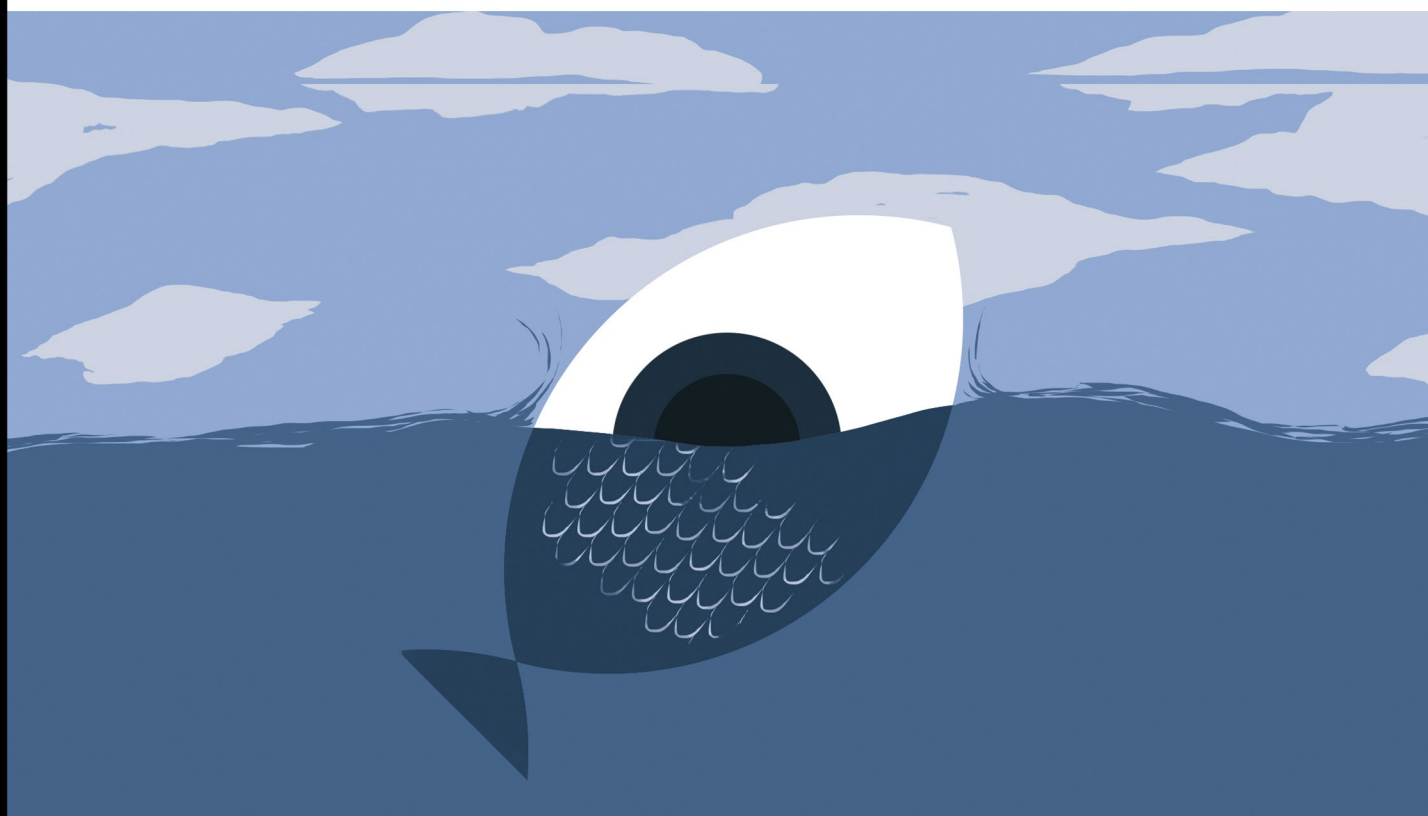


ANIMALI RAZIONALI

STUDI SUI CONFINI
E SULLE POSSIBILITÀ
DELLA RAZIONALITÀ

a cura di

Pierluigi Graziani
Giorgio Grimaldi
Massimo Sangoi



Isonomia Epistemologica

Isonomia – Epistemologica

Volume 8

ANIMALI RAZIONALI

STUDI SUI CONFINI E SULLE POSSIBILITÀ DELLA RAZIONALITÀ

Volume 1

Il realismo scientifico di Evandro Agazzi

Mario Alai (a cura di)

Volume 2

Complessità e riduzionismo

Vincenzo Fano, Enrico Giannetto, Giulia Giannini, Pierluigi Graziani (a cura di)

Volume 3

Oltre la fisica normale

Isabella Tassani (a cura di)

Volume 4

Mettere a fuoco il mondo

Elena Casetta, Valeria Giardino (a cura di)

Volume 5

Metaphor and Argumentation

Francesca Ervas, Massimo Sangoi (a cura di)

Volume 6

Forecasting the Future

Stefano Bordoni, Sara Matera (a cura di)

Volume 7

Teaching and Learning Mathematics

Laura Branchetti (a cura di)

Volume 8

Animali razionali

Pierluigi Graziani, Giorgio Grimaldi, Massimo Sangoi (a cura di)

ISONOMIA - Epistemologica Series Editor

Gino Tarozzi

gino.tarozzi@uniurb.it

ANIMALI RAZIONALI

STUDI SUI CONFINI E SULLE POSSIBILITÀ DELLA RAZIONALITÀ

A cura di

Pierluigi Graziani
Giorgio Grimaldi
Massimo Sangoi

© ISONOMIA – Epistemologica

ISSN 2037-4348

Direttore scientifico: Gino Tarozzi
Direttore editoriale: Pierluigi Graziani
Dipartimento di Scienze Pure e Applicate
P.za della Repubblica, 13 – 61029 Urbino (PU)

<http://isonomia.uniurb.it/>

Design by massimosangoi@gmail.com

Tutti i diritti sono riservati. Questa pubblicazione non può essere, neppure parzialmente, riprodotta, archiviata o trasmessa in qualsiasi forma o con qualsiasi mezzo, elettronico, meccanico, fotocopia, registrazione o altro, senza averne ottenuta l'autorizzazione scritta da parte dell'editore.

L'immagine di copertina è stata disegnata da Luca Di Battista.

Sommario

| | |
|--|-----|
| PIERLUIGI GRAZIANI, GIORGIO GRIMALDI, MASSIMO SANGOI <i>Introduzione</i> | 7 |
| CLAUDIA CASADIO, <i>Aspetti e limiti del ragionamento umano. Studi in psicologia e neuroscienze</i> | 17 |
| FRANCESCA ERVAS, <i>La “casa della ragione”. Razionalità e irrazionalità nei contesti conversazionali</i> | 47 |
| TIZIANA MIGLIORE, <i>La razionalità figurativa. Con alcune note su Philippe Ramette</i> | 69 |
| VALERIA GIARDINO, <i>I fondamenti cognitivi della matematica umana: dove cercarli?</i> | 85 |
| VINCENZO FANO, <i>Il dilemma del prigioniero iterato: un problema aperto</i> | 115 |
| DANIELE SANTORO, <i>Responsabilità e ragionamento controfattuale nel diritto</i> | 133 |
| SABRINA MORETTI, <i>L’agire razionale nelle spiegazioni sociologiche: dai “tipi ideali” ai modelli di simulazione</i> | 159 |
| STEFANO VASELLI, <i>Due economie della conoscenza a confronto (o della razionalità del realismo)</i> | 177 |
| ANTONIO PETAGINE, <i>Tra pensare e fare. Tommaso d’Aquino e Duns Scoto a proposito dell’irrazionalità pratica</i> | 203 |

| | |
|--|-----|
| ERCOLE ERCULEI, <i>Truth against Reason?</i> <i>Siger of Brabant and the Question of the Eternity of the World</i> | 229 |
| EMANUELA MACERA, « Vernünftige Selbstliebe ». <i>Universale e particolare nel giudizio politico kantiano</i> | 257 |
| GIORGIO GRIMALDI, <i>Mythos e logos. Alle origini del pensiero occidentale</i> | 291 |
| PATRIZIA PEDRINI, <i>Cosa causa l'irrazionalità motivata</i> | 309 |
| VERA TRIPODI, <i>Razionalità ed emotività.</i> <i>Il ruolo svolto dalle emozioni in epistemologia</i> | 321 |
| ADRIANO ANGELUCCI, MANUEL BASTIONI, PIERLUIGI GRAZIANI, MARIA GRAZIA ROSSI, <i>Alcune riflessioni sul fenomeno dell'Uncanny Valley</i> | 349 |

Introduzione

Pierluigi Graziani

Università degli Studi “G. d’Annunzio” di Chieti-Pescara
pierluigi.graziani@uniurb.it

Giorgio Grimaldi

Università degli Studi “G. d’Annunzio” di Chieti-Pescara
giorgio.grimaldi@yahoo.it

Massimo Sangoi

Università degli Studi di Urbino Carlo Bo
massimosangoi@gmail.com

Ogni pensiero, man mano che va costituendosi, traccia i propri confini. Se è autentico pensiero, e non rigida posizione in forma di dogma, ai suoi limiti vi è solo un solco visibile e attraversabile, mai degli steccati.

Per troppo tempo vi è stata una reciproca delegittimazione fra un pensiero che intende ispirarsi al moderno metodo scientifico e seguirne i passi e un pensiero che rifiuta di divenire solo “calcolante” e di applicare tale metodo a pressoché la totalità dei campi del sapere. Si è trattato di una battaglia a colpi di pregiudizi da cui nessuno è uscito vincitore, nonostante le grida di vittoria e i proclami entusiastici ora dell’una, ora dell’altra parte. La prima vittima di questa guerra persa in partenza, come si può immaginare, è stato il pensiero stesso, stratonato all’estremo da un lato verso sterili “astrazioni” scolastiche e dall’altro verso fascinosi espressioni che a volte si sono rivelate essere vere e proprie fumisterie concettuali. I due estremi, come spesso accade, hanno trovato un loro punto d’incontro: la mancanza di una profonda e rigorosa riflessione filosofica.

Da diversi anni si è allargato e rafforzato il fronte di coloro che non intendono barricarsi dietro le differenze di scuole di pensiero compiacenti verso se stesse e spesso del tutto autoreferenziali. È a questo fronte che appartiene il presente volume.

Che ci si disponga al dialogo non significa, tuttavia, che non vi siano negli interlocutori differenze. Al contrario, per rispetto dei dialoganti e in nome di un pensiero che mira a individuare buone ragioni invece di basarsi su cattive ideologie, le divergenze, come deve essere in ogni conversare franco, sono esposte in occasione di questo incontro sul confine. Ecco perché nelle pagine che seguono non vi è alcun *accordo* prestabilito, ma, laddove sussistano, le differenze sono presenti in tutta la loro portata, secondo tutta la forza di contrapposizione che esprimono. Seguendo William Blake vogliamo dire che *Opposition is true Friendship*, e aggiungere che tanto più siamo coinvolti autenticamente nella ricerca, quanto più questo implica confronti, contrasti, il sincero manifestarsi delle differenze e lo sviluppo critico del proprio pensiero.

Da questo spirito nasce il progetto *Animali Razionali. Studi sui confini e sulle possibilità della razionalità*, che intende analizzare la tematica della razionalità umana considerandola nelle sue diverse manifestazioni e metodiche di indagine.

Il primo saggio di **Claudia Casadio** offre esattamente una chiara e ampia introduzione a tali sfaccettature e metodiche considerandone i molteplici legami. Il contributo della Casadio analizza, ad esempio, i meccanismi che conducono i soggetti alla soluzione di problemi logici; i fattori che causano errori sistematici e *bias* e cosa ci dicono i *bias* sulla natura del ragionamento; gli aspetti del contenuto che condizionano l'abilità di ragionamento e cosa ci dicono sulla sua natura.

Al rapporto tra razionalità e sua manifestazione linguistica è dedicato il saggio di **Francesca Ervas** che, partendo dalle ricerche di Donald Davidson, si interroga sulla possibilità di *standard* di razionalità che ci permettano di stabilire il significato del messaggio comunicato dal parlante, e su quali sono i "requisiti minimi" di razionalità che potremmo condividere in un qualsiasi contesto conversazionale.

Partendo proprio da dove Davidson si è fermato, **Tiziana Migliore** si chiede perché, nell'offrire un'accezione di razionalità come "verità intersoggettiva", linguisticamente condivisa, Davidson scelga la procedura analogica, cioè la razionalità "figurativa". Su tale scorta la Migliore si interroga su quali sono i tratti del ragionamento figurativo rispetto ad altri tipi di ragionamento e cerca di rispondere a tale domanda avvalendosi degli spunti di un ciclo di fotografie di Philippe Ramette, le *Esplorazioni razionali*.

L'analisi della razionalità attraverso l'osservazione dei processi di rappresentazione simbolica continua con il saggio di **Valeria Giardino**. In questo articolo l'autrice discute dell'apporto delle scienze cognitive alla ricerca sui fondamenti cognitivi della matematica presentando brevemente alcuni importanti risultati sperimentali e considerandoli alla luce di tre livelli di competenze necessarie per fare aritmetica e geometria. In questa prospettiva, particolare importanza assumerà il livello in cui entrano in azione le capacità simboliche umane – che distinguono gli esseri umani da altre specie – e vengono creati e introdotti strumenti cognitivi. È a questo livello, infatti, che l'interazione tra natura e cultura diventa determinante per lo sviluppo della matematica come la conosciamo e di cui possiamo qui ritrovare i fondamenti cognitivi.

In continuità con la questione delle ricerche sui fondamenti cognitivi della matematica il volume presenta un'analisi di una delle più interessanti risorse matematiche messe in campo per analizzare le dinamiche delle interazioni umane: la Teoria dei Giochi. In modo particolare il saggio di **Vincenzo Fano** è diretto all'analisi della situazione nota come dilemma del prigioniero mettendo in luce una serie di aspetti filosofici, psicologici e matematici derivanti dal suo studio e offrendo una introduzione operativa a importanti tecniche di indagine dei ragionamenti tipiche della Teoria in questione.

Passando a un diverso dilemma, **Daniele Santoro** riflette su un importante problema nell'ambito della teoria dei controfattuali applicata al diritto. L'autore sostiene, infatti, che i modelli controfattuali conducono il teorico di fronte al seguente dilemma: o egli fornisce una teoria monistica della causalità lungo la falsariga del modello NESS, al costo però di non poter spiegare il ruolo della mera sorte, oppure salva la teoria, al costo di rivedere i giudizi sulla responsabilità degli agenti. In questo contesto Santoro mostra come il ruolo del ragionamento controfattuale nel diritto consista nel guidare il ragionamento nei giudizi di colpevolezza, in particolare nello stabilire ciò che l'agente avrebbe dovuto prendere in considerazione al momento dell'azione, e non esclusivamente nell'accertare le condizioni necessarie e sufficienti degli eventi.

I saggi di Fano e di Santoro hanno introdotto l'analisi della razionalità e delle scelte razionali in contesti giuridici, politici ed economici. Lo studio di **Sabrina Moretti** arricchisce tali analisi attraverso uno studio degli assunti sull'attore sociale propri di teorie come quella della scelta razionale e quella cognitiva della razionalità, mettendo in luce, proprio attraverso strumenti formali, il valore e i limiti di alcune metodiche di indagine di tale attore.

Il contributo di **Stefano Vaselli** analizza il tema della razionalità attraverso la domanda sul rapporto tra potere e conoscenza nell'attore sociale

come attore epistemico: che tipo di potere può rappresentare per un soggetto epistemico S entrare in possesso di una conoscenza che p , rispetto ad altri soggetti epistemici $S_1, S_2, S_3, \neq S$, o in anticipo su questi ultimi? L'analisi di Vaselli offre uno stimolo alla costruzione di un ponte tra epistemologia sociale ed epistemologia filosofica, un ponte edificabile sulla considerazione della conoscenza come risorsa economica. Di qui le domande a cui Vaselli cerca di rispondere nel testo: *perché* la conoscenza è una risorsa intrinsecamente economica? *come* ne è possibile una *distribuzione* ottimale? considerare l'oggetto della conoscenza da una prospettiva *realista* o, di contro, *comunitarista o costruttivista*, può rappresentare realmente un bivio dilemmatico tra due diverse e opposte *economie della conoscenza*?

Il tema della razionalità, quale emerge nei saggi finora descritti, è intimamente connesso a quello della volontà. Al legame tra questi due elementi e in particolare all'analisi del tema dell'irrazionalità pratica è rivolto il saggio di **Antonio Petagine**, che partendo da analisi storico-filosofiche, offre una stimolante indagine teorica della questione a partire dalla lezione di Tommaso d'Aquino e Duns Scoto: è davvero possibile considerare in modo cogente la volontà come una potenza intrinsecamente razionale se non si indica un qualche vincolo all'irrazionalità pratica?

L'intreccio fra riflessioni storico-filosofiche e teoretiche prosegue nei successivi tre saggi, attraverso i quali si può seguire un percorso che, dal pensiero medioevale, giunge al moderno e al contemporaneo per poi *tornare* all'antico.

Ercole Erculei, seguendo la complessa esperienza filosofica di Sigieri di Brabante (e della tradizione aristotelica), pone all'attenzione la difficoltà della polarità fede/ragione: entrambi i poli, lungi dal potersi liquidare a vicenda, proprio a vicenda si problematizzano, forse anche quando posti "in equilibrio".

Con **Emanuela Macera** siamo condotti a Kant, e a un Kant analizzato attraverso alcune sue opere celebri (su tutte la *Critica del giudizio*) per pensarne una possibile "*Critica della ragione politica*", mostrandone, quindi, tutto il portato politico a partire dalla riflessione teoretica.

Giorgio Grimaldi si concentra su Heidegger interprete di Parmenide, per giungere infine proprio a Parmenide, in vista e in funzione di una rilettura della radicalità di due elementi fondamentali da cui si diparte il pensiero occidentale: *mythos* e *logos*; entrambi sfuggono alla dicotomia (intesa come semplice e immediata) irrazionale/razionale.

Continuando con l'analisi del rapporto tra razionalità e irrazionalità, **Patrizia Pedrini** indaga nel suo saggio la tematica dell'irrazionalità motivata,

espressione che denomina quel fenomeno che si manifesta quando siamo sopravanzati da una passione e quindi agiamo, valutiamo situazioni e traiamo conclusioni mossi da ciò che proviamo, trascurando in parte o del tutto quello che la ragione ci consiglia. In particolare l'irrazionalità motivata viene indagata nella forma del fenomeno dell'autoinganno, di cui viene spiegata chiaramente la dinamica e illustrati i problemi che pone.

Vera Tripodi nel suo saggio continua l'analisi dei legami tra emotività e razionalità con l'obiettivo di mostrare come, alla luce di alcune obiezioni che l'epistemologia femminista ha sollevato, il nostro concetto di conoscenza e la dicotomia tra emotività e razionalità possano essere in parte riconsiderati fornendo una riabilitazione delle emozioni in campo epistemologico. In questa prospettiva la Tripodi mostra come alcuni studi sembrano smentire la classica tesi che le emozioni non hanno alcun peso o valore epistemico nel processo conoscitivo.

Infine, **Adriano Angelucci, Manuel Bastioni, Pierluigi Graziani e Maria Grazia Rossi** riflettono sul ruolo che alcune emozioni hanno all'interno del nostro rapporto con artefatti – materiali e digitali – umanoidi, nello specifico sulla rilevanza e natura del fenomeno noto come *uncanny valley* e come una sua adeguata spiegazione possa avere un ruolo importante per lo studio delle interazioni tra agenti sociali in società tecnologizzate.

Profilo degli autori

Adriano Angelucci è stato professore a contratto presso l'Università degli Studi di Urbino Carlo Bo e svolge attualmente attività di Post-Doc presso l'Università di Macao. Ha completato il suo Dottorato di Ricerca in Filosofia presso l'Università degli Studi di Verona nel 2012, in seguito a un soggiorno di due anni come *Visiting* alla Columbia University di New York. La sua ricerca si svolge principalmente in epistemologia e riguarda l'utilizzo di esperimenti mentali in filosofia analitica. Attualmente sta lavorando alla stesura di un volume dal titolo *Che cos'è un esperimento mentale*, di prossima pubblicazione per la casa editrice Carocci.

Manuel Bastioni, esperto di grafica 3D, ha ideato ed è stato direttore del progetto di modellazione umanoide MakeHuman. Nel 2004 ha ricevuto il prestigioso premio Suzanne per la grafica 3D open source. Collabora con numerosi progetti di ricerca nazionali ed internazionali sugli sviluppi e applicazioni della modellazione umanoide.

Claudia Casadio è professore ordinario di Logica e Filosofia della Scienza (M-FIL/02) e insegna presso l'Università "G. d'Annunzio" di Chieti-Pescara, nei corsi di laurea di Filosofia, Scienze dell'Educazione e Psicologia. Collabora con il Gruppo di Logica e Geometria della Cognizione dell'Università di Roma Tre, il Category Theory Group della McGill University di Montreal, i gruppi di logica e linguaggio di Utrecht OTS e Amsterdam ILLC. Profilo e pubblicazioni su Academia, Scopus e Google Scholar.

Ercole Erculei è assistente di ricerca presso la cattedra di Filosofia Medievale dell'Istituto di Filosofia alla *Rheinische Friedrich Wilhelms Universität* di Bonn (DE). Ha studiato filosofia a Chieti (con Giulio Lucchetta) e a Bonn (con Theo Kobusch). Nel 2011 ha vinto il premio come miglior studente straniero del DAAD (*Deutscher Akademischer Austauschdienst*). Nella primavera del 2013 ha vinto una borsa di studio presso la *Hanns Seidel Stiftung* di Monaco di Baviera (DE). Ha pubblicato in volumi collettanei e riviste specializzate in Italia e all'estero studi sulla tradizione patristica (e.g. Giovanni Crisostomo) e su autori medievali e rinascimentali (e.g. Meister Eckhart, Pico della Mirandola). Ha curato con Giorgio Grimaldi il volume *Politica, moralità, fortuna. Riflessioni storico-politiche sul Principe di Machiavelli* (Roma 2015). È membro della GANPh (*Gesellschaft für Antike Philosophie e.V.*) e della GMPR (*Gesellschaft für Philosophie des Mittelalters und der Renaissance*). Nel 2007 si è diplomato in violino presso l'Istituto Musicale "G. Braga" di Teramo.

Francesca Ervas è ricercatrice TD in Filosofia del Linguaggio presso l'Università di Cagliari. Ha conseguito il dottorato di ricerca in Filosofia e Teoria delle scienze umane presso l'Università Roma Tre ed è stata Post-Doc presso il Dipartimento di Linguistica – University College di Londra, l'Institut Jean Nicod, Ecole Normale Supérieure di Parigi e l'Università di Cagliari. I suoi interessi di ricerca riguardano, in particolare, la teoria della traduzione e la pragmatica sperimentale.

Vincenzo Fano è professore di Logica e filosofia della scienza nell'Università di Urbino Carlo Bo, autore di circa centocinquanta pubblicazioni di filosofia della fisica, della psicologia, storia della scienza, epistemologia e metafisica. Il suo ultimo libro è *I paradossi di Zenone*, Carocci, 2012.

Valeria Giardino ha conseguito il Dottorato di Ricerca in Logica ed Epistemologia all'Università di Roma La Sapienza. Ha lavorato all'Institut Jean Nicod, all'Universidad de Sevilla, all'Université de Lorraine e all'École

Normale Supérieure di Parigi. È stata *Academic Fellow* all'Italian Academy (Columbia University) e *Visiting Fellow* alla Freie Universität di Berlino. Da Ottobre 2015, è Chargée de Recherche al CNRS, presso gli Archivi Henri Poincaré di Nancy. I suoi interessi di ricerca riguardano il ruolo dei diagrammi e degli artefatti cognitivi nel nostro sistema di conoscenze.

Pierluigi Graziani è assegnista di ricerca in Logica e Filosofia della Scienza presso l'Università degli Studi "G. d'Annunzio" di Chieti-Pescara. I principali ambiti di ricerca di Graziani sono la storia e la filosofia della logica e la storia e la filosofia della matematica. Tra le sue più recenti pubblicazioni possono essere ricordate: "A structural and Foundational Analysis of Euclid's plane geometry: the case study of continuity", in V. Fano, F. Orilia and G. Macchia (eds), *Space and Time. A priori and A posteriori Studies*, De Gruyter-Ontos Verlag, 2014; il volume, curato assieme a Claudio Calosi, *Meontology and Sciences. Parts and wholes in the Contemporary Scientific Context*, Synthese Library, Springer 2014.

Giorgio Grimaldi è dottore di ricerca in Filosofia. Collabora con l'Università degli Studi "G. d'Annunzio" di Chieti-Pescara. È autore di saggi pubblicati in Italia e all'estero su temi e autori fra Settecento e Novecento (fra cui Hegel, Heidegger e Adorno) e delle monografie: *Leviatano o Behemoth. Totalitarismo e franchismo* (Perugia, 2009), *Tolleranza e diritto* (Perugia, 2012), *Oltre le tempeste d'acciaio. Tecnica e modernità in Heidegger, Jünger, Schmitt* (Roma, 2015). Inoltre, ha anche curato, con Ercole Erculei, il volume collettivo *Politica, moralità, fortuna. Riflessioni storico-filosofiche sul Principe di Machiavelli* (Roma, 2015) ed è autore, con Massimo Del Pizzo, del libro *L'Automa e i suoi automi* (Pescara, 2015). È membro della redazione della rivista internazionale "Itinerari". È socio dell'*Internationale Gesellschaft Hegel-Marx für dialektisches Denken*.

Emanuela Macera (1973) si è laureata in filosofia con lode presso l'Università degli Studi di Urbino con una tesi sul pensiero di Luigi Scaravelli ed è vincitrice di borse di studio presso l'Istituto Italiano per gli Studi Filosofici di Napoli. Ha conseguito il dottorato di ricerca presso l'Istituto "Pasquale Salvucci" di Urbino sull'opera etico-storico-politica di I. Kant riletta in chiave teoretico-trascendentale dal titolo: *La "Critica della ragion politica" da Immanuel Kant ad Hannah Arendt*, seguita dal Prof. Domenico Losurdo, con cui tuttora collabora per attività scientifica. È docente di filosofia presso i Licei, conseguendo nel 2010 un Master in Legislazione scolastica e Manag-

ment della negoziazione, presso l'Università degli Studi di Perugia. Attualmente continua la propria attività di ricerca attraverso pubblicazioni volte ad indagare lo statuto epistemologico del giudizio politico come giudizio storico.

Tiziana Migliore insegna Semiotica della percezione all'Università di Venezia Ca' Foscari, dove coordina il Centro Studi LISaV – Laboratorio Internazionale di Semiotica a Venezia. Vicepresidente dell'Associazione Internazionale di Semiotica Visiva (AISV-IAVS) e direttrice della collana editoriale *Riflessi* (Aracne, Roma), ha pubblicato più di sessanta articoli scientifici, in Italia e all'estero, tre monografie, fra cui *Miroglifici. Figura e scrittura in Joan Miró* (Et al. EDIZIONI, 2011) e una dozzina di curatele, fra cui *Ernst Gombrich on the Knowledge, Theory and Analysis of Art (Journal of Art Historiography, 2011)* e *Argomentare il visibile* (Esculapio, 2008). Scrive per *La Repubblica*, *Il Manifesto*, *Alfabeto2* e *Scenari*.

Sabrina Moretti è professore associato presso l'Università di Urbino Carlo Bo e insegna Metodologia e tecniche della ricerca sociale e Analisi delle reti sociali nei corsi di laurea in Informazione, Media e Pubblicità e Sociologia. Le sue attività di ricerca riguardano principalmente lo sviluppo dei modelli di simulazione ad agenti nelle scienze sociali e le analisi di rete negli studi organizzativi.

Patrizia Pedrini è abilitata Professore Associato in Filosofia Teoretica. È autrice di saggi in italiano e in inglese su problemi al confine tra filosofia, psicologia e psichiatria. È autrice dei volumi *Prima persona. Epistemologia dell'autoconoscenza*, ETS, Pisa 2009, e *L'autoinganno. Che cos'è e come funziona*, Laterza, Roma-Bari 2013. Attualmente lavora con Julie E. Kirsch al volume in inglese *Third-Person Self-Knowledge, Self-Interpretation, and Narrative*, sotto contratto con Springer, e al manoscritto *The Lives We Can't Live. A Study of Self-Deception*, in corso di stesura.

Antonio Petagine è ricercatore in Storia della Filosofia Medievale presso l'Università di Friburgo (Svizzera). Svolge attività di docenza seminariale presso l'Università Cattolica di Milano e l'Università degli Studi di Urbino Carlo Bo. È anche coordinatore dei corsi di Introduzione all'Etica ed Elementi sull'uomo del progetto Jump delle Residenze Universitarie della Fondazione Rui. Tra le sue monografie, ricordiamo: *Aristotelismo difficile. L'intelletto umano nella prospettiva di Alberto Magno, Tommaso d'Aquino e Sigieri di*

Brabante (milano2004), *Profili dell'umano. Lineamenti di Antropologia Filosofica* (Milano 2007) e *Matière, corps, esprit. La notion de sujet dans la philosophie de Thomas d'Aquin* (Fribourg 2014).

Maria Grazia Rossi ha un titolo di Dottorato in Scienze Cognitive, ottenuto presso l'Università di Messina. È ricercatore Post-Doc all'Università Cattolica di Milano e professore a contratto per l'insegnamento di Teoria dei linguaggi per il corso di Laurea in Lingue e Comunicazione dell'Università di Cagliari. Si è occupata di teoria delle emozioni in relazione al tema della decisione morale e politica; su questi temi ha pubblicato diversi lavori e un libro: *Il giudizio del sentimento* (Editori Riuniti, 2013).

Massimo Sangoi ha conseguito il Dottorato di Ricerca in Scienze Umanistiche: Teoria dei linguaggi presso il Dipartimento di Filosofia dell'Università degli Studi di Urbino Carlo Bo. Si è occupato specificamente di *metaphor studies*, ma i suoi interessi sono rivolti più in generale all'epistemologia e alle scienze cognitive. Con Francesca Ervas ha curato il volume *Metaphor and Argumentation* (2014).

Daniele Santoro è attualmente ricercatore presso il Centro Nazionale delle Ricerche (CNR-IRPPS) e membro del *Political Theory Group* dell'Università di Minho (Braga-Portogallo). Insegna inoltre Bioetica presso il Dipartimento di Scienze Politiche dell'Università Luiss di Roma. I suoi principali interessi di ricerca si concentrano sulle teorie della responsabilità, l'analisi pragmatica delle norme, le teorie dei diritti e il ruolo del dissenso in democrazia. I suoi contributi su questi temi sono apparsi in italiano e in inglese per *Philosophia*, *Philosophical Topics*, *Filosofia e Questioni Pubbliche*, e in saggi per *Routledge and Continuum*.

Vera Tripodi, membro del LabOnt e di APhEx, lavora all'Università di Torino. Ha conseguito il dottorato di ricerca in Logica ed Epistemologia presso l'Università di Roma Sapienza. È stata Visiting Scholar presso la *Columbia University* (NY) e *Post-Doctoral Research Fellow* presso il *Centre for Gender Research* dell'Università di Oslo e il centro di ricerca *Logos* dell'Università di Barcellona. Per Carocci ha pubblicato *Filosofia della sessualità* (2011) e *Filosofie di genere* (2015).

Stefano Vaselli (Roma, 1970) è Dottore di Ricerca in Logica ed Epistemologia presso l'Università di Roma La Sapienza, dottorando in Filosofia Teoretica per il PhD FINO (Dottorato di Filosofia del Nord Ovest), nonché dal 2013 membro del LabOnt (Laboratorio di Ontologia) dell'Università degli Studi di

Torino. Cofondatore del Portale di Filosofia APhEx (Analytical and Philosophical Explanations), è docente di ruolo di filosofia e storia presso il Liceo Scientifico Amedeo Avogadro di Roma. Si occupa di temi filosofici come il realismo, l'ontologia sociale, la filosofia delle scienze sociali e la filosofia della storia, argomenti su cui ha pubblicato numerosi articoli e saggi per *Rivista di Estetica*, *Epistemologia*, *Isonomia* e capitoli di testi collettanei. Questo è il suo primo libro monografico.

Aspetti e limiti del ragionamento umano

Studi in psicologia e neuroscienze

Claudia Casadio
Università degli Studi “G. d’Annunzio” di Chieti-Pescara
claudia.casadio@unich.it

Il ragionamento svolge un ruolo fondamentale in tutte le attività umane, da quelle specificamente cognitive, come apprendimento, sviluppo, formazione ed elaborazione delle conoscenze, a quelle creative e sociali. Le inferenze, che ci consentono di ricavare conclusioni dalle informazioni a nostra disposizione, sono il mezzo attraverso cui possiamo fare ricorso al vasto bagaglio di conoscenze che possediamo e abbiamo sviluppato nel corso della nostra esistenza, siano esse implicite, inconse o consapevoli. In questo senso, il ragionamento si collega in modo naturale alle dimensioni dell’intelligenza generale, della presa di decisione e della soluzione dei problemi.

1. Ragionamento e soluzione di problemi

Trovare una soluzione alla molteplicità di problemi che si devono affrontare quotidianamente richiede ragionamento: ne facciamo uso per applicare le nostre conoscenze a situazioni particolari e adeguarle, migliorarle, a nostro favore. In effetti, il rapporto tra ragionamento e conoscenze è ambivalente: da un lato, migliori sono le nostre conoscenze, più accurate sono le inferenze che possiamo trarre, i risultati che possiamo ottenere; dall’altro, conoscenze e credenze influenzano e interferiscono col ragionamento, dando luogo a condizionamenti, fraintendimenti ed errori.

1.1. Inferenze e automatismi

Gran parte del ragionamento che impieghiamo nello svolgimento delle nostre attività quotidiane non è esplicito e consapevole, ma automatico ed inconscio, allo stesso modo in cui lo sono il sistema percettivo e la svariata ed articolata gamma di funzioni cognitive messe in opera dal nostro cervello. Si considerino i seguenti esempi:

- (1) a. Il padre del mio amico Piero è anche il padre di sua sorella.
- b. Il libro che sto leggendo ora è lo stesso che leggevo ieri.
- c. Il mio vicino di casa è la persona che si muove al piano di sopra.

Il caso (1a) è esemplificato da una situazione in cui recandomi a casa del mio amico, sono ricevuto da sua sorella che mi comunica “Piero si trova in giardino con mio padre”; dalla conoscenza in mio possesso che la ragazza che mi ha aperto la porta è la sorella di Piero, inferisco che l’uomo che è in giardino è il padre di Piero. Questo tipo di inferenza è fondata su principi di natura antropologica e sociale, relativi alla struttura della famiglia occidentale. Nel secondo esempio (1b) entra in gioco la nostra concezione ingenua dello spazio e del tempo, che ci fa identificare un oggetto come lo stesso in circostanze diverse, anche se qualcuno avrebbe potuto sostituire il libro che sto leggendo con un altro identico. Infine, (1c) richiama una situazione in cui sento dei passi nell’appartamento sopra di me e penso che siano del mio coinquilino, se non so che è partito per le vacanze, nel qual caso potrebbe trattarsi di un ladro, o di un familiare che è venuto ad innaffiare le piante sul balcone.

La nostra giornata è piena di inferenze di questo tipo; rappresentano il sostrato su cui si fonda la coscienza della nostra collocazione nello spazio e nel tempo e la rete articolata con cui intessiamo le relazioni sociali in base a principi di economia e cooperazione¹. Consideriamo un altro esempio, che riguarda una inferenza *pragmatica*, poiché presuppone una situazione in cui il parlante, ad esempio la moglie, si rivolge all’ascoltatore, il marito, chiedendogli di fare qualcosa:

- (2) a. Prendi l’auto piccola. Le chiavi sono sul tavolo.
- b. Prendi l’auto piccola e chiudi bene la porta quando esci. Le chiavi sono sul tavolo.

¹ Si veda il *Principio di cooperazione* in Grice (1975: 41-58), e gli studi sulla rilevanza in Sperber, Wilson (1986).

Nel primo caso, sappiamo, senza fermarci a riflettere, che le chiavi sul tavolo sono quelle dell'auto; nel secondo caso, invece, si è portati a pensare che le chiavi siano quelle della porta². Lo scambio linguistico, le relazioni tra persone, dipendono da inferenze come queste che si basano sull'impiego economico delle risorse cognitive e sulla cooperazione tra soggetti, confermando che esiste uno stretto rapporto tra ragionamento e linguaggio.

1.2. Inferenze induttive vs. deduttive

Negli studi di logica e ragionamento sono considerati due tipi fondamentali, ed in un certo senso opposti, di inferenza: *deduttive* vs. *induttive*. In una inferenza deduttiva si ricava una conclusione che è latente, implicita, prevista nelle premesse; si tratta di una informazione di cui siamo in possesso e che viene evidenziata, isolata dal processo deduttivo; ad esempio, la deduzione delle proposizioni singolari (3b) dalle universali (3a):

- (3) a. Tutti i poligoni hanno un perimetro.
Tutti i cani hanno i polmoni.
- b. Questo poligono ha un perimetro.
Questo cane ha i polmoni.

In una *inferenza deduttiva* la verità della conclusione segue necessariamente dalla verità delle premesse. Al contrario, in una *inferenza induttiva* la conclusione aggiunge informazione alle premesse; chiamiamo ragionamento induttivo il processo di pensiero che consente di pervenire ad una conclusione a partire da una certa evidenza (possibilmente suffragata dall'esperienza). Un caso tipico è rappresentato dalla *generalizzazione empirica* che consente, ad esempio, di derivare dall'osservazione (4a) la conclusione (4b):

- (4) a. Tutti i cigni che ho visto e che vedo sono bianchi.
- b. Quindi tutti i cigni sono bianchi.

Questa conclusione, sebbene suffragata dall'alto numero di casi che soddisfano la premessa, non è logicamente valida e può essere falsificata nel momento in cui il soggetto vede un *cigno nero*. Infatti, la conclusione di una

² Questo è infatti il sintagma nominale più prossimo, quindi l'antecedente più probabile, al rimando anaforico.

inferenza (generalizzazione) induttiva non è necessariamente vera, ma possiede soltanto un certo (più o meno alto) grado di *probabilità*; per varie ragioni, dovute alla natura stessa del processo induttivo, possono intervenire informazioni non rilevanti, mentre informazioni rilevanti possono sfuggire o essere ignorate; possono subentrare nuove evidenze (il cigno nero dell'esempio) che contraddicono la conclusione. Così, se da un lato i procedimenti induttivi di ragionamento ci consentono di arrivare a risultati accettabili, fino a quando non vengono smentiti da nuove esperienze, sono i processi deduttivi che ci danno certezze e risultati sicuri su cui fare affidamento. Nella vita quotidiana le persone si trovano a dover compiere, spesso inconsapevolmente, una grande quantità di inferenze deduttive ed induttive, come mostrano gli esempi presentati in (1a) e (2), che possono essere così analizzati:

(5) *Inferenza deduttiva* (informazione parentale implicita)

- a. Matteo = il padre della sorella di Piero.
- b. Matteo → padre di Piero.

(6) *Inferenza induttiva* (di tipo pragmatico)

- a. Il marito vuole usare l'automobile della moglie.
- b. La moglie gli dice che le chiavi sono nel cassetto.
- c. Chiavi → automobile.

Ovviamente, nell'ambito della vita quotidiana, si possono incontrare difficoltà e limitazioni nella produzione di inferenze di questo genere: nell'esempio (5) Matteo potrebbe non essere il padre naturale di Piero, ma solo il suo patrigno, o non essere neppure sposato con la madre di Piero³; nella conversazione in (6) possono occorrere fraintendimenti, difficoltà di trasmissione del messaggio, per cui, ad esempio, la moglie può parlare delle chiavi di casa, invece che di quelle dell'automobile. Tuttavia, in generale, le persone, in quanto parlanti consapevoli della propria lingua, sono in grado di produrre e trarre nel modo migliore possibile una grande varietà di inferenze deduttive o induttive, provvedendo a correggere eventuali errori, difficoltà di comunicazione, ecc.

Da un punto di vista generale, la divisione tra inferenze (o forme di ragionamento) *deduttive* vs. *induttive* è utile, ma non va presa in modo rigido; la ricerca ha mostrato che vi sono interrelazioni e sovrapposizioni tra queste

³ L'antropologia ha mostrato che il concetto di *padre* varia secondo la cultura e l'organizzazione sociale; lo stesso avviene per altre categorie linguistiche e cognitive; si veda in proposito Bruner *et al.* (1956).

due dimensioni, con particolare riferimento agli errori o alle altre difficoltà incontrate nei processi di ragionamento. Non vanno poi dimenticate le inferenze chiamate *abduitive*, studiate in modo sistematico nell'ambito delle ricerche sulla spiegazione scientifica, che consentono di risalire da ipotesi di natura osservativa a conclusioni di carattere generale⁴.

2. Ricerche sperimentali sul ragionamento

Gli studi sul ragionamento hanno privilegiato le sperimentazioni in cui i soggetti sono invitati ad affrontare e risolvere compiti⁵ di tipo *deduttivo*, oppure di tipo *induttivo* e *probabilistico*. I compiti di ragionamento deduttivo hanno il vantaggio di offrire soluzioni standard, verificabili in base ai principi della logica proposizionale, ed in certi casi della logica predicativa. I compiti impiegati nella ricerca sperimentale sul ragionamento deduttivo possono essere di tipo *formale*, astratti, o non formalizzati, ma basati su formulazioni nel *linguaggio naturale* di schemi formali di ragionamento; ciascuno di questi tipi può a sua volta essere presentato in modo *esplicito* o in modo *implicito*. Possiamo così schematizzare questa tipologia, prendendo come esempio un compito basato sulla regola di inferenza proposizionale *Modus Ponens*, che abbreviamo con la sigla MP:

⁴ Il ruolo dell'abduzione nell'ambito della logica della scoperta scientifica è stato sottolineato dal filosofo statunitense Charles Sanders Peirce: si veda in proposito Magnani, Carnielli, Pizzi (2010).

⁵ Traduciamo con la parola "compito" il termine *task* abitualmente impiegato nella letteratura relativa alle sperimentazioni sul ragionamento, come in Evans, Newstead, Byrne (1993), un testo classico in questo campo; negli studi in lingua italiana viene impiegato anche il termine "problema"; si vedano Girotto (1994), Girotto, Legrenzi (2000).

| COMPITI CON CONTENUTI FORMALI | |
|--|--|
| COMPITO ESPLICITO | COMPITO IMPLICITO |
| $\frac{A \rightarrow B}{A} \\ \hline B$ | $\frac{A \rightarrow B}{\sim (\sim A)} \\ \hline B$ |
| COMPITI CON CONTENUTI NATURALI | |
| COMPITO ESPLICITO | COMPITO IMPLICITO |
| 1. Se Maria dorme, allora sogna. 2. Maria dorme. \therefore Maria sogna. | 1. Se Maria è sveglia, allora non sogna. 2. Maria non dorme. \therefore Maria non sogna. |

Nel caso di compiti con contenuto formale ai soggetti viene richiesto di derivare la conclusione di un'operazione o regola logica di inferenza, avendo reso note le premesse: ad esempio, di derivare, in base alla regola MP, la conclusione B dalle premesse $A \rightarrow B$, A (compito esplicito) o dalle premesse $A \rightarrow B$, $\sim(\sim A)$, (compito implicito); nel secondo caso il soggetto deve saper identificare la premessa A con la formula $\sim(\sim A)$ applicando la legge della *doppia negazione* $\sim(\sim A) \leftrightarrow A$, non esplicitamente menzionata nel compito. In alternativa, ai soggetti viene resa nota la conclusione ed una premessa e viene richiesto di indicare la premessa mancante.

I compiti di ragionamento formulati per mezzo di enunciati del linguaggio naturale procedono secondo le stesse modalità: i soggetti dovranno indicare la conclusione o una premessa mancante impiegando regole logiche di inferenza, sillogismi o operazioni con i connettivi. Tuttavia, questi compiti risultano più complessi perché entrano in gioco nuovi elementi come il *contenuto* delle premesse e le *conoscenze, credenze e aspettative* dei soggetti. Ad esempio, nel compito di ragionamento che impiega l'enunciato *se Maria è sveglia, allora non sogna*, bisogna essere consapevoli della regola *implicita* di significato⁶ [dormire = non essere sveglio] che consente di inferire la conclusione *Maria non sogna* in base all'applicazione della regola della doppia negazione:

⁶ Le ricerche sperimentali offrono numerosi argomenti a sostegno della tesi che la memoria *semantica* umana contenga "postulati di significato" cognitivi, come risultato dell'interazione di esperienze percettive e processi di categorizzazione; si vedano ad esempio Anolli, Legrenzi (2001); Zorzi, Girotto (2004).

- (7) a. dormire = *non* essere sveglio
b. *non* (*non* essere sveglio) \leftrightarrow essere sveglio

In generale, mentre i compiti di ragionamento formali sono formulati senza fare riferimento al contenuto di premesse e conclusioni, per cui le soluzioni non dipendono dalle conoscenze o credenze dei soggetti, nei compiti di ragionamento concreti, formulati per mezzo di enunciati della lingua naturale, questa dimensione diventa rilevante, influenzando i risultati.

Vi sono poi due modalità che vengono seguite nella ricerca sperimentale di formulazione ed assegnazione di compiti di ragionamento: *inferenziale* oppure di *valutazione*. In un compito di tipo inferenziale (*inferential task*) viene mostrato ai soggetti un insieme di proposizioni da interpretare come premesse e si chiede di indicare un'altra proposizione come conclusione valida da tali premesse. Obiettivi di questi esperimenti è verificare se i soggetti sanno riconoscere la forma delle inferenze logiche (in genere inferenze sillogistiche o proposizionali), sanno trarre conclusioni da premesse, rispettando le regole di inferenza della logica, sanno riconoscere una conclusione valida date certe premesse.

Un compito di valutazione (*evaluation task*) riguarda la conoscenza e comprensione delle operazioni logiche, i connettivi, i quantificatori e le loro funzioni di verità. Si può, ad esempio, chiedere di costruire una tavola di verità, o un altro sistema di rappresentazione formale (albero, grafo); oppure si chiede di valutare una tavola di verità o un grafo, o di completarli.

3. Tre problemi della ricerca sul ragionamento

Quando all'inizio del XX secolo cominciarono ad essere condotti i primi esperimenti sul ragionamento, furono subito scoperti due fatti fondamentali: il primo è che gli esseri umani fanno molti errori logici nei compiti assegnati in laboratorio; il secondo è che le loro risposte sono influenzate dal contenuto, o dal contesto, impiegato per formulare il problema, nonostante la non rilevanza logica di queste dimensioni. Richiamiamo brevemente tre problemi significativi che si sono presentati ai ricercatori.

3.1. Problema della competenza

Quali sono i meccanismi che conducono i soggetti alla soluzione di problemi logici?⁷ La questione fondamentale è rappresentata dalla *correttezza* logica: con quanta frequenza e regolarità i soggetti producono le inferenze definite dalla logica; quando sbagliano e natura degli errori. Il risultato più sorprendente ottenuto in questo ambito è che è stato dimostrato che tutti i tipi di compiti di ragionamento sono associati con *alte* frequenze di errori logici. Questo, tuttavia, non significa che i soggetti svolgano a caso i compiti di ragionamento che vengono loro presentati, perché le risposte corrette sono registrate con una frequenza molto più che casuale. Il problema piuttosto è che nelle maggioranza dei casi si manifesta una tipologia di errori sistematici (*bias*) causati da vari fattori.

3.2. Problema dei bias⁸

Quali fattori causano errori sistematici e *bias* e cosa ci dicono i *bias* sulla natura del ragionamento? Gli errori sono fenomeni casuali oppure sistematicamente influenzati (*biased*) da certi aspetti del compito? È oggetto di studio come tali *bias* si formano, se sono effetto di tentativi razionali o di altre cause, la relazione che hanno con il sistema logico usato. L'osservazione degli errori sistematici del ragionamento umano può offrire grandi opportunità nella conoscenza dei processi di pensiero, poiché le condizioni che producono errori sono informative dei processi cognitivi sottostanti. Si sono riscontrati vari tipi di effetti: quelli prodotti dal *contenuto* del problema, per esempio, quali termini compaiono nelle premesse; la presenza della *negazione*; le influenze causate dal *contesto*, ovvero lo scenario, sfondo o *frame*, in cui il problema si situa.

⁷ Particolare importanza hanno avuto gli esperimenti con sillogismi classici di Wilkins (1928); Woodworth, Sells (1935, trad. it. 294-325). La metodologia seguita in questi primi esperimenti si è conservata nelle ricerche sul ragionamento fino ai giorni nostri: i soggetti sono condotti in un laboratorio e vengono loro proposti degli argomenti logici con la richiesta di indicare se una proposizione presa singolarmente, o scelta da una lista, segue logicamente dalle premesse. Si vedano in proposito Evans, Newstead, Byrne (1993); Girotto (1994); Manktelow (1999).

⁸ Il termine *bias* si riferisce a comportamenti sistematicamente collegati ad aspetti di un compito che sono *irrilevanti* rispetto alla definizione e alle regole del compito stesso; nel caso del ragionamento si tratta di aspetti di natura non logica o comunque indipendenti dalle forme del ragionamento. Questo fenomeno, tuttavia, non va confuso con manifestazioni di irrazionalità.

3.3. Problema del contenuto

La terza questione che si presenta, a partire dalle precedenti, riguarda quali aspetti del contenuto condizionano l'abilità di ragionamento e cosa ci dicono sulla natura del ragionamento stesso. Contenuto e contesto possono interagire nel determinare il ruolo svolto dalla conoscenza precedente (*prior knowledge*) posseduta dagli esseri umani nel momento in cui affrontano un problema di ragionamento.

4. Modelli teorici del ragionamento

Nella tradizione filosofica, che possiamo fare risalire ad Aristotele, “ragionare” e “ragionamento” sono intesi come sinonimi di logicità e razionalità in coerenza con l'etimologia della parola⁹. Questa situazione cambia radicalmente all'inizio del Novecento quando la psicologia sperimentale comincia a formulare ipotesi proprie su pensiero e ragionamento. In una fase iniziale delle ricerche psicologiche¹⁰, il fatto che le persone mostrassero un profilo molto basso nell'esecuzione dei compiti sperimentali di ragionamento (in particolare sillogistico e proposizionale) portò i ricercatori a concludere che ci fosse qualcosa che non andava nei partecipanti, piuttosto che nella natura e nella presentazione di regole e inferenze. Così la conclusione di queste prime ricerche da parte degli psicologi fu che gli esseri umani fossero illogici, e dunque irrazionali, in contrapposizione con le tesi logiciste, che godevano di grande favore all'epoca, e sostenevano la naturale logicità delle leggi del pensiero. Un importante contributo alla comprensione della natura peculiare del ragionamento umano è stato dato dallo psicologo britannico Peter Wason¹¹ le cui ricerche hanno fornito argomenti a favore delle posizioni contrarie alla teoria piagetiana delle regole formali e ai modelli logicisti del pensiero.

⁹ La parola “ragione”, da cui “ragionamento” e “ragionare”, deriva dall'etimologia latina *ratio* (acc. *rationem*), che vuol dire *conto*, *calcolo*, quindi facoltà di calcolare, intelligenza, raziocinio; nel Medioevo “ragionare” assume anche il significato linguistico di *argomentare*, *parlare*, ancora presente nella lingua italiana moderna (Sabatini-Coletti, *Dizionario della Lingua Italiana*, Rizzoli, Milano 2003).

¹⁰ Tra i primi studi nella letteratura sul ragionamento troviamo gli esperimenti con il sillogismo di Wilkins (1928); Woodworth, Sells (1935).

¹¹ Peter Wason (1924-2003) è a ragione ritenuto il fondatore della moderna psicologia del pensiero e del ragionamento cui ha contribuito per tutta la vita con un'intensa attività di ricerca sperimentale; si veda la sua biografia in Evans, Johnson-Laird (2003: 177-84).

¹¹ Si vedano Wason (1960: 129-40) e Wason (1966: 135-51).

A lui si devono due celebri problemi, uno relativo al ragionamento *induttivo*, il problema del 2-4-6, e uno relativo al ragionamento *deduttivo*, il *compito di selezione* noto come test delle *quattro carte*¹². Le sue ricerche hanno dimostrato che gli esseri umani sono portati a *verificare* le loro ipotesi, piuttosto che *falsificarle*, atteggiamento che è stato poi definito *bias della conferma* (*confirmation bias*). Poiché gli errori sistematici compiuti dalle persone nei compiti sperimentali rappresentano evidenti violazioni delle regole della logica formale, Wason sostenne che gli esseri umani manifestano comportamenti non logici nei loro risultati cognitivi e nelle loro decisioni. L'equazione *non-logicità = irrazionalità* appare chiaramente nei suoi scritti; ad esempio:

The selection task reflects [a tendency toward irrationality in argument] to the extent that the subjects get it *wrong*. [...] It could be argued that irrationality rather than rationality is the norm. People all too readily succumb to logical fallacies¹³.

Una situazione simile si presentò negli anni Sessanta nell'ambito delle ricerche su decisione e probabilità, ambito di indagine all'incrocio tra psicologia ed economia in cui statistica e calcolo delle probabilità erano considerati sistemi normativi corretti per rappresentare le scelte razionali umane. La situazione ebbe un mutamento radicale negli anni Ottanta con lo sviluppo della cosiddetta *teoria delle euristiche e dei biases* formulata da Amos Tversky e Daniel Kahneman. In vari articoli diventati famosi, i due psicologi statunitensi hanno dimostrato che la razionalità umana è limitata da condizionamenti cognitivi che portano le persone, quando devono prendere decisioni in situazioni di incertezza, a fare uso di strategie semplificatrici, chiamate euristiche, le quali, a loro volta, sono collegate ad un'articolata serie di errori (*bias*) sistematici di giudizio¹⁴.

Vediamo così che gli studi moderni sul ragionamento presentano due fasi: una fase iniziale in cui vi è un comune consenso nel riconoscere la logica classica, in particolare la logica proposizionale e la logica sillogistica, come lo standard *normativo* rispetto al quale verificare le capacità di ragionamento delle persone e, più in generale, la razionalità umana; una seconda fase, che si estende fino all'epoca presente, in cui, dovendo far fronte al problema degli errori e delle deviazioni dei soggetti umani dal paradigma ipotetico deduttivo,

¹² Si vedano Wason (1960: 129-40) e Wason (1966: 135-51).

¹³ Wason (1983: 59).

¹⁴ Si vedano Kahneman, Tversky (1979: 263-291); Kahneman, Slovic, Tversky (1982)

si sono cercati percorsi e modelli diversi, più adeguati alla descrizione e rappresentazione di quella ampia, e in buona parte ancora indeterminata, dimensione che identifichiamo come razionalità umana.

5. Principali teorie del ragionamento

Prendiamo ora sommariamente in considerazione alcuni modelli teorici contemporanei, le cui caratteristiche mostrano lo sviluppo incontrato dalle ricerche sul ragionamento nel corso del XX secolo.

5.1. Teoria delle regole formali

La teoria delle regole formali, conosciuta anche come *teoria della logica mentale*¹⁵, sostiene che esiste una “logica nella mente” umana, basata su un sistema interno di regole formali, anche se i suoi principi e inferenze non corrispondono necessariamente a quelli della logica formale standard. Questa posizione rimanda, in particolare, al noto psicologo svizzero Jean Piaget¹⁶, secondo il quale lo sviluppo cognitivo umano si evolve in vari stadi e si compie nell’adolescenza con l’acquisizione delle operazioni formali, risultato che si manifesta nella capacità dell’adolescente di apprendere e svolgere le operazioni matematiche e altre operazioni astratte, incluse quelle logiche. Una formulazione particolare di questa teoria, che ha avuto vari sostenitori ed è tuttora oggetto di dibattito, è che la mente contenga sistemi di regole di inferenza in base ad una “logica naturale”: un sistema limitato di regole astratte di ragionamento che viene applicato inconsciamente per dedurre, nelle varie situazioni, appropriate conclusioni da insiemi di premesse.

5.2. Teoria dei modelli mentali

La teoria dei modelli mentali è stata formulata dallo psicologo anglosassone Philip Johnson-Laird¹⁷, e si richiama alla semantica logica e alla teoria formale dei modelli. Essa caratterizza il ragionamento come un processo *semantico*, piuttosto che sintattico, basato non su regole di inferenza formali,

¹⁵ L’esponente più noto di questo orientamento teorico è lo psicologo cognitivo RIPS, autore del volume *The Psychology of Proofs* (1994), MIT Press, Cambridge (MA); si vedano anche Braine (1990: 133-57); Johnson-Laird (1983).

¹⁶ Si veda Piaget (1956).

¹⁷ Si veda Johnson-Laird (2001: 434-42).

indipendenti dal contenuto, ma su procedure sistematiche per costruire e valutare *modelli mentali* delle proposizioni da cui sono derivate le conclusioni. Questi modelli rappresentano stati *parziali* del mondo in cui le premesse sono vere.

Secondo questa teoria, il ragionamento si sviluppa attraverso tre stadi fondamentali: (i) in un primo momento, i soggetti elaborano, in gran parte inconsciamente, un *modello mentale* per rappresentarsi un possibile stato del mondo (situazione) che sia consistente con le informazioni fornite dalle premesse; (ii) quindi, formulano una conclusione plausibile generando una descrizione del modello che sia semanticamente informativa (non una semplice ripetizione di una premessa o un enunciato meno informativo di una premessa); (iii) infine, la conclusione può essere verificata cercando di costruire come *controesempio* modelli alternativi in cui le premesse dell'argomento sono *vere*, ma la conclusione è *falsa*. Se nessun controesempio di questo genere viene trovato, si può inferire che la conclusione è *valida*. Questa teoria spiega come mai i soggetti hanno maggiori difficoltà nelle regole in cui compaiono premesse negative, come il *Modus Tollens* (MT), rispetto a regole come *Modus Ponens* (MP), in cui le premesse sono tutte affermative: risulta infatti più complesso creare modelli mentali di situazioni in cui gli oggetti indicati nelle premesse non sono presenti, come risultato della loro negazione. Si consideri il seguente esempio:

(7) MODUS PONENS

- a. Se la lettera è **A**, allora il numero è **3**.
- b. La lettera è **A**.
- c. Il numero è **3**.

RISULTATI: corretti da 100% (98%) a 90%

(8) MODUS TOLLENS

- a. Se la lettera è **A**, allora il numero è **3**.
- b. Il numero non è **3**.
- c. La lettera non è **A**.

RISULTATI: corretti da 75% a 56%

| | |
|--------------|--------------|
| A | 3 |
| A | Non 3 |
| Non A | 3 |
| Non A | Non 3 |

Per eseguire correttamente MP è sufficiente una rappresentazione parziale delle condizioni di verità di un enunciato condizionale, in base a cui antecedente **A** e conseguente **3** sono veri (prima riga della tavola).

Invece, per eseguire MT è necessario rendere *esplicite tutte* le condizioni di verità del condizionale, in particolare quelle in cui antecedente **A** e conseguente **3** sono *falsi*, quindi sono veri **Non 3** e **Non A** (quarta riga della tavola).

La teoria dei modelli mentali ha riscosso, e tuttora riscuote, molto successo nell'ambito degli studi sul ragionamento; vi sono tuttavia aspetti del comportamento dei "ragionatori umani" che non possono essere spiegati esclusivamente nei termini della generazione di appropriati modelli mentali perché coinvolgono il piano complesso e affascinante del rapporto dell'uomo con l'ambiente.

5.3. Teorie relative al contesto

Negli anni Ottanta del secolo scorso si sono sviluppati approcci teorici che spiegano il ragionamento facendo esplicito riferimento all'ambiente e al comportamento dei soggetti umani. Uno di questi è la *Teoria degli schemi pragmatici*¹⁸, che sostiene che il ragionamento è 'sensibile al contesto' (*context sensitive*) e dipende in modo significativo dagli scopi e dalle finalità che i soggetti si propongono di realizzare: vi sono *schemi pragmatici*, che sono attivati dal contesto e si traducono in regole che possono essere applicate per ragionare in un dominio particolare (*context* o *domain sensitive*). In essenza, la teoria sostiene che gli esseri umani imparano a ragionare in certi contesti o ambienti e producono, danno vita, a partire da tali esperienze, a schemi che permettono di fare astrazione dalle situazioni specifiche, rendendo la conoscenza più generale. Queste regole hanno frequentemente carattere *deontico*, assumendo la forma di *permessi*, *obblighi* ecc., ovvero regole di carattere sociale che indicano ciò che si deve (si può) o non si deve (non si può) fare. Per esempio, si può ipotizzare uno schema pragmatico di *permesso* che si applica a regole che assumono la forma generale:

(9) Se deve essere compiuta un'azione **A**, allora deve essere soddisfatta la condizione preliminare (o *precondizione*) **B**: azione **A** → condizione **B**

(10) a. Se una persona beve alcolici in un bar, allora deve avere più di 18 anni.
b. Jim ha 16 anni (= Jim non ha più di 18 anni)

¹⁸ La teoria è stata formulata in particolare da Cheng, Holyoak (1985: 391-416); ha avuto una varietà di sviluppi e applicazioni: cfr. Girotto (1994: 43-55); Girotto, Legrenzi (2000).

c. Dunque, Jim non può bere alcolici in un bar (MT)

RISULTATI Significativi miglioramenti nella performance di MT: da **40%** a **70%**

Applicazioni al compito di selezione di Wason: astratto **1%**, non modale **8%**, modale **74%**

Se si usa un enunciato condizionale come (10), in un contesto che può essere interpretato come un *permesso* o *una disposizione* sociale, legale, ecc., allora viene richiamato lo schema pragmatico rilevante cui sono associate le regole di ragionamento che possono essere applicate¹⁹.

5.4. Teorie evoluzionistiche

Il modello forse più popolare e discusso ai giorni nostri è la *psicologia evoluzionistica*²⁰ proposta dagli psicologi statunitensi Leda Cosmides e John Tooby; secondo questi autori, gli esseri umani non possiedono una logica indipendente dal contesto: il ragionamento si è evoluto attraverso strategie di soluzione dei problemi nell'ambiente naturale e, in particolare, sociale. Un ruolo fondamentale in questo sviluppo è stato assolto dalla facoltà, cognitiva e sociale, detta di *cheater detection*. Il ragionamento, e in genere la razionalità umana, dipendono dal contenuto del contesto umano e sociale, perché si sono sviluppati come conseguenza di specifiche pressioni dell'ambiente, in particolare i così detti "contratti sociali" in cui valgono inferenze come:

contratto sociale → accordo-coesione → successo → controllo di trasgressioni e violazioni

Leda Cosmides²¹ sostiene che il ragionamento riuscito (in accordo con i canoni della logica classica) si ha solo in casi in cui si presentano *contenuti* molto circoscritti e che questi contenuti sono stati selezionati dall'evoluzione: corrispondono a situazioni nel nostro ambiente a cui siamo particolarmente preparati, perché sono cruciali per la sopravvivenza. Le persone, trovandosi a dovere risolvere un problema che richiede ragionamento, prendono decisioni che riflettono le modalità quotidiane utilizzate per raggiungere i propri scopi. Il ragionamento logico corretto ha le sue origini nella coordinazione

¹⁹ I primi a lavorare in questa direzione sono stati Griggs, Cox (1982: 407-20).

²⁰ Traduciamo così il termine *evolutionary psychology* usato nella letteratura; l'opera più nota è Cosmides (1989: 187-276); si richiama anche Cosmides, Tooby (1992: 163-228).

²¹ Si veda Cosmides (1989).

sociale: per assicurare la sopravvivenza è essenziale saper individuare le controparti che non rispettano gli accordi, controparti che accettano il *beneficio* senza pagare il prezzo associato. D'altro lato, individuare l'altruismo, che rappresenta una sorta di disposizione *default* dei soggetti umani, non è necessario, per cui non si sono evoluti *altruist detectors*²².

In questo quadro teorico, non è sorprendente che il modello sperimentale preferito dalla psicologia evoluzionistica sia il compito di selezione di Wason: i risultati sperimentali hanno mostrato che tutti i ragionamenti riusciti nel *compito delle quattro carte* comportano *contratti sociali*; in altre parole, tali compiti hanno un sostanziale carattere *normativo*, in cui i partecipanti concordano sullo scambio di benefici; tali contratti includono provvedimenti, avvertimenti, vari tipi di modalità deontiche. Sommarariamente, possiamo così riassumere le categorie teoriche caratteristiche dell'approccio evoluzionistico:

1. *Contratti sociali* Il ragionamento riuscito è il risultato di contratti sociali in cui le diverse parti si accordano per scambiare benefici.
2. *Cooperazione* La corretta esecuzione (anche dal punto di vista della logica) di molti compiti di ragionamento (ad esempio MP) mostra che deve esserci stata una forma di adattamento in questa direzione. La buona riuscita nel ragionamento ha origine nella cooperazione sociale.
3. *Cheater detection* Per assicurare la sopravvivenza è necessario controllare che le parti rispettino i contratti, gli accordi presi; devono essere in grado di controllare che la controparte non intenda ingannare, ricevere i benefici senza pagare il prezzo previsto. La buona esecuzione di un compito di ragionamento dipende dall'attivazione di meccanismi di *cheater detection*, ovvero i meccanismi di controllo dell'inganno, della violazione, dell'imbroglio.

5.5. Teorie dualistiche

Un recente sviluppo delle ricerche su pensiero e ragionamento è dedicato all'ipotesi che operazioni e funzioni del ragionamento siano dovute a due distinti sistemi cognitivi in competizione per il controllo delle nostre azioni ed inferenze: un primo sistema, chiamato **Sistema 1**, molto antico dal punto di

²² Vi sono molte convergenze tra questo approccio e la teoria delle massime conversazionali e *principio di cooperazione* di Grice; cfr. Grice (1975).

vista dell'evoluzione e condiviso con gli altri rappresentanti del mondo animale; un secondo sistema, detto **Sistema 2**, più recente dal punto di vista evolutivo, e tipicamente umano. Queste teorie collocano «due menti in un solo cervello»²³ e hanno prodotto negli ultimi anni un insieme significativo di dati sperimentali che si avvalgono delle risorse tecniche e teoriche delle scienze fisiologiche e delle neuroscienze²⁴. Molte delle applicazioni delle teorie dualiste del ragionamento sono realizzate nelle aree scientifiche della teoria del giudizio e della decisione. Le differenze fra i due sistemi possono essere così sintetizzate:

SISTEMA 1. È descritto come una forma di cognizione universale condivisa da animali ed esseri umani; include i comportamenti istintivi innati e contiene un insieme di sotto-sistemi autonomi che contengono sia moduli innati che sistemi di conoscenze specifiche, acquisite attraverso meccanismi generali di apprendimento. I sostenitori del modello dualista ritengono che i suoi processi siano essenzialmente di tipo *associativo*, e le operazioni siano *rapide, parallele, automatiche e inconsce*.

SISTEMA 2. Più recente a livello evolutivo, sviluppato come prerogativa della specie umana, le sue modalità sono *lentezza e sequenzialità*; fa specificamente uso del sistema centrale della memoria di lavoro. Presiede al ragionamento *astratto* e al pensiero *ipotetico*, che non possono essere ottenuti dal Sistema 1, ma è controllato dalle capacità limitate della memoria di lavoro e si correla alle capacità di intelligenza generale.

Consideriamo un semplice esempio relativo alla presa di decisione: possiamo, e frequentemente lo facciamo, decidere le nostre azioni sulla base delle esperienze passate, ricordando ciò che è stato, o ha funzionato, meglio per noi. Queste decisioni *intuitive*, spontanee, che non richiedono molta riflessione, dipendono dal SISTEMA 1. Tuttavia, possiamo, e in molti casi critici siamo costretti a farlo, anche costruire modelli mentali o simulazioni di eventi futuri che ci aiutino a progettare le decisioni che dobbiamo prendere. Questa capacità di ragionare in modo *ipotetico*, che caratterizza il pensiero umano, con particolare riferimento alla elaborazione astratta, è assicurata dal SISTEMA 2.

²³ Evans (2003: 454).

²⁴ Si vedano Evans (2003: 454-59); Goel, Dolan (2001: 901-909).

Una delle metodologie preferite per dimostrare la presenza di processi dualistici nei compiti di ragionamento fa ricorso all'effetto del *contenuto* delle premesse chiamato *belief-bias* (bias della credenza). Negli esperimenti si cerca di creare un conflitto nei soggetti tra risposte basate su un processo di ragionamento logico e risposte derivate dalle credenze precedenti (*prior belief*) sulla verità delle conclusioni. In particolare i partecipanti vengono invitati a considerare argomenti le cui conclusioni sono *credibili o non credibili*. Come metodo tipico vengono proposti dei sillogismi di cui deve essere stabilita la validità (compito di valutazione); questi sillogismi rientrano in una delle seguenti quattro categorie²⁵:

BIAS DELLA CREDENZA

- (1) Argomento *valido*, conclusione *credibile* (NESSUN CONFLITTO)
 - a. Nessun cane da caccia è aggressivo.
 - b. Alcuni cani addestrati sono aggressivi.
 - c. Quindi, alcuni cani addestrati non sono cani da caccia.

- (2) Argomento *valido*, conclusione *non credibile* (CONFLITTO)
 - a. Nessun cosmetico è economico.
 - b. Alcuni prodotti di bellezza sono economici.
 - c. Quindi, alcuni prodotti di bellezza non sono cosmetici.

- (3) Argomento *invalido*, conclusione *credibile* (CONFLITTO)
 - a. Nessuna sostanza additiva è economica..
 - b. Alcune marche di sigarette sono economiche.
 - c. Quindi, alcune sostanze additive non sono sigarette.

- (4) Argomento *invalido*, conclusione *non credibile* (NESSUN CONFLITTO)
 - a. Nessun milionario è un lavoratore instancabile.
 - b. Alcuni uomini ricchi sono lavoratori instancabili.
 - c. Quindi, alcuni milionari non sono uomini ricchi.

²⁵ In questo caso si tratta di sillogismi della seconda figura, modo *Festino*, con ha la forma logica: Nessun A è B; Qualche C è B; Qualche C non è A; si veda Evans, Newstead, Byrne (1993).

