before the person is consciously aware of their decision." It states that in the future a woman named Jill agrees to wear the scanner for a month, during which time the neuroscientists are able to predict all of her decisions, even when she is trying to trick them, and including decisions about whom to vote for in an election. And it ends with a statement meant to suggest neuro-naturalism, in one version stating, "these experiments confirm that all human mental activity is entirely based on brain activity such that everything that any human thinks or does could be predicted ahead of time based on their earlier brain activity," and in another, using the phrase, 'all human mental activity just is brain activity'*³⁸⁶.

Come nei dati ricavati dal *team* di Nahmias, anche stavolta parte dei partecipanti, fra il 75-90%, rispose che la tecnologia ipotizzata non confliggeva con l'esistenza del libero arbitrio e della responsabilità morale, dando quale giustificazione che il *brain scanner* veniva considerato per ciò che si limitava a fare, ossia la registrazione dell'attività neurale dell'agente. Per cui, ne conseguiva, che la natura o la "fonte" del *free will* sarebbe stata costituita solo ed esclusivamente dall'agente stesso. In questo modo, Nahmias (e molti come lui che condividono le sue posizioni) condivide l'ipotesi secondo cui il proprio Sé sia la fonte del proprio *free will*:

Supposing most people are theory-lite in this way such that a common response to neuronaturalism is what I am calling the natural reaction, we might then wonder whether this is the correct reaction—whether there is a way to make sense of decision-making, even free

³⁸⁵ *IDEM*, pp. 10-11.

³⁸⁶ EDDY NAHMIAS, JASON SHEPARD E JANE. REUTER, *It's Ok if 'My brain Made Me Do It': People's Intuitions about Free Will and Neuroscientific Prediction*, in "Cognition", 133(2), 2014, p. 502.

*will, in a neuro-naturalistic framework. I will now sketch a positive answer to that question*³⁸⁷.

Proprio per questo motivo, sostengono Nahmias *et alii*, la maggior parte dei partecipanti non solo non ha sentito il rigido determinismo dell'esperimento come una minaccia al libero arbitrio, ma non ha avuto bisogno neppure di spiegarsi il *free will* in termini metafisici – facendo, cioè, ricorso a particolari poteri del Sé, così come sostiene Knobe³⁸⁸.

Per dirla con le parole di Nahmias:

*These patterns of results suggest that only a minority of our participants (typically just one in five) find the possibility described in the scenario inconsistent with their understanding of free will and responsibility, and most of this minority group seem to be interpreting the scenario to involve bypassing of the agents' mental states, perhaps because this minority group is employing a transcendent conception of the self. A significant majority of participants, however, do not see the naturalistic vision described in this scenario as suggesting bypassing nor do they take its potential truth to suggest that we would lack free will or responsibility*³⁸⁹.

Alla luce di tutte le ricerche condotte fino ad oggi, il *team* di Nahmias è incline a pensare che la *folk psychology* non interpreti il determinismo come una minaccia al *free will*, perché, come evidenziato nello studio del 2012³⁹⁰,

*they did not misunderstand determinism to mean that agents' beliefs, desires, and decisions play no role in what they end up doing (and most did not make that mistake in our brain-scan scenario)*³⁹¹,

³⁸⁷ EDDY NAHMIA, *Your Brain as the Source of Free Will Worth Wanting: Understanding Free Will in the Age of Neuroscience*, in GREGG CARUSO E OWEN FLANAGAN (A CURA DI), *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*, op. cit., p. 8.

³⁸⁸ Cfr. SHAUN NICHOLAS E JOSHUA KNOBE, *Moral responsibility and determinism: The cognitive science of folk intuitions*, in "Nous", 41(4), 2007, pp. 28-53.

³⁸⁹ EDDY NAHMIA, *A naturalistic vision of Free Will*, in ELIZABETH O'NEILL, EDOUARD MACHERY (A CURA DI), *Current Controversies in Experimental Philosophy*, op. cit., pp. 7-8.

³⁹⁰ Cfr. DYLAN MURRAY, EDDY NAHMIA, *Explaining Away Incompatibilist Intuitions. Philosophy and Phenomenological Research*, Vol. 88, 2012.

³⁹¹ EDDY NAHMIA, *A naturalistic vision of Free Will*, in ELIZABETH O'NEILL, EDOUARD MACHERY (A CURA DI), *Current Controversies in Experimental Philosophy*, op. cit., p. 10.

seppure gli stati mentali siano rigidamente determinati, ciò non escluderebbe il ruolo causale dei desideri o delle volizioni al fine del compimento di un'azione. Si assiste, cioè, alla percezione di una *causal competition* fra i fattori presenti nello scenario e negli stati mentali dell'agente³⁹² che Nahmias chiama *principio della competizione causale* (CCP):

*People will be reluctant to hold an agent responsible for behavior when they interpret her behavior as being fully caused by factors that do not include any of her reasons (or by processes that do not include any of her reasoning)*³⁹³.

Il CCP potrebbe spiegare anche i motivi per cui spiegazioni riduzionistiche sulla presa di decisione siano spesso percepite come una minaccia al libero arbitrio ed alla responsabilità morale: se una decisione o un atto dell'agente sono presentate come causate *solo ed esclusivamente* dall'attività neurale, escludendo qualsiasi ruolo alle volizioni o desideri dell'agente, allora ciò porterà, inevitabilmente, alla convinzione dell'illusorietà sia del *free will* sia della *moral responsibility*.

Potrebbe esser questa la ragione, secondo Nahmias, secondo cui le persone pensano, probabilmente, che le scoperte neuroscientifiche costituiscano una minaccia al libero arbitrio; lo *slogan*

my brain made me do it

viene percepito come una privazione totale di ciò che, per molti, è considerata “l'essenza umana”: la coscienza.

7. 3 Il libero arbitrio ed il ruolo della coscienza nella visione sperimentale del filosofo statunitense Eddy Nahmias

Immagina di dover prendere un'importante decisione di lavoro: accettare l'offerta A o quella B, sapendo che la tua vita sarà significativamente differente a seconda della scelta che farai. Certamente, sostiene Nahmias, dietro ogni riflessione ci saranno motivazioni e ragioni *consapevoli* da parte dell'agente che sembrerebbero far protendere verso un paradigma della “libera scelta” in cui la responsabilità e la ragionevolezza siano requisiti fondanti per il *free will*, ma:

³⁹² Si legge: “People who *do* make the bypassing mistake are likely doing so because they perceive there to be ‘causal competition’ between the factors presented in the scenario and the agent’s mental states”, in *IVI*.

³⁹³ *IDEM*, p. 11.

*what if I told you that all of the mental processes involved in making your choice – imagining the options, evaluating them, making the decision – all of these processes happened ... in your brain?*³⁹⁴.

Ognuno di questi processi mentali è (o è realizzato) da un complesso “impianto neurale” che causalmente interagisce in accordo con le leggi della natura; un processo che Nahmias chiama “neuro-naturalism”:

*Indeed, each of those mental processes just is (or is realized in) a complex set of neural processes which causally interact in accord with the laws of nature. Call this thesis about the relation of mental processes and neural processes ‘neuro-naturalism’*³⁹⁵.

Molti, sostiene Nahmias, trovano il “n-n” palesemente assurdo o, addirittura, impossibile, poiché sposano la concezione dualista della mente e del libero arbitrio; risulta a costoro difficile accettare l’ipotesi che la “mera” attività neurobiologica possa spiegare fenomeni quali la coscienza, l’immaginazione o la presa di decisione per i motivi che più volte, in questa sede, si è avuto modo di ribadire. Una reazione “antimonistica” denominata *dualist reaction*³⁹⁶.

Molti altri, che Nahmias definisce *pessimisti*³⁹⁷, interpretano il “n-n” della mente e dell’agentività come *angst-inducing*³⁹⁸: un’espressione che, con una certa fantasia, potrebbe essere tradotta con “induzione all’angoscia”, poiché (come i *dualist reaction*) questi ultimi rifiutano qualsiasi naturalizzazione del *free will*, ma, dall’altra, accettano il “n-n” come forma di riduzionismo.

La loro “angoscia” si spiegherebbe nel fatto che seppur dovesse, un giorno, essere evidente la verità “neuro-biologica” del libero arbitrio, dall’altra, un po' come il concetto di *sublime* in Kant, libererebbe gli stessi dall’angoscia provata per la nuova presa d’atto. Ciò, allo stesso tempo, avrebbe, però, il merito di liberare i più dalla condivisione molto attraente del dualismo cartesiano.

³⁹⁴ EDDY NAHMIAS, *Your Brain as the Source of Free Will Worth Wanting: Understanding Free Will in the Age of Neuroscience*, in GREGG CARUSO E OWEN FLANAGAN (A CURA DI), *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*, Oxford University Press, Oxford, 2016, p. 2.

³⁹⁵ *IVI*.

³⁹⁶ *IVI*.

³⁹⁷ *IVI*.

³⁹⁸ *IVI*.

Come la Rivoluzione copernicana e la teoria galileiana sembravano al di fuori della “regola” per spiegare la cosmologia dell’epoca, allo stesso modo il “n-n”, sostiene Nahmias, potrebbe essere la teoria che potrebbe spiegare in maniera più plausibile la natura del libero arbitrio; sembrando ragionevole che fosse il Sole a girare intorno alla Terra (poiché, fenomenologicamente e percettivamente, tale appariva), allo stesso modo, potrà esserlo ipotizzare e prender consapevolezza che il *free will* non abbia una natura “metafisica” o “dualistica”, ma “naturale”.

Similmente, dualisti e riduzionisti tendono a pensare che il “n-n” confligga con la concezione personale del Sé, ma, suggerisce Nahmias, molti, adottando la già citata *Theory-lite* (in cui la coscienza così come la volontà o la stessa mente non trovano una spiegazione adeguata delle loro inter-relazioni o della loro stessa natura) condividono una ipotesi di lavoro “poco impegnativa”.

Ad avviso della scrivente, in tale confusione concettuale, si sta arretrando dalla ricerca della *noesis* alla mera *doxa*.

Proseguendo sulla scia del discorso principale senza divagare, in questo momento, su considerazioni e convinzioni personali che distoglierebbero dalla spiegazione del nascente approccio, i filosofi Green e Cohen, ad esempio, sostengono che le neuroscienze rivendicano un riduzionismo di fondo, dando per scontato che le persone abbiano radicate credenze, implicite o esplicite, sulla mente come una “entità” non fisica e sul libero arbitrio come un particolare risultato positivo di “speciali poteri” libertari al di fuori delle leggi causali che regolano il mondo naturale. Essi pensano che questa tesi metafisicamente opaca, non contribuisca ad una spiegazione quanto più idonea e corrispondente allo *status quo* del libero arbitrio, abbandonando, in tal modo, il dualismo e qualsiasi accezione negativa del riduzionismo *tout court*.

Le neuroscienze, d’altro canto, hanno il pregio di constatare i meccanismi sottesi alla presa di decisione, sicuramente controintuitive per la morale condivisa e le pratiche legali, ma, certamente, possono anche contribuire a modificarle, riconoscendo alle stesse neuroscienze la possibilità di illuminare la “scatola nera della mente” per comprenderla “un po’ di più”. Si potrebbe, dice Nahmias, riconoscere *the transparent bottleneck* del cervello come una *connessione* che porta una notevole quantità di informazioni, sia da ciò che si è appreso nella vita sia dal presente che si vive:

We can recognize the transparent bottleneck of the brain as the complex nexus that brings together a remarkable amount of information from both the past (including genes,

*upbringing, and learning) and the present (including external stimuli and internal beliefs, desires, goals, etc.)*³⁹⁹.

Questo nesso, spiega Nahmias, servirebbe quale sorgente dell'attività coscienziale che integrerebbe (alcune di) queste informazioni per risultare, poi, nella presa di decisione:

*This nexus then serves as the source of the causal activity that integrates (some of) this information as we make decisions*⁴⁰⁰.

Tali processi integrativi, in accordo alla *natural reaction* del quadro del “n-n”, sarà il punto critico che qualsiasi teoria epistemologica di stampo naturale sul libero arbitrio dovrà, necessariamente, affrontare, sia per l'eventuale spiegazione della natura dello stesso sia per il ruolo che la coscienza e la ragione svolgono per il *decision making*:

*These integrative processes, according to the natural reaction to the neuro-naturalist picture, will indeed be the ones most critical for responsibility, such as our consciously weighing options and reasoning about what to do*⁴⁰¹.

Nahmias, infatti, è convinto che ogni connessione neurale – insita e costitutiva del nostro cervello – sia il sito di un'unica storia causale che aiuta a spiegare il perché ognuno di noi è ciò che è. In questo senso l'espressione

my brain made me do it

dovrebbe essere interpretata, ossia come individualità irripetibile ed unica che conferisce preziosità a qualsiasi vicenda umana⁴⁰²:

*Because each nexus of neural activity—that is, each of our brains— is the site of a unique causal history, this picture also helps to explain why each of us, along with our subjective experiences, is unique*⁴⁰³.

³⁹⁹ *IVI.*

⁴⁰⁰ *IVI.*

⁴⁰¹ *IVI.*

⁴⁰² Si ricorda la scelta operata dalla sottoscritta di circoscrivere il dibattito al solo genere umano per motivi di spazio.

⁴⁰³ EDDY NAHMIA, *Your Brain as the Source of Free Will Worth Wanting: Understanding Free Will in the Age of Neuroscience*, op. cit., p. 4.

Alla luce di una visione naturalistica dei processi cognitivi, è possibile che il ruolo della coscienza nell'agente derivi dalla complessa inter-relazione delle informazioni fornite, a loro volta, dal ruolo dei geni o dell'influenza subita dall'ambiente: potrebbe essere questa una possibile spiegazione al perché, in alcuni casi, si va incontro a delle

*faulty wiring and bad decisions*⁴⁰⁴,

a dei “cablaggi difettosi e cattive decisioni”.

Il primo problema nell'accettare una posizione naturalistica come il “n-n” è, senza dubbio, sostiene Nahmias, che i neuroscienziati siano particolarmente inclini a “vedere un conflitto” fra il loro riduzionismo e la spiegazione del comportamento nella psicologia di senso comune, che non si riferirebbe al medesimo *Mechanicism* da loro supportato; adottando una visione differente quale il “n-n” sarebbero “costretti” non solo a riconoscere che le spiegazioni neuroscientifiche non siano in grado di spiegare *ogni* comportamento umano, ma anche l'inspiegabilità del fenomeno coscienziale.

La versione assunta della teoria del “n-n” non vuole eliminare la variante psicologica, ma vuole constatare che devono esserci complessi processi neurali e psicologici che realizzano o influenzano la propria decisione finale. Presumibilmente, i neuroscienziati hanno scoperto e continuano a scoprire particolari percorsi neurali che, attivandosi, producono un certo comportamento manifesto. Con una futuristica tecnologia ontogenetica, forse, saranno anche in grado di intervenire su particolari processi per ottenere determinati effetti, ma, indipendentemente dalle ipotesi, ciò che conta è che essi si limitano a constatare l'esistenza di una *neural form*. Da ciò, due importanti conseguenze.

1. In molti casi, le variabili psicologiche hanno un'influenza maggiore sulle proprie decisioni rispetto alle sole variabili neurali, anche se le prime sono rese possibili solo grazie alle seconde e, quindi, è plausibile sostenere che le ultime siano la fonte delle proprie deliberazioni:

*First, in many cases psychological variables, such as my conscious imagining of options, will have a stronger invariance relation to my decision than neural variables, even the ones that realize those psychological variables, and hence they are plausibly understood to be the causal sources of my decision*⁴⁰⁵.

⁴⁰⁴ *IVI*.

⁴⁰⁵ *IDEM*, p. 10.

2. L'eventuale adozione del "n-n" potrebbe essere una risposta alla presumibile minaccia all'esistenza del *free will* messa in campo dal determinismo.

Se il determinismo è vero, allora c'è un insieme di variabili causali passate che, tutte insieme, in accordo alle leggi della natura, sono sufficienti a far scaturire la decisione; nondimeno, il determinismo non comporta che tutte queste condizioni possano essere la fonte causale della mia decisione. Piuttosto, in molti casi, gli eventi, si potrebbe dire, si manifestano "nel collo della bottiglia"; alla fine, cioè, di un lungo e complicato processo di cui sono ignoti – al momento – i percorsi neurali:

[...] *the integrated causal activity of my conscious deliberations, as instantiated in the complex nexus of neural activity, is the causal source of my decision. Hence, even if there are causally sufficient conditions in the distant past for my decision, the causal source of my decision lies within me, not in any of the causes in my distant past*⁴⁰⁶.

Il determinismo, vale a dire, non spoglia l'agente né dalla sua deliberazione cosciente né dalle sue responsabilità, poiché la fonte delle sue decisioni sarà sempre e solo il suo cervello.

Questo argomento, in realtà, potrebbe essere usato anche per rispondere al *manipulation argument* degli incompatibilisti, secondo cui un'azione risultante da un processo deterministico (i cui fattori sono al di fuori del controllo dell'agente), presenta sempre, in linea di principio, una minaccia alla responsabilità morale dell'agente stesso. Se esso, infatti, è causalmente determinato ad agire da altri agenti (i neuroscienziati che, in un futuro lontano, siano in grado di intervenire direttamente sull'attività neurale per indirizzare le decisioni dell'agente in un verso o in un altro), allora, intuitivamente, quest'ultimo non potrà essere responsabile della sua decisione perché non derivante propriamente da sé⁴⁰⁷.

Il *focus* dei *manipulation cases* o "casi di manipolazione" è di immaginare, come si ricorderà, una qualsiasi forma di manipolazione condotta su un agente, ignaro dell'influenza che sta subendo nell'essere indirizzato verso una o un'altra decisione⁴⁰⁸.

In tali casi, allora, non ci sarebbe differenza epistemica fra l'agente determinista e l'agente incompatibilista, poiché entrambi, per cause diverse, non sarebbero responsabili delle proprie decisioni⁴⁰⁹.

⁴⁰⁶ *IBIDEM*, p. 11.

⁴⁰⁷ Cfr. DEREK PEREBOOM, *Free will, Agency and Meaning in Life*, Oxford University Press, Oxford, 2014.

⁴⁰⁸ Cfr. ROBERT KANE, *The Significance of Free Will*, Oxford University Press, Oxford (NY), 1996.

⁴⁰⁹ Cfr. DEREL PEREBOOM, *Free Will, Agency and Meaning in Life*, op. cit.

In questa ottica, Nahmias sostiene che, nella disputa fra deterministi ed indeterministi sui *manipulation cases*, gli agenti deterministici dovrebbero essere considerati responsabili a discapito dei secondi perché la fonte, la ragione, la causa delle loro decisioni non è al di fuori di loro stessi (i neuroscienziati che, nel mondo incompatibilista, indirizzano le scelte, a livello neurale, di uno o più agenti), ma in loro: il “n-n”, allora, fornisce anche un modo di disfarsi della *Theory-lite*, inadeguata alla spiegazione dei fenomeni cognitivi.