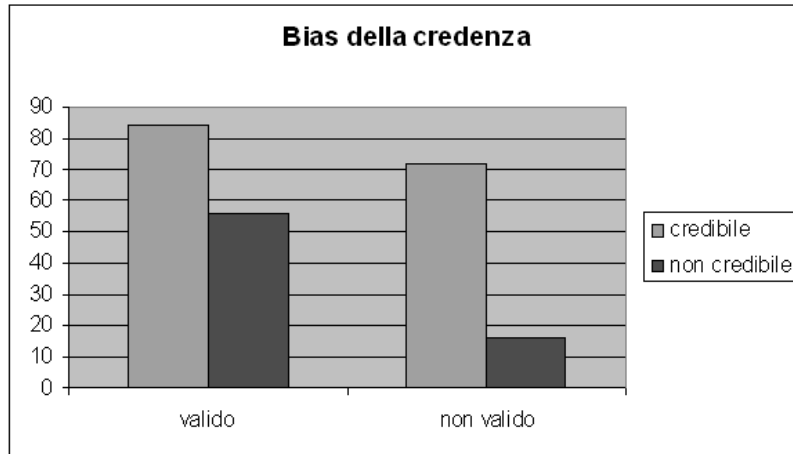


Figura 1

Trattandosi di un problema di tipo logico, bisogna indicare solo le conclusioni che seguono necessariamente dalle premesse; tuttavia i partecipanti, in genere studenti universitari, quindi persone adulte intelligenti, sono significativamente influenzati tanto dalla *credibilità* della conclusione proposta, quanto dalla *validità* dell'argomento. I risultati di un esperimento di questo tipo sono illustrati nella *Figura 1*, dove sono indicati i valori per le coordinate: sillogismo *valido* vs. *invalido* (asse delle x) e conclusione *credibile* vs. *non credibile* (colore chiaro vs. scuro)²⁶.

I partecipanti all'esperimento si dimostrano influenzati dalla credibilità della conclusione, esibendo un forte effetto di *bias* della credenza nel caso del sillogismo *invalido* con conclusione *credibile*. La spiegazione offerta dalle teorie dualistiche è che, sebbene i partecipanti tentino di ragionare logicamente in accordo con le istruzioni che hanno ricevuto, l'influenza delle credenze è molto forte, difficile da sopprimere e condiziona le risposte. Così la performance del compito viene influenzata dal *conflitto* tra i processi di ragionamento attivati dal Sistema 2, e le suggestioni delle credenze governate dal Sistema 1. Secondo Evans²⁷ la capacità di risolvere questo conflitto in favore della logica si collega alla misurazione delle abilità cognitive dei soggetti, in particolare l'intelligenza generale e la memoria a breve termine, e declina in modo significativo con l'età.

²⁶ Si veda Evans (1982: 455).

²⁷ Evans (2002: 978-96).

6. Studi sul ragionamento in psicologia e neuroscienze

6.1. Studi con i sillogismi

Il sillogismo è la forma di inferenza deduttiva più usata dagli psicologi nelle ricerche sperimentali sul ragionamento umano. I soggetti sperimentali accettano di buon grado di esprimere un giudizio in merito ai sillogismi loro proposti, ma vengono commessi molti errori. La ricerca in psicologia per tutto il secolo scorso è stata dedicata a capire la sorgente e la ragione di tali errori.

Esistono 24 sillogismi validi²⁸, ma gli psicologi li raddoppiano perché riconoscono anche le forme con premesse invertite, anzi nei test e compiti sperimentali in genere compaiono proprio questi sillogismi rovesciati; così la sequenza AB-BC è la I figura classica, con le premesse rovesciate; AB-CB è la II figura, BA-BC è la III figura, BA-CB è la IV figura, con le premesse rovesciate. La seguente presentazione illustra come queste figure offrano il vantaggio di mostrare in modo lineare la *transitività* dalla prima alla seconda premessa:

| I figura | II figura | III figura | IV figura |
|----------|-----------|------------|-----------|
| A - B | A - B | B - A | B - A |
| B - C | C - B | B - C | C - B |
| A - C | A - C | A - C | A - C |

Una conclusione di un'inferenza logica, e in particolare di un'inferenza sillogistica è *valida* se: è una conclusione che deve essere VERA se le premesse sono VERE; in altre parole, se le premesse di un'inferenza logica sono VERE, non è possibile derivare una conclusione FALSA. Negli studi sul sillogismo questo porta alla domanda: come possiamo decidere quali sillogismi hanno conclusioni valide? Questa risposta è tutt'altro che ovvia, e ricercatori diversi hanno raggiunto risultati differenti nel corso del tempo. I sillogismi, infatti, producono una vasta tipologia di risposte e soluzioni: alcune forme valide sono riconosciute come tali da tutti i soggetti, ad esempio la forma universale della I figura *Barbara* (detta anche *perfetta*):

| | |
|--------------------|---|
| Tutti gli A sono B | Tutti i deputati sono parlamentari |
| Tutti i B sono C | Tutti i parlamentari sono eletti dal popolo |
| Tutti gli A sono | C Tutti i deputati sono eletti dal popolo |

²⁸ Si veda Casadio (2005: 28-30).

Altre forme valide non vengono tuttavia comprese e ricevono al massimo risposte date a caso. D'altro lato, molti sillogismi non validi vengono riconosciuti come validi. Nell'esempio il primo e terzo sillogismo sono validi (I figura, *Darii*), mentre il secondo e il terzo sono invalidi:

| | |
|--|---|
| DARII | FALLACIA |
| Tutti gli A sono B Qualche C è A Qualche C è B | Tutti gli A sono B Qualche C è B *Qualche C è A |
| DARII | FALLACIA |
| Tutti i Francesi sono Europei Qualche regista è Francese Qualche regista è Europeo | Tutti i Francesi sono Europei Qualche regista è Europeo *Qualche regista è Francese |

In generale, l'impiego di sillogismi con premesse e conclusioni *concrete*, detti anche sillogismi *realistici*, ha ottenuto risultati in genere migliori, rispetto all'impiego di sillogismi *astratti*, basati su lettere o altro materiale formale. Nelle sperimentazioni sviluppate a partire dagli anni Trenta del secolo scorso, sono stati individuati alcuni significativi aspetti comuni: vi sono sillogismi molto facili, che le persone affrontano e risolvono senza difficoltà; vi è poi una gamma di variazioni da questa situazione di base fino a giungere a casi di sillogismi molto complessi, che nessuno riesce a risolvere; infine, la buona o cattiva esecuzione dei soggetti durante gli esperimenti dipende in modo significativo dalla *disposizione* e dal *contenuto* delle premesse. Lo studio di questi aspetti ha portato i ricercatori a formulare alcune ipotesi teoriche che ci aiutano a spiegare il ragionamento sillogistico.

La prima di queste ipotesi teoriche è conosciuta come *effetto atmosfera*: la modalità esibita dalla coppia delle premesse induce una disposizione generale, o atmosfera, che suggerisce conclusioni di un certo tipo. Ad esempio, due premesse *universali* suggeriranno una conclusione *universale*, due premesse *affermative* suggeriranno una conclusione *affermativa*; la presenza di almeno una premessa *particolare* suggerirà una conclusione *particolare*; e la presenza di almeno una premessa *negativa* suggerirà una conclusione *negativa*. Questo modo di affrontare il sillogismo assicura una buona resa dal

punto di vista logico, dal momento che molti degli argomenti validi sono consistenti con l'effetto atmosfera²⁹; vengono inoltre spiegati molti degli errori osservati.

Una seconda ipotesi viene detta dell'*associazione (matching effect)*: si preferisce che nella conclusione occorra lo stesso quantificatore che appare nelle premesse (quindi l'ampia accettazione dell'occorrenza di tre quantificatori universali nei sillogismi della I figura, modo *Barbara*); se nelle premesse compaiono quantificatori diversi, si preferisce quello più *conservativo*, ovvero quello che comporta la considerazione del numero minore di casi. Così *nessuno* è preferito a *qualche* che, a sua volta, è preferito a *tutti*. Analogamente si ipotizza che gli individui completino le inferenze sillogistiche scegliendo una conclusione che, avendo lo stesso modo della premessa, aggiunge meno informazione³⁰.

Sono stati poi individuati due tipi di effetti connessi all'interpretazione delle premesse: il primo, chiamato *conversione*, è stato proposto da Chapman e Chapman (1959)³¹: gli errori compiuti nel derivare le conclusioni dei sillogismi dipendono dal fatto che le persone interpretano le premesse come se implicassero le loro converse. Si hanno le seguenti possibilità:

| | | | |
|---|--------------------|----------|---------------------|
| A | Tutti gli A sono B | CONVERSA | * Tutti i B sono A |
| E | Nessun A è B | | Nessun B è A |
| I | Qualche A è B | | Qualche B è A |
| O | Qualche A non è B | | * Qualche B non è A |

Di fatto solo due di queste conversioni sono invalide, come indicano gli asterischi: infatti le proposizioni di tipo E (universale negativa) e di tipo I (particolare affermativa) possono essere validamente convertite. Gli errori di conversione nei compiti sperimentali riguardano, dunque, le proposizioni di tipo A (universale affermativa) ed O (particolare negativa). Questo dato è chiamato anche *effetto alla Grice*, perché spiega le interpretazioni più frequenti delle premesse sillogistiche e le relative conclusioni inferite dalle persone nelle situazioni normali, quotidiane, facendo ricorso alla teoria della

²⁹ L'effetto *atmosfera* ripropone le condizioni di buona formazione dei sillogismi già indicate da Aristotele: la presenza di una premessa negativa richiede una conclusione negativa; se una premessa è particolare anche la conclusione deve essere particolare; se entrambe le premesse sono universali, anche la conclusione deve esserlo.

³⁰ Chater, Oaksford (1999: 57-65).

³¹ Chapman, Chapman (1959, trad. it. 334-49).

conversazione e al *principio di cooperazione* formulati dal filosofo Paul Grice³².

Gli studi più recenti si sono concentrati sulla dimensione *dinamica* del ragionamento sillogistico, analizzandolo nei termini delle *strategie* messe in atto dagli individui per inferire le conclusioni dalle premesse o per valutare la validità degli argomenti. Particolare attenzione è stata di recente dedicata alla interazione tra le premesse con quantificatori universali e la premesse con quantificatori esistenziali (particolari)³³. Si è anche approfondito il ruolo del ragionamento sillogistico nella comprensione delle diverse funzioni mentali attivate dal ragionamento deduttivo quando si presenta in modo *astratto*, come processo essenzialmente logico-linguistico, e quando invece richiede manipolazione spaziale e ricerca di elementi *concreti*; i risultati hanno mostrato che il ragionamento sillogistico è attivato in due distinti sottosistemi cerebrali, in dipendenza dalla presenza o meno di contenuto semantico, come propongono le teorie dualistiche del ragionamento.

6.2. Ragionamento e contesto

Negli studi ad orientamento cognitivo il termine *frame* viene invocato per indicare una varietà di dimensioni come il contesto di riferimento, l'organizzazione concettuale, la struttura di particolari facoltà mentali o senso motorie³⁴. Nella proposta di George Lakoff³⁵ la nozione di *frame* assume il senso generale di struttura concettuale che coinvolge tutte le attività cognitive umane e si manifesta nella articolazione del contesto (messaggio, comunicazione, ambiente ...) in cui tali attività si esplicano. Assumendo questa prospettiva, possono essere individuati e caratterizzati anche *frames* del ragionamento, responsabili dei *bias* specifici di questa attività cognitiva:

REASONING FRAMES (*contesti di ragionamento*)

1. Negazione → *Negation Bias* (bias della negazione)
2. Credenze → *Belief Bias* (bias della credenza)
3. Corrispondenza → *Matching Bias* (bias della associazione)
4. Contenuto pragmatico → *Behaviors, Conventions* (comportamenti, convenzioni)

³² Grice (1975: 41-58).

³³ Si veda Abrusci, Casadio, Medaglia, Porcaro (2013: 1017-1027).

³⁴ Si vedano Minsky (1975: 211-277); Barsalou (1984).

³⁵ Lakoff & Johnson (1980).

5. Contenuto deontico → *Obligations, Preferences* (obblighi, preferenze)
6. Contenuto emozionale → Emozioni Positive vs. Negative
Pessimismo vs. Ottimismo
Atteggiamento collaborativo vs. non collaborativo

Secondo Lakoff³⁶ ogni parola può introdurre, suggerire, un *frame*: ad esempio, la parola ELEFANTE suggerisce il *frame* associato ad un certo insieme di tratti, frutto delle conoscenze condivise, socialmente e culturalmente alimentate: grandezza, peso, larghe orecchie, naso lungo e mobile. Tra gli aspetti che caratterizzano un *frame* sono rilevanti i seguenti: (i) organizzazione complessa: avere una articolazione interna, in particolare, un nucleo (*core*) e una periferia (ii) familiarità: evocare un *frame*, lo rafforza; (iii) atmosfera: la negazione di un *frame*, lo rinforza³⁷.

6.3. Ragionamento e neuroscienze

Recenti risultati sperimentali a favore della teoria dualistica del ragionamento sono stati presentati da Vinod Goel e colleghi³⁸, ricorrendo a metodi di studio in ambito neuropsicologico che fanno uso di strumenti di *brain imaging* come Elettroencefalografia (EEG) Risonanza Magnetica funzionale (fMRI), Magnetoencefalografia (MEG). Ad esempio, un esperimento disegnato da V. Goel e R. J. Dolan³⁹ ha fornito consistente evidenza a favore dell'esistenza di una differenziazione, a livello neurologico, tra il ragionamento compiuto con materiale astratto e quello eseguito con materiale concreto, dotato di significato e ambientato in contesti familiari ai soggetti. Gli autori concludono, in sintonia con le tesi del modello dualista, che il ragionamento è implementato in due sistemi cognitivi distinti, attivati in diverse aree cerebrali in dipendenza dell'uso, o meno, di contenuto semantico.

In particolare, gli esperimenti, che mettono a confronto le reazioni a sillogismi *astratti* vs. sillogismi concreti dotati di *contenuto semantico*, sottoponendo soggetti volontari all'impiego della risonanza magnetica funzionale (fMRI), hanno mostrato che il ragionamento basato sul contenuto attiva

³⁶ Lakoff (2004).

³⁷ Queste e simili proprietà avvicinano la teoria dei *frames* ad altre teorie popolari nella semantica cognitiva, in particolare la teoria del *prototipo* e le teorie della *categorizzazione*; si vedano Girotto, Legrenzi (2000); Zorzi, Girotto (2004).

³⁸ Goel, Dolan (2001: 901-909).

³⁹ Goel, Dolan (2004: 109-21).

un'area significativa dell'emisfero *sinistro* del cervello, mentre quello senza contenuto semantico attiva un'area del sistema *parietale*. Altre ricerche, realizzate in questa direzione⁴⁰, hanno individuato interessanti contrasti che coinvolgono contesti emozionali, dove le emozioni considerate possono essere emozioni *positive*, come allegria, gioia, gratitudine, empatia, altruismo, o emozioni *negative* come tristezza, paura, rabbia, vergogna. I seguenti sono alcuni esempi di sillogismi con contenuto emozionale, in questo caso emozioni negative, impiegati in test sperimentali in cui, presentate ai soggetti le due premesse (a, b), si chiede di scegliere tra le due conclusioni (c: corretta, c*: scorretta):

SILLOGISMI CON EMOZIONI NEGATIVE

PREMESSE

- a) Un uomo in crisi epilettica era vicino alla stanza dei medici.
- b) La stanza dei medici era al pronto soccorso.

CONCLUSIONI

- c) Un uomo in crisi epilettica era vicino al pronto soccorso.
- c*) La stanza dei medici era vicino ad un uomo in crisi epilettica

PREMESSE

- a) Nessun infermiere era in grado di occuparsi dei malati del terzo piano.
- b) Qualche giovane che si trovava al terzo piano era infermiere.

CONCLUSIONI

- c) Qualche giovane che si trovava al terzo piano non era in grado di occuparsi dei malati.
- c*) Nessun infermiere si trovava al quarto piano

Come è stato sottolineato da Pessoa (2008), questo genere di ricerche ha mostrato che regioni cerebrali tipicamente definite come “affettive”, in particolare l'*amigdala*⁴¹, sono coinvolte in modo significativo nei processi cognitivi; e regioni cerebrali tradizionalmente considerate “cognitive”, come la

⁴⁰ Si richiamano Pessoa (2008: 148-58); Brunetti (2013).

⁴¹ L'*Amigdala* è una parte del cervello che gestisce le *emozioni*, in particolare la paura. A livello anatomico ha forma ovoidale (in greco antico *amygdala* significa mandorla) ed è

corteccia frontale, sono impegnate in modo cruciale nelle emozioni. La conclusione è che *cognizione* ed *emozioni* costituiscono processi *integrati* nel cervello: le regioni cerebrali di natura affettiva intervengono nella cognizione dando luogo a comportamenti cognitivi-emozionali complessi (*complex cognitive-emotional behaviors*). La ricerca contemporanea, impegnata a fornire una più adeguata spiegazione della base neurale della cognizione, deve dunque trovare una spiegazione anche per la base neurale della integrazione di emozioni e cognizione.

Bibliografia

- Abrusci, V. M., Casadio, C., Medaglia, M. T., Porcaro, C. (2013), "Universal vs. particular reasoning: a study with neuroimaging techniques, *Logic Journal of the IGPL*, vol. 21, pp. 1017-1027.
- Anolli, L., Legrenzi, P. (2001), *Psicologia generale*, Bologna, Il Mulino.
- Barsalou, L. W. (1984), *Determination of graded structures in categories*, GA: Psychology Department, Atlanta, Emory University.
- Blanchette, I., Richards, A. (2010), "The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning", *Cognition & Emotion*, 24(4), 561-595.
- Braine, M. D. S. (1990), "The Natural Logic Approach to Reasoning", in Overton, W. F. (ed.), *Reasoning, Necessity, and Logic: Developmental Perspectives*, Hillsdale (NJ), Erlbaum, pp. 133-57.
- Bruner, J. S., Goodnow, J. A., Austin, G. A., Brown, R. W. (1956), *A Study of Thinking*, Wiley, New York (*Il pensiero: strategie e categorie*, trad. it. Di E. Rivero, Roma, Armando, 1973).
- Brunetti, M., Perrucci, M. G., Di Naccio, M. R., Ferretti, A., Del Gratta, C., Casadio, C., Romani, G. L. (2013), *Framing deductive reasoning with emotional content: an fMRI study*, preprint submitted.

un'area che fa parte del sistema limbico, nella regione rostromediale del lobo temporale, al di sotto del giro uncinato e anteriormente all'ippocampo.

- Casadio, C. (2005), *Logica e psicologia del pensiero*, Roma, Carocci.
- Chapman, J. P., Chapman, J. P. (1959), "Atmosphere Effect Re-examined", *Journal of Experimental Psychology*, 58, pp. 220-6; (*L'effettoatmosfera riesaminato*, trad. it. di P. Legrenzi, A. Mazzocco (a cura di), 1975, pp. 334-49).
- Chater, N., Oaksford, M. (1999), "Ten Years of the Rational Analysis of Cognition", *Trends in Cognitive Science*, 3, 2, pp. 57-65.
- Cheng, P. W., Holyoak, K. J. (1985), "Pragmatic Reasoning Schemas", *Cognitive Psychology*, 17, 391-416.
- Cosmides, L. (1989), "The Logic of Social Exchange: Has Natural Selection Shaped how Humans Reason? Studies with the Wason Selection Task", *Cognition*, 31, 187-276.
- Cosmides, L., Tooby, J. (1992), "Cognitive Adaptations for Social Exchange", in J. Barkow, L. Cosmides, J. Tooby (a cura di.), *The Adapted Mind*, New York, Oxford University Press, 163-228.
- Evans, J. St. B. T. (1982), *The Psychology of Deductive Reasoning*, London, Routledge & Kegan Paul.
- Evans, J. St. B. T. (2002), "Logic and Human Reasoning: An Assessment of the Deduction Paradigm", *Psychological Bulletin*, 128, 978-96.
- Evans, J. St. B. T. (2003), "In Two Minds: Dual-process Accounts of Reasoning", *Trends in Cognitive Science*, 7(10), 454-9.
- Evans, J. St. B. T., Johnson-Laird, P. N. (2003), "Peter Wason (1924-2003)", *Thinking and Reasoning*, 9(3), 177-84.
- Evans, J. St. B. T., Newstead, S. E., Byrne, R. M. J. (1993), *Human Reasoning: The Psychology of Deduction*, Mahwah (NJ), Lawrence Erlbaum Associates.
- Giroto, V. (1994), *Il ragionamento*, Bologna, Il Mulino.
- Giroto, V., Legrenzi, P. (2000), *Psicologia del pensiero*, Bologna, Il Mulino.

- Goel, V., Dolan, R. J. (2001), "Functional neuroanatomy of three-term relational reasoning", *Neuropsychologia*, 39 (9), 901–909.
- Goel, V., Dolan, R. J. (2004), "Differential involvement of left prefrontal cortex in inductive and deductive reasoning", *Cognition*, 3:B, 109-21.
- Grice, P. (1975), "Logic and Conversation", in Cole, P., Morgan, J. L. (a cura di), *Speech Acts*, Academic Press, pp. 41-58; (*Logica e conversazione*, trad. it. di P. Casalegno et al. (a cura di), *Filosofia del linguaggio*, Milano, Cortina, 2002).
- Griggs, R. A., Cox, J. R. (1982), "The Elusive Thematic Materials Effect in the Wason Selection Task", *British Journal of Psychology*, 73, pp.407-20.
- Houdè, O., Zago, L., Crivello, F., Moutier, S., Pineau, A., Mazoyer, B., Tzourio-Mazoyer, N. (2001), "Access to deductive logic depends on a right ventromedial prefrontal area devoted to emotion and feeling: evidence from a training paradigm", *Neuroimage*, vol. 14, 1486-1492.
- Johnson-Laird, P. N. (1983), *Mental Models*, Cambridge (MA), Cambridge University Press.
- Johnson-Laird, P. N. (2001), "Mental Models and Deductive Reasoning", *Trends in Cognitive Science*, 5, 434-42.
- Kahneman, D., Slovic, P., Tversky, A. (1982), *Judgment Under Uncertainty: Heuristics and Biases*, New York, Cambridge University Press.
- Kahneman, D., Tversky, A. (1979), "Prospect theory: An analysis of decisions under risk", *Econometrica*, 47(2), pp. 263–291.
- Lakoff, G. (2004). *Don't think of an Elephant! Know your values and frame the debate*, Chelsea Green Publishing Co, (*Non pensare all'elefante!*, trad. it. Di B. Tortorella, Roma, Fusi Orari, 2006).
- Lakoff, G., Johnson, M. (1980), *Metaphors we live by*, Chicago, University Press, (*Metafora e vita quotidiana*, trad. it. di P. Violi (a cura di), Milano, Bompiani, 1998).

- Magnani, L., Carnielli, W., Pizzi, C. (a cura di) (2010), "Model-Based Reasoning in Science and Technology Abduction, Logic, and Computational Discovery", Collana *Studies in Computational Intelligence*, Vol. 314, Heidelberg/Berlino, Springer.
- Manktelow, K. (1999), *Reasoning and Thinking, Cognitive Psychology, A Modular Course*, Hove, Psychology Press.
- Minsky, M. (1975), "A framework for representing knowledge", in P. Winston, ed., *The Psychology of Computer Vision*, New York, McGraw-Hill, pp. 211-277.
- Pessoa, L. (2008), "On the relationship between emotion and cognition", *Natural Reviews Neuroscience*, 9(2), 148-58.
- Piaget, J. (1956), *Logic and Psychology*, Manchester University Press, (*Logica e psicologia*, trad. it. di A. V. Visalberghi, Firenze, La Nuova Italia, 1969).
- Sperber, D., Wilson, D. (1986), *Relevance: Communication and Cognition*, Oxford, Basil Blackwell.
- Rips, L. J. (1994), *The Psychology of Proofs*, Cambridge (MA), MIT Press.
- Wilkins, M. C. (1928), "The Effect of Changed Material on the Ability to do Formal Syllogistic Reasoning", *Archives of Psychology*, 102.
- Woodworth, R. S., Sells, S. B. (1935), "An Atmosphere Effect in Syllogistic Reasoning", *Journal of Experimental Psychology*, 18, 451-60 ("Un effetto atmosfera nel ragionamento in forma sillogistica", trad. it. di P. Legrenzi, A. Mazzocco (a cura di), *Psicologia del Pensiero*, 1975, pp. 294-325).
- Wason, P. C. (1960), "On the Failure to Eliminate Hypotheses in a Conceptual Task", *The Quarterly Journal of Experimental Psychology*, 12, 129-40.
- Wason, P. C. (1966), «Reasoning», in B. Foss (ed.), *New Horizons in Psychology*, Harmondsworth, Penguin, pp. 135-51.

Wason, P. C. (1983), «Realism and Rationality in the Selection Task», in J. St. B. T. Evans (ed.), *Thinking and Reasoning*, Routledge, London, pp. 45-75.

Zorzi, M., Girotto, V. (2004), *Fondamenti di psicologia generale*, Bologna, Il Mulino.

La “casa della ragione”

Razionalità e irrazionalità nei contesti conversazionali

Francesca Ervas
Università di Cagliari
ervas@unica.it

La ragione è un’isola piccolissima nell’oceano dell’irrazionale.
(Immanuel Kant)

1. Una “casa della ragione”

Nei dialoghi di tutti i giorni cerchiamo di capire le parole dei nostri interlocutori che si presentano a volte come incoerenti, di difficile interpretazione. Capita spesso che le persone con cui conversiamo ci appaiano irrazionali, proprio a motivo di ciò che stanno dicendo. Che fare allora? Rinunciare a comprenderle o cercare di riportare quanto dicono all’interno di un discorso che possa sembrarci razionale? Alcuni filosofi del linguaggio risponderebbero a queste domande dicendo che siamo naturalmente tentati di rendere razionali i nostri interlocutori, ai nostri occhi, applicando alcuni “requisiti minimi” di razionalità che ci permettano di comprenderli. Quali sono i “requisiti minimi” di razionalità che potremmo condividere in un qualsiasi contesto conversazionale? E, prima ancora, ci sono degli *standard* di razionalità che ci permettano di stabilire il significato del messaggio comunicato dal parlante? In questo saggio si cercherà di mostrare alcuni interessanti tentativi di risposta a queste domande, per capire che cos’è la razionalità in contesti in cui si fa uso di una capacità specificamente umana, il linguaggio, per comunicare con i nostri simili.

A questo proposito Donald Davidson ha spesso messo in luce l’esistenza di uno spazio concettuale, tacitamente e intersoggettivamente condiviso, che coincide con alcuni fondamentali *standard* di razionalità¹. Tuttavia, non esiste

Il presente articolo è stato prodotto durante l’attività di ricerca di Francesca Ervas, finanziata con le risorse del P.O.R. SARDEGNA F.S.E. 2007-2013 – Obiettivo competitività regionale

Francesca Ervas, “La “casa della ragione”. Razionalità e irrazionalità nei contesti conversazionali”, in Pierluigi Graziani, Giorgio Grimaldi, Massimo Sangoi, *Animali razionali. Studi sui confini e sulle possibilità della razionalità*, pp. 47-68
© 2015 Isonomia, Rivista online di Filosofia – Epistemologica – ISSN 2037-4348
Università degli Studi di Urbino Carlo Bo
<http://isonomia.uniurb.it/epistemologica>

a suo parere una lista di principi basilari della razionalità che i parlanti possano decidere di accettare o meno, ma esiste piuttosto uno sfondo di razionalità “in un senso originario”, come «*condizione* per avere dei pensieri in assoluto»². Tuttavia lo stesso Davidson nota che, se la razionalità è la condizione del pensiero, ci si ritrova invischiati nel “paradosso dell'irrazionalità”: l'irrazionalità infatti – appartenendo al pensiero umano – non può sottrarsi allo sfondo di razionalità che ne è condizione. Con le parole di Davidson, il “paradosso dell'irrazionalità” sorge perché «l'irrazionale non è meramente il non-razionale, che giace al di fuori dell'ambito del razionale; l'irrazionalità è un fallimento all'interno della *casa della ragione*»³. Si tratta allora di capire come sia fatta la “casa della ragione”, secondo quale architettura e quali siano le staccionate che ne delimitano i confini. La domanda filosoficamente più interessante è dunque quella sulla natura della razionalità umana e, insieme, sulla natura dell'irrazionalità, che nella prospettiva davidsoniana sembrerebbe essere un “fallimento”.

La colonna portante della “casa della ragione” davidsoniana è il Principio di Carità, il principio *costitutivo* negando il quale si arriverebbe a «distuggere il fondamento di intelligibilità su cui riposa ogni interpretazione»⁴. Il Principio di Carità è *condizione di possibilità* dell'interpretazione e si basa su tre assunzioni⁵: a) che ci sia un *accordo* di fondo con il parlante; b) che il parlante abbia una rete di credenze per la maggior parte *coerenti e vere*; c) che il parlante sia *razionale*⁶. Il concetto di razionalità proposto da Davidson è dunque molto pretenzioso: che il nostro interlocutore debba essere coerente, avere credenze vere e in accordo con la maggior parte delle nostre pare restringere eccessivamente i requisiti necessari a descrivere il comportamento razionale dei soggetti in certe circostanze. Il Principio di Carità così definito non è in

e occupazione, Asse IV Capitale umano, Linea di Attività 1.3.1 “Avviso di chiamata per il finanziamento di Assegni di Ricerca”. Si desidera ringraziare Chiara Chelini e Roberto Cordeschi per aver discusso più volte il tema della “razionalità limitata”.

¹ Per un'introduzione al pensiero di Donald Davidson, cfr. C. Amoretti & Ervas, (2011).

² Davidson (2004: 189-198).

³ Davidson (2004: 169-187).

⁴ Davidson (1990: 279-328). Per un'analisi del principio di carità come *costitutivo*, cfr. McGinn (1977: 521-535); Ramberg (1989: 73-75); Spati (1994); Longato (1996: 91).

⁵ Per una critica delle tre formulazioni del Principio di Carità davidsoniano, si rimanda a Ervas (2008).

⁶ Nei diversi saggi in cui Davidson ha trattato quest'argomento, egli sottolinea o l'uno o l'altro aspetto, ma nella maggior parte dei casi dà una formulazione del Principio di Carità in cui vengono collegati due o più di essi. Cfr. solo come esempio 1: Davidson (1984, trad. it. 75, 206-208, 247, 228, 280-281, 284; 2: 75, 208, 247, 227-228, 281; 3: 208, 236, 238).

grado per esempio di guidare le scelte dell'interprete radicale nei singoli casi di interpretazione: non spiega i modi per identificare e valutare le credenze che possono essere considerate direttamente rilevanti per comprendere questa o quest'altra espressione di un parlante, quelle errate, quelle in accordo con il punto di vista dell'interprete, ecc. Queste assunzioni sembrano ancora più sconcertanti qualora si pensi che la razionalità e la coerenza del parlante, così come l'accordo fra un ampio insieme di credenze vere, vengono decisi dall'interprete. A parere di Davidson, gli *standard* di razionalità sono infatti tali *dal punto di vista dell'interprete*: «Se non riusciamo a trovare un modo per interpretare i proferimenti e altri comportamenti di un essere come rivelatori di un insieme di credenze ampiamente coerenti e vere *secondo i nostri criteri*, non abbiamo motivo di considerare *razionale* quell'essere, né di considerarlo in grado di avere credenze o di dire alcunché»⁷. In molti passi come quello appena citato⁸, Davidson sembra quindi concludere che i discorsi che l'interprete non riesce a riportare entro i confini imposti dal Principio di Carità non sono discorsi, sono suoni disarticolati e si potrebbe persino dubitare che chi li ha pronunciati sia un essere umano.

Ian Hacking ha pesantemente criticato Davidson a tal proposito, sostenendo che l'accordo tacito presupposto dal Principio di Carità riguarda credenze che possono servire solo all'interpretazione dei proferimenti più banali e grossolani dei parlanti. Non solo, a suo parere il Principio di Carità sarebbe lo strumento principale di un certo "imperialismo linguistico" che:

Ha un arsenale più fornito di quello militare, in quanto, tramite un ragionamento trascendentale, si potrebbe magari provare che se l'indigeno non condivide le nostre preferenze e i nostri desideri, molto semplicemente egli non partecipa alla razionalità umana ed è, nel migliore dei casi, un subumano (e per l'indigeno neanche questa è una novità)⁹.

L'obiezione di Hacking non coglie tuttavia nel segno perché l'accordo presupposto da Davidson va inteso come qualcosa di più basilare delle credenze sul vissuto quotidiano, qualcosa di più originario, ovvero uno sfondo comune che ci permette di identificare anche le inevitabili incomprensioni. Davidson prevede infatti che ci possa essere un *marginale di errore* nell'applicazione del punto di vista dell'interprete, ma questo disaccordo è comprensibile solo sullo sfondo del massiccio accordo di base presupposto dal Principio di Carità: «Il massimo grado di comprensione delle parole e dei pensieri altrui si raggiunge quando interpretiamo in modo da ottimizzare l'accordo (ciò non

⁷ Davidson (1973, trad. it. 193-211).

⁸ Ervas (2008).

⁹ Hacking (1975, trad. it. 182-183).

esclude, come abbiamo detto, l'*errore spiegabile*, cioè le *differenze di opinioni*)»¹⁰. Davidson vuole piuttosto escludere una situazione di *errore su larga scala*, perché un tale disaccordo non sarebbe comprensibile all'interprete ed eliminerebbe qualsiasi possibilità di dialogo con il parlante.

2. Razionalità e irrazionalità in Davidson

Che cos'è dunque questo sfondo comune che ci permette di cogliere anche le incomprensioni con il nostro interlocutore? Che cosa ci fa identificare il nostro interlocutore come razionale sebbene con opinioni divergenti? Rispondere a queste domande significherebbe cercare uno sfondo anche per la "casa della ragione", ma, come nota Davidson, ciò equivale ad interrogarci sull'emergere del pensiero, se si ritiene che il pensiero sia esclusivamente quello razionale. In particolare, nel caso di Davidson, la questione è ancor più complicata, perché si assume che pensiero e linguaggio articolato emergano insieme e manca perciò un vocabolario che descriva lo stadio precedente all'emergere del pensiero¹¹. Ci troviamo dunque di fronte a una *impasse* nella descrizione dell'emergere del pensiero razionale:

La difficoltà incontrata nel descrivere l'emergere dei fenomeni mentali è di natura concettuale: è la difficoltà di descrivere i primi stadi nella maturazione della ragione, gli stadi che precedono la fase in cui i concetti come intenzione, credenza e desiderio hanno chiara applicazione. Nell'evoluzione del pensiero, nella storia sia dell'umanità sia di un individuo, vi è uno stadio in cui non c'è pensiero, e subito dopo uno in cui vi sono pensieri. Descrivere l'emergere del pensiero vorrebbe dire descrivere il processo che conduce dal primo al secondo di questi stadi. Ci manca un vocabolario soddisfacente per descrivere i passaggi intermedi¹².

Ci si potrebbe chiedere allora se lo sfondo da cui emerge il discorso razionale non sia qualcosa di totalmente altro dalla razionalità: l'irrazionalità. Tuttavia non sembra essere questa la proposta davidsoniana: l'irrazionalità, infatti, può essere compresa solamente se rimane all'interno dei confini della razionalità: «Solo un essere razionale può essere irrazionale. L'irrazionalità è uno stato o un processo mentale – uno stato o un processo razionale – fallito»¹³. In questo senso, l'irrazionalità genera un paradosso: se si cerca di spiegarla, si "razionalizza". Se invece l'irrazionalità sta al di fuori della "casa

¹⁰ Davidson (1974, trad. it. 281).

¹¹ Davidson (2001, trad. it. 121-135).

¹² Davidson (2001, trad. it. 162).

¹³ Davidson (2004: 169).

della ragione”, «compromettiamo semplicemente la nostra capacità di diagnosticare l'irrazionalità, ritirando quello sfondo di razionalità che serve a giustificare qualsiasi diagnosi»¹⁴.

Se pensiero razionale e linguaggio articolato emergono insieme, si potrebbe allora ipotizzare che il pensiero irrazionale preceda l'avvento del linguaggio:

C'è una situazione prelinguistica e precognitiva che mi sembra costituire *una condizione necessaria* del pensiero e del linguaggio, una condizione che può esistere indipendentemente dal pensiero, e che dunque può precederlo. Sia nel caso degli animali non umani, sia in quello dei bambini piccoli, è una condizione che può essere osservata.¹⁵

Ci sarebbe dunque uno stadio preverbale della vita umana in cui non c'è ancora pensiero razionale. Tuttavia questa soluzione al paradosso dell'irrazionalità presenta alcune difficoltà. Innanzitutto, nel primo anno di vita gli esseri umani non possiedono un linguaggio articolato, ma i bambini piccoli hanno la possibilità di imparare un linguaggio articolato e di avere dunque un pensiero verbale. Da un punto di vista filosofico tuttavia, non sembra adeguato descrivere la natura di qualcosa (in questo caso l'irrazionalità del primo anno di vita), semplicemente constatando che ha la *possibilità di essere qualcos'altro*. Inoltre un bambino si trova nella situazione radicale di non possedere alcuna lingua e di doverne imparare una: dunque, se impara una prima lingua, *come può impararla se ancora non pensa?* Per Davidson infatti, la capacità di pensare è strettamente vincolata a quella di parlare. Dovremmo dunque concludere che il pensiero non sia solo quello che abita la "casa della ragione", ma che esista anche un pensiero "irrazionale", indispensabile anche all'apprendimento del linguaggio. Non esiste dunque – come sostiene Davidson – un passaggio tra uno stadio in cui il bambino non pensa a uno stadio in cui il bambino pensa e parla, quanto piuttosto un passaggio tra tipi diversi di pensiero, una lenta trasformazione da pensiero irrazionale a pensiero razionale che avviene quando un bambino impara la sua prima lingua. Non è questa purtroppo la soluzione di Davidson, che resta imbrigliato nell'impasse teorica dell'emergenza del pensiero: «il pensiero proposizionale richiede il linguaggio, il linguaggio richiede il pensiero proposizionale»¹⁶.

La razionalità descritta come principio *costitutivo* del pensiero e del linguaggio non è al centro dell'interesse solo della filosofia davidsoniana o di quella di stile analitico. Il pensiero occidentale ha proposto più volte questo

¹⁴ Davidson (2004: 184).

¹⁵ Davidson (2001: 163-164).

¹⁶ Davidson (1999: 43).

paradigma che si ritrova anche in filosofi “continentali”, come per esempio Hans-Georg Gadamer. Per Gadamer il linguaggio «è essenzialmente il linguaggio del dialogo» e «costituisce la propria realtà solo attraverso l’attuarsi effettivo della comunicazione»¹⁷. Questa comune eredità proviene dalla dialettica e dal dialogo platonici:

Nei greci si poteva imparare che il pensare della filosofia non deve seguire da un sistematico pensiero guida, o da una fondazione ultima nella forma di un principio supremo, che possa dar ragione di tutto, ma che esso ha già una guida: esso, nel continuare a pensare l’originaria esperienza del mondo, deve pensare a fondo la forza concettuale del linguaggio in cui viviamo. Voler insegnare questo è secondo me il segreto del dialogo platonico.¹⁸

Sia Davidson che Gadamer sono partiti dal *Filebo* di Platone e, seguendo due percorsi e due diversi stili di indagine filosofica, han finito per riconoscere il nucleo comune alle loro teorie dell’interpretazione¹⁹. Davidson, infatti, a partire dal saggio *Una graziosa confusione di epitaffi* (1986)²⁰, descriverà la comprensione altrui in modo molto simile a una gadameriana “fusione di orizzonti”, come un continuo aggiustamento del punto di vista dell’interprete a seconda del contesto conversazionale.

3. Razionalità e irrazionalità in Grice

Altri filosofi hanno invece parlato della razionalità come principio *regolativo*²¹. È questo il caso di Paul Grice²², secondo il quale la comunicazione è un processo inferenziale in cui l’ascoltatore deve ricostruire il significato del parlante basandosi non solo su ciò che il parlante ha detto ma anche su ciò che aveva intenzione di dire. Il significato di un enunciato è quindi il *significato del parlante*, cioè quel significato che il parlante *intende* comunicare a un ascoltatore nel proferire un dato enunciato in una determinata occasione. Gli atti comunicativi individuali dipendono infatti dalle intenzioni complesse dei parlanti volte a produrre vari stati mentali nei loro ascoltatori. Il parlante cerca di produrre un certo effetto sull’ascoltatore tramite il riconoscimento da parte di quest’ultimo della sua intenzione. Grice mette così in crisi le teorie semantiche che spiegano il significato dei proferimenti in termini di verità o

¹⁷ Gadamer (1960, trad. it. 909).

¹⁸ Gadamer (1993, trad. it. 463).

¹⁹ Davidson (2005: 261-275); Gadamer (1997); Hahn (1997: 433-435).

²⁰ Davidson (1986, trad. it. 59-85).

²¹ Levinson (1983).

²² Per un’introduzione al pensiero di Paul Grice, cfr. Labinaz (2012).

falsità degli enunciati. La conversazione eccede infatti questo tipo di analisi non appena si consideri come il significato convenzionale degli enunciati cambi a seconda del contesto conversazionale in cui compaiono²³.

Secondo Grice, nell'inferire il significato del parlante a partire dalla decodifica del suo proferimento e dall'informazione contestuale disponibile, l'ascoltatore è guidato dall'aspettativa che il proferimento soddisfi alcuni *standard* di razionalità. Nella prospettiva griceana, la conversazione è una forma di comportamento razionale e cooperativo guidato dal seguente principio: «Il tuo contributo alla conversazione sia tale quale è richiesto, allo stadio in cui avviene, dallo scopo o orientamento accettato dello scambio linguistico in cui sei impegnato»²⁴. Si tratta del Principio di Cooperazione che esclude alcune possibilità conversazionali come improprie, attraverso la ben nota serie di massime conversazionali di quantità, qualità, relazione e modo²⁵. Le massime spiegano come i parlanti possano dire una cosa ed essere correttamente compresi come se ne dicessero un'altra. A differenza del Principio di Carità davidsoniano, che non può essere violato in quanto condizione di intelligibilità, Grice osserva invece che chi partecipa a uno scambio comunicativo può fallire nel soddisfare una massima in vari modi: può violare il principio generale di cooperazione, può uscire dal raggio d'azione di una massima, può trovarsi davanti a un conflitto fra massime, può intenzionalmente "burlarsi" di una massima, ecc. Quando una persona viola il principio di cooperazione o una sua massima in uno di questi modi, la comunicazione non si interrompe. Essendo la comunicazione una forma di comportamento razionale, capiremo che quella persona vuole probabilmente veicolare un altro significato, che non è quello letterale, ma appunto il "significato del parlante"²⁶.

Quando l'interprete cerca una ragione alternativa a quella offerta dal significato letterale, dà adito a una implicatura conversazionale, ovvero a un aspetto del significato suggerito, inteso, implicato e mai esplicitamente detto. L'interprete "ragionerà" seguendo altre vie per riuscire a cogliere l'intenzione sottesa ai proferimenti espliciti del parlante. Infatti, come notato:

Sia che ci riferiamo alla capacità degli esseri umani di produrre significati attraverso la loro intenzione di comunicare qualcosa a qualcuno, sia che prendiamo in

²³ Atlas (2005); Avramides (1989); Bianchi (2003).

²⁴ Grice (2003: 229).

²⁵ I partecipanti alla conversazione si aspettano infatti che un interlocutore cooperativo dia un contributo né più né meno informativo di quanto è richiesto ("massima della quantità"), non dica ciò che ritiene falso o di cui non ha prove sufficienti ("massima della qualità"), dica qualcosa di pertinente ("massima della relazione") e si esprima in forma chiara, concisa e ordinata ("massima del modo"). Cfr. Grice (2003).

²⁶ Bianchi (2003).

esame la capacità dell'uditorio di ricostruire le intenzioni dei parlanti, le analisi proposte da Grice mostrano che produrre e comprendere significati sono attività guidate da ragioni.²⁷

Il desiderio di trovare delle ragioni per i propri contributi conversazionali è ciò che mette in luce la razionalità degli esseri umani. È il caso delle implicature conversazionali, che devono poter essere “calcolabili” in quanto, per essere attribuite a un parlante razionale, devono risultare “giustificabili” in un dato contesto conversazionale. Se si fa uso di proprietà molto generali del contesto, si ha un'implicatura conversazionale generalizzata, ovvero tipicamente associata alla proposizione espressa, valida quindi in lingue diverse. Per esempio, il proferimento “Giada ha avuto un bambino *e* si è sposata” ha una lettura temporale (*e*, per i più maliziosi, anche causale) così come nel suo equivalente francese “Giada a eu un bébé *et* s'est mariée”: prima Guada ha avuto un bambino e poi si è sposata. Se invece si fa uso di proprietà particolari del contesto, si ha un'implicatura conversazionale particularizzata. Per esempio, se il parlante A chiede a B: “Ci sarà Sally all'incontro di questo pomeriggio?” e il parlante B risponde “La sua macchina si è rotta”, il parlante A inferirà che Sally non verrà all'incontro. Per la non-distaccabilità delle implicature conversazionali particularizzate, questo tipo di implicature si preservano nella traduzione, se si preserva il tipo di contesto conversazionale. In entrambi i casi, comunque, quando un parlante viola una massima, il suo comportamento viene reinterpretato, con l'attribuzione di un significato non-letterale, che dipende da ciò che quella violazione implica nel contesto in cui è avvenuta²⁸.

L'innovativo contributo che Grice ha offerto relativamente alla nozione di “significato del parlante” va inquadrato dunque nel suo più ampio progetto di costruire una teoria dell'essere razionale, cioè una teoria che definisca ciò che contraddistingue gli esseri che possiedono come proprietà essenziale la razionalità²⁹. Secondo Grice, la razionalità degli esseri umani si manifesta proprio nel loro desiderio di motivare i propri comportamenti, siano essi di tipo linguistico o non linguistico. Gli esseri umani sono razionali perché tendono a cercare le *ragioni* del proprio e dell'altrui agire. Ciò non significa che riescano sempre a giustificare il loro comportamento e che lo facciano in maniera corretta. Come in Davidson, Grice lascia la possibilità di un margine di errore. Tuttavia, a differenza di quanto risulta dai requisiti posti da Davidson, gli esseri umani non sono necessariamente razionali. Lo diventano quando

²⁷ Labinaz (2012: 25-26).

²⁸ Ervas (2011: 53-64).

²⁹ Grice (1991); Grice (2001).

riconoscono nella razionalità la loro proprietà essenziale e nella misura in cui dimostrano un desiderio costante, non occasionale, di ricercare giustificazioni per i comportamenti propri e altrui. Si tenta così di raggiungere la “casa della ragione”, si desidera tracciare i suoi confini e la sua architettura.

In *Relevance: Communication and Cognition* (1986)³⁰, Dan Sperber e Deirdre Wilson criticano proprio l’idea griceana di razionalità come ideale regolativo, che di fatto non fornirebbe alcun criterio di identificazione per selezionare, all’interno di un insieme potenzialmente infinito di premesse presenti nel contesto conversazionale, un numero finito di premesse sulle quali basare l’inferenza da ciò che il parlante enuncia a ciò che egli intende. Inoltre, le regole di inferenza non vengono mai veramente esplicitate nel processo comunicativo: quando comunichiamo non facciamo quasi mai affidamento sulle lunghe catene inferenziali previste da Grice. Sperber e Wilson si interrogano infatti sulla *plausibilità psicologica* di alcune tra le più interessanti ipotesi di Grice, mettendo letteralmente “alla prova dei fatti” la sua teoria³¹. Cosa succede quando si passa dai modelli teorici a come i parlanti di fatto comunicano e si comprendono nei reali contesti conversazionali? La teoria della pertinenza vuole rispondere precisamente a questa domanda.

4. La razionalità dei parlanti nella teoria della pertinenza

Il modello costruito da Sperber e Wilson, che ha l’obiettivo di essere in primo luogo psicologicamente plausibile, è guidato da un principio cognitivo: il principio di pertinenza. Due sono le nozioni che definiscono il principio di pertinenza: la nozione di *effetto cognitivo* e la nozione di *sforzo di elaborazione*. Ad esempio, un proferimento verbale è pertinente se produce un effetto cognitivo positivo, che modifica una qualche rappresentazione del mondo. Tanto maggiore sarà la pertinenza quanto maggiori saranno gli effetti cognitivi ottenuti e minori gli sforzi di elaborazione richiesti rispetto a un dato contesto conversazionale. Il costo o lo sforzo cognitivo è giustificato dall’effetto cognitivo ottenuto tramite il processo inferenziale. Sperber e Wilson propongono una procedura di comprensione guidata dalla pertinenza: nel calcolare gli effetti cognitivi, l’interprete dovrebbe seguire un percorso che minimizza lo sforzo cognitivo, controllare le conclusioni (vale a dire le ipotesi interpretative dell’espressione che il parlante ha trasmesso) in ordine di accessibilità e fermarsi una volta soddisfatte le aspettative di pertinenza. In altri termini, l’interlocutore prende le mosse dal significato decodificato linguisticamente

³⁰ Sperber & Wilson (1986).

³¹ Noveck & Sperber (2004).

ma lo arricchisce con informazioni contestuali finché l'interpretazione non soddisfa la sua aspettativa di pertinenza³².

Se per Grice la cooperazione comunicativa e le massime si rendono necessarie soprattutto al livello implicito, là dove si profila un'implicatura conversazionale, secondo Sperber e Wilson il processo di comprensione inferenziale guidato dalla pertinenza è onnipresente e onnipervasivo: riguarda cioè qualsiasi atto comunicativo e va esteso anche al livello esplicito, là dove si dà un'esplicatura³³. La lingua vive sempre in condizioni di indigenza: almeno sotto il profilo di ciò che è lessicalizzato, c'è una sproporzione ineliminabile tra ciò che è esprimibile letteralmente e ciò che gli utenti della lingua possono aver bisogno di esprimere. È per questo che soltanto grazie a dei processi pragmatici di modulazione lessicale ("*narrowing*" e "*broadening*") avviene il passaggio dalla forma logica del proferimento al significato esplicito³⁴. La forma logica è una rappresentazione concettuale incompleta, risultato di un modulo *semantico* di decodifica linguistica, e costituisce l'input di un modulo *pragmatico* dal quale ricaviamo dei concetti *ad hoc* o occasionali, legati a un determinato contesto e a un determinato compito. Questa spiegazione del significato esplicito dell'enunciato permetterebbe di abbandonare la distinzione tradizionale tra usi letterali e usi figurati del linguaggio, perché la stessa "letteralità" non è altro che il risultato di un processo di costruzione di un concetto *ad hoc*. La comprensione non richiede necessariamente che parlante e destinatario condividano lo stesso concetto *ad hoc*: è sufficiente una somiglianza interpretativa, ovvero una parziale coincidenza di proprietà logiche ed enciclopediche dei concetti *ad hoc*.

Il fenomeno della sottodeterminazione semantica non può essere colmato con la sola risoluzione degli elementi di ambiguità o di indicialità, siano essi espressi in superficie come sostengono i minimalisti, o nascosti nella forma logica della frase come vorrebbero gli indicalisti. È più economico ipotizzare un solo sistema responsabile sia della proposizione completa sia della derivazione delle implicature conversazionali.

A tale scopo basta assegnare alla proposizione espressa alla proposizione espressa delle condizioni di verità intuitive, rispettando le intuizioni semantiche dei parlanti: comprendere un enunciato dichiarativo significa sapere in quali circostanze concrete esso sarebbe vero. Risultati sperimentali confermano inoltre che i tempi di elaborazione di enunciati contenenti usi figurati del linguaggio, come le metafore, non sono più lunghi di quelli dovuti alla

³² Sperber & Wilson (1986).

³³ Bianchi (2009).

³⁴ Carston (2002).

lettura e alla comprensione di enunciati letterali³⁵. Non ci sarebbe dunque un’elaborazione di default che parte dal significato letterale per arrivare al significato implicito dato dall’implicatura conversazionale: l’elaborazione pragmatica avverrebbe già al livello dell’esplicito, contribuendo a mettere in crisi la nozione stessa di “letteralità”, di cui non si era completamente sbarazzato nemmeno Grice³⁶.

È bene notare tuttavia che Grice intendeva proporre un metodo per la ricostruzione razionale dei percorsi di comprensione³⁷, per spiegare *perché* i parlanti si comportano in un certo modo nei contesti conversazionali, mentre Sperber e Wilson intendono piuttosto capire *come* di fatto l’interprete comprende il significato del parlante. I percorsi argomentativi proposti da Grice non hanno infatti l’ambizione di rispecchiare i processi psicologici effettivi di comprensione: i due progetti, processamento psicologico e giustificazione ideale o razionalizzazione, devono essere tenuti distinti. Va riconosciuto il carattere normativo e non psicologico delle implicature griceane, che si rifanno a una concezione *argomentativa* della razionalità (riguardante i *fini* dei nostri comportamenti linguistici o meno), più che una concezione *strumentale* (riguardante i *mezzi* utilizzati dal nostro apparato cognitivo), più propria dei teorici della pertinenza. Per i teorici della pertinenza la “casa della ragione” non è un ideale regolativo, è invece costruita con determinati materiali e strumenti ed è provvista di determinate strutture.

La teoria della pertinenza presenta tuttavia dei limiti³⁸. Innanzitutto non solo sembra molto difficile misurare la pertinenza degli stimoli, ma sembra poco plausibile che ci sia un solo modo per soddisfarla. Inoltre, la teoria della pertinenza sembra concentrarsi esclusivamente sugli effetti in termini informativi o cognitivi, senza considerare altri tipi di benefici – sociali, o in termini di cortesia – che sono cruciali nelle nostre interazioni comunicative. Per risolvere questi problemi, i teorici della pertinenza dovrebbero spiegare innanzitutto come avvenga il calcolo dei benefici a fronte dei costi cognitivi, calcolo costoso da un punto di vista delle risorse cognitive. Inoltre, i due principi – cognitivo e comunicativo – attraverso cui viene formulata la nozione di pertinenza sembrerebbero ineludibili, impossibili da violare. Ciò renderebbe infalsificabile la teoria della pertinenza: un grave limite per una teoria che pretende di essere empiricamente controllabile. Si rimprovera inoltre alla teoria della pertinenza, e più in generale al contestualismo, di utilizzare lo stesso genere di soluzione per fenomeni linguistici molto diversi tra loro, limitandosi

³⁵ Bambini & Resta (2012: 37-60).

³⁶ Bianchi (2009).

³⁷ Sbisà (2007).

³⁸ Bianchi (2009).

a fornire una spiegazione non di *tipi* di enunciati, ma di *occorrenze* di enunciati, con una conseguente perdita di potere esplicativo. Rimane poi ancora controverso il concetto stesso di inferenza, estesa a tutti i fenomeni pragmatici nella teoria della pertinenza. La soluzione proposta dai pertinentisti è l'esistenza di un modulo specificamente dedicato alla lettura delle intenzioni dei parlanti, ovvero la Teoria della Mente, ma resta comunque ancora da spiegare come operi la Teoria della Mente al proprio interno nella differenziazione di un così vasto insieme di compiti pragmatici³⁹. Un tentativo in questo senso è rappresentato dai recenti studi di Olivier Mascarò e Dan Sperber sulla vigilanza epistemica⁴⁰, che dimostrano che la Teoria della Mente è solo una parte di una più ampia capacità di filtrare le informazioni escludendo quelle false o "sbagliate" dai contenuti comunicati, insieme a una capacità metalogica e a una capacità morale-affettiva⁴¹.

5. Una "casa a misura d'uomo": la razionalità limitata

Il modello di razionalità alla base della procedura di calcolo proposta dai teorici della pertinenza è quello economico del miglior rapporto tra costi e benefici nel contesto conversazionale. Tale modello del processo decisionale, prevalente in economia classica e teoria dei giochi, è quello di un agente che prende la decisione migliore tra le infinite possibili. L'agire umano di tutti i giorni tuttavia non è quasi mai così perfetto e serve invece trovare una procedura di comprensione più realistica e allo stesso tempo "tollerabilmente buona", ovvero una strategia *selettiva* buona pur non essendo sempre la migliore possibile⁴². Una risposta a questa domanda potrebbe venire dal pensiero di Herbert Simon, che, a partire dal testo *Administrative Behavior* (1946)⁴³, cerca di dare una spiegazione del procedimento razionale di un agente umano reale nella sua relazione con l'ambiente esterno e con le proprie capacità interne. Secondo Simon, la razionalità umana è per propria natura doppiamente limitata: tale agente è infatti condizionato sia dai propri limiti cognitivi interni (es. la propria memoria, la propria capacità di elaborazione) sia dalla complessità dei compiti a cui si deve far fronte per interagire con l'ambiente esterno (anche e soprattutto con l'ambiente sociale)⁴⁴. Di fronte a tali limiti,

³⁹ Marraffa & Meini (2005).

⁴⁰ Mascarò & Sperber (2009: 367-380).

⁴¹ Ervas (2012: 28-33).

⁴² Shannon (1950: 256-275).

⁴³ Simon (1946).

⁴⁴ Cordeschi (2011: 197-215).

le decisioni razionali umane assomigliano più a quelle di un manager impegnato a risolvere problemi contingenti (*administrative man*) rispetto a quelle dell'onnisciente e idealizzato *Homo oeconomicus* dell'economia classica⁴⁵.

Non si deve pensare alla razionalità limitata come all'ottimizzazione dell'utile in un dato contesto o come il fallimento a cui le persone vanno incontro nel seguire i loro ideali normativi. Piuttosto la razionalità limitata è la chiave per capire come le persone prendano decisioni senza ottimizzare l'utile e consiste in una strategia che passo dopo passo guida l'agente umano sottoposto a limiti di tempo, conoscenze e capacità computazionali, ove una strategia ottimale sia presente o meno. Come spiega Paolo Labinaz:

Mentre la teoria della decisione razionale prevede che i decisori prendano in considerazione simultaneamente tutte le alternative e gli esiti a loro associati al fine di individuare l'opzione ottimale, Simon sostiene che gli esseri umani, in quanto limitatamente razionali, esplorano le alternative e i loro esiti in maniera sequenziale, non potendo contare fin dall'inizio su una conoscenza completa di quelle che sono le alternative effettivamente disponibili⁴⁶.

I limiti interni e i limiti imposti dalla complessità dell'ambiente esterno forzano un agente reale a usare nella scelta interpretativa non la strategia idealmente migliore, ma diverse strategie, o *euristiche*, che siano abbastanza soddisfacenti, o per essere più precisi: "satisficing"⁴⁷. Questo aggettivo è stato coniato da Simon per alludere al fatto che «un decisore, invece di andare alla ricerca di un'ottimizzazione spesso impossibile, sceglie tra le alternative disponibili quella che per prima soddisferà il suo livello di aspirazione»⁴⁸. Questo procedimento euristico riduce la complessità computazionale compatibilmente con una capacità di elaborazione ragionevole.

La "casa della ragione" non ha dunque confini ideali predeterminati, ma viene costruita a misura d'uomo. Si potrebbe aggiungere che la razionalità limitata è una casa completamente immersa nella natura. Newell e Simon descrivono infatti gli esseri umani come sistemi *adattivi*⁴⁹. Come ha suggerito Roberto Cordeschi⁵⁰, mentre l'ambiente esterno non cambia dal punto di vista oggettivo dell'osservatore, esso cambia costantemente dal punto di vista soggettivo dell'agente umano che si adatta con nuove soluzioni a seconda delle sue possibilità percettive e cognitive, a seconda dei propri limiti interni. Ciò

⁴⁵ Simon (1956: 129-138); Simon (1983).

⁴⁶ Labinaz (2013: 63).

⁴⁷ Simon (2000: 243-253).

⁴⁸ Labinaz (2013: 64).

⁴⁹ Newell & Simon (1972).

⁵⁰ Cordeschi (2008: 219-258).

che sembra più sorprendente è la capacità umana di rapportarsi all'ambiente esterno quando entrano in scena altri esseri umani⁵¹. La cooperazione è stata infatti spesso definita come un enigma evolutivo, proprio perché contraddice l'ipotesi dell'egoistica ottimizzazione dell'utile come regola di sopravvivenza⁵². La teoria classica della scelta razionale non ha una vera e propria risposta a tale riguardo e sono state avanzate soluzioni alternative, che si riferiscono alla capacità umana di cogliere caratteristiche salienti di un contesto conversazionale o di estrarre una regola che sembri essere mentalmente disponibile ai propri interlocutori per potersi coordinare e cooperare⁵³. Le persone rispondono comunque in modi differenti in relazione al contesto e considerano, nelle loro scelte, valori che contraddicono l'ipotesi dell'ottimizzazione. Come notato, «la teoria della razionalità limitata è legata alle prospettive pragmatiche dell'intelligenza sociale che enfatizzano il bisogno di valutare le competenze cognitive nei contesti di vita appropriati in cui quelle competenze si sviluppano»⁵⁴.

Di fronte all'evidenza che le persone di fatto cooperano, si può ipotizzare che esse anticipino le credenze, intenzioni e desideri altrui. Se la razionalità umana è limitata, ci devono essere dei meccanismi che ci permettono di fare una scelta tra molte soluzioni alternative. Nel caso della scelta razionale che implica la cooperazione con altre persone, gli esseri umani fanno spesso tacitamente appello a certe norme sociali condivise. Il concetto simoniano di razionalità sarebbe utile a rispondere proprio alla domanda sull'emergenza delle norme sociali:

Sappiamo ancora ben poco su come [le norme sociali] si siano formate, le forze che determinano il loro contenuto, come e perché cambiano, le loro strutture cognitive ed emotive, come si collegano ai valori, come influenzano la nostra percezione della giustizia e della sua violazione, e come sono influenzate da e influenzano la nostra architettura neuropsicologica. In breve, le norme sociali sono uno dei più grandi problemi irrisolti delle scienze comportamentali.⁵⁵

Le norme e le convenzioni sociali, e più generalmente i comportamenti che implicano il rispetto di regole, possono essere considerate delle “scorciatoie cognitive” nel processo di decisione che riducono lo sforzo di elaborazione. In questo senso, gli esseri umani seguono le norme sociali proprio

⁵¹ Ervas (2011: 216-220).

⁵² Henrich & Henrich (2006: 221-245).

⁵³ Schelling (1960); Sugden & Gold (2007: 109-137).

⁵⁴ Wright & Dawson (1988: 781).

⁵⁵ Fehr & Fishbacher (2004: 64).

perché la loro razionalità è limitata. Come spiega Cristina Bicchieri⁵⁶, una norma sociale viene seguita non come risultato della razionalità classica, ma proprio perché si dà un fallimento della razionalità classica. L'agente ha bisogno delle norme sociali come sfondo che le permetta un orientamento nella complessa struttura delle sue interazioni sociali⁵⁷.

Le norme sociali si evolvono in un processo in linea con la prospettiva simoniana di adattabilità: lì dove le norme si sono dimostrate più adatte a un particolare contesto, si diffondono, mentre altre norme cadono in disuso. Dati i limiti dell'informazione imposti sia dall'ambiente esterno sia dalla mente umana, le norme sociali persistono se producono risultati “satisficing” per il gruppo sociale come intero. Come dimostrato da recenti studi di Hugo Mercier e Dan Sperber, le norme sociali si instaurano ovunque perché in questo modo i singoli riescono a risolvere problemi che da soli non riuscirebbero a risolvere⁵⁸. Ciò non significa che la stessa norma sarebbe stata selezionata in un diverso ambiente o gruppo sociale: le connotazioni soggettive di spazio e tempo giocano un ruolo fondamentale. Infatti, la variabilità culturale emerge chiaramente negli esperimenti interculturali sull'emergenza delle norme di cooperazione⁵⁹. La ricerca fatta in questo campo mostra che, se la razionalità limitata prende forme diverse, ciò accade proprio perché è limitata. La razionalità “illimitata” dell'*Homo oeconomicus* non si trova in nessun luogo.

6. Breve nota di chiusura

La struttura della “casa della ragione” è stata indagata da più prospettive per dare una risposta al problema della comprensione tra parlanti in contesti conversazionali. Se Davidson ha dato una descrizione della razionalità come condizione stessa del pensiero, come principio costitutivo del pensare, Grice ne ha fatto un ideale regolativo. Entrambi hanno però discusso degli *standard* astratti di razionalità che non danno conto dell'agire umano quotidiano, concreto e che non si confrontano con le effettive capacità cognitive umane. Le due prospettive propongono un ideale di ottimalità nei procedimenti razionali difficilmente riscontrabile nei contesti conversazionali quotidiani oppure poco plausibile da un punto di vista psicologico. I teorici della pertinenza hanno invece cercato di testare la procedura razionale che guida l'interpretazione per dare una spiegazione psicologicamente plausibile dell'agire umano.

⁵⁶ Bicchieri (2006).

⁵⁷ Smith (2008).

⁵⁸ Mercier & Sperber (2011: 57-111).

⁵⁹ Henrich, Boyd, Bowles, Camerer, Fehr, Gintis, McElreath (2001: 73-78).

Tuttavia la loro procedura di calcolo del rapporto tra sforzo di elaborazione e benefici cognitivi si rifa ancora a un concetto di ottimalità che non rispecchia i percorsi del ragionamento umano, che sono spesso incompleti o sottoposti a limiti di tempo, memoria, conoscenze, ecc., tanto da sembrare a volte “irrazionali” ai nostri occhi.

L’irrazionalità è per Davidson un fallimento all’interno della “casa della ragione” e crea un paradosso insolubile a causa dei limiti invalicabili dati dagli *standard* della razionalità classica occidentale, comune, come si è visto, anche alla filosofia “continentale”. D’altra parte, per Grice, l’irrazionalità è una condizione sempre presente dal momento che gli esseri umani aspirano alla razionalità e tendono a giustificare ciò che fanno attraverso delle ragioni che fanno capo a massime sempre eludibili. Secondo Simon, invece, «ottimalità e irrazionalità sono i due estremi di una scala in cui la razionalità limitata sta nel mezzo [...]. Questa sua duttilità sembra avvicinarla più delle sue alternative a quella che è la razionalità che la gente manifesta nella vita di tutti i giorni»⁶⁰. I casi in cui l’agire umano si distanzia dai canoni della razionalità classica, conclude Simon, «non dovrebbero essere confusi con l’idea che le persone si comportino in modo generalmente “irrazionale” [...] Di solito le persone hanno delle ragioni per fare ciò che fanno»⁶¹. In ultima analisi dunque i confini della “casa della ragione” sono determinati sia da elementi razionali che da elementi irrazionali. Tuttavia, come direbbe lo stesso Simon, «l’area della razionalità è quella dell’adattabilità a questi elementi non razionali»⁶². Kantianamente, la razionalità diventerebbe così un’isola che emerge nell’oceano dell’irrazionalità.

⁶⁰ Labinaz (2013: 65).

⁶¹ Simon (1985: 293-304, cit. 297).

⁶² Simon (1946, trad. it. 241).

Bibliografia

- Amoretti, C., Ervas, F. 2011, "Donald Davidson", *APhEx. Portale italiano di filosofia analitica*, 3, pp. 190-210.
- Atlas, J. D. 2005, *Logic, Meaning and Conversation: Semantical Underdeterminacy, Implicature, and their Interface*, Oxford, Oxford University Press.
- Avramides, A. 1989, *Meaning and Mind: An Examination of a Gricean Account of Language*, Cambridge, The MIT Press.
- Bambini, V., Resta, D. 2012, "Metaphor and Experimental Pragmatics: When Theory Meets Empirical Investigation", in F. Ervas, E. Gola, "Philosophical Perspectives on Experimental Pragmatics", in *Humana. Mente Journal of Philosophical Studies*, 23, pp. 37-60.
- Bianchi, C. 2003, *Pragmatica del linguaggio*, Roma-Bari, Laterza.
- Bianchi, C. 2009, *Pragmatica cognitiva. I meccanismi della comunicazione*, Roma-Bari, Laterza.
- Bicchieri, C. 2006, *The Grammar of Society: The Nature and Dynamics of Social Norms*, Cambridge, Cambridge University Press.
- Carston, R. 2002, *Thoughts and Utterances: The Pragmatics of Explicit Communication*, Oxford, Blackwell.
- Cordeschi, R. 2008, "Steps Toward the Synthetic Method: Symbolic Information Processing and Self-Organizing Systems in Early Artificial Intelligence Modeling", in P. Husband, O. Holland, M. Wheeler (ed.), *The Mechanical Mind in History*, Cambridge, The MIT Press, pp. 219-258.
- Cordeschi, R. 2011, "Artificial intelligence and evolutionary theory: Herbert Simon's unifying framework", in C. Cellucci, E. Grosholz, E. Ippoliti (ed.), *Logic and Knowledge*, Cambridge, Cambridge Scholar Publishing, pp. 197-215.
- Davidson, D. 1973, "Radical Interpretation", in *Dialectica*, 27, pp. 313-328 ("Interpretazione radicale", in *Verità e interpretazione*, trad. it. di R. Brigati e E. Picardi (a cura di), Bologna, Il Mulino, 1994, pp. 193-211).

- Davidson, D. 1974, "On the Very Idea of a Conceptual Scheme", in *Proceedings and Addresses of the American Philosophical Association*, 47, pp. 5-20 ("Sull'idea stessa di schema concettuale", in *Verità e interpretazione*, trad. it. di R. Brigati e E. Picardi (a cura di), Bologna, Il Mulino, 1994, pp. 263-282).
- Davidson, D. 1984, *Inquiries into Truth and Interpretation*, Oxford, Clarendon Press, (*Verità e interpretazione*, trad. it. di E. Picardi (a cura di), Bologna, Il Mulino, 1994).
- Davidson, D. 1986, "A Nice Derangement of Epitaphs", in E. Lepore (ed.), *Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson*, Oxford, Blackwell, pp. 433-446 ("Una graziosa confusione di epitaffi", trad. it. in D. Davidson, I. Hacking, M. Dummett, *Linguaggio e interpretazione. Una disputa filosofica*, Milano, Unicopli, 1993, pp. 59-85).
- Davidson, D. 1990, "The Structure and Content of Truth", in *Journal of Philosophy*, 87, pp. 279-328.
- Davidson, D. 1999, "Interpretation: Hard in Theory, Easy in Practice", in M. De Caro (ed.), *Interpretations and Causes. New Perspectives on Donald Davidson's Philosophy*, London, Kluwer Academic Publishers, pp. 31-44.
- Davidson, D. 2001, "Rational Animals", in *Subjective, Intersubjective, Objective*, Oxford, Oxford University Press, pp. 95-105 ("Animali razionali", in *Soggettivo, intersoggettivo, oggettivo*, trad. it. di S. Levi, Milano, Raffaello Cortina Editore, 2003, pp. 121-135).
- Davidson, D. 2001, "The Emergence of Thought", in *Subjective, Intersubjective, Objective*, Oxford, Oxford University Press, pp. 123-134 ("L'emergere del pensiero", in *Soggettivo, intersoggettivo, oggettivo*, trad. it. di S. Levi, Milano, Raffaello Cortina Editore, 2003, pp. 157-171).
- Davidson, D. 2004, "Paradoxes of Irrationality", in *Problems of Rationality*, Oxford, Oxford University Press, pp. 169-187.
- Davidson, D. 2004, "Incoherence and Irrationality", in *Problems of Rationality*, Oxford, Oxford University Press, pp. 189-198.

- Davidson, D. 2005, "Gadamer and Plato's "Philebus", in *Truth, Language, and History*, Oxford, Oxford University Press, pp. 261-275.
- Ervas, F. 2008, Uguale ma diverso. Il mito dell'equivalenza nella traduzione, Macerata, Quodlibet.
- Ervas, F. 2011, "Equivalenza ed adeguatezza pragmatica nella traduzione", in S. Dal Maso, G. Massariello (a cura di), *I luoghi della traduzione*, Roma, Bulzoni, pp. 53-64.
- Ervas, F. 2011, "Discussion of Cordeschi R., Artificial intelligence and evolutionary theory: Herbert Simon's unifying framework", in C. Cellucci, E. Grosholz, E. Ippoliti (ed.), *Logic and Knowledge*, Cambridge Scholar Publishing, Cambridge, pp. 216-220.
- Ervas, F. 2012, "Sperimentare la pragmatica filosofica. Una ridefinizione delle competenze comunicative nel linguaggio figurato", in *Reti, Saperi, Linguaggi*, 4:1, pp. 28-33.
- Fehr, E., Fishbacher, U. 2004, "Third Party Punishment and Social Norms", in *Evolution and Human Behaviour* 25, pp. 63-87.
- Gadamer, H. G. 1960, *Wahrheit und Methode*, Tübingen, Mohr (*Verità e metodo*, trad. it. di G. Vattimo (a cura di), Milano, Bompiani, 2003).
- Gadamer, H. G. 1993, "Selbstdarstellung", in *Wahrheit und Methode 2: Ergänzungen*, Tübingen, Mohr, pp. 479-508 ("Autoesposizione", in *Verità e metodo 2*, trad. It. Di G. Vattimo (a cura di), Milano, Bompiani, 1996, pp. 462-463).
- Gadamer, H. G. 1997, "Reply to Donald Davidson", in L. Hahn (ed.), *The Philosophy of Hans-Georg Gadamer*, Chicago, Open Court Publishing Company, pp. 433-435.
- Grice, P. 1967, "Logic and Conversation", in P. Cole, J. Morgan (ed.), *Syntax and Semantics 3: Speech Acts*, New York, Academic Press, pp. 41-58 ("Logica e conversazione", trad. it. di P. Casalegno et al. (a cura di), *Filosofia del linguaggio*, Milano, Raffaello Cortina Editore, 2003, pp. 221-244).
- Grice, P. 1991, *The Conception of Value*, Oxford, Clarendon Press.

- Grice, P. 2001, *Aspects of Reason*, Oxford, Clarendon Press.
- Hacking, I. 1975, *Why does Language Matter to Philosophy?*, Cambridge, Cambridge University Press (*Linguaggio e Filosofia*, trad. it. B. Sassoli, Milano, Cortina, 1994).
- Henrich, J., Boyd, R., Bowles, S., Camerer, C. F., Fehr, E., Gintis, H., McElreath, R. 2001, "Cooperation, Reciprocity and Punishment in Fifteen Small-scale Societies", in *American Economic Review*, 91, pp. 73-78.
- Henrich, J., Henrich, N. 2006, "Culture, Evolution and the Puzzle of Human Cooperation", in *Cognitive Systems Research*, 7, pp. 221-245.
- Labinaz, P. 2012, "Paul Grice", APhEx. Portale italiano di filosofia analitica, 6.
- Labinaz, P. 2013, *La razionalità*, Roma, Carocci.
- Levinson, S. 1983, *Pragmatics*, Cambridge, Cambridge University Press (*La pragmatica*, trad. it. di M. Bertuccelli Papi, Bologna, il Mulino, 1993).
- Longato, F. 1996, *Comunicazione e interpretazione. Itinerari filosofici*, Trieste, Stampa Tipografica Triestina.
- Marraffa, M., Meini, M. 2005, *La mente sociale. Le basi cognitive della comunicazione*, Roma-Bari, Laterza.
- Mascaro, O., Sperber, D. 2009, "The Moral, Epistemic, and Mindreading Components of Children's Vigilance towards Deception", in *Cognition*, 112, pp. 367-380.
- McGinn, C. 1997, "Charity, Interpretation and Belief", in *Journal of Philosophy*, 74, pp. 521-535.
- Mercier, H., Sperber, D. 2011, "Why Do Humans Reason? Arguments for an Argumentative Theory", in *Behavioural and Brain Sciences*, 34, pp. 57-111.
- Newell, A., Simon, H. A. 1972, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs.

- Noveck, I., Sperber, D. (ed.) 2004, *Experimental Pragmatics*, Basingstoke, Palgrave Macmillan.
- Ramberg, B. I. 1989, *Donald Davidson's Philosophy of Language. An Introduction*, Oxford, Blackwell.
- Sbisà, M. 2007, *Detto e non detto. Le forme della comunicazione implicita*, Roma-Bari, Laterza, 2007.
- Shannon, C. E. 1950, "Programming a Computer for Playing Chess", in *Philosophical Magazine*, 41, pp. 256-275.
- Schelling, T. C. 1960, *The Strategy of Conflict*, Cambridge, Harvard University Press.
- Simon, H. A. 1946, *Administrative Behavior*, New York, Macmillan, (*Il comportamento amministrativo*, trad. it. di S. Cimmino, Bologna, Il Mulino, 1958).
- Simon, H. A. 1956, "Rational Choice and the Structure of the Environment", in *Psychological Review*, 63, pp. 129-138.
- Simon, H. A. 1983, *Reason in Human Affairs*, Stanford, Stanford University Press.
- Simon, H. A. 1985, "Human Nature in Politics: The Dialogue of Psychology and Political Science", in *The American Political Science Review*, 79, pp. 293-304.
- Simon, H. A. 2000, "Barriers and Bounds to Rationality", in *Structural Change and Economic Dynamics*, 11, pp. 243-253.
- Smith, V. L. 2008, *Rationality in Economics: Constructivist and Ecological Forms*, Cambridge, Cambridge University Press.
- Sparti, D. 1994, *Sopprimere la lontananza uccide. Donald Davidson e la teoria dell'interpretazione*, Firenze, La Nuova Italia.
- Sperber, D., Wilson, D. 1986, *Relevance. Communication and Cognition*, Oxford, Blackwell (*La pertinenza*, trad. it., Milano, Anabasi, 1993).

Sugden, R., Gold, N. 2007, "Collective Intentions and Team Agency", *Journal of Philosophy*, 104, 2007, pp. 109-137.

Wright, J. C., Dawson, V. L. 1998, "Person Perception and the Bounded Rationality of Social Judgment", in *Journal of Personality and Social Psychology*, 55, pp. 780-794.

La razionalità figurativa. Con alcune note su Philippe Ramette

Tiziana Migliore
Università Cà Foscari di Venezia
tiziana.migliore@unive.it

Il saggio di Donald Davidson che dà il titolo a questa raccolta si chiude con una confessione e un problema irrisolto: «l'unico *modo* in cui avere il contrasto fra soggettivo e oggettivo è attraverso il concetto di verità intersoggettiva. Devo però ammettere che non so come mostrare questo punto».

Il filosofo, per arrivare il più vicino possibile a una spiegazione, propone, “al posto dell'argomento”, “un'analogia”: «se venissi lanciato sulla Terra, non avrei alcun modo per determinare la distanza che mi separa dai vari oggetti. Saprei solamente che giacciono su una linea che parte da me e va verso di loro».¹ Davidson applica cioè, alla nozione di “oggettività”, la sua teoria della triangolazione, che consiste nell'asserire la centralità e la dipendenza logica dei contenuti mentali dal *legame comunicativo, linguistico*, fra la nostra mente, il mondo esterno, e la mente di un'altra persona. «La razionalità è un tratto sociale. Solo coloro che comunicano la possiedono».² E di nuovo in maniera analogica: «il nostro senso di oggettività è la conseguenza di un altro tipo di triangolazione, che richiede due soggetti. Ciascuno interagisce con un oggetto, ma cosa dà a ciascuno il concetto del modo in cui le cose oggettivamente sono, è la linea di fondo formata dal linguaggio fra le persone».³

¹ Davidson (1982a: 131). Viceversa, «non abbiamo alcun modo (né possiamo aspettarci di trovarne) per porre in corrispondenza eventi descritti nel vocabolario fisico con eventi descritti nel vocabolario mentale». Cfr. Davidson (1999, trad. it. 162). Sull'argomento vedi soprattutto il capitolo “Epistemologia esternalizzata”, in Davidson (2001: § 13).

² Davidson (1982a: 131).

³ Davidson (1982a: 131).

Una visione “esternalista”, epistemologica più che ontologica, esplicativa della tesi del “monismo anomalo”, per cui non esistono leggi psico-fisiche” causali che rendano identici tipi di eventi fisici e tipi di eventi mentali.⁴ Questa visione implica un cambiamento radicale della definizione di “linguaggio”: non più un insieme di norme convenzionali situate al di fuori dei parlanti, ma un sistema che si attualizza al momento dell’incontro fra gli interlocutori, negli aggiustamenti reciproci. Ne è la prova il fenomeno del “malapropismo”.⁵ Soprattutto, secondo Davidson,

Condividere un’abilità precedentemente appresa non basta per la riuscita di una comunicazione verbale [...]. La comprensione reciproca è raggiunta attraverso l’esercizio dell’immaginazione, l’appello alla conoscenza generale del mondo e la consapevolezza degli interessi e delle inclinazioni umane.⁶

Vale la pena di ripartire da dove Davidson si è fermato. E di chiedersi perché, nell’offrire un’accezione di razionalità come “verità intersoggettiva”, linguisticamente condivisa, Davidson scelga la procedura analogica, cioè la razionalità “figurativa”.⁷ Quali sono i tratti del ragionamento figurativo rispetto ad altri tipi di ragionamento? Obiettivo del contributo è di rispondere a questa domanda, valendosi in seguito degli spunti di un ciclo di fotografie di Philippe Ramette, le *Esplorazioni razionali*.

1. Credere e sapere

La riflessione di Davidson sulla razionalità mette a fuoco aspetti della dimensione cognitiva nei quali la semiotica strutturalista e generativa si ritrova. Anzitutto il filosofo postula che «poiché l’identità di un pensiero non può essere separata dalla sua collocazione nella rete logica degli altri pensieri, questo non può essere ricollocato in un altro posto della rete senza diventare un altro pensiero». ⁸ L’estensione a maglie così descritta, che Davidson radicalmente

⁴ Cfr. Davidson (1970, trad. it. 79-102).

⁵ «Il linguaggio, così come lo intende la maggior parte dei filosofi, non esiste». Cfr. Davidson (1977, trad. it. 59-85).

⁶ Davidson (1994, traduzioni nostre, 2).

⁷ Davidson (1982a: 131).

⁸ Davidson (1982a: 123). Per Davidson i contenuti mentali hanno un rapporto di tipo olistico fra loro, analogo a quello esistente fra gli enunciati di un linguaggio: così come un enunciato significa solo grazie alle sue connessioni con l’intero sistema linguistico, un contenuto mentale può esistere solo in quanto logicamente connesso a un’intera rete di contenuti mentali, il che ne garantisce una certa “opacità” rispetto alla semplice causazione fisica. Cfr. Davidson (1970).

legittima solo in virtù del possesso di un linguaggio – «un essere vivente non può avere un pensiero a meno di non avere un linguaggio»⁹ – è identificabile come una “struttura”, se si considerano i criteri posizionali e differenziali emergenti.¹⁰

Ritiene, poi, che «condizione necessaria e sufficiente per la razionalità sia l’ubicazione dei pensieri all’interno di una fitta rete di credenze collegate».¹¹ Qui il confronto con la semiotica generativa, dove «il credere è l’istanza cruciale del discorso scientifico»,¹² è particolarmente rilevante. Algirdas J. Greimas, indagando l’universo cognitivo sotto forma di “competenze modali”, ha notato che il sapere non circola soltanto, ma tende a un fare persuasivo. «Il far-sapere è in primo luogo un far-credere».¹³ Il “patto fiduciario” ha un ruolo fondamentale nei processi di comunicazione intersoggettiva, anche se sapere e credere sembrano appartenere a una struttura elastica: possono arrivare a confondersi, quando la tensione si allenta o, viceversa, a polarizzarsi e a produrre un’opposizione categoriale.¹⁴

Sul lato del fare interpretativo si tratta di “controllare” l’adeguamento del nuovo e dello sconosciuto rispetto al vecchio e conosciuto, quindi di congiungersi (nel caso di riuscita) o di disgiungersi (nel caso di scacco) con esso. Per Greimas

Il ri-conoscimento, contrariamente alla conoscenza, è un’operazione di comparazione tra ciò che ci viene “proposto” (= la proposizione logica, nel senso di “proposta” in quanto cioè suggestione e offerta) e ciò che già si sa o si crede. Il

⁹ Davidson (1982a: 126).

¹⁰ Gilles Deleuze, nell’articolo del 1967 “Da che cosa si riconosce lo strutturalismo?”, ha definito alcuni criteri formali del movimento. Il secondo, topologico o di posizione, e il quarto, la differenziazione, coincidono con l’idea di rete che ha Davidson. È noto inoltre che a fondamento dello strutturalismo c’è il linguaggio. «Non esiste struttura se non di ciò che è linguaggio, fosse pure un linguaggio esoterico o addirittura non verbale. Non esiste una struttura dell’inconscio se non nella misura in cui l’inconscio parla ed è linguaggio. Non esiste una struttura dei corpi se non nella misura in cui si ritiene che i corpi parlino, con un linguaggio che è quello dei sintomi». Cfr. Deleuze (1967 trad. it. 91-110).

¹¹ Davidson (1982a: 122).

¹² Cfr. la voce “credere” in Greimas & Courtés (1979).

¹³ Greimas (1983, trad. it. 111).

¹⁴ Greimas (1983: 112). La dicotomia occidentale fede/ragione scaturisce, secondo Sapir-Whorf, da un’elaborazione secondaria, posta a livello superficiale della categorizzazione. Ma è anche un fenomeno intraculturale che si è prodotto davvero. Greimas ricorda la testimonianza di un vecchio stalinista, il quale parlava del suo “sdoppiamento” in questi termini: «ormai esisteva un me che *sapeva* e un altro me stesso che *credeva*. Tra loro la corrente era spezzata. Neanche le loro memorie comunicavano». Cfr. Cathala (1982, trad. it. 121).

riconoscimento, in quanto confronto, comporta necessariamente un'*identificazione*, nell'enunciato offerto, della totalità o delle briciole di "verità" di cui già eravamo in possesso.¹⁵

Davidson aggiunge, in risonanza con Greimas, che nei percorsi di scoperta il credere motiva la sorpresa, «consapevolezza del contrasto tra ciò che credevo e ciò che vengo a credere». E chiama in ballo «credenze circa la correttezza delle proprie credenze».¹⁶

Ma la chiave di volta della questione è l'oggetto del patto epistemico, cioè le *condizioni di corrispondenza tra linguaggio e percezione*. Il credere fonda il ri-conoscimento del "mondo", dalla certezza referenziale fino allo stupore estetico dell'inatteso.¹⁷ La forza della razionalità figurativa consiste nel fatto che il percepito, l'apparenza del mondo naturale, è di ordine figurativo. Anche il livello manifestato dei nostri discorsi nei sistemi di modellizzazione – la significazione figurativa – ha radici fenomenologiche.¹⁸

2. La "funzione figurativa"

La fenomenologia ci ha insegnato a vedere *come* lo stadio figurativo della significazione avviene: qui, infatti, la presa del visibile non è diretta a valle, verso gli effetti prodotti e la loro strutturazione, ma guarda a monte della figurativizzazione, alle condizioni del suo apparire. "*Funzione presentativa*" [*Darstellende*] è il nome dato da Husserl ai «dati sensibili animati dalle "apprensioni" e che, nell'animazione, costituiscono ciò che noi diciamo "*il manifestarsi del*" colore, "*della*" figura, ecc...». La compagine della percezione

¹⁵ Greimas (1983: 115).

¹⁶ Davidson (1982a: 129).

¹⁷ Dopo Edmund Husserl, la credenza percettiva può essere incentrata I) sull'oggetto che si forma nella percezione, dotato allora di caratteri ("noematici") assimilabili alle modalità aleatiche dei logici: il reale, il necessario, il possibile, il problematico, il contingente, il dubbioso o l'irreale; oppure II) sulla credenza del soggetto e sul suo modo di assumere e far proprio l'oggetto che percepisce. I caratteri di tale credenza ("noetici") sono assimilabili alle modalità epistemiche della logica: il certo, il probabile, l'improbabile, il congetturale, il supputativo. Cfr. Husserl (1913, trad. it. 233-242).

¹⁸ Ecco perché, nel riflettere sul linguaggio e sul nostro modo di essere al mondo, alcuni dei più grandi pensatori del XX secolo – Husserl, Saussure, Wittgenstein, Hjelmslev – si sono valse dell'impiego metaforico del gioco degli scacchi. Gli spostamenti spaziali delle figure sulla scacchiera non sono meri atti di gioco, ma manifestazioni litotiche del faccia a faccia di due soggetti cognitivi variamente competenti nell'interpretare le mosse dell'interlocutore. Sequenze in *trompe-l'œil* di far-credere e quindi far-fare. Cfr. Greimas (1980, trad. it. 215-220).

implica un complesso sistema di manifestazioni e adombramenti, che è coscienza di un'unica e medesima cosa grazie al riunirsi, in una *unità di apprensione*, e grazie alla possibilità di *sintesi dell'identificazione*.¹⁹ Lo strato di senso che avvolge le cose nell'appercezione si articola con la messa in discorso delle figure depositate nei linguaggi e trasmesse dagli usi.²⁰

Per Merleau-Ponty la materia stessa della conoscenza ha origine nello schema dinamico della percezione, quando ci si chiede “che cosa vedo veramente”.²¹ Così,

Il significato motorio dei colori può essere compreso solo se essi cessano di essere degli stati chiusi in sé o delle qualità indescrivibili offerte alla constatazione di un soggetto pensante, se raggiungono in me un certo montaggio generale in virtù del quale io mi adatto al mondo; se mi invitano a una nuova maniera di valutarlo, e se d'altra parte la motilità cessa di essere la semplice coscienza dei miei cambiamenti di luogo presenti o prossimi per divenire la funzione che stabilisce, in ogni momento, stabilisce i miei campioni di grandezza, l'ampiezza variabile del mio essere al mondo. Il blu è ciò che sollecita da me un certo modo di guardare, ciò che si lascia tastare da un movimento definito dal mio sguardo. È un certo campo o una certa atmosfera offerta alla potenza dei miei occhi e di tutto il mio corpo.²²

Un movimento di costituzione reciproca fra soggetto percipiente e oggetto percepito, instabile e talvolta fragile, il cui orientamento può sempre invertirsi.²³ Husserl e Merleau-Ponty introducono l'idea di una “corrispondenza” non univoca né meccanica o necessaria fra percezione e figure del mondo: tutto accade come se vi fossero due entità “imperfette”, cariche di potenzialità, che vanno ognuna incontro all'altra. La leggibilità si fa a prezzo di convenzioni vincolanti, ma provvisorie e precarie. Nelle parole di Greimas, la griglia culturale si proietta in avanti, ma è anche possibile riconoscere, nelle Gestalten iconizzabili, corrispondenze normalmente invisibili, formanti più o meno sfigurati, latori di nuove significazioni. «La figuratività non è un semplice ornamento delle cose; essa è quello schermo dell'apparire la cui virtù consiste nel dischiudere, nel lasciar intravedere, grazie o a causa della sua imperfezione, la possibilità di un senso ulteriore».²⁴

Di qui l'azzardo del ragionamento figurativo. A differenza della razionalità deduttiva, che organizza cause e conseguenze, gerarchie, relazioni logiche

¹⁹ Husserl (1914, trad. it. 98-99).

²⁰ Vedi Bertrand (2000, trad. it. 148-149).

²¹ Merleau-Ponty (1945, trad. it. 307).

²² Merleau-Ponty (1945, trad. it. 287).

²³ Cosicché, *in extremis*, a decidere del “soggetto” non è una natura originaria, ma la prospettiva come assunzione di un punto di vista. Cfr. Migliore (2014).

²⁴ Cfr. Greimas (1987, trad. it. 57).

fra le parti e il tutto, l'argomentazione figurativa organizza e sfrutta specularmente una verità che si suppone immanente al sensibile. "Scienza del concreto".²⁵ Si distinguono, in generale, due diverse forme di pensiero scientifico, non riconducibili a fasi diseguali dello sviluppo umano o a stadi dell'evoluzione del sapere, ma a livelli strategici con cui è possibile pensare la natura: il primo, "segnico", è prossimo alla percezione e all'intuizione, pronto a cogliere le segnature delle cose; il secondo, "concettuale", è esemplificato dalla mentalità "discosta" della scienza moderna, che frappone una distanza tra percipiente (*res cogitans*) e percepito (*res extensa*). Laddove il concetto esige un'integrale trasparenza alla realtà, il segno pretende che alla realtà si incorpori una certa densità umana. *It addresses somebody*.²⁶ In questa razionalità si scruta l'apparenza del mondo e si postula che le caratteristiche visibili siano il segno di proprietà specifiche, ma celate. Il segreto è una promessa di rfigurazione.

3. Il discorso quotidiano

Radicata nell'esperienza senso-motoria, la razionalità figurativa è anche inscritta nei linguaggi, come una sorta di catacresi generalizzata. È la cosiddetta "profondità del figurativo": eccede dal letterale della rappresentazione per dotarsi di significazioni astratte. George Lakoff e Mark Johnson hanno dimostrato che il discorso quotidiano è intriso di configurazioni funzionanti per analogia diretta, in grado di organizzare vasti campi concettuali. In particolare molti contenuti semantici grammaticalizzati sottendono "primitivi" culturali, validi almeno in Occidente, del nostro vivere la spazialità, condiviso. Attraverso metafore di orientamento, la razionalità figurativa situa, per esempio, la gioia in alto ("essere al settimo cielo") e la tristezza in basso ("avere il morale a terra"); la salute in alto (è "all'apice" della forma) e la malattia in basso ("è caduto ammalato"); la virtù in alto ("è al di sopra di ogni sospetto")

²⁵ L'espressione "scienza del concreto" è di Claude Lévi-Strauss. Cfr. Lévi-Strauss (1962).

²⁶ Lévi-Strauss considera emblematici di queste due mentalità il bricoleur e l'ingegnere. L'ingegnere interroga l'universo e, in forma negativa, urta contro una resistenza con cui deve scendere a patti. Opera mediante concetti. Viceversa il bricoleur opera mediante segni, che sono intermediari tra percetti e concetti. Caratteristica del pensiero mitico, come del bricolage, è di elaborare insiemi strutturati non direttamente per mezzo di altri insiemi strutturati, ma utilizzando residui e frammenti di eventi. Il pensiero mitico elabora strutture combinando assieme residui di eventi, mentre la scienza crea, sotto forma di eventi, i suoi strumenti e i suoi risultati, grazie alle strutture che fabbrica senza posa e che sono le sue ipotesi e le sue teorie. L'uno costruisce eventi per mezzo di strutture, l'altro strutture per mezzo di eventi. Cf. Lévi-Strauss (1962: 28-35).

e il vizio in basso (“non mi abbasserò a far questo”); la ragione in alto (“un elevato livello intellettuale”) e la passione in basso (“dominare le proprie emozioni”).²⁷ La stessa struttura concettuale delle passioni – e il corpo come contenitore di trasformazioni – è ricavabile, nelle culture, riordinando in maniera narrativa le metafore presenti nei dizionari: per la definizione di “collera” si hanno il “ribollire”, il “dare in escandescenze”, l’“esplodere”, il “mandare in pezzi”...²⁸

3.1. Concetti e strumenti di analisi

A partire da queste ricerche la semiotica ha sviluppato una teoria e una metodologia di indagine della razionalità figurativa. Ha preso in carico le categorie, nate dalla psicologia della percezione, di esterocezione – il mondo esterno – di interocezione – il mondo interiore – e di propriocezione – l’interfaccia tra i due mondi: la risonanza sensibile. Il modo in cui tali nozioni sono transitate nella teoria del senso è chiaro. Il concetto di figuratività, che indica la transcodifica delle figure dell’espressione del mondo naturale in figure del contenuto delle lingue naturali, ha sostituito quello di proprietà esterocezione. Il concetto di astrazione, allusivo alle categorie che non consentono tale transcodifica, ha preso il posto dell’interocezione. Infine il concetto di timia, che, con la sua polarità euforia/disforia, dà conto dell’incorporazione del sensibile, è subentrato alla propriocezione. Già nel 1983 Greimas osservava che:

Le figure del mondo hanno una duplice funzione: in quanto apparire della sua realtà, esse ci servono da referente, intra- o extra-discorsivo; in quanto figure del linguaggio, esse sono là per dire altre cose che se stesse [...]. Qui non è tanto l’articolazione sintagmatica del discorso figurativo ad attrarre la nostra attenzione, ma piuttosto la sua attitudine a proiettare una duplice referenza, la prima *in profondità* e creatrice di una isotopia tematica più astratta, e la seconda *in lateralità*, in grado di sviluppare una nuova isotopia figurativa parallela.²⁹

Così procedono l’allegoria e la parabola, metafore estese al livello del racconto. L’allegoria riveste le astrazioni di rappresentazioni figurative. È il

²⁷ Lakoff & Johnson (1980). «Nessuna definita struttura concettuale emerge direttamente dal funzionamento emotivo stesso. Dal momento che vi sono *sistematiche correlazioni* fra le nostre emozioni (come la felicità) e le nostre esperienze senso-motorie (come la posizione eretta), queste formano le basi dei contenuti metaforici di orientamento (come contento è su)». *Ibidem*. p. 79.

²⁸ Cfr. Lakoff & Kövecses (1987: 195-221).

²⁹ Greimas (1983: 127).

caso dell'unità significante nel pensiero medievale, che, nell'allegoria, associa materiale e spirituale (es., le tre fiere, la lonza, il leone e la lupa, ostacolo al cammino di Dante nella "selva oscura", *Inferno*, canto I, vv. 31-60). La parabola, sviluppando il racconto figurativo di azioni di vita quotidiana, trasforma l'evento raccontato in un interrogativo che resta in sospeso (es., "Il Buon Semiatore" nei *Vangeli*), perché spetta all'enunciatario dare la risposta. Questi è invitato a costruire una significazione altra, fondata su un'isotopia più teorica di quella che lo svolgimento concreto gli consente di percepire. Allegoria e parabola rendono confrontabili regime scientifico e regime letterario, strappando la scienza a un'immagine puramente inferenziale.³⁰ E problematizzano lo iato fra scienze "dell'uomo" e scienze "inumane", scienze naturali e scienze "innaturali". La letteratura scientifica, trasformazione verbale degli esperimenti nei laboratori, si modella spesso sulla letteratura letteraria, tanto che il discorso scientifico avrebbe bisogno di una filologia e una critica sue proprie. Comunque, sarebbe meglio evitare di fissare a priori il tipo di razionalità, sulla base del regime di riferimento, per affidarsi invece alla descrizione testuale come criterio di scoperta dei tipi di forme discorsive che impiegano tipi diversi di razionalità.³¹ In fatto di metodo, il ragionamento analogico si è usurato con l'uso. La semiotica gli preferisce l'omologazione, che adotta la proporzione matematica nella valutazione dei rapporti fra isotopie ritenute parallele.³² Tipica dei sistemi semisimbolici – a una coppia del piano dell'espressione si correla una coppia del piano del contenuto – è operativa nelle *Esplorazioni razionali* di Ramette.

Un tema dirimente, nello studio delle razionalità, è il rilevamento della distanza costitutiva dell'oggetto. C'è una giustezza dell'osservazione modificando la quale l'oggetto stesso si trasforma. Si pensi agli insiemi di Mandelbrot, frattali che da vicino risultano essere una miriade di punti (e di numeri complessi). Cambiare distanza significa cambiare l'oggetto e sconvolgere un modo di descrizione per assumerne un altro.³³ Al suo interno qualsiasi discorso appare come "raddoppiato" da un osservatore che lo articola: un'istanza, di natura cognitiva, che non coincide né dipende dal contenuto del

³⁰ Sull'iconografia e la retorica del discorso scientifico vedi Bastide (2001). Bastide discute il testo della conferenza di un celebre fisiologo, Homer Smith ("The Fate of Sodium and Water in the Renal Tubules", 1958), che per confessare un cambiamento di credenza sul funzionamento del rene, inserisce nella sua argomentazione una parabola. L'insuccesso del tentativo di un fisiologo nel guadagnarsi il Paradiso è di fatto una riflessione sugli inconvenienti di una certezza data per assoluta. *Ibidem*. pp. 139-165.

³¹ Cfr. P. Fabbri (1987: 47-48).

³² Greimas (1983: 128).

³³ Fabbri (1987: 51-52).

discorso, ma ne attualizza l'enunciazione strategica. Se vi sono degli osservatori, vi sono inoltre delle cose da osservare e quindi delle "fonti" di informazione. A meno di non ridurre il linguaggio a una mera trasmissione di dati, *osservatori* e *informatori*, non necessariamente antropomorfi, costruiscono un "oggetto di valore". Fanno circolare saperi e credenze nei discorsi.³⁴

4. Razionalità del testo, razionalità dello sguardo

Il salto di qualità dell'approccio semiotico, merleau-pontiano, sta nel conferire alle operazioni percettive e interpretative lo stesso grado di realtà che hanno le operazioni produttive. Alla razionalità come *proprium* dell'insieme significante si accompagna una razionalità della lettura, istruita da questo insieme. Jacques Geninasca ha individuato tre tipi di presa. La presa "molare" è una percezione di grandezze organizzate secondo legami di dipendenza unilaterale, da cogliere attraverso una razionalità inferenziale. È compatibile con uno spazio euclideo, misurabile e isotropo. La presa "semantica" concerne relazioni multilaterali fra grandezze contrastive, catturabili tramite un ragionamento mitico, cioè profondo, in verticale, stratificato. Infine, nella presa "impresiva", soggetto e oggetto sono ritmicamente integrati in una totalità euforica che infrange le regole stabilite. L'oggetto, situato nello spazio oggettivo, è vissuto allo stato di soggetto, come nell'immaginazione musicale.³⁵ Presa semantica e presa impresiva risponderrebbero meglio a una razionalità testuale figurativa.

In termini di forme del "comprendere", è anche possibile delineare una «tipologia dei percorsi contrattuali della significazione figurativa». Si distinguono il "credere accolto", cioè una lettura consenziente, talora ingenua; il "credere ricusato", che rifiuta la visione ricevuta o la altera con rovesciamenti ironici; il "credere critico", dove il ragionamento arriva a modificare lo statuto della figuratività; il "credere in crisi", effetto, viceversa, di una testualità che fa crollare l'illusione figurativa e mette in dubbio la credenza alla base della condivisione semantica. Smembra il discorso che ne promuove le impressioni e degrada la stessa affidabilità sensibile della lingua.³⁶

³⁴ Cfr. Greimas (1988: 54). Nell'analisi del testo narrativo Gérard Genette ha distinto la focalizzazione "zero" ("la Terra è rotonda") dalla focalizzazione "esterna" ("in genere si pensa che la Terra sia rotonda") dalla focalizzazione "interna" ("io, Galileo, dico che la Terra è rotonda" oppure "vista da Sirio, la Terra è rotonda"). Genette (1972). Per un approfondimento in chiave semiotica, cfr. Fontanille (1989).

³⁵ Cfr. Geninasca (1984, trad. it. 244-249). Vedi anche Geninasca (2003).

³⁶ Cfr. Bertrand (2000: 256-259).

5. Philippe Ramette. L'arte cosa sperimentale

La “cassetta di attrezzi” per l'analisi della razionalità figurativa vale al di là dei linguaggi verbali, è cioè applicabile ad altre forme espressive. In quest'ultimo paragrafo se ne saggerà l'efficacia sul testo visivo, che reciprocamente potrà fornire nuove indicazioni di senso. Il caso studio scelto è il ciclo di fotografie di Philippe Ramette dal titolo *Esplorazioni razionali: esercizi di visione* paragonabili agli esperimenti cari alla fisica – Galileo si chiedeva come sarebbero caduti i corpi se un giorno la resistenza dell'aria fosse scomparsa; Einstein immaginava di vedere il mondo viaggiando a cavallo su un raggio di luce. La differenza è che Ramette questi esperimenti li fa davvero, anche se ce ne offre soltanto il risultato finale.

“Diciamo che” la superficie della Terra o del mare resta orizzontale, mentre gli alberi, i mobili e le rocce obbediscono al peso verticale, con un personaggio che vi si muove leggero, sostenuto da una forza perpendicolare che gli permette di scalare il tronco o il muro in piedi, orizzontalmente. Oppure “diciamo che” il personaggio resta soggetto alla consueta gravità, verticale, ma in un mondo dove il mare diventa parete e il suolo scogliera. O anche “diciamo che”, con un quarto di giro in più, arriva minacciante la caduta verso il cielo, vecchia storia dell'esperienza del brivido per il terrore davanti alla sorte che attendeva gli eventuali visitatori dei due poli.³⁷

Sono dimostrazioni per assurdo in cui l'artista francese si prende gioco delle abitudini umane. Sfida le leggi della fisica, più costrittive delle convenzioni sociali, non per instaurare un caos dove tutto è plausibile, ma per dirigere le forze della natura a suo modo. La logica resta quella del sistema solare. Il “*what if*” – una finzione regolamentata *in toto* – punta piuttosto sul potere di liberare l'uomo dal campo gravitazionale, opponendo all'evidenza del peso la resistenza discreta dei materiali. Certo, ruotando la Terra di 180°, si rovesciano correlazioni semisimboliche classiche, tanto che

alto, cielo, chiaro : basso, terra, scuro = pesantezza : leggerezza

La domanda sembra essere: “perché l'uomo dovrebbe obbedire alle leggi di massa passiva dei corpi oggetto?” Non potrebbe camminare su un muro – come le mosche o le lucertole – o vivere sott'acqua?

Tacciare il quesito di irrazionalità è un errore, dato che «l'irrazionalità non è qualcun altro che fallisce nel credere, sentire o fare quel che noi riteniamo ragionevole, ma il fallimento di coerenza o consistenza della singola

³⁷ Lévy-Leblond (2010, traduzioni nostre, 104).

persona nel suo schema di credenze, di attitudini, di emozioni, intenzioni e azioni» (Davidson 1982b: 169-170, trad. ns.),³⁸ «ferma restando la sua concezione di razionalità, quando mostra appunto un'interna inconsistenza o incoerenza».³⁹ Invece le situazioni di Ramette attualizzano l'impossibilità fattuale con coerenza.

Il "personaggio" è un simulacro della sua persona, rappresentato di profilo, di spalle o tre quarti e ripreso dall'alto, dal basso o di scorcio. Non guarda mai lo spettatore, perché è coinvolto nella dimensione spazio-temporale del nuovo assetto, assistente-partecipante con ruoli pragmatici e cognitivi, di focalizzazione e di sanzione positiva dell'universo che abita. Può limitarsi ad ammirarlo, mettendolo in prospettiva, ma può anche percorrerlo. Sempre in abito scuro, si astiene dall'esprimere passioni singolari, ma con la mimica – le braccia conserte o le dita chiuse sul mento – coltiva ironicamente il beneficio del dubbio. Sta facendo esercizio, secondo l'accezione di Davidson:

Chiedete a un uomo perché si esercita [...]. Come minimo la spiegazione richiamerà due fattori: un valore, obiettivo, volontà o attitudine dell'agente e la credenza che, agendo in un certo modo, egli potrà promuovere il valore o l'obiettivo di rilievo o agire in accordo con la sua propria attitudine. Da un lato l'azione, dall'altro la coppia credenza-desiderio che dà ragione dell'azione: questi due fattori si spiegano solo se messi in relazione in due modi molto diversi. Per prima cosa ci dev'essere una relazione logica. Credenze e desideri hanno un significato e questi significati devono essere tali da implicare qualcosa di valido o di desiderabile nell'azione. Perciò un uomo che desidera essere in forma e crede che l'esercizio lo mantenga in forma trarrà la conclusione che c'è qualcosa di desiderabile nell'esercizio, che spiega perché fa esercizio. Secondariamente le ragioni dell'agire di un agente, per poter spiegare l'azione, devono essere le ragioni sulle quali agisce; le ragioni devono giocare un ruolo *causale* nell'occorrenza dell'azione.⁴⁰

Nello specifico la ragione che giustifica l'azione di Ramette, mossa da una credenza-desiderio, non è una satira delle pretese della scienza, come in *The Weighing House* (1763), l'incisione di Hogarth contro la gravità di Newton. Ramette, invece, priva il soma dei suoi limiti e lo dota di agilità per compiere, in equilibrio e senza sforzo, imprese estreme. Se solitamente si

³⁸ Davidson (1982b, traduzioni nostre, 169-187).

³⁹ Davidson (1985, traduzioni nostre, 189-198).

⁴⁰ Davidson (1982b, traduzioni nostre, 173). Ragioni mosse e causate, però, da credenze e desideri. «L'azione intenzionale non può emergere prima di credenze e desideri, poiché un'azione intenzionale è un evento spiegato dalle credenze e dai desideri che l'hanno causato; le credenze non possono emergere una alla volta, poiché il contenuto di una credenza dipende dalla posizione che essa occupa nella rete di altre credenze». Davidson (1999: 162).

contrasta il peso del corpo con l'allenamento, una fatica che rinforza la capacità plastica di metamorfosi,⁴¹ qui la sensazione disforica della caduta è evitata ipotizzando di ovviare alla norma. Agilità ha la medesima radice di "agio", "modo di sentirsi libero nello spazio", ed è un parasonimo di "gioco".⁴² L'esito, però, non è l'ebbrezza della levitazione. La sagoma, in piedi o seduta, conserva il distacco fisico dai fenomeni che esperisce, la giusta distanza per la costituzione del sapere. E ha sempre un piano di sostegno o un appiglio adiuvanti nella performance. Quando è in verticale, ha i piedi ben piantati... alla parete.

Lo statuto teorico del concessivo – il controfattuale dell'attrazione che ci ancora al suolo – richiede una tenuta e una solidità maggiori delle usuali formule scientifiche. Serve a definire un quadro di visione che, nel negare l'esistente, sottolinea quanto vi siamo assuefatti. La razionalità figurativa scatenava una reazione contro l'adattamento: immagina un altro ordine di cose, mandando in crisi il credere rassegnato alla gravità. Accade nello spessore del mondo.

Il fisico, per prassi, dimentica che termini impiegati in una scienza antica come la meccanica – parole quali peso, tensione, forza – prima di diventare nomenclature di elementi, provengono dalla lingua comune e designano esperienze sensibili. Una forza non è soltanto il nome del simbolo F nel principio di Newton, $F = ma$. E peso e tensione non sono unicamente esseri matematici eterei, ma spinte, scontri, sfide, minacce. Ecco che si sveglia il contrasto fra umano e non umano come "verità intersoggettiva", l'essenza del razionale secondo Davidson.⁴³ Raggiunge, in Ramette, il "sublime del pensiero".

Una sorta di Caspar David Friedrich in prospettiva mentale e non della natura: rispetto alla teoria dei mondi possibili che ipostatizza e tiene separati, artificialmente, "finzione" e "realtà", la mentalità figurativa, tessitrice di una verità relazionale fra soggetti e oggetti, consente di prevedere e immaginare gli aggiustamenti umani a una realtà diversamente impostata.

⁴¹ «La discesa prudente vuole che ci si provi a trattenere contro la pesantezza, che ci si lanci con ardore verso il vuoto di attrazione, ma dominando la sua legge. Invertire dunque il lavoro dei muscoli, mettere la schiena al posto della faccia, sviluppare la forza anziché concentrarla su di sé, in estensione più che in trazione». Serres (1999, traduzioni nostre, 58).

⁴² Greimas (1980: 219-220).

⁴³ Davidson (1982a, trad. it. 131).

Bibliografia

- Bastide, F. (2001), *Una notte con Saturno*, Roma, Meltemi.
- Bertrand, D. (2000), *Précis de sémiotique littéraire* (Basi di semiotica letteraria, trad. it., Roma, Meltemi, 2002).
- Davidson, D. (1970), "Mental Events", ("Eventi mentali", trad. it. di A. De Palma, G. Pareti (a cura di), in *Mente e corpo. Dai dilemmi della filosofia alle ipotesi della neuroscienza*, Torino, Bollati Boringhieri, 2004, pp. 79-102).
- Davidson, D. (1977), "A nice derangement of Epitaphs" ("Una graziosa confusione di epitaffi", trad. it. di L. Perissinotto (a cura di), *Linguaggio e interpretazione. Una disputa filosofica*, Milano, Unicopli, 1993, pp. 59-85).
- Davidson, D. (1982a), "Rational Animals" ("Animali razionali", trad. it. di S. Gozzano (a cura di), *Mente senza linguaggio*, Roma, Editori Riuniti, 2001, pp. 119-132).
- Davidson, D. (1982b), "Paradoxes of Irrationality", poi in Id., *Problems of Rationality*, Oxford, Oxford University Press, 2004, pp. 169-187.
- Davidson, D. (1985), "Incoherence and Irrationality", poi in Id., *Problems of Rationality*, Oxford, Oxford University Press, 2004, pp. 189-198.
- Davidson, D. (1994), "The Social Aspect of Language", in B. McGuinness, G. Oliveri (a cura di), *The Philosophy of Michael Dummett*, Dordrecht, Kluwer, pp. 1-16).
- Davidson, D. (1999), "The Emergence of Thought" ("L'emergere del pensiero", trad. it. in Id., *Soggettivo, Intersoggettivo, Oggettivo*, Milano, Cortina, 2003, pp. 157-171).
- D. Davidson, D., (2001a), "The Emergence of Thought", in *Subjective, Intersubjective, Objective*, Oxford, Oxford University Press, 2001, pp. 123-134 (*L'emergere del pensiero*, trad. it., in *Soggettivo, intersoggettivo, oggettivo*, Milano, Raffaello Cortina Editore, 2003, pp. 157-171).

- Davidson, D. (2001b), “Epistemology externalized”, in *Subjective, Intersubjective, Objective*, Oxford, Oxford University Press, 2001 (“Epistemologia esternalizzata”, trad. it. in Id., *Soggettivo, Intersoggettivo, Oggettivo*, Milano, Cortina, 2003, pp. 245-259).
- Deleuze, G. (1967), “À quoi reconnaît-on le structuralisme?” (“Da che cosa si riconosce lo strutturalismo?”), trad. it. in P. Fabbri, G. Marrone (a cura di), *Semiotica in nuce*, Roma, Meltemi, 2000, pp. 91-110).
- Fabbri, P. (1987), “Moduli e parabole: ragionare per figure”, in Id., *Elogio di Babele*, Roma, Meltemi, 2003, pp. 44-52).
- Fontanille, J. (1989), *Les espaces subjectifs, Introduction à la sémiotique de l'observateur*, Paris, Hachette.
- Genette, G. (1972), *Figures III (Figure III. Discorso del racconto*, trad. it. L. Zecchi, Torino, Einaudi, 1976).
- Geninasca, J. (1984), “Le regard esthétique” (“Lo sguardo estetico. Analisi di un testo di Stendhal”, trad. it. di in Id., *La parola letteraria*, I. Pezzini e M. P. Pozzato (a cura di), Milano, Bompiani, 2000, pp. 224-254).
- Geninasca, J. (2003), “El logos del formato” (“Il logos del formato”, trad. it., *Documenti di Lavoro del Centro Internazionale di Semiotica e Linguistica di Urbino*, Roma, Aracne, 2013).
- Greimas, A. J. (1980), “A propos du jeu” (“A proposito del gioco”, trad. it. in Id., *Miti e figure*, F. Marsciani (a cura di), Bologna, Esculapio, 1995, pp. 215-220).
- Greimas, A. J. (1983), “Le savoir et le croire: un seul univers cognitif” (“Il sapere e il credere: un solo universo cognitivo”, trad. it., in Id., *Del senso 2. Narrativa, Modalità, Passioni*, Milano, Bompiani, 1984, pp. 111-129).
- Greimas, A. J. (1987), *De l'Imperfection; (Dell'imperfezione*, trad. it. di G. Marrone, Palermo, Sellerio, 1988).
- Greimas, A. J. (1988), “Per una semiotica del discorso”, *Carte Semiotiche*, 4/5, numero monografico *Punto di vista e osservazione. Analisi e tipologia dei discorsi*, I. Pezzini, C. Marmo (a cura di), XIV Convegno AISS, Bologna 1986, pp. 52-57.

- Greimas, A. J., Courtés, J. (1979), *Semiotique. Dictionnaire raisonné de la theorie du langage* (Semiotica. Dizionario ragionato della teoria del linguaggio, trad. it. di P. Fabbri (a cura di), Milano, Bruno Mondadori, 2009).
- Husserl, E. (1913), *Ideen zu einer reinen Phänomenologie und phänomenologische Philosophie* (Idee per una fenomenologia pura e per una filosofia fenomenologia, trad. it. di E. Filippini (a cura di), Torino, Einaudi, 1965).
- Lakoff, G., Johnson, M. (1980), *Metaphor We Live By* (*Metafora e vita quotidiana*, trad. it. di P. Violi (a cura di), ed. Milano, Bompiani, 2007).
- Lakoff, G., Kövecses, Z. (1987), "The cognitive model of Anger inherent in American English", in D. Holland, N. Quinn (a cura di), *Cultural models in language and thought*, New York, Cambridge University Press, 1987, pp. 195-221.
- Latour, B. (1991), *Nous n'avons jamais été modernes* (Non siamo mai stati moderni. Saggio di antropologia simmetrica, trad. it. di G. Lagomarsino, Milano, Elèuthera, 1995).
- Lévy-Leblond, J. M. (2010), *La science n'est pas l'art. Brèves rencontres*, Paris, Hermann.
- Lévi-Strauss, Cl. (1962), *La pensée sauvage; (Il pensiero selvaggio*, trad. it. di P. Caruso, Milano, Il Saggiatore, 1964).
- Merleau-Ponty, M. (1945), "Le sentir", in Id., *Phenomenologie de la perception* ("Il sentire" in *Fenomenologia della percezione*, ed. in *Semiotica in nuce II. Teoria del discorso*, trad. it. di P. Fabbri, G. Marrone (a cura di), Roma, Meltemi, 2001, pp. 285-307).
- Migliore, T. (2014), "Sul prospettivismo", EC, *Rivista dell'Associazione Italiana di Semiotica*.
- Serres, M. (1999), *Variations sur le corps*, Paris, Le Pommier.

I fondamenti cognitivi della matematica umana: dove cercarli?

Valeria Giardino
Institut Jean Nicod, Paris
Valeria.Giardino@ens.fr

In questo articolo discuto dell'apporto delle scienze cognitive alla ricerca sui fondamenti cognitivi della matematica, e presento brevemente alcuni dei più importanti risultati sperimentali sulle capacità aritmetiche e geometriche degli animali, dei bambini, e dei membri di popolazioni che possiedono una lingua priva di parole per numeri superiori al 4. Sosterrò che per comprendere appieno sia il valore sia i limiti di queste evidenze sperimentali è utile individuare tre livelli di competenze necessarie per fare aritmetica e geometria, che vanno dalla percezione alla conoscenza astratta. In questa prospettiva, particolare importanza assumerà il livello II, in cui entrano in azione le capacità simboliche umane – che distinguono gli esseri umani da altre specie – e vengono creati e introdotti strumenti cognitivi. È a livello II che l'interazione tra natura e cultura diventa determinante per lo sviluppo della matematica come la conosciamo e di cui possiamo qui ritrovare i fondamenti cognitivi.

1. Introduzione

La cognizione matematica è un tema molto discusso nelle attuali scienze cognitive. L'obiettivo della ricerca in quest'ambito è svelare quali processi cognitivi sottendano le nostre competenze matematiche, ovvero in che modo la matematica – nelle forme che ci sono familiari – abbia potuto svilupparsi a partire dalle nostre capacità più spontanee, in particolare dalla percezione. È evidente che le scienze cognitive non sono certo l'unica né tantomeno la prima disciplina a essersi fatta domande sull'origine e sulle caratteristiche

della matematica. Al contrario, si può affermare che questi temi sono da sempre al centro della ricerca in filosofia.

Per fare solo due esempi tra i tanti a cui si può pensare, consideriamo in primo luogo quello che Locke racconta delle capacità dei componenti di una tribù di Americani – i ‘Tououpinambos’ – che, sebbene fossero «di spirito assai pronto e razionale», «non potevano, come noi, contare fino a 1000, né avevano idea distinta di quel numero, sebbene potessero calcolare molto bene fino a 20»¹. Quando li si interrogava su numeri molto grandi, i Tououpinambos si indicavano i capelli per far riferimento a una moltitudine che non erano in grado di enumerare. Tuttavia, sebbene non possedessero nomi per numeri superiori a 5, «qualsiasi numero al di là di quello era raffigurato mostrando le loro dita e quelle degli altri che fossero presenti». Secondo Locke, casi come questo fanno sospettare che le parole comunemente usate per contare – ‘uno’, ‘due’, ‘tre’, ‘quattro’, e così via dicendo – «conducono a farci contare bene», ma non sono, perlomeno entro certi limiti, strettamente necessarie all’esecuzione di questo compito, inteso come tener traccia di numerosità differenti. Inoltre, nel passaggio riportato, Locke pone l’accento sull’apporto cognitivo costituito non solo delle parole ma anche dell’uso delle dita per il fine specifico di contare. Potremmo a questo punto chiederci: qual è la relazione tra gli ‘strumenti cognitivi’ di cui spontaneamente facciamo uso, come le parole o le nostre stesse dita, e la nostra capacità di ‘far di conto’?

Domande analoghe possono essere sollevate anche per altri ambiti della matematica. Come noto, Kant ha analizzato in profondità la nostra capacità di rappresentare lo spazio e le connessioni tra questa capacità e la geometria. Secondo Kant, la rappresentazione geometrica «originariamente deve essere un’intuizione; perché da un semplice concetto non possono derivare proposizioni che lo oltrepassino, come accade in geometria»². A suo modo di vedere, questa intuizione sarebbe precedente a ogni atto percettivo, ovvero un’intuizione pura e non empirica.

Ci troviamo dunque, sia nel caso dell’aritmetica che in quello della geometria, di fronte a quesiti a cui è difficile trovare risposta: Che cosa determina – prima ancora dell’esperienza e dei concetti – la nascita e lo sviluppo di queste due discipline? Quali sono le condizioni per la loro comparsa e come sono diventate quelle che conosciamo? Infine, qual è il loro rapporto con il nostro sistema linguistico e concettuale? Le scienze cognitive sembrano oggi essersi fatte carico di queste domande, ereditate dalla filosofia del passato, pur

¹ Locke (1690: 249). Devo questa citazione a De Cruz *et al.* (2010).

² Kant (1781, trad. it. 57-58).

avendo sviluppato strategie molto diverse – in qualche caso persino incompatibili – rispetto a quelle proposte dalla riflessione filosofica tradizionale; in ogni modo, anche per le scienze cognitive, l’obiettivo è comprendere quali siano quelli che vengono attualmente chiamati i ‘fondamenti cognitivi’ della matematica.

Vorrei soffermarmi a questo proposito su cosa si debba intendere per ‘fondamenti cognitivi’. La questione potrebbe sembrare solo terminologica ed è invece a mio parere di importanza cruciale. In generale, va innanzitutto precisato che l’obiettivo delle scienze cognitive è quello di individuare le basi biologiche di un numero sempre maggiore di specifiche capacità cognitive – come per esempio saper distinguere tra diverse numerosità o avere determinate intuizioni spaziali. La ricerca di queste basi biologiche s’intreccia con l’identificazione delle cause che ne hanno sollecitato nel corso dell’evoluzione la comparsa e lo sviluppo. Un primo passo in questa indagine è considerare se queste capacità siano specifiche a un unico dominio, o in altre parole in che misura esse siano indipendenti rispetto alla presenza di altre capacità cognitive. Di particolare interesse, è valutare se una capacità scaturisca o meno dal linguaggio. Vedremo come questo interrogativo si presenterà nel caso della matematica.

In questo quadro generale, è necessario tuttavia far notare come l’etichetta ‘cognizione’ sia stata finora utilizzata in contesti distinti, spesso con diversi significati. Una parte degli studiosi sembra ritenere che i fondamenti ‘cognitivi’ delle nostre credenze e dei nostri valori siano determinati da un numero esiguo di sistemi, presenti nel nostro cervello e indipendenti l’uno dall’altro, in grado di adattarsi a contesti complessi e multiformi, sui quali si baserebbero nuove e più composite costruzioni – ad esempio i concetti³. La ricerca empirica avrebbe fornito ormai le prove dell’esistenza di almeno quattro di questi cosiddetti sistemi ‘di base’, e indicherebbe la presenza di un quinto. Questi sistemi ci metterebbero nelle condizioni di rappresentare: (i) gli oggetti inanimati e le loro interazioni meccaniche; (ii) gli agenti intorno a noi e le loro azioni, quando finalizzate a uno scopo specifico; (iii) gli insiemi e alcune loro relazioni numeriche – ordine, addizione, e sottrazione; (iv) gli elementi all’interno di una struttura spaziale e le loro relazioni geometriche. Inoltre, un quinto sistema sarebbe alla base delle nostre interazioni sociali, perché ci renderebbe in grado di riconoscere e distinguere tra gli altri: (v) i membri del gruppo sociale a cui apparteniamo. Non è necessario assumere che questi sistemi siano ‘innati’ o ‘a priori’; è sufficiente piuttosto accettare

³ Kinzler *et al.* (2007).

che di fatto essi si attivino in maniera del tutto spontanea nella nostra interazione con il mondo, non importa quali siano le circostanze particolari, l'età degli individui (entro certi limiti), la loro appartenenza culturale e in alcuni casi persino la loro specie (questi sistemi sono infatti almeno in parte comuni agli esseri umani e ad alcuni animali non umani).

Tuttavia, l'idea che si debba intendere la cognizione come 'conoscenza di base' contrasta con un altro approccio alla cognizione che ha invece come obiettivo la considerazione non solo dell'aspetto biologico degli individui, della loro 'natura', ma anche dell'influenza che la formazione, la loro 'cultura', può avere e di fatto ha sullo sviluppo delle varie capacità cognitive. In questa prospettiva, l'attenzione è rivolta al modo in cui i processi storici e le conquiste culturali hanno modellato e modulato i sistemi d'interazione spontanei a cui abbiamo accennato poco sopra, conducendo all'introduzione di numerose tecnologie e allo sviluppo della stessa scienza. A mio parere, l'elemento storico e culturale non può essere trascurato – se l'obiettivo finale vuole essere determinare le basi cognitive di attività umane complesse come ad esempio la matematica. In queste attività, una forma di ragionamento astratto si combina con l'utilizzo di un certo numero di sistemi di rappresentazione e di tecnologie cognitive molto sofisticati. In particolare, difenderò l'idea che la cognizione debba essere intesa come 'distribuita': i processi cognitivi vanno compresi nei termini della propagazione e della trasformazione di rappresentazioni che sono pubbliche, ovvero che possono essere condivise, modificate e in certi casi persino abbandonate se non più efficaci. Trovo a questo proposito del tutto attendibile la tesi che sostiene che gli eventi cognitivi superino i confini fisici della superficie epidermica o della calotta cranica di un individuo, e possano essere considerati come distribuiti in almeno tre modi: (i) tra i membri di uno stesso gruppo sociale; (ii) nella coordinazione di una struttura interna e di una esterna – che può essere ricavata dall'ambiente circostante o costruita con un obiettivo specifico; (iii) nel tempo – in quanto i prodotti di eventi precedenti sono in grado di trasformare la natura degli eventi successivi⁴.

Torniamo ora a discutere della metodologia delle scienze cognitive, per precisare sulla base di quali prove ricavano le loro teorie. La loro prima fonte di evidenza scientifica è l'etologia e lo studio della cognizione in animali non umani. L'individuazione della presenza di alcune abilità cognitive specifiche e complesse nei nostri precursori evolutivi, nonché in altri animali, fornirebbe prove a favore della determinazione biologica di queste abilità che si sco-

⁴ Cfr. Hutchins (2001).

prono comuni a specie differenti; gli animali non sono certo in grado di parlare, eppure mostrano di possedere – come vedremo – alcune capacità matematiche di base. Una seconda fonte è costituita dagli studi sui neonati o sui bambini prescolari; la comparsa precoce di certe capacità fornirebbe un'evidenza comportamentale a favore della loro specificità rispetto a un dominio di problemi particolare e dunque della loro indipendenza rispetto ad altre capacità. Infatti, il comportamento di neonati e di bambini in età prescolare non è ancora – perlomeno non in maniera massiccia – condizionata dall'istruzione e dalla cultura. Un'altra fonte di evidenza è l'antropologia cognitiva e lo studio di popolazioni che non sono dotate di scrittura e/o che utilizzano una lingua più povera di quella propria alla nostra comunità e ad altre comunità analoghe. Il fatto che queste popolazioni siano capaci di trovare soluzioni a determinati compiti cognitivi che sono conformi a quelle fornite da individui alfabetizzati, fornisce ulteriori ragioni per concludere che alcune abilità cognitive sono in effetti indipendenti dal nostro linguaggio e dalla nostra cultura.

Nel resto dell'articolo, presenterò brevemente una selezione dei risultati empirici che sono stati ottenuti sull'aritmetica e sulla geometria a partire da queste tre fonti, tralasciandone una quarta che si basa sullo studio di soggetti che hanno subito delle perdite a livello cognitivo, per esempio a seguito di un'ischemia cerebrale. Il motivo per cui non ritengo necessario nel presente contesto introdurre questi risultati, è che non ho come obiettivo – perlomeno non in questa sede – di rispondere ad alcuna domanda specifica sul funzionamento del cervello o sulle basi neurologiche delle capacità di cui discuterò. Non mi occupo qui di esplorare i fondamenti neurofisiologici della matematica, quanto invece di tracciare il percorso cognitivo che conduce gli esseri umani dalle loro competenze percettive allo sviluppo della matematica che conosciamo⁵.

Vorrei infine segnalare un'altra possibile confusione nell'uso dell'etichetta 'fondamenti cognitivi della matematica'. Va detto che molto di quello che è stato testato finora – perlomeno negli articoli che ho potuto leggere fino ad oggi – non è in senso stretto 'matematica'. L'indagine sperimentale non si è spinta al di là della valutazione delle competenze in aritmetica e in geometria euclidea elementare. Come vedremo, gli studi che presenterò si interrogano in primo luogo sull'esistenza di sistemi di rappresentazioni mentali che ci permettano di ragionare su numerosità e forme senza ricorrere a parole o a

⁵ Per una raccolta piuttosto esaustiva dei diversi studi empirici sulla cognizione matematica, si veda Campbell (2005).

simboli. Per mettere in evidenza questo limite nella ricerca empirica, individuerò tre livelli corrispondenti alle competenze richieste nell'esecuzione di specifici compiti aritmetici o geometrici, che vanno da capacità di carattere più tipicamente percettivo ad altre che ci permettono di accedere alla conoscenza astratta. La mia proposta è di valutare quale di questi livelli sia in gioco nei diversi scenari sperimentali presentati. Una tale distinzione in livelli è a mio parere di aiuto per comprendere in che misura i risultati empirici ci dicono davvero qualcosa sui fondamenti cognitivi della matematica. Il mio approdo finale sarà di concentrare l'attenzione sul secondo livello di competenza che – contrariamente al primo – non coinvolge unicamente la percezione ma anche la familiarità con una particolare tecnologia cognitiva, e – contrariamente al terzo e ultimo livello – non richiede ancora il possesso di concetti astratti. Secondo la mia interpretazione, è proprio a questo secondo livello che vediamo interagire natura e cultura, e per questo motivo esso si rivela un passaggio cruciale per la cognizione umana. Ora è più chiaro dove cercare i fondamenti cognitivi della matematica; lo studio dettagliato dell'interazione natura-cultura è tuttavia obiettivo di una ricerca ulteriore.

Nei paragrafi che seguono, presenterò brevemente una selezione dei risultati empirici a partire dalle diverse fonti che ho menzionato, e discuterò per ciascuno di essi fino a che punto si possano considerare risolutivi rispetto alla domanda sui fondamenti cognitivi della matematica umana.

2. Cognizione e aritmetica

2.1. Animali

In un esperimento ormai celebre degli anni '50 del secolo scorso, è stato dimostrato che i corvi sono in grado di eseguire correttamente un compito aritmetico⁶. Gli sperimentatori hanno posto davanti agli uccelli una carta con sopra disegnato un certo numero di pallini neri; i corvi dovevano poi scegliere tra varie scatole sul cui coperchio vi erano disegnati dei pallini neri, quella con il coperchio con il numero di pallini neri corrispondenti al numero di pallini sulla carta. Per provare che i corvi basavano la loro scelta sulla numerosità dei pallini e non sulla loro disposizione spaziale, sia la disposizione dei pallini sul coperchio che quella dei pallini sulla carta sono state controllate e variate sistematicamente. I risultati mostrano che i corvi sono in grado di distinguere fino a 6 pallini. In esperimenti successivi si è visto che gli scimpanzé possono fare anche di più. A seguito di un addestramento appropriato, essi arrivano

⁶ Köhler (1951).

persino a creare corrispondenze biunivoche tra numerali arabi fino a 6 mostrati loro su una carta e il numero di oggetti che vedono apparire su uno schermo, per esempio matite o cucchiaini⁷. Sia i corvi che gli scimpanzé sarebbero dunque in grado di distinguere numerosità. Inoltre, nel caso dei primati non umani, essi saprebbero costruire corrispondenze biunivoche tra simboli che rappresentano numeri e gruppi di oggetti. Si noti tuttavia che il limite – 6 pallini o 6 elementi in un insieme – è il medesimo in entrambe le specie.

Altri studi si sono occupati di valutare se gli animali siano in grado di tener traccia di numerosità anche quando esse vengano date in modalità diverse da quella visiva⁸. In un esperimento, dei topi sono stati addestrati a distinguere tra 2 o 4 suoni e tra 2 o 4 luci intermittenti. Gli animali sono stati prima abituati alle seguenti regolarità: dopo 2 lampi di luce o 2 suoni, dovevano premere la leva a sinistra; dopo 4 lampi di luce o 4 suoni, quella a destra. Messi di fronte a un compito nuovo, nel quale apparivano 2 luci e immediatamente dopo seguivano 2 suoni, i topi spontaneamente reagivano premendo la leva a destra, mostrando così di mettere insieme come in un'addizione eventi discreti temporalmente contigui – le luci e i suoni – sebbene di natura diversa.

Possiamo chiederci a questo punto che tipo di evidenza viene fornita da questi studi relativamente ai fondamenti cognitivi della matematica. Senza dubbio, i risultati mostrano che la capacità di distinguere tra numerosità è rilevante per la sopravvivenza, esattamente come lo è il riconoscimento di altre caratteristiche dell'ambiente, che si tratti di colori o – come vedremo più avanti – di forme. Si possono quindi rintracciare delle ragioni evolutive per lo sviluppo di queste capacità: essere in grado di distinguere dove si trovi la quantità maggiore di cibo è sicuramente d'importanza cruciale per la sopravvivenza. Tuttavia, interrogiamoci su quelli che sono i limiti delle prestazioni degli animali. Per prima cosa, in ciascuno di questi esperimenti gli animali sono stati sottoposti a un lunghissimo addestramento, in alcuni casi durato dalla loro nascita al loro decesso, e ciononostante essi hanno imparato a trattare solo con numeri molto piccoli⁹. Per numeri oltre il 6, le loro reazioni cessano di essere accurate. In fondo, è discutibile anche la stessa affermazione che questi esperimenti mostrerebbero che gli animali sono in grado di contare: considerata la soglia molto bassa del numero di elementi o di eventi di cui

⁷ Matsuzawa (1985).

⁸ Church & Meck (1984).

⁹ Ricordo a questo proposito che in un suo intervento al *Collège de France* Elisabeth Spelke fece notare come il pappagallo grigio africano Alex, di cui tanto si è scritto, è morto prima di riuscire ad arrivare a contare fino 7, mentre basta osservare un bambino che impara a contare per rendersi conto di trovarsi di fronte a una situazione molto diversa.

riescono a tener traccia, un'interpretazione alternativa è che il loro comportamento e le loro risposte si basino su indizi di tipo percettivo – che magari non sono per noi facili da individuare – senza che questo comporti l'attribuzione o la distinzione da parte degli animali di alcuna numerosità particolare agli insiemi di punti o di oggetti.

2.2. Neonati

Un'altra fonte interessante di evidenza scientifica sono i neonati, che non hanno ancora sviluppato il linguaggio e non sono ancora condizionati in maniera massiccia dalla cultura del proprio gruppo di appartenenza. Un esperimento ormai molto conosciuto dei primi anni '90 coinvolge bambini di 4 mesi e mezzo. Secondo l'interpretazione dell'autrice, questo esperimento dimostrerebbe che già a pochi mesi i neonati sono in grado di estrarre dall'ambiente informazioni relative alla numerosità¹⁰.

L'ingegnoso design dell'esperimento è il seguente. Un pupazzo viene spostato dal lato al centro di un piccolo palco posto davanti ai neonati (Fig. 2.1a). Quindi, uno schermo cala nascondendo il pupazzo alla vista dei neonati (Fig. 2.1b). Subito dopo, un secondo pupazzo, identico al primo, viene mostrato ai neonati e poi posizionato dietro lo schermo (Fig. 2.1c). A questo punto, lo schermo viene sollevato, e le possibilità sono due: o ci sono due pupazzi (Fig. 2.1e), o ce n'è uno solo (Fig. 2.1f). La predizione della sperimentatrice è che, se i neonati fossero in grado di tener traccia del numero degli oggetti che percepiscono, allora dovrebbero considerare il primo caso (in cui 1 giocattolo + 1 giocattolo fa 2 giocattoli) come meno sorprendente del secondo (in cui 1 giocattolo + 1 giocattolo fa 1 giocattolo). L'assunzione metodologica in esperimenti di questo genere aventi come soggetti neonati, è che il tempo di fissazione dello sguardo dei neonati sia indicativo del loro stupore: più osservano la scena, più mostrano di essere sorpresi da quello che vedono, perché esso viola le loro aspettative. I risultati mostrano che in effetti i neonati guardano sistematicamente più a lungo il risultato scorretto ($1+1=1$) rispetto a quello corretto ($1+1=2$), indipendentemente da altri cambiamenti nell'oggetto in altre proprietà di controllo, come la loro forma o il loro colore.

¹⁰ Wynn (1992).

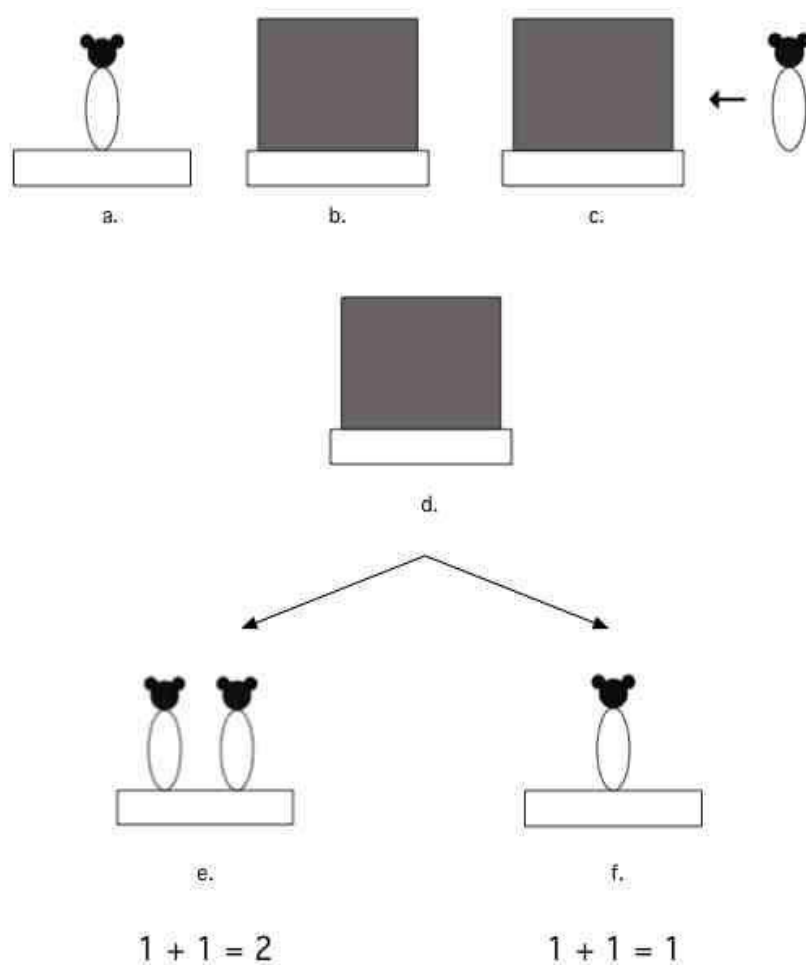


Fig. 2.1 Il design dell'esperimento di Wynn

Un esperimento più recente, con soggetti più grandi tra i 10 e i 12 mesi, cerca di valutare se bambini così piccoli siano in grado di effettuare operazioni aritmetiche elementari. Gli sperimentatori mostrano ai bambini due contenitori vuoti e iniziano poi a riempirli di biscotti, senza mai superare un massimo di 6 biscotti per contenitore¹¹. Ovviamente, i contenitori sono opachi, e di conseguenza i bambini non possono vedere dal di fuori se essi contengano biscotti e in quale quantità. Inoltre, i contenitori non vengono riempiti nello stesso ordine e si alterna quale dei due – a sinistra o a destra – contenga più biscotti, in modo che l'unica variabile mantenuta costante è il numero dei

¹¹ Feigenson *et al.* (2002).

biscotti. I risultati dell'esperimento indicano che i bambini sono capaci di scegliere il contenitore che contiene il maggior numero di biscotti, a condizione che i biscotti non siano più di 4. Di nuovo, troviamo un numero limite vicino a quello già discusso per gli animali.

Chiediamoci ora che tipo di evidenza viene fornito da questi esperimenti relativamente ai processi cognitivi legati alla matematica. Gli sperimentatori mettono in evidenza che i loro risultati darebbero torto a Piaget e alla sua ricostruzione dello sviluppo delle capacità aritmetiche nei bambini¹². Secondo Piaget, queste capacità si sviluppano tardi, poiché dipendono da altre capacità logiche che appaiono prima, come ad esempio quella di costruire corrispondenze biunivoche. Tuttavia, il comportamento di questi bambini farebbe credere che essi in qualche modo sono già in grado di comprendere le relazioni numeriche e hanno determinate aspettative rispetto all'aritmetica degli oggetti che li circondano. Detto questo, interroghiamoci nuovamente sul tipo di risposta che viene fornito da questi esperimenti alla domanda sui fondamenti cognitivi dell'aritmetica.

Una prima preoccupazione è la seguente: nell'esperimento con i pupazzi, la sperimentatrice è davvero sicura di sapere che cosa i neonati guardano più a lungo e trovano quindi sorprendente? Altre interpretazioni al di là delle aspettative aritmetiche sono possibili. I neonati potrebbero essere interessati a proprietà extra numeriche: per esempio, potrebbero avere maggiore familiarità con alcuni oggetti piuttosto che con altri, oppure potrebbero essere incuriositi da altre loro proprietà, come la loro posizione sul palco, la superficie totale che coprono, il volume che occupano e così via dicendo. Alcuni critici dell'esperimento hanno proposto un modello alternativo alla numerosità basato sulla rappresentazione che i neonati mostrerebbero di possedere degli oggetti reali che vengono loro presentati¹³. Quello che si può concludere con sicurezza è quindi solo che i neonati distinguono tra due oggetti diversi, ma non possiamo essere sicuri sulla base di quale informazione. Inoltre, nel caso

¹² Piaget (1952).