Una visione naturalistica del libero arbitrio può contribuire, così, alla comprensione di rilevanti capacità ed altrettanti limitazioni dell’uomo, spiegando perché alcuni possano essere più o meno responsabili di altri. Ciò, inoltre,

*can explain when and why someone with particular neural deficits (perhaps due to genes or upbringing) thereby lacks the cognitive and emotional capacities to evaluate relevant reasons or control their actions in such a way that it is appropriate to mitigate blame and punishment*⁴¹⁰.

Il *focus* principale, dunque, dell’argomentazione di Nahmias è che, come precedentemente sostenuto, la scienza non solo non sia in grado di sostenere l’assunzione per cui non esiste il libero arbitrio (proprio in quanto non è capace, per sua stessa natura, di comprendere il senso in cui l’uomo in quanto tale posseda il libero arbitrio), ma che, partendo dai dati del sentire comune si possa avviare un’analisi più adeguata delle variabili in gioco sui problemi *free will; free will/determinism; free will/indeterminism...* affinché si possa giungere, poi, alla presa d’atto di false premesse e di successivi falsi ragionamenti. Partendo, ad esempio, dalla *folk psychology*, Nahmias sarebbe stato in grado di procedere alla comprensione della teoria del “n-n” ed alla presa d’atto che il sentire comune è ben lontano (aspetto negativo, aggiungerebbe la sottoscritta) da qualsiasi sofisticazione scientifica nei ragionamenti condotti a proposito dei temi in oggetto. Ciò comporta la consapevolezza che qualsiasi eventuale processo neuronale si dovesse scoprire quale causa del proprio volere salvaguarderà il libero arbitrio dal punto di vista della *folk psychology*.

Del medesimo parere si trova anche Adina Roskies⁴¹¹, la quale conviene che, in effetti, ultimamente le neuroscienze stanno incrementando la preoccupazione che la comprensione del funzionamento cerebrale è una minaccia sia al libero arbitrio sia alla responsabilità

⁴¹⁰ EDDY NAHMIA, *Your Brain as the Source of Free Will Worth Wanting: Understanding Free Will in the Age of Neuroscience*, in GREGG CARUSO E OWEN FLANAGAN (A CURA DI), *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*, op. cit., p. 13.

⁴¹¹ Cfr. ADINA ROSKIES, *Neuroscientific challenge to free will and responsibility*, in “Trends Cogn Sci.”, 10(9): 2006, pp. 419-423.

morale. Per la filosofa statunitense, invece, una buona ragione per sostenere che tali preoccupazioni siano fuori luogo derivano dal fatto che la nozione comune sul *free will* è sempre esistita, indipendentemente dai reali/dai presunti tali studi neuroscientifici.

7.4 Una nuova fenomenologia per il libero arbitrio?

Da quanto si è avuto modo di apprendere dall'approccio fin qui condotto dal *team* di Nahmias, non dovrebbe essere una novità quella secondo cui quest'ultimo stia cercando una "terza via" all'*impasse* determinismo/indeterminismo per spiegare la natura del libero arbitrio. Probabilmente, una certa sorpresa la si potrebbe riscontrare nel fatto che il primo passo da compiere in quella direzione sarebbe avviare uno studio che si concentri, come visto, anche sul modo in cui il *sentire comune* percepisca o spieghi il *free will*. Da ciò, secondo Nahmias, si sarà, poi, in grado di capire, effettivamente, la veridicità o della tesi deterministica o di quella indeterministica o di nessuna delle due.

Secondo il suo *team*, la fenomenologia finora condotta sul libero arbitrio è errata perché si baserebbe sulle sole personali introspezioni dell'interrogante (condotte in prima persona) e ben lontane dalla fenomenologia "oggettiva" dell'evento studiato; i filosofi, vale a dire, avrebbero sbagliato laddove avrebbero presupposto che la propria personale introspezione "dell'evento libero arbitrio" potesse essere universalmente riconosciuto come loro stessi lo avessero percepito, provato, sentito.

Un discorso "a metà", verrebbe da dire, da dover incrementare con la parte mancante: la psicologia sperimentale. Una riconciliazione tra Filosofia e Psicologia che potrebbe essere in grado di aiutare il dibattito grazie all'ausilio dei dati empirici, fenomenici e non solamente speculativi; solo in questo modo, secondo Nahmias, si potrà giungere ad un'analisi fenomenologica attendibile e corretta del problema trattato.

Un primo dato derivante dal nuovo approccio sarebbe quello di comprendere che, contrariamente a quanto finora si è pensato, la tesi sul libero arbitrio maggiormente condivisa dal senso comune non sarebbe quella indeterministica/libertaria, ma deterministica/compatibilistica.

In genere, le teorie sul *free will* sono maggiormente plausibili quando riescono a "catturare" adeguatamente le intuizioni e le esperienze sul senso di agentività che l'uomo sente nell'atto di scegliere. In merito a ciò, il tradizionale dibattito filosofico sul tema si è suddiviso, come ormai si saprà, in *compatibilisti* e *libertari*: entrambi d'accordo nel ritenere che l'uomo

possieda libertà di scelta, ma in disaccordo per la fenomenologia e per il modo in cui il libero arbitrio si espliciterebbe nell'agente e, contestualmente, nel mondo. Se da una parte, allora, entrambi gli schieramenti concordano sul fatto che la medesima fenomenologia possa essere ricavata tramite ricerche e descrizioni introspettive, dall'altra si discostano proprio su tale punto; d'altronde, se esistono diverse teorie sul libero arbitrio, evidentemente proprio la sua introspezione "oggettiva" è problematica.

Non è da escludere, inoltre, che l'introspezione in prima persona compiuta dal filosofo è oggetto della particolare interpretazione con cui lo stesso "interroga" il concetto medesimo; ciò si evince anche dalle due importanti distinzioni fenomenologiche compiute dai compatibilisti e dai libertari, secondo cui il libero arbitrio o discenderebbe direttamente dal nesso causale che governa il mondo oppure dalla medesima volontà dell'agente in un reale del tutto contingente⁴¹².

Il dato che, immediatamente, si evincerebbe dal nuovo approccio psico-filosofico sarebbe quello di notare che l'analisi fenomenologica condotta dai libertari è ben lontana dalle assunzioni della gente comune:

*More importantly, libertarian phenomenological descriptions set more demanding veridicality conditions, since their theories require more demanding metaphysical conditions than compatibilisti theories do. This raises the questions: why should we believe we need libertarian free will unless we feel like we have it? Why shouldn't we be satisfied with the less demanding compatibilisti conception of free will it is in fact consistent with our experiences?*⁴¹³

Nahmias *et alii* tendono a precisare che il ricorso alla fenomenologia della *folk psychology* non sia quello di *risolvere*, tramite l'ausilio della ricerca psicologica, un dibattito tanto controverso quanto millenario, ma quello di incentivare uno studio sul libero arbitrio che non tralasci lo stesso campo del sentire comune: proprio da quest'ultimo si potrebbero ricavare dei dati il più vicino possibile all'"oggettività", poiché sarebbe oggetto di analisi non soltanto da

⁴¹² Cfr. EDDY NAHMIA, *Your Brain as the Source of Free Will Worth Wanting: Understanding Free Will in the Age of Neuroscience*, in GREGG CARUSO E OWEN FLANAGAN (A CURA DI), *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*, op. cit.

⁴¹³ Si legge: "Di notevole importanza è che le descrizioni fenomenologiche dei libertari richiedono condizioni di verità esigenti più di quanto siano richieste dai compatibilisti. Ciò pone una domanda: perché dovremmo credere di aver bisogno del libero arbitrio libertario quando sentiamo di possederlo in termini diversi dal modo in cui gli stessi lo descrivono? Perché non dovremmo sentirci soddisfatti della definizione del libero arbitrio descritto dai compatibilisti, metafisicamente più semplice e più vicino alla nostra esperienza del libero volere", in *IDEM*, p. 163.

una prospettiva in “prima persona” (il filosofo), ma anche di quella in “terza persona” (il senso comune). Solo dopo l’analisi e la raccolta di questi dati si potrà impiantare un dibattito proficuo sul *free will*, seguendo il filone di ricerca del nuovo approccio che si baserà sulle *folk intuitions* (intuizioni comuni⁴¹⁴), i cui scopi saranno quelli di attirare l’attenzione sull’importanza del sentire comune su temi quali il libero arbitrio e la responsabilità morale per valutare quanto il compatibilismo e la teoria libertaria siano comunemente condivise e per quale motivo⁴¹⁵. Solo in tal modo si potrà offrire, da una parte, un’adeguata descrizione fenomenologica dell’oggetto di studio; dall’altra, la risoluzione alle discrepanze tra le affermazioni teoretiche dei filosofi e le loro stesse asserzioni fenomenologiche sul libero arbitrio.

Ritornando alle differenze fenomenologiche che dividono i compatibilisti dai libertari, il filosofo statunitense evidenzia come tali posizioni siano in disaccordo, dapprima, sull’esperienza, successivamente, sul concetto inerente al libero arbitrio, individuando due punti di contrasto:

1. l’interpretazione della nozione del “dover fare altrimenti”;
2. se le libere azioni siano causate da speciali poteri dell’agente derivanti dai suoi stati mentali (forza di volontà, libero arbitrio...) oppure se esse siano la conseguenza di cause antecedenti alla medesima esistenza dell’agente.

Per quanto riguarda il primo punto, i libertari sostengono che un’azione è libera solo se

[...] *given all conditions as they are at and up to until the moment of choice*⁴¹⁶,

mentre i compatibilisti sostengono che la capacità di fare altrimenti può essere intesa (come si ricorderà) sia come “capacità condizionale” (ossia la possibilità di fare altrimenti di come *de facto* il nesso necessario porta l’agente ad agire) sia come la capacità generale di rispondere in modo adeguato alle circostanze che accadono necessariamente.

Stando all’interpretazione fenomenologica del libero arbitrio data dai libertari, sembra – sostengono Nahmias *et alii* – che se è vero che l’agente fino all’attimo in cui sceglie di compiere o di non compiere l’azione sia in grado di cambiare idea (poiché frutto delle proprie credenze, desideri e simili) allora ciò significherebbe, implicitamente, ammettere che l’agente

⁴¹⁴ *IBIDEM*, p. 164.

⁴¹⁵ Cfr. EDDY NAHMIA, JUSTIN D. COATES E TREVOR KVARAN, *Free Will, Moral Responsibility and Mechanism: Experiments on Folk Intuitions*, in “Midwest Studies in Philosophy”, 31 (1), 2007, pp. 214-242.

⁴¹⁶ Si legge: “[...] date tutte le condizioni in cui si trovano (ad agire), possono scegliere (di fare altrimenti) fino allo stesso momento della scelta”, in EDDY NAHMIA, STEPHEN MORRIS E THOMAS NADELHOFFER, *The Phenomenology of Free Will* in “Journal of Consciousness Studies”, Vol. 11, n. 7’8, 2004, p. 165.

subirebbe necessariamente (causalmente) le cause pregresse (educazione ricevuta, esperienze avute...), indipendentemente dalle ragioni che pertengono quella data scelta in quel dato momento. Per tale motivo i libertari si rifanno a teorie di stampo metafisico, che poco hanno a che fare con l'approccio prettamente fenomenologico al libero arbitrio e, di conseguenza, secondo Nahmias, la confutazione della tesi compatibilista sulla nozione "poter fare altrimenti" spetterebbe a loro. I libertari, inoltre, aggiungono alla spiegazione della possibilità di scelta la nozione confusa di *causal oomph*⁴¹⁷ che, se letteralmente significa una "grinta causale", *de facto* significa che i sentimenti che l'agente percepisce nell'atto della scelta inclinerebbero verso una decisione piuttosto che in un'altra.

Ciò porta all'annosa domanda se l'agente sia realmente la causa del suo agire ed in che senso causi il suo stesso agire: una risposta che, come si ricorderà, viene fornita dall'argomento dell'*agent causation*⁴¹⁸, secondo cui l'azione dell'agente derivi direttamente da se stesso, essendo fortemente attivo nella presa di decisione cosciente che lo investe nel momento della scelta; per cui l'azione, per rispondere alla domanda, deriva direttamente da lui stesso. Il ricorso al primo motore immobile aristotelico è emblematico: ogni azione dell'agente corrisponde all'origine di infiniti altri effetti, la cui causa risiede solo ed esclusivamente nello stesso.

Al contrario, i compatibilisti sostengono che l'azione è sempre il risultato di una serie pregressa e infinita di cause necessarie che portano il soggetto a compiere ciò che di fatto compirà; in tal modo, qualsiasi spiegazione fenomenologica della presa di decisione cosciente sarà giustificata dal nesso causale che regola la sua stessa volontà ed il mondo. L'introduzione di speciali poteri dell'agente che siano in grado di dar conto di un'azione derivante direttamente da se stessi sarebbe il risultato di una visione fortemente metafisica del libero arbitrio, che si allontanerebbe *in toto* dalla realtà pratica. Proprio per tale motivo Nahmias *et alii* sostengono che l'onere della prova – il mostrare, cioè, che la fenomenologia del libero arbitrio avviene secondo le intenzioni o le volizioni dell'agente secondo il *source incompatibilism* – spetterebbe proprio ai libertari.

Il problema, come si sarà capito, resta la volontarietà dell'atto.

Come precedentemente sostenuto, l'esperimento di Libet ha sollevato numerose questioni etiche sull'effettiva esistenza o inesistenza del *free will*. Innanzitutto, ciò che emerse fu, da un lato, l'intento riduzionista di spiegare il libero arbitrio su basi neurali; dall'altro, la controversa diatriba sulla legittimità della scenografia dell'esperimento.

⁴¹⁷ *IDEM*, p. 166.

⁴¹⁸ Cfr. il capitolo V.

Se in Libet la paura di “ingabbiare” l’uomo all’interno di uno scenario riduzionista fece sì che quest’ultimo tentasse di salvaguardare la libertà di scelta da parte dell’agente, appoggiando una visione dualistica sul rapporto mente/cervello ed introducendo il concetto di “volontà di veto”, anche oggi l’idea di indagare neuroscientificamente la volontà, l’intenzione ed il libero arbitrio ha fatto nascere in molti la convinzione che tali concetti siano incompatibili con le medesime nozioni che le sorreggono. Se, cioè, da una parte sembrerebbe lecito naturalizzare l’uomo, dall’altra sembrerebbe impossibile ridurre la volontà cosciente ad una serie di sinapsi ben articolate; questa distonia nascerebbe dal fatto che molti filosofi sono convinti che l’esperienza in prima persona che l’agente compie, essendo consapevole delle proprie intenzioni, è assolutamente un dato non indagabile scientificamente. Si è così alimentata, nello scenario collettivo, la visione della volontà cosciente come quel principio, forse metafisico, che lega causalmente la mente alle azioni, tale da far risultare la medesima cosa

l’esperienza di voler compiere [consapevolmente, n.d.r.] un’azione con la causazione dell’azione da parte della mente [cosciente, n.d.r.]⁴¹⁹.

Riprendendo le considerazioni di Wegner, si è visto come la maggior parte degli esseri umani comprende le altre entità come viventi, poiché le riconosce dirette verso uno scopo, essendo in grado di pensare ai possibili obiettivi che vogliono raggiungere. Applicando tale capacità anche a se stessi, l’uomo risulta in grado di pensarsi come un agente causale, essendo in grado di comprendere sé e gli altri tramite credenze, desideri e progetti che contribuiscono alla convinzione di possedere una volontà cosciente.

Tale sensazione, a sua volta, determina l’esperienza in prima persona di possedere una volontà cosciente, proprio in quanto si è in grado di immaginare le azioni da compiere prim’ancora che si compiano. Ciò contribuisce ad alimentare la causazione mentale sul mondo fisico, mascherando, però, il suo carattere prettamente illusorio: proprio come il gioco di prestigio risulta all’occhio dello spettatore il risultato misterioso di particolari capacità del mago – ignorando completamente i meccanismi fisici capaci di creare l’illusione – nel medesimo modo la sequenza causale reale di stampo neurale che soggiace al comportamento umano coinvolge un insieme enormemente complicato di meccanismi. Si sviluppa, in tal modo, l’idea di una volontà cosciente che precede le azioni fisiche del corpo di cui, oltretutto, il soggetto agente ne fa esperienza.

⁴¹⁹ DANIEL WEGNER, *L’illusione della volontà cosciente*, in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI, (A CURA DI), *Siamo davvero liberi?*, op. cit., p. 97.

Se, allora, la volontarietà dell'agente sembra essere solamente un gioco di prestigio che la mente farebbe a se stessa, il problema di come definirsi moralmente responsabili resta insoluto e l'unica alternativa possibile sarebbe quella di separare la libertà dal concetto stesso di responsabilità morale. In che modo ciò sarebbe possibile?

Al di là delle discrepanze esistenti fra compatibilisti e libertari, il *focus* centrale dei diversi argomenti riportati è sempre quello secondo cui la ricerca psicologica (nonostante sia ben lontana dall'esser risolutiva) sul libero arbitrio potrebbe aiutare notevolmente il dibattito ancora in corso, perché fornirebbe al filosofo, come detto, un dato "oggettivo" sull'esistenza o meno del *free will*. Tale apporto verrebbe dal sentire comune, dalle *folk intuitions*, e proprio da tale assunto ha preso piede il lavoro dello psicologo Roy F. Baumeister⁴²⁰, il quale intende coniugare i dati empirici derivanti dalle neuroscienze con le intuizioni comuni sul libero arbitrio per avviare uno studio maggiormente adeguato di quest'ultimo – sulla scia di Nahmias ed il suo *team*.

Da un punto di vista psicologico, il *free will* può essere inteso quale strategia adattativa umana all'ambiente e come una forma evoluta del controllo dell'azione che incorpora prescrizioni culturali, tradotte poi in comportamento:

*Free will, in a psychologically useful and adaptive sense, is therefore an advanced form of action control that can incorporate cultural prescriptions and execute them behaviourally*⁴²¹.

Da ciò, emergerebbe, secondo Baumeister, un fatto di notevole importanza: non solo molte persone credono nel *free will* in termini metafisici, assumendo dei comportamenti in accordo alle loro credenze, ma è altresì semplice manipolare le loro credenze. Il lavoro di Vohs e Schooler condotto nel 2008, ad esempio, ha evidenziato che alcuni studenti erano inclini a cambiare completamente credenza sull'esistenza del libero arbitrio a seconda che si mostrassero loro prove o argomentazioni a favore della sua illusorietà o meno. L'aspetto più importante delle ricerche condotte da Vohs e Schooler è stato che – come constatato anche da Knobe e Nicholas – anche il comportamento sociale degli studenti sottoposti all'esperimento mutò profondamente: se essi erano indotti a credere nel determinismo, ad esempio,

⁴²⁰ Cfr. ROY F. BAUMEISTER, *Understanding Free Will and Consciousness on the Basis of Current Research Findings in Psychology*, in ROY F. BAUMEISTER, ALFRED R. MELE E KATHLEEN D. VOHS (A CURA DI), *Free will and consciousness: How might they work?*, Oxford University Press, New York, 2010.