¹³ Uller *et al.* (1999). Faccio notare che Susan Carey propone una lettura basata su quello che lei definisce *object file* anche per l'esperimento dei biscotti in Feigenson *et al.* (2002). Secondo la sua interpretazione, i risultati dell'esperimento condurrebbero a pensare che esista un secondo sistema cognitivo dopo quello che vedremo con i Mundurucu e i Pirahã di aritmetica approssimata (*analog magnitude system*) che sia dedicato al riconoscimento di individui. Questo secondo sistema, che la Carey chiama *parallel individuation system*, non sarebbe specifico alla rappresentazione numerica ma alla rappresentazione esplicita di individui. Carey propone infine di considerare un terzo sistema, dedicato ai quantificatori ripresi poi dal linguaggio naturale (*natural language quantifiers*). Questi tre sistemi, presi insieme, costituirebbero le basi sulle quali verrà poi costruita la rappresentazione del numero naturale. Cfr. Carey (2009) e nota 22.

dell'esperimento con i due contenitori, i bambini di 10 e 12 mesi sembrano invece essere capaci di discriminare tra diverse numerosità, ma solo per numeri molto piccoli. È quindi possibile pensare che il loro comportamento si basi, come nel caso degli animali, su competenze ancora legate alla percezione, ovvero sulla capacità di tener traccia di alcune regolarità nello spazio percettivo.

2.3. Popolazioni con poche parole e bambini che contano

In anni recenti, due popolazioni aborigene della foresta dell'Amazzonia – i Pirahã e i Mundurucu – sono state le protagoniste della ricerca in antropologia cognitiva¹⁴. La lingua di entrambe queste popolazioni ha poche parole per indicare numeri: i Pirahã possiedono parole precise solo per i numeri 1 e 2, mentre si riferiscono a 3 e 4 in maniera più irregolare; i Mundurucu usano invece con coerenza le parole per i numeri 1, 2 e 3, ma non sono coerenti quando parlano di quantità corrispondenti a 4 e 5. Gli esperimenti effettuati hanno tuttavia dimostrato che nonostante essi possiedano un numero di parole limitato a confronto con i nostri numerali, sono invece del tutto in grado di confrontare e addizionare – in maniera non precisa ma approssimata – numeri più grandi che superano le quantità per cui hanno parole. Nei due test cosiddetti della *distanza* e della *grandezza* – presentati con stimoli non simbolici come ad esempio insiemi di punti – Pirahã e Mundurucu si comportano in maniera del tutto simile a quella dei soggetti di controllo occidentali¹⁵. Questa circostanza suggerisce che i processi cognitivi che sottendono le loro manipolazioni numeriche – non simboliche – siano equivalenti a quelli di altre popolazioni dotate di lingue più sofisticate, perlomeno per quanto riguarda i termini introdotti per indicare numeri. Di conseguenza, lo studio del comportamento di queste popolazioni dimostra che gli essere umani alfabetizzati condividono con quelli illetterati – e in parte con gli animali non verbali – una rappresentazione delle quantità numeriche che risulta essere indipendente dalle loro capacità linguistiche.

¹⁴ Cfr. Gordon (2004) e Pica *et al.* (2004).

¹⁵ È stato dimostrato che indicare il più grande tra due numeri o tra due insiemi di oggetti – sia che vengano usati stimoli simbolici o non-simbolici – è più veloce e più accurato quando i due numeri o la dimensione dei due insiemi sono molto lontani fra loro. Ci si riferisce comunemente a questo effetto come all' 'effetto distanza' (*distance effect*). Inoltre, il tempo impiegato per indicare il più grande tra due numeri o due insiemi dipende anche dalla grandezza del numero o della dimensione dell'insieme. Ci si riferisce comunemente a questo effetto come all' 'effetto grandezza' (*size effect*). Cfr. Cappelletti & Giardino (2007).

Tuttavia, pur accettando che questi risultati dimostrino che Pirahã e Mundurucu hanno una qualche rappresentazione – seppure approssimata – di numerosità anche grandi, cerchiamo di capire quale sia il limite delle loro competenze. Per esempio, sanno contare? A giudicare da ciò che si ricava da questi studi, sembrerebbe che perlomeno i Pirahã non abbiano mai sviluppato una pratica analoga alla nostra che li metta nella condizione di contare, e non solo perché non posseggono le parole per i numerali, ma soprattutto perché, come si vede in alcuni nuovi test dove viene loro richiesto di effettuare corrispondenze biunivoche, non sembrano nemmeno essere in grado di utilizzare spontaneamente le dita per contare. Per qualche ragione, pare infatti che essi si fermino alla capacità di distinguere tra numerosità, ma non riescano a tener traccia di numerosità attraverso lo spazio e il tempo¹⁶. In effetti, suggeriscono gli sperimentatori, è solo quando una cultura sviluppi una pratica con il fine di contare che oggetti esterni alla mente come le stesse dita o semplici tacche segnate su un ramoscello possono assumere il ruolo di ‘segnaposti’, che tengono traccia ‘per noi’ delle diverse numerosità. Lasciamo per un attimo da parte operazioni aritmetiche più complesse e guardiamo semplicemente al modo in cui impariamo a contare: le parole che utilizziamo – ‘uno’, ‘due’, ‘tre’, ... – non sono in fondo nient’altro che i migliori segnaposti che la nostra cultura abbia prodotto con il fine di tener traccia di numerosità, e dunque in un certo senso la migliore tecnologia cognitiva a nostra disposizione. Avremmo certo potuto continuare a contare sulle dita – ma sarebbe stato poi difficile andare oltre 10 – oppure affidarci a dei segni – ma avremmo allora dovuto portare sempre con noi un gessetto e una lavagnetta. Si potrebbe obiettare che sebbene le popolazioni che abbiamo visto posseggano una lingua più povera della nostra almeno per quanto riguarda i possibili segnaposti per le numerosità, questo non voglia dire che non abbiano una propria cultura: una trappola in cui gli antropologi cognitivi rischiano ripetutamente di cadere è di imporre il punto di vista occidentale alla considerazione di pratiche culturali differenti. Certamente, perlomeno in linea di principio, queste popolazioni avrebbero potuto sviluppare pratiche del tutto dissimili dalle nostre, come ad esempio l’uso sistematico delle parti del corpo per contare¹⁷. Tuttavia, non sembra che questo sia il caso dei Pirahã, che, seguendo il resoconto dei loro osservatori, di fatto non mettono in pratica nessun tipo di strategia per tener traccia degli oggetti da contare, per esempio utilizzando le dita o i movimenti

¹⁶ Frank *et al.* (2008).

¹⁷ Cfr. Butterworth (1999).

oculari. In breve, essi non hanno sviluppato alcun tipo di strategia non verbale che li aiuti a contare¹⁸.

Vediamo invece come si comportano i bambini prescolari, che hanno già imparato a contare ma a cui non è stato ancora insegnato come risolvere operazioni aritmetiche. Un esperimento ha messo alla prova l'aritmetica 'spontanea' di bambini tra i 5 e i 6 anni, di diversa estrazione economica e sociale. Questa forma di aritmetica si baserebbe non sull'istruzione ma (i) sulla capacità dei bambini a distinguere numerosità e (ii) sulle strategie che essi hanno sviluppato per contare¹⁹. Ai bambini vengono dati da risolvere problemi come il seguente (Fig. 2.2):

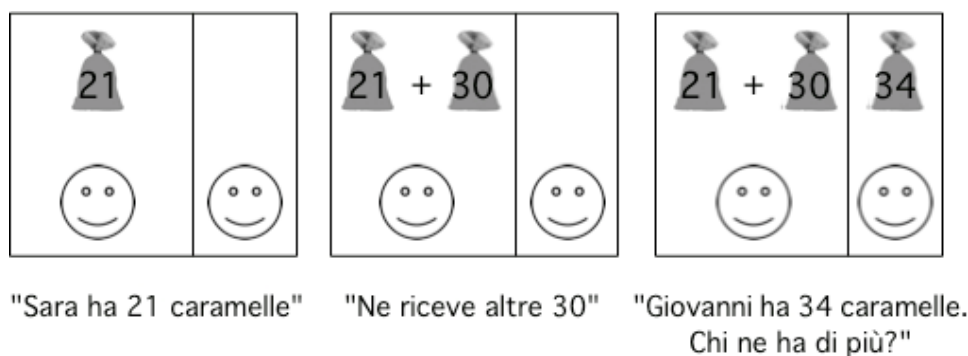


Fig. 2.2. Un esempio di prova di aritmetica prescolare

L'obiettivo degli sperimentatori è di capire se i bambini messi di fronte a problemi di questo genere fanno affidamento sulle loro rappresentazioni non simboliche e approssimate, quando anche i numeri vengano loro presentati attraverso numerali arabi. I risultati hanno mostrato che le risposte date dai bambini a questi problemi simbolici sono caratterizzate dai medesimi limiti che si sono già osservati in altri bambini, come anche nei neonati e negli ani-

¹⁸ Molto interessante a questo proposito è il libro di Daniel Everett che racconta la sua esperienza decennale di ricercatore (e in un primo momento missionario) tra i Pirahã. Nel libro, il linguista argomenta a favore dell'ipotesi che questa popolazione trasmetta il sapere attraverso sequenze di azioni o attraverso parole che però devono provenire da chi è stato testimone diretto o da chi ha ascoltato un testimone diretto dei fatti. Secondo Everett, questo principio, da lui chiamato *immediacy of experience principle*, spiegherebbe la mancanza presso i Pirahã di una tradizione orale e di qualsiasi rituale, nonché di parole per generalizzazioni specifiche come i numeri o i colori. Cfr. Everett (2008).

¹⁹ Gilmore *et al.* (2007).

mali non umani, per operazioni analoghe con stimoli non simbolici: la precisione delle risposte diminuisce man mano che il rapporto tra i due numeri da sommare o da confrontare si avvicina a 1. Le operazioni di addizione e di confronto sono ugualmente accurate, mentre l'operazione di sottrazione risulta più difficile. La corrispondenza tra questi due risultati in contesti diversi mostrerebbe che i bambini prescolari, a cui ancora non sono state insegnate le operazioni aritmetiche, fanno affidamento sul loro sistema non simbolico di rappresentazione, già utilizzato per distinguere tra numerosità, per risolvere problemi presentati simbolicamente. È sorprendente come questo permetta loro – per quanto entro certi limiti – di trovare la soluzione corretta a operazioni aritmetiche presentate tramite numerali arabi, prima ancora che queste operazioni aritmetiche vengano loro insegnate utilizzando simboli. L'ipotesi che voglio fare qui è che questo accade perché i bambini hanno già: (i) memorizzato i numerali nel loro ruolo di segnaposti verbali per contare; (ii) imparato a creare una corrispondenza tra i numerali e la propria rappresentazione non simbolica dei numeri a cui i numerali si riferiscono. A queste condizioni, essi sviluppano spontaneamente procedure di ragionamento che si basano sul proprio sistema non simbolico; queste procedure permettono loro di manipolare quantità anche quando esse siano presentate simbolicamente. Se questo è vero, allora il loro sistema non simbolico viene attivato ogni volta che venga loro richiesto di creare le corrispondenze biunivoche che hanno imparato a costruire contando, senza che questo dipenda da quanta aritmetica sia stata loro insegnata.

In un esperimento più recente, è stato mostrato che i bambini sono in grado di creare corrispondenze tra simbolico e non simbolico in entrambe le direzioni (sebbene siano meno precisi nella corrispondenza dal non simbolico al simbolico) e che questa capacità si sviluppa tra i 6 e gli 8 anni ed è collegata alla loro riuscita scolastica²⁰. Questo dimostrerebbe che le rappresentazioni non simboliche sottese al ragionamento simbolico avrebbero un ruolo nello sviluppo delle conoscenze matematiche del bambino. Secondo gli sperimentatori, i risultati ottenuti su questa forma di aritmetica spontanea, approssimata e non simbolica, avrebbero interessanti conseguenze sui modi in cui bisognerebbe insegnare l'aritmetica nei primi anni di scuola elementare.

²⁰ Cfr. Mundy & Gilmore (2009).

2.4. Tre livelli di competenza in aritmetica

Per dare un'interpretazione ai risultati che ho presentato brevemente fin qui, propongo di distinguere tra tre diversi livelli di competenza che sembrano essere stati individuati nel corso degli esperimenti e che sono necessari per fare aritmetica. Al livello I, troviamo la capacità di *distinguere* tra *numerosità* diverse, ovvero di paragonare due insiemi. Questa capacità sembra riferirsi ai medesimi processi cognitivi sia che le numerosità vengano presentate attraverso insiemi di punti o di oggetti, sia attraverso simboli come i numerali arabi – nel secondo caso, è ovviamente richiesta una competenza ulteriore. A questo livello, la matematica che riusciamo a fare è una matematica solo 'approssimata': due quantità vengono messe a confronto e si giudica quale sia la più grande o la più piccola. Questo tuttavia non vuol dire essere in grado di individuare il valore assoluto delle due quantità: perché l'aritmetica diventi 'precisa', è necessario salire al livello II, che, a mio modo di vedere, costituisce un ponte tra l'aritmetica approssimata e l'aritmetica precisa che potrà essere effettuata a livello III.

Al livello II, troviamo infatti la capacità di *contare*: è proprio contando che si assegnano numerosità precise agli elementi di un insieme, anche andando oltre il raggio della 'subitizzazione'²¹. Se consideriamo il contare come 'processo', riconosciamo che esso fa uso dei numeri come mere posizioni all'interno di una sequenza ordinata: contando, si costruisce una corrispondenza biunivoca tra gli oggetti da contare e degli specifici segnaposti – ad esempio le dita – che vengono introdotti al posto degli oggetti da contare²². Una volta che una pratica per contare si sia stabilizzata, i numeri naturali – o meglio, i numerali che ad essi si riferiscono – costituiscono i migliori segnaposti che la nostra cultura ci abbia trasmesso.

Al livello III, abbiamo infine l'aritmetica come la conosciamo: i numeri diventano entità astratte, e chi pratica l'aritmetica ha ormai sviluppato per essi

²¹ È stata definita 'subitizzazione' la capacità di riconoscere immediatamente, senza contare in maniera progressiva e seriale, di quanti elementi è costituito un insieme fino a insiemi di 3-4 elementi. Cfr. Dehaene (2010).

²² Si potrebbe obiettare che il contare non è ordinale come affermo qui, dal momento che il risultato del contare è un risultato cardinale: alla fine del processo arriviamo al numero totale di oggetti contati (Gallistel & Gelman (2005)). Carey si basa su questa concezione cardinale del contare e sui *counting principles* definiti in Gelman & Gallistel (1978) per sostenere la sua teoria di un (*Quinean*) *bootstrapping* che permetterebbe ai bambini di costruire una lista di numerali per rappresentare i numeri naturali (Carey (2009), cfr. nota 13). Per ragioni di spazio, non posso difendere qui l'idea che vi sono ragioni per considerare il contare un'attività ordinale; tuttavia, non mi pare che una tale obiezione, anche se fondata, invalidi l'argomento complessivo presentato in questo articolo.

concetti. A questo livello, vengono effettuate operazioni più complesse – siamo ormai oltre il semplice confronto tra quantità e oltre il contare – e in questo senso il linguaggio formale gioca un ruolo cruciale, poiché permette di avere un accesso più facile e più rapido alle diverse possibili operazioni.

La Tabella 2.1 riassume i tre livelli di competenza per l'aritmetica.

| Aritmetica |
|---------------------------------------|
| I. distinguere numerosità |
| II. contare |
| III. comprendere l'aritmetica formale |

Tabella 2.1: i tre livelli di competenza per l'aritmetica.

Se accettiamo questa distinzione, possiamo concludere che gli animali, i neonati, i bambini fino a un anno, e le popolazioni di cui abbiamo parlato non sembrano andare oltre il livello I, mentre bambini più grandi che abbiano sviluppato una pratica del contare raggiungono il livello II. Tornerò su questi tre livelli nel paragrafo conclusivo dell'articolo.

3. Cognizione e geometria

3.1. Animali e bambini

Altri studi si sono interessati a svelare le conoscenze geometriche di animali e bambini molto piccoli.

L'esperimento sugli animali considerato forse il più significativo è quello in cui si è dimostrato che i topi si basano su indizi unicamente di carattere geometrico – come ad esempio la forma della stanza – per orientarsi e ritrovare del cibo²³. In un primo momento, gli animali vengono abituati a scoprire del cibo sempre in un medesimo angolo di una stanza rettangolare; in seguito, quando il cibo viene nascosto altrove all'interno della stanza, i topi mostrano di andare a cercarlo nell'angolo dove erano stati abituati a ritrovarlo, oppure nel suo equivalente geometrico – un errore che viene chiamato *errore rotazionale* (Fig. 3.1). Studi successivi hanno mostrato che sebbene i topi siano in grado di distinguere anche altre proprietà caratteristiche della stanza, per

²³ Cheng (1986).

qualche ragione non le usano per ritrovare il cibo, attenendosi unicamente a proprietà geometriche²⁴.



Fig. 3.1 L'errore rotazionale

Tuttavia, dal fatto che le capacità cognitive degli animali si siano adattate all'ambiente in modo da permetterne una migliore e più efficace navigazione non si può concludere che i topi o altri animali posseggano abilità spaziali sofisticate: il loro comportamento ci indica che essi sono certamente in grado di utilizzare elementi geometrici come indizi per navigare lo spazio – e lo fanno in maniera sistematica e coerente – ma questo può semplicemente dipendere dalla loro capacità di riconoscere invarianze percettive in un ambiente che hanno in precedenza già esplorato.

L'esperimento è stato replicato con bambini tra i 18 e i 24 mesi che, in maniera del tutto analoga ai topi, utilizzano indizi geometrici per cercare un giocattolo che è stato nascosto in una stanza rettangolare, anche quando abbiano a disposizione non solo informazioni di carattere geometrico. I bambini vedono che gli sperimentatori nascondono un oggetto in un angolo della stanza e poi vengono disorientati. Devono riorientarsi basandosi sulla forma della stanza e in certi casi anche grazie a specifici indizi non geometrici, come il colore di uno dei muri. Diversamente dagli adulti, i bambini fino a 2 anni si orientano solo facendo riferimento alla geometria della stanza, e cadono per questo nel consueto errore rotazionale²⁵. Bambini più grandi, tra i 5 e i 7 anni, mostrano invece di acquisire gradualmente un comportamento che si adatta maggiormente alle caratteristiche dell'ambiente, e di fare sempre maggiore

²⁴ Va tuttavia notato che non tutte le specie fanno riferimento esclusivamente alla geometria per orientarsi. Altri animali, come per esempio le galline, i pesci rossi e le lucertole, si orientano utilizzando sia elementi geometrici che altre proprietà caratteristiche. Cfr. De Cruz (2007).

²⁵ Hermer & Spelke (1994); Hermer & Spelke (1996).

attenzione per orientarsi non solo alle relazioni geometriche ma anche a punti di riferimento di altro genere²⁶. Questi risultati indicano che i bambini crescendo imparano a basarsi su più di un tipo d'informazione, guardando alla geometria della stanza come anche ad altre caratteristiche dei suoi elementi, come poi continueranno a fare anche da adulti; tuttavia, è curioso notare che, prima di sviluppare un'attenzione ad altre proprietà, si comportano esattamente come i topi.

3.2. Forme e mappe

In anni più recenti, alcuni nuovi studi si sono occupati di valutare le capacità anche geometriche dei Mundurucu. In una prima prova, l'obiettivo è stato di valutare la comprensione intuitiva da parte dei membri di questa popolazione dei concetti di base della geometria: punti, linee, parallelismi, figure, congruenze e simmetrie. Per ciascuno di questi concetti, vengono loro mostrati sullo schermo di un computer sei oggetti che esemplificano tutti – fatta eccezione di uno – il concetto in questione (si veda la Fig. 3.2).

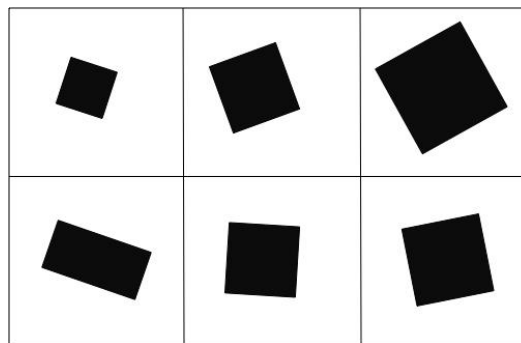


Fig. 3.2. Immagine sullo schermo corrispondente al concetto 'quadrato'

Viene quindi loro chiesto di indicare quale – tra i sei che vedono sullo schermo – sia l'oggetto 'strano', o 'brutto'²⁷. I risultati mostrano che i Mundurucu adulti – che appartengono dunque a questa cultura isolata e non hanno studiato – non hanno alcun problema con i concetti di base concernenti la topologia, la geometria euclidea, e le figure geometriche di base; le prime

²⁶ Hermer-Vazquez *et al.* (2001).

²⁷ Dehaene *et al.* (2006).

difficoltà si presentano per loro nel riuscire a individuare simmetrie e proprietà metriche, anche se ancora riescono a rispondere al compito. Non sono invece affatto in grado di comprendere numerose trasformazioni geometriche, nelle quali si richiede di immaginare di manipolare mentalmente una forma per ottenerne un'altra, e poi di riflettere sulla natura di questo cambiamento. Questo pattern di risposte coincide con quelle date dai bambini Munduruku – che ora frequentano la scuola e parlano portoghese – nonché dai bambini americani di controllo: tutti questi soggetti sono in grado di estrarre invarianti geometriche genuine dalle forme degli oggetti che vengono loro mostrati. Questa capacità sembra essere analoga a quella di distinguere tra numerosità diverse; inoltre, gli sperimentatori sottolineano come anche qui i risultati sembrano andare ancora una volta contro Piaget, che nella sua teoria sosteneva che nello sviluppo del bambino, si procedesse dalla topologia alla geometria proiettiva arrivando infine alla geometria euclidea²⁸.

Tuttavia, la capacità di estrarre invarianze non sembra ancora implicare che queste invarianze possano poi essere usate per rappresentare simbolicamente lo spazio circostante, come accade quando si costruisce o si utilizza una mappa. Con un nuovo esperimento, gli antropologi decidono di mettere nuovamente alla prova i Munduruku, valutando la loro capacità di applicare la conoscenza geometrica 'astratta' che hanno mostrato di possedere trasferendola dalla percezione a un contesto del tutto nuovo. Voglio segnalare fin da ora che trovo tuttavia inappropriato parlare, come fanno gli sperimentatori, di conoscenza geometrica 'astratta' già al livello del riconoscimento di invarianti geometrici. Come vedremo, sosterrò che anche in questo caso come nel caso della distinzione tra numerosità, siamo ancora all'interno di un contesto percettivo dal quale non si può prescindere.

I Munduruku devono ora risolvere un problema basandosi su una mappa. Il design dell'esperimento prevede tre contenitori, in uno dei quali vi è un oggetto nascosto. Ai partecipanti viene data una mappa sulla quale è rappresentata la disposizione dei contenitori. Una stella indica il contenitore che nasconde l'oggetto. Anche qui, come per l'esperimento precedente, le risposte di adulti e bambini Munduruku coincidono con quelle dei bambini americani, mentre gli americani adulti rispondono significativamente meglio. In ogni modo, secondo gli sperimentatori, il fatto che i Munduruku comprendano spontaneamente e intuitivamente concetti geometrici e mappe può farci concludere che un sistema di conoscenze geometriche di base, come quello per l'aritmetica elementare, è un costituente universale della mente umana.

²⁸ Piaget (1948/1960).

Ricordiamo a questo proposito che già Las Casas aveva fornito un'evidenza aneddotica a quest'affermazione, raccontando la storia di come il re del Portogallo avesse deciso di interrogare dei nativi per capire da loro se ci fossero altre isole nel Nuovo Mondo da esplorare (e da sfruttare)²⁹. Il re convocò un nativo – che non parlava portoghese – e gli fece portare dei fagioli perché li utilizzasse per costruire una mappa delle isole da dove proveniva. A questa richiesta, l'uomo, in maniera del tutto naturale e senza alcuna esitazione, collocò a terra i fagioli in modo che rappresentassero le isole di Ispaniola, di Cuba, le Lucayas e altre ancora. Il re chiese quindi a un altro nativo lì presente di aggiungere se necessario altri fagioli per mostrare le terre che si trovavano nell'arcipelago da cui Colombo li aveva prelevati, e questo secondo aborigeno integrò la configurazione delle isole disposta dal compagno con altri fagioli – e quindi altre isole – spiegando nella sua lingua di quali isole si trattasse – senza che ovviamente nessuno dei portoghesi fosse in grado di capire quello che andava raccontando.

Tuttavia, è a mio avviso necessario mettere in evidenza due questioni. La prima è che nell'interpretare il comportamento dei Munduruku o dei nativi d'America nel racconto di Las Casas, bisogna far attenzione a non confondere la capacità di riconoscere invarianze spaziali o quella di utilizzare mappe con la familiarità con concetti geometrici. Al contrario, anche in questo caso è utile distinguere tra diversi livelli di competenza. Nel primo esperimento, i Munduruku mostrano di essere in grado di estrarre invarianti geometriche, nel confronto tra diverse forme; nel secondo, di sapere utilizzare i medesimi invarianti in un compito simbolico come la comprensione di una mappa. Non c'è però ancora alcuna evidenza che essi posseggano una conoscenza geometrica astratta. In secondo luogo, i risultati dell'esperimento mostrano che la cultura e l'educazione hanno di fatto un peso sulle risposte dei soggetti, dal momento che gli adulti americani estraggono invarianti con maggiore facilità, e sono anche più veloci sia nell'individuare trasformazioni geometriche che nell'orientarsi con una mappa. Quello che voglio suggerire è che le prestazioni degli adulti occidentali sono migliori perché questi ultimi si sono formati in un percorso d'istruzione, sono immersi in una cultura che fa uso costante di artefatti cognitivi come le mappe, e infine padroneggiano una lingua che ha dato un nome alle diverse forme, ai diversi invarianti geometrici che sono comunque in grado di riconoscere. Tutti questi elementi non sono solo accessori, ma al contrario permettono loro di costruire un edificio più robusto sulla base delle loro intuizioni fondamentali³⁰.

²⁹ De las Casas (1875).

³⁰ Newcombe & Uttal (2006).

Come nel caso dell'aritmetica, anche per quanto riguarda la geometria è per questo cruciale prendere in considerazione lo sviluppo all'interno di una comunità di pratiche complesse come quella dell'uso di mappe e di analoghi strumenti cognitivi ai quali ancora una volta demandiamo del lavoro cognitivo, perché ci siano da supporto per orientarci nell'ambiente che ci circonda. Da quello che si legge nell'articolo, i Munduruku, sebbene vivano in numerosi villaggi sparsi per la foresta tra i quali sono in grado di orientarsi senza problemi, e sebbene utilizzino utensili di vario genere, non fanno 'tipicamente' uso di mappe e tantomeno si adoperano 'spontaneamente' per disegnare le loro case, i loro villaggi o l'ambiente in cui vivono. Attività come quella di tracciare un cerchio sulla sabbia per rappresentare un villaggio sono state osservate solo 'occasionalmente', e in ogni modo si tratta di disegni molto grossolani, che non preservano l'informazione metrica sugli angoli o le distanze tra i riferimenti come invece richiesto dall'esperimento per potersi orientare con la mappa³¹. Ci sembra dunque che queste informazioni siano sufficienti per concludere che i Munduruku non hanno mai sviluppato una pratica di rappresentazione dello spazio che li circonda tramite mappe (analogamente a quello che abbiamo visto per i Pirahã riguardo al contare).

Di nuovo, uno sguardo alla psicologia dello sviluppo può essere rilevante. In un esperimento recente, il medesimo test con i Munduruku è stato ripresentato a un gruppo di bambini di 4 anni. Come affermano gli sperimentatori, sebbene tutti gli animali siano in grado, come abbiamo visto nel paragrafo precedente, di rappresentare in qualche modo lo spazio che li circonda con il proposito di trovare strategie per orientarsi, solo gli esseri umani sono capaci – e lo fanno spontaneamente – di integrare queste rappresentazioni mentali creando e utilizzando mappe prodotte al di fuori della loro mente³². Il compito che i bambini devono risolvere è simile a quello proposto ai Munduruku, solo che in questo caso non si tratta di ritrovare un oggetto ma di sistemare un giocattolo nel luogo indicato dallo sperimentatore su una mappa. Per esempio, lo sperimentatore dice al bambino: «Questo disegno ti dice dove si vuole sedere Ranocchietto. Ranocchietto vuole sedersi dove c'è la stella. Mi puoi indicare la stella?» ... «Bravissimo!» ... «Ora accontenta per favore Ranocchietto e mettilo lì dove si vuole sedere». Altra differenza con l'esperimento condotto con i Munduruku è che ai bambini non viene detto se quello che hanno scelto sia davvero il posto dove Ranocchietto voleva sedersi.

³¹ Dehaene *et al.* (2006, 383).

³² Shusterman *et al.* (2008).

I risultati mostrano che i bambini sono in effetti in grado di eseguire il compito correttamente, e questo vuol dire che sviluppano molto presto e spontaneamente la capacità di stabilire e riconoscere corrispondenze geometriche tra la disposizione di oggetti nelle tre dimensioni e una mappa bidimensionale. Inoltre, non hanno alcuna difficoltà a utilizzare queste corrispondenze per orientarsi. Se accettiamo questa conclusione, allora sembrerebbe che alcune proprietà metriche estratte dall'ambiente siano elaborate e utilizzate spontaneamente anche da bambini prescolari, ai quali difatti non è stato ancora insegnato come risolvere un problema geometrico presentato attraverso simboli. Anche in questo caso, come in quello dei Mundurucu, non possiamo tuttavia ricavarne che i bambini posseggano di conseguenza una conoscenza astratta dei concetti geometrici; piuttosto, potremmo chiederci se le loro competenze siano già in qualche modo diverse da quelle della popolazione amazzonica, solo per il fatto che essi stanno crescendo in una cultura che fa un uso comune e costante di mappe.

3.3. Tre livelli di competenza in geometria

Per dare un'interpretazione ai risultati presentati fin qui, propongo di distinguere tra tre livelli di competenza anche per quanto riguarda la geometria. Al livello I, troviamo la capacità di *estrarre invarianze* dallo spazio circostante, capacità che ci permette di eseguire compiti come quello di imparare a orientarci in un ambiente che abbiamo in precedenza esplorato. Al livello II, troviamo l'*utilizzo di mappe* e di altri strumenti cognitivi analoghi che ci forniscono una *rappresentazione esterna* – prodotta al di fuori della nostra mente – sulla quale facciamo affidamento per orientarci in un ambiente, risparmiando costi cognitivi. Al livello III, abbiamo infine la geometria come siamo abituati a pensarla: le forme sono diventate entità astratte, e per esse sono stati sviluppati concetti. A quest'ultimo livello, nuovi e più sofisticati artefatti cognitivi – sistemi di rappresentazioni geometriche e linguaggi formali – vengono introdotti per permetterci di accedere e di manipolare più facilmente e con più sicurezza gli oggetti geometrici. I tre livelli sono riassunti nella Tabella 3.1.

| Geometria |
|--|
| I. estrarre invarianze |
| II. utilizzare mappe |
| III. comprendere la geometria astratta |

Tabella 3.1: i tre livelli di competenza per la geometria

Gli animali – compresi i topi – non sembrano andare oltre il livello I. I bambini prescolari raggiungono senz'altro il livello II, come anche molto probabilmente le popolazioni non dotate di una lingua come la nostra, sebbene non sembrano sempre in grado di sviluppare spontaneamente pratiche stabili che permettano loro di condividere questi strumenti con il resto della comunità. In tutti questi casi, siamo ancora lontani dal livello III.

4. Il livello II: siamo la specie simbolica

In quest'ultima sezione, vorrei riassumere la discussione dei risultati sperimentali che ho presentato brevemente fin qui e concentrarmi sul secondo livello di competenza che ho identificato (Tabella 4.1).

| Aritmetica | Geometria |
|-------------|----------------------|
| II. contare | II. utilizzare mappe |

Tabella 4.1: il secondo livello di competenze in aritmetica e geometria

La caratteristica interessante di questo livello è che l'attività cognitiva che lo riguarda fa riferimento sia nel caso dell'aritmetica sia della geometria all'introduzione e all'utilizzo di uno strumento cognitivo o di una tecnologia cognitiva – che si tratti di una tacca per tener traccia di un elemento di un insieme o di uno schizzo per tener traccia della struttura di un ambiente. Vediamo cosa questo significa. Innanzitutto, è proprio al livello II che l'interazione tra natura e cultura appare in tutta la sua forza: le abilità cognitive richieste a livello I sono condizione necessaria per accedere al livello II, poiché essere in grado di discernere numerosità è necessario per contare così come lo è essere in grado di estrarre invarianze geometriche per utilizzare delle mappe. Tuttavia, questa è condizione necessaria ma non sufficiente, poiché per raggiungere il livello II bisogna possedere un'ulteriore capacità, ovvero comprendere che oggetti come le tacche sul muro o le linee in un disegno bidimensionale possono essere usate per *'rappresentare'* qualcos'altro, per

‘*stare al posto di*’ qualcos’altro, sia che si tratti nel caso dell’aritmetica degli elementi di un insieme o della disposizione degli oggetti in un ambiente tridimensionale nel caso della geometria. Ovviamente, far parte di una comunità nella quale si sia stabilita una pratica come l’utilizzo di un particolare sistema numerico o di mappe, per esempio mappe orientate secondo i quattro punti cardinali, è senz’altro d’aiuto per raggiungere questo livello. È dunque a livello II che osserviamo l’influenza reciproca dell’elemento cognitivo-naturale del ragionamento su quello cognitivo-culturale, nella creazione, l’introduzione e la stabilizzazione di una particolare tecnologia cognitiva³³. Dal punto di vista della natura, sembra accadere qualcosa di distintivo nei bambini che hanno raggiunto i 4 anni: a questa età, essi sono in grado – e non sembrano esserlo prima – di contare e di utilizzare una mappa in maniera coerente. Dal punto di vista della cultura, una volta che una pratica si sia stabilizzata in una comunità, di fatto essa integra e potenzia le nostre capacità non simboliche. Mi sembra inoltre che la comparsa di strumenti cognitivi sia un passo cruciale anche per lo sviluppo delle competenze richieste a livello III: una volta che si sia acquisita familiarità con i vari simboli e i vari sistemi di rappresentazioni esterne grazie ai quali risparmiamo costi cognitivi, velocizzando e rendendo maggiormente precisa la risoluzione dei problemi che affrontiamo ogni giorno, questi simboli e queste rappresentazioni possono poi evolvere in tecnologie più complesse. In questo modo, si apre la strada allo sviluppo del livello III, al quale sono tuttavia richieste anche altre capacità, in primo luogo le nostre competenze linguistiche.

Vorrei ora soffermarmi sull’ipotesi che a livello II sia necessario riconoscere che qualcosa può essere utilizzato come ‘*rappresentazione di*’ qualcos’altro, e per questo tener traccia di oggetti ed eventi attraverso lo spazio e il tempo. Ciò che mi interessa sottolineare è che se questo è vero, allora è proprio il livello II a costituire le basi cognitive della matematica, poiché coinvolge un’attività cognitiva intermedia che va al di là dell’estrazione di

³³ Alcuni studi recenti hanno cercato di dimostrare che la rappresentazione numerica non ha un fondamento spaziale, facendo vedere in che modo i soggetti non hanno alcuna difficoltà a costruire corrispondenze anche tra una linea e stimoli non numerici, come ad esempio dei rumori. Questo mostrerebbe che esiste un ‘senso della grandezza’ che può manifestarsi sia spazialmente che non spazialmente, e che quindi la diffusa e studiata corrispondenza tra numeri e linea numerica è una costruzione culturale (Nunez *et al.* (2011)). Si andrebbe in questo modo contro l’idea che esista una sorta di linea numerica innata, come suggerito da Galton (1880). Effetti come lo *SNARC* (*Spatial numerical association of response code*) effect, ovvero l’effetto per cui i tempi di reazione nel rispondere a un numero alto con la mano destra sono minori che rispondere a un numero basso con la stessa mano e viceversa (Dehaene *et al.* (2003)), potrebbero probabilmente spiegarsi attraverso l’influenza che la scrittura ha sulla nostra rappresentazione dei numeri naturali.

caratteristiche concrete e percettive dall'ambiente circostante ma non è ancora simbolica in senso pieno, come accade nel caso del livello III. Il livello II fornisce quindi un primo livello di astrazione che è tuttavia coadiuvato da un supporto esterno, fisico, che ci permette di spingerci oltre la percezione verso la conoscenza astratta. Oltretutto, il livello II sembra essere tipicamente umano e fare esplicitamente riferimento alle nostre potenzialità simboliche: gli esseri umani sono 'la' specie simbolica. Come è stato suggerito, sebbene condividiamo il medesimo pianeta con milioni di altre specie viventi, viviamo anche in altri luoghi ai quali nessuna altra specie ha accesso: abitiamo un mondo pieno di astrazioni, di impossibilità, di paradossi³⁴. Deacon sostiene che sia stata l'evoluzione del linguaggio ad aprirci la porta che conduce verso questo mondo virtuale, perché il linguaggio non è solo un canale di comunicazione, ma anche l'espressione esteriorizzata di un modo di pensare inusuale in natura: la rappresentazione simbolica. Senza questo processo di simbolizzazione, questi luoghi inaccessibili ad altre specie sarebbero per noi irraggiungibili e persino impossibili da concepire. Nel modo in cui leggo Deacon, linguaggio va qui inteso in senso ampio, come appunto processo di rappresentazione simbolica: l'utilizzo di tacche per contare o di mappe per orientarsi è già rappresentazione simbolica, sebbene non si possa considerarlo linguaggio in senso stretto. In effetti, è vero che la familiarità con il linguaggio – in senso stretto – influisce sulle nostre capacità a livello II, per esempio sul contare o sull'orientarsi, e tuttavia non è strettamente necessaria perché esse si manifestino: come suggerisce Locke nel passaggio citato nell'introduzione, il linguaggio – di nuovo in senso stretto – contribuisce alla nostra capacità di far di conto, ma non ne è un prerequisito.

Se l'immagine che propongo è verosimile, allora essa ha come conseguenza che lo studio della cognizione in generale e delle basi cognitive della matematica in particolare debba concentrarsi sul livello II e sull'introduzione di rappresentazioni esterne e artefatti cognitivi, e strutturarsi in un'ottica interdisciplinare: come Bruner metteva in evidenza già nel 1990, se la rivoluzione cognitiva si concentrasse sulle attività simboliche proprie degli esseri umani nel costruire e dar senso non solo al mondo ma anche a se stessi, allora la psicologia avrebbe l'occasione di mettere insieme le sue forze con quelle delle discipline interpretative che le sono sorelle, le scienze umane e sociali³⁵. Le basi cognitive della matematica verrebbero così ritrovate davvero nell'in-

³⁴ Deacon (1997).

³⁵ Bruner (1990).

tersezione tra natura e cultura, tra mente e tecnologia, e nell'esercizio di capacità cognitive pre-esistenti accompagnato e influenzato da pratiche culturali e dall'uso di strumenti cognitivi.

5. Conclusioni

In questo articolo, ho voluto mostrare che c'è un interesse nel considerare gli studi empirici per comprendere quali siano i fondamenti cognitivi della matematica umana, ma che bisogna anche riconoscere quali siano i limiti di questi studi, e interrogarsi su cosa ci dicono delle nostre capacità. Per parlare di 'sistemi di base' riguardanti la conoscenza aritmetica e geometrica, è utile per prima cosa definire diversi livelli di competenza. Nella prospettiva che propongo, i livelli di competenza sono tre: il primo livello è direttamente legato a questi sistemi di base e riguarda l'organizzazione della nostra percezione, mentre il terzo livello fa riferimento a costruzioni culturali e sociali sofisticate e alla conoscenza formale e astratta. In questa ricostruzione, assume rilevanza il secondo livello, poiché è a questo livello che natura e cultura si trovano a interagire in maniera più evidente e attiva: è qui che si sviluppano pratiche simboliche, che vengono creati artefatti cognitivi, e che gli esseri umani si affidano a rappresentazioni esterne e potenzialmente pubbliche, che permettono loro non solo di affinare le loro capacità, ma anche di condividerle con il resto della comunità e in questo modo aprire la strada a un loro ulteriore sviluppo.

Altro lavoro è necessario per comprendere appieno la relazione tra questi livelli differenti, e quale influenza abbiano eventualmente l'uno sull'altro. Ritengo in ogni modo che solo una prospettiva di questo genere renda possibile determinare quali siano i fondamenti cognitivi della matematica umana e cosa ci porti dalla percezione di numerosità e forme intorno a noi alla matematica che conosciamo.

Bibliografia

- Bruner, J. (1990), *Acts of Meaning*, Cambridge (MA), Harvard University Press.
- Butterworth, B. (1999), *The Mathematical Brain*, London, Macmillan (*Intelligenza matematica*, trad. it. di C. C. Apararo e S. Mancini, Milano, Rizzoli, 1999).
- Campbell, J. (ed.) (2005), *Handbook of mathematical cognition*. New York (NY), Psychology Press.
- Cappelletti, M., Giardino, V. (2007), “The cognitive basis of mathematical knowledge”, in M. A. Leng, M. Paseau & Potter (a cura di), *Mathematical Knowledge*, Oxford, Oxford University Press, pp. 73 – 83.
- Carey, S. (2009), “Where our number concepts come from”, *Journal of Philosophy*, CVI, 4, pp. 220–254.
- Cheng, K. (1986), “A purely geometric module in the rats spatial representation”, *Cognition*, XXIII, pp. 149 – 178.
- Church, R. M., Meck, W. H. (1984), “The numerical attribute of stimuli”, in Roitblat H. L., Bover, T. G., Terrace, H. S. (a cura di), *Animal Cognition*, Hillsdale, N. J., Lawrence Erlbaum Associates, pp. 445–464.
- Deacon, T. W. (1997), *The Symbolic Species*, New York e Londra, W. V. Norton Company.
- Dehaene, S. (1999), *The Number Sense: How the Mind Creates Mathematics*, Oxford, Oxford University Press (*Il pallino della matematica. Scoprire il genio dei numeri che è in noi*, trad. it. di M. L. Vesentini Ottolenghi e G. Guerriero, Nuova edizione, Milano, Raffaello Cortina Editore, 2010).
- Dehaene, S., Bossini, S., Giraux, P. (1993), “The mental representation of parity and number magnitude”, *Journal of Experimental Psychology*, CXXII, 3, pp. 371 – 396.
- Dehaene, S., Izard, V., Pica, P., Spelke, E. (2006), “Core Knowledge of Geometry in an Amazonian Indigene Group”, *Science*, CCCXI, pp. 381 – 384.

- De Cruz, H. (2007), “An Enhanced Argument for Innate Elementary Geometric Knowledge and its Philosophical implications”, in B. Van Kerkhove (ed.), *New Perspective on Mathematical Practices*, World Scientific, pp. 185 – 206.
- De Cruz, H., Neth, H., Schlimm, D. (2010), “The cognitive basis of arithmetic”, in B. Lôwe, T. Müller (a cura di), *PhiMSAMP*, College Publications, pp. 58 – 106.
- De las Casas, B. (1875), *Historia de las Indias* (Libro I, Cap. 74).
- Everett, D. L. (2008), *Don't Sleep, There Are Snakes: Life and language in the Amazonian Jungle*, Vintage Departures.
- Feigenson, L., Carey S., Hauser M. (2002), “The representations underlying infants' choice of more: object files versus analog magnitudes”, *Psychological Science*, XIII, 2, pp. 150 – 156.
- Frank, M. C., Everett, D. L., Fedorenko, E., Gibson, E. (2008), “Number as a cognitive technology: Evidence from Pirahã language and cognition”, *Cognition*, CVIII, 3, pp. 819 – 824.
- Gelman, R., Gallistel, C. R. (1978), *The Child's Understanding of Number*, Cambridge (MA), Harvard University Press.
- Gallistel, C. R., Gelman, R. (2005), “Mathematical Cognition”, in K. Holyoak, R. Morrison (a cura di), *The Cambridge handbook of thinking and reasoning*, Cambridge University Press, pp. 559 – 588.
- Gilmore C. K., McCarthy, S. E., Spelke, E. S. (2007), “Symbolic arithmetic knowledge without instruction”, *Nature*, CDXLVII, pp. 589 – 591.
- Gordon, P. (2004), “Numerical cognition without words: Evidence from Amazonia”, *Science*, CCCVI, pp. 496 – 9.
- Hermer, L., Spelke, E. S. (1994), “A geometric process for spatial reorientation in young children”, *Nature*, CCCLXX, pp. 57 – 59.
- Hermer, L., Spelke, E. S. (1996), “Modularity and development: The case of spatial reorientation”, *Cognition*, LXI, pp. 195 – 232.

- Hermer-Vazquez, L., Moffet A., Munkholm, P. (2001), “Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks”, *Cognition*, LXXIX, pp. 263 – 299.
- Hutchins, E. (2001), “Distributed cognition”, *The International Encyclopedia of the Social and Behavioral Sciences*, pp. 2068 – 2072.
- Kant, I. (1781), *Critica della ragion pura*, Roma-Bari, Laterza, 2000, pp. 57-58.
- Kinzler, K. D., Spelke, E. S. (2007), “Core systems in human cognition”, *Progress in Brain Research*, CLIV, pp. 257 – 264.
- Köhler, O. (1951), “The ability of birds to count”, *Bulletin of Animal Behaviour*, IX, pp. 41 – 45.
- Locke, J. (1690), *Saggio sull’Intelletto umano*, Torino, UTET, 1971.
- Matsuzawa, T. (1985), “Use of numbers by a chimpanzee”, *Nature* CCCXV, pp. 57 – 59.
- Mundy, E., Gilmore, C. K. (2009), “Children’s mapping between symbolic and nonsymbolic representations of number”, *Journal of Experimental Child Psychology*, CIII, pp. 490-502.
- Newcombe, N. C., Uttal, D. H. (2006), “Whorf versus Socrates, Round 10”, *Trends in Cognitive Science*, X, 9, pp. 394 – 396.
- Núñez, R., Doan, D., Nikoulina, A. (2011), “Squeezing, striking, and vocalizing: Is number representation fundamentally spatial?”, *Cognition*, CXX, 2, pp. 225–35.
- Piaget, J. (1948/1960), *The Child’s Conception of Geometry*, Londra, Routledge and Kegan Paul.
- Piaget, J. (1952), *The Child’s Conception of Number*, Londra, Routledge and Kegan Paul.
- Pica P., Lemer C., Izard V., Dehaene, S. (2004), “Exact and approximate arithmetic in an Amazonian indigene group”, *Science*, CCCVI, pp. 499 – 503.

- Shusterman, A., Lee, S. A., Spelke, E. S. (2008), “Young children’s spontaneous use of geometry in maps”, *Developmental Science*, XI, 2, F1 – F7.
- Uller, C., Carey, S., Huntley-Fenner G., Klatt, L. (1999), “What representations may underlie infant numerical knowledge”, *Cognitive Development*, XIV, pp. 1 – 36.
- Wynn, K. (1992), “Addition and subtraction by human infants”, *Nature*, CCCLVIII, pp. 749 – 50.

Il dilemma del prigioniero iterato: un problema aperto

Vincenzo Fano
Università degli Studi di Urbino Carlo Bo
vincenzo.fano@uniurb.it

1. Introduzione

Sono stati soprattutto i libri di Brian Skyrms¹, professore all'Università di Irvine, che hanno attirato l'attenzione dei filosofi sull'importanza per la filosofia politica della teoria dei giochi. Ancor più significativa è stata, per il grande pubblico, la divulgazione di molti importanti risultati promossa da Matt Ridley, i cui libri ora sono tradotti anche in italiano².

In questo breve saggio ci limiteremo a esaminare uno degli aspetti più controversi e affascinanti di questa disciplina, cioè il cosiddetto “dilemma del prigioniero”. Con ogni probabilità gli ideatori del problema sono stati Merrill Flood e Melvin Dresher, nell'ambito di ricerche sulle strategie militari della Guerra fredda (anni Cinquanta). Poco dopo il matematico Albert Tucker propose la storiella e il nome che sono poi diventati caratteristici. Due banditi vengono imprigionati con l'accusa di avere commesso un grave reato, colti però sul fatto di un reato lieve. Il magistrato va da entrambi e fa loro la seguente proposta: caro Prigioniero, se denunci il tuo complice del reato più grave ti abbuono sia la pena per il reato lieve che quella per quello grave, però, se anche l'altro prigioniero ti denuncia, ti dò un po' meno del massimo, perché hai collaborato, ma devo condannarti anche per il reato grave, 8 anni; se, invece, non denunci il tuo collega, allora, se lui ti denuncia, ti prendi 10 anni, se invece anche lui tace, posso condannarvi entrambi solo per il reato minore, 1 anno. A questo punto a entrambi i prigionieri conviene defezionare, cioè denunciare il complice, ma così facendo entrambi ottengono un risultato decisamente peggiore che se entrambi cooperassero.

¹ Skyrms (2004); Id. (1996).

² Ridley (2013); Id. (2012).

Nelle prossime pagine illustreremo una serie di risultati e analisi di questo dilemma³. Presupporremo un po' di dimestichezza con la teoria delle decisioni e dei giochi⁴.

2. Formulazione del dilemma

Consideriamo la tavola:

| | | |
|----------------|----------------|----------------|
| | B _c | B _d |
| A _c | -1,-1 | -10,0 |
| A _d | 0,-10 | -8,-8 |

Dove A e B sono due giocatori, c e d sono due possibili strategie e in ogni casella prima compare il payoff di A e poi quello di B.

A e B si trovano in una situazione in cui potrebbero cooperare (c) o defezionare (d). Se entrambi cooperano ottengono di più che se entrambi defezionano. Tuttavia se uno coopera e l'altro defeziona, quello che defeziona ottiene il massimo e quello che coopera il minimo. La tavola può anche essere così rappresentata in generale:

| | | |
|----------------|----------------|----------------|
| | B _c | B _d |
| A _c | C, C | S, O |
| A _d | O, S | D, D |

O sta per "payoff dell'opportunist", D sta per "payoff di chi defeziona se l'altro defeziona", C sta per "payoff di chi coopera se l'altro coopera" e S sta per "payoff dello sprovveduto". Affinché il gioco sia un dilemma del prigioniero deve essere $S < D < C < O$.

È chiaro che sia per A che per B defezionare domina⁵ su cooperare, per cui l'equilibrio⁶ è A_d, B_d. Tuttavia è difficile negare la forte sensazione di rammarico dovuta al fatto che non si è raggiunto l'accordo A_c, B_c, che avrebbe

³ Ci basiamo in modo essenziale sulla voce di Kuhn (2007). Fondamentale è anche il trattato di Binmore (2007).

⁴ Un'ottima introduzione resta Resnik (1990).

⁵ Una strategia è dominante per A, se qualsiasi strategia B scelga, non ci sono alternative migliori per A e almeno una è peggiore.

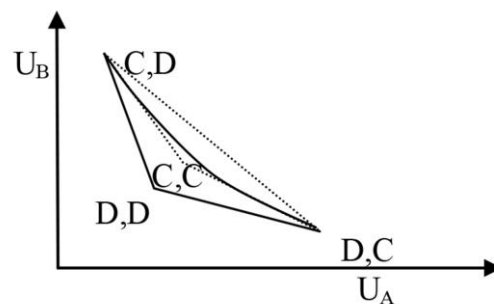
⁶ L'equilibrio detto anche "di nash" è definito come il punto in cui nessun giocatore può migliorare la propria utilità con un cambiamento di strategia unilaterale.

dato un' utilità maggiore a entrambi. In particolare la scelta A_c, B_c ha la caratteristica che nessun giocatore, cambiando strategia, può migliorare la sua utilità senza peggiorare quella dell'altro. In effetti entrambi i giocatori, restando il fatto che l'altro coopera, potrebbero migliorare defezionando; però questo comporterebbe un peggioramento per l'altro. Questa situazione viene chiamata "ottimo paretiano". La differenza fra equilibrio e ottimo paretiano è che nell'equilibrio i giocatori sarebbero liberi di peggiorare l'utilità dell'altro, mentre nell'ottimo no. Nel nostro caso l'equilibrio non è un ottimo paretiano, perché i giocatori potrebbero migliorare il loro payoff cooperando. È chiaro che se un equilibrio fosse un ottimo paretiano saremmo a cavallo. Il dilemma posto dalla tavola che stiamo esaminando consiste proprio nel fatto che l'equilibrio non è un ottimo paretiano⁷.

Notiamo anche che non tutti i dilemmi del prigioniero hanno un ottimo paretiano che coincide con le mosse c,c. Si consideri la tavola:

| | | |
|-------|-------|-------|
| | B_c | B_d |
| A_c | 5, 5 | 2, 10 |
| A_d | 10, 2 | 4, 4 |

Mettiamo i quattro punti su un grafico nel quale in ascissa abbiamo l'utilità di A e in ordinata l'utilità di B:



Si vede che C,C è sotto la linea che congiunge C,D con D,C. In questo caso esistono strategie miste⁸ di A e B che sono paretianamente migliori di C,C. Cioè alcune di quelle che si trovano sulla linea concava che unisce C,D

⁷ In questo sito è possibile trovare una divertente spettacolarizzazione del dilemma del prigioniero: <http://youtu.be/S0qjK3TWZE8>.

⁸ Una strategia mista rispetto all'insieme di n possibili strategie, s_1, \dots, s_n , consiste nello scegliere la propria strategia con una lotteria probabilistica del tipo c_1s_1, \dots, c_ns_n , con $c_1 + \dots + c_n = 1$ e $0 < c_i < 1$.

a D,C. Dunque qui l'ottimo paretiano non è più C,C ma una coppia di strategie miste. Si noti che queste strategie miste sono individuali e non *coordinate*. Se A e B potessero mettere in pratica strategie miste coordinate, allora sarebbero possibili strategie ancora migliori paretianamente, fino a tutte quelle che stanno sulla linea diritta fra C,D e D,C.

Se vogliamo essere sicuri che i due prigionieri non abbiano strategie miste paretianamente migliori di C,C, oltre alla condizione $S < D < C < O$, deve valere:

$$C \leq (S+O)/2,$$

assumendo che il dilemma sia simmetrico.

Ricordiamoci anche che il dilemma del prigioniero non è un gioco perfettamente competitivo, cioè un gioco a somma zero, cioè non vale che tutto ciò che perde B guadagna A e viceversa.

3. Considerazioni psicologiche

Proviamo a esaminare una possibile soluzione del dilemma. Il prigioniero A ritiene che il prigioniero B si comporterà come lui. A ha una forte propensione a collaborare, diciamo 0,9 contro 0,1. Inoltre la tavola potrebbe essere:

| | B _c | B _d |
|----------------|----------------|----------------|
| A _c | 10, 10 | 0, 15 |
| A _d | 15, 0 | 2, 2 |

A potrebbe ragionare così: è quasi sicuro che B cooperi, quindi a me conviene cooperare così ho un'utilità attesa di 9, cioè $0,9 \times 10$.

Da questa analisi sembra che ad A convenga cooperare. Siccome il gioco è simmetrico, lo stesso varrà per B. Sappiamo però che la teoria dei giochi ci dice che la strategia di defezionare è dominante. Sembra quindi che ci sia una contraddizione fra queste considerazioni psicologiche e la teoria dei giochi.

In realtà non c'è alcuna contraddizione, dato che proprio perché A sa che B si comporterà in un certo modo, cioè come lui si comporterebbe, gli conviene defezionare. Il calcolo dell'utilità attesa sarà il seguente:

$$U(c) = p(B_c) \times 10 + p(B_d) \times 0 = 0,9 \times 10 = 9$$

$$U(d) = p(B_c) \times 15 + p(B_d) \times 2 = 0,9 \times 15 + 0,1 \times 2 = 13,7$$

Da cui si vede che valutando l'utilità attesa ad A conviene defezionare.

4. Asincronia

Hume⁹ osserva che se due contadini, che coltivano poderi vicini, hanno grano che matura in tempi diversi e non riescono a raccoglierlo tutto da soli, potrebbero collaborare: prima assieme raccolgono il grano del podere che matura per primo e poi assieme raccolgono l'altro quando esso in seguito maturerà. Tuttavia agenti razionali non ragionano così. Il primo contadino, chiamiamolo A, riflette: se aiuto B (coopero=c) oggi, di certo a B domani conviene non aiutarmi (defeziona=d). Quindi fin da oggi mi conviene defezionare. A questo punto è chiaro che B defezionerà.

Da questa storiella si vede che il dilemma del prigioniero rimane tale anche se le due mosse dei giocatori sono asincrone. L'equilibrio resta A_d, B_d .

Sappiamo invece che nelle nostre campagne i contadini vicini di podere per lo più non si comportano a questa maniera, ma di fatto cooperano. Quale è la spiegazione di questa discrepanza fra il modello e la realtà? Il contadino ha fiducia nel comportamento del vicino? Sì e no: il contadino sa bene che il vicino fa quello che gli conviene fare, per cui non è solo la fiducia che lo porta a scegliere di cooperare. Il contadino sa invece, che il suo cooperare aumenta la probabilità che l'altro contadino cooperi, cioè c'è un vero e proprio effetto causale.

In questo caso si calcola così l'utilità attesa:

$$U(A_c) = p(A_c \text{ causi } B_c) \times C + p(A_c \text{ non causi } B_c) \times S$$

$$U(A_d) = D$$

È facile che se $p(A_c \text{ causi } B_c)$ è vicino a 1, essendo $C > D$, convenga collaborare.

Vedremo più avanti da dove deriva questo effetto causale.

⁹ Hume (1739-40).

5. Trasparenza

Si potrebbe provare ad affrontare il dilemma del prigioniero mediante una strategia condizionale. Una delle più note è quella proposta da Danielson¹⁰ che si chiama “cooperazione reciproca”.

“Se sei sicuro che l’altro coopera se tu cooperi e defeziona se tu defezioni, allora coopera.”

Ottima idea, ma la questione è: come fai a essere sicuro che l’altro adotta quella strategia di imitazione, che a volte viene indicata con la lettera I? Questo è possibile solo se il gioco del prigioniero viene giocato in situazioni di trasparenza; cosa che talvolta succede, soprattutto nei rapporti fra gli stati, dove spesso le scelte di politica estera devono passare attraverso l’approvazione degli organi di governo. In pratica, a volte uno dei giocatori è informato delle *intenzioni* dell’altro. Purtroppo però, da un punto di vista razionale, le intenzioni dell’altro giocatore non valgono molto, poiché l’altro potrebbe cambiare idea se gli conviene, come spesso di fatto succede¹¹. In altre parole, essere sicuro della strategia dell’altro è qualcosa di più di conoscere le sue intenzioni.

Bisogna stare attenti a non fare confusioni. Vissuti negativi dei giocatori, se defezionano, cambiano le utilità del gioco, magari trasformandolo da prigioniero in qualcosa d’altro; quindi, quando si presenta la tavola del prigioniero, essi devono essere già compresi. Un discorso analogo vale per eventuali sanzioni per chi defeziona, che potrebbero cambiare le utilità del gioco. Infine molto complicata è la questione dei giochi ripetuti. Fin qui stiamo esaminando solo il prigioniero nel caso di un’unica partita, per cui le eventuali conseguenze nelle partite successive non vengono prese in considerazione. Ne parleremo più avanti.

Dunque anche la trasparenza non cambia molto il dilemma del prigioniero.

¹⁰ Danielson (1992: cap. 5).

¹¹ È proprio quello che è accaduto fra Italia e India recentemente. Due soldati italiani, imprigionati in India con l’accusa di omicidio volontario, sono tornati in Italia per votare alle elezioni con l’impegno del nostro Governo di rimandarli in India, e invece, arrivati in Italia il Governo non li ha riconsegnati sulla base di un cavillo. Solo le minacce del Governo indiano al nostro ambasciatore hanno convinto il nostro Governo a rimandarli indietro.

6. La tragedia dei beni comuni

In un paese dove d'estate c'è scarsità d'acqua, occorrerebbe che una parte significativa dei residenti fosse moderata nei consumi, altrimenti le fonti si esaurirebbero. Esaminiamo questa situazione dal punto di vista della teoria dei giochi¹². Immaginiamo un gioco con molti giocatori: se n cooperano allora il bene si mantiene, se meno di n cooperano allora il bene si consuma. Ne segue che la tavola di scelta per l' *Aesimo* giocatore diventa:

| | Almeno n cooperano | Meno di n cooperano |
|-------|----------------------|-----------------------|
| A_c | $C+V$ | V |
| A_d | C | 0 |

V è l'utilità che si paga per cooperare, per cui V è un numero negativo. C è l'utilità che si ottiene se almeno n cooperano; deve essere $C+V > 0$. Si vede subito che per A la strategia dominante è defezionare.

Se il numero di persone coinvolte è sufficientemente grande la probabilità che A sia proprio l'ennesimo cooperatore è bassa. In questo ultimo caso in effetti defezionare non sarebbe più dominante.

Si noti anche che nella tragedia dei beni comuni la cooperazione universale non è necessaria. Infatti si ha la pareto-ottimalità anche solo per n cooperatori.

Pettit¹³ ha notato che nella tragedia dei beni comuni se è necessario che n sia uguale al numero dei giocatori allora la defezione assume un carattere diverso rispetto a quando n è decisamente minore. Nel primo caso si può parlare di un vero e proprio gioco sporco (*foul dealer*), mentre nel secondo è il solito opportunismo (*free-rider*) del prigioniero. Un esempio di gioco sporco potrebbe essere quello di un concorrente a un posto di lavoro che millanta nel suo curriculum. La probabilità di essere scoperto è minima, tuttavia dichiarando il falso il giocatore sporco si assicura il posto a danno di tutti gli altri.

7. La caccia al cervo

Bisogna stare attenti a non confondere il dilemma del prigioniero con il gioco della caccia al cervo¹⁴. Quest'ultimo ha la seguente tabella di utilità:

¹² Hardin (1968: 1243-1248).

¹³ Pettit (1986: 361-379).

¹⁴ Rousseau (1762, trad. it. 116).

| | | |
|-------|------------|------------|
| | B_c | B_d |
| A_c | C, C | S_A, O_B |
| A_d | O_A, S_B | D, D |

A e B sono i 2 giocatori; c sta per cooperare e d per defezionare; C è l'utilità se entrambi cooperano, O l'utilità di defezionare se l'altro coopera, S l'utilità dello sprovveduto se coopera e l'altro defeziona, D l'utilità se entrambi defezionano. Il gioco è simmetrico.

Nel prigioniero $O > C > D > S$. Nella caccia al cervo $C > O$, $O = D$ e $D > S$. Un esempio:

| | | |
|-------|-------|-------|
| | B_c | B_d |
| A_c | 4,4 | 0,3 |
| C_d | 3,0 | 3,3 |

Il nome deriva da un apologo di Rousseau nel *Discorso sull'origine della disuguaglianza*. Due cacciatori possono o collaborare e riuscire a catturare un cervo, oppure defezionare e cacciare ognuno per conto proprio una lepre. Si noti che questo non è un gioco a somma 0, in cui quello che guadagna A perde B e viceversa.

Nella caccia al cervo, pur restando $C > O$ e $D > S$, potrebbe essere anche $O > D$, se ad esempio ci sono poche lepri per cui è meglio se l'altro si affanna a cacciare il cervo da solo. Oppure $D > O$, cioè è meglio che entrambi caccino le lepri da soli, perché se qualcuno caccia il cervo da solo fa un gran baccano e le lepri scappano.

Da notare che nella caccia al cervo non esiste una strategia dominante. Infatti, se si sceglie d conviene solo se l'altro sceglie d, se, invece, si sceglie c, conviene solo se l'altro ha scelto c. Quindi in questo gioco ci sono due equilibri, cioè c,c e d,d e il primo di questi coincide con l'ottimo paretiano.

In questo gioco, per avere un criterio razionale di scelta, occorrerebbe essere in una situazione di rischio¹⁵, e quindi calcolarsi l'utilità attesa:

$$U(A_c) = p(B_c) \times C + p(B_d) \times S,$$

$$U(A_d) = p(B_c) \times O + p(B_d) \times D,$$

e massimizzarla.

¹⁵ Cioè conoscere le probabilità che l'altro defezioni e cooperi.

Si noti che se $U(A_c) < U(A_d)$ allora il gioco della caccia al cervo diventa nuovamente un dilemma, poiché sarebbe razionale non collaborare e quindi si perde l'ottimo paretiano.

8. Il conflitto delle volontà

Consideriamo la seguente tavola:

| | teatro | cinema |
|--------|--------|--------|
| Teatro | 2,1 | 0,0 |
| Cinema | -1,-1 | 1,2 |

A vuole andare a teatro e B al cinema, né A né B vogliono restare soli, per cui anche se vanno dove non vogliono, se c'è anche l'altro è un po' meglio. Il gioco ha due punti di equilibrio, che però non sono equi, perché avvantaggiano rispettivamente uno dei due giocatori. Chiediamoci se esista un equilibrio fra strategie miste più equo. Non esiste, poiché se A adotta la strategia mista $\frac{1}{2}(A_1)$, $\frac{1}{2}(A_2)$ la sua utilità attesa, se l'altro sceglie la stessa, sarà $\frac{1}{2}$, ma se l'altro risponde con la strategia pura B_2 , la sua utilità resterà $\frac{1}{2}$, mentre quella di B sarà 1.

La situazione del conflitto delle volontà è molto instabile, e può essere affrontato meglio come gioco *cooperativo*.

9. Giochi cooperativi

Consideriamo un tipo particolare di conflitto delle volontà:

| | teatro | cinema |
|--------|--------|--------|
| Teatro | 2,1 | 0,0 |
| Cinema | 0,0 | 1,2 |

Dove non c'è il desiderio di stare assieme. Immaginiamo che A e B possano decidere una strategia *coordinata*. Ad esempio potrebbero lanciare una moneta e se viene croce andare al cinema e se viene testa andare al teatro. Questo è possibile solo nei giochi *cooperativi*. Supponiamo che la tavola sia la seguente:

| | B ₁ | B ₂ |
|----------------|-------------------------------------|-------------------------------------|
| A ₁ | u _{A11} , u _{B11} | u _{A12} , u _{B12} |
| A ₂ | u _{A21} , u _{B21} | u _{A22} , u _{B22} |

Una strategia coordinata mista M avrebbe allora questo andamento:

$$M = x_1(A_1, B_1), x_2(A_2, B_2), x_3(A_1, B_2), x_4(A_2, B_1)$$

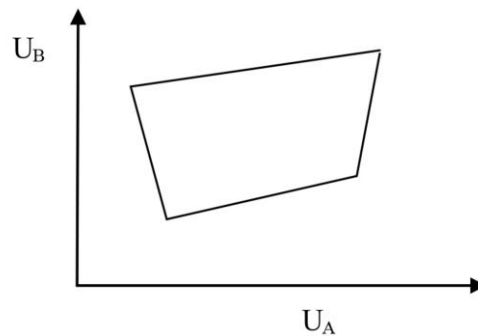
Dove la somma degli x_i deve dare 1 e ognuno di essi deve essere compreso fra 0 e 1. I numeri 1 e 2 indicano che A e B scelgono rispettivamente la strategia 1 o la strategia 2.

Le utilità attese di M saranno:

$$U_M(A) = x_1 \times u_{A11} + x_2 \times u_{A22} + x_3 \times u_{A12} + x_4 \times u_{A21}$$

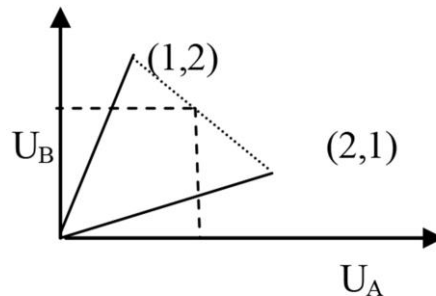
$$U_M(B) = x_1 \times u_{B11} + x_2 \times u_{B22} + x_3 \times u_{B12} + x_4 \times u_{B21}$$

Questa situazione si può rappresentare graficamente:



I quattro vertici sono le strategie pure, mentre tutti gli altri punti del quadrilatero rappresentano strategie cooperative miste.

Nel caso della tavola all'inizio di questo paragrafo il grafico sarà:



Si noti che tutti i punti sulla linea punteggiata sono pareto-ottimali, cioè ciascuno dei due giocatori, per fare meglio, deve necessariamente peggiorare la situazione dell'altro.

10. Il prigioniero iterato: il punto di vista matematico

Dalla nostra precedente analisi è risultato che la strategia d_i (defeziona incondizionatamente) è dominante nelle situazioni tipo prigioniero. È anche risultato che molte situazioni sociali significative hanno proprio la struttura del prigioniero. In generale solo in pochi casi il prigioniero diventa un gioco cooperativo, cioè nel quale è possibile una strategia coordinata. Strategia che a sua volta può essere cercata dai due giocatori mediante una trattativa¹⁶. Nella maggior parte delle nostre decisioni che riguardano il bene pubblico ci troviamo senza cooperazione confinati all'interno di un dilemma del prigioniero e quindi conviene defezionare. D'altra parte defezionare ci porta spesso al punto di equilibrio D,D che non è un ottimo paretiano.

Sarebbe dunque molto interessante dal punto di vista politico-morale, se si riuscisse a dimostrare che in certe situazioni di interazione sociale la strategia d_i non sia la migliore.

La prima cosa che viene in mente è quella di due giocatori che giocano il prigioniero più volte. Si potrebbe pensare che in un caso del genere la strategia d_i non sia più dominante. Purtroppo è facile dimostrare il contrario. Mettiamo che A e B si incontrino n volte. All'ennesima partita conviene defezionare, perché è l'ultima e quindi è come un prigioniero a un solo giro. A questo punto consideriamo la $n-1$ esima partita. È diventata l'ultima e quindi conviene defezionare, tanto qualsiasi cosa farà l'altro abbiamo già deciso di defezionare nell'ultima. Con questa procedura *induttiva a ritroso* si arriva a

¹⁶ Su questo argomento vale ancora la pena leggere lo splendido Schelling (1960).

dimostrare che l'equilibrio per entrambi i giocatori in un prigioniero ripetuto n volte con n finito è sempre d_i, d_i .

Tuttavia è importante notare che la scelta della strategia d_i nel prigioniero ripetuto si basa su un assunto molto forte, cioè che l'altro si comporti in modo ripetutamente razionale qualsiasi cosa succeda. Il che è molto irrealistico.

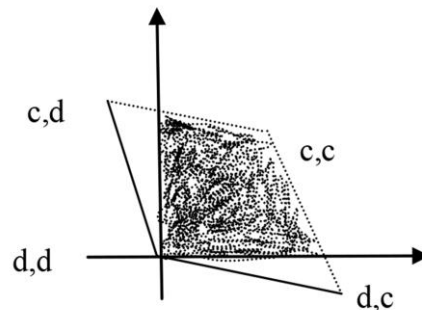
Inoltre consideriamo la possibilità che il prigioniero venga reiterato *indefinitamente*. Non un numero infinito di volte, che porterebbe a utilità infinite e quindi incomparabili fra loro, ma un numero *indefinito* di volte. Facciamo l'ipotesi, ad esempio, che ci sia sempre la probabilità p che ci si incontri ancora una volta. Per cui ci sarà una probabilità p^n che ci si incontri n volte. E per nessun n la probabilità sarà 0. In questo caso non si può applicare l'induzione a ritroso¹⁷. Inoltre si può dimostrare che esistono altri punti di equilibrio. Ad esempio supponiamo che A adotti la cosiddetta strategia GRIM, che consiste nel cooperare sempre; se però l'altro defeziona da quel momento in poi con lui si defeziona sempre. È facile vedere che se A e B entrambi giocano GRIM a nessuno dei due conviene cambiare, per cui GRIM, GRIM è un punto di equilibrio. Infatti se ad esempio A defeziona alla partita n senza che B abbia defezionato, porterà anche B a defezionare e perderà un'utilità pari a $(p^{n+1} + p^{n+2} + \dots) \times C$ rispetto a un guadagno $p^n \times O$. In molti casi sarà:

$$(p^{n+1} + p^{n+2} + \dots) \times C > p^n \times O$$

E allora GRIM, GRIM diventa un punto di equilibrio.

Tuttavia è molto più realistico considerare che semplicemente quando giochiamo un prigioniero *ignoriamo* se incontreremo ancora il nostro avversario, cioè in pratica p è sconosciuta. Possiamo assimilare questa situazione a un prigioniero giocato infinite volte. Dobbiamo tuttavia evitare che le utilità vadano all'infinito. Per fare questo si procede nella maniera seguente. Consideriamo il solito prigioniero:

¹⁷ Per quanto per n sufficientemente elevato la probabilità della serie è praticamente 0.



Immaginiamo che sia giocato infinite volte da un automa che ha una prima mossa determinata e poi una reazione stereotipata a come l'altro si è comportato nei giri precedenti. Ad esempio prima mossa c, poi se l'ultima mossa dell'altro è stata c allora c e se è stata d allora d. Indichiamo questo automa con c, $c \rightarrow c$, $d \rightarrow d$. I possibili automi che si basano solo su un giro precedente saranno:

c, $c \rightarrow c$, $d \rightarrow d$
 c, $c \rightarrow d$, $d \rightarrow d$
 c, $c \rightarrow c$, $d \rightarrow c$
 c, $c \rightarrow d$, $d \rightarrow c$
 d, $c \rightarrow c$, $d \rightarrow d$
 d, $c \rightarrow c$, $d \rightarrow c$
 d, $c \rightarrow d$, $d \rightarrow d$
 d, $c \rightarrow d$, $d \rightarrow c$

Qui abbiamo considerato gli automi finiti che si basano solo su un giro precedente. Ovviamente esistono infiniti automi finiti che possono prendere in considerazione un numero n grande a piacere di round precedenti.

Siccome gli automi sono finiti ogni gioco sarà di fatto un ripetersi infinito di uno stesso ciclo. Indichiamo come utilità attesa di A e B quella media del ciclo corrispondente alle strategie scelte, cioè agli automi designati dai due giocatori. In questo modo si evita che l'utilità vada all'infinito.

Riprendiamo ora in considerazione la figura precedente. Sappiamo che d,d è il punto di equilibrio per entrambi i giocatori. Chiamiamo l'area punteggiata "regione di cooperazione". Essa consiste in tutti quei possibili risultati in cui sia A che B hanno almeno il loro maximin. A questo punto possiamo formulare un caso particolare del fondamentale "teorema popolare" (*folk theorem*) della teoria dei giochi:

In un dilemma del prigioniero fra due automi finiti ripetuto un numero sufficiente di volte da ottenere la realizzazione di un ciclo, nella regione di cooperazione esistono infiniti punti di equilibrio se l'utilità viene calcolata facendo la media sul ciclo.

In pratica questo vuol dire che non solo la strategia d_i, d_i è in equilibrio, ma lo è anche la strategia GRIM, GRIM, come abbiamo visto e pure la strategia c_i, c_i , cioè la cooperazione incondizionata e altre infinite. Come ci saremmo aspettati, nel prigioniero reiterato indefinitamente – che è molto più realistico – la strategia d_i, d_i non è l'unico equilibrio.

Si può dimostrare un teorema ancora più forte che ha grande significato politico. Nella regione di cooperazione esistono anche infiniti punti che non solo sono di equilibrio, ma che sono anche di equilibrio per ogni *subgioco*, cioè per un gioco che si ottiene tagliando una parte lunga quanto si vuole dei primi round. Questo è molto importante, perché quei punti garantiscono che chiunque si allontani dalla strategia di equilibrio venga sempre punito. L'esempio più semplice di questo equilibrio *perfetto* è la coppia GRIM, GRIM. In pratica i giocatori di questo tipo di gioco si regolano reciprocamente evitando le deviazioni e facendo un lavoro di reciproco controllo¹⁸.

11. Il prigioniero iterato: il punto di vista empirico

Nel 1981 un fondamentale articolo di un economista (Axelrod) e un biologo teorico (Hamilton)¹⁹ ha dato l'avvio a un tipo di studio molto diverso del dilemma del prigioniero. I due autori hanno simulato al computer una serie di incontri fra strategie diverse di fronte al prigioniero. Il risultato empirico è stato che d_i non è la migliore strategia, cioè quella che ottiene più utilità. Questo dipende dal fatto che se i giocatori giocano strategie diverse da d_i è chiaro che d_i non è più la migliore. Nella realtà non tutti giocano d_i in un prigioniero ripetuto, per cui questo risultato ha un certo interesse anche a livello normativo.

La strategia che ha dato risultati migliori nella simulazione di Axelrod e Hamilton è stata la cosiddetta pan-per-focaccia (PPF), in inglese tit-for-tat (TFT). Essa consiste nel cooperare la prima volta che si incontra qualcuno e al giro successivo cooperare se l'altro al giro precedente ha cooperato e defezionare se ha defezionato.

¹⁸ È chiaro che queste considerazioni favoriscono una visione della convivenza civile non contrattualistica.

¹⁹ Axelrod & Hamilton (1981: 1390-96).

Notiamo che questo risultato non ha validità teorica, poiché, come abbiamo visto, in sede puramente a priori ci sono infiniti punti di equilibrio. Tuttavia esso ha un certo interesse, perché coglie un aspetto della realtà. Occorre però dire che, se scendiamo dal piano teorico a quello della simulazione, dovremmo essere il più possibile concreti. Allora dovremmo anche simulare il fatto che i reali realizzatori di una strategia possano qualche volta sbagliare. Ora, se mettiamo in campo un giocatore che pratica una PPF imperfetta, cioè con errori, i vantaggi di questa strategia diminuiscono significativamente, perché facilmente si possono creare lunghe catene di defezioni reciproche. In questi contesti una strategia più “misericordiosa” sembra essere maggiormente adeguata, cioè una pan-per-focaccia più generosa (PPFG), che non defeziona subito alla prima defezione dell’altro giocatore. In generale conviene adottare strategie che valutano l’intero percorso dei comportamenti precedenti dell’altro, piuttosto che solo quello dell’ultimo incontro.

Particolarmente interessante sembra essere la cosiddetta “strategia pavloviana”. Essa funziona così. Si consideri un qualche numero naturale n . Si attribuisca all’altro giocatore una certa probabilità p che cooperi²⁰, dopo di che si prevede che al nuovo round egli coopererà con probabilità²¹ $p+1/n$ se al turno precedente ha ricevuto l’utilità C (della cooperazione); $p-1/n$ se²² ha ricevuto D (l’utilità della defezione); $p+2/n$ se²³ ha ricevuto O (l’utilità dell’opportunist); $p-2/n$ se²⁴ ha ricevuto S (l’utilità dello sprovveduto). Il vantaggio di questa strategia chiamata P_n è che essa tiene conto implicitamente dell’intera storia precedente degli incontri utilizzando nel calcolo però solo l’ultimo comportamento. La strategia P_n inoltre sembra simulare bene ciò che capita nel mondo naturale. Si può stabilire a priori che P_n con $n>1$ contro la strategia “a caso” fa meglio di PPF. Si tenga anche conto che più n è grande più occorre tempo affinché la strategia pavloviana aggiusti il tiro.

In generale possiamo comunque dire che nella simulazione del prigioniero reiterato nessuna strategia è la migliore in ogni ambiente²⁵.

Negli studi più recenti sul prigioniero si esaminano anche i casi in cui i giocatori possono cambiare la loro strategia (evoluzione). In queste situazioni

²⁰ A lungo andare questo valore iniziale diventerà irrilevante.

²¹ È chiaro che se $p+1/n>1$ lo si pone =1.

²² Ovviamente se $p-1/n<0$, allora lo si pone =0.

²³ È chiaro che se $p+2/n>1$ lo si pone =1.

²⁴ Ovviamente se $p-2/n<0$, allora lo si pone =0.

²⁵ In questo sito <http://www.iterated-prisoners-dilemma.net/> si trova un divertente gioco che consente di simulare diverse strategie giocando il prigioniero reiterato.

evolutive si cerca di capire sia con calcoli a priori sia con simulazioni a posteriori quale siano le strategie più stabili a seconda dell'ambiente in cui vengono immesse.

12. Un recentissimo risultato

In un articolo²⁶ pubblicato da poco il fisico Dyson e il teorico dei giochi Press hanno dato un nuovo importante contributo matematico all'indagine del prigioniero iterato. Essi hanno dimostrato che una strategia che prende in considerazione più di un round precedente in generale non fa meglio di una che prende in considerazione solo l'ultimo giro. Hanno poi provato che esiste una classe di strategie miste – che hanno chiamato ZD – fra le quali alcune hanno proprietà importanti. Ad esempio un giocatore A può costringere l'altro B a una determinata utilità media alla lunga, indipendentemente dalla strategia che B sceglie. Quest'ultimo, cambiando la sua strategia, può al massimo diminuire l'utilità media di A. Nel caso in cui sia A che B conoscono le strategie di tipo ZD essi possono seguire strategie di equilibrio che hanno un'utilità maggiore della PPF e che vincolano l'altro a una certa utilità determinata. Infine Stewart e Plotkin²⁷ hanno provato a lanciare alcune strategie ZD nella simulazione di Axelrod e Hamilton, dimostrando che esse possono fare anche meglio di PPF.

Questo risultato mostra che il dilemma del prigioniero è un problema che va ulteriormente sviscerato, sia dal punto di vista teorico che empirico.

²⁶ <http://www.pnas.org/content/early/2012/05/16/1206569109.abstract>.

²⁷ <http://nr.com/whp/StewartPlotkinExtortion2012.pdf>.

Bibliografia

- Axelrod, R., Hamilton, W. D. (1981), "The evolution of cooperation", *Science*, 211, pp. 1390-96.
- Binmore, K. (2007), *Playing for real*, Oxford, Oxford University Press.
- Danielson, P. (1992), *Artificial Morality: Virtual Robots for Virtual Games*, London, Routledge, cap. 5.
- Hardin, G. (1968), "The Tragedy of the Commons", *Science*, 162, (13 December 1968), pp. 1243-1248).
- Hume, D. (1739-40), (*Trattato sulla natura umana*, trad. it. di P. Guglielmoni (a cura di), Roma-Bari, Laterza, 1982, p. 552).
- Kuhn, S. (2007), "Prisoner's Dilemma", *Stanford Encyclopedia of Philosophy*. On-line:
<http://plato.stanford.edu/entries/prisoner-dilemma/>
- Pettit, P. (1968), "Free Riding and Foul Dealing," *Journal of Philosophy*, 83, pp. 361-379.
- Resnik, M. D. (1990), *Scelte. Introduzione alla teoria delle decisioni*, Padova, Muzzio.
- Ridley, M. (2012), *Le origini della virtù. Gli istinti umani e l'evoluzione della cooperazione*, Milano IBL libri.
- Ridley, M. (2013), *Un ottimista razionale. Come evolve la prosperità*, Codice, Pomba.
- Rousseau, J. J. (1762), (*Discorsi e contratto sociale*, trad. it. di R. Mondolfo (a cura di), Bologna, Cappelli, 1977, p. 116).
- Schelling, T. C. (1960), (*Strategie del conflitto*, trad. it. di , Milano, Bruno Mondadori, 2006.
- Skyrms, B. (1996), *Evolution of the Social Contract*, Cambridge, Cambridge University Press.

Skyrms, B. (2004), *The Stag Hunt and the Evolution of Social Structure*, Cambridge, Cambridge University Press.

Sitografia

<http://www.iterated-prisoners-dilemma.net/>

<http://www.pnas.org/content/early/2012/05/16/1206569109.abstract>

<http://nr.com/whp/StewartPlotkinExtortion2012.pdf>

Responsabilità e ragionamento controfattuale nel diritto¹

Daniele Santoro
Centro Nazionale delle Ricerche, CNR-IRPPS, Roma
daniele.santoro@irpps.cnr.it

1. Introduzione: i controfattuali nel diritto

La nozione di causa svolge un ruolo fondamentale nell'ambito del diritto in quanto costituisce una condizione necessaria per l'attribuzione di responsabilità. Una persona infatti non può essere ritenuta responsabile per azioni che non ha causato e che non era in suo potere evitare. L'analisi più comune della causalità giuridica è quella condizionalistica. Questa analisi si basa sul ragio-

¹ L'idea originaria di questo saggio fu presentata ad un seminario sul realismo organizzato dal gruppo APhEx nel maggio 2007 presso l'Università La Sapienza di Roma. Ringrazio la Hoover Chaire d'etique et economie sociale dell'Université Catholique de Louvain per avermi permesso di completare questo lavoro durante un soggiorno di ricerca nel semestre invernale 2015-2016. Da allora ho avuto modo di discutere queste idee in diverse occasioni. In particolare, ringrazio Jean Boutier e Allain Trannoy per l'invito a tenere un seminario presso l'ÉHÉSS di Marsiglia nel giugno del 2012; Achille Varzi e John Collins per aver commentato una versione precedente di questo saggio durante un soggiorno a New York nel semestre invernale 2010-2011; Frédéric Nef e Otto Pfersmann per l'invito a tenere una lezione su questi argomenti nell'ambito di un seminario di dottorato presso l'ÉHÉSS di Parigi nel maggio 2010. Ringrazio infine Gabriele Calamo per una attenta lettura della versione finale di questo saggio. Ho avuto modo di presentare versioni precedenti di questo lavoro a diverse conferenze, tra cui quelle della Società Europea di Filosofia Analitica dell'agosto 2008 a Cracovia, della Società Francofona di Filosofia Analitica del settembre 2009 a Ginevra, e alla Inland Northwest Philosophy Conference (INPC XII) presso l'Università dell'Idaho nel marzo 2012. Alcune delle tesi difese nella prima parte di questo lavoro sono apparse in un saggio dal titolo "Legal Responsibility: A Pragmatic Perspective", in *Pragmatism, Law, and Language*, a cura di Hubbs e Lind (2013).

namento controfattuale, il quale stabilisce il tipo di nesso causale che intercorre tra azioni ed eventi². Ad una attenta analisi tuttavia, i controfattuali non sono in grado di spiegare una vasta gamma di conflitti fra le attribuzioni causali e le attribuzioni di colpevolezza³. Scopo di questo saggio è di analizzare alcuni esempi di questo tipo allo scopo di mostrare i limiti del ragionamento controfattuale. Partirò dall'analisi controfattuale delle condizioni necessarie e sufficienti, discutendo alcune obiezioni standard a questo tipo di criteri. Lo scopo di questa discussione è di proporre un modello più inclusivo che sia in grado di risolvere alcuni classici *puzzle* della responsabilità. Limiterò la mia analisi al ragionamento coinvolto nei giudizi di responsabilità penale e mostrerò che questo modello più inclusivo – le cosiddette condizioni NESS – risolvono alcuni casi difficili delle teorie condizionalistiche. Mostrerò tuttavia che il criterio NESS non è sufficiente, di per sé, a risolvere i problemi posti dagli effetti della mera sorte. Ciò tuttavia non significa che l'analisi controfattuale non abbia un ruolo nel ragionamento legale. Al contrario, il ruolo proprio dei controfattuali nel diritto consiste nel guidare il ragionamento nei giudizi di colpevolezza, in particolare nello stabilire ciò che l'agente avrebbe dovuto prendere in considerazione al momento dell'azione, e non esclusivamente nell'accertare le condizioni necessarie e sufficienti degli eventi.

2. Causalità e colpevolezza nella responsabilità penale

I giudizi di responsabilità nell'ambito del diritto penale si basano su due condizioni: la volontarietà della condotta e il nesso casuale tra la condotta e

² Intenderò qui per ragionamento legale un ragionamento esplicito basato su regole inferenziali in cui o le premesse o le conclusioni esprimono una modalità deontica codificata nell'ordinamento. Di conseguenza per qualsiasi inferenza, è sempre decidibile se le sue premesse o conclusioni costituiscono una norma codificata nell'ordinamento, o possono essere derivate da una norma siffatta. Poiché si tratta di inferenze decidibili, le regole inferenziali e le norme di riferimento dell'ordinamento devono poter essere esplicitate. Un esempio di ragionamento legale esplicito è, ad esempio, un semplice sillogismo di questo tipo: (P) Chiunque cagiona la morte di un uomo è punito con la reclusione non inferiore ad anni ventuno (art. 575 CP); (p) Il 1 ° gennaio 1996, Tizio uccise intenzionalmente Caio, (c) Tizio è colpevole di omicidio. La deduzione della conclusione è data dalla sostituzione di un termine singolare al posto del quantificatore universale. Si veda Fletcher (1998: 8).

³ Per chiarire il differente ruolo concettuale degli elementi costitutivi della responsabilità, distinguerò tra 'responsi' o 'resoconti' causali da un lato e 'attribuzioni' di colpevolezza dall'altro. La risultante di questi due elementi sono i 'giudizi' di responsabilità. Nei casi in cui sottolineo la natura processuale del ragionamento legale, mi riferirò a questi giudizi come 'verdicti' e 'imputazioni' a seconda del contesto.

l'evento giuridicamente rilevante. Nell'ambito del diritto di tradizione civilistica, queste due condizioni sono ricomprese negli elementi soggettivi e nella fattispecie obiettiva del reato commissivo. Similmente, nell'ambito della tradizione di *common law*, le due condizioni vengono individuate attraverso le categorie della *mens rea* e dell'*actus reus*.

L'ipotesi da cui muoverò in questo saggio è che una teoria complessiva della responsabilità legale non sia una teoria *additiva* delle condizioni necessarie e sufficienti per i giudizi di responsabilità, ma piuttosto esige che queste due condizioni siano in reciproco equilibrio. In che modo? L'analisi del nesso causale deve fornire una descrizione, o un insieme di descrizioni possibili, coerenti con l'elemento soggettivo della colpevolezza. Il verdetto di responsabilità penale sarà dunque l'esito di un processo di equilibrio riflessivo tra colpevolezza e causa. L'equilibrio riflessivo è un metodo comune di giustificazione nell'ambito della teoria etica e della filosofia della scienza. Per i sostenitori dell'equilibrio riflessivo, giustificare una teoria significa rendere reciprocamente coerenti i nostri giudizi particolari con i principi della nostra teoria. Per raggiungere questo equilibrio, a volte occorrerà modificare i nostri giudizi allo scopo di raggiungere la coerenza; in altri casi, come ad esempio in presenza di giudizi particolarmente esemplari, saremo piuttosto disposti a modificare i nostri principi per adattarli ai giudizi particolari. Attraverso questo lavoro di bilanciamento, giungiamo ad un equilibrio in cui il contenuto dei giudizi particolari riflette il contenuto dei principi generali della teoria, e dalla teoria è possibile dedurre i nostri giudizi. Una procedura analoga è all'opera nei verdetti di responsabilità penale. In questo caso però, l'equilibrio non è tra principi e giudizi, ma tra responsi causali e attribuzioni di colpevolezza.

L'analisi causale fornisce un resoconto di come siano andate le cose formulando un ventaglio di ipotesi possibili su come le cose potrebbero essere andate. L'analisi dell'elemento di colpevolezza ricostruisce poi le intenzioni dell'agente o forma delle ipotesi di moventi possibili. Nel giudicare una persona responsabile di un reato, il giudice si comporterà – almeno idealmente – come uno scienziato il cui compito è stabilire se una certa teoria si accorda con i fatti. Il giudice parte da un resoconto possibile di come siano andate le cose e individua quello che meglio si accorda con le prove o gli indizi disponibili. A partire da questa ipotesi stabilisce quale movente potrebbe aver mosso l'imputato ad agire. Se il resoconto causale è corroborato da prove, il collegio sarà restio a prendere in considerazione resoconti alternativi e cercherà piuttosto il movente più plausibile che si accorda con i fatti. Ma non sempre i resoconti causali sono sufficientemente corroborati. In quel caso, la formulazione di ipotesi su moventi alternativi, eventualmente di altri soggetti,

può fornire una guida per la ricerca del giusto resoconto causale. Attraverso questo processo di reciproco aggiustamento, andando avanti e indietro tra responsi causali e giudizi di colpevolezza, è possibile raggiungere un equilibrio in cui i due elementi convergono su uno stesso verdetto. Tuttavia non è la sola coerenza che giustifica un verdetto di responsabilità: per quanto in equilibrio, un verdetto che incolpi un capro espiatorio rimarrà un cattivo verdetto. I limiti alla variabilità dei nostri giudizi non può essere pertanto dato esclusivamente dalla ricerca di un equilibrio tra colpevolezza e cause, ma richiede un criterio indipendente di valutazione fattuale. In quest'ottica, la legge giustamente prevede che una persona non possa essere ritenuta colpevole per un reato, "se l'evento dannoso o pericoloso, da cui dipende l'esistenza del reato, non è conseguenza della sua azione od omissione"⁴. Occorre dunque specificare che la concezione della responsabilità da cui muovo è una forma di equilibrio riflessivo tra colpevolezza e cause in cui la valutazione di queste ultime è logicamente prioritaria rispetto ai responsi di colpa. Le ipotesi di colpevolezza possono essere una guida per la scoperta delle cause, soltanto laddove l'analisi causale ammette che le cose siano andate, o siano quantomeno potute andare, come le nostre attribuzioni di colpevolezza ci inducono a ritenere⁵.

Chiarita l'ipotesi generale, possiamo ora analizzare le relazioni possibili tra responsi di causa e i giudizi di colpevolezza. In primo luogo occorre dire che il nesso causale deve essere connesso alla volontarietà della condotta in modo tale che sia sempre decidibile se l'evento causato sia una conseguenza intenzionale o almeno prevedibile della condotta. In altre parole, deve essere possibile accertare che tipo di atto (volontario o meno, e – se volontario –, di quale fattispecie) abbia causato l'evento dannoso. Possiamo rappresentare queste due condizioni in uno schema di questo tipo:

⁴ Articolo 40, Codice Penale Italiano, Comma 1.

⁵ Sulla nozione di equilibrio riflessivo, si vedano Rawls (1971, trad. it. 56-58); Goodman (1955, trad. it. 5-33).

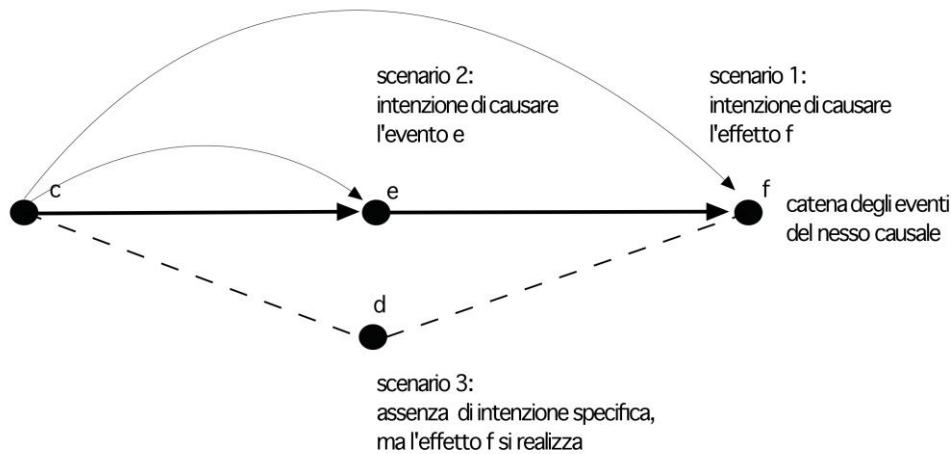


Fig. 2.1.: Intenzioni e nesso causale: tre scenari

Per spiegare lo schema rappresentato nella Figura 2.1, si consideri un esempio. Un soggetto *A* compie una azione *c* allo scopo di ottenere un effetto *e*. Poniamo che *A* colpisca *B* con un pugno. Poniamo che *B*, a seguito del pugno, cada dalle scale procurandosi una lesione. Chiamiamo questo effetto ulteriore *f*. Il nesso causale tra l'azione originaria *c* e l'effetto ulteriore *f* è diretto; non vi sono altri fattori, come l'intervento di un altro soggetto o di una condizione preesistente (ad esempio, una complicanza fisica) che contribuiscano alla morte di *B*.

Consideriamo ora l'elemento della colpevolezza. Abbiamo qui tre scenari alternativi possibili. Nel primo scenario, quello più semplice, *A* ha colpito *B* con l'effettivo intento di ucciderlo. L'intenzione dell'azione coincide in questo caso con l'effetto ultimo realizzatosi (*f*). Abbiamo qui che l'attribuzione di colpevolezza e il responso causale convergono sullo stesso verdetto di responsabilità per omicidio: l'agente ha infatti *volutamente* causare il danno provocato.

Nel secondo scenario, *A* colpisce *B* con l'intenzione di produrre l'effetto *e*, ad esempio di fargli del male, ma senza l'intenzione ulteriore di causarne la morte. In questo caso, possiamo affermare che l'intenzione *non* coincide con l'effetto determinatosi. L'evento delittuoso è piuttosto una conseguenza non prevista dell'atto e il verdetto di responsabilità volgerebbe piuttosto verso la fattispecie dell'omicidio preterintenzionale. Certo, chi prende a pugni una persona si mostra negligente o gravemente imprudente nel non tener conto del rischio eventuale di causare lesioni alla vittima, ma ciò non vale in tutti

casi. In un terzo scenario si può immaginare che, invece di prendere *B* a pugni, *A* avesse semplicemente mimato l'atto di dargli un pugno. Ora, sebbene mimare un pugno abbia comunque provocato la caduta, sarebbe assai più complicato ascrivere ad *A* un comportamento negligente. Qui l'intenzione associata alla condotta che ha innescato la catena causale non sarebbe quella di causare un danno alla vittima e – *a fortiori* – causarne la morte.

Questa breve analisi illumina un aspetto cruciale del problema della responsabilità penale, ossia che la fattispecie di reato ascritta all'imputato dipende dal tipo di descrizione che diamo degli eventi. Questa tesi, difesa da Elisabeth Anscombe⁶, è particolarmente saliente nei casi dei giudizi di responsabilità poiché l'imputazione di cui il soggetto è chiamato a rispondere è una sottoclasse specifica di tutte le possibili descrizioni di una sequenza di eventi, esattamente quella sottoclasse prevista come reato nel codice penale. Assumerò per ipotesi che questo giudizio sia sempre decidibile, ossia che vi sia sempre una risposta corretta alla domanda se un agente sia responsabile per aver compiuto un'azione. Per dare una risposta a questa domanda, abbiamo detto che i due elementi di una teoria della responsabilità devono essere in equilibrio, ossia che i giudizi di colpevolezza devono risultare coerenti con l'accertamento della storia causale dell'evento. Per stabilire se questa coerenza vi sia, occorre che il ragionamento attraverso cui viene accertata la responsabilità dell'agente segua un ordine logico: le attribuzioni di colpevolezza devono rispondere ai responsi causali, ma non viceversa. Se una attribuzione di colpevolezza risponde ad un responso causale, allora la catena di eventi è descrivibile come una condotta di cui il reo è responsabile. Una teoria generale della responsabilità penale è dunque una teoria dell'azione applicata ad un contesto specifico. Essa stabilisce quale condotta conti come azione volontaria, e quali conseguenze causali di quella condotta siano riconducibili alla responsabilità dell'agente. Nel prossimo paragrafo, mi concentrerò sugli aspetti causali della responsabilità penale.

3. L'analisi della responsabilità causale e i suoi limiti

Due eventi sono in una relazione di dipendenza causale se l'uno non si sarebbe verificato in assenza dell'altro. Il ragionamento sotteso alla dipendenza causale è di tipo controfattuale. Cos'è un controfattuale? È una asserzione condizionale il cui l'antecedente è una negazione di uno stato di cose (è dunque *contro i fatti*). I controfattuali stabiliscono una relazione tra come le cose

⁶ Anscombe (2000).

sarebbero andate se le condizioni fossero state diverse da quelle che effettivamente sono⁷. Ora, il criterio classico per identificare, nell'ambito del diritto, la presenza di un nesso causale tra una condotta e le sue conseguenze, è stabilire se la condotta messa in atto sia una condizione necessaria senza la quale l'evento dannoso non sarebbe occorso. Questo criterio è appunto di tipo controfattuale⁸. Chiamerò questa la tesi della causalità necessaria (CN):

(CN) Una azione *c* compiuta da un agente *A*, è una causa necessaria di un evento successivo *e*, se e solo se, se *c* non si fosse verificato, neanche *e* si sarebbe verificato⁹.

La tesi della causalità necessaria è una forma di dipendenza controfattuale tra eventi che si ottiene sostituendo una descrizione contenente un predicato d'azione al posto della descrizione dell'evento che figura come causa nella formulazione originaria. Sebbene vi siano dei limiti a questo tipo di formalizzazione¹⁰, la formulazione controfattuale appare in generale una strategia legittima per stabilire in che senso il contributo causale delle azioni sia rilevante per i giudizi di responsabilità. Un aspetto cruciale della causalità necessaria è che essa stabilisce un criterio assai stringente per tali giudizi. Non vi può essere infatti colpevolezza (*mens rea*) senza che l'agente abbia contribuito in modo necessario all'evento (o alle conseguenze di esso)¹¹.

Argomentazioni controfattuali di questo tipo sono essenziali nello stabilire le condizioni preliminari del dolo e della colpa, così come la dottrina delle scusanti e delle cause di giustificazione. Nell'ambito del diritto penale italiano, la tesi della causalità necessaria è solitamente formulata facendo appello al cosiddetto processo di eliminazione mentale: «un'azione è *condicio*

⁷ Lewis (1973). Si veda anche Menzies (2009).

⁸ Per la formulazione classica della dipendenza controfattuale, si veda Lewis (1973). Si noti che Lewis considera la dipendenza controfattuale come una relazione tra mondi possibili. Per un'analisi filosoficamente informata della metafisica dei controfattuali nel diritto, si veda Moore (2009: 371-470).

⁹ Per una discussione dettagliata delle condizioni necessarie del sistema di *common law*, si vedano Hart & Honoré (1985²:109-129); Moore (2009: 89-90, 396-397, 419-420). Per una ricognizione dei sistemi civilistici, si vedano Fiandaca & Musco, (2004: 200-223); Antolisei (1991: 211-214). Per una difesa della causalità nel diritto penale secondo il modello delle leggi di copertura scientifiche, si veda Stella (1975).

¹⁰ Sulla tesi che le azioni siano eventi la cui descrizione contiene un predicato d'azione, si vedano: Davidson (1963, trad. it. 41-62); *Id.* (1967, trad. it. 163-214).

¹¹ Per un riferimento al diritto italiano, si veda di nuovo l'articolo 40 del Codice Penale.

sine qua non per un evento, se non può essere mentalmente sottratta senza che l'evento cessi di esistere.^{12»}

Vi sono tuttavia alcune note obiezioni al criterio di causalità necessaria. Esse riguardano i casi in cui la catena causale primaria è interrotta da una catena causale secondaria (causalità *fail-safe*) oppure quando il verificarsi dell'effetto è anticipato da una catena secondaria (causalità *preventiva*). Ecco alcuni casi.

(1) I teorici della cospirazione avevano ragione a proposito della morte di John Fitzgerald Kennedy. Harry Lee Oswald non era solo sul tetto del Texas School Book Depository il 22 novembre 1963, quando decise di uccidere il Presidente degli Stati Uniti. Un secondo cecchino era presente sulla scena e aveva ricevuto l'ordine di uccidere Kennedy nel caso Oswald avesse fallito. Il secondo cecchino era più vicino all'auto presidenziale di quanto non lo fosse Oswald e, con tutta certezza, non avrebbe mancato il bersaglio qualora avesse sparato.

Nonostante lo scenario in (1) include la presenza di un secondo potenziale omicida, questo elemento non muta il ruolo causale di Oswald. È evidente – si dirà – che avendo Oswald sparato, resti lui responsabile per l'omicidio di Kennedy. Ma, per quanto l'intuizione sia corretta, non è facile stabilirlo se assumiamo (CN) come criterio di 'responsabilità causale'¹³. Ho premesso infatti che il cecchino avesse una elevata probabilità di uccidere Kennedy nel caso Oswald avesse fallito. Si potrebbe obiettare che, poiché Kennedy sarebbe stato comunque ucciso dal secondo cecchino, la condotta di Oswald non può essere considerata una causa necessaria nella catena di eventi che conduce alla morte di Kennedy e, dunque, Oswald non è causalmente responsabile nel senso richiesto dal criterio (CN). Basterebbe, in altri termini, stabilire la validità della seguente inferenza:

(1') Se Oswald non avesse sparato a Kennedy, allora il secondo cecchino lo avrebbe fatto;

Se il secondo cecchino avesse sparato, Kennedy sarebbe rimasto ucciso;

Dunque: se anche Oswald non avesse sparato, Kennedy sarebbe rimasto ucciso.

¹² Si veda Fiandaca & Musco (2004: 204).

¹³ Nella terminologia di Hart. Si veda Hart (1967: 211).

L'argomento è ovviamente fallace, ma – si potrebbe dire – solo perché il nostro atteggiamento nei confronti dei controfattuali discrimina ingiustificatamente tra eventi attuali ed eventi possibili. Se siamo disposti ad attribuire una alta probabilità al verificarsi di un evento, questa considerazione dovrebbe avere qualche importanza nella valutazione delle cause di quegli eventi. Secondo questo punto di vista, le condizioni di asseribilità di un asserto controfattuale dipendono dal grado di probabilità soggettiva che il conseguente sia vero, data la verità dell'antecedente¹⁴. Una alta probabilità che l'evento si verifichi rende il controfattuale asseribile e la conclusione giustificabile. Potrebbe essere questo dunque un argomento a difesa di Oswald? La risposta dipende da quanto siamo disposti a credere nella ineluttabilità della morte di Kennedy data l'evidenza a noi disponibile. Per probabilità prossime alla certezza, Oswald smette di essere causalmente responsabile secondo il criterio della causalità necessaria, poiché saremmo ragionevolmente certi che, se anche Oswald avesse mancato il bersaglio, si sarebbe innescata una catena causale alternativa che avrebbe prodotto lo stesso effetto. Tuttavia, sappiamo che questa conclusione contrasta con il corretto giudizio secondo cui è stato Oswald a causare materialmente la morte di Kennedy. Il criterio di causalità necessaria prevede pertanto che non vi sia una maggiore responsabilità causale per atti che rappresentano cause materiali rispetto alla responsabilità causale per eventi probabili, almeno per probabilità prossime alla certezza. Eppure questa differenza è cruciale per una teoria generale della responsabilità poiché in sua assenza dovremmo ammettere che possa essere ritenuto responsabile di un reato anche colui che probabilmente lo commetterà, sebbene non lo ha ancora compiuto.

L'interpretazione probabilistica dei condizionali controfattuali salva il criterio di causalità necessaria, ma al prezzo di negare il principio per cui non si può essere puniti per reati che non si sono cagionati. Tale principio riflette del resto una concezione profondamente radicata nella teoria dell'azione in generale: l'idea che non si possa essere autori di azioni di cui non si ha il controllo, e dunque che non si possa esserne chiamati a rispondere. L'analisi della logica sottostante la causalità necessaria spiega perché questo criterio non sia in grado di catturare i casi *fail-safe*. Il criterio di causalità necessaria si basa infatti su una errata interpretazione del principio di transitività causale. Tale principio stabilisce che

¹⁴ Più esattamente, se P è la funzione di probabilità bayesiana, e $(c \square \rightarrow e)$ la formula controfattuale per “ se fosse il caso che c , allora sarebbe il caso che e ”, allora $P(c \square \rightarrow e) = P(e|c)$ per tutti i c e gli e nel dominio di P , con $P(c) > 0$.

Si veda Stalnaker (1968: 98-112). Per una discussione di questa interpretazione della probabilità controfattuale, si veda anche Lewis (1976: 297-315).

(T) se ‘ c causa d ’ e ‘ d causa e ’, allora ‘ c causa e ’.

Combinando (T) con il criterio di causalità necessaria (CN), otteniamo:

(TCN) se ‘ c è una causa necessaria di d ’ e ‘ d è una causa necessaria di e ’, allora ‘ c è una causa necessaria di e ’,

da cui otteniamo a sua volta la formulazione controfattuale

(T \square) ‘se fosse stato il caso che c , allora sarebbe stato il caso che d ’, e
 ‘se fosse stato il caso che d , allora sarebbe stato il caso che e ’,

Dunque: ‘se fosse stato il caso che c , allora sarebbe stato il caso che e ’.

Facendo ora le dovute sostituzioni in (1), otteniamo:

(1 \square) Se (contrariamente a quanto avvenuto) Oswald non avesse sparato, allora lo avrebbe fatto il secondo cecchino;

se il secondo cecchino avesse sparato, allora Kennedy sarebbe stato ucciso;

Dunque: se anche Oswald non avesse sparato, Kennedy sarebbe stato ucciso.

Ho sostenuto che (1 \square) non è un buon argomento per ragioni *prima facie* epistemiche, ma in realtà l’errore è di carattere logico. Come ha sostenuto David Lewis¹⁵, mentre il principio (T) di transitività causale è valido, non è valida la sua controparte controfattuale (T \square). Lewis sostiene che l’inferenza (T \square) sia una ‘fallacia da transitività’¹⁶. Secondo Lewis infatti, anche qualora assumiamo che le premesse di (T \square) siano valide, non ne segue la conclusione poiché e potrebbe essersi verificato per qualche altra causa che c . In altri termini, c non può essere una causa necessaria di e poiché non vi è un legame controfattuale necessario tra c ed e . Tuttavia, anche se non necessaria, la dipendenza causale tra eventi attuali è sufficiente per la causazione.

Cosa ci dice questo a proposito del nostro caso di causazione *fail-safe*? Ne segue che non possiamo accettare la conclusione in (1 \square) poiché, anche laddove assumessimo che la probabilità di successo del secondo cecchino sia

¹⁵ Lewis (1973: 32-35). Si veda anche Id. (1976: 297-315) e Id. (1973: 556-67).

¹⁶ Lewis (1973: 32).

prossima alla certezza, non vi è un legame causale tra la condotta di Oswald e quella del secondo cecchino, e dunque – *a fortiori* – non può esservi un legame tra la condotta di Oswald e la morte di Kennedy. Il legame è piuttosto dato dal fatto che il secondo cecchino *intende* portare a termine la missione qualora Oswald fallisca, ma non è l'eventuale omissione di Oswald a causare il secondo sparo.

La diagnosi del caso *fail-safe* può essere dunque riassunta così: l'argomento in difesa di Oswald assume che la causalità necessaria sia il criterio corretto nei verdetti di responsabilità, e sfrutta l'idea che tale criterio implichi una forma di dipendenza causale (T□). In base a questo criterio, la condotta di Oswald non risulta necessaria a procurare la morte di Kennedy poiché anche la sua omissione avrebbe contribuito egualmente allo stesso effetto. Ma l'argomento appunto è fallace: vi è un legame causale tra l'azione di Oswald e la morte di Kennedy che non vi è nel caso della sua omissione.

Il criterio della causalità necessaria si rivela inadeguato anche nei casi di cosiddetta causazione preventiva (*pre-emptive*), ossia laddove una catena causale alternativa previene lo svolgimento della catena principale. A questo riguardo, si consideri una variante leggermente differente del caso (1), assai diffusa tra i teorici del complotto:

(2) Non è stato Oswald ad uccidere Kennedy. Mentre aveva già il bersaglio in vista, ha sentito un colpo di fucile partire dalla sua sinistra. Il rumore lo ha distratto facendogli mancare il bersaglio. Se il cecchino non lo avesse anticipato, Oswald avrebbe sparato.

Si noti che la differenza tra i due scenari non riguarda le intenzioni di Oswald, che sono identiche in ambedue i casi, ma la responsabilità causale di quest'ultimo. Più esattamente, nello scenario (2), delle due condizioni per la responsabilità penale, soltanto quella della colpevolezza (*mens rea*) è presente, mentre è assente il nesso causale. Il criterio di causalità necessaria esclude dunque che Oswald abbia causato la morte di Kennedy. Tuttavia, la nostra intuizione va in senso contrario: Oswald non ha solo contribuito nel ruolo di complice, ma ha tentato egli stesso di compiere l'omicidio. Ora, la differenza tra omicidio compiuto e tentato omicidio è spesso problematica (tornerò su questo punto nel §4) ed una teoria della responsabilità penale deve poter fornire una giustificazione per distinguere tra i due casi. Il punto cruciale è che il criterio della causalità necessaria non è sufficientemente sensibile a stabilire questa differenza, per ragioni simili a quelle già esposte a proposito del caso (1). Secondo tale criterio infatti, poiché Oswald non ha

contribuito causalmente alla morte di Kennedy, ne dovremmo concludere che egli non abbia contribuito affatto; un argomento evidentemente inaccettabile.

Anche in questo caso possiamo fornire una analisi della logica sottostante l'argomento. Riformulando il caso secondo lo schema di dipendenza causale, otteniamo che

Se il secondo cecchino non avesse sparato, Oswald l'avrebbe fatto (2□);

se Oswald avesse sparato, Kennedy sarebbe stato probabilmente ferito a morte;

Nel caso di *pre-emption* sappiamo che lo sparo del secondo cecchino ha causato la distrazione di Oswald nel momento cruciale. La premessa maggiore di (2□) è pertanto valida. La premessa minore invece non lo è per le stesse ragioni richiamate nell'analisi del caso precedente: non può essere giudicato responsabile di un delitto colui che non ha commesso il fatto. Fin qui l'argomento segue. Dove si intoppa è al passo successivo. Il criterio (CN) è in questo caso troppo stringente: escludendo l'imputazione di omicidio, esso esclude però anche plausibili ragioni di carattere epistemico che giustificano l'imputazione di tentato omicidio. Torna qui in campo l'interpretazione probabilistica dei controfattuali. La premessa minore di (2□) è infatti tanto più asseribile quanto più alta è la probabilità che Oswald avrebbe sparato se non fosse stato disturbato. È questa probabilità che giustifica l'imputazione di tentato omicidio. L'inferenza in (2□) è pertanto valida, se interpretata in senso probabilistico, allo scopo di stabilire reati legati a tentativi d'azione o atti di negligenza, grave imprudenza, od omissioni colpevoli. Il criterio di causalità necessaria non licenzia una conclusione di questo tipo.

Tirando le somme: se lo scopo dell'analisi filosofica è quello di catturare la prassi dell'argomentazione giuridica, dovremmo respingere la condizione necessaria come del tutto impraticabile, o formulare una teoria più ampia della causalità che non escluda la responsabilità causale per i casi *fail-safe* e di *pre-emption*.

Un criterio di causalità più inclusivo è quello della causalità sufficiente. Una formulazione standard di tale criterio è questa:

(CS) Per ogni evento e , vi è una causa sufficiente c .

La resa controfattuale di (CS) può essere così formulata:

Per ogni evento e , vi è una causa sufficiente tale che (CS□):
Ogni qual volta si verifica e , si verifica anche c ;
se e si verificasse in assenza di c , allora un altro evento d dovrà essersi verificato.

Il criterio di causalità sufficiente (CS) è più inclusivo di quello di causalità necessaria (CN) perché prevede esplicitamente la presenza di una possibile catena alternativa nella descrizione del contesto causale dell'evento. In questo modo, (CS) rende facilmente conto dei casi delle catene causali secondarie nei casi suddetti. Tuttavia, tale inclusività è ottenuta ad un certo prezzo. (CS) si presenta come infatti come una formulazione molto generale che risulta troppo poco informativa per i casi di *sovra-determinazione* causale. Si consideri la seguente variante del caso originario:

Né Oswald, né il secondo cecchino sanno della presenza dell'altro quel giorno a Dallas. Nascosti dietro le finestre a diversi piani di distanza l'uno dall'altro, entrambi colpiscono l'auto presidenziale, virtualmente nello stesso istante. Entrambi i colpi sono sufficienti a causare la morte di Kennedy (3).

Il caso (3) descrive una situazione in cui due azioni indipendenti sono cause singolarmente sufficienti per lo stesso evento¹⁷. È evidente che il criterio di causalità necessaria non è in grado di individuare la responsabilità coinvolta in queste circostanze: essendo entrambe le azioni sufficienti a causare l'effetto, nessuna delle due risulta necessaria, costringendoci così ad conclusione che appare aberrante: entrambi gli agenti potrebbero sostenere che il loro comportamento non sia stato causalmente rilevante. Occorre pertanto rafforzare il criterio di sufficienza causale. Una proposta è quella di combinare i criteri di necessità e sufficienza causale.

4. NESS: un modello controfattuale più inclusivo

Si consideri una variante del caso (3):

¹⁷ Cfr. Fiandaca & Musco (2004: 201), per una discussione di questo punto in merito al diritto penale italiano.

(4) Sia Oswald che il secondo cecchino colpiscono Kennedy, virtualmente nello stesso istante. Tuttavia, in questo caso, entrambi i colpi, presi singolarmente, non sono sufficienti ad uccidere Kennedy.

Il caso (4) è un esempio di causalità congiunta (*joint causation*, nella comune espressione inglese). Poiché nessuna delle due catene causali è sufficiente di per sé a produrre l'effetto, il criterio di causalità sufficiente non è in grado di distinguere il peso dei due distinti contributi causali. Il caso mostra piuttosto che un resoconto complessivo della responsabilità causale deve rendere conto non solo delle condizioni necessarie o sufficienti, ma delle condizioni necessarie e assieme sufficienti per un evento dannoso.

Una proposta teorica in questa direzione viene dall'ambito delle teorie regolariste della causalità, in particolare dal filosofo John Mackie. Mackie sostiene che non vi sono cause necessarie indipendentemente dal contesto in cui si verifica una particolare relazione causale. Una causa – secondo Mackie – è individuata da quelle che egli chiama condizioni INUS. Una condizione INUS è una condizione di per sé *Insufficiente* ma *Non-ridondante* nel generare un insieme che, sebbene *Non-necessario* (Unnecessary), è tuttavia *Sufficiente* per un effetto in un particolare contesto causale¹⁸. Il seguente esempio può chiarire l'idea alla base delle condizioni INUS:

(5a) Se sfrego un fiammifero ben asciutto, si accenderà;

(5b) Se sfrego un fiammifero ben asciutto, e la stanza è priva di ossigeno, il fiammifero non si accenderà;

(5c) Se non sfrego il fiammifero, ma lo avvicino ad un filo incandescente, e la stanza contiene ossigeno, il fiammifero si accenderà.¹⁹

Anche se sfregare un fiammifero può apparire una condizione necessaria e sufficiente perché si accenda (5a), possiamo alterare le condizioni dell'antecedente causale sottraendo ossigeno, in modo da alterare l'effetto (5b). In questo modo, possiamo notare che la frizione esercitata sul fiammifero si rivelerà non più sufficiente. Perché la frizione sia sufficiente per la combustione, occorrerà specificare la condizione aggiuntiva per cui l'ossigeno deve essere presente nella stanza. Ma si noti: lo sfregamento non è necessario di

¹⁸ Si veda Mackie (1980: 62-63).

¹⁹ Si tratta di una versione modificata di un esempio molto discusso, originariamente presentato da Goodman (1995: 5-33). Si veda anche la discussione del caso in Brandom (2000: 88).

per sé per la combustione. Rafforzando l'antecedente, il ruolo causale dell'azione di sfregamento muta a seconda dei casi: in alcuni, l'azione sarà sufficiente, in altri no. Nel caso (5c) la combustione avviene senza frizione, ma solo per contatto con un oggetto incandescente. Ne deduciamo che lo sfregamento del fiammifero non è una condizione necessaria. Lo è piuttosto la presenza di ossigeno.

Attraverso l'analisi di tipo INUS per i fenomeni naturali, possiamo specificare (almeno in principio) quale ruolo causale ogni elemento giochi all'interno di un dato insieme. Secondo alcuni²⁰, le condizioni INUS offrono uno schema di analisi causale a grana fine che può essere adattato all'ambito del diritto. È questa la teoria NESS della causalità giuridica. NESS, nell'acronimo inglese, stabilisce che una condotta (omissiva o commissiva) è un atto di reato quando costituisce un Elemento Necessario di un Set di condizioni Sufficienti a determinare l'evento dannoso. Dal punto di vista della responsabilità penale, possiamo dunque affermare che la causa di un reato deve figurare come causa necessaria per generare l'insieme di condizioni che, nell'insieme, sono sufficienti. Soltanto un resoconto congiunto di questo tipo può essere sufficientemente sensibile da individuare il ruolo dei diversi contributi nel determinare la causa *materiale* del delitto. Una formulazione plausibile del criterio NESS è la seguente:

Un azione *c*, è un elemento necessario di un insieme *S* di condizioni congiuntamente sufficienti per un evento *e* se e solo se, se *c* non si fosse verificato, allora l'insieme *S* non sarebbe sufficiente a causare *e* (NESS).

La domanda a questo punto è se possiamo ottenere, attraverso l'impiego di criteri NESS, lo stesso risultato per la condotta umana dell'analisi INUS per i fenomeni naturali. Ad una prima ricognizione, NESS sembra includere l'intera gamma di casi analizzati finora, in quanto permette di stabilire il ruolo causale di ciascun elemento di un insieme. In particolare, spiega la sufficienza causale come un aggregato di cause in un contesto particolare d'azione, e spiega la necessità causale come condizione necessaria non per l'effetto in quanto tale, ma per l'insieme delle cause sufficienti per quell'effetto. In questo modo, NESS non solo include tutto ciò che i criteri di causalità necessaria e causalità sufficiente sono in grado di spiegare, ma rende disponibile l'uso di entrambi criteri a seconda dei casi analizzati. Vi è inoltre un ulteriore van-

²⁰ Sul criterio NESS si vedano Wright (1991), Id. (2004), e la più recente rivisitazione della teoria in Id. (2011). Si vedano anche Fumerton & Kress (2001: 101-22); e, in italiano Stella (2004).

taggio: la principale differenza tra la formulazione NESS e il criterio di causalità necessaria è che NESS non individua cause modalmente robuste, ossia non ci dice quali cause siano necessarie in ogni mondo controfattualmente possibile, ma ci permette soltanto di stabilire quali elementi dell'insieme non possano essere sottratti senza rendere quell'insieme non più sufficiente per l'effetto²¹. Ciò che NESS permette di fornire è una analisi comparativa dei contributi causali di ciascun elemento di un insieme in condizioni *ceteris paribus*, ossia variando un elemento del set per volta e tenendo fermi gli altri, per poi stabilire in quali casi l'effetto si sarebbe verificato.

Ora, per poter contare come modello generale di responsabilità causale, NESS dovrebbe essere in grado di giustificare l'imputazione di responsabilità anche in casi complessi, e tuttavia comuni, che coinvolgono le conseguenze non intenzionali delle azioni. Si considerino, ad esempio, i casi di omicidio colposo, in cui l'elemento soggettivo della colpevolezza non dipende da un intento doloso. Sono questi i casi risultanti da una condotta negligente, come nel seguente esempio:

John prende a pugni Michael e lo lascia cadere sul ciglio della strada senza aiutarlo a rialzarsi; Michael muore dopo essere stato investito da un'auto che sopraggiunge (6).

La questione è se John sia responsabile della morte di Michael. Secondo il resoconto NESS, l'aver preso a pugni Michael è un elemento necessario di un insieme di cause concorrenti sufficienti a causarne la morte. L'intuizione di senso comune è infatti che, se John non avesse colpito Michael, quest'ultimo non sarebbe morto. Questa conclusione va tuttavia qualificata contro due possibili obiezioni. In primo luogo, il criterio NESS è formulato in termini controfattuali e pertanto, come abbiamo chiarito nel paragrafo precedente, non ammette la transitività causale. Non possiamo, in sostanza, affermare che, se John non avesse preso a pugni Michael, allora Michael non sarebbe morto. A questa obiezione si può rispondere notando che il problema della mancata transitività controfattuale non coinvolge il metodo di ragionamento impiegato in NESS. Secondo NESS infatti, la condotta di John è causalmente rilevante per la morte di Michael, ma non in quanto causa necessaria della catena causale 'pugno-caduta-incidente', ma piuttosto come causa necessaria a rendere

²¹ Si noti che il significato dell'espressione 'insieme sufficiente' in questo contesto si riferisce all'idea che l'insieme di condizioni può essere sufficiente per un effetto in una particolare circostanza, ma non in altre (ossia tale insieme è solo *contingentemente* sufficiente). Dunque, nelle formulazioni del tipo 'ci deve essere un insieme sufficiente di condizioni' l'uso di "deve" è esclusivamente *epistemico* e non esprime una forma di *necessità modale*.

l'insieme "pugno & caduta & incidente" sufficiente per la morte di Michael in quella particolare circostanza.

La seconda obiezione insiste sullo stesso punto, ma da un'angolatura differente. Si potrebbe sostenere che NESS non rispetti il requisito secondo cui il responso causale deve riflettere l'elemento soggettivo della colpevolezza: non vi è dubbio infatti che essendo John causalmente responsabile per l'evento originario (l'aver preso a pugni Michael), questo lo renda anche responsabile per l'evento dannoso (la morte di Michael), ma da ciò non segue che la colpevolezza per il primo evento possa essere estesa anche al secondo. In risposta a questa critica, occorre notare che John si è comportato in modo gravemente negligente, tralasciando con noncuranza il rischio di un possibile incidente²², e omettendo peraltro di fornire dovuta assistenza ad una persona in uno stato evidentemente confusionale. Sono questi requisiti normativi della colpevolezza che giustificano l'accusa di omicidio colposo. Il criterio NESS riguarda soltanto la valutazione delle cause, non dell'elemento normativo della condotta e pertanto non è sufficiente, di per sé, a fornire un responso per un verdetto incontestabile di responsabilità.

Abbiamo finora analizzato tre modelli di causalità nel diritto: il modello di necessità causale, il modello di causalità sufficiente, e il modello NESS, che integra i precedenti all'interno di una concezione più ampia della causalità. Lo scopo del modello NESS è di fornire una spiegazione della causalità umana senza presupporre né l'esistenza di relazioni nomiche, né di formulazioni legiformi²³. L'assunto che sottende questo tipo di analisi è che il ragionamento controfattuale serve a stabilire una correlazione tra eventi in condizioni limitate, segnatamente in presenza di condizioni *ceteris paribus*. Vi sono tuttavia casi in cui il criterio NESS non rende conto delle nostre attribuzioni di colpevolezza. A titolo d'esempio, si consideri il seguente esperimento mentale, discusso da Joel Feinberg in un articolo di qualche anno fa.

²² Su questo punto, si veda l'incisiva analisi di Duff (2007: 92): «Che in alcuni casi io possa essere ritenuto *noncurante* (*reckless*) nell'assumermi o nel procurare un rischio di cui non ero a conoscenza [non esclude che] possa essere ritenuto negligente (*negligent*) nell'assumermi o nel procurare quel rischio, e pertanto essere ritenuto responsabile per aver procurato quel rischio e per il danno eventuale. Alcuni teorici hanno notoriamente sostenuto che la responsabilità penale (*criminal liability*) per negligenza, per quanto utile, non sia corretta, dal momento che la negligenza non è una specie di colpa (*fault*). Una ragione per questo tipo di argomenti potrebbe essere l'idea che non posso avere controllo su ciò di cui non sono a conoscenza, nel qual caso non posso legittimamente essere ritenuto responsabile per quella cosa.» (T.d.A.). Per una recente discussione sul ruolo di responsabilità nei casi di negligenza, si veda anche Raz (2009b).

²³ Peraltro, il modello NESS non è un modello per scopi predittivi. Ciò che l'analisi NESS fornisce è soltanto una valutazione *a posteriori* del ruolo causale dei diversi eventi coinvolti.

La scena presenta due soggetti. Il soggetto A è ancora una volta un cecchino, e B è la sua vittima. Ci sono due casi:

Nel primo caso il cecchino A ha l'obiettivo di uccidere B e, a tale scopo, prende B di mira con il suo fucile, preme il grilletto e tutto va come previsto. Il proiettile colpisce B in un punto vitale e B come risultato muore. Nel secondo caso, la situazione è identica, eccetto per un particolare del tutto indipendente dalla volontà e dalla condotta di A, che fa sì che B sfugga alla morte. B potrebbe ad esempio indossare un giubbotto antiproiettile, oppure mentre A sta per premere il grilletto una zanzara gli si posa sul naso, o ancora il bersaglio si sposta casualmente all'ultimo secondo. In tutte le versioni di questa storia, A perde la mira e fallisce nel suo tentativo, sfuggendo pertanto all'incriminazione per omicidio. Stante i fatti, A sarà colpevole solo di tentato omicidio e soggetto ad una pena assai inferiore rispetto a quella prevista per un omicidio consumato²⁴. (T.d.A.)

Secondo Feinberg, questa differenza nella gravità della pena non è giustificata, poiché la decisione circa la punizione che merita una persona non dovrebbe essere basata su motivi puramente circostanziali, in quanto ciò esporrebbe quella persona agli effetti della mera sorte²⁵. Sebbene la prima preoccupazione di Feinberg sia di sostenere una riforma del diritto penale, le sue conclusioni sono rilevanti anche per valutare la capacità dell'analisi controfattuale di rendere conto del ruolo della sorte nelle attribuzioni di responsabilità penale.

Si consideri, ad esempio, una possibile descrizione controfattuale del caso di omicidio compiuto:

(Oc) Se A non avesse sparato, B non sarebbe morto.

Si consideri ora il caso di tentato omicidio: A colpisce B, ma B sfugge alla morte grazie ad un fattore *casuale c*. La resa controfattuale di questa circostanza può essere così espressa:

(Ot) Avendo A sparato, se *c* non fosse intervenuto, B sarebbe morto.

Si noti che il controfattuale in (Ot) contiene una differenza significativa rispetto a quello in (Oc). Il fattore casuale *c* che esprime il ruolo della mera sorte è qui introdotto nell'antecedente del condizionale, sulla falsariga di quanto avevamo fatto nell'esempio (5) per spiegare il metodo NESS di ricerca

²⁴ Il caso è tratto da Feinberg (2003: 78). Per un'analisi del ruolo della sorte nella definizione della nozione di *agency*, si veda Raz (2009b).

²⁵ Feinberg (2003: 78).

delle cause. Secondo questa analisi, ad ogni variazione di un elemento dell'insieme (ad esempio introducendo un nuovo elemento, o sottraendone uno), possiamo stabilire se l'insieme rimane sufficiente nel produrre l'evento. Nello stesso modo procediamo qui: introducendo il fattore casuale nell'antecedente del controfattuale, lo sparo si rivela non più sufficiente a produrre l'effetto. Piuttosto lo sparo si rivela una condizione esclusivamente necessaria. Se allora ciò che l'analisi NESS mostra è che i due casi hanno in comune un elemento di necessità, cosa giustifica una pena differente? Certamente non può essere una differenza nella volontà dolosa dell'agente, poiché abbiamo assunto che l'intenzione sia la medesima nei due casi. Feinberg sostiene che un'analisi a grana fine della causalità dovrebbe fornire una 'spiegazione ragionevole' per tale discrepanza. Tuttavia, una tale spiegazione non sembra esservi, né il criterio NESS è in grado di offrirne una.

I casi in stile Feinberg mostrano che i criteri di responsabilità causale non sono sempre in grado di distinguere tra atti le cui conseguenze divergono per il solo intervento della mera sorte. Un tale *vulnus* teorico, sebbene possa apparire remoto (in fondo – si potrebbe argomentare – i casi Feinberg sono alquanto rari), rivelano un dilemma più profondo: o ci fidiamo delle attribuzioni di colpevolezza e ammettiamo che le nozioni di causa impiegate nel diritto possano variare con i contesti decisionali (i tribunali), oppure dobbiamo essere disposti a rivedere i nostri giudizi. La prima soluzione, sebbene ammissibile, si scontra però con la tendenza, nella teoria del diritto (particolarmente quella di orientamento civilistico) ad evitare dei modelli alternativi di causalità. La seconda esige una revisione a volte profonda delle nostre intuizioni circa le nozioni di colpevolezza. In ogni caso, quando vi è una divergenza tra attribuzioni di colpevolezza e responsi causali, vi è anche un problema per il diritto. Ho mostrato che una lettura debole dell'analisi controfattuale offre responsi affidabili, sebbene non sufficienti, che hanno un importante ruolo euristico nel distinguere tra i diversi contributi causali degli agenti. Pertanto, una volta in possesso di un criterio a grana fine come NESS, occorrerà rivedere le attribuzioni di colpevolezza, aggiustandole rispetto ai responsi causali secondo il metodo per l'equilibrio riflessivo suggerito nel primo paragrafo. L'uso del ragionamento controfattuale può contribuire anche qui a chiarire anche aspetti importanti della nostra concezione della colpevolezza.

5. I controfattuali e la colpa

Vorrei difendere brevemente l'ipotesi che, nella prassi del diritto, così come nell'ambito più ampio del ragionamento pratico, assumersi e attribuire responsabilità per le azioni coinvolge una dimensione normativa, ossia la capacità di essere sensibili alle ragioni²⁶. In particolare, le imputazioni di responsabilità in sede penale non esprimono esclusivamente giudizi di natura causale, ma riguardano anche le ragioni che l'agente può addurre nel processo per giustificare o scusare le proprie azioni, o ancora le ragioni che l'agente avrebbe dovuto prendere in considerazione, ma che non ha considerato, al momento dell'azione. Le distinzioni paradigmatiche tra i diversi tipi e gradi di colpevolezza, così come la dottrina delle scusanti e delle cause di giustificazione possono essere così spiegate nei termini delle ragioni che l'agente non aveva e avrebbe dovuto considerare nel corso dell'azione.

In questo senso, possiamo affermare che il *locus* della formazione della responsabilità non sia la mente dell'agente, ma – per dirla con Wilfrid Sellars²⁷ – lo spazio delle ragioni in cui si articola la prassi del dare giustificazioni per la propria condotta e chiedere giustificazioni per la condotta altrui. L'approssimazione più fedele di tale prassi nell'ambito del diritto è ovviamente il processo penale. È all'interno di questa prassi governata da regole che la responsabilità dell'agente viene *costruita*, più che *scoperta*. La dimensione normativa della responsabilità non può essere ridotta ai motivi per cui gli agenti hanno agito, perché essere sensibili alle ragioni richiede più di una semplice attribuzione di motivi. Perché un agente agisca responsabilmente, egli deve poter esercitare un controllo sulle proprie azioni. Tale capacità richiede non solo dei moventi per agire, ma anche l'essere in grado di rispondere, di poter esercitare un controllo su di essi ed effettuare una scelta deliberata circa il perseguirli o meno²⁸. L'idea secondo cui la colpevolezza richieda il mero controllo dell'azione va dunque specificata: un atto di cui un

²⁶ Vi è un'ampia letteratura sulla concezione della responsabilità come capacità di essere sensibili alle ragioni (o la cosiddetta *responsiveness to reason*). Alcune importanti proposte teoriche in questa direzione sono: Fischer & Ravizza (1999); Scanlon (1998a) e Id. (1988b: 352-371); Watson (2001: 374-394). Anche Brandom delinea una teoria della responsabilità doxastica in termini di rispondenza alle ragioni sviluppando l'originaria concezione di Wilfrid Sellars. Si vedano, in particolare, in "Norms, Selves, and Concepts" e "Autonomy, Community, and Freedom" in Brandom (2006).

²⁷ L'espressione è di Robert Brandom, che attribuisce a Sellars la paternità del concetto. Si veda Brandom (1994: capitolo 3); Sellars (1956: § 36).

²⁸ Sullo stretto legame tra rispondenza alle ragioni e controllo, si veda di nuovo Fischer & Ravizza (1999: capitolo 2).

soggetto è chiamato a rispondere in un'aula di tribunale deve essere necessariamente volontario, ma la volontarietà non è essa stessa sufficiente ad stabilire la responsabilità per quell'atto. È esattamente questo il punto dei casi in stile Feinberg che coinvolgono il ruolo della sorte. Nel nostro esempio, l'elemento volontario della colpevolezza è identico in entrambi gli scenari, ma il fattore casuale che interviene è esterno al controllo dell'agente. L'idea che la colpevolezza sia un insieme di stati psicologici e disposizioni appare pertanto troppo debole per giustificare le imputazioni di responsabilità.

Un aspetto cruciale di questa concezione è che la prassi del dare e chiedere ragioni esige l'esercizio di un ragionamento controfattuale. In questo contesto però, il ragionamento controfattuale non si focalizza sulle proprietà modali degli eventi (se essi siano necessari, sufficienti, o meramente contingenti)²⁹, ma sulla giustificazione delle proprie azioni. I casi di negligenza e imprudenza sono esemplari di questo esercizio di ragionamento. Nel decidere se un agente è responsabile per un evento dannoso, non possiamo soltanto chiederci se la sua azione sia causalmente connessa all'evento, ma se egli fosse in grado di comprendere che, evitando di agire così come ha agito, avrebbe potuto così evitare di causare l'evento dannoso che gli viene imputato. Il problema non riguarda quale elemento causale avrebbe fatto la differenza se le cose fossero andate diversamente, né se l'autore aveva un motivo per agire come ha agito, ma piuttosto se egli avesse potuto formarsi una ragione per evitare la propria condotta delittuosa³⁰.

Secondo il punto di vista che qui difendo, perché qualcuno sia responsabile di un'azione, non basta che egli abbia un motivo per agire. Un movente può essere un elemento necessario, ma non sufficiente per l'attribuzione di responsabilità. Oltre all'elemento della volontarietà, occorre provare che (i) l'agente aveva una comprensione di ciò che è accaduto come qualcosa che era in suo potere evitare e che (ii) era in grado di concepire una ragione per non compiere quell'atto. È la capacità di pervenire a queste ragioni e di rispondere della loro assenza che contraddistingue l'attribuzione di responsabilità nel ragionamento pratico, così come nell'ambito della legge. I requisiti

²⁹ La mera sorte è un evento contingente. Con il termine 'contingente' non intendo qui la mera possibilità che un evento si verifichi, ma piuttosto il caso in cui un evento è attuale (ossia, è un evento che si verifica in questo mondo), e dunque possibile (ciò che si verifica in questo mondo deve essere anche possibile), ma tuttavia non è necessario (non tutto ciò che avviene, avviene infatti in tutti i mondi possibili). Si veda su questo punto l'introduzione a Gendler & Hawthorne (2002: 5).

³⁰ Ovviamente questa domanda non coincide col chiedersi se egli avrebbe potuto trovare *in quel momento* un motivo per agire o evitare di agire. Il problema è piuttosto se l'agente abbia le capacità di pervenire ad una tale ragione. In altre parole, cerchiamo una ragione giustificativa della sua condotta, non una spiegazione psicologica.

del controllo e della competenza sono riconosciuti in modo esplicito nel diritto quando si presentano casi in cui si sollevano eccezioni circa la capacità di intendere e di volere. Tali requisiti definiscono la capacità di essere agenti in generale e sono all'opera, sebbene implicitamente, nei giudizi che hanno come oggetto i reati di negligenza, imprudenza, ed omissione. In tutti questi casi, l'oggetto di giudizio non è la presenza di un movente criminoso, ma l'assenza di una dovuta considerazione dei rischi legati alla condotta. Pertanto, nonostante i termini di 'intenzione' e 'motive' nell'ambito del giudizio penale sembrano richiamare stati psicologici, queste nozioni sono normative e rendere esplicite queste proprietà nell'ambito del giudizio penale richiede l'impiego dei controfattuali³¹.

6. Conclusione

In questo saggio ho sostenuto che i modelli controfattuali conducono il teorico ad un dilemma: o egli fornisce una teoria monistica della causalità lungo la falsariga del modello NESS, però al costo di non poter spiegare il ruolo della mera sorte, oppure salva la teoria, al costo di rivedere i giudizi sulla responsabilità degli agenti. Ho proposto di seguire questa seconda strada, sostenendo che il ruolo dei controfattuali all'interno ragionamento giuridico ha due funzioni: è una guida euristica nella ricerca delle cause e spiega le caratteristiche normative dei nostri giudizi di colpevolezza. Ho sostenuto che, piuttosto che accertare ciò che gli agenti avrebbero fatto in circostanze particolari, l'obiettivo di un'analisi controfattuale è di accertare, nell'ambito del processo, la responsabilità degli agenti in risposta a ciò che avrebbero dovuto prendere in considerazione al momento dell'azione. La funzione euristica del ragionamento controfattuale è un aspetto cruciale della responsabilità penale. Ho sostenuto che, contro la concezione tradizionale che concepisce la responsabilità essenzialmente come imputazione di un atto causale, il concetto di responsabilità richiama le capacità razionali coinvolte nel rendere conto della condotta propria ed altrui. Questa capacità è la condizione di possibilità del discorso razionale stesso, ossia del gioco linguistico che si sviluppa nelle prassi del 'dare e chiedere ragioni'. Poiché i tribunali e le corti di giustizia

³¹ Questa lettura della colpevolezza ricalca peraltro la concezione normativa dominante in ambito penale. Ad esempio Fiandaca ed Musco affermano: «La concezione normativa oggi dominante afferma che è colpevole un soggetto imputabile, il quale abbia realizzato con dolo o colpa la fattispecie obbiettiva di un reato, in assenza di circostanze tali da rendere necessitata l'azione illecita. La colpevolezza in senso normativo è, pertanto, un concetto necessariamente complesso» (Fiandaca & Musco (2004: 289)).

sono la migliore approssimazione di questo tipo di prassi, l'accertamento della responsabilità nei consessi di giustizia consiste innanzitutto nell'accertamento delle ragioni appropriate per una certa condotta e della capacità richieste per formare ragioni di quel tipo. I risultati di questa analisi sono rilevanti per la stessa teoria della pena: nonostante i tentativi di giustificare la pena sulla base sicura di una teoria generale della causalità, ho sostenuto che tale giustificazione debba basarsi piuttosto sulla valutazione degli atteggiamenti normativi coinvolti nella pratica discorsiva e razionale del diritto.