⁴²¹ Scrive Baumeister: "Il libero arbitrio, in un senso psicologicamente utile e adattativo, è perciò una forma avanzata del controllo dell'azione che incorpora prescrizioni culturali per poi tradurle in *comportamento*, *IDEM*, p. 26.

aumentavano notevolmente gli atteggiamenti antisociali nella convinzione che, se il loro agire era necessitato, essi non potevano dirsi responsabili delle azioni compiute.

Tali studi hanno ulteriormente consolidato in Baumeister l'idea che la credenza nel *free will* non costituisce un accidente, ma una strategia adattativa evolutasi nell'arco della storia umana; di conseguenza, indipendentemente dal fatto che le neuroscienze possano mostrarne la sua effettiva esistenza, l'uomo sarà sempre portato a crederci.

Proprio per la natura evolutiva del libero arbitrio, Baumeister sostiene che quando comunemente si parla di libero volere, innanzitutto, si fa riferimento all'insieme di diversi modelli comportamentali che implicano diverse categorie di atti come quelli che persistono nonostante pressioni esterne, quelli che perseguono uno scopo a lungo termine, così come quelli che lasciano trasparire una riflessione consapevole. Tutti questi atti sono tradizionalmente considerati liberi e ciò implica che istintivamente, naturalmente le persone classifichino e ritengano alcuni atti più liberi di altri.

Come sostiene Baumeister:

*Variation in freedom of action is centrally important to my understanding*⁴²²,

poiché

*A fourth empirical fact is that the inner, psychological processes that produce the actions regarded as free differ from those that produce the relatively unfree actions*⁴²³.

La convinzione secondo cui il libero arbitrio sarebbe “niente più” che una strategia adattiva, Baumeister la denomina come *ego depletion*, un “impoverimento dell'Io”, per riferirsi, anche, agli stati comportamentali che risultano dall'esercizio dell'autocontrollo e dalla presa di decisione cosciente⁴²⁴.

Attualmente, non pochi filosofi stanno tentando di inquadrare il *free will* nei “nuovi” aspetti fenomenologici ed evolutivi del problema e, a tal proposito, D. A. Pizarro e Erik G. Helzer sostengono che le persone sono fondamentalmente motivate a valutare le azioni degli altri come morali, ritenendoli naturalmente responsabili dei loro atti. È come se l'uomo sociale fosse di per sé uno *stubborn moralists*, un “ostinato moralista”, che guidi anche la propria

⁴²² Scrive Baumeister: “La variante nella libertà dell'azione è di fondamentale importanza per la mia comprensione”, *IBIDEM*, p. 28.

⁴²³ Si legge: “I processi psicologici interiori, responsabili delle azioni considerate libere, differiscono da quelli che generano le azioni non libere”, *IDEM*, p. 29.

⁴²⁴ *IBIDEM*, p. 30.

moralità⁴²⁵. Mentre potrebbe sembrare che tutte le credenze morali varino nel tempo e nello spazio, in realtà alcune credenze basilari sull'assunto che alcuni atti siano moralmente vietati e che le persone ne siano responsabili, appaiono un tratto comune dell'essere umano. Le tesi di Baumeister, così, non sembrerebbero molto lontano "dal vero".

La tesi di Pizarro ed Helzer, come si è accennato, ricalca quella di Nahmias, secondo cui discostarsi *in toto* dalle assunzioni psicologiche/concettuali della gente ordinaria su temi quali libero arbitrio e responsabilità morale significherebbe allontanarsi, completamente, da uno studio adeguato dei problemi analizzati: l'analisi delle intuizioni comuni, in altre parole, potrebbe aiutare il filosofo a capire meglio i processi cognitivi della mente umana ed il modo migliore per giungere alla conoscenza di tali processi sarebbe proprio la ricerca psicologica.

A tal proposito Nahmias *et alii* scrivono:

*Indeed, by mapping folk intuitions about FW (free will) and MR (moral responsibility), we can better understand the cognitive processes that generate this intuition. And this will put philosophers in a better position to provide plausible explanations of why certain views are attractive even if they are mistaken – and to understand how mistaken intuitions might be corrected*⁴²⁶.

Seppur Nahmias, Coates e Kvaran appoggino in parte le tesi di Nicholas e Knobe⁴²⁷, è pur vero che a loro avviso l'interpretazione dei dati appresi dallo studio dei due filosofi statunitensi necessiterebbe di una diversa interpretazione: l'influenza esercitata sull'attribuzione o meno di responsabilità negli esempi sottoposti agli studenti non dovrebbe esser trattata come un mero "errore interpretativo", ma, piuttosto, come un giudizio

⁴²⁵ Cfr. DAVID A. PIZARRO, ERIK G. HELZER, *Stubborn Moralism and Freedom of the Will*, in in ROY F. BAUMEISTER, ALFRED R. MELE E KATHLEEN D. VOHS (A CURA DI), *Free will and consciousness: How might they work?*, op. cit.

⁴²⁶ Scrivono gli Autori: "Infatti, dalla mappatura delle intuizioni comuni sul libero arbitrio e sulla responsabilità morale, possiamo comprendere meglio i processi cognitivi che generano queste intuizioni. Ciò porterà i filosofi in una condizione migliore per fornire spiegazioni plausibili al perché alcune tesi siano più attrattive di altre nonostante siano errate – e a capire come le intuizioni erronee possano essere corrette", cit. in EDDY NAHMIA, JUSTIN D. COATES E TREVOR KVARAN, *Free will, moral responsibility and Mechanism: Experiments on folk intuitions*, in "Midwest Studies in Philosophy", 31(1), 2007, p. 217.

⁴²⁷ Nicholas e Knobe hanno suggerito che l'imbarazzo teorico che le persone provano di fronte al problema del libero arbitrio è, forse, dovuto all'esistenza di due diversi processi di ragionamento che spingono l'analisi in due direzioni distinte. Da una parte, risiederebbe l'aspetto emotivo; dall'altro, quello razionale. Ciò comporterebbe che, a seconda del modo in cui si presenta ad un individuo un problema morale sul libero arbitrio e sulla responsabilità, la sua risposta sull'esistenza o meno della libera volontà dipenderebbe sia dal grado di cultura personale sia dalle informazioni implicite fornite dal modo in cui si pone la domanda. A seconda, cioè, che si faccia prestare attenzione sul fatto che ci si trovi in un mondo deterministico o indeterministico o che non si faccia riferimento a nessun caso specifico – ma solo all'elaborazione in astratto di un problema – le persone rispondono in maniera diversa, a seconda che enfatizzino con il soggetto della storia oppure no.

prettamente razionale sul tema trattato⁴²⁸. Contrariamente a quanto suggerito dai due filosofi statunitensi – secondo cui le emozioni non limiterebbero la formulazione di un giudizio razionale sulla responsabilità morale – Nahmias è d'accordo nel ritenere che proprio le emozioni dovrebbero essere considerate “fattori abilitanti” (*enabling factors*⁴²⁹) che supportano ed incoraggiano i processi cognitivi/razionali e che proprio l'interpretazione erronea dei dati ricavati da Nicholas e Knobe – lontani dall'approccio delle *folk intuitions* – avrebbe contribuito ad una visione erronea della natura del libero arbitrio.

I due filosofi, infatti, dai dati in loro possesso, hanno ricavato che, intuitivamente, il sentire comune sul *free will* si avvicina alla tesi incompatibilista e non a quella compatibilista/deterministica: un risultato che, secondo Nahmias, è ben lontano dall'essere vero.

Proprio per lo stallo fra le due tesi in campo per la comprensione del libero arbitrio, come precedentemente si ricordava, Nahmias e Murray sostengono uno studio approfondito delle *folk intuitions*⁴³⁰ per mostrare che, contrariamente a quanto gli stessi incompatibilisti sostengono, il sentire comune – fenomenologico – sul libero arbitrio non è una genuina intuizione incompatibilista, ma una cattiva interpretazione di quella compatibilista.

Alla base di questo fraintendimento concettuale risiederebbe una presentazione fuorviata della stessa teoria compatibilista; secondo cui le credenze, i desideri o le decisioni dell'agente non fanno alcuna differenza in ciò che *necessariamente deve accadere*:

*Many ordinary persons come to interpret determinism to threats to free will that it does not in fact entail. We predict that laypersons often mistakenly take determinism to mean that everything that happens is inevitable – it will happen no matter what – or that agent's decisions, desires, or beliefs make no difference to what they end up doing and that such mistakes then generate people's intuitions about agent's lacking free will and moral responsibility*⁴³¹.

⁴²⁸ IDEM.

⁴²⁹ IBIDEM, p. 235.

⁴³⁰ Cfr. NAHMIA EDDY E DYLAN MURRAY, *Experimental Philosophy on Free Will: An Error Theory for Incompatibilist Intuition*, in JESUS AGUILAR, ANDREI BUKAREFF E KEITH FRANKISH (A CURA DI), *New Waves in Philosophy of Action*, Palgrave-Macmillan, 2010.

⁴³¹ Scrivono gli autori: “La gente comune interpreta il determinismo come una minaccia per il libero arbitrio. Noi sosteniamo che spesso il sentire comune erroneamente intende tale tesi come sinonimo che ciò che accade sia inevitabile, indipendentemente dalle decisioni, i desideri, le credenze dell'agente e che tale erronea interpretazione genera nel senso comune la perdita del libero arbitrio e della responsabilità morale”, in EDDY NAHMIA E DYLAN MURRAY, *Experimental Philosophy on Free Will: An Error Theory for Incompatibilist Intuitions*, in JESUS AGUILAR, ANDREI BUCKAREFF E KEITH FRANKISH (A CURA DI), *New Waves in Philosophy of Action*, op. cit., p. 191.

La tesi sostenuta da Nahmias e Murray è che molti di coloro che credono di avere un'intuizione incompatibilista sul libero arbitrio, in realtà concordano con una tesi erronea del determinismo, definita del *bypassing*⁴³², ossia l'idea base secondo cui le proprie azioni sono causate da una intuizione che *travalica, sorpassa*, la propria consapevolezza. In particolare:

*It is the thesis that one's actions are produced in a way that bypasses the abilities compatibilists typically identify with free will, such as rational deliberation, conscious consideration of beliefs and desires*⁴³³.

Il *Bypassing* suggerisce che gli agenti coscienti non abbiano alcun controllo sulle loro azioni perché essi stessi non giocano un ruolo fondante sulle scelte causali che compiono. Un tentativo iniziale di indagare gli argomenti del *bypassing* è stato intrapreso nel lavoro di Nahmias, Murray, Coates e Kvaran – precedentemente citato – in cui hanno riscontrato che la maggior parte delle persone sosteneva l'esistenza del libero arbitrio e della responsabilità morale in un mondo deterministico se le azioni degli agenti fossero completamente causate dai loro stessi desideri, credenze e pensieri:

[...] *if the scenery describe agent's decisions as completely caused by the specific thoughts, desires and plans occurring in our minds*⁴³⁴.

Laddove, invece, lo scenario dell'esempio suggeriva una visione riduzionistica del mentale, in cui, cioè, si spiegavano le azioni in base a reazioni chimiche e processi neurali, gli agenti sembravano unanimemente d'accordo nel suggerire che in un mondo determinista sia il libero arbitrio sia la responsabilità morale siano impossibili.

Un errore di *bypassing*, riprendendo quanto più sopra sostenuto, è stato commesso anche da Nicholas e Knobe, i quali, con una serie di espressioni ambigue e di omissioni durante l'esecuzione del loro esperimento, hanno contribuito a far passare l'idea che, effettivamente, il determinismo significhi assenza di libero arbitrio.

In particolare, Nahmias e Murray si riferiscono, oltre alla scelta di Nicholas e Knobe di non impiegare il termine *free will* nei questionari sottoposti – in quanto il termine specifico avrebbe potuto, a loro parere, falsare il giudizio di chi si fosse prestato alla compilazione del

⁴³² *IBIDEM*.

⁴³³ Scrivono gli autori: "È la tesi secondo cui un'azione è prodotta in un modo che bypassa le abilità compatibiliste che comunemente identificano il libero arbitrio, come una deliberazione razionale e cosciente delle proprie credenze e dei propri desideri", *IDEM*, p. 192.

⁴³⁴ *IDEM*, p. 198.

questionario –, anche a quelle di utilizzare espressioni come *fully morally responsible* per raccogliere, “genuinamente”, le percezioni e le intenzioni degli studenti stessi. Secondo Nahmias e Murray, infatti, le “etichette” adoperate richiederebbero ai soggetti coinvolti un’accurata o, quantomeno, approfondita riflessione sul rapporto libero arbitrio/responsabilità morale, che esulerebbe *in toto* dai giudizi intuitivi sul *free will*. Scopo, poi, dello stesso esperimento di Nicholas e Knobe.

Altro errore si potrebbe individuare nelle spiegazioni che gli stessi hanno utilizzato per definire l’universo causale in cui si lascia trapelare una concezione errata dello stesso determinismo; frasi del tipo:

*In Universe A, given the past, each decision has to happen the way that it does*⁴³⁵

hanno lasciato presagire una forma di fatalismo quale *ratio* del contesto e non una necessitazione causale tipica del determinismo, suggerendo o che gli agenti non abbiano nessun controllo su ciò che fanno, oppure che le loro azioni, indipendentemente da ciò che sia potuto accadere nel passato, saranno sempre le medesime nel futuro prestabilito.

Dall’analisi dei casi di *bypassing*, Nahmias e Murray concludono che, se i dati da loro raccolti sono corretti, ciò mostrerebbe il perché le *folk intuitions* mostrerebbero l’incompatibilismo quale tesi che maggiormente si avvicina a ciò che la gente comune intende come libero arbitrio, ma che, una volta sgombrato il campo da un’interpretazione erronea, la genuinità dell’intuizione compatibilista si mostrerebbe senza alcun problema di *bypassing*.

E l’ausilio delle neuroscienze?

Secondo Nahmias, intuitivamente si è portati a pensare che la spiegazione in termini fisici/biologici/neuroscientifici del comportamento umano significhi comprendere l’uomo solo in termini di reazioni chimiche, privandolo di ciò che comunemente si associa alle sue capacità cognitive: libero arbitrio e responsabilità morale. Invece, contrariamente a quanto si sia portati a credere, la *folk psychology* sostiene una visione del libero arbitrio che ben si presta alla felice convivenza di una spiegazione riduzionistica con l’esistenza dello stesso⁴³⁶.

Gli esperimenti condotti hanno dimostrato che nella maggior parte dei casi quando la gente si interroga sull’esistenza del libero arbitrio in relazione al compatibilismo o all’incompatibilismo, seppur *prima facie* sembra sostenere il puro incompatibilismo, in realtà

⁴³⁵ Si legge: "In un Universo A, dato il passato, qualsiasi decisione deve avvenire nel medesimo modo", *IBIDEM*, p. 196.

⁴³⁶ EDDY NAHMIA, JUSTIN D. COATES E TREVOR KVARAN, *Free will, moral responsibility and Mechanism: Experiments on folk intuitions*, in “Midwest Studies in Philosophy”, 31 (1), 2007, p. 215.

la tesi che sposa è quella di una compatibilità fra *Mechanism Incompatibilism* (MI), libero arbitrio e responsabilità morale.

A differenza del “puro incompatibilismo”, MI sostiene, ad esempio, che se anche le interazioni nel nostro universo possono essere riconducibili ad eventi quantici (quindi a livello submolecolare), ciò non significa che il mondo sia regolato da un principio causale necessario, così come sostiene il determinismo. Se anche, vale a dire, nel primo contesto si presentano leggi causali a livello quantico, ciò non comporta l'estensione del principio causale ad ogni azione in quanto MI e determinismo sono due tesi diverse. Ciò conduce Nahmias, ancora una volta, a sostenere che il ricorso alla ricerca psicologica è un passo necessario e primario per la delucidazione dei concetti da analizzare, sia perché il filosofo sarà così in grado di ripartire da un dato empirico per le sue dissertazioni sia perché solo dalla interazione costante la gente sarà in grado di correggere le proprie assunzioni erronee.

Per Nahmias, quindi, il libero arbitrio è un'abilità, una capacità psicologica che si dà all'agente a seconda delle effettive possibilità che ha di poterlo esercitare; *free will* e *free action* sono *gradi di nozioni* che si esercitano grazie alla *opportunità di esercitare* tali capacità nel prendere decisioni e nell'agire:

For an agent to have free will is for her to possess the psychological capacities to make decisions - to imagine alternatives for action, to select among them, and to control her actions accordingly - such that she is the author of her actions and can deserve credit or blame for them. For an agent to choose freely (or act of her own free will) is for her to have had (reasonable) opportunity to exercise these capacities in making her decision and acting [...] Free will and free action are graded notions: we have degrees of freedom, and accordingly, degrees of responsibility⁴³⁷.

Così come qualsiasi altra abilità cognitiva o manuale posseduta, allo stesso modo il libero arbitrio è il risultato di componenti *normative* e *strutturali*: delle prime, data la complessità di funzionamento del cervello, non si è al momento in grado di catalogare o individuare “leggi universali” che regolino l'esplicazione del *free will*, ma, sicuramente, così come esistono forme migliori o peggiori di svolgere un compito, allo stesso modo – continua Nahmias – esistono modi migliori e peggiori di prendere una decisione. Delle seconde, al contrario, le

⁴³⁷ EDDY NAHMIA, *Free Will as a Psychological Accomplishment*, in DAVID SCHMIDTZ E CARMEN PAVEL (A CURA DI), *Oxford Handbook on Freedom*, Oxford University Press, Oxford, 2016.
Articolo reperibile on-line p. 2 del link

<https://philarchive.org/archive/NAHFWA-4v1>.

capacità sono strutturalmente formate da ulteriori abilità interconnesse fra loro in parti e meccanismi che contribuiscono alla complessità ed al buon funzionamento della capacità cognitiva o non. Nel caso del libero arbitrio, non si può negare che alcune di esse, strutturalmente, affondano le loro radici nei processi neurali.

Il libero arbitrio troppo spesso è ricondotto all'esercizio di speciali poteri metafisici che tralasciano un ventaglio di abilità che contribuiscono alla sua esistenza: la capacità di immaginare un corso d'azione futuro, la capacità di controllo della propria azione e, quindi, la capacità di considerare tali alternative. Il primo passo del processo deliberativo non è, sostiene Nahmias, dire "Sì" o "No" all'esecuzione di un'azione, ma considerare, *in primis*, le varie possibilità, il loro valore così da esercitare, *in secundis*, la propria esperienza di libertà di scelta nel preferire un'alternativa ad un'altra. È l'insieme di queste fasi che forma il libero arbitrio, perché non è solo la scelta di un'azione, ma l'insieme delle proprie abilità di plasmare sia la propria attività non conscia e non deliberativa sia di creare situazioni in cui agire in maniera coerente agli scopi prefissati, anche senza esercitare la presa di decisione cosciente:

Free will is not just about selecting among options, but also about imagining alternatives from which to choose (Sripada ms). Free will also includes our abilities to shape our non-deliberative, non-conscious mental activity and to shape our situations so that we act in accord with our goals even without exercising the capacities for conscious decision-making⁴³⁸.