Bibliografia

- Anscombe, G. E. M. (2000), *Intention*, Cambridge Mass., Harvard University Press.
- Antolisei, F. (1991), *Manuale di Diritto Penale. Parte Generale*, Milano, Giuffrè.
- Brandom, R. B. (1994), *Making It Explicit. Reasoning, Representing, and Discursive Commitment*, Cambridge Mass.-London, Harvard University Press.
- Brandom, R. B. (2000), *Articulating Reasons*, Cambridge Mass., Harvard University Press.
- Brandom, R. B. (2006), *Reason in Philosophy. Animating Ideas*, Cambridge Mass., Harvard University Press.
- Davidson, D. (1963), "Actions, Reasons, and Causes" ("Azioni, ragioni e cause", in *Id, Azioni ed eventi*, trad. it. di R. Brigati, Bologna, Il Mulino, 1992, pp. 41-62).
- Davidson, D. (1967) "The Logical Form of Action Sentences" ("La forma logica degli enunciati d'azione", in *Id, Azioni ed eventi*, trad. it. di R. Brigati, Bologna, Il Mulino, 1992, pp. 163-214).
- Duff, R. A. (2007), *Answering for Crime. Responsibility and Liability in the Criminal Law*, Portland, Hart Publishing.
- Feinberg, J. (2003), *Problems At the Roots of Law*, Oxford – New York, Oxford University Press.

- Fiandaca, G., Musco, E. (2004), *Diritto Penale. Parte Generale*, Milano, Zanichelli, 2^a edizione.
- Fischer, J. M., Ravizza, M. (1999), *Responsibility and Control. A Theory of Moral Responsibility*, Cambridge, Cambridge University Press.
- Fletcher, G. (1998), *Basic Concepts of Criminal Law*, Oxford University Press.
- Fumerton, R., Kress, K. (2001), "Causation and the Law: Preemption, Lawful Sufficiency, and Causal Sufficiency", *Law and Contemporary Problems*, Vol. 64, No. 4, pp. 101-22.
- Gendler, T. S., Hawthorne, J. (a cura di), (2002), *Conceivability and Possibility*, Oxford, Oxford University Press.
- Goodman, N. (1955), *Fact, Fiction, and Forecast (Fatti, Ipotesi, Previsioni*, trad. it. di C. Marletti, Roma-Bari, Laterza, 1985).
- Hart, H. L. A. (1967), *Punishment and Responsibility*, di J. Gardner, a cura di, Oxford - New York, Oxford University Press, 2008 2nd edition.
- Hart, H. L. A., Honorè, T. (1985), *Causation in the Law*, Oxford, Oxford University Press.
- Lewis, D. (1973), *Counterfactuals*, Oxford (UK), Basil Blackwell.
- Lewis, D. (1973), "Counterfactuals and Comparative Possibility", *Journal of Philosophical Logic*, 2, pp. 418-446.
- Lewis, D. (1973), "Causation," *Journal of Philosophy*, 70, pp. 556-67.
- Lewis, D. (1967), "Probabilities of Conditionals and Conditional Probabilities," *Philosophical Review*, 85, pp. 297-315.
- Mackie, J. L. (1980), *The Cement of the Universe: A Study of Causation*, Oxford, Clarendon at Oxford University Press.
- Menzies, P. (2009), "Counterfactual Theories of Causation", in E. Zalta (a cura di) *The Stanford Encyclopedia of Philosophy*.
(On-line: <http://plato.stanford.edu/archives/fall2009/entries/causation-counterfactual/>)

- Moore, M. J. (2009), *Causation and Responsibility. An Essay in Law, Morals, and Metaphysics*, Oxford-New York, Oxford University Press.
- Rawls, J. (1971), *A Theory of Justice (Una teoria della giustizia*, trad. it. di U. Santini, Milano, Feltrinelli, 1999).
- Raz, J. (2009), "Agency and Luck", in Ulrike Heuer & Gerald Lang (a cura di), *Luck, Value, and Commitment: Themes From the Ethics of Bernard Williams*, Oxford University Press, pp. 133-162.
- Raz, J. (2010), "Responsibility & the Negligence Standard", *Oxford Journal of Legal Studies*, 30, 1, pp. 1-18.
On-line: <http://ssrn.com/abstract=1436022>
- Robinson, P. H., Dubbe, M. D. (2007), "The American Model Penal Code: A Brief Overview", *New Criminal Law Review*, Vol. 10, pp. 319-341.
- Scanlon, T. M. (1998a), *What We Owe to Each Other*, Cambridge Mass., Harvard University Press.
- Scanlon, T. M. (1998b), "The Significance of Choice", in G. Watson (a cura di), *Free Will*, Oxford-New York, Oxford University Press, 2003, pp. 352-371.
- Stalnaker, R. (1968), "A Theory of Conditionals", in *Studies in Logical Theory*, di N. Rescher (a cura di), London, Blackwell, pp. 98-112.
- Stella, F. (1975), *Leggi scientifiche e spiegazione causale nel diritto penale*, Milano, Giuffrè.
- Stella, F. (2004) "La vitalità del modello della sussunzione sotto leggi. A confronto il pensiero di Wright e Mackie", in Id. (a cura di), *I saperi del giudice: la causalità e il ragionevole dubbio*, Milano, Giuffrè.
- Sellars, W. (1956), *Empiricism and the Philosophy of Mind*, di R. Brandom (a cura di), Cambridge Mass. - London, Harvard University Press, 1997.
- Watson, G. (2001), "Reasons and Responsibility", *Ethics*, 111, pp. 374-394.

- Wright, R. W. (1991), "Causation in the Tort Law", *The International Library of Essays in Law and Legal Theory, Tort Law*, di E. Weinrib (a cura di), Dartmouth Publishing Co., 1991.
- Wright, R. W. (2004), "Causation, Responsibility, Risk, Probability, Naked Statistics, and Proof: Pruning the Bramble Bush by Clarifying the Concepts", in F. Stella, (a cura di), *I saperi del giudice: la causalità e il ragionevole dubbio*, trad. it., Milano, Giuffrè.
- Wright, R. W. (2011), "The NESS Account of Natural Causation: A Response to Criticisms", in R. Goldberb (a cura di), *Perspectives on Causation*, Hart Publishing.

L'agire razionale nelle spiegazioni sociologiche: dai “tipi ideali” ai modelli di simulazione

Sabrina Moretti
Università degli Studi di Urbino Carlo Bo
sabrina.moretti@uniurb.it

1. Introduzione

Lo sviluppo del concetto di razionalità è stato fondamentale nella storia della sociologia. La concezione illuministica di una ragione unica e universale a cui deve tendere l'azione umana è andata in crisi proprio con la nascita del pensiero sociologico, che ha posto l'attenzione su una idea di razionalità condizionata dalle situazioni storiche e sociali¹.

Tra le varie concezioni di razionalità, quella che ha avuto un impatto maggiore in sociologia, specialmente in ambito metodologico, è quella strumentale che ha origine in economia. Le scienze economiche si sono concentrate su una idea di agire umano che vede l'individuo perfettamente razionale, capace di calcolare i rischi e i profitti di ciascuna azione con la finalità di conseguire sempre i risultati migliori. Una razionalità strumentale, quindi, tesa a raggiungere determinati fini egoistici e facilmente valutabili. Il concetto di strumentalità dell'azione è stato poi generalizzato da Weber², che ha teorizzato l'azione orientata allo scopo, cioè quella azione che valuta i mezzi più adeguati per conseguire l'obiettivo desiderato. Per Weber, l'azione strumentale è quella in cui i mezzi sono adeguati a raggiungere determinati fini. Egli non si pone il problema della razionalità di tali fini, ammettendone l'arbitrarietà: anche le attività magiche possono essere considerate razionali rispetto allo scopo. Questa concezione di razionalità si è poi evoluta dando vita al paradigma dell'individualismo metodologico, secondo il quale la spiegazione

¹ Coccozza (2005).

² Weber (1961).

di un fenomeno sociale deve ricercarne le cause nelle azioni individuali delle quali è il prodotto.

Una parte della cultura del ventesimo secolo ha invece posto l'attenzione sull'importanza dei fattori inconsci. Seguendo questa prospettiva, Marx e Mannheim definiscono agire razionale quello che è in grado di riconoscere e controllare questi fattori. Solo attraverso la presa di coscienza dei condizionamenti sociali si può giungere ad una razionalizzazione delle azioni³.

Habermas dà un'altra interpretazione del concetto di razionalità, legandola a quella di agire comunicativo. Egli definisce il concetto di "razionalità comunicativa", intesa come "il conseguimento dello scopo illocutivo di intendersi"⁴. Pertanto, essa contraddistingue quelle situazioni in cui si è disposti a sostenere una tesi per mezzo di argomentazioni e a confrontarsi con quelle degli altri. Se non si ha la possibilità di argomentare i principi imposti, ma si ricorre ad affermazioni dogmatiche, siamo nel regno dell'irrazionalità.

Il concetto di razionalità, in sociologia, non è stato solo legato ad una visione individualistica, ma per alcuni studiosi come Luhmann esso può essere utilizzato per caratterizzare il sistema sociale. La sociologia non deve considerare la razionalità dell'azione, ma la razionalità sistemica. L'agire razionale è inteso come esperienza che contribuisce "alla soluzione dei problemi sistemici e dunque alla conservazione di strutture riduttive entro un mondo estremamente complesso"⁵.

L'obiettivo del presente lavoro è analizzare le implicazioni metodologiche connesse con le definizioni di agire sociale razionale sviluppate nell'ambito del pensiero sociologico e, in particolare, il ruolo che queste hanno nella costruzione dei modelli di simulazione sociale basati su sistemi multi-agente. Verrà preso in considerazione il concetto di agire razionale maturato principalmente nell'ambito dell'individualismo metodologico, poiché è proprio in questo contesto che si sono evolute le idee e i concetti che hanno favorito la nascita e lo sviluppo dei modelli di simulazione ad agenti in sociologia.

Molto semplicemente, un modello basato sui sistemi multi-agente è un oggetto computazionale composto da agenti autonomi che sono in grado di interagire tra loro, che hanno una parziale rappresentazione del loro ambiente e che sono capaci di prendere decisioni in dipendenza di valutazioni relative alla situazione in cui si trovano ad operare. La realizzazione di tali sistemi, dunque, richiede la definizione di modelli di comportamento degli agenti, in risposta ai mutamenti dell'ambiente nel quale vivono, e delle loro modalità di

³ Izzo (1995).

⁴ Habermas (1986: 65).

⁵ Luhmann (1983: 92).

interazione. Il problema si sposta sulla ricerca di come ogni agente coordini la propria conoscenza e le azioni da effettuare con quelli degli altri⁶.

Per capire meglio la relazione tra il concetto di agire razionale e lo sviluppo dei modelli di simulazione sociale basati su sistemi multi-agente, si ritiene importante fare riferimento a due questioni che hanno caratterizzato il dibattito all'interno dell'individualismo metodologico: la prima è quella che vede l'evoluzione dell'*homo sociologicus* dall'*homo œconomicus*; la seconda è relativa alla relazione tra agire individuale e organizzazione sociale, che rimanda alla ben nota dicotomia "tra approccio olistico e approccio individualistico"⁷.

2. *Homo sociologicus e homo œconomicus*

Secondo il modello razionalistico dell'economia, le scelte individuali seguono il criterio della massima utilità attesa. Ciò significa che l'individuo ha delle preferenze stabilite a priori ed è in grado di valutare esattamente le conseguenze di ciascuna scelta disponibile e di decidere di attivare quei comportamenti che puntano a massimizzare la soddisfazione delle sue preferenze. L'individuo, posto davanti a una serie di alternative, è in grado di raccogliere tutte le informazioni possibili senza l'assunzione di nessun costo e può decidere in piena libertà, valutando le alternative che gli garantiranno i migliori risultati.

Pareto⁸ fu il primo a criticare questa impostazione, sostenendo che l'agire umano non può essere spiegato solo attraverso elementi economici. Egli, infatti, stabilendo una distinzione fra azioni logiche e azioni non-logiche, intendeva rilevare ed enfatizzare il ruolo degli elementi non-razionali dell'agire sociale, che a suo parere caratterizzano la maggior parte delle azioni. Il pensiero di Pareto si distanzia nettamente dal modello classico della teoria della scelta razionale, in quanto si basa sull'assunzione che non si può spiegare l'agire umano ricorrendo al concetto di razionalità.

Altri studiosi criticano la teoria della scelta razionale, senza però respingerla completamente, ma introducendo dei "correttivi sociologici".⁹ Homans¹⁰, ad esempio, pur condividendo la concezione strumentale dell'azione,

⁶ Moretti (2002).

⁷ Cocozza (2005: 25).

⁸ Pareto (1964).

⁹ Izzo (1995: 199).

¹⁰ Homans (1975).

sposta il problema della razionalità verso una prospettiva sociologica, sostenendo che la soddisfazione non riguarda solo aspetti economici, ma anche ricompense emotive e valoriali, come ad esempio l'approvazione sociale. Egli, quindi, introduce una più ampia gamma di finalità che non sono più esclusivamente relative al denaro e ai bisogni materiali.

Simon¹¹ critica il realismo dei modelli classici dell'economia basati sull'*ottimo paretiano*, introducendo il concetto di razionalità limitata. L'individuo, secondo Simon, è limitato sia perché generalmente non possiede tutte le informazioni necessarie per valutare esattamente le conseguenze di ciascuna azione alternativa sia a causa della sua incapacità computazionale nel calcolare tali conseguenze. Occorre quindi dare una nuova definizione di agire razionale collegata alle teorie psicologiche e cognitive. La razionalità denota uno "stile" di comportamento che è appropriato al raggiungimento di un dato obiettivo, nei limiti imposti da certe condizioni o vincoli, un comportamento che mira a una soluzione soddisfacente piuttosto che ottimale. Le decisioni vengono prese seguendo un processo euristico anziché eseguire un calcolo costi-benefici. Anche le istituzioni sociali possono favorire la razionalità, in quanto forniscono un repertorio di esperienze, valori e principi che aiutano a prendere decisioni quando le facoltà limitate non riescono a stimare una realtà complessa.

Secondo Boudon¹², la sociologia si distanzia dall'economia relativamente alla concezione di razionalità dell'azione, che non si deve basare esclusivamente su criteri utilitaristici, ma includere anche tutte quelle azioni "comprensibili" e dotate di senso. Un'azione può essere considerata razionale quando si è in grado di identificare delle "buone ragioni" che muovono l'attore. Per Boudon, molte azioni etichettate come irrazionali dalla sociologia, in realtà, sono solo frutto di un mancato riconoscimento delle motivazioni ultime dei soggetti che le hanno compiute. L'irrazionalità non è completamente negata, ma è associata ad azioni mosse da stati emotivi che lo stesso attore è in grado di individuare e di riconoscere. La concezione boudoniana prevede anche una razionalità assiologica, in quanto considera razionali anche quelle azioni guidate da valori. L'attore sociale, anche quando non ricava utilità dalle sue azioni, può essere in grado di darne una spiegazione: "un giudizio di valore deve essere comunicabile e argomentabile [...] bisogna che una valutazione sia trans-soggettiva". Ciò è possibile in quanto i soggetti appartenenti ad un sistema sociale condividono in genere lo stesso insieme di

¹¹ Simon (1984).

¹² Boudon (1985; 1987).

valori e di conseguenza “molte azioni e credenze dell’attore sociale sono per vocazione pubbliche”¹³.

Per Boudon, la possibilità di dare un senso alle azioni è una funzione della struttura della situazione, quindi la razionalità ha un significato relativo. L’osservatore deve essere in grado di trovare le motivazioni che spingono l’attore a compiere l’azione in una determinata situazione.

Comprendere, nell’accezione weberiana, significa stabilire, tra la situazione dell’attore e le sue motivazioni ed azioni, relazioni tali per cui l’osservatore possa concludere – e convincere i suoi lettori a concludere – che, nella stessa situazione, egli si sarebbe senza dubbio comportato come l’attore¹⁴.

Boudon introduce il concetto di razionalità soggettiva: le buone ragioni hanno lo status di congetture, principi e teorie che la maggior parte delle persone con lo stesso livello di informazione e/o interesse in questione approverebbero se dovessero affrontare la stessa situazione. Esse dipendono dalle credenze dell’attore, dalle esperienze passate e dal contesto sociale.

Concludendo, le varianti sociologiche della teoria della scelta razionale hanno portato a estendere il significato di razionalità. Mentre l’*homo æconomicus* agisce esclusivamente in termini utilitaristici, l’*homo sociologicus* agisce in modo dotato di senso e mosso da buone ragioni. La razionalità non deve riguardare la natura dei fini dell’azione, ma l’adeguatezza dei mezzi; essa è un fatto puramente strumentale. Le finalità non vanno messe in discussione e la dimensione sociale e normativa assume ora un ruolo rilevante. Per lo scienziato sociale le buone ragioni possono essere molteplici, spesso associate a contesti di riferimento normativi più o meno istituzionalizzati. Si introduce così una “concezione pluralistica di razionalità”¹⁵: ci sono ragioni che in un contesto danno vita ad azioni razionali, mentre in altri non sono considerate buone ragioni.

3. Agire razionale e azione collettiva

L’analisi dello sviluppo del concetto di razionalità non può prescindere dalla dicotomia tra la visione olistica della società e quella individualistica che ha da sempre caratterizzato il dibattito all’interno del pensiero sociologico. Secondo la prospettiva olistica, le istituzioni e le strutture sociali condizionano

¹³ Boudon (1997: 177, 374).

¹⁴ Boudon (1985: 73).

¹⁵ La Spina (2003: 441).

l'agire umano. Durkheim¹⁶, rifiutando ogni tipo di riduzionismo, ha sottolineato come il fatto sociale sia una realtà *sui generis* che ha anche carattere coercitivo. Lo strutturalismo, nella sua versione più rigida, ha portato alla concezione che a livello sociale esista una struttura relativamente stabile di posizioni e relazioni che influenzano le motivazioni e le azioni degli attori sociali. Infine, il funzionalismo e la teoria dei sistemi considerano il "sistema" sociale determinante per capire le azioni individuali, in quanto ogni soggetto ne mantiene l'equilibrio grazie all'interiorizzazione di un insieme di norme che apprende durante la fase di socializzazione.

Le teorie che fanno riferimento all'approccio individualistico considerano la società come la somma delle azioni individuali. La teoria della scelta razionale e l'individualismo metodologico pongono la centralità sull'azione umana, che deve essere l'elemento chiave su cui fondare le spiegazioni sociali¹⁷. Ma l'individuo non agisce in maniera isolata, per cui occorre anche affrontare la questione della spiegazione delle azioni collettive; in particolare, è necessario cercare di rendere conto degli effetti dell'interazione sociale e dei processi di formazione, sopravvivenza ed evoluzione dei gruppi organizzati¹⁸.

La teoria dei giochi cerca di descrivere le interazioni che avvengono tra individui che agiscono in maniera utilitaristica, considerando la situazione in cui le conseguenze di una decisione dipendono anche dalle decisioni prese dagli altri. Attraverso un formalismo matematico, cerca di modellizzare in modo particolarmente efficace proprio le situazioni d'interazione, rendendo espliciti nei ragionamenti degli attori che il loro agire deve tener conto delle conseguenze dovute all'interdipendenza delle scelte.

Due o più individui che devono compiere delle scelte hanno determinate preferenze per quanto riguarda i risultati, oltre ad una certa conoscenza delle scelte disponibili a ciascuno e delle rispettive preferenze. Il risultato dipende dalle scelte di entrambi, o di tutti, se le persone sono più di due. Non esiste una soluzione ottimale «indipendente»: bisogna vedere cosa fanno gli altri.¹⁹

“La teoria dei giochi ha dimostrato quanto sia intrattabile il problema della prescrizione di un'azione razionale in una situazione multipersonale dove gli interessi sono opposti”²⁰. Essa dimostra che il comportamento puramente egoistico, puntuale e di breve periodo produce un risultato irrazionale,

¹⁶ Durkheim (1963).

¹⁷ Cocozza (2005: 11).

¹⁸ Collins (1992).

¹⁹ Schelling (1989: 10-11).

²⁰ Simon (1985: 355).

nel senso che non porta alla situazione ottimale, e richiama pertanto l'attenzione sul ruolo del sistema dei valori, delle norme e delle abitudini che incidono fortemente sul comportamento razionale individuale. Le scelte, infatti, dipendono da come si sviluppano le interazioni sociali, per cui l'attore ha delle aspettative sul comportamento altrui e sa che l'altro ha a sua volta delle aspettative sul suo comportamento. Inoltre, la teoria dei giochi si deve confrontare anche con il problema delle differenze culturali che rendono il gioco asimmetrico²¹. Se gli individui provengono da culture diverse avranno anche diversi criteri di valutazione dei risultati. Occorre così individuare insiemi distinti di *payoff* che siano in grado di rappresentare la pluralità delle preferenze²².

Boudon²³ affronta il problema della relazione tra micro e macro livello, sostenendo che le strutture sociali si formano come conseguenze inattese e impreviste delle azioni individuali. I fenomeni sociali devono essere visti come effetto di aggregazione di comportamenti dei singoli attori; da ciò deriva la necessità di studiare la società partendo da una prospettiva individualistica. Boudon riprende in qualche modo il pensiero di Hayek²⁴, il quale introduce il concetto di ordine sociale spontaneo, che si forma per evoluzione ed è in grado di svilupparsi e mantenersi grazie a forze endogene. Secondo Boudon, le spiegazioni sociologiche non devono basarsi sulla ricerca di leggi causali, ma sulla teoria dell'azione. Rifiutando la visione strutturalista e olistica sostiene che ogni fenomeno sociale *M* deve essere considerato come una funzione delle azioni individuali *m*, che a loro volta hanno origine da una determinata situazione *S* dipendente da elementi macrosociali *M'*:

$$M = M(m(S(M')))$$

Hedström²⁵ introduce il concetto di “meccanismo generatore”: la spiegazione sociologica deve essere in grado di individuare i meccanismi che “causano” un determinato fenomeno. Un meccanismo generatore è un insieme di entità e di attività capaci di generare regolarmente un certo effetto. Egli sostiene che la teoria dell'azione deve essere alla base di ogni spiegazione, ma poiché la sociologia si interessa delle conseguenze sociali, è inevitabile fare riferimento anche alle caratteristiche collettive dell'azione. Ciò significa che

²¹ Cfr. Oechssler & Schipper (2003); Sacco & Zarri (2003).

²² Wildavsky (1992).

²³ Boudon (1985).

²⁴ Hayek (1988).

²⁵ Hedström (2006).

per individuare i meccanismi generatori occorre considerare, oltre alle proprietà degli elementi micro, anche le caratteristiche relazionali. Le motivazioni che determinano le azioni degli individui vanno ricercate considerando l'insieme delle credenze, delle preferenze e delle opportunità dell'attore sociale. La teoria della scelta razionale spiega soltanto una piccola parte delle ragioni che spingono gli individui ad agire, ma non può dar conto da sola di tutte le azioni umane. Occorre ogni volta ricercare le ragioni intelligibili, che inevitabilmente dipendono anche dalle credenze e dalle preferenze.

È da notare che con l'evoluzione dell'individualismo metodologico si introducono dei limiti alle spiegazioni riduzioniste, in quanto le azioni dei singoli devono fare riferimento anche a caratteri sociali quali le norme e le istituzioni. Cherkaoui²⁶, pur sostenendo la necessità di individuare i meccanismi generatori dei fenomeni, sottolinea la complessità della relazione tra i livelli micro e macro, nel momento in cui si considera un agente non perfettamente razionale. Non è sempre possibile effettuare un pieno riduzionismo, soprattutto quando le distribuzioni e le strutture delle relazioni condizionano pesantemente le scelte. Questo, se da una parte genera maggior realismo, dall'altra introduce una contraddizione rispetto ai presupposti che stanno alla base di questo paradigma, riducendone l'efficacia e la rilevanza metodologica.

4. Razionalità e modelli di simulazione

Il concetto di razionalità ha avuto un ruolo strumentale per lo studio della società. Per Weber²⁷, l'agire razionale è un tipo ideale che ha anche una funzione metodologica: i tipi ideali sono infatti concetti astratti capaci di individuare regolarità, somiglianze e ricorrenze, costruiti dallo scienziato sociale per cercare di mettere ordine in un mondo multiforme; approssimazioni che consentono di semplificare la realtà empirica e renderla "analizzabile".

La teoria della scelta razionale standard è basata su assunti a priori e su un insieme di preferenze definite *ex ante* che determinano l'azione; gli attori sociali non violano mai tali assunti. L'applicazione di questa teoria è basata su una collezione di assiomi che ne garantiscono la coerenza interna²⁸. In ambito economico, l'utilità soggettiva può essere quantificata senza problemi,

²⁶ Cherkaoui (2005).

²⁷ Weber (1961).

²⁸ Cocozza (2005: 109).

per cui i risultati di ogni scelta possono essere calcolati attraverso una funzione matematica, permettendo di costruire dei modelli capaci di trovare la soluzione ottima.

Quando ci si sposta da una concezione esclusivamente economicistica a una sociologica, che prevede preferenze e obiettivi non materiali, si pone il problema di come “calcolare” l'utilità soggettiva. Emerson, ad esempio, sostiene che per comparare i benefici di tutte le possibili scelte occorre “un concetto di valore o di utilità soggettiva che posseda un'origine non arbitraria e un'unità di misura”²⁹. Senza un sistema di misurazione non è possibile definire una funzione che calcola l'utilità soggettiva e quindi costruire un modello matematico.

Simon, introducendo il concetto di razionalità limitata, ricorre a modelli cognitivi definiti nell'ambito della scienza dell'informazione e dell'intelligenza artificiale. I processi decisionali umani sono in realtà molto più complessi rispetto a quelli ottimizzanti della scelta razionale, poiché tendono a seguire percorsi soggettivi che dipendono da esperienze, situazioni e aspirazioni diverse, ma comunque non rinuncia all'idea che possano essere formalizzati e rappresentati da modelli. Egli infatti ritiene necessario ed euristica-mente fruttuoso fare riferimento a questi strumenti: partendo dall'assunto di razionalità limitata è possibile definire differenti modelli formali di attore razionale, ognuno dei quali riproduce una particolare configurazione di vincoli che strutturano la situazione nel quale opera l'attore. Secondo Simon, la scienza dell'informazione fornisce gli strumenti necessari per costruire modelli adeguati. “Non si può dar conto del comportamento umano, anche se razionale, con una manciata di variabili”³⁰. La tecnica migliore “per trattare situazioni troppo complesse per l'applicazione dei metodi conosciuti di ottimizzazione è la simulazione”³¹. Seguendo la concezione cognitivista, i processi cognitivi, e quindi anche quelli decisionali, possono essere convertiti in programmi per computer; nonostante l'hardware di un computer funzioni in modo diverso dal cervello dell'uomo, è possibile stabilire un'analogia tra il processo di “ragionamento umano” e il processo di elaborazione di un computer.

La razionalità soggettiva di Boudon riprende le assunzioni della sociologia comprendente di Weber. Esistono tante “ragioni” che giustificano un comportamento, scoprire tali ragioni significa “comprendere” le motivazioni e i significati che gli individui danno alle loro azioni. L'analisi sociale deve

²⁹ Emerson (1987: 12).

³⁰ Simon (1985: 317).

³¹ *Ivi* p. 275.

inoltre far ricorso alla sociologia formale definita da Simmel, che elabora costrutti non riferibili a particolari casi empirici, ma utilizzabili “per spiegare situazioni molto diverse, a condizione di fornire le opportune precisazioni in ogni caso osservato”³².

Boudon sostiene che le scienze sociali non sono in grado di trovare delle “leggi universali” di spiegazione, come in altre scienze, per cui si devono concentrare sulla definizione di modelli. Il modello ha validità locale, piuttosto che generale, e rappresenta una “situazione ideale”, un’*approssimazione della realtà non riscontrabile empiricamente*; ogni volta che vogliamo spiegare un fenomeno empirico attraverso il modello dobbiamo introdurre ulteriori informazioni, occorre cioè trasformare in costanti alcune variabili. Con il concetto di modello Boudon intende ciò che Simmel intendeva con sociologia formale. I modelli – o teorie formali secondo il linguaggio di Simmel – propongono un quadro concettuale che può spiegare una particolare classe di fenomeni empirici differenti, “nel momento in cui le variabili incluse nelle funzioni proposizionali che la compongono sono sostituite da costanti pertinenti, adattate cioè alle caratteristiche della situazione considerata in ciascun caso”³³ e offrono un linguaggio utile per l’analisi di determinati processi.

Questa concezione di modello si adatta al concetto di razionalità “*soggettiva*”. È infatti possibile costruire modelli differenti di attore razionale, ognuno dei quali basato su una logica di azione peculiare, che racchiude gli aspetti legati ai valori, alle credenze, alla socializzazione e all’ambiente di riferimento. Ciò consente di rappresentare attori e situazioni idealtipiche e di spiegare una certa classe di eventi determinati da queste logiche di azione.

Secondo la sociologia analitica, per trovare i meccanismi generatori che causano certe ricorrenze occorre partire da una teoria dell’azione che ponga l’attenzione sulle caratteristiche collettive, come i reticoli di interazione, le credenze diffuse e le norme sociali³⁴. Le spiegazioni devono considerare attori idealtipici, le cui intenzioni non sono necessariamente “*identiche alle intenzioni di alcun attore concretamente esistente, ma dovrebbero essere viste piuttosto come intenzioni caratteristiche che potrebbero motivare attori tipici a comportarsi in modi specifici*”³⁵. La sociologia analitica evidenzia la necessità di definire spiegazioni sociologiche che condividano certi standard di “*formalizzazione concettuale*”: al “*fine di comprendere le relazioni complesse che esistono fra azione, interazione ed effetti sociali emergenti, è quindi necessario formalizzare adeguatamente la logica astratta del processo*

³² Boudon (1985: 252).

³³ *Ivi* p. 254.

³⁴ Hedström (2006: 7).

³⁵ *Ivi* p. 48.

che si suppone sia all'opera"³⁶. I modelli di simulazione basati su agenti sembrano i migliori strumenti per definire formalizzazioni di questo tipo, poiché consentono di rappresentare diverse tipologie di strutture di interazione e di logiche di azione. Essi consentono di definire una popolazione di agenti che rappresentino degli attori sociali idealtipici e che agiscono seguendo una certa "razionalità" in una situazione specifica, anch'essa idealtipica.

Le applicazioni dei modelli multi-agente alle scienze sociali, pur utilizzando una grande varietà di approcci teorici, condividono con l'individualismo metodologico l'obiettivo di individuare quali siano i processi micro che determinano un fenomeno a livello macro. Si parte da un "fatto stilizzato"³⁷ a livello macro, poi si definiscono delle ipotesi sui processi micro che potrebbero determinarli e si verificano tali ipotesi con la simulazione, dando origine a quella che Epstein chiama una "spiegazione generativa"³⁸. Il modello, quindi, dovrà specificare dei protocolli che definiscono il comportamento e le interazioni di ogni agente, in dipendenza della situazione in cui si troveranno ad operare durante la simulazione.

In letteratura troviamo una grande varietà di applicazioni che interpretano in maniera diversa la relazione tra il micro e il macro livello sociale. Macy e Flache³⁹ sostengono che i modelli possono essere divisi in due categorie rispetto alle logiche di azione implementate: quelli basati sulle scelte e quelli basate sulle regole di comportamento. Nel primo caso, gli agenti hanno delle preferenze o degli scopi precisi che vengono esplicitati e le azioni sono scelte tra una serie di alternative, ognuna delle quali è valutata in maniera strumentale. Anche se l'agente ha informazioni incomplete o capacità limitate di valutazione, le scelte vengono fatte seguendo un criterio di soddisfazione personale. Nel secondo caso, gli agenti mettono in atto comportamenti di routine, sulla base di regole che forniscono una reazione standard ad una determinata condizione. Ciò si verifica, ad esempio, quando si vuole rappresentare un comportamento dettato da una norma.

Un modello molto semplice che segue una logica di azione di tipo strumentale è il famoso modello di Schelling⁴⁰ sulla segregazione razziale. In questo caso, infatti, ogni agente ha come obiettivo quello di avere una certa percentuale di vicini omofili; le azioni scelte sono quelle che permettono di raggiungere tale scopo.

³⁶ *Ivi* p. 96.

³⁷ Heine *et al.* (2005).

³⁸ Epstein (1999).

³⁹ Macy & Flache (2009).

⁴⁰ Schelling (1971).

Anche i modelli di simulazione basati sulla teoria dei giochi seguono la logica strumentale. Il gioco è caratterizzato da procedimenti che cercano di definire i possibili esiti di interazioni strategiche tra più agenti aventi una chiara gerarchia di preferenze. Tutti i comportamenti scelti dagli agenti rispondono ad un'esigenza utilitaristica, compreso il comportamento imitativo; l'imitazione infatti viene adottata ogni volta che si reputa possa portare al soddisfacimento dei propri desideri. L'imitazione è un mezzo per raggiungere uno scopo e le regole di interazione sono endogene, nel senso che le strutture di interazione sono determinate dalle decisioni razionali dei singoli attori sociali. Nei modelli basati sulla teoria evoluzionistica dei giochi⁴¹ i comportamenti normativi emergono come conseguenza imprevedibile delle interazioni di gioco strategico. I meccanismi interazionali descritti dalla teoria dei giochi e basati sull'assunto dell'agire utilitaristico sono i fattori che generano le norme sociali.

Molti modelli che studiano la diffusione delle idee sono invece basati su regole di comportamento routinarie; gli agenti non hanno preferenze e scopi definiti esplicitamente, decidono di adottare un comportamento imitativo semplicemente perché sono definiti come "conformisti".⁴² L'imitazione in questo caso è un fine dell'attore sociale e non un mezzo⁴³ e le strutture di interazione sono esogene, sono cioè un input del modello e non emergono dalle decisioni individuali. Nella maggior parte dei casi, tali regole di comportamento non vengono definite sulla base di regolarità empiriche stilizzate, ma piuttosto sono degli assunti specificati *ex ante* che si ipotizza possano generare il fenomeno in questione. A volte l'origine di questi assunti non è chiara; in particolare non si capisce se i comportamenti vengano intesi come frutto di norme e credenze sviluppate all'interno di un certo contesto sociale e apprese attraverso un processo di apprendimento e di socializzazione, o siano considerate come un comportamento "naturale" o come prodotto di un'evoluzione culturale⁴⁴. Nel primo caso questo tipo di azione può essere considerata razionale secondo la concezione di Boudon, poiché discende da un principio normativo; l'attore X crede in Z e di conseguenza ha delle buone ragioni per agire in coerenza con Z. L'obiettivo della spiegazione generativa che ne consegue è verificare se queste norme facciano emergere situazioni, regolarità e ordine sociale; in caso affermativo si può dedurre che esse sono funzionali al sistema. Le regole di interazione rispondono così ad esigenze di

⁴¹ Cfr. Axelrod (1997; 1984).

⁴² Chavalarias (2006).

⁴³ Hedström (1998).

⁴⁴ Trajkovski & Collins (2009); Henrich & Boyd (1998).

razionalità collettiva piuttosto che di razionalità individuale. Se invece consideriamo il comportamento imitativo come un aspetto “naturale” o generato da un processo evolutivo, allora siamo di fronte ad azioni alle quali è difficile attribuire una motivazione.

5. Conclusioni

La teoria della scelta razionale è basata su assunti ben precisi sul comportamento umano: l'attore sociale è dotato di una razionalità perfetta e di abilità di calcolo illimitate, per cui è capace di valutare le conseguenze delle proprie azioni senza commettere errori. Inoltre, poiché tutti gli individui hanno le stesse caratteristiche e le stesse capacità, è possibile fare riferimento ad un solo agente ‘rappresentativo’ e analizzare matematicamente i problemi in esame.

Quando si passa dall'*homo oeconomicus* all'*homo sociologicus*, in cui l'agente cognitivo rimpiazza l'agente perfettamente razionale, la simulazione ad agenti diventa uno strumento rigoroso che sostituisce il calcolo analitico. L'agente viene rappresentato da algoritmi computazionali in grado di agire secondo l'assunto di razionalità limitata, basandosi anche su fenomeni di adattamento, apprendimento e imitazione. I modelli simulativi permettono così di rappresentare con maggior realismo l'attore sociale “tipicizzato” per mezzo di strumenti standardizzati e intersoggettivamente accessibili, in quanto consentono di esprimere la razionalità soggettiva attraverso strumenti formali, allo stesso tempo flessibili e rigorosi.

D'altra parte però la teoria cognitiva della razionalità appare poco chiara e a volte confusa, a differenza della teoria della scelta razionale che è semplice, ben definita e completamente trasparente. Di conseguenza, ha generato una molteplicità di approcci disomogenei che hanno a loro volta determinato una produzione di modelli di simulazione spesso difficilmente comparabili tra loro. È difficile giungere ad un lavoro organico e anche la discussione e il confronto sui risultati diventano problematici.

È da notare, inoltre, che la “spiegazione generativa” basata sull'agente cognitivo rischia di dar vita a una struttura logica circolare. Infatti, mentre la teoria della scelta razionale è incentrata su un completo riduzionismo che stabilisce *ex ante* le preferenze e riconduce le spiegazioni alle intenzioni individuali, la teoria dell'azione basata sull'agente cognitivo fonda le spiegazioni anche su elementi sociali come le norme. La condotta degli agenti può dipendere da regole di comportamento che a loro volta sono il risultato di qualche processo di interazione ripetuta degli agenti con il loro ambiente sociale.

Per uscire da questa circolarità, le spiegazioni che si avvalgono dei modelli di simulazione devono a volte ricorrere alle idee sviluppate dal funzionalismo sociologico e dalla teoria dei sistemi, allontanandosi così dalla visione riduzionista dell'individualismo metodologico. Nelle simulazioni in cui l'azione è basata su un comportamento normativo, l'obiettivo delle spiegazioni è di verificare quali siano le conseguenze di tali norme all'interno dell'ambiente sociale in cui operano gli agenti. Le regole di comportamento vengono considerate delle norme interiorizzate e la simulazione non si preoccupa di analizzarne i processi di formazione e trasformazione, ma piuttosto di determinare se sono funzionali al sistema sociale. Gli elementi di base della spiegazione generativa sono "tipi ideali" di interazione sociale, in questo caso dettate da una norma, e le simulazioni hanno l'obiettivo di verificare le proprietà che emergono come conseguenze inattese di tali interazioni.

Concludendo, dunque, per risolvere il problema della relazione micro-macro che si è creato all'interno dell'individualismo metodologico, a volte i modelli di simulazione effettuano un riduzionismo i cui atomi delle spiegazioni sono interazioni sociali "tipicizzate" che includono contemporaneamente sia requisiti sociali che individuali, senza ricorrere alle finalità e alle intenzioni degli agenti. Spesso, però, rimane non risolto il problema di come definire le tipologie di interazione, in particolare se esse debbano avere un carattere ontologico o fenomenologico.

Bibliografia

- Axelrod, R. (1984), *The Evolution of Cooperation*, New York, Basic Books.
- Axelrod, R. (1997), *The Complexity of Cooperation. Agent-Based Models of Competition and Collaboration*, Princeton, University Press.
- Boudon, R. (1985), *Il posto del disordine*, Bologna, Il Mulino, ed. or. 1984.
- Boudon, R. (1987), "Razionalità e teoria dell'azione", *Rassegna Italiana di Sociologia*, XXVIII, 2, pp. 175-203.
- Boudon, R. (1997), *Il vero e il giusto*, Il Mulino, Bologna, ed. or. 1995.
- Chavalarias, D. (2006), "Metamimetic Games: Modeling Metadynamics in Social Cognition", *Journal of Artificial Societies and Social Simulation*, vol. 9, no. 2. On line: <http://jasss.soc.surrey.ac.uk/9/2/5.html>

- Cherkaoui, M. (2005), *Invisible Codes: Essays on Generative Mechanisms*, Oxford, Bardwell Press.
- Cocoza, A. (2005), *La razionalità nel pensiero sociologico tra olismo e individualismo*, Milano, Franco Angeli.
- Collins, R. (1992), *Sociological Insight: An Introduction to Non-Obvious Sociology*, New York, Oxford University Press.
- Durkheim, E. (1963), *Le regole del metodo sociologico*, Milano, Comunità, ed. or. 1895.
- Emerson, T. M. (1987), "Toward a Theory of Value in Social Exchange", in K. S. Cook (a cura di), *Social Exchange Theory*, Sage, Newsbury Park, pp. 11-46.
- Epstein, J. M. (1999), "Agent-Based Computational Models and Generative Social Science", in *Complexity*, vol. 4, no. 5, pp. 41-60.
- Habermas, J. (1986), *Teoria dell'agire comunicativo*, Bologna, Il Mulino, ed. or. 1981.
- Hayek, F. von (1988), *Conoscenza, mercato, pianificazione, Saggi di economia e di epistemologia*, Bologna, Il Mulino.
- Hedström, P. (1998), "Rational imitation", in R. Swedberg (a cura di), *Social Mechanisms: An Analytical Approach to Social Theory*, Cambridge, Cambridge University Press, pp. 306-27.
- Hedström, P. (2006), *Anatomia del sociale. Sui principi della sociologia analitica*, Milano, Bruno Mondadori, ed. or. 2005.
- Heine, B. O., Meyer, M., Strangfeld, O. (2005), "Stylised Facts and the Contribution of Simulation to the Economic Analysis of Budgeting", in *Journal of Artificial Societies and Social Simulation*, vol. 8, no. 4. On line: <http://jasss.soc.surrey.ac.uk/8/4/4.html>.
- Henrich, J., Boyd, R. (1998), "The Evolution of Conformist Transmission and the Emergence of Between-Group Differences", in *Evolution and Human Behavior*, vol. 19, pp. 215-241.

- Homans, G. C. (1975), *Le forme elementari del comportamento sociale*, Milano, Franco Angeli, ed. or. 1961.
- Izzo, A. (1995), *I percorsi della ragione. Il tema della razionalità nella storia del pensiero sociologico*, Roma, NIS.
- La Spina, A. (2003), "Lo spazio della persona. Alcune precisazioni sulle teorie del comportamento razionale e sull'individualismo metodologico", in *Studi di Sociologia*, Anno 41, Fasc. 4, pp. 439-454.
- Luhmann, N. (1983), *Illuminismo sociologico*, Milano, Il Saggiatore, ed. or. 1970.
- Macy, M., Flache, A. (2009), "Social Dynamics From The Bottom Up: Agent-Based", in P. Hedström e P. Bearman (a cura di), *The Oxford Handbook of Analytical Sociology*, Oxford, Oxford University Press.
- Moretti, S. (2002), "Computer simulation in sociology: what contribution?" in *Social Science Computer Review*, 20, 1, pp. 43-57.
- Oechssler, J., Schipper, B. (2003), "Can you Guess the Game You're Playing?", in *Games and Economic Behavior*, vol. 43, no. 1, pp. 137-152.
- Pareto, V. (1964), *Trattato di sociologia generale*, Milano, Comunità.
- Sacco, P. L., Zarri, L. (2003), "Complessità motivazionale, interazione strategica e gestione del conflitto", in *Etica ed Economia*, voll. 1-2, pp. 73-89.
- Schelling, T. C. (1971), "Dynamic models of segregation", in *Journal of Mathematical Sociology*, vol. 1, pp. 143-86.
- Schelling, T. C. (1989), "Un primo approccio alla teoria dei giochi", in G. E. Rusconi (a cura di), *Giochi e paradossi in politica*, Torino, Einaudi, pp. 9-42.
- Simon, H. A. (1984), *La ragione nelle vicende umane*, Bologna, Il Mulino, ed. or. 1983.
- Simon, H. A. (1985), *Causalità, razionalità, organizzazione*, Bologna, Il Mulino.

Trajkovski, G., Collins, S. G. (a cura di) (2009), *Handbook of Research on Agent-Based Societies: Social and Cultural Interactions*, Hershey (PA), Information Science Reference.

Weber, M. (1961), *Economia e società*, Milano, Comunità, ed. or. 1922.

Wildavsky, A. (1992), "Indispensable Framework or Just Another Ideology?", in *Rationality and Society*, vol. 4, pp. 8-23.

Due economie della conoscenza a confronto (o della razionalità del realismo)

Stefano Vaselli

Università degli Studi di Torino, LabOnt, Consorzio Dottorato FINO
stefano.vaselli@gmail.com

1. Il problema del realismo e degli antirealismi e delle loro, rispettive, razionalità

Il tema che si vorrebbe esaminare in questo saggio, sebbene incardinato in un argomento a favore del realismo ontologico, ha la sua naturale sorgiva in una domanda che ha visto solcare l'epistemologia contemporanea come un vero e proprio *leitmotiv* negli scritti di studiosi entrati a far parte della *koinè* anglosassone come Kuhn e Feyerabend, ma anche di teorici più eterodossi e decisamente "anti-analitici", come Latour e Foucault: che tipo di rapporto può sussistere tra potere e conoscenza? In termini più prosaici ma anche più diretti: che tipo di potere può rappresentare per un soggetto epistemico S entrare in possesso di una conoscenza che p , rispetto ad altri soggetti epistemici $S_1, S_2, S_3, \neq S$, o in anticipo su questi ultimi? La conoscenza, non necessariamente di tipo scientifico e naturalistico, come attività o processo sociale, è in grado, in qualche modo, di originare delle intrinseche concentrazioni di "ricchezza conoscitiva", delle sperequazioni cognitive, o addirittura dei soggetti talmente ricchi di conoscenze in grado di poter controllare il processo di formazione e distribuzione delle informazioni in modo strategicamente esiziale per tutti gli altri operatori epistemici loro *competitor*? Quello che seguirà, così, è un argomento non perorato solo a sostegno di un'idea di "conoscenza come

Stefano Vaselli, "Due economie della conoscenza a confronto (o della razionalità del realismo)", in Pierluigi Graziani, Giorgio Grimaldi, Massimo Sangoi, *Animali razionali. Studi sui confini e sulle possibilità della razionalità*, pp. 177-202

© 2015 Isonomia, Rivista online di Filosofia – Epistemologica – ISSN 2037-4348

Università degli Studi di Urbino Carlo Bo

<http://isonomia.uniurb.it/epistemologica>

risorsa economica”¹ ma a favore, altresì, della proposta per cui tale “equivalenza tassonomica” andrebbe accolta non solo dalle scienze sociali ma anche, e soprattutto, dall’epistemologia filosofica. La conoscenza, così, non risulterebbe essere solo un vero e proprio “*taxon* merceologico” umano, ma anche uno tra i beni più importanti sia per la sopravvivenza, sia per la stessa “dignità razionale” della nostra specie, essendo il suo più inestimabile tesoro evolutivo. Ciò non di meno, detto “tesoro” non è inesauribile e, anche se la sua inarrestabile crescita, come osserva il premio Nobel Douglass C. North, continua ad esservi anche in assenza di incentivi sociali, non si produce “da solo”. *Fatti non foste a viver come bruti/ma per seguir virtute e canoscenza* non sono due versi in grado di confutare quell’*aforisma*, sicuramente più prosaico (ma non meno profondo) attribuito a M. Friedman, per cui «nessun pasto è gratis». Considerare quest’ipotesi ci aiuterebbe, di contro, a scegliere un sentiero argomentativo che, strada facendo, conduce ad un importante bivio problematico: *perché* la conoscenza è una risorsa intrinsecamente economica? *Come* ne è possibile una *distribuzione* ottimale? Considerare l’oggetto della conoscenza da una prospettiva *realista* o, di contro, *comunitarista* o *costruttivista*, può rappresentare realmente un bivio dilemmatico tra due diverse e opposte tra loro *economie della conoscenza*?

La prima ipotesi è questa: ciò che contraddistingue sin dai suoi albori, come una specie unica *homo sapiens sapiens*, sono le sue peculiari, se non uniche, capacità e competenze cognitive e, in più, la sua abilità nel saper produrre e distribuire la conoscenza come una vera e propria e assai elaborata, “risorsa sociale”. Ma allora, se seguissimo (solo metaforicamente) la tesi della produzione/alienazione di Marx², l’uomo si identificherebbe, sì, con il proprio “conoscere” e con le modalità *simbolico-culturali* per via delle quali conosce quel che conosce, e quindi con la propria derivante identità storico-evolutiva (che non è *solo* una mera memoria storica³). Diversamente da altri organismi, però (e *contra* Marx) gli esseri umani non sarebbero “*solo*” qualcosa in più di quel che producono; possono identificarsi con *ciò che conoscono* e con il “come lo conoscono”, dal momento che la conoscenza, comunque la si pensi

¹ Ci riferiamo a conoscenza intendendola come credenza vera e giustificata, tesi risalente a Platone, Menone §97 e Teeteto, §203c, e uscita malconcia dall’analisi di Gettier (1963: 121-23) e, risposta in una bella ricostruzione analitica da Vassallo (2003).

² Marx, in questo non fece altro che seguire un’intuizione e una speranza, come si è visto, di Smith (1776); Roncaglia (2005).

³ Pomian, (1999).

al riguardo, coinvolge (secondo alcuni già al livello percettivo⁴) capacità e competenze concettuali e, quindi, simboliche.

Di qui, assunta l'unicità "simbolica" della nostra specie⁵, risulterebbe abbastanza semplice mostrare perché il tratto fondamentale della nostra capacità di sviluppo sarebbe quella basata sulla natura epistemica di *homo sapiens-sapiens*. Ma se quest'ipotesi fosse corretta ne deriverebbero delle conseguenze molto interessanti anche per il nostro usuale modo di intendere l'equilibrio delle forze interne ai processi di formazione e distribuzione della conoscenza, equilibrio che, lungi dal dipendere solo da fattori logicamente *endogeni* (la natura delle teorie, il loro potere esplicativo, ecc.), potrebbe dover dipendere anche da fattori più *esogeni*, come l'affermarsi di veri e propri *monopolisti* o *oligopolisti* della conoscenza o, come vennero chiamati mezzo secolo fa dall'economista K. J. Arrow, da dei veri e propri "dittatori della scelta sociale" in questo caso della scelta epistemica⁶. Questa seconda ipotesi è quella su cui verranno forniti più lumi, essendo anche la più difficile da esporre in modo non opaco. Com'è possibile, infatti, che da un processo di produzione e distribuzione della conoscenza, possano derivare delle condizioni di squilibrio, per cui *il potere conoscitivo* risulterebbe *ipso facto* più intrappolato nelle "maglie esplicative" di certi "pescatori di conoscenza" e delle loro "reti", ossia delle loro teorie, che nelle reti conoscitive di altri, *a parità* (del tutto teorica, ovviamente) *di potere di accesso*? Inoltre: perché è così importante per la crescita della conoscenza che le cose debbano andare proprio in questo modo?

Per rispondere a tutte queste domande, pertanto, è necessario istituire un vero e proprio confronto tra due concezioni della conoscenza e della verità oggi tornate prepotentemente in conflitto tra loro da un dibattito che, verso la fine del secolo scorso, sembrava, invece, destinato ad esaurirsi: il realismo oggettivo (soprattutto considerato nella sua "nuova" veste di Nuovo Realismo o dei "nuovi realismi" filosofici) e le posizioni ad esso alternativo e da esso fortemente criticate. Ma tale confronto dovrà essere attuato in maniera nuova: si tratta di vedere se, soggiacenti a queste due differenti concezioni ontologiche dell'oggetto della conoscenza, vi siano due differenti "economie del conoscere", due metodi affatto e radicalmente differenti di vedere nella conoscenza un prodotto non solo della nostra creatività e dei nostri schemi

⁴ McDowell (1994). Agli antipodi di McDowell, ma sempre sensibile ad una natura "economica" in quanto "ecologica" della percezione, è Gibson (1979) dove il padre della psicologia ecologica parla espressamente di "disponibilità o furniture percettive all'uso sensorio-motorio di un organismo" o *affordances* (da *to afford*: fornire).

⁵ Jablonka & Lamb (2005).

⁶ Arrow (1951).

razionali ma anche delle risorse che siano (o non) “esternalità” pre-economiche rispetto al *trade-off* della nostra concettualizzazione, ovvero provenienti (o meno) da un dittatoriale mondo esterno, sempre in grado di imporci le sue leggi e le sue invarianti naturali.⁷

2. La favola dei coleotteri

In un’ingegnosa e kafkiana *reductio ad absurdum* della possibilità stessa di un “linguaggio privato”, Wittgenstein invita ad immaginare l’angusta situazione d’incomunicabilità fenomenologico-cognitiva dei partecipanti ad un improbabile *linguistic game*, consistente nel tenere, ciascuno, uno scarabeo prigioniero in una propria scatola:

Supponiamo che ciascuno abbia una scatola in cui c’è qualcosa che noi chiamiamo “coleottero”. Nessuno può guardare nella scatola dell’altro; e ognuno dice di sapere che cos’è un coleottero soltanto guardando il suo coleottero. Ma potrebbe ben darsi che ciascuno abbia nella sua scatola una cosa diversa. Si potrebbe addirittura immaginare che questa cosa mutasse continuamente. Ma supponiamo che la parola “coleottero” avesse, tuttavia, un uso per queste persone! Allora non sarebbe quello della designazione di una cosa. La cosa contenuta nella scatola non fa parte in nessun caso del giuoco linguistico; nemmeno come un qualcosa; infatti, la scatola potrebbe anche essere vuota. No, si può «dividere per» la cosa che è nella scatola; di qualunque cosa di tratti, si annulla⁸.

Poniamo, adesso, che in questo “gioco”, solo una persona, molto più fortunata degli altri, abbia un reale coleottero nella propria scatola (mentre, gli altri: «una cosa diversa o mutante» o il nulla) e che, trovando e consultando un atlante entomologico, venga a scoprirlo *per primo*. Libro alla mano, costei potrebbe scegliere di aprire la propria scatola e di illustrarne il contenuto agli altri, violando, perciò, le rigide restrizioni poste dal filosofo austriaco. Così facendo, però, questa persona si rivelerebbe l’unica ad aver *realmente* giocato allo “scarabeo prigioniero” o, in più, vincerebbe il gioco stesso? Di certo, da quel momento in poi, non si potrebbe più «dividere per» il contenuto della sua scatola, senza pure «dividere per» le nomenclature entomologiche vigenti. Ergo, questo giocatore, con la sua “innovazione”, instaurando tra sé e gli altri una *nuova configurazione* nell’equilibrio economico della conoscenza, (in questo caso dei coleotteri), diventerebbe il nuovo padrone di questo gioco linguistico, e le certezze di chi, in esso, si illudeva di sapere cosa fosse un coleottero, o di chi pensava di poter ingannare gli altri giocatori,

⁷ Ferraris (2012).

⁸ Wittgenstein (1953).

magari bluffando con una scatola vuota, potrebbero essere deluse o smascherate. Poiché ogni *scoperta*, spesso, rivela che qualcuno (noi, gli altri) ignorasse qualcosa, alcuni individui si scopriranno, prima o poi, impoveriti da perdite di pseudo-conoscenza, cioè da errori, auto-inganni o *bluff*. Questo, appena illustrato, è un tipico esempio di “economia epistemica” che, in questa sede, sarà definito *realista*. La *reale* suddivisione in natura delle specie di insetti, in questo modello è detta, in termini tecnici, un’*esternalità economica*, non suscettibile di ulteriori contrattazioni esplicative, che il proprietario del vero scarabeo può limitarsi a gestire con un *management* euristico, dove, da un gratuito incognito (l’evoluzione naturale degli insetti), può ricavare un fortuito *primum cognitum*, per cui *quell* insetto è proprio uno scarabeo.

Ma, questa non è la sola economia cognitiva attuabile in questo “gioco”. Un’alternativa un po’ fiabesca (nel senso della *Fable of Bees* di Mandeville), che si potrebbe definire *costruttivista* o *comunitarista*, vede, all’apertura della scatola del proprietario del coleottero, un altro giocatore alzare la mano e chiedere di giustificare i dati contenuti nell’atlante del fortunato vincitore avanzando il sospetto che *anche la moderna teoria zoologica* potrebbe essere un altro, gigantesco “scatolone” teorico, chiuso e sigillato a sua volta. Anzi, ché basarsi su dei mitici “dati esterni”, che potrebbero esser frutto, a propria volta, di altre contrattazioni o “costruzioni” concettuali, l’intera *comunità* degli “entomologi del senso comune” potrebbe decidere, per mettere tutti d’accordo, di permettere a ciascuno di crearsi una propria teoria dell’insetto che ognuno tiene rinchiuso, negoziandone, a scatola chiusa, ciascun dettaglio anatomico (*pace* Wittgenstein). Infine, lo scarabeo *standard* sarà quello vincente la sfida del libero mercato della conoscenza, il cui equilibrio sarà stato in grado di soddisfare il maggior numero di domande e offerte di giustificazioni per credere che quel “qualcosa” fosse proprio uno scarabeo.

La conoscenza, con i suoi operatori e le loro teorie, formerebbe, dunque, un mutevole e competitivo gioco economico? Già in passato, Lord Bacon, anche lui ispirato da insetti (e aracnidi), ed E. Mach, si posero la stessa domanda, rispondendo positivamente. Il primo, esaltando le virtù economiche dell’ape induttivista, allegorico ibrido delle indoli produttive delle formiche empiriste e dei ragni a-prioristi⁹, il secondo, chiosando con mirabile sintesi, che «tutta la scienza ha lo scopo di sostituire, ossia di *economizzare*, esperienze mediante la riproduzione e l’anticipazione dei fatti nel pensiero¹⁰». Su

⁹ Bacon (1620, trad. it. 72-3).

¹⁰ Mach (1883, trad. it. 470). Peirce ha ben parafrasato questa tesi: «Ogni verità che ci *risparmierà* di essere sorpresi da un fatto percettivo futuro, che ci fornirà i mezzi per prevederlo, o

questa scia, si potrebbe ipotizzare che *ogni* contenuto epistemico non banale sia *sempre* stato, *di per sé*, una risorsa economica, ovvero «un mezzo scarso applicabile ad usi alternativi¹¹», sin dai tempi più remoti in cui *homo sapiens-sapiens* iniziava a farne i primi «usi alternativi», gettando, anche sotto il profilo tecnologico, le basi del proprio progresso culturale. Ed è un'ipotesi, questa, di *economicità ontologica della conoscenza*, che potrebbe spalancare imprevedibili scenari epistemologici. Ogni equilibrio socio-economico, infatti, lungi dal dipendere solo da meccanismi *esogeni*, come l'origine storica delle risorse, è sensibile anche a fattori più tipicamente *endogeni*, come l'inevitabile affermarsi di monopolisti o dittatori della scelta sociale. Un'idea, questa, del Premio Nobel per l'economia 1972, K. J. Arrow, che è corollario di un suo famoso risultato: il teorema di possibilità (o impossibilità) di funzioni pubbliche di scelta sociale per sistemi di scelta elettiva o di libera concorrenza¹². Di esso, un altro economista iridato dal Nobel ha detto: «*la devastante scoperta di Arrow è, per la politica, ciò che il teorema di Gödel è per la matematica*¹³». Una stima non mal riposta, trattandosi della dimostrazione dell'incompatibilità *logica* di ogni metodo di scelta, con l'esprimibilità collettiva di preferenze individuali, (politiche, economiche, sociali), per più di 2 opzioni, vale a dire del fatto che *ogni* meccanismo sociale di scelta, entro certi ampi limiti, rende, *ceteris paribus*, la libertà di scelta più intrappolabile in certe «scatole di detentori di preferenze» piuttosto che in altre. Una funzione sociale distributiva di scelte vertenti su opzioni teoriche di ipotesi esplicative o predittive, anziché su voti o beni di consumo, valendo l'ipotesi di economicità ontologica della conoscenza, andrebbe, perciò, incontro agli stessi vincoli logici del teorema di Arrow, e si potrebbe sviluppare, così, senza adire a sospetti di *imperialismo economicistico*¹⁴, un confronto valutativo tra le due economie epistemiche poc'anzi visualizzate (realismo e comunitarismo) al fine di individuarne pregi e difetti concettuali.

i mezzi per prevedere condizionalmente che cosa percepirebbe qualcuno che fosse nella situazione di percepirlo; questo [...] è ciò che la scienza prende in considerazione» Peirce (1901: 485-538, 498).

¹¹ Robbins (1932, trad. it. 20).

¹² Arrow (1951: 63).

¹³ Samuelson (1974: 120).

¹⁴ Come quello attaccato da Guala (2006: 95-111).

3. Quattro definizioni e una tesi

Il primo punto nell'agenda di questa tesi è l'adozione del concetto di conoscenza come *credenza vera e giustificata*, risalente a Platone, uscita malconcia dall'analisi di E.L. Gettier, ancor oggi difesa da autori come P. Boghossian e D. Marconi¹⁵, ma rifiutata dalla teoria *sociologica della conoscenza scientifica*, dai cui approcci, pertanto, questa tesi rivendica la propria estraneità¹⁶.

Con *teoria economica della conoscenza*, perciò, si indicherà uno specifico ambito delle scienze sociali, il cui campo di studio è la teoria della produzione, della distribuzione, e della valutazione dei sistemi e delle strutture della conoscenza umana, che risulti articolata in modo tale da rendere compatibile la definizione di conoscenza sopra adottata, con il concetto di *risorsa*, cioè di mezzo scarso per usi *epistemici* alternativi. Pertanto, questa teoria non va confusa con la *teoria della conoscenza economica*, sebbene, sotto molti aspetti, i due ambiti di ricerca risultino fortemente interconnessi¹⁷. Per il primo economista moderno ad aver dato un contributo significativo all'economia della conoscenza, F.A. Von Hayek, il problema che essa pone allo scienziato sociale è «quello di trovare un metodo che non soltanto utilizzi con maggiore efficacia la *conoscenza dispersa* tra i singoli membri della società, ma che sfrutti anche meglio le loro capacità di scoprire e di esplorare nuove cose»¹⁸. La teoria hayekiana della dispersione della conoscenza afferma, così, che ogni tentativo di concentrare quelle scarse conoscenze, determinanti in ognuno un'ignoranza *antropologica* di tutte le opportunità date, finirà comunque, per aumentare tale ignoranza relativamente ad ogni individuo¹⁹. Tuttavia, pur essendo tale non-onniscienza o *ignoranza relativa*, la base di libere, individuali, scelte cognitive, «l'aumento dello stock di conoscenze all'interno della cornice istituzionale data si è verificato lungo tutto il corso della storia, in conseguenza della costante tendenza degli esseri umani ad inventare e ad innovare anche in assenza di incentivi istituzionali²⁰». La scarsa distribuzione della conoscenza, perciò, va distinta dalla sua, sempre crescente, produzione ma, l'eventuale obiezione, per cui le potenzialità cognitive sarebbero già, in

¹⁵ Platone, *Menone* § 97 e *Teeteto*, § 203c; Gettier (1963: 121-23); Boghossian (2006: cap. 2); Marconi (2007: cap. 1).

¹⁶ Bloor (1994: 7).

¹⁷ Guala (2006).

¹⁸ VonHayek (1979: 190). Si veda anche Id. (1937: 35-56), e Id. (1952), come esempio di «spiegazione a mano invisibile» della cognizione umana.

¹⁹ North (2005, trad. it. 103-4).

²⁰ *Ibidem*. pp. 38 e 78.

linea di principio, *infinite* (non essendo, così, delle risorse, ma mere esternalità), sfuma di fronte alla “finitezza relativa” degli orizzonti cognitivi umani. Questo è il dominio proprio dell’economia della conoscenza.

Il realismo *esternista*, è qui inteso come una tesi ontologica di *esistenza* d’una realtà mondana, in linea di principio del tutto autonoma e dai *processi* cognitivi del pensiero, individuale o collettivo, e dai *contenuti* concettuali degli stessi, *compresi i sistemi di riferimento storico-temporali*. “Realismo” non indica, perciò, nel prosieguo, una teoria *epistemica* della verità, ma una tesi in difesa dell’esistenza del *reale*, come *esternalità ontologica*: «Il realismo è la tesi secondo la quale c’è un modo in cui le cose sono che è logicamente indipendente da tutte le rappresentazioni umane. Il realismo non dice come sono le cose, ma solo che c’è un modo in cui esse sono»²¹.

Costruttivismo o *comunitarismo* epistemico indicherà, invece, il “retaggio analitico” proprio d’un orientamento filosofico risalente a Peirce, di filosofi neo-pragmatisti, o “internisti”, come Sellars, Rorty, Brandom, e McDowell²². Seppur diversi, e talvolta discordanti, questi autori si contraddistinguono per un convergente stile di rappresentazione della *realtà*: chi posizionandola all’interno di uno “spazio logico delle ragioni”, appannaggio di una comune dimensione sociale di padronanza linguistica²³, chi facendo leva sul fatto che l’intera vicenda della ricerca umana sia una storia in divenire di diverse comunità epistemiche²⁴, chi situandone la percezione nel dominio del “concettuale *senza confini*”²⁵, chi elaborando la suddetta idea di “spazio delle ragioni” in un sistema neo-hegeliano di tipo inferenzialista²⁶. In tutti questi casi, la realtà è vista, *a là* Peirce, come meta *ideale* di una diacronica, indefinitamente lunga, *impresa storico-collettiva*, dove i liberi sforzi di ogni ricercatore, di epoca in epoca, vedranno infine convergere i propri risultati: «l’autentica origine del concetto di realtà mostra che questo concetto implica essenzialmente la nozione di una COMUNITÀ senza limiti definiti, e capace di un *incremento indefinito di conoscenza*»²⁷.

Mostrando come e perché questi due indirizzi teorici citati rappresentino due tipologie di economia epistemica del tutto incompatibili tra loro, scopo finale di un loro confronto sarà quello di capire quale tra esse possa supportare, alla luce del teorema di Arrow e dei suoi sviluppi, una *funzione di scelta*

²¹ Searle (1995, trad. it. 176).

²² Per una rassegna storiografica sulla *Peirce’s Legacy* si veda Murphy (1990).

²³ Sellars (1957).

²⁴ Rorty (2004).

²⁵ McDowell (1994).

²⁶ Brandom (2000).

²⁷ Peirce (1868, trad. it. 107).

epistemica “a prova di dittature”. Se l’ipotesi di partenza è corretta, trova fondamento il dubbio che, tra le due teorie considerate, quella sedicente “garante” del perdurare di un equilibrio ottimale, tra le scelte *individuali* epistemiche e l’unanime risultato *comunitario* della ricerca, sia la destinata ad avere la peggio.

4. La conoscenza è l’esito di processi di scelta

Un’obiezione di principio a questa teoria potrebbe essere, però, la seguente: pur riconoscendo un ruolo alle preferenze e alle scelte in ogni vera economia della conoscenza, come potrebbe mai dipendere da scelte soggettive e individuali, quell’importante modulo della dimensione conoscitiva, noto come *metodo scientifico*, avente nell’oggettività *impersonale* la propria meta? Non si potrebbe, di contro, usare un esempio di portata così vasta, proprio per dimostrare quanto e come la conoscenza non possa essere assimilata ad esiti di preferenze o scelte, ancorché sociali?

Preferenza e scelta sociale sono due concetti reciprocamente implicati ma ben distinti. Con *preferenza*, in questo caso, s’intende l’attitudine (abbastanza libera dall’inferenza logica²⁸) che ogni individuo *ha* nell’interagire con uno o più panieri di singole opportunità a partire dai propri desideri, credenze e aspettative. Di contro, una *scelta* è quel che, lo stesso individuo, in quanto agente, *fa* a partire da ogni singola, possibile, opzione d’un paniere, agendo con un determinato atteggiamento di preferenza. Ogni scelta implica almeno una preferenza, ma non è vero l’inverso, poiché è possibile non solo concepire, ma anche sperimentare l’esistenza di individui economicamente non paradigmatici, per il classico modello di “ottimalità decisionale” dell’*homo œconomicus*²⁹ i quali, pur disponendo di ogni informazione utile per decisioni razionali e ottimizzanti, date le proprie preferenze, *non scelgono di conseguenza*, senza per questo risultare affetti da strane “sindromi dell’asino di Buridano”, ma solo perché il modello dell’*homo œconomicus* risulta inadeguato per spiegare le loro *performances*³⁰. Invece, *ogni* conoscenza, essendo per ipotesi un contenuto *proposizionale*, consiste sempre in una *credenza* vera e giustificata, ed è perciò il risultato di almeno una preferenza *doxastica*. Anche il metodo sperimentale, perciò, tende a considerare conoscenze solo quelle preferenze doxastiche che, oltre ad essere state vagliate con una griglia di

²⁸ Zajonc (1980: 151-75).

²⁹ Pareto (1899: 145-57).

³⁰ Per i limiti sperimentali dell’*homo œconomicus* si vedano i saggi di Kahneman & McFadden (2005).

giustificazioni, risultano predittivamente affidabili, o plausibilmente vere³¹, offrendo contenuti intenzionali di prezioso, *ergo* economico, interesse teorico, empirico, applicativo. Attribuirsi uno stato conoscitivo coinciderebbe, così, con l'atto intenzionale della scelta epistemica che, in quanto scelta, non è altro che l'esercizio di una preferenza doxastica, orientata da una o più scoperte, riflessioni, critiche, argomentazioni, purché non circolari e conclusive. Già a partire dagli eventi che il ricercatore sceglie di eleggere a fatti salienti, poiché dotati di certe potenzialità d'uso, egli è posto dinanzi a incroci di scelte iniziali ed esiziali, come osservò una volta Poincaré:

Che sia necessario fare una scelta questo è fuor di dubbio; quale che sia la nostra attività, i fatti vanno più svelti di noi, e noi non saremmo mai in grado di acciuffarli. [...] Gli uomini di scienza credono che esiste una gerarchia dei fatti e che fra questi sia possibile operare una scelta sensata³².

La scelta epistemica, perciò, è un fatto comprovabile tanto dall'analisi concettuale quanto dalle riflessioni svolte dagli scienziati stessi. Dunque, la teoria della scelta sociale potrebbe, fertilmente, occuparsene, applicando ad essa risultati come quelli di Arrow.

5. La “possibilità” di Arrow

Il teorema di Arrow e il significato della sua dimostrazione, hanno un proprio antefatto storico nel noto “paradosso sociale” di Condorcet, detto *paradosso della votazione*³³, per il quale l'esito di un voto, contrariamente alle più intuitive assunzioni del senso comune sul rapporto tra scelta ed esito finale, dipende solo dall'ordine della consultazione, e non dalle sommatorie dei voti espressi, per ogni possibile ordinamento di preferenze. Un *ordinamento* è un'operazione logica per cui, da un insieme non ordinato di n elementi, si passa a una *struttura di ordine operazionale n* (o *n -pla*). In questo caso gli elementi sono preferenze³⁴. Per ovviare al paradosso della votazione, Arrow definisce una *funzione di scelta sociale R* :

³¹ «Inferenza, nel senso più ampio, è coestensivo con *deliberata adozione*, sotto ogni aspetto, di un'asserzione come *vera*». Peirce (1901: 499-500).

³² Poincaré (1909, trad. it. 9-10).

³³ De Caritat-Condorcet (1785).

³⁴ Odifreddi (2001: 210-1).

Una funzione di scelta sociale R è un processo o una regola che, in corrispondenza di ogni insieme di ordinamenti individuali R_1, \dots, R_n per situazioni sociali alternative (un ordinamento per ciascun individuo) stabilisce un ordinamento sociale di situazioni sociali alternative R ³⁵.

R specifica, così, un ordinamento sociale su tutti i gruppi di preferenze sociali, per ogni ordinamento preferenziale individuale. Il problema è fare di R una *funzione di scelta pubblica*, che trasformi l'insieme delle libere preferenze individuali in un ordinamento accettabile per *tutta la società*. All'uopo, Arrow individua nei seguenti 5 punti, altrettante ineludibili clausole³⁶:

1 *Non-dittatorialità*: La funzione R non deve seguire l'ordinamento delle preferenze d'un solo individuo o d'un solo sottoinsieme di individui, ignorando nel contempo le altrui preferenze.

2 *Monotonicità, o associazione positiva tra valori individuali e sociali*: Se un individuo cambia il proprio ordinamento preferenziale, optando per una nuova scelta, la funzione R deve anch'essa preferire tale opzione o restare invariata, ma non può assegnare a tale opzione una *preferenza minore*: se un'opzione risulta vincente in una votazione venendo preferita da un numero n di scelte possibili, risulterà vincente anche in ogni votazione in cui otterrà $n+x$ scelte di voto. La transitività delle preferenze individuali, così, ("se O preferisce x , cui preferisce y , cui preferisce z , allora z è preferita da O anche a x ") deve *monotonicamente* essere rispettata anche a livello di scelta sociale.

3 *Non-imposizione o sovranità del votante*: Qualsiasi possibile preferenza sociale deve essere raggiungibile *solo* a partire da un appropriato insieme di preferenze individuali e ogni risultato deve poter essere raggiunto in qualche maniera.

4 *Universalità o dominio non ristretto*: La funzione R deve creare un ordinamento delle preferenze sociali *deterministico*, a-ciclico (non avente valori di partenza ed arrivo coincidenti) e completo, a partire da qualsiasi insieme iniziale di preferenze individuali.

5 *Indipendenza dalle alternative irrilevanti*: Se si limita l'attenzione ad un sottoinsieme di opzioni (poniamo, una coppia), e la funzione di scelta sociale è applicata ad esse soltanto, il risultato deve essere compatibile con il caso in cui la funzione di scelta sociale sia applicata all'intero insieme di alternative possibili. Perciò, il modo in cui un insieme sociale ordina una coppia di stati alternativi x e y dovrà dipendere solamente dalle preferenze individuali relative a *quella* coppia, e non dal modo in cui vengono ordinate le altre irrilevanti alternative come, invece, nel paradosso di Condorcet.

³⁵ Arrow (1951: 25).

³⁶ *Ibidem.* pp. 24-35.

Arrow, ha dimostrato, così, che se il gruppo di soggetti votanti comprende almeno 2 individui, e l'insieme delle alternative possibili almeno 3 opzioni, *non è possibile costruire una funzione di scelta sociale R che soddisfi, al contempo, tutte le clausole (1-5)*³⁷. Esiste, dunque, un singolo votante o gruppo di votanti, chiamato in seguito (1970), da Amartya Sen, "il dittatore", che *da solo* determina il risultato della scelta. Egli è l'*intersezione insiemistica* di tutti i gruppi formati dagli elettori decisivi³⁸. *Rebus sic stantibus*, delle due l'una: o accettiamo il paradosso di Condorcet, (l'esito di ogni scelta sociale dipende dall'ordine in cui vengono espresse le singole preferenze) o, in un sistema che escluda questa ipotesi (e in presenza di almeno 3 possibili opzioni), ogni *insieme decisivo comprende un dittatore*, cioè un insieme di votanti (anche di un solo individuo) che *da solo* determina il risultato del voto. Ambedue le opzioni sono in conflitto con l'idea-base di un sistema di scelta equamente rappresentativo, così come, mutando, rispettivamente, gli elettori attivi e passivi in *consumatori* o in *produttori/distributori*, con l'essenza di un *libero mercato*. Due idee risultanti, perciò, logicamente irrealizzabili.

6. Pareto-efficienza epistemica

Nella teoria economica neo-classica v'è, però, un meccanismo in grado di generare e descrivere predittivamente l'emergere di equilibri autoregolativi che, nel caso del gioco linguistico del coleottero potremmo definire come *microeconomici*: la Pareto-efficienza. Sviluppata da V. Pareto, all'interno della sua già citata teoria dell'*homo œconomicus*, la situazione descritta da questo modello si realizza quando, in un qualsiasi assetto distributivo, l'allocatione delle risorse è tale che non è possibile migliorare la condizione di un soggetto economico senza peggiorare la condizione di un altro³⁹. Nel caso *realista* del gioco del coleottero, detto equilibrio sarà detto "risultato *a somma zero*", dove cioè, *a priori*, è previsto che qualcuno vinca e qualcun altro perda⁴⁰. Se sussistesse, per ipotesi, un *isomorfismo operativo*⁴¹ tra l'algoritmo di equilibrio di un ambiente distributivo paretiano, e un ambiente epistemico dato, ne de-

³⁷ *Ibidem*. p. 62-3; Sen (2002, trad. it. 195-234).

³⁸ *Ibidem*. p. 157.

³⁹ Pareto (1896, trad. it. 51-70).

⁴⁰ VonNeumann & Morgestern (1974).

⁴¹ Un *isomorfismo* è una relazione che permette di *mappare* due strutture complesse una nell'altra, in modo che per, ogni parte di una struttura, c'è una parte corrispondente nell'altra, dove *corrispondente* indica che le due parti giocano ruoli simili nelle loro rispettive strutture. Cfr. Enderton (1972: 6-7).

riverrebbe che l'allocazione di quei contenuti cognitivi di tutti e soli quei soggetti individuali o collettivi formanti l'ambiente, sarebbe tale da non permettere il miglioramento della condizione epistemica di un soggetto, senza causare l'impoverimento epistemico di (almeno) un altro membro dell'ambiente stesso (la situazione già vista nella "favola del coleottero" liberamente tratta da Wittgenstein). Sotto l'ipotesi, discutibile ma ampiamente accettata, che *ogni* conoscenza consista almeno in una proposizione vera⁴² o, nelle menti non-umane più evolute, in suo omologo semiotico, diviene così possibile definire la seguente ipotesi di Pareto-efficienza epistemica (*IP-EE*):

Se in un ambiente epistemico AE composto da almeno due operatori epistemici, il contenuto epistemico di un'ipotesi H corrisponde a T e quello di un'ipotesi I a $\neg T$ allora, se un operatore O "riconosce" che H, giustificando tale ipotesi con la scoperta T, tale "ri-allocazione" andrà a sfavore di I, nel qual caso, un altro operatore O*, non potrà più conoscere I scoprendo $\neg T$ (O non conosce I, ex hypothesis, avendo scoperto T, sotto l'ipotesi H). Il risultato di tale allocazione sarà detto nuovo Pareto-equilibrio epistemico di AE (il simbolo \neg indica la negazione logica di una proposizione).

Infatti, se ogni conoscenza è *almeno* una proposizione vera, e se T è una conoscenza, allora, banalmente, essendo T vera, $\neg T$ è falsa. Quindi, se O* avrà creduto di conoscere $\neg T$, ora non potrà più "ri-conoscerla", perché se $\neg T$ è falsa, a partire dalla ri-allocazione paretiana di T, $\neg T$ non potrà più essere considerata una conoscenza, perdendo *valore epistemico*, non permettendo più di *risparmiare esperienze possibili*. *IP-PE* è una legge di equilibrio⁴³, vale a dire una legge che descrive un comportamento *temporalmente* vincolato: anche se O* riuscisse al tempo t_n a riprendere terreno, dopo il consolidamento del profitto epistemico di O in t_{n-1} , quest'ultimo avrebbe ancora un vantaggio netto tutto da spendere. Nel lungo periodo, tale vantaggio, a meno di rovesci e sorpassi, potrebbe accumularsi. Le plusvalenze così ottenute, potranno essere espresse in termini di affidabilità predittiva di T, seguendo un approccio pragmatista, o dal "valore aletico" della *verità* di T (che potrebbe essere una teoria, ma anche un *corpus* di credenze del senso comune) se, invece, si propendesse per il realismo esternista. Il realismo esternista, infatti, intende *reale* come *indipendente dalla conoscibilità da parte di ogni mente* e ha, perciò, nella *verità*, intesa come corrispondenza strutturale ai fatti, un mero indicatore di *plusvalore epistemico*.

⁴² Williams (2001, cap. 1).

⁴³ Ingrao & Israel (1987: 118-19).

Il falsificazionismo popperiano, come è stato osservato di passaggio da Hayek⁴⁴, farebbe parte, come caso rilevante, di questo secondo modello. Per il realista oggettivo Popper, infatti, una teoria è non-pseudoscientifica nella misura in cui non vada quotandosi come la *più preziosa* nel proprio campo, lasciando, così, “porte aperte” ad una propria, potenziale, deprezzabilità, tramite falsificazione. Derivante dal *fallibilismo* peirceano⁴⁵, questa “concezione libero-scambista della conoscenza” (come si vedrà, non è l’unica) consiste nell’accettare una libera competizione da parte delle teorie concorrenti, ed essere pronti a batterle come ad esserne battuti, restando così (punto ineludibile, per il realismo) aperti all’eventualità di scoperte improbabili o imprevedibili. Equilibri epistemici realisti di tipo paretiano, perciò, si caratterizzano per il fatto che, date le usuali ipotesi sulle preferenze o sull’euristica utilizzata per ottenerli, sia impossibile, *per ogni singolo problema risolto*, avere migliori livelli conoscitivi individuali, o una migliore distribuzione di altra conoscenza⁴⁶. In tal modo, applicando l’equilibrio paretiano alla situazione in cui si trovano dei competitori o dei cooperatori epistemici, l’*IP-EE* farà sì che, trovato l’equilibrio tra le conoscenze individuali dell’ambiente dato, il conoscitore che vorrà essere *più ricco* dovrà, necessariamente: falsificare le teorie o le ipotesi degli altri, impoverendoli *direttamente*; o pervenire per primo a informazioni utili anche ad altri a sviluppare conoscenze; o ostacolare (slealmente) la capacità di altri competitori epistemici, sempre all’interno dello stesso *paniere* euristico, impoverendoli, così, *indirettamente*.

7. L’equilibrio epistemico costruttivista

È a Wilfrid Sellars, però che spetta il merito di aver definito per primo un tipo di ambiente razionale per lo scambio cognitivo del tutto simile, all’equivalente epistemologico di un mercato liberal-paretiano. Il concetto in questione, illustrato in alcune famose *Lectures* del 1956⁴⁷ prende il nome di “spazio logico delle ragioni” o, per usare la terminologia sellarsiana, di *logical space of warrants* (*garanzie*). In esso, la giustificazione epistemica è definita da Sellars come un «*dare e chiedere ragioni*» il che è significativamente isomorfo ad un sistema di *domanda e offerta di ragioni*: un mercato, in più, che essendo *logico*, può anche essere di tipo *impersonale*⁴⁸. A differenza del modello di

⁴⁴ Popper (1969: 11); VonHayek (1968: 9-23).

⁴⁵ Peirce (1897: 51-3).

⁴⁶ Pareto (1896, trad. it. 61-70).

⁴⁷ Sellars (1956, §36).

⁴⁸ *Ibidem.* § 10.

Popper, però, il modello di libero scambio delle ragioni (o garanzie) razionali di Sellars, si caratterizza subito per il suo pragmatismo. Nell'argomentare, infatti, a favore dello *space of warrants*, Sellars assegna un ruolo di primissimo piano alla confutazione del cosiddetto *mito del dato*⁴⁹, attingendo, così, a piene mani, alla critica pragmatista (peirceana) dell'intuizione cartesiana e del *primum cognitum* immediato, che ha come obiettivo la dimostrazione della natura integralmente mediata in modo *inferenziale e concettuale* della conoscenza⁵⁰. Il corollario paretiano dell'economia dello *space of warrants* potrebbe essere che, per incrementare la propria produttività i membri di una "comunità di conoscitori" partecipino al sistema di dare-avere ragioni, cooperando, così, fino al punto detto di equilibrio *Pareto*-efficiente, oltre il quale "ricavi" ulteriori costeranno, necessariamente, la messa in atto di *giochi non cooperativi*, tra le diverse agenzie conoscitive considerate (scienziati, ricercatori, gruppi di lavoro, ecc.), e le loro teorie. In questo caso, proprio questi giochi, anziché a *somma zero*, come nel caso dell'entomologo *realista*, potrebbero essere a *somma positiva*, tendenti, cioè, ad un equilibrio di Cournot-Nash, (dove, sempre *a priori*, è raggiungibile un accordo tra le parti)⁵¹. Se in un mercato concorrenziale, un paniere di risorse vedrà riflettere nei loro prezzi la loro scarsità in natura, secondo il principio della domanda e dell'offerta, allora, in una condizione di libero accesso alle informazioni, il valore di un aggregato di conoscenze sarà più alto laddove la scarsità *distributiva* delle stesse in un dato ambiente epistemico, sarà valutabile per mezzo di *inferenze verso le spiegazioni migliori*, in grado di fornire credenze predittivamente affidabili nell'ambiente stesso. In una *libera concorrenza perfetta*, se il bene è scarso, e se quindi la domanda è superiore all'offerta, il prezzo sarà elevato, e viceversa. *Ipsa modo*, se in un ambiente epistemico paretiano comunitario i contenuti in questione sono di rara disponibilità, il valore che i competitori saranno disposti ad attribuire a tali contenuti, sarà più alto. Se un mercato paretiano permette un'efficienza completa, ponendo i beni in mano a chi è disposto a pagare di più, e quindi a "saperli sfruttare meglio", allora, per isomorfismo una siffatta concorrenza epistemica comunitaria garantirebbe un'efficienza completa, rendendo le ipotesi più accurate un'acquisizione di chi sia disposto a lavorare di più (analizzando, argomentando, comunicando) per ottimizzare qualità e quantità della conoscenza, mostrandosi, così, più in grado di saperla utilizzare. *Fin qui*, gli ipotetici isomorfismi proposti sembrano funzionare abbastanza bene per ambedue i modelli.

⁴⁹ Sviluppata da McDowell (1994).

⁵⁰ Rorty introduzione italiana a Sellars (1956); Peirce (1868, trad. it. 317-32).

⁵¹ Nash (1950a 48-9); Id. (1950b: 155-62).

8. L'unanimità epistemica paretiana è impossibile

Ampliando, però, il risultato d'impossibilità di Arrow, Sen⁵² ha dimostrato che un sistema che voglia imporre, nel contempo, efficienza paretiana e libertà (intesa come un dominio dove le sole preferenze dell'individuo determinano la sua scelta), può sviluppare situazioni in cui, al più, un *solo individuo* ha piena garanzia di una libera scelta. In questo modo è dimostrata l'inapplicabilità dell'efficienza allocativa paretiana a modelli (neoliberisti) d'ordine spontaneo a "mano invisibile"⁵³. Ma, se questi sviluppi del risultato di Arrow fossero applicabili anche al dominio d'una funzione pubblica di scelta *qualsiasi*, allora, sotto l'ipotesi di isomorfismo qui adottata, e provata l'esistenza di funzioni di scelta *epistemiche*, si può già ipotizzare la dimostrabilità d'un limite alla *libera* possibilità distributiva anche nelle pareto-efficienze epistemiche. Incluse quelle, realiste o comunitariste, qui prese in esame.

Secondo la ri-formulazione del teorema sviluppata da Sen, la clausola 2 di monotonicità di Arrow, infatti, può essere rimpiazzata proprio dall'unanimità paretiana: se ogni singolo individuo preferisce una certa scelta *x* all'opzione *y*, allora *x* deve essere preferita a *y* anche dall'intera funzione di scelta sociale, e così via. Quest'ultima formulazione è più restrittiva, in quanto ipotizza che le clausole 2 e 5 (monotonicità e *indipendenza dalle alternative irrilevanti*) dell'enunciato originale di Arrow, implichino, assieme, anche l'efficienza paretiana⁵⁴. La definizione paretiana di libertà, infatti, riflettendo quella più tipica dell'individualismo liberale, accettata da Sen, prevede che un individuo possieda una *sfera personale riconosciuta*, in cui le sue preferenze siano rilevanti nella determinazione delle preferenze sociali. Si assuma allora che la scelta epistemica della comunità appartenga alle sfere personali riconosciute di agenti epistemiche particolari, *gli operatori cognitivi*. Per ipotesi, essi saranno esempi di *homo æpistemicus*, ossia di una tipologia epistemicamente isomorfa all'*homo æconomicus* di Pareto: puri soggetti teorici e non necessariamente reali, rappresentabili nella teoria da una propensione continua alla produttività marginale della conoscenza, incarnati nella storia da scienziati naturali o sociali o ricercatori di qualsiasi campo, *purché dotati di una sfera personale socialmente nota*, che li renda operatori *pubblicamente* epistemiche; pena la ricaduta in situazioni estreme, come quella dei "proprietari di coleotteri".

⁵² Sen (1970: 152-57); e Id. (1971: 1406-07).

⁵³ Sen (1970: 283-286).

⁵⁴ *Ibidem*, pp. 199-203.

Stabilita questa possibilità anche per un singolo operatore epistemico, si potrebbe definire, per induzione matematica, un principio paretiano per l'efficienza allocativa epistemica, affermando che se ciascun *homo æpistemicus* preferisce un ordinamento teorico T ad un altro ordinamento T^* , allora T può essere socialmente preferito a T^* . Siffatta funzione di scelta epistemica determinerebbe delle preferenze cognitive complete e coerenti, definite sull'insieme degli ordinamenti cognitivi alternativi, per ogni insieme di preferenze, fissando, così, un ordinamento R per operatore cognitivo. Tale funzione avrebbe un "dominio non ristretto", e sarebbe perciò *libera*, se e solo se risultasse compatibile con ogni n -esimo gruppo di ordinamenti possibili di n preferenze individuali. *Ebbene: ciò non è possibile*. Per quel che si è già rilevato, infatti, si può enunciare un *principio di impossibilità dell'agente (cognitivo) liberal-paretiano*, che non è altro che l'applicazione epistemica dell'omonimo principio logicamente derivato da Sen dal risultato di Arrow:

Per ogni operatore (cognitivo) e per ogni ordinamento di preferenze (doxastiche) non può esistere una funzione di scelta (epistemica) che soddisfi, nello stesso tempo, l'esistenza di un dominio non ristretto, il principio d'efficienza paretiana, e la clausola di libertà minimale di (di conoscenza)⁵⁵.

Il senso più forte di questo "lemma" è il seguente: se si volesse godere realmente di un dominio non-ristretto di scelta nell'ambito epistemico e, simultaneamente, del rispetto di una condizione di *unanimità paretiana* (l'efficienza allocativa della conoscenza), allora è necessario, *logicamente*, fare a meno della *libertà* di scelta epistemica. Tra tutti gli operatori, *al più*, solo un operatore, che si chiamerà il *dittatore epistemico* avrà pieno potere di imporre la propria visione teorica al resto degli operatori dell'ambiente epistemico della sua comunità razionale di appartenenza.

9. Confronto finale. La convenienza del realismo

Quanto è realmente bizzarro o paradossale questo risultato di non-unanimità paretiana della scelta epistemica? Quale "economia della conoscenza", tra realismo e costruttivismo comunitarista, ha più da perdere, *filosoficamente*, dalla scoperta che, come per la scelta politico-economica, anche per quella epistemica, funzioni di scelta pubblica non-dittatoriali e liberal-paretiane sono logicamente impossibili?

⁵⁵ Sen (1970: 237-9).

Invero, per un realista, che si tratti di *scegliere* tra l'ipotesi astronomica per cui il sistema solare ha 9 pianeti, e quelle per cui ne avrebbe 12 o 8, o tra i preparati omeopatici e le più avanzate terapie antivirali per un'efficace cura dell'AIDS, non sarà certo un'unanime "funzione pubblica" a determinare una libera scelta *realmente* epistemica, ma le *reali* scoperte di *nuove* proprietà dei corpi celesti o di tragici reperti epidemiologici. L'impossibilità di Arrow-Sen è, in ogni caso, innocua per il realista, perché se egli volesse "scegliere epistemicamente", tra Bellarmino, Galileo e Keplero, non conterebbe nulla per lui (secondo un noto esempio), se il primo avesse o no il diritto di opporre il proprio fondamentalismo alla teoria di Galileo. Altrimenti, *questa* scelta sarebbe, come pensava Rorty, alla stregua della «questione [...] tra Kerensky e Lenin, o tra la Royal Academy [...] e Bloomsbury⁵⁶», una mera questione di gusti. Mentre, di contro, tra tutti e tre, Keplero e Galileo furono i soli *realmente* in fertile scambio epistemico tra loro. Che la funzione di scelta partiana crei equilibri di ordinamenti preferenziali determinanti, a loro volta, consensi tanto unanimi quanto dispotici è, come si è visto, una ben dimostrata verità logica. Quel che, perciò, si può paventare è che *nicchie di mercato* epistemico (le cui preferenze più sensibili all'unanimità culturale potrebbero decidere della validità stessa d'una teoria), finiscano, col tempo, per affermarsi. Aggregandosi tramite intersezione insiemistica, queste nicchie potrebbero, a propria volta, imporre dittature epistemiche *à la* Sen, sotto forma di *opinion-governance*, mass-mediatiche, ideologiche, accademiche o religiose. Questo, forse, è il prezzo relativistico⁵⁷, che i comunitaristi *devono* pagare, per affermare che l'oggettività «non consiste nell'obbedienza a standard che gli epistemologi auspicherebbero di codificare ma, piuttosto, in una partecipazione di dare-avere a un progetto sociale cooperativo⁵⁸» o che essa, come disse Peirce, è «ciò in cui, presto o tardi, alla fine si risolveranno le informazioni e il ragionamento», poiché «l'opinione sulla quale, fatalmente tutti coloro che indagano si troveranno in definitiva d'accordo è ciò che intendiamo per verità, e l'oggetto rappresentato in questa opinione è il reale⁵⁹». Questa visione si nutre, infatti, d'un sincero, e sperequato, ottimismo economico verso la dispersione *distributiva* della conoscenza (crescente, per Hayek in *proporzione diretta* ai tentativi di ridurla), e verso un suo, ancor più improbabile, riequilibrio.

Il realista, pertanto, accetterà *sempre* la clausola 1 del teorema di Arrow, rifiutando *scientificamente e culturalmente* ogni forma istituzionale o non di

⁵⁶ Rorty (1979, trad. it. 251).

⁵⁷ Ben denunciato da Boghossian (2006, cap. 6).

⁵⁸ Rorty introduzione italiana a Sellars (1956: 4).

⁵⁹ Peirce (1878, trad. it. 391).

monopolio epistemico, perché, da Galileo in poi, rifiutare i *principia auctoritatis* è tra i “primi comandamenti” del realismo. Questo rifiuto delle dittature scientifico-culturali, non negoziabile dal realista post-galileiano (diversamente da quello scolastico), permette di accogliere, ma solo *provvisoriamente*, anche le clausole 2, 3, 4 di Arrow, nell’analisi di ogni singolo problema. Adottando un provvisorio e settoriale *pluralismo*, implicito in processi euristici di competizione per la scoperta e l’innovazione, contro un *relativismo* epistemico perenne, il realista rende, così, socialmente trasparente alle comunità di ricerca l’esercizio metodologico della sua funzione di scelta, a patto che dare, prima o poi, risposte fattive ad ogni singolo problema rimanga il vero *fine* della ricerca stessa.

Il realista non potrà accettare, invece, la clausola 5 (*indipendenza dalle alternative irrilevanti*) di Arrow, per non trascurare l’ordine *reale* in cui gli eventi da selezionare come fatti salienti vanno presentandosi alla sua ricerca. Nessuna opzione, infatti, risulta irrilevante, per chi, in stato di *serendipity* (apertura euristica all’imprevisto) è pronto a fare scoperte inattese⁶⁰; pertanto il realista *non* riterrà un paradosso, per la scelta epistemica, il risultato di Condorcet. Il ritmo diacronico delle scoperte, anzi, può spiegare perché i risultati della ricerca, di là da ogni finalismo storico, abbiano sempre dietro di sé un “sentiero storico-causale” o *path-dependence*⁶¹, utile per capire come essi siano apparsi proprio in dati momenti e modi, nonostante il “cieco” avanzare per prove ed errori della ricerca. Diversamente da chi vede nella realtà qualcosa di «non necessariamente indipendente dal *pensiero in generale*, ma solo da ciò che voi, o io, o un qualsivoglia numero *finito* di uomini può pensare di essa⁶²», il realismo oggettivo non può promettere scelte epistemiche che siano finali e al contempo libere come quelle di partenza, perché ad essere autonoma da qualsiasi preferenza doxastica *in generale* è solo la realtà.

Il realismo è, in effetti, una tesi di “auto-inerzia causale”. Come il moto inerziale di un corpo può cambiare solo per mezzo di altre forze esterne, così un *fatto reale* può ricollocarsi nello “spazio logico delle ragioni” solo per l’intervento *causale* di altri, nuovi, fatti reali. Le ragioni concettuali non possono, simmetricamente, essere parimenti autonome dalla realtà esterna, almeno non nella conoscenza *reale*, poiché anch’essa, in quanto *reale*, è causalmente inerte solo a fatti reali; altrimenti non sarebbe neppure una *risorsa*.

Le leggi naturali; il *fatto* che esistano da prima della comparsa del loro, forse unico, descrittore intelligente, che n’è, in gran parte, un frutto evolutivo;

⁶⁰ Merton & Barber (1992).

⁶¹ North (2005, trad. it. 33-44, 109).

⁶² Peirce (1878: 391).

le loro forme matematiche: tutte queste “cose”, per il realista sono esternalità ontologiche, e sono perciò dei *trade off*, libere, cioè, dalle leggi e dai costi del mercato di un “dare e chiedere” ragioni; fino alla scoperta di *nuove*, opposte, esternalità.

Per i comunitaristi, invece, ciò che garantisce alla conoscenza di essere esercitata in base ad una sorta di *laissez-faire* epistemico, è ciò che J.McDowell ha ben definito (criticandola) la “certezza della fine della ricerca”, vale a dire: «uno stato di cose in cui la ricerca, compresa la meta-ricerca sulle credenziali di ciò stesso che attualmente passa per *ricerca*, non sarebbe più necessaria⁶³». Questo *limite ideale* garantirebbe il valore dell’investimento comune di tutte le generazioni, nella *free pursuit of Knowledge*, poiché la “realtà costruttivistica” è autonoma solo dai processi mentali d’ogni individuo, ma non dalla pensabilità concettuale d’una COMUNITÀ ideale e storica. Affinità, queste, con l’idealismo hegeliano che, per Sellars e lo stesso McDowell, finiscono per essere programmatiche⁶⁴. La funzione di scelta epistemica comunitarista *deve*, così, attuare *sempre* ogni clausola (1-5) di Arrow, perché soddisfare un processo liberal-comunitario di ricerca è l’essenza della sua agenda filosofica: un’idea fiduciosa nelle virtù autoregolatrici del mercato epistemico, che *ipso facto* non ammette alcuna esternalità e, per il quale, “realtà” è solo un concetto *inter alia*, totalmente parte del *trading* (logico) dello spazio delle ragioni. La fine della ricerca, come una sorta di “mano invisibile hayekiana”, dovrebbe avere, così, l’*omologo* dell’equilibrio ottimale della COMUNITÀ di ricerca, nell’esito finale dell’unanime scelta epistemica.

Come previsto da Sen, però, tale unanimità è impossibile. Perciò, il comunitarista coerente non ha molte scelte: qualcosa, nel suo “mercato”, dovrà fungere, prima o poi, da *limite reale* (e perciò da “esternalità”) allo spazio logico o storico del dare-chiedere ragioni. L’*homo epistemicus* comunitarista è, altrimenti, destinato ad essere un inconsapevole – fintantoché internista – suddito del dittatore del settore di ricerca di turno, oppure, viceversa, il dittatore epistemico *a là* Sen dello stesso settore. Quest’ultimo caso avrebbe davvero del paradossale per questo *homo epistemicus*, dal momento che, da un lato, in quanto internista, egli negherà l’autonomia della realtà dal “pensiero in generale” della sua COMUNITÀ, seguendo Peirce, Sellars e Rorty, dall’altro, e *ideo tempore*, egli, in quanto dittatore epistemico, non potrà non comportarsi come l’oracolo di quel che è stato chiamato, in una nota parodia dell’esternismo, il punto di vista dell’occhio di Dio⁶⁵.

⁶³ McDowell (1994, L.2, §8).

⁶⁴ *Ibidem*. L.4; Sellars (1956, §§ 1, 20).

⁶⁵ Putnam (1981, capitolo 1).

La morale della favola epistemologica del coleottero vede, così, l'operatore realista (curioso di sapere cos'ha realmente "in scatola") accettare un pluralismo epistemico provvisorio, che vigerà fino all'apertura delle scatole, e il comunitarista (desideroso di giocare *ad libitum* al gioco linguistico del coleottero) favorire un *vincitore fisso* del gioco, nel tentativo, utopico, di far vincere tutti.

L'economia realista si rivela, così, e *a priori*, la più coerente e conveniente, nonché la più "in attivo" delle due, perché stimando l'esistenza di eternalità oggettive, diversamente dalla rivale, non avrà mai bisogno, "in barba al rasoio" di Ockam, di *moltiplicare i suoi enti teorici oltre necessità*, postulando "limiti ideali", o paventando "Miti del Dato". «Quando ho esaurito le giustificazioni arrivo allo strato di roccia – scrive Wittgenstein – e la mia vanga si piega. Allora sono disposto a dire: "Ecco, agisco proprio così"»⁶⁶. Ai *limiti ideali*, uno "strato di roccia" sarà sempre preferibile come "punto di arrivo" di una ricerca *criticamente libera*.

Bibliografia

- Arrow, K. J. (1951), *Social Choices and Individual Values*, New York (Scelte sociali e valori individuali, trad. it., Milano, Etas, 2003).
- Arrow, K. J. (1984), *The Economics of Information*, Cambridge (MA), Harvard.
- Bacon, F. (1620) *Novum Organum*, London ("Aforismi sull'interpretazione della natura", in *Novum Organum I*, trad. it., Roma-Bari, Laterza, 1968, 72–3).
- Bloor, D. (1994), *La dimensione sociale della conoscenza*, Milano, Cortina, 7.
- Boghossian, P. (2006), "Fear of Knowledge", in *Against Relativism and Constructivism*, Oxford, Oxford University Press, Cap. 2.
- Brandom, R. (2000), *Articulating Reasons: An Introduction to Inferentialism*, Cambridge (MA), Harvard University Press.

⁶⁶ Wittgenstein (1953, § 217).

- De Caritat-Condorcet, N. (1785), *Essay sur l'application del'analyse à la probabilité de la decision rendues à la pluralite des voix*, Paris.
- Damasio, A. R. (1994), *Descartes' Error: Emotion, Reason, and the Human Brain*, New York, Grosset/Putnam.
- Enderton, H. (1972), *A Mathematical Introduction to Logic*, Orlando, Academic Press.
- Foray, D. (2000), *L'èconomie de la connaissance*, Paris, La Découverte.
- Galeotti, A. E. (1988), "Individuale e collettivo", in *L'individualismo metodologico nella teoria politica*, Milano, Mondadori.
- Gettier, E. L. (1963), "Is Justified True Belief Knowledge?" in *Analysis*, 23, 121-23.
- Gibson, J. J. (1979), *An Echological Approach to the Visual Perception*, New York, Accademic Press.
- Goldman, A. I. (1970), "A Theory of Action", Princeton (NJ), e in *Philosophical Applications of Cognitive Science*, Boulder, 1993 (Applicazioni filosofiche della scienza cognitiva, trad. it. di M. Riccucci, Bologna, Il Mulino, 1996).
- Guala, F. (2006), *Filosofia dell'economia*, Bologna, Il Mulino, 2006.
- Hayek, F. A. von (1937), "Economics and Knowledge", in *Individualism and Economic Order*, Chicago, University of Chicago Press, 1948.
- Hayek, F. A. von (1952), *The Sensory Order. An Inquiry into the Foundations of Theoretical Psychology*, Chicago, University of Chicago Press.
- Hayek, F. A. von (1968), "Competition as a Discovery Procedure" in *Quarterly Journal of Australian Economy*, 5, 3.
- Hayek, F. A. von (1979) *Law, Legislation and Liberty*, London, The Free Press of Glencoe, p. 190.
- Hofstadter, D. (1979), *Gödel, Escher, Bach. Un'eterna ghirlanda brillante*, Milano, Adelphi, 49.

- Ingrao, B., Israel, G. (1987), *La mano invisibile; l'equilibrio economico nella storia della scienza*, Roma-Bari, Laterza, pp. 118-19.
- Jablonka, E., Lamb, M. (2005), *Evolution in Four Dimension; Behavioral and Symbolic Variation in the History of Life*, Boston, MIT Press (L'evoluzione a quattro dimensioni; variazione genetica, epigenetica, comportamentale e simbolica, nella storia della vita, trad. it. N. Colombi, Torino, Utet, 2007).
- Kahneman, D. McFadden, D. Smith V. (2005), *Critica della ragione economica*, di A. Motterlini (a cura di), M. Piattelli-Palmarini, Milano, Il Saggiatore.
- Kusch, M. (1995), "Psychologism: a case study", in the *Sociology of Philosophical Knowledge*, London, Routledge.
- Mach, E. (1883), *La meccanica nel suo sviluppo storico-critico* (trad. it. di A. D'Elia, Torino, Bollati-Boringhieri, 1977).
- May, K. (1952), "A Set of Independent, Necessary and Sufficient Conditions for Simple Majority Decisions" in *Econometrica*, 20, pp. 680-684.
- Marconi, D. (2007), *Per la verità; filosofia e relativismi*, Torino, Einaudi.
- McDowell, J. (1994), *Mind and World*, Harvard, Harvard University Press., Lect 1-3.
- Merton, R. K., Barber, E. G. (1992), *The Travels and Adventures of Serendipity. A Study in Historical Semantics and Sociology of Science*, Princeton, Princeton University Press.
- Murphy, J. P. (1990), *From Peirce to Davidson*, Boulder, Westview Press.
- Nash, J. F. (1950a), "Equilibrium Point in N-person Games" in *Proceedings of the USA National Academy*, 36, pp. 48-9,
- Nash, J. F. (1950b), "The Bargaining Problem", *Econometrica*, 18, pp.155-62.
- North, D. C. (2005), *Understanding the Process of Economic Change*, Princeton, Princeton University Press (*Capire il processo di cambiamento economico*, trad. it. di E. Felice, Bologna, Il Mulino, 2006).

- Odifreddi, P. (2001), *C'era una volta un paradosso. Storie di illusioni e verità rovesciate*, Torino, Einaudi, pp. 210-11.
- Pareto, V. (1896), "La corbe de la répartition de la richesse" in *Recueil publi par la Faculté de Droit de l'Université de Lausanne à l'occasion dell'Exposition Nationale de 1896*, Lausanne, pp. 373-87 ("La curva di ripartizione della ricchezza", trad. it. in *Le diseguaglianze economiche*, Torino, Bollati-Boringhieri, 1995).
- Pareto, V. (1897), *Cours d'économie politique*, 2 vv. Lausanne (*Corso di economia politica*, trad. it., Torino, 1949).
- Pareto, V. (1899), "I Problemi della Sociologia" in *Rivista Italiana di Sociologia*, Marzo, pp. 145-57.
- Peirce, C. S. (1868), *Collected Papers*, 5, pp. 213-63 ("Questioni concernenti certe pretese facoltà umane", trad. it. in *Opere*, di A. Massimo Bonfantini (a cura di) e G. Proni, Milano, Bompiani Il Pensiero Occidentale, 2003, pp. 317-32).
- Peirce, C. S. (1868), "Pensiero-segno-uomo", trad. it., in *Opere*, di A. Massimo Bonfantini (a cura di) e G. Proni, Milano, Bompiani Il Pensiero Occidentale, Milano, 2003, p. 107
- Peirce, C. S. (1878), *Collected Papers*, 5, p. 407 ("Come rendere chiare le nostre idee", trad. it. in *Opere*, di A. Massimo Bonfantini (a cura di) e G. Proni, Milano, Bompiani Il Pensiero Occidentale, Milano, p. 391).
- Peirce, C. S. (1897), *Collected Papers*, 7, pp. 186, 1.8-1.14 ("Autopresentazione", trad. it., in *Opere*, di A. Massimo Bonfantini (a cura di) e G. Proni, Milano, Bompiani Il Pensiero Occidentale, Milano, 2003, pp. 51-3).
- Peirce, C. S. (1901), *Collected Papers*, 7, p. 186 ("Abduzione e storia", trad. it., in *Opere*, di A. Massimo Bonfantini (a cura di) e G. Proni, Milano, Bompiani Il Pensiero Occidentale, pp. 485-538).
- Peirce, C. S. (1905), *Collected Papers* 5, p. 412 ("Che cos'è il Pragmatismo", trad. it., in *Opere*, di A. Massimo Bonfantini (a cura di) e G. Proni, Milano, Bompiani Il Pensiero Occidentale, pp. 399-414).

- Platone, “Menone” §97 e “Teeteto”, §203c, in *Opere*, Roma-Bari, Laterza, 1924.
- Poincaré, J. H. (1909), “The Choice of Facts”, in *The Monist*, 19, pp. 232–39, 232-33 (“La scelta dei fatti” trad. it. in *Scienza e metodo*, Torino, Einaudi, 1997).
- Pomian, K. (1999), *Sur l’histoire*, Paris, Gallimard (*Che cos’è la storia*, trad. it. di M. Di Sario, Milano, Bruno Mondadori, 2001)
- Popper, K. (1969), *Conjectures and Refutations*, London, Routledge (*Congetture e confutazioni*, trad. it. di G. Pancaldi, Bologna, Il Mulino, 1972, Capp.1, 3, 11).
- Putnam, H. (1981), *Reason, Truth and History*, Cambridge (MA), Cambridge University Press.
- Rifkin, J. (2000), *The Age of Access*, New York, Tarcher.
- Robbins, L. (1932), *An Essay on the Nature and Significance of Economic Science*, London, 1932-1935 (Saggio sulla natura e il significato della scienza economica, trad. it., Torino, UTET, 1953).
- Roncaglia, A. (2005), *Il mito della mano invisibile*, Roma-Bari, Laterza, cap 1.
- Rorty, R. (1997) “Introduction”, in W. Sellars, *Empiricism and Philosophy of Mind*, Cambridge (MA), Harvard University Press, p. 1-15.
- Rorty, R. (1979), *Philosophy and the Mirror of Nature*, Princeton (NJ), Princeton University Press.
- Samuelson, P. (1974), *Scientific American*, Ott., p. 120.
- Searle, J. R. (1983), *Intentionality*, Cambridge (MA), Cambridge University Press (*Dell’intenzionalità*, trad. it. di D. Barbieri, Milano, Bompiani, 1985).
- Searle, J. R. 2006, *The Construction of Social Reality*, Cambridge (MA), Cambridge University Press, 1995 (*La costruzione della realtà sociale*, trad. it. di A. Bosco, Torino, Einaudi, 2006).

- Sellars, W. (1956), *Empiricism and Philosophy of Mind*, con una "Introduction" di R. Rorty, e uno Study Guide di R. Brandom, Cambridge (MA), University Press, 1997 (*Empirismo e filosofia della mente*, trad. it. di E. Sacchi, Torino, Einaudi, 2004, Lez. 1).
- Sen, A. K., "The Impossibility of a Paretian Liberal", *Journal of Political Economy*, 78, 1970, 152-57 ("L'impossibilità di un liberale paretiano", *Scelta, benessere, equità*, trad. it. di F. Delbono, G. Gozzi, V. Denicolò, Bologna, Il Mulino 2006).
- Sen, A. K. (1971), "The Impossibility of a Paretian Liberal: Reply", in *Journal of Political Economy*, 79, pp. 1406-07.
- Sen, A. K. (2002), "Rationality and Social Choice" in *Rationality and Freedom*, Cambridge (MA), University Press ("Razionalità e scelta sociale" in *Razionalità e libertà*, Bologna, Il Mulino, 2005, 195-234).
- Simon, H. (1984), *La ragione nelle vicende umane*, Bologna, Il Mulino.
- Smith, A. (1776), *An Inquiry on Nature and Causes of the Wealth of Nations*, London, in R. H. Campbell e A. S. Skinner (a cura di), Oxford, University Press (*La Ricchezza delle Nazioni*, trad. it. di F. Bartoli, C. Camporesi, S. Caruso, Roma, Newton-Compton, 1995, Cap II, § 1).
- Vaselli, S. (2007), "Due varietà di realismo" in *Prospettive filosofiche: il realismo*, Napoli, IISF, pp. 115-127.
- Vassallo, N. (2003), *Teoria della conoscenza*, Roma-Bari, Laterza.
- Neumann, Von J., Morgestern, O. (1947), *Theory of Games and Economic Behaviour*, Princeton, Princeton University Press.
- Williams, M. (2001), *Problems of Knowledge*, Princeton, Princeton University Press.
- Wittgenstein, L. (1953), *Philosophical Investigations*, Oxford, 1, § 293.
- Zajonc, R. B. (1980), "Feeling and Thinking: Preferences need no Inferences" in *American Psychologist*, 35.

Tra pensare e fare Tommaso d'Aquino e Duns Scoto a proposito dell'irrazionalità pratica

Antonio Petagine
Università de Fribourg (SUISSE)
antonio.petagine@unifr.ch

Perché il possesso della ragione non ci dà la garanzia di compiere sempre le azioni migliori? Si aggiunga poi l'impressione – quasi un luogo comune – secondo cui gli uomini sembrano sollecitati ad agire dai propri gusti, dalle emozioni o dai desideri, ben più che da valutazioni prettamente razionali. Come spiegare dunque il fenomeno dell'*akrasia*, ossia il fatto, già ben messo in luce da Aristotele nell'*Etica Nicomachea* e da san Paolo nella *Lettera ai Romani*, secondo il quale gli uomini possono agire intenzionalmente contro ciò che pur ritengono più razionale e più giusto¹?

In un noto studio, William Charlton ha sostenuto che, fino a Gilbert Ryle, il problema aristotelico dell'*akrasia* fosse sostanzialmente scomparso nella filosofia occidentale, a causa della nozione «giudeo-cristiana» di volontà². Sebbene questa tesi non sia esatta dal punto vista storico³, essa rivela comunque un fatto reale: dopo il *De libero arbitrio* di Agostino, i pensatori medievali e moderni hanno abitualmente attribuito alla *volontà* il ruolo fondamentale, nel produrre l'azione⁴. La «scoperta della volontà» permette di spiegare in modo davvero convincente lo scarto tra pensare il bene e farlo, tra vedere ciò che è migliore e metterlo in pratica: infatti cogliere il bene non basta, perché dobbiamo anche volerlo e deciderci per esso. È l'esigenza di

¹ Cfr. Aristotele, *Etica Nicomachea*, I, 7, c. 2, 1145b22-30; *Rm*, 7, 15-24. Si ricordi anche l'adagio ovidiano «uideor meliora proboque, deteriora sequor» (Ovidio, *Metamorfosi*, VII, 20-21, p. 10).

² Cfr. Charlton (1988).

³ Si vedano, al riguardo, le giuste osservazioni di Saarinen (1994: 1-4).

⁴ Si veda in particolare Agostino, *De libero arbitrio*, III, c. 1, nn. 13-14, p. 276.

quest'atto di libera decisione che rende non scontata la razionalità dell'azione. Essa resta sempre, infatti, sotto la nostra responsabilità.

Uno spunto potente allo sviluppo di questa prospettiva è certamente fornito dal contesto culturale cristiano. Per il cristiano, infatti, il rapporto tra Dio e l'uomo è fondato sull'amore: la creazione e la rivelazione sono atti di donazione gratuita da parte di Dio, ossia una manifestazione primigenia d'amore a cui l'uomo non deve fare altro che corrispondere, con le proprie azioni. Ora, la corrispondenza d'amore, per essere genuina, non può che essere libera: infatti, non si può costringere qualcuno ad amare. Chi concede all'altro la capacità di dire a tutti gli effetti di «sì», gli concede ad un tempo la drammatica possibilità di dire di «no», anche in modo irrevocabile. Nella prospettiva cristiana, questo vale, come vedremo, non solo nel caso dell'uomo, ma anche in quello dell'angelo.

All'interno del panorama scolastico, le posizioni di Tommaso d'Aquino e quelle di Giovanni Duns Scoto hanno goduto di una particolare autorevolezza e hanno segnato in modo decisivo il dibattito filosofico dei secoli successivi, circa la natura dell'azione umana e i dinamismi che essa sottende. Per le ragioni che stiamo per mettere in luce, è stato facile etichettare la posizione di Tommaso come «intellettualismo» e quella di Duns Scoto come «volontarismo». Eppure, come vedremo, proprio il tema dell'irrazionalità pratica si rivela come un campo in cui l'adeguatezza di queste etichette viene messa alla prova.

1. L'«intellettualismo» di Tommaso d'Aquino

1.1. L'uomo come essere «soltanto» razionale

Secondo Tommaso d'Aquino, la possibilità di non agire sempre nel modo più razionale si radica nel fatto stesso di essere dotati di ragione. Infatti, la «ragione» (*ratio*) è quella forma di intelligenza che appare contrassegnata, tanto in ambito speculativo, quanto in ambito pratico, da un'originaria e inaggirabile indeterminatezza. Aderendo senza esitazioni a quanto diceva Aristotele nel *De anima*, Tommaso sottolinea che l'intelletto umano è come una *tabula rasa*, su cui nulla è scritto⁵. Ciò significa che l'uomo parte sempre cognitivamente da zero e viene a conoscenza della realtà solo progressivamente e con

⁵ Cfr. Aristotele, *L'anima*, l. 3, c. 5, 430a10-11, 18-19. Ispirandosi ad Averroè, Tommaso accetta di paragonare l'intelletto alla materia prima: come la materia non possiede di per sé la forma di alcuna cosa, ma è pura potenza, così l'intelletto umano è detto «possibile» per lo stesso motivo, ossia perché non possiede alcuna natura, se non quella «di essere possibile».

il tempo. Nemmeno la conoscenza di sé sfugge a questa condizione: l'uomo si conosce soltanto conoscendo⁶.

Sviluppando questa prospettiva, Tommaso rimarca anche la necessità, per il nostro intelletto, dell'apporto della conoscenza sensibile. La ricezione sensoriale viene rielaborata dalla sensibilità, producendo così la rappresentazione sensibile – il cosiddetto «fantasma». Il fantasma è l'oggetto del quale l'intelletto ha bisogno per produrre l'universale, separando l'elemento formale dai caratteri materiali e individuanti⁷. Ciò significa non solo che l'universalità di cui l'uomo è capace non può che essere *a posteriori*, ma anche che l'intelletto umano ha *sempre* bisogno della sensibilità, anche quando ha già elaborato la specie intelligibile⁸: il fantasma è infatti indispensabile per l'intelletto, così come i colori lo sono per la vista⁹.

Cfr. Averroé, *Commentarium Magnum in Aristotelis de Anima libros*, l. 3, § 5, p. 387, ll. 13-16 et 23-26; *Ibid.*, § 17, pp. 436-437, ll. 19-27; Tommaso d'Aquino, *Scriptum super libros Sententiarum magistri Petri Lombardi episcopi Parisiensis*, liber 2 (= *In II Sent.*), d. 3, q. 1, a. 3, *sol.*, p. 94; *Id.*, *In IV Sententiarum*, d. 49, q. 2, a. 1, ad 10, p. 684; *Id.*, *De ente et essentia*, c. 4, pp. 315-381, p. 377, ll. 180-185; *Id.*, *Liber de veritate catholicae Fidei contra errores infidelium seu Summa contra Gentiles* (= SCG), l. 2, c. 76, § 1563, p. 222; *Ibid.*, c. 96, § 1816, p. 267; *Ibid.*, c. 98, §§ 1828 et 1839, pp. 268 e 270-271; *Ibid.*, l. 3, c. 39, § 2173, pp. 45-46; *Id.*, *Summae Theologiae* (= ST), I, q. 14, a. 2, ad 3, t. 4, p. 169; *Ibid.*, q. 87, a. 1, *resp.*, t. 5, p. 355; *Ibid.*, I-II, q. 50, a. 6, *resp.*, t. 6, p. 323; *Id.*, *Quaestiones disputatae de veritate* (= QDV), q. 8, a. 6, *resp.*, p. 238, ll. 148-158; *Ibid.*, q. 10, a. 8, *resp.*, p. 322, ll. 258-269, *Ibid.*, ad sc. 6, p. 325, ll. 509-516; *Ibid.*, q. 14, a. 1, *resp.*, p. 436, ll. 83-91; *Ibid.*, q. 18, a. 7, *resp.*, p. 556, ll. 110-119; *Id.*, *Quaestiones disputatae de anima* (= QD *De an.*), q. 2, ad 17, p. 21, ll. 490-499; *Id.*, *Sentencia libri de Anima* (= SDA), l. 3, c. 3, pp. 216-217, ll. 90-102.

⁶ Cfr. Tommaso d'Aquino, SCG, l. 3, c. 26, § 2080, p. 35; *Ibid.*, c. 46, § 2233, p. 60; *Id.*, ST, I, q. 14, a. 2, ad 3, t. 4, p. 169; *Ibid.*, q. 89, a. 2, *resp.*, t. 5, p. 375; *Id.*, QDV, q. 8, a. 6, *resp.*, p. 238, ll. 167-172; *Ibid.*, q. 10, a. 2, *resp.*, pp. 300-301, ll. 105-121; *Id.*, *Quaestiones disputatae de potentia* (= QDP), q. 7, a. 9, *resp.*, pp. 207-208; *Id.*, QD *De an.*, q. 2, ad 5, pp. 19-20, ll. 389-392.

⁷ Cf. Tommaso d'Aquino, *In II Sent.*, d. 3, q. 1, a. 2, ad 3, p. 91; *Ibid.*, q. 3, a. 3, ad 1, pp. 121-122; *Id.*, ST, I, q. 76, a. 2, ad 3, t. 5, p. 217; *Id.*, SCG, l. 2, c. 77, § 1581-1582, p. 226; *Ibid.*, c. 100, § 1856, p. 273; *Id.*, QDV, q. 8, a. 6, ad 5, p. 239, ll. 200-206; *Id.*, QD *De an.*, q. 4, *resp.*, p. 33, ll. 115-125.

⁸ Cfr. Tommaso d'Aquino, SDA, l. 3, c. 2, p. 209, ll. 32-50; *Ibid.*, l. 3, c. 7, p. 236, ll. 90-96; *Id.*, QDP, q. 3, a. 9, ad 22, p. 68; *Id.*, *Compendium theologiae seu Brevis compilatio theologiae ad fratrem Raynaldum*, l. 1, c. 82, p. 107; *Id.*, SCG, l. 2, c. 60, § 1387, p. 193.

⁹ Cfr. Tommaso d'Aquino, *In II Sent.*, d. 3, q. 3, a. 1, ad 2, p. 114; *Ibid.*, d. 17, q. 2, a. 1, *sol.*, p. 424; *Id.*, ST, I, q. 75, a. 2, ad 3, p. 197; *Id.*, QDV, q. 5, a. 10, *resp.*, p. 170, ll. 167-172; *Id.*, *Questio Disputata De Spiritualibus creaturis* (= QD *De sp. creat.*), a. 2, *resp.*, p. 25, ll. 209-211; *Id.*, *Quaestiones de quolibet VII*, q. 5, a. 1, ad 3, p. 25, ll. 87-91. Il paragone con i colori in rapporto alla vista è abitualmente attribuito ad Aristotele, ma non si trova letteralmente nel testo dello Stagirita, per quanto esso renda bene quanto Aristotele vuol dire, quando paragona

Per quanto l'indeterminatezza e la necessità dell'apporto della sensibilità siano caratteri specifici della nostra intelligenza, non c'è dubbio che l'elemento più proprio, che rende l'intelligenza umana «ragione», è la *discorsività*. Noi non giungiamo intuitivamente alla verità, come se il pensare si identificasse con un puro vedere: la ragione procede, collegando premesse a conclusioni, separando ed associando tra loro rappresentazioni, idee, concetti – *componendo et dividendo*¹⁰. A noi, enti dotati di ragione, tocca *costruire* il giudizio sulle cose: l'attività umana di pensiero si manifesta come un percorso laborioso, che esige impegno, implicando sforzo e dedizione.

Tommaso d'Aquino rimarca che questa condizione ci espone tanto facilmente all'errore, che non ci si deve meravigliare di trovare, all'interno della rivelazione cristiana, anche delle verità che sarebbero di per sé già alla nostra portata: nella sua bontà, Dio ha dato il sigillo della sua autorità a certe verità fondamentali per la nostra vita, verità che, con la forza della semplice ragione, solo pochi uomini particolarmente intelligenti avrebbero potuto raggiungere, e non senza dover dedicare molto tempo alla loro ricerca¹¹.

Gli elementi che abbiamo brevemente richiamato ci permettono di comprendere perché, usando un linguaggio di matrice neoplatonica, Tommaso abbia affermato che l'uomo è «solo» un essere razionale¹²: noi possiamo

l'attività intellettuale sensibile a quella sensibile. L'origine è piuttosto Averroè, *Commentarium Magnum in Aristotelis de Anima libros*, l. 3, § 5, p. 401, ll. 410-417. Cfr. Cos, *Nota* alla linea 211 di Tommaso d'Aquino, *QD De sp. creat.*, a. 2, *resp.*, p. 25.

¹⁰ «Intellectus humanus necesse habet intelligere componendo et dividendo. Cum enim intellectus humanus exeat de potentia in actum, similitudinem quandam habet cum rebus generabilibus, quae non statim perfectionem suam habent, sed eam successive acquirunt. Et similiter intellectus humanus non statim in prima apprehensione capit perfectam rei cognitionem; sed primo apprehendit aliquid de ipsa, puta quidditatem ipsius rei, quae est primum et proprium obiectum intellectus; et deinde intelligit proprietates et accidentia et habitudines circumstantes rei essentiam. Et secundum hoc, necesse habet unum apprehensum alii componere vel dividere; et ex una compositione vel divisione ad aliam procedere, quod est ratiocinari» (Tommaso d'Aquino, *ST*, I, q. 85, a. 5, *resp.*, t. 5, p. 341).

¹¹ Cfr. *Id.*, *SCG*, I.1, c. 4, §§ 23-26, p. 6. Si vedano anche *Id.*, *QDV*, q. 14, a. 10, *resp.*, p. 467, ll. 184-209; *Id.*, *ST*, I, q. 1, a. 1, *resp.*, t. 4, p. 6; *Id.*, II-II, q. 2, a. 4, t. 7, pp. 30-31.

¹² Cfr. Tommaso d'Aquino, *Scriptum super libros Sententiarum magistri Petri Lombardi episcopi Parisiensis*, t. 1 (= *In I Sent.*), d. 25, q. 1, a. 1, ad 4, p. 604; *Id.*, *In II Sent.*, d. 3, q. 1, a. 6, *sol.* et ad 2, pp. 104-106; *Ibid.*, d. 9, q. 1, a. 4, *sol.*, pp. 239-240; *Ibid.*, d. 9, q. 1, a. 8, ad 1, pp. 249-250; *Ibid.*, d. 16, q. 1, a. 3, *sol.*, p. 402; *Id.*, *ST*, I, q. 58, a. 3, *resp.*, t. 5, p. 83; *Ibid.*, q. 59, a. 1, ad 1, p. 92; *Id.*, *QDV*, q. 15, a. 1, ad 7, p. 482, ll. 500-511; *Id.*, *QD De sp. creat.*, a. 2, ad 12, p. 32, ll. 430-438; *Id.*, *In Librum Beati Dionysii de Divinis nominibus Expositio*, c. 5, lect. 1, § 615, p. 232. La divisione dei viventi in esseri intelligenti, razionali e sensibili è mutuata dal trattato *Sui Nomi divini* dello pseudo-Dionigi: cfr. Pseudo-Dionigi Areopagita, *De divinis nominibus*, p. 2, s. 1, pp. 167-176. Tommaso riprende, come una sorta di adagio,

concepire, ma non possedere, forme più raffinate e più perfette di intelligenza. Ecco perché «intelletto», nel senso più proprio e più pieno, è solo Dio. Inoltre, Tommaso ritiene filosoficamente necessario stabilire l'esistenza di una forma di intelligenza intermedia tra Dio e l'uomo, che coinciderebbe con quella degli angeli, di cui parla la Bibbia. Gli angeli sono infatti concepiti come intelligenze pure, ma anche, allo stesso tempo, finite¹³.

Al di là della conclusività che possiamo accordare oggi agli argomenti con cui Tommaso sostiene la necessità filosofica che gli angeli esistano¹⁴, non si può negare il valore teorico che l'angelologia possiede nel quadro speculativo dell'Aquinate: in effetti, quando tratta della natura della nostra razionalità, Tommaso propone frequentemente un confronto tra la nostra intelligenza e quella degli angeli. In un tale quadro teorico, l'angelologia svolge l'importante funzione di «limite di salvaguardia», che impedisce di attribuire all'uomo modelli di intelligenza che non possono competergli, senza che questo significhi che tali modelli siano falsi o inconsistenti¹⁵. Così, è l'angelo, non l'uomo, che possiede la capacità di essere immediatamente cosciente di sé; è l'atto di pensiero dell'angelo, non quello dell'uomo, ad essere compiutamente intuitivo, possedendo da sempre, in modo connaturale e innato, le «idee chiare e distinte» delle cose¹⁶; è l'angelo, non l'uomo, che quando agisce *secundum naturam* persegue infallibilmente il bene razionale che coglie¹⁷.

anche l'affermazione di Isaac Israeli, secondo la quale la ragione umana «è prodotta all'ombra dell'intelligenza»: Tommaso d'Aquino, *In II Sent.*, d. 3, q. 1, a. 6, *sol.*, p. 104; Id., SCG, l. 2, c. 46, § 1233, p. 167; *Ibid.*, c. 91, §§ 1774-1775, pp. 260-261; Isaac Israeli, *De definitionibus*, p. 313, ll. 26-27.

¹³ Cfr. Tommaso d'Aquino, QD *De sp. creat.*, a. 5, *resp.*, pp. 59-62, ll. 131-258; Id., SCG, l. 2, c. 91, §§ 1777-1780, pp. 261-262; Id., ST, I, q. 50, a. 1, *resp.*, t. 5, p. 4; Id., *In librum De causis expositio* (= *In De causis*), lect. 2, §§ 57-61, pp. 15-16; *Ibid.*, lect. 7, §§ 180-195, pp. 51-52; *Ibid.* lect. 9, §§ 211-222, pp. 60-61.

¹⁴ Sono diversi gli studiosi tomisti che hanno visto negli argomenti proposti da Tommaso soltanto delle suggestioni, prive di autentica apoditticità. Cfr. Elders (1993: 235-236); Doolan (2012: 13-44). È ancora più radicale la lettura di Bernardo Carlos Bazán: oltre a non essere apodittica, la dimostrazione dell'esistenza di sostanze create puramente intellettuali avrebbe avuto degli effetti negativi sull'antropologia filosofica di Tommaso, perché avrebbe fatto ricadere anche sull'anima umana, che è invece essenzialmente unita al corpo, l'immagine di una sostanza intellettuale separata, quale essa non è. Cf. Bazán (2010: 47-85).

¹⁵ Sottolineano questo aspetto Gilson (1989⁶: 209); Collins (1947 : xiii); Vernier (1986: 78-79); Putallaz (1991: 226-233); Suarez-Nani (2002: 26-53, 72-73); Id., (2006: 21-29); Scribano (2006: 27).

¹⁶ Cfr. in particolare Id., *In II Sent.*, d. 3, q. 3, a. 2, ad 1, p. 117.

¹⁷ «Homo secundum suam naturam non statim natus est ultimam perfectionem adipisci, sicut angelus. Et ideo homini longior via data est ad merendum beatitudinem, quam angelo» (Tommaso d'Aquino, ST, I, q. 62, a. 5, ad 1, t. 5, p. 115).

Nell'ottica di Tommaso, la nostra forma di intelligenza si chiama dunque «ragione» più propriamente che «intelletto», proprio perché essa non è in grado di possedere evidenze immediate, se non quella dei primi principi¹⁸. È certamente suggestivo il fatto che un lettore attento di Tommaso d'Aquino come Jacques Maritain si sia opposto al *cogito* cartesiano vedendo in esso un «peccato di angelismo»¹⁹.

1.2. L'inaghirabile volere

L'analisi della ragione umana che abbiamo proposto fino a qui ci permette di cogliere una prima risposta di Tommaso alla questione che abbiamo posto all'inizio, ossia perché il possesso della ragione non ci dà la garanzia di compiere sempre le azioni migliori. Secondo l'Aquinate, la *possibilità* di ingannarsi e di produrre valutazioni non adeguate sul da farsi è inscritta nella struttura stessa della ragione umana, in quanto caratterizzata da tre elementi fondamentali: l'indeterminatezza originaria e insormontabile; la dipendenza, almeno *ex parte obiecti*, da facoltà non razionali; la possibilità di arrivare alla verità solo attraverso un procedimento laborioso²⁰.

¹⁸ «In homine est intellectus; non tamen propter hoc in ordine intellectualium proprie ponitur: quia illa substantia intellectualis dicitur cujus tota cognitio secundum intellectum est, quia omnia quae cognoscit, subito sine inquisitione sibi offeruntur: non autem ita est de cognitione animae, quia per inquisitionem et discursum rationis ad notitiam rei venit: et ideo rationalis dicitur, quia ejus cognitio secundum terminum tantum et secundum principium intellectualis est: secundum principium, quia prima principia sine inquisitione statim cognoscit, unde habitus principiorum indemonstrabilium intellectus dicitur; secundum terminum vero, quia inquisitio rationis ad intellectum rei terminatur; et ideo non habet intellectum ut naturam propriam, sed per quamdam participationem. Ratio autem et de Deo et de angelis dicitur; tamen alio modo sumitur, secundum quod scilicet omnis cognitio immaterialis ratio potest dici, prout dividitur ratio contra sensum, et non contra intellectum» (Tommaso d'Aquino, *In II Sent.*, d. 3, q. 1, a. 6, ad 2, pp. 105-106). Cfr. anche *Ibid.*, d. 24, q. 2, a. 3, *sol.*, p. 610; *Id.*, *ST*, I, q. 58, a. 3, *resp.*, t. 5, p. 83; *Id.*, *Expositio libri Posteriorum*, t. 1*/2, II, lect. 20, p. 246-247, ll. 292-340.

¹⁹ «Tâchons de trouver aux choses leur nom véritable: le péché de Descartes est un péché d'*angélisme*, il a fait de la Connaissance et de la Pensée une Perplexité sans remède, un abîme d'inquiétude, parce qu'il a conçu la Pensée humaine sur le type de la Pensée angélique» (Maritain (1984 : 487).

²⁰ «In nobis sunt quaedam potentiae et habitus cognoscitivi in quibus nunquam falsitas esse potest, sicut sensus, et scientia, et intellectus principiorum; quidam vero in quibus potest esse falsum, sicut imaginatio, et opinio, et extimatio. Ex hoc autem falsitas accidit in aliqua cognitione, quod non est ita in re sicut apprehenditur; unde si aliqua vis cognoscitiva est talis quod nunquam in ea est falsitas, oportet quod suum cognoscibile nunquam deficiat ab eo quod de ipso cognoscens apprehendit» (Tommaso d'Aquino, *QDV*, q. 2, a. 12, *resp.*, p. 83, ll. 153-163). Si veda al riguardo il contributo di Roland-Gosselin (1923).

Tuttavia, questo tipo di risposta non è ancora sufficiente: l'indicazione della possibilità di ingannarsi non dice ancora che cosa provochi il *fatto* che ci inganniamo, né perché ci accada di agire «acraticamente», ossia contro le nostre stesse convinzioni. Al riguardo, come stiamo per vedere, Tommaso sostiene senza esitazioni che dietro l'irrazionalità pratica c'è sempre una qualche falsa opinione che ci facciamo sulle cose e dunque una cognizione imperfetta di ciò che dovremmo fare: non commetteremmo azioni contrarie a ragione, se cogliessimo con piena evidenza il bene da perseguire. Questo significa allora che Tommaso ha sostenuto una dottrina di tipo «intellettualistico»?

Non si può certamente dire che l'Aquinate abbia schiacciato il problema su una questione limitatamente cognitiva. Egli infatti distingue l'ignoranza vera e propria (*ignorantia*) dal semplice non-sapere (*nescientia*). Questa precisazione è importante, perché l'ignoranza non indica la semplice mancanza di conoscenza da cui tutti gli uomini, senza eccezioni, partono, prima di conoscere qualcosa. In senso stretto, per Tommaso, si è ignoranti quando non si conosce ciò che si sarebbe tenuti a conoscere. Per questo, il soggetto è per lo più responsabile della propria ignoranza, perché per compiere l'azione più razionale bisogna *volere* conoscere le cose in modo esauriente e volerle prendere in considerazione *accuratamente*, secondo i diversi aspetti che le caratterizzano²¹.

Questa idea è coerente con il fatto che l'azione umana, per come Tommaso la concepisce, è l'effetto di una sinergia complessa di cognizione e appetizione, che sorge dall'intreccio non districabile di contributi distinti di ragione pratica e volontà. Esse si influenzano reciprocamente, tanto al livello dell'intenzione e della scelta, quanto a quello dell'esecuzione e dell'espletamento effettivo dell'azione. Così, l'intenzione (della volontà) sta alla base della deliberazione (della ragione) che precede la scelta (della volontà). Solo a quel punto, il frutto della scelta viene indicato imperativamente come da farsi (dalla ragione) e quindi messo in atto (dalla volontà)²².

Pur senza entrare nei dettagli, possiamo almeno segnalare che la dinamica dell'azione descritta da Tommaso prende le mosse da due punti assolutamente fermi, universali e necessari: uno sta dal lato della volontà e consiste nel fatto che la felicità, in quanto fine ultimo di qualsiasi azione, viene voluto

²¹ Cfr. Tommaso d'Aquino, *Quaestiones disputatae de malo* (=QD *De malo*), q. 8, a. 1, ad 7, pp. 195-196, ll. 406-430; Id., ST, I-II, q. 6, a. 8, *resp.*, t. 6, pp. 62-63; Id., q. 76, aa. 1-4, pp. 52-58; si veda al riguardo Lottin (1949 : 9-51).

²² Cfr. Tommaso d'Aquino, ST, I-II, qq. 6-17, t. 6, pp. 55-126. Per un'analisi di questa dinamica, rimandiamo almeno a Forlivesi (1996); Pinckaers (1997).

assolutamente e necessariamente²³; l'altro sta dal lato della ragione pratica e si identifica con il possesso della sinderesi, ovvero con la capacità dell'intelletto pratico di cogliere *infallibilmente* il principio primo *bonum est faciendum, malum vitandum*²⁴.

A partire da questi due punti fermi, la volontà interviene, scegliendo liberamente ciò che porta al fine, attraverso la considerazione inquisitiva della ragione²⁵. Così, tutti gli altri momenti dell'azione sono caratterizzati da una certa indeterminatezza, in cui la volontà può certamente spingere la ragione verso la produzione di un giudizio retto e compiuto, che consideri a fondo tutti gli aspetti, ma può anche limitarne la spinta veritativa, facendo stazionare il giudizio su conclusioni affrettate, inadeguate o unilaterali²⁶. In forza di queste considerazioni, Tommaso può sostenere che la possibilità di compiere azioni non conformi a ragione dipende sempre da un qualche errore di valutazione, senza che questo significhi minimizzare il fatto che sia la volontà a spingere o meno la ragione ad andare fino in fondo nel giudizio sul bene da compiersi²⁷.

Questo tipo di soluzione fa tuttavia sorgere un problema. Tommaso sostiene chiaramente che la volontà è pur sempre un *appetito razionale*: il suo

²³ «Quilibet volens de necessitate vult suum ultimum finem: sicut homo de necessitate vult suam beatitudinem, nec potest velle miseriam» (Tommaso d'Aquino, SCG, l. 1, c. 80, § 678, p. 93). Vedi anche Id., ST, I-II, q. 13, a. 6, *resp.*, t. 6, p. 103; Id., QD *De malo*, q. 3, a. 3, *resp.*, p. 72, ll. 171-184.

²⁴ «Sicut autem ens est primum quod cadit in apprehensione simpliciter, ita bonum est primum quod cadit in apprehensione practicae rationis, quae ordinatur ad opus, omne enim agens agit propter finem, qui habet rationem boni. Et ideo primum principium in ratione practica est quod fundatur supra rationem boni, quae est, *Bonum est quod omnia appetunt*. Hoc est ergo primum praeceptum legis, quod *bonum est faciendum et prosequendum, et malum vitandum*. Et super hoc fundantur omnia alia praecepta legis naturae: ut scilicet omnia illa facienda vel vitanda pertineant ad praecepta legis naturae, quae ratio practica naturaliter apprehendit esse bona humana. Quia vero bonum habet rationem finis, malum autem rationem contrarii, inde est quod omnia illa ad quae homo habet naturalem inclinationem, ratio naturaliter apprehendit ut bona, et per consequens ut opere prosequenda, et contraria eorum ut mala et vitanda» (Tommaso d'Aquino, ST, I-II, q. 94, a. 2, *resp.*, t. 6, p. 170). Sulla sinderesi, vedi Id., ST, I, q. 79, a. 12, *resp.* et ad 3, t. 5, pp. 280-281.

²⁵ Cfr. Tommaso d'Aquino, ST, I, q. 19, a. 5, *resp.*, t. 4, p. 239; *Ibid.*, q. 58, a. 7, ad 2, p. 90; *Ibid.*, q. 83, a. 4, *resp.*, p. 303; Id., ST, I-II, q. 8, a. 3, *resp.* et ad 3, t. 6, p. 72; Id., QDV, q. 5, a. 1, *resp.*, p. 139, ll. 151-170; Id., q. 22, a. 6, *resp.*, pp. 627-628, ll. 68-138.

²⁶ Cfr. Tommaso d'Aquino, QD *De malo*, q. 16, a. 6, ad 6 e ad 11, p. 311, ll. 352-356, 384-391. Vedi anche Id., *In II Sent.*, d. 24, q. 3, a. 1, *sol.*, p. 617; Id., ST I-II, q. 8, a. 1, *resp.*, t. 6, p. 68; *Ibid.*, q. 9, a. 1, *resp.*, p. 74.

²⁷ Cfr. Id., ST, I-II, q. 6, a. 8, *resp.*, t. 6, p. 63; Id. QD *De malo*, q. 6, a. un., *resp.*, p. 150, ll. 425-440; *Ibid.*, q. 7, a. 5, ad 7-8, pp. 173-174, ll. 235-256; Id., *Ibid.*, q. 8, a. 1, ad 7, p. 196, ll. 425-430.

oggetto in effetti è il bene colto cognitivamente e il suo atto è intrinsecamente legato a quel che le presenta il giudizio pratico della ragione²⁸. Coerentemente con ciò, Tommaso precisa che la volontà muove se stessa, dal punto di vista dell'esecuzione dell'atto, ma è l'intelletto a muovere la volontà quanto alla specificazione di ciò che viene voluto²⁹. Ora, se le cose stanno in questo modo, come è possibile che la volontà possa dare il proprio assenso a ciò che non è razionale?