Nella visione di Nahmias, le scienze della mente possono aiutare a scoprire *come* lavora il libero arbitrio:

Psychology and neuroscience will increasingly illuminate the mechanisms that underlie the capacities essential to free will, as they have with complex human capacities such as perception, language, and memory. These sciences are also discovering that we have less free will than we tend to think— that our capacities for decision-making do not always meet the normative criteria we think or hope they do⁴³⁹.

⁴³⁸ *IDEM*, p. 3.

⁴³⁹ *IVI*.

Di contro, le scienze della mente possono anche aiutarci a sviluppare tali capacità in modo rilevante tramite l'educazione ed il loro costante esercizio, sfatando quelle posizioni secondo cui solo l'uomo sarebbe simile al *Primo Motore Immobile* di aristotelica memoria e restituendoci al processo evolutivo naturale a cui l'uomo e tutte le specie sono soggette.

7.5 Prospettive future

Gli ultimi cinquanta anni di dibattito sul libero arbitrio sono stati attraversati dalla diatriba delle posizioni di Peter van Inwagen, il *Consequence Argument* a difesa dell'incompatibilismo, e di Harry Frankfurt⁴⁴⁰, *Possibilità alternative e libertà morali* a difesa del compatibilismo.

Seppure affascinanti ed intriganti, tali diatribe hanno tralasciato l'eventualità di un terreno comune da cui far partire un'analisi ulteriore del *free will* tale da includere tanto le capacità cognitive di ognuno in grado di considerare consciamente l'alternativa ad un'azione quanto operare una scelta secondo proprie ragioni.

Il nascente approccio sin qui tracciato ad opera di Nahmias, da lui stesso definito “psico-filosofico” – come detto – consta proprio in ciò: ricercare un giusto equilibrio fra la nostra esperienza ordinaria di possedere il libero arbitrio e la sua connessione con le capacità cognitive, esplorate neuroscientificamente, così da:

[...] *in part to understand the degree to which humans, and different individuals, possess them, and to consider their normative components, in part to understand how we might improve them and our opportunities to exercise them*⁴⁴¹.

L'esercizio del libero arbitrio e della libera scelta includono, secondo il filosofo statunitense, alcune o tutte quelle componenti seguenti, ognuna delle quali, a sua volta, richiede specifiche capacità psicologiche:

1. Riconoscere che una decisione ha bisogno di esser presa;
2. Immaginare varie opzioni fra ipotetici percorsi d'azione;

⁴⁴⁰ Cfr. HARRY FRANKFURT, *Alternate Possibilities and Moral Responsibility*, in “Journal of Philosophy”, 66, 1969, pp. 828-839.

⁴⁴¹ EDDY NAHMIAS, *Free Will as a Psychological Accomplishment*, in DAVID SCHMIDTZ E CARMEN PAVEL (A CURA DI), *Oxford Handbook on Freedom*, op. cit., p. 6.

3. Valutare le singole opzioni ed immaginare i risultati alla luce dei propri desideri, valori, obiettivi. Questo processo include una capacità di focalizzare l'attenzione verso o via da considerazioni più o meno rilevanti;
4. Prendere una decisione, selezionando un'opzione;
5. Formarsi intenzioni rilevanti per attuare la propria decisione; ad esempio, decidere quando e come agire, stilando una propedeuticità delle azioni da perseguire per raggiungere l'obiettivo;
6. Quando necessaria, esercitare la forza di volontà per agire in accordo alle decisioni, alle intenzioni, ai piani di azione.

Una settima capacità ci consentirebbe, secondo Nahmias, di agire liberamente in accordo alle decisioni o motivazioni prese o accettate e di essere, potenzialmente, moralmente responsabili senza attuare le sei fasi appena descritte:

7. Assuefazione ed auto-vincolo; si portano avanti azioni che aumentano significativamente la probabilità che si agisca in accordo a decisioni o precedenti deliberazioni nel momento stesso dell'azione, includendo, anche, gli eventuali sforzi per cambiare le abitudini, per regolare i propri stati emotivi, evitare le tentazioni (ritenute tali)⁴⁴².

In genere, tali capacità, come già ricordato, sarebbero trascurate nel dibattito sul *free will*, forse perché:

*We sometimes engage in these processes in more active ways, consciously considering what decisions must be made today, trying to come up with more options to consider, actively shifting attention away from distracting or irrelevant options, and imagining specific action plans in order to prepare oneself to carry them out automatically (e.g., as when athletes or musicians imagine performances ahead of time)*⁴⁴³.

In ogni caso, non ci si dovrebbe limitare a considerare il libero arbitrio solo come espressione manifesta delle “capacità attive” dell'agente (il controllo, la volontà, la consapevolezza...), ma anche di quelle “passive”; ossia di tutti quei processi inconsapevoli, neurobiologici che esprimono il tipo di scelta e di persona che vogliamo essere. In particolare, Nahmias si concentra sull'immaginazione.

⁴⁴² *IBIDEM*, p. 7.

⁴⁴³ *IDEM*, p. 8.

Una volta che si riconosce che una decisione deve esser presa (punto n. 1), ci immaginiamo, di conseguenza, i possibili scenari (punto 2); considerando, poi, tali alternative, spesso ci si ferma ad immaginare (a rappresentarci, consapevolmente) il modo in cui potrebbero svolgersi e come agiremmo se ci trovassimo, realmente, a viverle. In tali operazioni, spesso, lasciamo vagare liberamente la mente, ipotizzando, di contro, altri scenari alternativi che dipenderebbero dalle decisioni assunte in uno dei tanti “mondi” immaginari.

La capacità di immaginare percorsi d’azione alternativi è ciò che spiega, secondo Nahmias, non solo la nostra esperienza di libertà, ma anche il sentimento di *sentirsi libero*, poiché attualizzeremmo, seppur virtualmente, la scelta e l’opzione preferita. Di contro, c’è da constatare che siamo incapaci a scegliere un’opzione mai considerata:

*il fallimento dell’immaginazione limita la nostra libertà*⁴⁴⁴,

sostiene Nahmias.

Queste capacità cognitive legate all’immaginazione sarebbero compatibili con il determinismo in quanto se all’indeterminismo non si recrimina la mancanza di alternative al primo si; *punctum dolens* dello stesso.

La “libertà epistemica”, resa tale dalla facoltà immaginativa, sarebbe possibile proprio dalle lezioni neuropsicologiche relative alla medesima immaginazione, che offrono, di conseguenza, una spiegazione alla nostra esperienza di libero arbitrio; assumendo che i processi neurali coinvolti siano l’origine di quell’effetto domino che conduce alla presa di decisione cosciente allora sarà inappropriato sostenere che il libero arbitrio sia un’illusione⁴⁴⁵.

Le capacità cognitive legate all’immaginazione sono importanti, sostiene Nahmias, per diversi motivi:

1. Forniscono la *latitude for self expression*⁴⁴⁶: possibilità di immaginare se stessi in diversi contesti o situazioni possibili con relativi biasimo o lodi per il proprio comportamento;
2. escludendo improbabili scoperte neurobiologiche che suggeriscono l’irrilevanza causale dell’attività neurale che asservisce l’immaginazione, le nostre opzioni immaginative o il non riuscire a farlo letteralmente aprono o chiudono le possibilità di azioni future. In genere, non è possibile eseguire con successo azioni complesse, piani

⁴⁴⁴ *IBIDEM*, p. 9.

⁴⁴⁵ Cfr. EDDY NAHMIAE E DYLAN MURRAY, *Explaining away incompatibilist intuitions* in “Philosophy and Phenomenological Research”, 88, 2014, pp. 434-467.

⁴⁴⁶ Cfr. MARTIN SELIGMAN, PETER RAILTON, ROY BAUMEISTER E CHANDRA SRIPADA, *Navigating into the future or driven by the past*, in “Perspectives on Psychological Science”, 8, 2013, pp. 119-141.

molto meno complessi, a meno che uno non abbia immaginato come farlo. Le opzioni inimmaginabili non sono realmente opzioni: ciò che immaginiamo è, quindi, una causa che fa differenza in ciò che finiamo col fare⁴⁴⁷.

3. L'immaginazione è "naturalmente" connessa alla responsabilità morale: conseguenze dannose o deplorevoli sono il risultato del fallimento dell'immaginazione.

Una volta compreso che il libero arbitrio sia un insieme complesso di capacità cognitive, si possono considerare i casi in cui esiste un suo *deficit*, come, ad esempio, i disordini psichiatrici. L'opportunità di esercitare la propria capacità di immaginare opzioni future può essere limitato da diverse cause (esterne o interne all'agente stesso), che portano Nahmias a sostenere che ognuno possiede *gradi diversi* di libertà e di libero arbitrio in concomitanza alla situazione o alla storia personale che vive. È proprio su tale constatazione che si incentrano i criteri normativi del *free will* che Nahmias, ad oggi, li racchiude in tre:

1. L'immaginare opzioni alternative;
2. Riuscire a realizzare ciò che si è immaginato come un percorso d'azione fattibile, attuabile, secondo i propri desideri, obiettivi, valori, interessi;
3. L'attuazione pratica del percorso scelto.

Gli individui che possiedono un grado di libero arbitrio più basso rispetto alla norma sono, in genere, gli emarginati o coloro che vivono un disagio sociale, educativo o familiare per cui non dovrebbero essere ritenuti responsabili delle scelte che operano.

Le ragioni di tale asserzione, però, non sono motivate dallo stesso Nahmias, suggerendo solamente di essere "giustificazionisti" per coloro che, nella società, delincono.

Tale discorso è, decisamente, pericoloso: per una *china scivolosa*, dove finirebbe la pena per i trasgressori?

Nonostante quanto finora detto, alcune ricerche suggeriscono che le azioni non siano sempre coerenti con le ragioni o motivazioni che, consciamente, consideriamo giuste: le ricerche di psicologia sociale suggeriscono che spesso siamo influenzati da fattori inconsapevoli e che rifiutiamo di riconoscere, laddove lo analizzassimo o ce lo facessero notare, che, oggettivamente, siamo stati influenzati da questi fattori; esempio, l'immobilismo in un gruppo nell'aiutare qualcuno che chiede aiuto, l'influenza della folla etc.

L'idea, comunemente condivisa, che una naturalizzazione della mente possa comportare la successiva perdita delle peculiarità del tutto umane minaccia l'esistenza del libero arbitrio e

⁴⁴⁷ EDDY NAHMIAS, *Free Will as a Psychological Accomplishment*, in DAVID SCHMIDTZ E CARMEN PAVEL (A CURA DI), *Oxford Handbook on Freedom*, op. cit., p. 11.

Nahmias chiama tale posizione *argument from rationalization*⁴⁴⁸, riassumibile in tale sillogismo:

1. Il libero arbitrio richiede che un'azione sia il risultato delle ragioni consciamente analizzate dall'agente;
2. La Scienza sta mostrando che le nostre azioni non provengono direttamente dalle ragioni conscienti dell'agente, ma, piuttosto, le azioni sono prodotte da altri fattori inconsci che, *post hoc*, li razionalizziamo come fossero motivi coscienti all'azione;
3. La Scienza sta mostrando che l'uomo ha perso il libero arbitrio (e la responsabilità morale).

Molte teorie sul dibattito sul *free will* sposano o sarebbero d'accordo con la plausibilità della premessa 1, così come alcune evidenze del *modular epiphenomenalism* supportano la premessa 2; entrambe non considerano che, come già detto, da sole le neuroscienze non saranno in grado di redimere controversie o donare una forte spinta risolutrice al problema sul *free will*, poiché sono solo parte “della verità”, non la sua totalità. Già, ad esempio, le ricerche sulla filosofia morale o la psicologia sociale offrono un aspetto “contenutistico” e non “formale” al comportamento che il naturalismo o lo stesso determinismo non possono, per loro natura, offrire.

Si sta presentando, cioè, l'ipotesi che l'uomo posseda meno libero arbitrio di quanto creda: *gradi di libertà* al posto del libero volere; “libero” anche da ciò che non sarebbe in grado di volere.

Ecco, allora, secondo Nahmias, allo *status quo*, che cosa sappiamo e che cosa non sappiamo del libero arbitrio:

1. Non sappiamo se il determinismo sia vero. Sappiamo che la scienza che studia il cervello ed il comportamneto umano è inadeguata a stabilirlo. Abbiamo anche buone ragioni di credere che le risposte alle domande sulla verità del determinismo siano meno importanti per le persone comuni più di quanto suggerisca la tradizione.
2. Non sappiamo se il naturalismo sia vero. Non sappiamo, infatti, se ci siano menti “non fisiche” o leggi psicologiche al di fuori di quelle che governano l'universo. Ma abbiamo buone ragioni di credere che il metafisico, in genere, e le scienze cognitive forniscano prove induttive a favore del naturalismo.

⁴⁴⁸ EDDY NAHMIAS, *Is Free Will an Illusion?*, in WALTER SINNOTT-ARMSTRONG (A CURA DI), *Moral Psychology*, vol. 4: *Freedom and Responsibility*, op. cit., p. 18.

3. Non abbiamo una teoria della coscienza che spieghi la relazione fra la mente cosciente ed il cervello. Nella visione di Nahmias, è questa carenza di conoscenza che induce le persone, inclusi alcuni scienziati, a pensare che il naturalismo sia una minaccia per l'esistenza del libero arbitrio. O, per lo meno, lo escluda del tutto. Quando si parla, infatti, di processi neurale si tenta di spiegare un dato comportamento, ma, allo stesso tempo, è "oggettivo" che manchi la relazione fra lo stesso ed il processo cosciente, che soggiace o conduce il medesimo. Un particolarismo che Nahmias sospetta sia presente *anche* nelle scienze cognitive.
4. Non sappiamo quanto il nostro comportamento sia influenzato dai correlati neurali della coscienza, pur avendo evidenze che essi giochino un ruolo causale in alcuni comportamenti.
5. Non sappiamo quanto il nostro comportamento sia razionale e coerente con i nostri giudizi morali. Le nostre capacità di agire coerenti con le nostre ragioni non sappiamo se siano state accettate come tali o provengano da un'analisi personale: se le ricerche future suggeriranno che tali capacità razionali siano limitate più di quanto si creda allora saranno limitate anche il raggio d'azione del libero arbitrio e della responsabilità morale.

Alla luce di quanto detto, esiste, allora, il *free will*?

Nahmias, ad esempio, è convinto che l'uomo abbia il libero arbitrio e, seppur limitato nell'esercitarlo, avrebbe bisogno di imparare il modo in cui poterlo sviluppare ed usarlo con (maggiore, n.d.r.) saggezza⁴⁴⁹, confidando nel fatto che *se e quando* avremo una teoria della coscienza sarà chiara la relazione fra essa ed i processi neurali sottesi e, di conseguenza, saremo vicini ad un'adeguata comprensione dello stesso.

Per quanto riguarda la scrivente, ad oggi, ha più dubbi che certezze non solo sulla sua esistenza, ma anche su ciò che significa la stessa espressione "libero arbitrio".

⁴⁴⁹ *IBIDEM*, p. 21.

Conclusioni

Stilare le conclusioni di un lavoro è opera ardua, complicata e difficile.

Se l'introduzione è ciò che il lettore affronta con incertezze e curiosità, le conclusioni rappresentano, ai suoi occhi, "la risposta" definitiva a quello scetticismo ed a quel desiderio di sapere con cui si è lasciato condurre nella direzione di quel filo rosso che l'autore, consciamente, ha tentato di sbrogliargli nel corso dell'opera. È in quel momento, allora, che se il lettore, da una parte, avverte l'esigenza (forse, "il diritto") di veder riconosciuto il suo sforzo nell'essersi fidato, dall'altra, l'autore intravede nelle conclusioni il *locus* in cui "si è disvelato". Infatti, "farsi leggere" rappresenta un atto di completa confidenza privata dell'intimità dello stesso; il quale, in ciò che ha scritto, ha trasferito completamente il suo Sé. Non si tratta di un'ermeneutica fra testo/lettore alla Gadamer, ma, semplicemente, di una "presa d'atto" delle difficoltà cui la scrivente incorrerà per ciò che le stesse *Conclusioni* rappresentano.

Si è esordito, all'inizio del lavoro, sostenendo che il *punctum dolens* degli studi che sottostanno a queste pagine è stato proprio il concetto di empatia. Una difficoltà che sembrava insormontabile è stata, come detto, tracciare le linee o mettere dei punti fermi al fenomeno, poiché nella letteratura odierna il concetto di *empathy* è utilizzato un po' come il *jolly* nei giochi di carte: a seconda del contesto, quest'ultimo diviene uno o un altro fenomeno con caratteristiche diverse o parziali rispetto a quello che si riscontrerebbe in un altro contesto con altre o altrettante caratteristiche.

Ma per quali motivi?

La difficoltà principale, come si può immaginare, è stato *capire* quale fosse "l'essenza stessa" dell'empatia tale da poter delineare con chiarezza gli immaginari paletti concettuali entro cui descriverla; uno dei quali, ad esempio, rappresentato dal grave limite strutturale della stessa (che Hoffman più volte ricorda nelle sue ricerche): la *deriva egoistica*⁴⁵⁰.

Il problema di come interpretarla non è di facile risoluzione, poiché, personalmente, a seconda del modo in cui si intenda, il lettore potrebbe trovarsi di fronte ad una conclusione che, *prima facie*, sembrerebbe essere o troppo poco drastica per un verso o troppo perentoria per un altro.

⁴⁵⁰ Si rimanda il lettore al capitolo II.

La deriva egoistica, si ricorderà, è quel “fenomeno” secondo il quale l’osservatore, posto di fronte alle sofferenze di una vittima, immagina – attraverso uno sforzo cognitivo – il “mettersi al posto dell’altro”, rievocando, ad esempio, ricordi personali che possono far passare in secondo piano la situazione emotiva del malcapitato. Hoffman spiega tale fenomeno attraverso il ricorso al meccanismo dell’*assunzione di ruolo centrata su di sé*, in cui l’Ego dell’osservatore prende il sopravvento su quello altrui, tanto da spostare l’attenzione dell’osservatore stesso dal Tu all’Io; la conseguenza diretta è il dissolvimento completo dell’empatia per la vittima, in virtù di una situazione emotiva che ponga al centro dell’attenzione il solo osservatore (in questo senso diviene *centrata su di sé*).

La letteratura in merito spiega il fenomeno attribuendo all’*empathy* gravi limiti strutturali (esempio è la stessa *abituazione*) per cui, prendendone atto, cerca di scoprirne degli ulteriori con altrettanti particolarismi, ipotizzando o un funzionamento parziale del sistema neurobiologico che la renderebbe possibile o ricorrendo ad ipotesi di varia natura che si spiegherebbero con la medesima complessità del fenomeno.

Su tali approcci non sono d’accordo.

Sul primo, come ho già avuto modo di ribadire *in itinere*, non si può contare sulla rigosità di qualsivoglia spiegazione, poiché gli attuali limiti strutturali delle tecnologie a disposizione in questo particolare periodo storico, uniti alla complessità evidente del funzionamento cerebrale (di cui, sempre stando ad oggi, la Comunità scientifica sembra essere ben lontana dal “disvelamento” dell’ “enigma cervello”), non rendono possibile avanzare una teoria duratura e solidamente ben strutturata su qualsiasi anomalia delle aree cerebrali coinvolte nell’empatia di base. Di conseguenza, essendo qualsivoglia spiegazione complementare alla “dottrina scientifica”, mancando la completezza della seconda manca la completezza anche della prima. Si dovrebbe andare “oltre”, cercando una spiegazione del fenomeno che coincida non sulle modalità entro cui esso avvenga, ma sul perché un grave limite strutturale come il dissolvimento o lo spegnimento dell’empatia possa avvenire. Ci si inserisce, allora, nel secondo approccio al problema.

A tal proposito, non volendo interessare l’ambito di interpretazione neuroscientifica/biochimica per le ragioni che più volte si è avuto modo di ribadire e perché, sostanzialmente, il presente aspira ad essere un lavoro di approccio morale al “problema empatia” (individuando nell’aspetto cognitivo l’origine di ogni fraintendimento o perplessità in merito), qualsiasi “soluzione” in ambito delle scienze umane si possa fornire si troverebbe, inevitabilmente, di fronte a due possibilità:

1. la deriva egoistica, essendo la risultante di un atto immaginativo, probabilmente avrebbe luogo, poiché la stessa “forza immaginativa” impiegata per “immedesimarsi nella situazione altrui” non sia così abile a superare il proprio istinto di sopravvivenza, tipico di qualsiasi specie vivente; di conseguenza, “esaurendosi” la propulsione dell’immaginazione, l’osservatore cesserebbe l’atto cognitivo con protagonista la vittima per trasformarlo, in tal modo, in un altrettanto atto cognitivo con protagonista se stesso.

Sarebbe così spiegato il dissolvimento, lo “spegnimento” dell’empatia per il Tu.

Da ciò si potrebbe dedurre che l’uomo, sostanzialmente, sarebbe falsamente altruistico, riprendendo un tema caldo di questi anni, per cui empatico, sì, ma fino ad un certo punto.

La deriva egoistica rappresenterebbe, allora, quel “punto di non ritorno” o quel “punto oltre il quale non poter andare” ad essere disponibili verso l’altro; una “disponibilità”, si tenga in mente, che significhi “essere in grado di aprirsi all’opzione di soffrire *quanto* la vittima e, allo stesso tempo, tentare di *capirla*”; in questo senso, la “disponibilità” dell’uomo sarebbe sia “limitata”, poiché *immedesimarsi nei panni altrui* significherebbe soffrire o gioire, a seconda dell’emozione dell’altro, sia egoisticamente “determinata”.

Un limite di concessione alla sfera emotiva del Tu che, forse per istinto di conservazione e/o sopravvivenza, limita l’accesso all’intimità propria della vittima per salvaguardarla da sé e dall’altro.

La deriva egoistica, allora, rappresenterebbe proprio quella concessione *naturale* cui l’animale umano sarebbe oggetto.

Si potrebbe dire che il *Sapiens Sapiens* non sia altruistico?

Sinceramente non sarei d’accordo, se la spiegazione fosse ciò che si è appena letto. I concetti di *egoismo* ed *altruismo* sono più complessi di una mera banalizzazione di non concedersi all’altro; probabilmente, anzi certamente, ad un’attenta analisi, si giungerebbe all’identificazione di parole più adeguate della mera contrapposizione entro cui si vuole o trasformare l’uomo in un essere fortemente altruistico o, di contro, fortemente egoistico.

2. Una seconda interpretazione sarebbe quella secondo cui la deriva egoistica rappresenterebbe non tanto il grave limite strutturale dell’empatia cognitiva, ma l’empatia cognitiva stessa; non ci si troverebbe, cioè, di fronte ad un limite della stessa, ma al sintomo che la *cognitive empathy* non sarebbe il fenomeno che si creda sia.

Non si tratta di spiegare un limite del fenomeno, ma essere accorti nel capire che sia un sintomo o un’anomalia (riprendendo il significato di Khun) dello stesso; essendo la deriva

egoistica la risultante di un atto cognitivo, perché non cambiare punto di vista e non vedere che l'osservatore, posto di fronte alla sofferenza della vittima, cognitivamente *sceglie consapevolmente, consciamente* di "lasciarsi andare" alla situazione emotivamente forte del Tu per poi giungere, finché non cali o si sposti l'attenzione dalle situazioni fattuali che si trovi a vivere, a quelle personali, intime che l'osservatore ha già vissuto o *immagina* potrebbe vivere?

Volendo, cioè, stilare un prima da un dopo, sarei incline ad esplicitare questa seconda interpretazione come segue:

scelta consapevole dell'osservatore → *attivazione consapevole* dell'immaginazione → "immedesimazione nella situazione altrui" → ricordi personali (*memoria*) → spostamento dell'emotività dall'altro a sé.

Non si tratterebbe di un limite del fenomeno empatico, ma la risultante di un processo che partirebbe dal libero arbitrio proprio dell'agente di volersi cimentare in una situazione (emotiva o esperienziale) che ben si allontanerebbe dalle spiegazioni solite che, in genere, si attribuiscono o alla deriva egoistica o alla stessa empatia.

Stando a tale interpretazione (ancora ben lontana da una strutturazione sistematica e teorica solida, ma in costruzione) ne deriverebbe che la *cognitive empathy* non soltanto non risulterebbe essere il fenomeno che interesserebbe l'ipotetico rapporto fra osservatore e vittima, ma che tutte le spiegazioni successive relative ai limiti della stessa dovrebbero essere riconfigurate come parte integrante una spiegazione di altra natura; addirittura di un altro fenomeno, che non si configurerebbe, come preannunciato, nell'empatia stessa.

Si spiegherebbero, in questo modo, anche i motivi per cui, giunti ad un certo punto, l'immedesimazione emotiva verso e per l'altro cesserebbe; potrebbe subentrare, e nulla vieta di sostenerlo, quell'istinto di sopravvivenza o di difesa personale cui si è accennato poco più su, facendo cessare, conseguentemente, il *mindreading* verso la vittima stessa.

Una interpretazione, quest'ultima, che deriva dalla convinzione che nel corso della storia del concetto empatico si sia assistito ad una incomprensione dello stesso che non riguardi le modalità di attivazione o di presenza del fenomeno, ma la sua stessa natura; un *qui pro quo* i cui effetti hanno impedito la sua chiarificazione reale, ben lontana da quanto oggi molti pensano di sapere sull' "empatia".

Come può esser possibile, infatti, dare così tante definizioni ad un fenomeno, concedendo, a sua volta, che quest'ultimo si plasmi, costantemente, in modi differenti negli ambiti di studio che ne costituiscono l'oggetto di indagine?

Un problema, dal punto di vista dell'interrogante, che diviene necessario approfondire, poiché, a quel punto, ciò potrebbe significare:

1. Il fenomeno che si sta descrivendo non è quello che si crede di descrivere (si è di fronte, cioè, ad un fenomeno diverso, scambiato per un altro);
2. Il fenomeno che si sta descrivendo non esiste;
3. Il fenomeno che si sta descrivendo non soltanto non è ciò che si pensa, ma, sostanzialmente esso non esiste.

Tenendo presente che i tempi del pensiero non sono quelli della vita dell'uomo, ma che le idee, i concetti, i dubbi, i dibattiti sui problemi posti da uno o da un altro filosofo non si esauriscono con la fine della vita dello stesso, ma che essi influenzano le epoche successive, si deve tener presente che nel XVII secolo Cartesio poneva la domanda se la natura umana fosse solo *res cogitans* o solo *res extensa*, gettando le basi di un dualismo di sostanza ancora ben presente nell'odierno dibattito sulla filosofia della mente.

Il ricorso al problema *mind/brain* ed al modo in cui sia possibile una loro inter-relazione (ammettendo che siano due entità separate) non è casuale: rappresenta, infatti, l'inizio della chiave di volta di un *qui pro quo* che ha dato origine ad un dibattito che, sostanzialmente, è mal posto; l'*empathy* in quanto tale.

Mentre in Occidente la tradizione filosofica si è incentrata, nei secoli, sugli aspetti più ermeneutici o esistenzialistici del pensiero stesso, negli Stati Uniti, a partire dal Positivismo Logico (che, al pari di quello ottocentesco, ambiva alla formalizzazione di un monismo metodologico fra le scienze umane e quelle naturali), la ripresa del concetto di empatia, nel significato di *Verstehen* fu, caldamente, osteggiata. Solo Dilthey, volendo semplificare il discorso, ne parlò quale "trasposizione psichica" del soggetto di "comprendere le volontà dell'altro", giungendo a Weber quale riconoscimento dell'empatia come strumento epistemico per la comprensione razionale dei motivi che soggiacciono all'azione in vista di uno scopo; per un'attenzione ed una rivalutazione del concetto si dovranno attendere gli anni Settanta.

In questo periodo, infatti, dalle teorie dell'azione ispirate alla filosofia di Wittgenstein e dai successivi dibattiti scaturiti dall'*action theory*, si arriva a Putnam e Nozick, i quali fanno dell'empatia lo spunto per brevi osservazioni, l'uno; oggetto di intense analisi, l'altro.

Per Putnam, l'empatia non può essere considerata uno strumento epistemologico *tout court*, in grado, cioè, di fornire una conoscenza formalizzabile e razionale secondo il classico modello

hempeliano di leggi di copertura e processo nomologico-deduttivo per la *conclusio* del sillogismo, ma, certamente, non nega che l'empatia, appellandosi ai criteri di *semplicità* e *plausibilità*, possa fornire "opinioni corrette" da cui, successivamente, scaturirà "conoscenza". Putnam, cioè, non si interroga, a differenza di Nozick, dei motivi per cui l'*empathy* possa fornire opinioni corrette (o, osando, "conoscenza"), poiché quest'ultima è utilizzata "solo" come spunto di riflessione per altri tipi di problemi che non rientrano nel modo in cui "ci si metta nei panni altrui, provando, a sua volta, le medesime emozioni osservate". Inoltre, il filosofo statunitense individua sul "concetto empatia" il problema del controllo: seppur egli ammette la plausibilità del tipo di "intuizioni" che essa fornisce sulla *ratio* del comportamento altrui, in che modo queste potranno essere sottoposte al vaglio del controllo?

La risposta fornita dallo stesso Putnam è che anche quest'ultimo sarà, a sua volta, intuitivo così come le "conoscenze" di cui è l'espressione; che cosa significhi, però, che la natura del controllo sia essa stessa intuitiva non risulta essere molto chiaro: sembra, infatti, che neppure il filosofo statunitense abbia ulteriormente chiarito tale aspetto, aggiungendo, "semplicemente", che la conoscenza intuitiva, basata su un sapere pratico, sarà a sua volta oggetto di controllo da parte di dati altrettanti intuitivi.

Alcuni interpreti hanno visto in tale circolarità il ricorso alla *folk psychology*, poiché un "sapere pratico", un "controllo intuitivo" o "conoscenza intuitiva" lascerebbe pensare a molti che si tratti di un tipo di "sapere" non caratterizzato da complesse riflessioni logico-razionali; bensì da una natura "più semplice", condivisa dai membri di un medesimo gruppo in uno stesso contesto, tipica, cioè, della *psicologia di senso comune*.

Iniziano, allora, a configurarsi due elementi centrali che porteranno, ad avviso della scrivente, ad una incomprensione del fenomeno:

- l'idea che l'empatia possa fornire una forma, seppur semplice, di "conoscenza";
- il ricorso alla *folk psychology*.

Il quadro non si farà più chiaro con l'avvento di Nozick, il quale, dando seguito alle riflessioni di Putnam, "andrà oltre", conferendo legittimità epistemica all'empatia, grazie all'ausilio di due argomentazioni centrali:

- a. i concetti di *plausibilità* e *semplicità* delle "opinioni intuitive" di putnamiana memoria;
- b. essendo il *Verstehen* inteso come l'atto immaginativo attraverso il quale l'osservatore si pone dal punto di vista dell'altro (vedendolo come simile a sé), il ragionamento ricorda quello della logica induttiva: essendo possibile riproporlo attraverso le *leggi di*

copertura “alla Hempel” allora avrà la medesima natura di qualsiasi ragionamento scientifico.

Ricapitolando, finora nel discorso si sono presentate tre premesse:

1. Il dualismo cartesiano;
2. L'introduzione, nel dibattito, della *folk psychology*;
3. L'empatia come inferenza “conoscitiva” dell'osservatore nei confronti dell'altrui soggettività.

Prima di aggiungere la quarta premessa, vorrei proporre alcune riflessioni.

Volendo esemplificare il discorso per giungere al *focus* dell'argomentazione, si potrebbe sostenere che dal XVII secolo ad oggi (momento storico che rimanda la memoria alla speculazione di Cartesio) si è dibattuta, in una certa corrente del pensiero, la difficoltà di “stabilire” la natura “monista” o “dualista” dell'uomo *tout court*. È anche evidente che le risposte e le articolazioni alla domanda posta dal filosofo francese hanno avuto infinite sfaccettature ed altrettante innumerevoli risposte fino a quando, dall'avvento del Positivismo Logico e dal Neopositivismo Logico, il modo di *fare* filosofia, volente o nolente, è mutato. I vecchi problemi che attanagliano tuttora la tradizione occidentale sono stati reinterpretati o in parte abbandonati in quella analitica; ma un problema, però, restava: come dar conto della natura umana? È mente e cervello o solo uno dei due?

Per non cadere, nuovamente, nella “trappola cartesiana” di dover dar conto della natura di una o di due *res* con ciò che, necessariamente qualsiasi tipo di risposta avrebbe comportato, ritorna all'attenzione del dibattito filosofico il concetto di *humeana memoria*, la *simpatia*, che se in Hume era indicativa di una *naturale benevolenza*, con Titchner e con coloro i quali si affideranno ad essa diventerà *dentro le passioni* con tutti i significati e le sfaccettature note.

Da qui nasce il *qui pro quo*.

Il ripristino, se si vuole, dell'empatia ha costituito il mezzo attraverso il quale la tradizione analitica, rigettando qualsiasi ricorso al trascendente (le due *res*, ad esempio) ed approcciandosi al mondo in maniera del tutto trascendentale/naturale, ha fornito una risposta al problema posto già da Cartesio: che fosse una *naturale benevolenza* o una intuizione sentimentale attraverso la quale “capire l'altro” il fenomeno avrebbe *sempre e comunque* avuto un *quid* naturale, non sostanziale o duale.

Fornito, allora, il ricorso al mezzo per ovviare al problema cartesiano, restava insoluto non tanto il problema di come spiegare l'inter-relazione fra uomini (Hume, d'altro canto, ne avrebbe già fornito la risposta), quanto l'ausilio di un intermediario; un collante fra soggetti

inseriti in una società per fornire quella plausibilità e quella legittimità che Putnam, prima, e Nozick, dopo, e molti dopo di loro, hanno tentato o proposto di dare all'empatia.

Sellars, a questo punto, è stato in grado di fornire quel *traid d'union* di cui si aveva bisogno attraverso il ricorso a ciò che di più trascendentale possa esistere, non indagando o non fornendo il destro per ulteriori indagini che potessero fuorviare dalla “natura naturale” del mezzo stesso: il *sensu comune*. Il “si dice, si pensa...”, che Heidegger aveva identificato quale *esistentivo* ben lontano dalla ricerca autentica dell'essere, torna, nella filosofia analitica, sotto le medesime spoglie, ma con un accrescitivo valoriale ben diverso: esso costituisce il mezzo attraverso il quale la relazione umana può aver luogo. Si ovvia, in tal modo ad un secondo problema: la comunicazione fra monadi di Husserl.

Il binomio *empatia-folk psychology*, allora, diviene la spiegazione trascendentale alla natura umana, mettendo, “apparentemente”, da parte Cartesio e l'insieme dei problemi trascendenti cui, inevitabilmente, il ricorso o il rifiuto delle due *res* richiama; quantomeno nel fornire una risposta per l'uno o per l'altra parte.

C'è, però, un'ulteriore precisazione da fare: perché l'utilizzo dell'avverbio “apparentemente”? Perché, sostanzialmente, per quanto la filosofia analitica utilizzi un metodo distinto rispetto a quello continentale, la maggior parte dei problemi resta la stessa: il *mind/brain problem*, ad esempio è uno fra questi; senza contare quelli relativi alla natura del mentale o alla coscienza. Cambiano i contesti, ma non i problemi.

Esemplificativo di questa interpretazione che si sta delineando è l'atteggiamento dei filosofi (alcuni dei quali sopramenzionati) che hanno, progressivamente, avallato, contribuito e supportato questo *escamotage* fino a quando non si è giunti, addirittura, a parlare di *mindreading*.

La storia del pensiero in merito alla *lettura della mente* identifica il saggio di Premack e Woodruff quale inizio simbolico per l'attenzione dei filosofi (nuovamente) ad una teoria della mente; in realtà, si dovrebbe parlare di un *ripristino* o, volendo essere più precisi, di un *continuum* che non è mai stato interrotto.

Sinceramente, il saggio dei due etologi, da un punto di vista speculativo prettamente dedito all'animale umano, non aggiunge o non toglie nulla di quanto già si sapesse o si ipotizzasse sulla mente umana; ciò, però, che è stato in grado di fare è stato quello (volontariamente o non) di inserirsi in quella costruzione progressiva dell'*escamotage empatia* per “risolvere” il problema dell'intersoggettività. Serviva, cioè, un elemento che fosse in grado di spiegare, con plausibilità e legittimità, i motivi per cui un agente, inserito all'interno di un contesto dato (*folk psychology*), fosse in grado di “capire l'altro”. Il ricorso ai mezzi già noti (empatia e

credenze condivise) non potevano essere sufficienti, poiché si cercava di fornire una giustificazione razionale al tipo di conoscenza *epistemica* che l'empatia potesse fornire e senza un supporto "scientifico" quest'ultima appariva, semplicemente, mera intuizione. Ecco, allora, che il saggio di Premack e Woodruff appare sulla letteratura internazionale, fornendo, a quel punto, l'"ancora di salvataggio" dell'intero complesso che si stava delineando: il *mindreading*, il *mentalizing*, la *lettura della mente*.

Un ingranaggio all'interno della catena che, ovviamente, non fu fornito solo ed esclusivamente dal saggio menzionato, poiché in tale discorso anche Quine ha fornito il destro per indirizzare/incrementare l'inquinamento della natura "empatica" attraverso l'introduzione della *traduzione radicale*.

Eccoci, allora, giunti, alla quarta premessa.

È noto ai molti l'esempio tipico del filosofo statunitense riguardo le difficoltà cui incorrerebbe un ipotetico linguista che, in un contesto estraneo con aborigeni che parlano una lingua a lui ignota, si trovi nelle condizioni di tradurre il termine *gavagai* in una parola a lui familiare (*coniglio*).

Al di là degli aspetti rilevanti della sua filosofia, nell'esempio Quine opera un'analogia o un'associazione fra il comportamento ed il concetto di empatia, giungendo, semplificando, alla "risoluzione" del problema dell'attribuzione intenzionale:

*mettendo in rilievo la legittimità empirica, se non proprio quella scientifica, del processo psicologico di empatizzazione, del quale, pur nel rifiuto di ogni ontologia mentalistica, segnala l'inevitabilità e l'efficacia per lo meno nelle attribuzioni intenzionali della folk psychology*⁴⁵¹.

L'uso della *psicologia di senso comune*⁴⁵² ha costituito e costituisce tuttora una parte significativa del dibattito sul processo empatico negli sviluppi e nelle modalità già delineate.

L'empatia, vale a dire, si fa *psicologia di senso comune*: quel mezzo trascendentale attraverso cui ovviare al problema del dualismo cartesiano di come comprendere l'altro, trasformandosi, a quel punto, nel "problema della lettura della mente". Non si tratta, cioè, di una sofisticazione del pensiero o di un corollario al problema principale (che cos'è l'empatia), ma della traslitterazione trascendentale del problema del dualismo cartesiano: anziché spiegare la natura umana attraverso il ricorso alle due *res*, si spiega quest'ultima attraverso il

⁴⁵¹ ANTONIO RAINONE, *La riscoperta dell'empatia. Attribuzioni intenzionali e comprensione nella filosofia analitica*, Bibliopolis, Napoli, 2005, p. 75.

⁴⁵² Si rimanda al capitolo IV.

mindreading in relazione alla *metacognition*: “sostanza pensante” in relazione alla “sostanza estesa”.

Tale spiegazione si evince anche dai primi concetti di chiarificazione, presentati nel capitolo V, su che cosa si debba intendere per “modello sulla priorità della lettura della mente”: la base di partenza per il lavoro di tale abilità resta l’ambito sensoriale o la diretta percezione di *input* esterni a riprova del fatto che l’intento di fondo, sia di Carruthers sia di molti altri autori, è quello di ovviare o “risolvere” il problema classico delle due *res*. Qualsiasi tipo di argomentazione apportata dall’uno o dall’altro modello resta, tuttavia, pur sempre la guida unica per dare una risposta di stampo trascendentale alla natura della mente, poiché per quanto si possa essere d’accordo sulla sofisticazione progressiva del discorso relativo al *mindreading/metacognition*; interpretazione/introspezione, si manifesta, costantemente, il timore di non fornire elementi che possano far pensare ad un dualismo di sostanza lontano dal “naturale”. Non a caso anche nello stesso modello di Gopnik quest’ultima sostiene, come detto, che tutte le attribuzioni degli stati mentali a se stessi sarebbero basati su una teoria, ad eccezione di ciò che l’Autrice chiama “ronzio cartesiano”, il “Cartesian buzz”: il problema, allora, diviene costruire una teoria attendibile che sposi il progetto di naturalizzazione del mentale che, nulla in contrario, non dovrebbe stupire o sollevare enormi polveroni, almeno per chi, come la scrivente, non vede tale intento come una *diminutio* della natura umana; anzi! Ma al di là delle considerazioni personali su questo punto, si vuole mettere in luce un altro aspetto, non secondario, del discorso che sin qui si è tenuto fra *metacognition/mindreading*: che fine ha fatto l’empatia (cognitiva)? L’ipotetico lettore potrebbe rispondermi, se siamo riusciti nell’intento, che il problema del *mindreading* è l’aspetto controverso e rilevante dell’empatia cognitiva (si è esordito in tal modo all’inizio del capitolo terzo), ma in accordo a quanto appena evidenziato, il concetto di qualsiasi tipo di empatia si tratti è lontano (o, forse, assente, mi azzardo a dire) dagli autori considerati. Parlare di *mindreading* o di *metacognition* non significa per il dibattito attuale, che si sta svolgendo in queste ore, trovare una spiegazione a quello sforzo cognitivo con cui interagisco con l’altro, poiché rifletto sulla sua situazione emotiva, ma significa interrogarsi, ancora una volta, sulla natura del mentale.

Lo stesso ausilio della *folk psychology* e di ciò che tale nozione è diventata nell’evoluzione del dibattito da Sellars ad oggi, 2019, mette in mostra questo aspetto e, soprattutto, che l’empatia è divenuta lo specchietto delle allodole per parlare di altro; il problema diviene, però, capire il perché non se ne parla.

Sono sempre più convinta che nel processo di sofisticazione avuto dal Neopositivismo Logico ad adesso il concetto originario di empatia (qualsiasi fenomeno questa denotasse) è stato

denaturalizzato per servire a dare una risposta quanto più trascendentale al problema classico delle due *res*. È stato un processo generatosi per gradi: dapprima la ripresa da Lipps del concetto di empatia; dopodiché l'introduzione, da parte di Sellar, del "Mito di Jones"; poi, ancora, il ricorso alla *folk psychology*; la suddivisione delle risposte fornite dalla *Theory Theory* e dalla *Simulation Theory* per finire, almeno per il momento, con l'ultimo tassello del quadro, consistente nella relazione fra *mindreading* e *metacognition*.

Altre strade, da ciò che sto apprendendo, si stanno per aprire e nuovi interrogativi stanno nascendo, a tal proposito, ma ciò evidenzia, ancora una volta, che, progressivamente, il denotato (l'empatia) si è svuotato di senso per dare spazio ad altri denotati che nulla hanno a che fare con quello originario. In tale percorso, ovviamente, non posso tralasciare l'impatto che la scoperta del *sistema mirror* ha avuto sulla riflessione in oggetto, in quanto ciò ha fornito "nuova linfa vitale" al dibattito, corroborandolo della convinzione che "si era sulla giusta strada" per "abbandonare" Cartesio.

Si pensi, ad esempio, alle "origini" della *Simulation Theory*: dapprima un "riproporre internamente" alla Lipps l'emozione o l'azione osservata, dopo, negli anni '90, una "simulazione cerebrale" dell'osservatore di quell'emozione o di quell'azione.

A prescindere dagli sviluppi che la ST e la TT stanno avendo, resta sempre il fatto che l'uso della psicologia di senso comune e dello strumento dell'empatia sono stati i mezzi per risolvere un problema annoso, che da Cartesio continua ad esistere e ad interessare le riflessioni attuali.

Di conseguenza, il concetto stesso di empatia, se in Hume, ad esempio, poteva trovare una sua *ratio* quale *benevolenza* verso il prossimo alla base della morale umana, oggi è un concetto vuoto: svuotata completamente di senso, è la falsa spiegazione del modo in cui dar conto dell'intersoggettività non ad un livello base, ma sempre cognitivo. Il fatto stesso che si plasmi in tanti modi quanti sono gli interroganti nei diversi contesti dovrebbe aver posto/porre un notevole campanello d'allarme.

Si pensi, ad esempio, agli esperimenti di Batson (tanto per rimanere in ciò che è stato oggetto di chiarificazione nelle pagine precedenti).

A fondamento delle sue spiegazioni sembra ci siano solo considerazioni generali su che cosa indichi, sommariamente, l'etichetta empatia. Sembrerebbe, cioè, che non ci sia un'analisi approfondita sul significato, sull'uso del termine e sullo *status* ontologico del fenomeno stesso. Mi si potrebbe obiettare che l'intento dell'autore è capire le motivazioni all'altruismo, ma tale ricerca sembra a chi scrive manchevole dalle fondamenta: Batson dà per scontato che

tanto l'empatia esista quanto la sfera dei comportamenti adottata, derivanti, ipoteticamente da essa, siano veritieri.

Il problema, oltre ad essere di ordine gnoseologico, è divenuto anche metodologico: studiare gli effetti di un fenomeno senza preoccuparsi della sua effettiva esistenza sembrerebbe poco affidabile anche perché gli esperimenti che Batson ripropone nell'arco della sua ricerca sono, per la maggior parte, basati sulla visione, ossia sull'osservazione per mezzo di uno schermo televisivo, della sofferenza inferta alla vittima. Ciò significa, in linea con le scoperte del sistema specchio, prendere atto che l'uomo "funziona" in un certo modo: dove si troverebbe, infatti, quello sforzo cognitivo del porsi nella situazione osservata affinché, di riflesso, l'osservatore stesso *senta* la sofferenza dell'altro?

L'aspetto controverso del fenomeno empatico, vale a dire, è completamente estromesso dalle indagini esperienziali per cui, ad avviso di chi scrive, le ricerche di Batson testimonierebbero l'ulteriore veridicità del *sistema mirror* in cui alla visione di un fenomeno si reagisce, neurobiologicamente, in un dato modo. E non la propensione "naturale" o "fittizia" all'aiuto reciproco.

È difficile, perciò, dedurre da quanto appena esposto che il soggetto "provi empatia": certamente quella di base, ma tali evidenze sono ben lontane dallo stabilire se in condizioni diverse, non visive e non dirette, si possa provare *un* "fenomeno empatico".

Il lettore, allora, potrebbe chiedere, legittimamente, che cosa effettivamente sia l'empatia.

Seguendo, innanzitutto, quell'intuizione principe che mi ha condotto ad affrontare questo percorso, sono convinta che l'*empathy* (per come oggi, comunemente, si definisce) sia, in realtà, la manifestazione "tangibile", cognitivamente parlando, di un altro fenomeno: il libero arbitrio.

La motivazione che mi sono data non è, al momento, sufficiente a dar conto di un'asserzione così impegnativa, poiché moltissime sono le domande a cui io stessa devo trovare una risposta per poterne dare una convincente a riguardo. Ma di una "verità", se mi si concede questa parola forte, sono certa: che l'espressione "empatia cognitiva" non denota, come detto, il fenomeno ai molti conosciuto proprio per quelle motivazioni che ho tentato di spiegare poco più su.

Se tale intuizione fosse vera, molte spiegazioni a diverse teorie dovrebbero essere rivisitate; ad esempio, sempre per tornare a Batson, le sue ricerche (alla luce di quanto detto) non si spiegherebbero più come preoccupazione empatica alla base dell'altruismo, ma, probabilmente, che i comportamenti pro-sociali *non deriverebbero* dall' "empatia", ma da una

scelta consapevole del soggetto di intervenire a favore dell'altro. Il libero arbitrio, cioè, spiegherebbe quell'ampia gamma di comportamenti pertinenti la *già* empatia.

Anche sul *free will*, però, il discorso non mi sembra molto limpido.

A proposito, ad esempio, della causalità indeterministica⁴⁵³ non mi sembra molto chiaro.

Si ricorderà che con tale concetto i libertari vogliono dar conto dell'origine di ogni nuova causazione *indeterminata*, affinché a quell'agente e non ad altri ne sia riconosciuta la genesi; in questo senso l'utilizzo del concetto di *primo motore immobile*, da parte sia dei compatibilisti sia dei libertari, pone un dubbio, poiché in entrambe le tesi il ricorso al concetto aristotelico diviene quasi un punto centrale per l'eventuale risoluzione di problemi cui entrambe incorrono.

Nel primo scenario, ad esempio, l'aspetto da enfatizzare sarebbe – stando alla spiegazione fornita dai compatibilisti – la capacità che l'agente abbia nel produrre/generare nuove e differenti catene causali con le successive conseguenze già evidenziate; nel secondo, invece, l'aspetto da evidenziare sarebbe non l'origine delle catene causali, ma l'intenzionalità dell'agente stesso dalla quale, a ritroso, si potrebbe dedurre l'appartenenza di quell'azione a quell'agente e non ad un altro con le modalità ed i particolarismi già incontrati.

L'analisi, vale a dire, si sposterebbe sugli *effetti* dell'eventuale libero arbitrio nell'individuo e non sulla sua stessa natura, *punctum dolens* dell'intero dibattito, sia nel primo sia nel secondo caso; nel primo perché, a causa della causazione transeunte/immanente l'agente sarà considerato come un primo motore immobile che, di conseguenza, diviene l'effetto e non l'origine. Nel secondo caso, come detto, si parte dalle azioni compiute per risalire alle intenzioni, che divengono effetti di cause da ricercare; ancora una volta, però, si incorre in una grave difficoltà: se ci si trova in un contesto a-causale che natura dare alle intenzioni?

Sembra, cioè, che la coniugazione in un modo o in un altro del concetto di *primo motore immobile* equivalga ad un *escamotage* per ovviare al problema dolente dell'una e dell'altra tesi: come dar conto del libero arbitrio dell'agente in un mondo necessitato, la prima; come salvaguardarlo da una deriva scettica, nel secondo.

Non è un caso che i contrari alla naturalizzazione umana si facciano portavoce di posizioni radicali o, alle volte, di posizioni che inducano il lettore ad una rivisitazione sostanziale delle proprie posizioni per paura di cadere in un nichilismo senza ritorno; è il caso, come visto dello stesso Libet, il quale ricorre alla “volontà di veto” per salvaguardare il concetto di libero arbitrio laddove i suoi dati, se interpretati in un certo modo, avrebbero evidenziato la sua totale illusorietà o, peggio, la sua non esistenza. È su questa scia, allora, che frasi tipo “non si

⁴⁵³ Si rimanda alla lettura del capitolo V.

possono compiere scelte diverse rispetto a quelle dettate dal proprio cervello” vengono usate come titoli allarmistici sui principali mezzi di diffusione “culturale”, o come spiegazione al fatto che l’uomo *tout court* o non sia libero o non responsabile o chissà quale altra manchevolezza possieda.

Contro tali tendenze e contro tali posizioni vorrei soffermarmi solo un attimo per cercare di ragionare insieme.

Sostenere che “non si possano compiere scelte diverse rispetto a quelle dettate dal proprio cervello” è una posizione mal posta.

Significherebbe:

1. considerare il cervello un’entità distinta dalla propria volontà; si cadrebbe, come è evidente, nel dualismo cartesiano e del modo in cui dar conto della natura delle due *res* che, se ricordo bene, fu un problema su cui anche lo stesso Cartesio traballò, fornendo, infine, la “soluzione” della ghiandola pineale – a tutti ben nota.

A di là delle difficoltà oggettive, coloro i quali appoggiassero una posizione radicale in tal senso dimostrerebbero, però, di non considerare l’esistenza di molteplici teorie esistenti nel panorama culturale odierno che spiegano il mentale in termini o di sopravvenienza o di epifenomeno il rapporto fra la mente ed il cervello. Quanto meno, i sostenitori delle due *res* sarebbero chiamati a fornire una o più spiegazioni plausibili, ragionevoli e scientifiche della loro natura, della loro inter-relazione e simili.

2. Sostenere che non si possa compiere un’azione diversa rispetto a quella dettata dal proprio cervello pone nelle condizioni di pensare che in un qualsiasi soggetto esistano due volontà: una del *brain* ed un’altra del *mind*.

A meno che non ci si trovi di fronte a specifiche patologie di ordine psichiatrico o psicologico, nella consuetudine ordinaria dubito che qualsiasi agente normodotato consideri se stesso come scisso in due ed espressioni come quella sopra presentata non hanno altro scopo che alimentare falsi dubbi e false perplessità, sia sulla natura umana sia su quella relativa al *free will*.

Non ci si può, infatti, alienare da sé, consapevolmente, pensando che il supporto materiale di noi stessi, quindi, noi stessi, sia racchiuso all’interno di una “prigione causale” (il cervello) dalla quale non si possa uscire; desiderare ciò che si vuole equivale, ad avviso della scrivente, a desiderare ciò che il nostro cervello/”il nostro Io” desidera.

Mi rendo conto che è difficilissimo, forse quasi impossibile, pensare di poter “risolvere” il problema del libero arbitrio attraverso la constatazione di una bio-chimica del *brain*, che, per tutte le ragioni già incontrate nel percorso fin qui condotto, possa fornire un aiuto o un

supporto a qualsivoglia teoria sul *free will*, poiché sono incline a pensare che, indipendentemente dalla Scienza *tout court*, il problema resta e resterà prettamente etico. Il ruolo del *filosofo in poltrona*, vale a dire, è ben lontano dall'aver esaurito il suo compito; ragion per cui mi trovo da dissentire *in toto* su qualsiasi metodo o approccio alternativo al problema *libero arbitrio* si fornisca: mi riferisco, infatti, a quello di Eddy Nahmias.

Il ricorso all'utilizzo della *folk psychology* nei modi e nelle maniere che si è già avuto modo di delucidare⁴⁵⁴ è ciò che di più lontano esista da uno studio *scientifico* del problema.

Utilizzare la *doxa* per approcciarsi a problemi riguardanti la *noesis* significa annullare l'intera *Cultura* che verrà; a quel punto le medesime Istituzioni dedite alla formazione/alfabetizzazione sarebbero inutili senza contare che si cadrebbe in ciò che Nahmias stesso ha criticato: nell'errore delle Scienze di travalicare il proprio campo di indagine, commettendo errori categoriali e concettuali di cui la stessa non si sarebbe resa conto se non vi fosse stato chi (in questo caso, Nahmias stesso), ragionando con le categorie del pensiero e non con meri dati o meri "saperi comuni", non lo avesse evidenziato.

È una contraddizione in termini e in fatti porre in un unico approccio l'unione, la fusione di una scienza, la *Filosofia*, con ciò da cui essa si è sempre allontanata, l'opinione (la *doxa*, appunto); si torna, allora, all'ovvietà ed all'insieme di ovvietà di cui Sellars, prima, e Lewis, dopo, hanno ipotizzato con la *folk psychology*. In questo modo non solo la Filosofia verrebbe denaturalizzata, depauperata della sua stessa *ratio* (usando un termine forte, uso la parola *essenza*), ma diverrebbe alla stregua di quanti, lontani dalla scientificità del pensiero, si arrogerebbero il diritto di "entrare anch'essi" nel dibattito; la parola di un filosofo o quella di uno scienziato (tanto per rimanere nelle categorie già presentate), a quel punto, sarebbero uguali a quelle del senso comune.

Avrebbero lo stesso peso.

È, in effetti, ciò che sta accadendo oggi, in cui il rispetto delle competenze è divenuto una chimera e dove la negazione del riconoscimento dei ruoli sta comportando danni sociali, epistemici e scientifici che, probabilmente, faranno slittare l'umanità in un nuovo *oscuro Medioevo*.

Ecco ciò che significa, ad avviso della scrivente, il ricorso alla *folk psychology* che Nahmias *et alii* hanno intrapreso: una strutturazione piatta del "sapere" (che, a quel punto, non potrebbe più definirsi tale) in cui l'analfabeta quanto il dotto avrebbe l'eguale peso nell'istituzione di un ragionamento che sarebbe/sarà ben lontano dall'essere scientifico.

⁴⁵⁴ Si rimanda al capitolo VII.

Sarebbe la fine della filosofia stessa, rinnegando, completamente, i grandi che ci hanno permesso di giungere dove siamo. Un approccio siffatto è solo da condannare.

Dovendo trarre, dunque, le conclusioni dalle conclusioni si può sostenere che quanto fin qui esposto costituisce non “la fine”, ma la consapevolezza di un nuovo inizio con cui approcciarsi al *falso* problema empatico, da una parte, ed a quello del *free will*, dall'altra; sperando, in un futuro non troppo lontano, di licenziare alle stampe il *Volume II* del presente lavoro come “atto dovuto”, non solamente all'*amore per il sapere*, ma anche a quell'ipotetico lettore che ha avuto la pazienza di seguire, dall'inizio, il delinarsi di un percorso (fin qui) non irto di ostacoli.

Bibliografia

- ABBAGNANO NICOLA, *Dizionario di Filosofia*, Utet Libreria, Milano, 2006.
- AKAISHI REI, KIM UMEDA, K. SAKAI, *Autonomous mechanism of internal choice estimate underlies decision inertia*, in "Neuron", 81 (1), 2014, pp. 195-206.
- ANDERSON MICHAEL L. E DON PERLIS, *What puts the "meta" in metacognition?*, in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009, pp. 138-139.
- ANDERSON MICHAEL E TIM OATES, *A Review of Recent Research in Metareasoning and Metalearning*, in "AI Magazine", Vol. 28, n°1, Spring, 2007, pp. 7-16.
Articolo reperibile on-line al seguente indirizzo:
<https://www.aaai.org/ojs/index.php/aimagazine/article/view/2025>
- AUNE BRUCE, *Hypotheticals and Can: Another Look*, Vol. 27, n. 6, Oxford University Press, Oxford, 1968, pp. 191-195.
- BARGH JONATHAN, *Free will is un-natural* in JOHN BAER, JAMES KAUFMANN AND ROY BAUMEISTER (A CURA DI), *Are we free? Psychology and Free Will*, in "International Journal of Philosophical Studies", 20 (5), pp. 725-734.
- BARON-COHEN SIMON, ALAN LESLIE ED UTA FRITH, *Does the Autistic Child have a Theory of Mind*, in "Cognition", n. 21, 1985, pp. 37-46.
- BARON-COHEN SIMON, *La scienza del male. L'empatia e le origini della crudeltà*, Raffaello Cortina Editore, Milano, 2012.
- BARON-COHEN SIMON, *Mindblindness: An Essay on Autism and Theory on Mind*, MIT Press, Massachusetts, 1995.
- BARTSCH KEIN E HENRY M. WELLMAN, *Children Talk about the Mind*, Oxford University Press, Oxford, 1995.
- BATSON DANIEL ET ALII, *Is empathic emotion a source of altruistic motivation?*, in "Journal of Personality and Social Psychology", Vol. 40, n. 2, 1981, p. 291.
- BATSON DANIEL, *Altruism in humans*, Oxford University Press, New York, 2011.
- BATSON DANIEL, LAURA L. SHAW, *Evidence for altruism: toward a pluralism of Prosocial Motives*, in "Psychological Inquiry", Vol. 2, n. 2, 1991.
- BATSON DANIEL, *The altruism question: toward a social psychological answer*, Lawrence Erlbaum, Hillsdale, 1991.

- BATSON DANIEL, *The empathy-Altruism Hypothesis*, Oxford Handbook Online, Oxford, 2014.
- BATSON DANIEL, *The Naked Emperor: Seeking a more plausible genetic basis for psychological altruism*, in “Economics and Philosophy”, 26, 2010, pp. 149–164.
- BATSON DANIEL, *These things called empathy: Eight related but distinct phenomena*, in JEAN DECETY E WILLIAM J. ICKES (A CURA DI), *The social neuroscience of empathy*, MIT Press, Massachusetts, 2009.
- BATSON DANIEL, *Two forms of perspective taking: Imagining how another feels and imagining how you would feel*, in KEITH D. MARKMAN, WILLIAM M. KLEIN E JULIE A. SUHR (A CURA DI), *Handbook of imagination and mental simulation*, Psychology Press, New York, 2009.
- BATHYANY ALEXANDER, *Mental causation and free will after Libet and Soon: reclaiming conscious agency*, in ALEXANDER BATHYANY, AVSHALOM C. ELITZUR (A CURA DI), “Irreducibly Conscious. Selected Papers on Consciousness”, Winter, Heidelberg, 2009.
- BAUMEISTER ROY F., *Understanding Free Will and Consciousness on the Basis of Current Research Findings in Psychology*, in ROY F. BAUMEISTER, ALFRED R. MELE E KATHLEEN D. VOHS (A CURA DI), *Free will and consciousness: How might they work?*, Oxford University Press, New York, 2010.
- BENGSON JESSE J., TODD A. KELLEY, XIAOKE ZHANG, JANE-LING WANG, GEORGE R. MANGUN, *Spontaneous neural fluctuations predict decisions to attend*, in “J. Cogn. Neurosci”, 26, 204, pp. 2578-2584.
- BERING JESSE, *The folk psychology of soul in Behavioral and Brain Sciences*, vol. 29 (5), 2006, pp. 453-462.
- BLASS ELLIOTT M., JUDITH R. GANCHROW E JACOB E. STEINER, *Classical conditioning in newborn humans 2-48 hours of age*, in “Infant Behavior and Development”, 7, 1984, pp. 223-235.
- BLOOM PETER, *Descarte’s Baby: How the Science of Child Development Explains What Makes Us Human*, Basic Books, New York, 2004.
- BODE STEFAN, CARSTEN MURAWSKI, CHUN SION-SOON, PHILIPP BODE, JUTTA STAHL, PHILIP L. SMITH, *Demystifying ‘free will’: The role of contextual information and evidence accumulation for predictive brain activity*, in “Neuroscience and Biobehavioral Review”, 47, 2014, pp. 636-645.
- BONIOLO GIOVANNI, GABRIELE DE ANNA E UMBERTO VINCENTI (a cura di), *Individuo e persona. Tre saggi su chi siamo*, Bompiani, Milano, 2007.

- BRENTANO FRANZ, *La psicologia da un punto di vista empirico*, Biblioteca Universale Laterza, Roma-Bari, 1997.
- BUCKNER CAMERON, ADAM SHRIVER, STEPHEN CROWLEY E COLIN ALLEND, *How “weak” mindreaders inherited the earth*, in “Behavioral and Brain Sciences”, Vol. 32, Issue 2, 2009.
- BYRNE RICHARD W. E ANDREW WHITEN, *Machiavellian Intelligence*, Oxford University Press, Oxford, 1997.
- CALISI ASTRO, *Oltre gli orizzonti del conosciuto. La sfida della mente alla scienza del XXI secolo*, Editrice Uni Service, Trento, 2011.
- CAPIELLO IDA, *La morale della simpatia di David Hume*, Liguori Editore, Napoli, 1983.
- CARTESIO RENATO, *Meditazioni Metafisiche*, Bompiani, Firenze, 2001.
- CARRUTHERS PETER, *How we know our minds: The relationship between mindreading and metacognition*, in “Behavioral and Brain Sciences”, 32, 2009, pp. 121-138.
- CARRUTHERS PETER, *Mindreading underlies metacognition*, in “Behavioral and Brain Sciences”, Vol. 32, Issue 2, 2009.
- CATANIA CHARLES, *Cognitive science at fifty*, in “Behavioral and Brain Sciences”, Vol. 32, Issue 2, 2009.
- CHISHOLM ROBERT, *Human Freedom and self* in GARY WATSON (A CURA DI), *Free Will*, Oxford University Press, USA, 2003.
- COUCHMAN JUSTIN J., MARIANA V. C. COUNTINHO, MICHAEL J. BERAN, J. DAVID SMITH, *Metacognition is prior* in “Behavioral and Brain Sciences”, Vol. 32, Issue 2, 2009.
- CRANE TIM, *Elements of Mind*, Oxford University Press, Oxford, 2001.
- CRANE TIM, *Fenomeni mentali. Un'introduzione alla filosofia della mente*, Cortina Raffaello, Milano, 2003.
- CRAVER CARL, *Explaining the Brain: Mechanism and the Mosaic Unity of Neuroscience*, Oxford University Press, New York, 2007.
- CUNNINGHAM SUZANNE, *What is a mind? An introduction to the Philosophy of Mind*, Hackett Publishing Company, Inc., Cambridge, 2000.
- CURRIE GREGORY E IAN RAVENSCROFT, *Recreative Minds: Imagination in Philosophy and Psychology*, Oxford University Press, Oxford, 2002.

- DAMASIO ANTONIO, *Emozione e Coscienza*, Biblioteca Scientifica Adelphi, Milano, 2007.
- DAMASIO ANTONIO, *L'errore di Cartesio. Emozione, ragione e cervello umano*, Adelphi Edizioni, Milano, 1995.
- DARWALL STEPHEN, *Empathy, Sympathy, Care* in "Philosophical Studies", 89, 1998, pp. 261-282.
- DAVIES MARTIN E TONY STONE, *Mental Simulation, tacit theory and the threat of collapse*, in "Philosophical Topics", 29, nn1/2, 2001, pp. 127-173.
- DAWKINS RICHARD, *Il gene egoista. La parte immortale di ogni essere vivente*, Edizioni Mondadori, Milano, 1995.
- DE CARO MARIO, *Il libero arbitrio, Una introduzione*, Editori Laterza, Milano, 2011.
- DE CARO MARIO, *Libero arbitrio e neuroscienze*, in ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Neuroetica*, Il Mulino, Bologna, 2011.
- DE MONTICELLI ROBERTA, *Che cos'è una scelta?*, in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, Codice Edizioni, Torino, 2019.
- DE ROSA MARIA ROSARIA, *Theodor Lipps. Estetica e critica delle arti*, Guida Editori, Napoli, 1990.
- DE VIGNEMONT FRÉDÉRIQUE, *Drawing the Boundary Between Low-Level and High-Level Mindreading*, in "Philosophical Studies", 144(3), 2009, pp. 457-466.
- DENNETT DANIEL, *Coscienza. Che cosa è*, Editore Laterza, Roma-Bari, 2009.
- DENNETT DANIEL, *L'Atteggiamento Intenzionale*, Il Mulino, Bologna, 1993.
- DENNETT DANIEL, *L'evoluzione della libertà*, Raffaello Cortina Editore, Milano, 2004.
- DENNETT DANIEL, *The Intentional Stance*, MIT Press, Massachusetts, 1987.
- DENNETT DANIEL E BRYCE HUEBNER, *Banishing 'I' and 'we' from accounts of metacognition*, in "Behavioral and Brain Sciences", 32, 2009, pp. 148-149.
- DI FRANCESCO MICHELE, *Introduzione alla filosofia della mente*, Carocci Editore, Roma, 2002.
- DONALD MERLIN, *Consciousness and the Freedom of Act*, in ROY F. BAUMEISTER, ALFRED R. MELE & KATHLEEN D. VOHS (A CURA DI), *Free Will and Consciousness: How Might They Work?*, Oxford University Press, Oxford, 2010.

- ECKARDT BARBARA VON, (1997) *The Empirical Naivete of the Current Philosophical Conception of Folk Psychology*, in MARTIN CARRIER E PETER MACHAMER (A CURA DI), "Mindscapes", University of Pittsburgh Press). Pittsburgh, 1997.
- EKMAN PAUL, WALLACE V. FRIESEN, MAUREEN O'SULLIVAN ED ANTHONY CHAN, *Universals and cultural differences in the judgments of facial expressions of emotion*, in "Journal of Personality and Social Psychology", 53, 1987, pp. 712-717.
- EVANS GARETH, *Semantic Theory and Tacit Knowledge*, in SAM HOLTZMAN E CHRISTOPHER M.. LEICH (A CURA DI), "Wittgenstein: To Follow a Rule", Routledge & Kegan Paul, Londra, 1981.
- EVANS JONATHAN ST. B. T., *Introspection, confabulation and dual-process theory*, in "Behavioral and Brian Sciences", Vol. 32, Issue 2, 2009.
- EVANS JONATHAN E DAVID OVER, *Rationality and Reasoning*, Psychology Press Publishers, East Sussex (UK), 1996.
- FARAH MARTHA J. E ANDREA HEBERLEIN, *Personhood and Neuroscience: Naturalizing or Nihilating?*, in "American Journal of Bioethics, Neuroscience", Vol. 7, 2007, pp. 37-48.
- FIALA BRIAN E SHAUN NICHOLS, *Confabulation, confidence and introspection* in "Behavioral and Brian Sciences", Vol. 32, Issue 2, 2009.
- FISHER JUSTIN C., *Does Simulation Theory Really Involve Simulation?*, in "Philosophical Psychology", 19(4), 2006, pp. 417-432.
- FISHER JOHN, *My Way*, Oxford University Press, Oxford, 2006.
- FRANKFURT HARRY, *Alternate Possibilities and Moral Responsibility*, in "Journal of Philosophy", 66, 1969, pp. 828-839.
- FRANKFURT HARRY, *Freedom of the Will and The Concept of Person* in "The Journal of Philosophy", Vol. 68, 1971, pp. 5-20.
- FRANKFURT HARRY, *Possibilità alternative e libertà morale*, in M. DE CARO (A CURA DI), *La logica della libertà*, Meltemi Editore, Bari, 2002.
- FRANKISH KEITH, *How we know our conscious minds: Introspective access to conscious thoughts*, in "Behavioral and Brian Science", Vol. 32, Issue 2, 2009.
- FRANKISH KEITH, *Mind and Supermind*, Cambridge University Press, Cambridge, 2007;
- FISHER JOHN E MARK RAVIZZA, *Responsability and Control: A Theory of Moral Responsibility*, Cambridge University Press, Cambridge, 1998.

- FRIED ITZHAK, ROY MUKAMEL, GABRIEL KREIMEN, *Internally generated preactivation of single neurons in human medial frontal cortex predicts volition*, “Neuron” n. 69, 2011, pp. 548-562.
- FRIEDMAN ORI E ADAM R. PETRASHEK, *Non interpretative metacognition for true beliefs*, in “Behavioral and Brain Sciences”, Vol. 32, Issue 2, 2009, pp. 146-147.
- GALLAGHER SHAUN, *Simulation Trouble*, in “Social Neuroscience”, 2(3–4), 2007.
- GALLESE VITTORIO, *Before and Below ‘Theory of Mind’: Embodied Simulation and the Neural Correlates of Social Cognition*, in “Phil. Trans R. Soc. B.”, 362, pp. 659-669.
- GAZZANIGA MICHAEL, *Chi comanda? Scienza, mente e libero arbitrio*, Codice Edizioni, Milano, 2013.
- GAZZANIGA MICHAEL, *The Split Brain in Man*, in “Scientific American Inc”, 1967, pp. 24-29.
Articolo reperibile al link:
<https://pdfs.semanticscholar.org/d8dd/bd1721eb8b98fcf5119b5b39fcb91044a8d2.pdf>
- GINET CARL, *On Action*, Cambridge University Press, Cambridge, 1990.
- GINET CARL, *Reasons Explanation of Action: An Incompatibilist Account* in “Philosophical Perspective”, Vol. 3, 1989.
- GOLDMAN ALVIN, *The psychology of folk psychology*, in “Behavioral and Brain Sciences”, Vol. 16, Issue 1, 1993, pp. 15-28.
- GOLDMAN ALVIN, *Replies to Perner and Brandl, Saxe, Vignemont, and Carruthers*, in “Philosophical Studies”. 144(3): 2009, pp. 477–491.
- GOLDMAN ALVIN, *Simulation Mind. The Philosophy, Psychology and Neuroscience of Mindreading*, Oxford University Press, Oxford, 2006.
- GOPNIK ALISON E ANDREW N. MELTZOFF, *Words, Thoughts and Theories*, Mit Press, Massachusetts, 1997.
- GOPNIK ALISON ED HENRY WELLMAN, *Why the Child's Theory of Mind Really Is a Theory*, in “Mind and Language”, n. 7, 1992, pp. 145–71.
- GORDON ROBERT M., *‘Radical’ Simulationism*, in PETER CARRUTHERS E PETER K. SMITH (A CURA DI), “Theories of Theories of Mind”, Cambridge University Press, Cambridge, 1996.

- GREEN JOSHUA, *The secret joke of Kant's soul* in WALTER SINNOTT-ARMSTRONG (A CURA DI), *Moral Psychology. The neuroscience of Morality: Emotion, Disease and Development*, MIT Press, Cambridge, 2007, pp. 35-79;
- GREEN JOSHUA E JONATHAN COHEN, *For the law, neuroscience changes nothing and everything* in *Philosophical Transactions of the Royal Society of London B*, 359, 2007, pp. 1775-1778.
- HAGGARD PATRICK, *Human volition: towards a neuroscience will*, in "Nat. Rev. Neurosci", 9, 2008, pp. 934-946.
- HAGGARD PATRICK E MARTIN EIMER, *On the relation between brain potentials and the awareness of voluntary movements*, in "Exp. Brain Res." n. 126, 1999, pp. 128-133.
- HAIDT JONATHAN, *The emotional dog and its rational tail: a social intuitionist approach to moral judgment*, in "Psychology Review", 108, 2001, pp. 814-834.
- HARLOW JOHN, *Recovery from the passage of an iron through the head*, in "Publications of the Massachusetts Medical Society", 2, 1868, pp. 327-347.
- HARRIS SAM, *Free will*, New York University Press, New York, 2012.
- HAYNES DAVID, *Posso prevedere quel che farai*, in ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Neuroetica*, Il Mulino, Bologna, 2011.
- HAYNES JHON DYLAN, *Case closed for free will?* In *Science NOW Daily News*, aprile 2008.
- HEAL JANE, *Simulation vs Theory-Theory: What is at Issue?*, in CHRISTOPHER PEACOCKE (A CURA DI), "Objectivity, Simulation, and the Unity of Consciousness: Current Issues in the Philosophy of Mind", Oxford University Press, Oxford, 1994, pp. 129-144. Articolo reperibile on-line al seguente indirizzo:
<https://www.thebritishacademy.ac.uk/pubs/proc/volumes/pba83.html>
- HEGEL GEORGE WILHELM FRIEDRICH, *Fenomenologia dello Spirito*, Bompiani, Milano, 2000.
- HOFFMAN MARTIN, *Empatia e sviluppo morale*, Il Mulino, Bologna, 2008.
- HONDERICH TIM, *Sei davvero libero? Il problema del determinismo*, Il Saggiatore, Milano, 1996.
- HUME DAVID, *Ricerca sui principi della morale*, Editori Laterza, Roma-Bari, 2018.
- HUME DAVID, *Trattato sulla Natura Umana*, Editori Laterza, Roma- Bari, 2004.
- HURLBURT RUSSELL E SARAH AKHTER, *Consciousness and Cognition*, in "Elsevier Journa", 17, 2008, pp. 1364-1374.

- HURLEY SUSAN, *Understanding Simulation*, in "Philosophy and Phenomenological Research", 77(3), 2008, pp. 755-774.
- INWAGEN PETER VAN, *An essay on free will*, Oxford University Press, USA, 1986.
- JEFFREY DAVID, *Empathy, sympathy and compassion in healthcare: is there a problem? Is there a difference? Does it matter?* In "Journal of Royal Society of Medicine", Vol. 109 (12), pp. 446-452.
- KAHNEMAN DANIEL AND AMOS TVERSKY, *On the Reality of Cognitive Illusions*, in "Psychological Review", 103, 1996, pp. 582-591.
- KANE ROBERT, *Free Will: new direction for an ancient problem* in "Journal of Philosophy", 96, pp. 222-248.
- KANE ROBERT, *The Significance of Free Will*, Oxford University Press, Oxford (NY), 1996.
- KANT IMMANUEL, *Fondazione metafisica dei costumi*, UTET, Torino, 1995.
- KHUN THOMAS, *La struttura delle rivoluzioni scientifiche*, Piccola Biblioteca Einaudi, Torino, 2009.
- KNOBE JOSHUA, *Esperimenti sul pensiero*, articolo reperibile on line al seguente link:
<http://download.kataweb.it/mediaweb/pdf/espresso/scienze/2012/01/31/115530199-3cebd776-d977-417c-bf9c-dc7a1a1ee84e.pdf>
- KNOBE JOSHUA E SHAUN NICHOLS, *Experimental Philosophy. Volume 2*, Oxford University Press, Oxford, 2014.
- KORIAT ASHER, HILIT MA'AYAN E RAVIT NUSSINSON, *The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior*, in "Journal of Experimental Psychology: General", 135(1), 2006, pp. 36-69.
- KREBS DANIEL L., *Empathy and Altruism*, in "Journal of personality and social psychology", 1975, XXXII, pp. 1134-1146.
- LAIRD JASON. D., JAY J. WAGENER, MICHAEL. HALAL E MICHELLE SZEDGA, *Remembering what you feel: Effects of emotion and memory*, in "Journal of Personality and Social Psychology", 42, pp. 646-675.
- LAVAZZA ANDREA, *Neuroscienze e filosofia morale*, in "Rivista di Filosofia", Vol. XCVIII, n. 3, dicembre 2007, pp. 327-358.
- LAVAZZA ANDREA, SARTORI GIUSEPPE (A CURA DI), *Neuroetica*, Il Mulino, Bologna, 2011.

- LEE CRONBACH J. E PAUL E. MEEHL, *Construct validity in psychological tests* in “Psychological Bulletin”, n. 52, 1955, pp. 281-387.
- LESLIE ALAN M., *How to acquire a representational theory of mind*, in DAN SPERBER (A CURA DI), “Metarepresentations: A Multidisciplinary Perspective”, Oxford university Press, Oxford, 2000, pp. 197-224.
- LEVY NEIL, *Neuroetica. Le basi neurologiche del senso morale*, Apogeo Editore, Milano, 2009.
- LEWIS DAVID, *Psychological and theoretical identification*, in “Australian Journal of Philosophy”, n. 50, 1972.
- LEWIS MICHAEL E JEANNETTE M. HAVILAND-JONES, *Handbook of Emotions*, Guilford Press, Ney York, 2018.
- LIPPS THEODOR, *Fonti della conoscenza. Empatia*. Articolo reperito dalla Biblioteca di Ateneo dell’Università di Bologna, *Alma Mater Studiorum*.
- LISHNER DAVID A., DANIEL C. BATSON, ELISABETH HUSS, *Tenderness and Sympathy: Distinct mpathic Emotions Elicited by Different Forms of Need*, in “Personality and Social Psychology Bulletin”, 37(5), pp. 614-625.
- LUCAS JASON R., *The Freedom of the Will*, Clarendon Press, Oxford, 1970.
- MADONNA FEDERICA, *Il problema aperto della mente in relazione alle neuroscienze. Che cosa ne direbbe, oggi, Cartesio?*, Aracne Editore, Roma, 2018.
- MADONNA FEDERICA, *Considerations about cognitive empathy*, in “American Journal of Humanities and Social Science Research”, Vol. 3, Issue 4, pp. 34-40.
- MADONNA FEDERICA, *Mentalizing: the cognitive aspect of empathy* in “Journal of Advance Research in Social Science & Humanities”, Vol. 5, Issue 2, 2019, pp. 69-78. Articolo reperibile on-line al seguente indirizzo:
<https://jiaats.com/ojs31/index.php/ssh>
- MAIBOM HEIDI L., *Feeling for Others: Empathy, Sympathy and Morality*, in Special Issue “Other Minds and Moral Agency”, “Inquiry”, 52, 2009, pp. 483-499.
- MELE ALFRED, *Another scientific threat to free will?* in “The Monist”, 95, 2012, pp. 422-440.
- MELE ALFRED, *Free Will and Luck: Replies to My Critics* in “Philosophical Explorations”, n. 10, 2007, pp. 195-210.

- MONROE ANDREW E. E BERTRAM F. MALLE, *From uncased will to conscious choice: the need to study, not speculate about people's folk concept of free will* in *Review of Philosophy and Psychology*, 1(2), 2010, pp. 211-224;
- MONTAGUE READ, *Free will*, in "Current Biology", 18, 2008, pp. 584-585.
- MOORE GEORGE E., *Ethics*, Oxford University press, London, 1912; traduzione italiana di M.V. PREDAVAL MAGRINI (A CURA DI), *Etica*, Franco Angeli, Milano, 1982.
- MORRIS STEPHEN, *The impact of the Neuroscience on the Free Will Debate* in "Florida Philosophical Review", Vol. 9, Issue 2, 2009, pp. 56-78.
- MURPHY DOMINIC, *Varieties of self-explanation*, in "Behavioral and Brain Sciences", Vol. 32, Issue 2, 2009.
- NAHMIAS EDDY, *A naturalistic vision of Free Will*, in ELIZABETH O'NEILL, EDOUARD MACHERY (A CURA DI), *Current Controversies in Experimental Philosophy*, Routledge, New York, 2014.
- NAHMIAS EDDY, *Free Will as a Psychological Accomplishment*, in DAVID SCHMIDTZ E CARMEN PAVEL (A CURA DI), *Oxford Handbook on Freedom*, Oxford University Press, Oxford, 2016. Articolo reperibile on line al link:
<https://philarchive.org/archive/NAHFWA-4v1>.
- NAHMIAS EDDY, *Is Free Will an Illusion? Confronting Challenges from the Modern Mind Science*, in WALTER SINNOTT-ARMSTRONG (A CURA DI), *Moral Psychology: Free Will and Moral Responsibility*, MIT Press, Massachusetts, 2014.
- NAHMIAS EDDY, *Scientific Challenge to Free Will* in CONSTANTINE SANDIS E TIMOTHY O'CONNOR (A CURA DI), *A Companion to the Philosophy of Action*, Wiley-Blackwell, Hoboken, New Jersey, 2010.
- NAHMIAS EDDY, *When consciousness matters: a critical review of Daniel Wegner's The Illusion of Conscious Will*, in "Philosophical Psychology", 15, 2002, pp. 527-541.
- NAHMIAS EDDY, *Why "Willusionism" lead to "Bad Results": Comments in Baumeister, Crescioni and Alquist*, in "Neuroethics", 4, N. 1 , 2011, pp. 17-24.
- NAHMIAS EDDY, *Your Brain as the Source of Free Will Worth Wanting: Understanding Free Will in the Age of Neuroscience*, in GREGG CARUSO E OWEN FLANAGAN (A CURA DI), *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*, Oxford University Press, Oxford, 2016.
- NAHMIAS EDDY, JUSTIN D. COATES E TREVOR KVARAN, *Free Will, Moral Responsibility and Mechanism: Experiments on Folk Intuitions*, in "Midwest Studies in Philosophy", 31 (1), 2007, pp. 214-242.

- NAHMIAS EDDY, STEPHEN MORRIS E THOMAS NADELHOFFER, *The Phenomenology of Free Will* in “Journal of Consciousness Studies”, Vol. 11, n. 7-8, 2004. Pp. 162-179(18).
- NAHMIAS EDDY, STEPHEN MORRIS, THOMAS NADELHOFFER, JASON TURNER, *Is incompatibilism intuitive?* In “Philosophy and Phenomenological Research”, 72 (1), 2006, pp. 28-53.
- NAHMIAS EDDY E DYLAN MURRAY, *Experimental Philosophy on Free Will: An Error Theory for Incompatibilist Intuition*, in JESUS AGUILAR, ANDREI BUKAREFF E KEITH FRANKISH (A CURA DI), *New Waves in Philosophy of Action*, Palgrave-Macmillan, 2010.
- NAHMIAS EDDY E DYLAN MURRAY, *Explaining Away Incompatibilist Intuitions. Philosophy and Phenomenological Research*, Vol. 88, 2012, pp. 434-467.
- NAHMIAS EDDY, STEPHEN MORRIS, THOMAS NADELHOFFER AND JASON TURNER, *The Phenomenology of Free Will*, in “Journal of Consciousness Studies”, 1 (7-8), 2004, pp. 162-179.
- NAHMIAS EDDY, JASON SHEPARD E JANE REUTER, *It's OK if 'My Brain Made Me Do It': People's Intuitions about Free Will and Neuroscientific Prediction* in “Cognition”, 133(2), 2014, pp. 502-513.
- NANNINI SANDRO, *Naturalismo Cognitivo. Per una teoria materialistica della mente*, Quodlibet Studio, Macerata, 2007.
- NICHOLAS SHAUN E JOSHUA KNOBE, *Moral responsibility and determinism: The cognitive science of folk intuitions*, in “Nous”, 41(4), 2007, pp. 28-53.
- NIETZSCHE FRIEDRICH, *Così parlò Zarathustra*, Bompiani, Milano, 2010.
- NOZICK ROBERT, *Philosophical Explanation*, Clarendon Press, Oxford, 1981.
- PATERNOSTER ALFREDO, *Introduzione alla filosofia della mente*, Laterza, Roma-Bari, 2014.
- PEREBOOM DERK, *Lo scetticismo ottimistico su libertà e responsabilità* (traduzione di Andrea Lavazza) in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Quanto siamo responsabili? Filosofia, Neuroscienze e Società*, Codice Edizioni, Torino, 2013, pp. 127-155.
- PEREBOOM DERK, *Free will, Agency and meaning of Life*, Oxford University Press, Oxford, 2014.
- PEREBOOM DERK, *Living without free will (Cambridge Studies in Philosophy)*, Cambridge University Press, Cambridge, 2001.

- PIZARRO DAVID A., ERIK G. HELZER, *Stubborn Moralism and Freedom of the Will*, in ROY F. BAUMEISTER, ALFRED R. MELE E KATHLEEN D. VOHS (A CURA DI), *Free will and consciousness: How might they work?*, Oxford University Press, New York, 2010.
- POCKET SUSAN E SUZANNE PURDY, *Are voluntary movements initiated preconsciously? The relationship between readiness potential, urges and decisions*, in WALTER SINNOTT-ARMSTRONG E LYNN NADEL (A CURA DI), *Conscious Will and Responsibility: A Tribute to Benjamin Libet*, Oxford University Press, New York, 2010.
- PREMACK DANIEL E GUY WOODRUFF, *Does the chimpanzee have a theory of mind?*, in “Behavioral and Brain Sciences”, 4, 1978, pp. 515-629.
- PROUST JOËLLE, *Overlooking metacognitive experience*, in “Behavioral and Brain Sciences”, Vol. 32, Issue 2, 2009, p. 158.
- PUGLIESE ALICE, *La dimensione dell’intersoggettività. Fenomenologia dell’estraneo nella filosofia di Edmund Husserl*, Mimesis Editore, Milano, 2004.
- PUTNAM HILARY, *Mind, Language and Reality*, Philosophical Paper, Vol. 2, Cambridge University Press, Cambridge (Mass.), traduzione italiana di ROBERTO CORDESCHI, *Mente, Linguaggio e Realtà*, Adelphi, Milano, 1987.
- QUINE WILLARD VAN ORMAN, *Three Indeterminacies*, in R. BARRETT E R. GIBSON (A CURA DI), *Perspective on Quine*, Basil Blackwell, Oxford, 1990, pp. 1-16, cit. p. 4
- RAINONE ANTONIO, *La riscoperta dell’empatia. Attribuzioni intenzionali e comprensione nella filosofia analitica*, Bibliopolis, Napoli, 2005.
- RAINONE ANTONIO, *La riscoperta dell’empatia. Attribuzioni intenzionali e comprensione nella filosofia analitica*, Bibliopolis, Napoli, 2005.
- RAVENSCROFT IAN M., *Folk Psychology as a Theory* in STANDFORD ENCYCLOPEDIA OF PHILOSOPHY, consultabile al seguente indirizzo internet:
<https://stanford.library.sydney.edu.au/archives/spr2010/entries/folkpsych-theory/>
- RAWLS JOHN, *A theory of justice*, Harvard University Press, Harvard, 1971.
- REALE GIOVANNI, *Socrate. Alla scoperta della sapienza umana*, Rizzoli, Milano, 2001.
- RIGONI DAVIDE E MARCEL BRASS, *La libertà: da illusione a necessità*, in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, Codice Edizioni, Torino, 2019
- RIZZOLATTI GIACOMO E CORRADO SINIGAGLIA, *So quel che fai. Il cervello che agisce e i neuroni specchio*, Raffaello Cortina Editore, Milano, 2005.

- ROBBINS PHILIP, *Guilt by dissociation: Why mindreading may not be prior to metacognition after all*, in “Behavioral and Brain Sciences” Vol. 32, Issue 2, 2009.
- ROCHAT PHILIPPE, *Social-affective origins of mindreading and metacognition*, in “Behavioral and Brian Sciences”, Vol. 32, Issue 2, 2009.
- ROGERS CARL, *Client-centered therapy*, Houston-Mifflin, Boston, 1951.
- ROSEN JEFFREY, *The brain on the stand*, in “The New York Times”, 11 marzo 2007.
- ROSKIES ADINA, *Neuroscientific challenges to free will and responsibility*, in “Trends in Cognitive Sciences”, Vol. 10, n. 9, pp. 419-423.
- RUFO FABRIZIO, ANDREA BORRI, *Il cervello non è più quello di una volta*, in ANDREA CERRONI, FABRIZIO RUFO (A CURA DI), “Neuroetica tra etica e società”, Utet, Torino, 2013.
- SAVULESCU JULIAN E NICK BOSTROM (A CURA DI), *Human Enhancement*, Oxford University Press, NY, 2009.
- SCHELER MAX, *Il formalismo nell’etica e l’etica materiale dei valori*, Bompiani, Torino, 2013.
- SCHNEIDER LOGAN, ELISE HOUDAYER, OU BAI E MARK HALLETT, *What we think before a voluntary movement*, in “J. Cognitive Neuroscience” n. 25, 2013, pp. 822-829.
- SCIENCENOW, edizione del 12 giugno 2007.
- SEARLE JOHN, *Libertà e neurobiologia*, Mondadori Bruno, Milano, 2005.
- SEIFERT JOSEF, *La morte cerebrale non è più la morte di fatto. Argomentazioni filosofiche* in ROSANNA BARCARO E PAOLO BECCHI (A CURA DI), *Questioni mortali. L’attuale dibattito sulla morte cerebrale e il problema dei trapianti*, Edizioni Scientifiche Italiane, Napoli, 2004.
- SELIGMAN MARTIN, PETER RAILTON, ROY BAUMEISTER E CHANDRA SRIPADA, *Navigating into the future or driven by the past*, in “Perspectives on Psychological Science”, 8, 2013, pp. 119–141.
- SELLARS WILFRID, *Philosophy and the scientific image of man*, in ROBERT COLODNY (A CURA DI), *Science, perception and Reality*, Humanities Press, Ridgeview, 1963.
- SHAUN NICHOLS E STEPHEN P. STICH, *Mindreading. An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*, Oxford University Press, Oxford, 2003.

- SINGER PETER, *Ripensare la vita. La vecchia morale non serve più*, Il Saggiatore, Milano, 1994.
- SKINNER BORING, *The Operational Analysis of Psychological Terms*, in “Behavioral and Brain Sciences”, Vol. 7, Issue 4, 1984, pp. 547-553.
- SMITH ADAM, *Teoria dei sentimenti morali*, Bur Rizzoli, Milano, 2016.
- SOON CHUN SION, AH HE, STEPHAN BODE, JOHN-DYLAN HAYNES, *Predicting free choices for abstract intentions*, in “Proc. Natl. Acad. Sci. USA, 110 (15), 2013, pp. 5733-5734.
- SOON CHUN SIONG, MARCEL BRASS, HANS-JOCHEN HEINZE E DYLAN HAYES, *Unconscious determinants of free decisions in the human brain* in *Nature Neuroscience*, 11, 2008, pp. 543-545.
- SPENCER DAN E DEIRDRE WILSON, *Pragmatics, Modularity and Mindreading*, in “Mind and Language”, Wiley, 17 (1), 2002, pp. 3-33.
- STANFORD ENCYCLOPEDIA OF PHILOSOPHY, consultabile on -line al seguente indirizzo:
<https://stanford.library.sydney.edu.au/entries/folkpsych-simulation/> .
- STITCH STEPHEN E SHAUN NICHOLS, *Cognity penetrability, rationality and restricted simulation*, in “Mind and Language”, 12, nn3/4, 1997, pp. 297-326.
- STITCH STEPHEN E SHAUN NICHOLS, *Folk Psychology: simulation or tacit theory?*, in “Mind and Language”, n. 7, pp. 35-71;
- STONE DAVIES, *The mental simulation debate: a progress report*, in PETER CARRUTHERS E PETER K. SMITH (A CURA DI), “Theories of Theories of Mind”, Cambridge University Press, Cambridge, 1996.
- STRAWSON PETER F., *Libertà e risentimento*, in MARIO DE CARO (A CURA DI), *La logica della libertà*, Meltemi Editore, Bari, 2002.
- STUEBER KARSTEN, *L’empatia*, Il Mulino, Bologna, 2010.
- STUEBER KARSTEN, *Rediscovering Empathy. Agency, Folk Psychology and the Human Sciences*, The MIT Press, Massachusetts, 2010.
- TEMPIA FRANCESCA, *Decisioni libere*, in MARIO DE CARO, ANDREA LAVAZZA, GIUSEPPE SARTORI (A CURA DI), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, Codice Edizioni, Torino, 2019.
- WALLACE JAY, *Responsability and the Moral Sentiments*, Harvard University Press, Cambridge (MA), 1994.

- WEGNER DANIEL, *The Illusion of Conscious Will*, The MIT PRESS, Cambridge, 2002.
- WISPÈ LAUREN, *The distinction between sympathy and empathy*, in “Journal of personality and social psychology”, Vol. 50, n. 2, 1986, pp. 314-321.
- WITTGENSTEIN LUDWIG, *Tractatus logico-philosophicus*, Einaudi Editore, Torino, 1996.