Tommaso risolve la questione ispirandosi ad Aristotele: l'oggetto proprio della volontà, colto infallibilmente da parte sua, è la nozione di bene in universale, mentre l'azione ha sempre una natura singolare e contingente³⁰. Questo scarto, nell'ottica dell'Aquinate, si rivela decisivo. Poiché il principio primo della ragione pratica – *bonum faciendum, malum vitandum* – è colto in maniera infallibile, nessuno fa deliberatamente il male in quanto male, né si comporterebbe in modo contrario alla ragione, se cogliesse fino in fondo l'irrazionalità del proprio agire. Tuttavia, non è sufficiente giudicare qualcosa come un bene, perché diventi oggetto della decisione concreta: esso deve manifestarsi come il bene adatto (*conveniens*) a questa precisa situazione, e quindi tale da essere scelto *hic et nunc*. Il punto cruciale sta allora nel fatto che nessuno dei beni operabili, compresi quelli consoni alla ragione, può apparirci come la cosa migliore da farsi universalmente e sotto ogni aspetto, ma sempre solo sotto un *certo* aspetto, limitato e parziale³¹.

Ciò significa che ogni azione presenta degli elementi che, all'interno della singola situazione concreta, possono indurre a non compierla. Pensiamo a quando cogliamo che un'azione, per quanto possa apparirci razionale, provocherà però anche la perdita di altri beni o la rinuncia a certi piaceri e vantaggi; aggiungiamoci poi la presenza in noi della presunzione o dell'accidia, che ci impediscono di produrre una valutazione corretta delle cose, perché, per ragioni diverse, non ci fanno prendere in considerazione tutti gli elementi

²⁸ Cfr. Tommaso d'Aquino, ST, I-II, q. 6, a. 1, *resp.*, t. 6, p. 56; *Ibid.*, I-II, q. 8, a. 1, *resp.*, p. 68. Vedi anche Id., ST, I-II, q. 9, a. 1, ad 3, t. 6, p. 78.

²⁹ Cfr. Tommaso d'Aquino, ST, I-II, q. 9, a. 1, *resp.*, t. 6, p. 74; Id., ST I, q. 82, a. 4, t. 5, p. 303; Id., QDV, q. 22, a. 12, pp. 642-643, ll. 54-154; SCG, l. 3, c. 26, §§ 2078-2080, pp. 34-35; Id., QD *De malo*, q. 6, a. un., *resp.*, p. 149, ll. 339-391.

³⁰ Cfr. Tommaso d'Aquino, *In III Sent.*, d. 26, q. 1, a. 2, *resp.*, pp. 816-817; Id., ST, I-II, q. 1, a. 2, ad 3, t. 6, p. 9; vedi anche Id., ST, I, q. 82, a. 4, *resp.*, t. 5, p. 303; Id., ST, I-II, q. 2, a. 6, *resp.*, t. 6, p. 22; *Ibid.*, q. 9, a. 1, *resp.*, p. 74; Id., ST I-II, q. 76, a. 1, *resp.*, t. 7, p. 52; Id., ST, II-II, q. 24, a. 1, *resp.*, t. 8, p. 174; Aristotele, *Etica Nicomachea*, l. 7, c. 3, 1147a1-10.

³¹ Cfr. Tommaso d'Aquino, QD *De malo*, q. 3, a. 9, ad 7, p. 88, ll. 286-289; *Ibid.*, q. 6, a. un., *resp.*, p. 148, ll. 284-292; *Ibid.*, p. 150, ll. 425-440.

necessari per capire quale sia la cosa migliore da fare³². Per questo Tommaso sottolinea la necessità di coltivare la virtù aristotelica della prudenza: essa infatti è la capacità di fare la cosa più giusta non in generale, o astrattamente, ma nella particolarità dell'azione concreta, tenendo conto della vita intesa come un tutto³³.

1.3. Tra ragione e passioni: le radici dell'*akrasia*

La messa in luce dello scarto tra dimensione universale del bene appreso e dimensione particolare dell'azione da farsi fornisce a Tommaso la chiave per spiegare il fenomeno dell'*akrasia*. Aristotele ne fa una manifestazione fondamentale dell'incontinenza: è l'eccesso di passione che offusca la ragione e impedisce di fare il bene di cui pure, in qualche modo, si ha cognizione³⁴. Abbiamo già ricordato che nel pensiero medievale, si è piuttosto messo l'accento sulla debolezza della volontà. Dal canto suo, Tommaso fa certamente dell'*akrasia* la manifestazione di una volontà incapace di promuovere il bene razionale che pure coglie; tuttavia, per spiegare come l'*akrasia* sia possibile, egli non manca di mettere in rilievo, sulla scia di Aristotele, l'importanza che ha nell'uomo il rapporto tra ragione e passioni. Secondo Tommaso, le passioni sono capaci di esercitare una certa pressione sulla volontà, in rapporto all'azione concreta, perché esse possiedono, in comune con le azioni da farsi, una natura immediatamente singolare e contingente, mentre la volontà apprende il bene cognitivamente e dunque secondo il carattere dell'universalità³⁵. Per usare un'immagine, potremmo dire che la volontà, che è razionale, per produrre l'azione deve parlare una lingua, quella della contingenza e della singolarità, rispetto alle quali sono le passioni ad avere le capacità oratorie migliori.

Questo richiamo alla dimensione passionale non significa che Tommaso voglia dare una connotazione (im)morale alle passioni: esse sono, in senso proprio, dei fenomeni psicologici – moti dell'anima – che si verificano in noi

³² «Quia peccatum in actu consistit, error manifeste habet rationem peccati. Non enim est absque praesumptione, quod aliquis de ignoratis sententiam ferat, et maxime in quibus periculum existit» (Id., *QD De malo*, q. 3, a. 7, resp., p. 81, ll. 97-100); sull'accidia come radice dell'ignoranza volontaria *Ibid.*, q. 8, a. 1, ad 7, p. 196, ll. 426-430.

³³ Cfr. Tommaso d'Aquino, ST, I-II, q. 57, aa. 4-6, pp. 367-371; Id., *Sententia libri Ethicorum* (= *Sent. Ethic.*), l. 6, c. 4, p. 345, ll. 16-23.

³⁴ Aristotele, *Etica Nicomachea*, l. 7, c. 2, 1145b22-31.

³⁵ «Ratio est propinquior voluntati quam passio, set particulari appetibili est propinquior passio quam ratio universalis» (Tommaso d'Aquino, *QD De malo*, q. 3, a. 9, ad 13, p. 88, ll. 317-320).

per il semplice fatto che siamo dotati di anima e corpo³⁶. Piuttosto, egli ricorda che nell'uomo la volontà deve svolgere un compito che non può essere eluso: dato che l'uomo non è un angelo, una *felicità integralmente umana* non è pensabile senza che tutte le dimensioni umane, comprese quelle non razionali, siano condotte verso i beni indicati dalla ragione³⁷. Verrebbe da dire che, nella prospettiva di Tommaso, il «vantaggio pratico» delle passioni sulla volontà svolge un ruolo positivo, nella salvaguardia della genuina integrità umana, perché esso impedisce alla volontà di trascurare il corpo e la sensibilità: in quanto guida dell'intero umano, la ragione non può fruire pienamente dei suoi stessi beni, se non spinge anche la sensibilità verso l'ordine razionale della *beatitudo*.

Al riguardo, richiamando un passo della *Politica* di Aristotele, Tommaso spiega che la ragione ha con le passioni un rapporto *politico* (nel senso di *democratico*), non *dispotico*³⁸. In effetti, in quanto espressioni specifiche della sensibilità, le passioni godono di una certa autonomia rispetto alla ragione, perciò non possono venire costrette «dispoticamente» a fare ciò che la ragione indica. Prendere atto che la natura delle passioni è «non razionale» non significa però concepirle come «irrazionale» o «antirazionale». Si tratta piuttosto di prendere atto che la volontà, una volta colto ciò che è secondo ragione, può e deve impegnarsi a *persuadere* la sfera passionale della bontà di ciò che è razionale – come dovrebbe fare un buon *leader* democratico – producendo

³⁶ Sulla natura delle passioni in Tommaso e sulla loro portata antropologica, si vedano Meyer (1994); Laghi (2000); Vecchio (2002: 5-18); Id. (2009: 81-86).

³⁷ «Hoc autem est bonum in communi, in quod voluntas naturaliter tendit, sicut etiam quaelibet potentia in suum obiectum, et etiam ipse finis ultimus, qui hoc modo se habet in appetibilibus, sicut prima principia demonstrationum in intelligibilibus, et universaliter omnia illa quae conveniunt volenti secundum suam naturam. Non enim per voluntatem appetimus solum ea quae pertinent ad potentiam voluntatis; sed etiam ea quae pertinent ad singulas potentias, et ad totum hominem. Unde naturaliter homo vult non solum obiectum voluntatis, sed etiam alia quae conveniunt aliis potentiis, ut cognitionem veri, quae convenit intellectui; et esse et vivere et alia huiusmodi, quae respiciunt consistentiam naturalem; quae omnia comprehenduntur sub obiecto voluntatis, sicut quadam particularia bona» (Tommaso d'Aquino, ST, I-II, q. 10, a. 1, *resp.*, t. 6, p. 83).

³⁸ «Sicut Philosophus dicit in I *Polit.*, ratio, in qua est voluntas, movet suo imperio irascibilem et concupiscibilem, non quidem *despotico principatu*, sicut movetur servus a domino; sed *principatu regali seu politico*, sicut liberi homines reguntur a gubernante, qui tamen possunt contra movere. Unde et irascibilis et concupiscibilis possunt in contrarium movere ad voluntatem. Et sic nihil prohibet voluntatem aliquando ab eis moveri» (Tommaso d'Aquino, ST, I-II, q. 9, a. 2, ad 3, t. 6, p. 77); Cfr. Id., ST, I, q. 81, a. 3, ad 2, t. 5, p. 191; Id., ST, I-II, q. 17, a. 7, *resp.*, t. 6, p. 123; *Ibid.*, q. 56, a. 4, ad 3, p. 359; *Ibid.*, q. 58, a. 2, *resp.*, p. 373; Id., *Sent. Ethic.*, l. 10, c. 10, p. 583, ll. 21-27; Id., QD *De malo*, q. 3, a. 9, ad 14, p. 88, ll. 322-328. Cfr. Aristotele, *Politica*, l. 1, c. 3, 1254b5-6.

così nella sensibilità quella che possiamo chiamare *la passione per il bene razionale*. Questo, agli occhi di Tommaso, è il modo sicuro per coprire con successo la distanza tra l'universalità del bene e la singolarità dell'azione.

Letta alla luce di questa relazione tra passioni e ragione, l'idea che *l'uomo per essere razionale debba essere virtuoso* è perfettamente coerente con la presa d'atto della sua natura intrinsecamente corporea: la ragione non ha solo il compito di determinare in astratto l'azione migliore da compiersi, ma deve anche «prendere dimora» presso la sensibilità. Coerentemente con ciò, Tommaso presenta la virtù morale come disposizione, acquisita nel tempo e attraverso un certo numero di atti, grazie alla quale la ragione è in grado di permeare la sensibilità, così che il soggetto agente «percepisca», «viva» e «senta» l'azione conforme a ragione come conveniente e attraente, facendo *con piacere* la cosa migliore e detestando *visceralmente* le condotte irrazionali³⁹.

Possiamo allora comprendere perché per Tommaso è certamente il vizio, piuttosto che la passione in se stessa, a costituirsi come ostacolo potente al compimento delle azioni migliori. Al vizioso una certa azione potrebbe anche apparire come la migliore, ma essa, a causa del vizio, non gli si potrà presentare immediatamente e facilmente come «conveniente». Per questo, il vizio può anche non impedire al soggetto agente di comprendere astrattamente quale sia la condotta più razionale, ma gliela farà certamente apparire come inadatta, «lontana» dalla singolarità dell'azione da compiere. Così, spiega Tommaso, il vizio mette l'uomo, ente razionale, nella condizione innaturale di seguire ciò che non è razionale⁴⁰. Non deve stupire allora che agli occhi di

³⁹ «Sed si aliquis timeat et audeat et sic de aliis quando oportet et in quibus oportet et ad quos oportet et cuius gratia oportet et eo modo quo oportet, hoc erit medium in passionibus et in hoc consistit optimum virtutis» (Tommaso d'Aquino, *Sent. Ethic.*, l. 2, c. 6, p. 96, ll. 48-54). «Nam radix virtutis consistit in ipsa rectitudine rationis secundum quam oportet actiones et passiones dirigere. Aliter tamen sunt dirigibiles actiones, quam passiones» (*Ibid.*, c. 8, p. 102, ll. 37-39). «Secundum hoc, aliquantulum verum est quod Socrates dixit, quod scientia praesente, non peccatur, si tamen hoc extendatur usque ad usum rationis in particulari eligibili. Sic igitur ad hoc quod homo bene agat, requiritur quod non solum ratio sit bene disposita per habitum virtutis intellectualis; sed etiam quod vis appetitiva sit bene disposita per habitum virtutis moralis. Sic igitur ad hoc quod homo bene agat, requiritur quod non solum ratio sit bene disposita per habitum virtutis intellectualis; sed etiam quod vis appetitiva sit bene disposita per habitum virtutis moralis. Sicut igitur appetitus distinguitur a ratione, ita virtus moralis distinguitur ab intellectuali. Sicut igitur appetitus distinguitur a ratione, ita virtus moralis distinguitur ab intellectuali. Unde sicut appetitus est principium humani actus secundum quod participat aliquantulum rationem, ita habitus moralis habet rationem virtutis humanae, in quantum rationi conformatur» (Id., *ST*, I-II, q. 58, a. 2, *resp.*, t. 6, p. 373).

⁴⁰ Cfr. Id., *ST*, I-II, q. 71, a. 2, ad 3, t. 7, p. 5.

un vizioso è il fatto di agire in modo conforme a ragione ad apparire, paradossalmente, come qualcosa di innaturale.

Tommaso mette dunque in luce con chiarezza quanto il vizio limiti la capacità di produrre una valutazione corretta dell'azione da compiersi. Tuttavia, egli non smette di ricordare che il vizio non cancella la libertà di giudicare e di agire altrimenti; il vizioso però, per mettere in atto l'azione conforme a ragione, dovrà adoperarsi in un duro sforzo, da cui la virtù lo avrebbe certamente dispensato.

1.4. Il peccato degli angeli e il paradosso dell'eccellenza razionale

L'importanza di questa dottrina nel contesto del pensiero di Tommaso emerge anche quando l'Aquinate, avendo assunto l'idea che gli angeli di cui parla la Bibbia sono intelligenze pure e perfette, si trova a spiegare il loro peccato. Si ricordi che Tommaso ha affermato che gli angeli, proprio grazie alla loro intelligenza intuitiva, non sono affetti dall'indeterminatezza che caratterizza l'intelletto umano, né rispetto al fine, né rispetto a ciò che è più razionale per raggiungerlo⁴¹. Inoltre, l'angelo non deve misurarsi con le passioni, come capita invece all'uomo. Per queste ragioni, l'angelo agisce in modo *infallibilmente* razionale, circa i beni naturali⁴². Come è possibile allora che delle creature dotate di un'intelligenza tanto perfetta possano avere commesso un'azione irrazionale, come il peccato?

Tommaso spiega che il peccato angelico non ha per oggetto un bene naturale, ma l'adesione all'intimo volere di Dio. Gli angeli conoscono perfettamente che Dio esiste e come Dio opera nelle creature; tuttavia, l'intimo contenuto della volontà divina appartiene ad un ordine *sconosciuto* all'angelo, non meno che all'uomo. Si tratta dunque di un ordine soprannaturale, nel senso che nessun angelo potrebbe conoscere, con i propri mezzi naturali, ciò che Dio vuole intimamente, se Dio stesso non glielo rivelasse. Per questo motivo, nel momento in cui tale rivelazione avviene, l'angelo si trova, per una volta, nella stessa situazione in cui l'uomo si trova abitualmente: deve prendere una *decisione particolare*, aderendo o meno ad una rivelazione ben determinata della volontà divina. Non possedendo la ragione, ma un intelletto vero e proprio, l'angelo deciderà una volta per tutte, in modo irrevocabile:

⁴¹ «Homo secundum suam naturam non statim natus est ultimam perfectionem adipisci, sicut angelus. Et ideo homini longior via data est ad merendum beatitudinem, quam angelo» (Tommaso d'Aquino, ST, I, q. 62, a. 5, ad 1, t. 5, p. 115).

⁴² Cfr. Id., ST, I, q. 63, a. 4, *resp.*, t. 5, p. 129; Id., QD *De malo*, q. 16, a. 6, *resp.*, p. 310, ll. 233-254.

perciò, o produrrà un'adesione indefettibile alla volontà divina, oppure un rifiuto irrecuperabile⁴³.

A questo punto, Tommaso suggerisce quello che potremmo chiamare il «paradosso dell'eccellenza razionale»: dato che il bene più alto coincide con ciò che è sommamente intelligibile, chi è eccellente dal punto di vista razionale dovrebbe essere nelle condizioni migliori per afferrarlo. Eppure, proprio chi possiede l'intelligenza più grande è esposto ad un particolare rischio, che potrebbe portare l'intelligenza eccellente a deviare irrimediabilmente da ciò che è razionale. Di fronte all'appello che viene da un bene ordinato secondo un criterio che supera il proprio intelletto, l'angelo può decidere di non (af)fidarsi, ma di giudicare esclusivamente secondo le proprie regole, che, essendo di ordine inferiore, si rivelano inadatte a valutare correttamente la volontà divina. Tommaso esprime al riguardo un concetto che non stupirebbe il lettore della *Fenomenologia dello Spirito* di Hegel: lo spirito è fatto per cercare il proprio compimento oltre se stesso, dirigendosi verso ciò che supera i confini della propria natura⁴⁴; l'angelo peccatore blocca questo movimento, compiacendosi della propria eccellenza e fermandosi ad essa, pensando così di poter giudicare con i propri criteri ciò li supera. In tal modo, la propria eccellenza razionale viene usata in modo unilaterale e fuorviante: il peccato angelico è innanzitutto un peccato di superbia⁴⁵.

Secondo una tale prospettiva, proprio chi è molto intelligente può allora correre il pericolo di chiudere in modo paradossale le porte della propria intelligenza e di farlo in modo addirittura irrevocabile, interpretando la propria eccellenza razionale come una forma di autosufficienza e di autoreferenzialità, che si oppone radicalmente all'apertura inesauribile, che deve invece caratterizzare l'intellettualità. Stando così le cose, non è strano che Tommaso simpatizzi per l'ipotesi di Gregorio Magno, secondo la quale il primo angelo peccatore potrebbe essere stato un cherubino, cioè un angelo dotato della scienza più perfetta: solo chi è eccellente, osserva Tommaso, può peccare di

⁴³ Cfr. Tommaso d'Aquino, QD *De malo*, q. 16, a. 5, *resp.*, pp. 305-306, ll. 290-340; Id., ST, I, q. 63, t. 5, aa. 4-7, pp. 129-136; Id., SCG, I, 3, c. 109, §§ 2046-2049, pp. 165-167.

⁴⁴ «Nomine beatitudinis intelligitur ultima perfectio rationalis seu intellectualis naturae : et inde est quod naturaliter desideratur, quia unumquodque naturaliter desiderat suam ultimam perfectionem. Ultima autem perfectio rationalis seu intellectualis naturae est duplex. Una quidem, quam potest assequi virtute suae naturae [...] Sed ultimam beatitudinem [...] non est aliquid naturae, sed naturae finis» (Tommaso d'Aquino, ST, I, q. 62, a. 1, *resp.*, t. 5, p. 110). «Felicitas consistit in applicatione ad id quod superius est» (*Ibid.*, q. 64, a. 1, ad 1, p. 140).

⁴⁵ Tommaso d'Aquino, QD *De malo*, q. 16, a. 3, *resp.*, p. 294, ll. 209-246; *Ibid.*, a. 6, *resp.*, pp. 310, ll. 281-286; Id., ST, I, q. 63, a. 2, t. 5, p. 123, *Ibid.*, a. 3, p. 126; Id., SCG, I, 3, c. 109, §§ 2047-2048, p. 166.

superbia, perché la superbia non potrebbe nemmeno nascere, senza la fondata consapevolezza della propria grandezza⁴⁶.

Queste riflessioni, pur venendo dall'angelologia, possiedono, a cascata, un evidente valore antropologico. Esse fungono certamente, in prima battuta, da esortazione morale, ricordando che l'eccellenza intellettuale deve andare insieme all'umiltà: il peccato degli angeli funge da monito, che rammenta l'esito fallimentare del tentativo di dissociarle. Tuttavia, la tesi di Tommaso sottende un assunto più fondamentale: il fatto che l'eccellenza intellettuale abbia un suo modo proprio di essere esposta all'ignoranza e all'agire non conforme a ragione significa che per Tommaso la questione dell'irrazionalità pratica ha sempre come base, anche per un'intelligenza straordinaria, l'assenza di una cognizione intuitiva del bene. Essa, però, non viene intesa da Tommaso come una mera mancanza di perfezione, ma viene presentata come condizione strutturale, affinché le creature che possiedono l'intelligenza *dispongano della propria eccellenza*, per poter aderire in modo libero e responsabile, e non meramente naturale, ai beni più alti.

2. Liberare la volontà: la linea di Giovanni Duns Scoto

Lo scarto tra l'universalità del bene colto dalla ragione e la particolarità dell'azione da farsi, richiamato da Tommaso per giustificare la possibilità dell'irrazionalità pratica, suppone certamente che la volontà libera sia il principio per cui mettiamo in atto le nostre azioni. Abbiamo visto che Tommaso coniuga però la libertà della volontà con l'intrinseca indeterminatezza dell'intelletto nel cogliere il bene. Ciò sembra supporre l'idea che noi, di fronte a un bene colto come conforme a ragione sotto ogni aspetto, non potremmo non perseguirlo, come dimostrano per un verso l'intrascendibilità della tensione alla felicità e per l'altro verso l'infallibilità naturale degli angeli.

È proprio su questo punto che la posizione di Giovanni Duns Scoto si presenta come un'alternativa teorica alla proposta di Tommaso. Per Scoto, infatti, è solo ed esclusivamente la libertà che rende la volontà capace di un potere autonomo di autodeterminazione, rispetto a qualsiasi oggetto, compresi i dettami della ragione. Se per Tommaso una piena evidenza razionale toglierebbe la possibilità di agire contro ragione, Scoto ribatte che, impostando le cose in questo modo, il senso per cui possiamo dire di essere liberi

⁴⁶ Cfr. Id., *In II Sent.*, d. 6, q. 1, a. 1, *sol.* et ad 1, p. 161; Id., *ST*, I, q. 63, a. 7, *resp.* et ad 1, t. 5, p. 136; Gregorio Magno, *Homiliae in Evangelia*, l. 2, hom. 34, § 7, pp. 305-306; Id., *Moralia in Job. Libri XXIII-XXXV*, l. 33, cc. 22-23, §§ 46-48, pp. 1663-1667.

verrebbe fortemente impoverito, perché si farebbe dipendere il potere di autodeterminazione della volontà da una certa limitatezza della ragione⁴⁷.

Questo non significa che Scoto voglia slegare la volontà libera dalla ragione o minimizzarne il legame. Nemmeno, ovviamente, si potrebbe dire che Scoto voglia fare della libertà una forza irrazionale: egli ha affermato con chiarezza che la volontà ha un'inclinazione naturale nei confronti del bene e del giusto, indicati dalla retta ragione⁴⁸ e che il carattere specifico della volontà sta nel fatto di avere la ragione come propria misura⁴⁹.

Non si può quindi pensare che Scoto voglia indebolire il legame tra volontà e ragione. Egli tiene però a sottolineare che se la volontà si autodetermina davvero e se è la libertà il suo carattere più proprio, allora non è possibile che un qualche oggetto possa presentarsi come irresistibile: ciò significherebbe affermare contraddittoriamente che qualcosa sia ad un tempo libero e determinabile secondo necessità⁵⁰. Per questo, per spiegare come la ragione concorra alla produzione dell'atto di volizione, Scoto ricorre emblematicamente a questo paragone: così come abbiamo bisogno dell'attività della sensibilità per pensare, ma non per questo il pensare si rivela come un'attività sensibile, così la cognizione delle cose è necessaria al volere, ma non entra nell'atto proprio della volontà, che si determina liberamente⁵¹.

⁴⁷ Così si esprime Scoto nella *Lectura*: «ratio non potest demonstrative syllogizare nisi de oppositis, sed de altero sophisticæ et defectivæ; igitur si propter talem indifferentiam in intellectu 'quia potest iudicare de oppositis', esset libertas in voluntate, sequeretur quod libertas esset in voluntate 'quia habet cognitionem defectibilem', et hæc misera esset libertas!» (Duns Scoto, *Lectura in librum secundum Sententiarum* (= *Lectura*), II, d. 25, q. un., 35, p. 239,9-14).

⁴⁸ Cfr. Duns Scoto, *Ordinatio*, II, d. 6, q. 2, 49, p. 49, ll. 381-387.

⁴⁹ Cfr. Duns Scoto, *Lectura*, II, d. 25, q. un., 79-80, pp. 256,11-257, 8; Id., *Reportata Parisiensia*, II, d. 25, q. un., 20, p. 128. Si vedano, al riguardo, le riflessioni di Todisco (2003: 337-338) e soprattutto di Ingham (2004).

⁵⁰ Cfr. Duns Scoto, *Quodlibet XVI*, 4-7, pp. 184-185, pp. 188-189; Id., *Reportata Parisiensia*, II, d. 25, p. 120 e pp. 128-129.

⁵¹ Cfr. Duns Scoto, *Reportata parisiensia*, II, d. 25, q. un., 16-18, pp. 125-126; *Ibid.*, 20, p. 128. Nella *Lectura*, egli si era opposto a chi, come Enrico di Gand, sembrava ridurre la cognizione dell'oggetto a *conditio sine qua non* dell'atto di volizione, senza riconoscere il suo ruolo di causa vera e propria, per quanto solo parziale. Cfr. Duns Scoto, *Lectura*, II, d. 25, q. un., 69, p. 253,1-10; Enrico di Gand, *Quodlibet*, I, q. 15, *sol.*, p. 93, ll. 58-61; Id., *Quodlibet X*, q. 9, ad 1, pp. 289-290, ll. 87-97; Id., *Quodlibet XII*, q. 26, *sol.*, pp. 151,80-156,18. In modo convincente, Stephen D. Dumont ha mostrato che nei *Reportata Parisiensia*, scritti certamente dopo la *Lectura*, Scoto esibisce un'attitudine più decisamente «volontarista», avvicinandosi, così, a quanto sosteneva Enrico di Gand, più di quanto gli studiosi degli ultimi decenni abbiano voluto ammettere. Cfr. Dumont (2001: 719-794). Non è questo il luogo per discutere nel dettaglio questo argomento, uno dei più controversi, all'interno della dottrina di

A partire da questo assunto, Duns Scoto prende le distanze da una delle idee più fondamentali dell'etica aristotelica, assunte e sviluppate da Tommaso d'Aquino: non è vero che l'uomo non può non voler essere felice. Chi sostiene una posizione, del genere, argomenta Scoto, ragiona sostanzialmente in questo modo:

il fine ultimo si presenta alla volontà come l'oggetto più perfetto;

non potendo trovare in un tale oggetto alcun difetto, la volontà non può vedere in esso alcun male;

non potendo riconoscere al fine ultimo alcuna ragione di male, la volontà non può che volerlo.

Scoto spiega di non avere alcuna difficoltà a concedere la prima e la seconda affermazione, ma ritiene che la terza contenga un salto logico: il fatto che sia impossibile provare repulsione per la felicità, dato che non le può essere attribuita alcuna ragione di male, non significa che essa debba necessariamente essere voluta. Ad esempio, il fatto che io non odi qualcuno o qualcosa, non significa allora che *ipso facto* lo ami. Tra il *velle* e il *nolle* c'è ancora lo spazio per il *non velle*, per il «non-movimento», per la mancanza di adesione⁵².

Scoto ricorda anche che, se chiamiamo «natura» il regno delle cose necessitate, dovremmo riconoscere che l'intelletto opera *per modum naturae*: infatti esso non ha la libertà di non pensare, quando ha di fronte a sé il proprio oggetto, né di non giungere a certe conclusioni, partendo da certe premesse⁵³.

Duns Scoto. Ci limitiamo ad osservare che, dal punto di vista della coerenza filosofica, probabilmente la posizione dei *Reportata* si armonizza molto di più con quanto stiamo vedendo sostenere da parte di Scoto, a proposito della libera volontà.

⁵² «Ista posset concedi, quod voluntas non potest resilire ab objecto, sive nolle objectum, in quo non ostenditur aliqua ratio mali, nec aliquis defectus boni, quia sicut bonum est objectum hujus actus, qui est *velle*, ita malum vel defectus boni, quod pro malo reputatur, est objectum hujus actus, qui est *nolle*, et non sequitur ultra, non potest nolle hoc, igitur necessario vult hoc, quia potest hoc objectum neque nolle, neque velle» (Duns Scoto, *Quodlibet* XVI, 5, p. 188). Vedi anche Id., *Reportata Parisiensia*, II, d. 25, pp. 128-129; Id., *Lectura*, II, d. 25, q. un., 74, pp. 255,1-8.

⁵³ Cfr. Duns Scoto, *Quaestiones super libros Metaphysicorum. Libri VI-IX*, I, 9, q. 15, 36, pp. 684,13-685,1; Id., *Lectura*, II, d. 25, 37, p. 240,16-18. Nell'ottica di Scoto, la volontà è certamente in grado di distogliere l'intelletto dal proprio oggetto, inducendolo a procedere in modo sofistico, ma può farlo solo «forzandolo», per così dire, dall'esterno. Cfr. su questo aspetto, l'analisi di Hoffmann (2004 : 494-496).

Poiché dunque la relazione tra l'intelletto e il proprio oggetto presenta i caratteri di una tale necessità, se noi cogliessimo infallibilmente il fine ultimo, allora, secondo Scoto, noi coglieremmo di necessità anche ciò che dobbiamo fare in funzione del fine e non potremmo che metterlo in atto. A quel punto, l'azione non conforme a ragione rimarrebbe senza una valida spiegazione⁵⁴.

Argomentando in questo modo, Scoto non vuole negare che ci possa anche capitare di sbagliare, senza per questo essere stati negligenti o in cattiva fede, ma intende sottolineare che non è sul campo dell'ignoranza e della debolezza della ragione che si gioca principalmente la partita dell'azione non conforme a retta ragione. Per Scoto, noi possiamo andare contro ciò che sappiamo essere giusto, per il semplice motivo che non è sufficiente vedere nitidamente la cosa migliore da fare, per metterla in atto. Per compiere l'azione conforme a ragione serve un amore gratuito e libero per essa, che il mero possesso della verità non basta a garantire.

3. Conclusioni: il nodo della volontà razionale

Leggendo Aristotele, ma non senza san Paolo e sant'Agostino, sia Tommaso d'Aquino che Duns Scoto interpretano la possibilità dell'irrazionalità pratica in termini non primariamente difettivi, ma prevalentemente *attivi*: noi siamo innanzitutto responsabili di ciò che facciamo; per fare il meglio, la volontà deve mettersi in movimento, *decidendosi* per l'adesione all'ordine della ragione. Per entrambi, dunque, noi possiamo compiere anche azioni non conformi a ragione, pur essendo esseri razionali, perché una tale possibilità è iscritta nelle condizioni stesse del libero esercizio della volontà.

Distinta dalla dimensione meramente conoscitiva, la volontà è però pur sempre, per entrambi, una potenza di natura razionale. Proprio questo fatto rende del tutto stringente, tanto per Tommaso, quanto per Scoto⁵⁵, la necessità di spiegare l'irrazionalità pratica: perché la volontà, essendo fatta per realizzarsi seguendo i dettami della retta ragione, segue ciò che razionale non è?

La risposta a questa domanda diventa allora la spia per comprendere meglio ciò che più genuinamente distingue le proposte di Tommaso e di Duns Scoto, ben al di là della contrapposizione tra un Tommaso «intellettualista» e uno Scoto «volontarista»⁵⁶: dato che entrambi, come speriamo di avere mostrato, affermano che la volontà è intrinsecamente legata alla ragione, allora la differenza sta piuttosto nell'indicazione di *come* questo legame si realizzi.

⁵⁴ Cfr. Duns Scoto, *Quodlibet* XVI, 4, p. 185.

⁵⁵ Si veda al riguardo Hoffmann (2004: 487-489).

⁵⁶ Cfr. ad esempio Landry (1922: 227-231).

Tommaso inserisce l'irrazionalità pratica nel quadro della natura discorsiva e inquisitiva della nostra ragione. Letto sotto questa luce, l'intellettualismo socratico appare incompleto, più che sbagliato e Tommaso stesso non manca di riconoscerlo: Socrate ha minimizzato, certo, la complessità della dimensione appetitiva umana, trascurando la spinta esercitata sulla volontà dalle passioni, ma aveva ragione, quando sosteneva che il fondamento intrinseco del nostro volere sta nella visione del bene⁵⁷.

Scoto ha messo invece in primo piano una certa nozione di volontà libera, in virtù della quale l'idea di un qualsiasi orientamento intrinseco e necessario, perfino quello alla felicità, segnalato da Aristotele e adottato da Tommaso, risulta inaccettabile. L'errore, il male, il peccato sono per Scoto una drammatica testimonianza della nostra libertà, che non è mai orientabile univocamente, a nessun livello: per questo, né una perfetta *scientia*, né il coglimento del fine ultimo rendono irresistibile l'assenso, perché la questione fondamentale che anima le nostre azioni è un problema di amore, di dedizione, di *cura*, più che di cognizione adeguata del bene⁵⁸. L'azione conforme a ragione si presenta allora come il frutto di quella verace e drammatica adesione che, secondo Duns Scoto, l'«intellettualismo» di certi suoi contemporanei impediva di mettere compiutamente a fuoco.

Come spiega in modo convincente François Loiret, Scoto sembra dunque aver operato una «denaturalizzazione», una «definalizzazione» e una «desubordinazione» della volontà rispetto all'intelletto⁵⁹. Ciò significa allora che la prospettiva filosofica di Scoto abbia anticipato il grande tema kantiano dell'autonomia della volontà? Non è questo il contesto per offrire nel dettaglio un confronto tra Scoto e Kant, tuttavia non possiamo non riconoscere che per Kant la volontà agisce in modo razionale quando agisce in modo autonomo, possedendo *in se stessa* la regola dell'agire. Per Scoto, non meno che per Tommaso d'Aquino, è proprio questa la condizione che non può realizzarsi nella nostra volontà: la misura e la regola razionale della volontà umana *non* sono in lei e non possono esserlo, altrimenti non ci sarebbe nessuna differenza tra il nostro volere e quello di Dio.

⁵⁷ «Aliqualiter verum est quod Socrates dixit, quod scientia praesente, non peccatur, si tamen hoc extendatur usque ad usum rationis in particulari eligibili. Sic igitur ad hoc quod homo bene agat, requiritur quod non solum ratio sit bene disposita per habitum virtutis intellectualis; sed etiam quod vis appetitiva sit bene disposita per habitum virtutis moralis» (Tommaso d'Aquino, ST, I-II, q. 58, a. 2, *resp.*, t. 6, p. 373).

⁵⁸ «Voluntas determinat potentiam rationalem, ita quod voluntas nullam determinationem capit a scientia nec ab alio» (Duns Scoto, *Lectura*, II, q. 25, q. un., 37, p. 240, ll. 23-25).

⁵⁹ Cfr. Loiret (2003: 50-283).

Sarebbe quindi un errore – al quale il *cliché* dello Scoto irriducibile volontarista non è estraneo – proiettare anacronisticamente sulla nozione scotista di volontà quella moderna, in particolare quella kantiana. Piuttosto, dovremmo dire che Duns Scoto, per difendere il carattere intrinsecamente libero della nostra volontà, rifiuta di vincolare l'esercizio della nostra volontà alla «qualità» della nostra cognizione del bene, come Tommaso ha invece ritenuto necessario fare. Vale la pena notare che Tommaso, in virtù di un tale vincolo, suggerisce non solo e non tanto che la libertà abbia dei limiti, ma che ce li abbia l'irrazionalità pratica: quello per l'errore, per il peccato, per l'incapacità della volontà di dirigere le passioni verso i beni razionali può essere uno spazio più o meno ampio, a seconda soprattutto dell'impegno che noi mettiamo nella retta valutazione delle cose e nella costruzione delle virtù; tuttavia, non potrà essere uno spazio illimitato, totalmente indeterminato.

Così, il confronto tra le prospettive di Tommaso e di Scoto quanto al tema dell'irrazionalità pratica ci interpella su una questione essenziale, che non è solo di ordine storico: è davvero possibile considerare in modo cogente la volontà come una potenza intrinsecamente razionale (come entrambi i nostri autori fanno, al di là delle etichette di «volontarismo» e «intellettualismo»), se non si indica un qualche vincolo all'irrazionalità pratica, quali sono, nell'ottica di Tommaso, l'intrinseca tensione alla felicità e la visione non intuitiva del bene? Il punto decisivo stava e sta ancora, a nostro avviso, nella necessità di indicare un «limite di salvaguardia» che, senza cancellare la possibilità dell'irrazionalità pratica, permetta comunque di inquadrarla e di circoscriverla.

Bibliografia

Autori antichi e medievali

Agostino, *De libero arbitrio*, cura et studio W. M. Green, Turnhout, Brepols, 1970, pp. 211-321.

Aristotele, *Opera Omnia*, ex recensione I. Bekkeri, Academia Regia Borussica, Berlin, Reimer, 1831-1836.

Averroè, *Commentarium Magnum in Aristotelis De Anima Libros*, ed. F. Stuart Crawford, Cambridge, The Mediaeval Academy of America, (Mass.) 1953.

- Gregorio Magno, *Homiliae in Evangelia*, cura et studio R. Étaix, Turnhout, Brepols, 1999.
- Gregorio Magno, *Moralia in Job. Libri XXIII-XXXV*, cura et studio M. Adriaen, Turnhout, Brepols, 1985.
- Giovanni Duns Scoto, *Lectura in librum secundum Sententiarum* [dd. 7-44], in *B. Ioannis Duns Scoti Opera Omnia*, t. XIX, studio et cura Commissionis Scotisticae, Città del Vaticano, Typis Vaticanis, 1993.
- Giovanni Duns Scoto, *Ordinatio* [dd. 4-44], in *B. Ioannis Duns Scoti Opera Omnia*, t. VIII, studio et cura Commissionis Scotisticae, Città del Vaticano, Typis Vaticanis, 2001.
- Giovanni Duns Scoto, *Reportata Parisiensia*, in *B. Duns Scoti Opera Omnia*, t. XXIII, editio nova juxta editionem Waddingi XII tomos continentem a patribus franciscanis de observantia accurate recognita, Paris, Vivès, 1894.
- Giovanni Duns Scoto, *Quaestiones Quodlibetales XIV-XXI*, in *B. Duns Scoti Opera Omnia*, t. XXVI, editio nova juxta editionem Waddingi XII tomos continentem a patribus franciscanis de observantia accurate recognita, Paris, Vivès, 1895.
- Giovanni Duns Scoto, *Quaestiones super libros Metaphysicorum. Libri VI-IX*, in *B. Ioannis Duns Scoti Opera Philosophica*, t. IV, ed. R. Andrews, G. Etzkorn, G. Gàl, R. Green, F. Kelley, G. Marcil, T. Noone, R. Wood, The Franciscan Institute, St. Bonaventure (N. Y.), St. Bonaventure University, 1997.
- Enrico di Gand, *Quodlibet I*, in *Henrici de Gandavo Opera Omnia*, V, ed. R. Macken, Leuven-Leiden, Leuven University Press-Brill, 1979.
- Enrico di Gand, *Quodlibet X*, in *Henrici de Gandavo Opera Omnia*, XIV, ed. R. Macken, Leuven-Leiden, Leuven University Press-Brill, 1981.
- Enrico di Gand, *Quodlibet XII*, in *Henrici de Gandavo Opera Omnia*, XVI, ed. R. Macken, Leuven, Leuven University Press, 1987.
- Isaac Israeli, *De definicionibus*, éd. J. T. Munckle, «Archives d'Histoire Doctrinale et Littéraire du Moyen-Âge» 11 (1937-38), pp. 299-340.

- Ovidio, *Metamorfosi*, t. 4, a cura di E. J. Kennedy, testo critico basato sull'edizione di R. Tarrant, Milano, Fondazione Valla - A. Mondadori, 2011.
- Ps.-Dionigi l'Areopagita: P. Chevallier et al., *Dionysiaca. Recueil donnant l'ensemble des trad. latines des ouvrages attribués au Denys de l'Aréopage etc.*, Paris, Desclée de Brouwer, 1937 [rist. anast. Stuttgart-Bad Cannstadt, Frommann-Hoolzbog, 1989].
- Tommaso d'Aquino, *Opera omnia iussu impensaue Leonis XIII P. M. edita*, cura et studio Fratrum Praedicatorum, Roma, 1882 et sqq.
- Tommaso d'Aquino, *Scriptum super libros Sententiarum magistri Petri Lombardi episcopi Parisiensis*, libri 1-2, ed. P. Mandonnet, Paris, Lethielleux, 1929.
- Tommaso d'Aquino, *Scriptum super libros Sententiarum magistri Petri Lombardi episcopi Parisiensis*, liber 3, 2 voll., ed. M. F. Moos, Paris, Lethielleux, 1956.
- Tommaso d'Aquino, *In quattuor libros Sententiarum*, in *S. Thomae Aquinatis Opera Omnia*, v. 1, curante R. Busa, Stuttgart-Bad Cannstadt, Frommann-Hoolzboog, 1980.
- Tommaso d'Aquino, *Liber de veritate catholicae Fidei contra errores infidelium seu Summa contra Gentiles*, tt. 2-3, ed. P. Marc - C. Pera - P. Caramello, Torino-Roma, Marietti, 1961.
- Tommaso d'Aquino, *In librum De causis Expositio*, cura et studio C. Pera, Torino-Roma, Marietti, 1955.
- Tommaso d'Aquino, *In librum Beati Dionysii De divinis nominibus Expositio*, cura et studio C. Pera, P. Caramello, C. Mazzantini, Torino-Roma, Marietti, 1950.
- Tommaso d'Aquino, *Quaestiones disputatae de potentia*, in *S. Thomae Aquinatis Quaestiones disputatae*, t. 2, ed. P. M. Pession, Torino-Roma, Marietti 1965¹⁰, pp. 1-276.

Studi

- Barad, J. (1998), "Aquinas Assent/Consent Distinction and the Problem of Akrasia", *New Scholasticism*, 62, pp. 98-111.
- Bazán, B. C. (2010), "On Angels and Human Beings: did Thomas Aquinas succeed in demonstrating the Existence of Angels?", *Archives d'Histoire Doctrinale et Littéraire du Moyen Âge*, 77, pp. 47-85.
- Bettoni, E. (1966), *Duns Scoto filosofo*, Milano, Vita e Pensiero.
- Charlton, W. (1988), *Weakness of Will. A Philosophical Introduction*, Oxford, Blackwell.
- Collins, J. (1947), *The Thomistic Philosophy of the Angels*, Washington D. C. The Catholic University of America Press.
- Dumont, S. D. (2001), "Did Duns Scotus Change His Mind on the Will?", in J. A. Aersten, K. Emery Jr., A. Speer (a cura di), *Nach der Verurteilung von 1277. Philosophie und Theologie an der Universität von Paris im letzten Viertel des 13. Jahrhunderts. Studien und Texte*, Berlin-New York, De Gruyter, pp. 719-794.
- Elders, L. J. (1993), *The Metaphysics of Being of St. Thomas Aquinas in a Historical Perspective*, Leiden-New York-Köln, Brill.
- Dahl, N. O. (1984), *Practical reason, Aristotle and the Weakness of the Will*, Minneapolis, University of Minnesota.
- Doolan, G. T. (2012), "Aquinas on the Demonstrability of Angels", in T. Hoffmann (éd.), *A Companion to Angels in Medieval Philosophy*, Leiden-Boston, Brill, pp. 13-44.
- Forlivesi, M. (1996), "I rapporti tra intelletto e volontà nell'opera di Tommaso d'Aquino", *Divus Thomas*, 99, pp. 222-258.
- Gilson, É. (1989), *Le Thomisme. Introduction à la philosophie de saint Thomas d'Aquin*, Paris, Vrin.
- Henry, J. (1925), "L'imputabilité de l'erreur d'après saint Thomas d'Aquin", *Revue Néo-Scholastique de Philosophie*, 27, pp. 225-242.

- Hoffmann, T. (2004), “L’«Akrasia» selon Duns Scot”, in O. Bulnois, E. Krager, J.-L. Solère, G. Sondag (a cura di), *Duns Scot à Paris, 1302-2002. Actes du Colloque de Paris, 2-4 septembre 2002*, Turnhout, Brepols, pp. 487-516.
- Hoffmann, T. (ed.) (2008), *Weakness of Will from Plato to Present*, Washington D. C., The Catholic University of America.
- Hoffmann, T., Müller, J., Perkams, M. (a cura di) (2006), *The problem of Weakness of Will in Medieval Philosophy*, Louvain, Peeters.
- Ingham, M. B. (2004), “La genèse de la volonté rationnelle: de la Lectura à la Reportatio II, 25”, in O. Boulnois, E. Krager, J.-L. Solère, G. Sondag (a cura di), *Duns Scot à Paris, 1302-2002. Actes du Colloque de Paris, 2-4 septembre 2002*, Turnhout, Brepols, pp. 409-423.
- Kainz, H. P. (1972), *Active and Passive Potency in Thomistic Angelology*, The Hague, Nijhoff.
- Kent, B. (1995), *Virtues of the Will. The Transformation of Ethics in the Late Thirteenth Century*, Washington D. C., The Catholic University of America Press.
- Laghi, M. (2000), “Passio e “passione” nella letteratura tomista. Riflessioni in merito allo Status Quaestionis”, *Divus Thomas*, 103, pp. 59-92.
- Landry, B. (1922), *Duns Scot*, Paris, Alcan.
- Legrand, J. (1946), *L’univers et l’homme dans la philosophie de saint Thomas*, t. 2, Bruxelles-Paris, L’Édition Universelle-Desclée de Brouwer.
- Loiret, F. (2003), *Volonté et infini chez Duns Scot*, Paris, Kimé.
- Lottin, O. (1949), “La nature du péché d’ignorance depuis le XII^e siècle jusqu’au temps de saint Thomas d’Aquin”, in Id., *Psychologie et morale aux XII^e et XIII^e siècles*, t. 3, seconde partie, Louvain-Gembloux, Abbaye du Mont César-Duculot, pp. 9-51.
- Maritain, J. (1984), “Descartes ou l’incarnation de l’ange”, in Id., *Trois Réformateurs. Luther – Descartes – Rousseau. Œuvres complètes*, v. III,

- édition publiée par le Cercle d'Études Jacques et Raïssa Maritain, Fribourg (Suisse)-Paris, Éditions Universitaires-Saint Paul, pp. 485-521.
- Meyer, M. (1994), "Le problème des passions chez Saint Thomas d'Aquin", *Revue internationale de philosophie*, 189, 48, pp. 363-374.
- Pégahaire, J. (1936), *Intellectus et ratio selon S. Thomas d'Aquin*, Paris, Vrin.
- Pegis, A. C. (1940), "In Umbra Intelligentiae", *The New Scholasticism*, 14, pp. 146-180.
- Pinckaers, S. (1997), *Appendice II à Saint Thomas d'Aquin, Somme Théologique. Les Actes Humains. Tome premier. 1a-2ae, Questions 6-17*. Traduction française par H.-D. Gardeil, notes et appendices par S. Pinckaers, Paris, Les Éditions du Cerf, pp. 405-449.
- Putallaz, F. X. (1991), *Le sens de la réflexion chez Thomas d'Aquin*, Paris, Vrin.
- Robiglio, A. A. (2002), *L'impossibile volere. Tommaso d'Aquino, i tomisti e la volontà*, Milano, Vita e Pensiero.
- Rousselot, P. (1924²), *L'intellectualisme de saint Thomas (L'intellectualismo di san Tommaso)*, trad. it. di M. Pastrello, Milano, Vita e Pensiero.
- Roland-Gosselin, M. D. (1923), *La théorie thomiste de l'erreur*, in *Mélanges thomistes*, Paris, Vrin, pp. 253-274.
- Saarinen, R. (1994), *Weakness of the Will in Medieval Thought. From Augustine to Buridan*, Leiden-New York-Köln, Brill.
- Scribano, E. (2006), *Angeli e beati. Modelli di conoscenza da Tommaso a Spinoza*, Roma-Bari, Laterza.
- Stegman, T. D. (1989), "Saint Thomas Aquinas and the problem of Akrasia", *The Modern Schoolman*, 66, pp. 117-128.
- Stroud, S., Tappolet, C. (2003), *Weakness of the Will and Practical Irrationality*, Oxford, Clarendon.
- Suarez-Nani, T. (2002), *Les Anges et la Philosophie. Subjectivité et fonction cosmologique des substances séparées à la fin du XIII^e siècle*, Paris, Vrin.

- Suarez-Nani, T. (2006), "Tommaso d'Aquino e l'angelologia: ipotesi sul suo significato storico e la sua rilevanza filosofica", in A. Ghisalberti, A. Petagine, R. Rizzello (a cura di), *Lecture e interpretazioni di Tommaso d'Aquino oggi: cantieri aperti. Atti del Convegno Internazionale di studio (Milano, 12-13 settembre 2005)*, Torino, Quaderni di Annali Chieresi, pp. 11-29.
- Todisco, O. (2003), *Lo stupore della ragione. Il pensare francescano e la filosofia moderna*, Padova, Messaggero.
- Vecchio, S. (2002), *Introduzione a Tommaso d'Aquino, Le passioni dell'anima*. Dalla Somma di Teologia, I, II, questioni 22-48, traduzione e introduzione di S. Vecchio, Firenze, Le Lettere, pp. 5-18.
- Vecchio, S. (2009), "Il piacere da Abelardo a Tommaso", in C. Casagrande, S. Vecchio (a cura di), *Piacere e dolore. Materiali per una storia delle passioni nel Medioevo*, SISMEL-Edizioni del Galluzzo, pp. 67-86.
- Vernier, J. M. (1986), *Les Anges chez Saint Thomas d'Aquin*, Paris, Nouvelles Éditions Latines.
- Williams, T. (1997), "Reason, Morality and Voluntarism in Duns Scotus: a Pseudo-Problem Dissolved", *The Modern Schoolman*, 74, pp. 73-94.
- Williams, T. (1998), "The Unmitigated Scotus", *Archiv für Geschichte der Philosophie*, 80, pp. 162-181.

Truth against Reason? Siger of Brabant and the Question of the Eternity of the World

Ercole Erculei
Rheinische Friedrich-Wilhelms-Universität Bonn
erculei@uni-bonn.de

1. Prolegomena: *fides quaerens intellectum* and the traumatic reception of Aristotle's philosophy

Directly connected with every possible discussion of the capabilities and boundaries of reason, the problem of its relation with faith has certainly represented a *vexata quaestio* in the history of European philosophy, science and religion. Furthermore, this issue does not lack in topicality, if one considers the increased use/abuse of religious topics in current political and social debates or the abundant polemics in the mass media between “supporters” of a religious world view and those of a “scientific” one – unfortunately often conducted at a very ideological level –, and even the fact that, in the last 15 years, two popes – John Paul II and Francis I – have dedicated specific encyclicals to the attempt of sketching a framework for fruitful coexistence, dialogue and mutual strengthening between faith and reason based upon the oneness of the Truth¹.

If the issue of the distinction between “just believing” and “actually knowing”, between *doxa* and *epistêmê*, has always belonged to the fundamental questions and tasks of philosophical research since Plato and Aristotle, the urgency of a specific thematization of the relation between religious faith and philosophy/science (as the highest expression of natural reason in its efforts toward comprehending and explaining the world) increased enormously during the affirmation in the Roman Empire of a religion such as Christianity:

¹ Cfr. Ioannes Paulus II, Littera Encyclica “*Fides et ratio*”; Franciscus, Littera Encyclica “*Lumen Fidei*”.

which on the one hand was based on a sacred book containing revealed Truths, but on the other strongly claimed universal validity, i. e. that universality which is the distinguishing feature of the conclusions of rational argument.

Without going into detail, I would like to briefly mention two possible fundamental positions regarding the relation between faith and natural reason. They are very helpful as a historic background for the next introduction of the figure of Siger of Brabant and his dealing with the topic of the eternity of the world. A first model is a harshly “incompatibilism with conflict” between faith and reason²: it not only implies that the conclusions of the one are different/incompatible with the conclusions of the other, but also that they are rivals at the same level, in a “fight” where the “winner” will emerge as the *only* source of truth, while every bond with the truth will be denied of the other, “losing” part. Often mentioned as emblematic examples of this model (with faith as the “winner”) are the famous passage of first letter by St. Paul to the Corinthians, where he stresses that the revealed Truth contained in the sacred book is «folly» (*môria*) from the perspective of the «wisdom (*sophia*) of the world», while all human, natural, earthly wisdom is folly for God’s wisdom³, or the position of Tertullian, who polemically points out that a philosopher and a Christian have as much in common as the «friend and the enemy of error, the corrupter and the restorer and teacher of the truth, its thief and its custodian»⁴. With his rhetorical taste for disconcerting provocations, the Christian apologist goes as far as to speak of the credibleness of the death of God’s Son because of its inconceivableness, of the certainty of His resurrection because of its impossibility, a position that the posterity has summarized in the formula *credo quia absurdum*⁵.

Nevertheless, this model of incompatibilism with its denial of the capabilities of reason in order to discover the truth has been never dominant in Christian thought, neither in the Patristic nor during Medieval period the – a point that would not need to be stressed, if the hackneyed myth of the “dark” and fideistic, irrational Middle Ages (or of the Christian religion at all) was

² Cfr. Swindal (2013).

³ Cfr. 1 *Cor.*, 1.19-25.

⁴ Tertullianus, *Apologeticum*, XLVI, 18, p. 162: Adeo quid simile philosophus et Christianus, Graeciae discipulus et caeli, famae negotiator et vitae, verborum et factorum operator, et rerum aedificator et destructor, amicus et inimicus erroris, veritatis interpolator et integrator et expressor, et furator eius et custos? [trans. EE].

⁵ Id., *De Carne Christi*, V, 4, ll. 27-29, p. 881: Et mortuus est dei filius; credibile est, quia ineptum est. Et sepultus resurrexit: certum est, quia impossibile.

not still quite diffused today – especially among non-scholars – despite several decades of excellent historical research on the philosophical, scientific, technic, etc. contributions during this epoch. In this regard, it is worth considering Origen’s position in the prologue of his *De Principiis*, where he recognizes the importance and goodness of the exercise of natural reason, legitimizing and encouraging research into the *rationes* of divine sentences in the Holy Scripture by those humans who are well-endowed with «language, wisdom and knowledge»⁶. Through the prophets, the apostles and the Holy Scripture, God primarily wanted to make clear the basic ethical principles and rules for a moral conduct of one’s own life; therefore, this practical level of Scripture must be quickly understandable for all hearers, even those «who are capable of receiving nothing more than an exhortation to believe»⁷. Nevertheless, they are not the only addressees of the Holy Scripture, which cannot be reduced to its literal meaning. The plurality of enigmas, metaphors, parables, obscure passages, etc. in Scripture not only shows the necessity of a multilevel allegorical interpretation of it, but it actually urges the *logos* to go beyond the level of the symbols in order to try to grasp, to *understand* the truth⁸. Christians cannot simply limit their announcement of the “good news” into an invitation to believe; those among them who can should try to use the demonstrative approach, through the dialectical method of «questioning and answering»⁹.

Several passages of other Christian authors can be legitimately connected to this compatibilistic Origenian strongly compatibilistic model¹⁰. For example, Lactantius stresses the impossibility of a *true* religion without wisdom, as well as *true* wisdom without religion¹¹, «because it is the same God, who ought to be understood, which is the part of wisdom, and to be honored, which is the part of religion»¹², and because «the fountain of wisdom and religion is God; and if these two streams shall turn aside from Him, they must be dried up»¹³. The metaphor of a singular divine *fons* as the ground of an intrinsic accordance between reason and faith can be found in John Scotus Eriugena, who does not hesitate to recommend that authority should not dissuade us from

⁶ Origenes, *De principiis*, lib. 1, praef., 44-57, par. 3, pp. 78-80.

⁷ Id., *Contra Celsum*, VI, 10,26-33, p. 202.

⁸ *Ibidem*. VII, 10,16-25, pp. 36-38.

⁹ *Ibidem*. VI, 10,32-33, p. 202.

¹⁰ Cfr. Swindal (2013).

¹¹ Lactantius, *Divinae Institutiones*, III, 11, (2), p. 204; Cfr. Id. *Epitome Divinarum Institutionum*, 36, 4-5, p. 712.

¹² Id., *Divinae Institutiones*, IV, 4, 2, p. 282.

¹³ *Ibidem*. IV, 4, (4), p. 282.

what we are *rationaly* persuaded of¹⁴. Indeed, for him, true philosophy is true religion, while true religion also cannot be anything other than true philosophy. The idea is an explicit reference to Augustine's *De vera religione*¹⁵, an author whose distance in essential topics from Eriugena and Origen (e.g. the destiny of the soul after the death, the existence of the hell, the cosmology and eschatology, the relations within the three persons of the Trinity, the concept of freedom and its relation with divine grace, etc.) is well known and need not be stressed here. Nevertheless, the Bishop of Hippo as well as such an "Augustinian" author as Anselm of Canterbury also describe, defend and show – directly through their works – the strong desire to *understand* what they as Christians already believe¹⁶, criticizing the negligence of those who do not try to understand what they already believe¹⁷, or the fideistic hate against rationality, which is actually God's gift for us and our distinguishing feature in comparison to the animals¹⁸.

This compatibilistic-cooperative attitude and project have been summarized in the famous formulas «*crede ut intelligas*» (Augustine)¹⁹, «*fides quaerit, intellectus invenit*» (Augustine)²⁰, «*credo ut intelligam*» (Anselm)²¹, and – the most emblematic – «*fides quaerens intellectum*» (Anselm)²². They

¹⁴ Iohannes Scotus Eriugena, *Periphyseon*, lib. I, ll. 2973-2977, p. 96: Nulla itaque auctoritas te terreat ab his quae rectae contemplationis rationabilis suasio edocet. Vera enim auctoritas rectae rationi non obsistit, neque recta ratio uerae auctoritati. Ambo siquidem ex uno fonte, diuina uidelicet sapientia, manare dubium non est.

¹⁵ Id., *De divina praedestinatione*, cap. 1,16-18, p. 5: Conficitur inde ueram esse philosophiam ueram religionem conuersim que ueram religionem esse ueram philosophiam. Cfr. Augustinus Hipponensis, *De vera religione*, V, 8,13-14, p. 193: [...] non aliam esse philosophiam, id est sapientiae studium, et aliam religionem [...].

¹⁶ Cfr. Augustinus Hipponensis, *De libero arbitrio*, II, cap. 2, 5.16, l. 65, p. 239: Sed nos id quod credimus nosse et intellegere cupimus; Id., *Contra Academicos*, III, 20, 43, ll. 21-22, p. 61: [...] quid sit uerum non credendo solum sed etiam intellegendo apprehendere impatienter desiderem. Cfr. Anselmus Cantuariensis, *Cur Deus Homo, Commendatio operis ad Urbanem Papam II*, pp. 39-41; Id., *Proslogion, Prooemium*, pp. 93-94.

¹⁷ Anselmus Cantuariensis, *Cur Deus Homo*, I, 2, p. 48,16-18: Sicut rectus ordo exigit ut profunda Christianae fidei prius credamus, quam ea praesumamus ratione discutere, ita negligentia mihi videtur, si, postquam confirmati sumus in fide, non studemus quod credimus intellegere.

¹⁸ Augustinus Hipponensis, *Epistulae II, Epist. 120*, 3, p. 706,18-21: Absit namque ut hoc in nobis Deus oderit, in quo nos reliquis animantibus excellentiores creauit. Absit, inquam, ut ideo credamus, ne rationem accipiamus sive quaeramus; cum etiam credere non possemus, nisi rationales animas haberemus.

¹⁹ Id., *Sermones de vetero testamento, sermo XLIII*, 7-9, ll. 141-185, pp. 511-512.

²⁰ Id., *De Trinitate*, lib. XV, cap. 2,27, p. 461.

²¹ Anselmus Cantuariensis, *Proslogion*, cap. 1, p. 100,18.

²² *Ibidem*. p. 94,7.

cannot be misunderstood as a sign of mistrust of the capabilities of reason and its relation to truth: on the contrary, they imply a strong optimism concerning the power of natural reason and its boundaries, which are actually so far that it can include as own research subjects topics such as the demonstration of the immortality of the soul, the demonstration of the existence of God, or the investigation simply on the basis of pure reason into dogmas such as the Trinity or the Incarnation.

An eventual dissonance between the conclusions of rational investigation and the articles of faith cannot become an excuse for negligence or surrendering to fideism: on the contrary, it should reflect the motivation and impulse for a *better* investigation, since there is only one truth and the source of the articles of faith is nothing other than God/Christ, who is also the *Logos*, the Truth²³. Consequently, a rational investigation that apparently leads to an idea in contrast with the truth of faith *must* be wrong, *must* be a bad philosophical investigation, an arrogant, immature, failing attempt to understand what one believes: the option of *questioning the truth* of the articles of faith due of the results of philosophical inquiry is out of the question²⁴.

Today, such a “unshakable” presupposition of the truth of the articles of faith, even as a signal of the insufficiency of a philosophical enterprise, would probably appear as the employment of an external, heterogeneous criterion in the evaluation of the conclusions of rational investigation: our habitual perspective is the systematic distinction between religion/theology on the one hand and philosophy/science on the other as different and *autonomous*

²³ Cfr. John 14.6: «Egô eimi hê hodos kai hê alêtheia kai hê zôê»

²⁴ Cfr. Augustinus Hipponensis, *De gratia et libero arbitrio*, cap. I, pp. 881-882: Itaque, dilectissimi, ne vos perturbet hujus quaestionis obscuritas, moneo vos primum, ut de iis quae intelligitis, agatis Deo gratias: quidquid est autem quo pervenire nondum potest vestrae mentis intentio, pacem inter vos et charitatem servantes, a Domino ut intelligatis orate; et donec vos ipse perducatur ad ea quae nondum intelligitis, ibi ambulate quo pervenire potuistis. [...] Ambulando quippe in quod pervenimus, et quo nondum pervenimus pervenire poterimus, Deo nobis revelante si quid aliter sapimus, si ea quae iam revelavit non relinquamus.

Cfr. Anselmus Cantuariensis, *Epistola De Incarnatione Verbi*, I, pp. 6,10-7,4; p. 8,2-4; p. 10,13-16: Nullus quippe Christianus debet disputare, quomodo quod catholica ecclesia corde credit et ore confitetur non sit; sed semper eandem fidem indubitanter tenendo, amando et secundum illam vivendo humiliter quantum potest quaerere rationem quomodo sit. Si potest intelligere, deo gratias agat; si non potest, non immittat cornua ad uentilandum sed submittat caput ad venerandum. [...] illi non habent fidei firmitatem, qui quoniam quod credunt intelligere non possunt, disputant contra eiusdem fidei a sanctis patribus confirmatam veritatem. [...] Haec dixi ne quis, antequam sit idoneus, altissimas de fide quaestiones praesumat discutere; aut si praesumpserit, nulla difficultas aut impossibilitas intelligendi valeat illum a veritate cui per fidem adhaesit excutere.

fields/disciplines, with different principles, different subjects, different aims, different criterions of validity of their assertions, etc. However, we cannot forget that it did not represent the point of view of an author such as Augustine and his *rationalistic* understanding of the faith/the *auctoritas* and the figure of Christ/*Logos*/Truth²⁵, a decisive element that should always be considered when approaching authors prior to the 13th-14th centuries. Indeed, these centuries were the time of the «final separation of the provinces of the philosophy and the theology»²⁶, a consequence of that «trauma», caused by the reception of the whole Aristotelian *corpus*²⁷ and which has found in the figure of Siger of Brabant one of its most important protagonists.

The rediscovery of the whole Aristotelian *corpus*, as well as the confrontation with the main works of Jewish and Muslim philosophy, represented not only a powerful arsenal of arguments *pro* or *contra* concerning a huge quantity of metaphysical, physical, cosmological, and ethical topics but quickly imposed new and higher intellectual standards, e.g. regarding the investigation of the natural world (that could not simply be reduced to a series of moral symbols, but rather became the field of a reality whose phenomena could be well explained through constant, non even necessary, causal connections), or in regard to philosophical terminology and the conceptual instruments, regarding the concept of science, of demonstrative knowledge and the distinction of different philosophical disciplines²⁸. Under the impact of this Aristotelian avalanche, the fundamentals of the compatibilistic model of *fides quaerens intellectum* and the patristic and medieval ideal of a *sapientia christiana* – in which the sacred Scripture, philosophy and the seven artes liberales (even the *artes mechanicae* in Hugo of S. Viktor)²⁹ constituted the organic whole of Christian wisdom, of the Truth, without any irreconcilable internal contrast or splitting – had been radically shaken. The philosopher of Stagira indeed offered a series of well-grounded arguments for theses, which were the exact negation of some essential assertions of the Christian faith, and whose validity or falsity now had to be questioned/discussed *within* the boundaries of a single discipline, i.e. on the basis of the principles and meth-

²⁵ Augustinus Hipponensis, *De Trinitate*, lib. XIII, cap. 19, 24, ll. 50-51, cit., p. 416: Scientia ergo nostra Christus est, sapientia quoque nostra idem Christus est. Cfr. Horn (1995: 127-28).

²⁶ Gombocz (1997: 232).

²⁷ Vanni Rovighi (1978: 41).

²⁸ Cfr. Bianchi (1990: 5sqq).

²⁹ Cfr. Hugo de Sancto Victore, *Didascalicon (De studio legendi)*, I, cap. IV, cap. V, cap. VIII, pp. 11-16.

odologies of the single discipline to which the single “dissonant” thesis actually belonged: for instance, 1) the thesis of a soul as the first *entelecheia*/form of the corruptible body which therefore seems to be mortal; 2) the thesis that the world is eternal.

In the present paper, I will focus my attention on the second thesis, and particularly how Siger of Brabant dealt with that Aristotelian arguments *pro aeternitate mundi*, which represented such a thorny question for the theologians and philosophers of the 13th century, especially concerning the relation of reason and faith, that they can be legitimately compared with the impact of Copernicus heliocentric theory in 16th and 17th centuries or the Darwinian theory of evolution in the 19th and 20th centuries. E. Grant writes in this sense: «[b]ecause Aristotle based his natural philosophy on the firm conviction that the world is eternal, it posed a major threat to the creation account in *Genesis*»³⁰. What faith asks to believe as true (that the world was created by God *ex nihilo* and it will have an end) now seems not to be “explainable” through reason, but rather contradicted by it: natural reason on an Aristotelian basis cannot “demonstrate” that the world had a beginning; on the contrary, it seems to show that it is eternal. Does reason lead – without committing mistakes – to falsehood? Does rational validity not imply truth? Is reason not a source of truth, but rather of falsity? Or is the truth irrational? Does the Christian religion impose believing the irrational and absurd? Shall one subsequently adopt the position of the *credo quia absurdum*? Shall the philosophers «take captive the intellect in the respect of Christ», as firmly wanted by the Bishop of Paris Étienne Tempier in his condemnation³¹? Has the *fides-quaerens-intellectum*-model and the ideal of *sapientia christiana* been irretrievably jeopardized by the Aristotelian plurality of circumscribed yet autonomous disciplines and – among them – of physics in comparison with revealed theology?

With these epochal questions in the background, I can finally proceed with the introduction of Siger’s reception of Aristotle’s eternalistic arguments.

³⁰ Grant (1996: 74).

³¹ Cfr. H. Denifle-A. Chatelain (edd.), “Opiniones ducentae undeviginti [...] a Stephano episcopo Parisiensi de consilio doctorum Sacrae Scripturae condemnatae 1277 [...]”, in: *Charitularium Universitatis Parisiensis*, t. I, n. 473, Prop. 18, Paris, Delalain, 1889, p. 544: [...] etiam philosophus debet captivare intellectum in obsequium Christi [trans. EE].

2. Four main eternalistic arguments

In the course of his intellectual production, Siger has the opportunity on several occasions to deal with Aristotle's eternalistic arguments, even in non-cosmological/physical contexts. Their constant presence in the work of the Brabantine philosopher serves not only to connect with the already mentioned trauma of the rediscovery of Aristotelian physics/cosmology, which explicitly contradicts such a main pillar of the Christian faith as the *creatio ex nihilo*, but also with the other minefield of the 13th century: a psychology and anthropology constructed on the basis of Aristotle's *De anima*, an issue about which Siger, in his famous q. 9 of the *Quaestiones in tertium de anima*, sustains the scandalous thesis of the numerical oneness of the *intellectus possibilis*. As acutely remarked by A. Petagine,

non è nel *De anima* che Aristotele ha guadagnato le categorie più adatte a definire la natura intellettuale: gli strumenti concettuali adeguati per definire la natura e le operazioni dell'intelletto non sono infatti quelli con cui si definiscono le altre anime, bensì quelli con cui si definiscono le sostanze prodotte direttamente dalla Causa Prima.

Ciò significa che per comprendere e definire la natura dell'intelletto in sé, si deve piuttosto venire a contatto con gli argomenti contenuti nel *De coelo* o nella *Fisica*, uscendo dalle anguste maglie del *De anima*³².

Hence, Siger directly uses the Aristotelian arguments *pro aeternitate mundi* to demonstrate the eternity of the intellect, and such an eternity – or, to be more precise, perpetuity – becomes the basis for one of Siger's demonstrations of the numerical oneness of the *intellectus possibilis*: something eternal does not need a «*multiplicatio individuorum*» in order to “save” its «*esse specificum*»³³.

Consequently, several interesting passages using Aristotelian materials about the eternity of the world can be found in Siger's version of arguments *pro aeternitate intellectus/intelligentiae*³⁴. Furthermore, it is actually one of these passages *pro aeternitate intellectus* – i.e. q. 2 in the *Quaestiones in tertium de anima* – that offers us concrete help in the organization and structuring of Siger's material concerning the topic of eternalism. Here, he presents

³² Petagine (2004: 124).

³³ Sigerus de Brabantia, *Quaestiones in tertium de anima* [= *InIIIAn*], q. 9,31-34, pp. 26-27.

³⁴ Cfr. *Ibidem*. q. 2, pp. 4-8; Id., *De anima intellectiva* [= *AnInt*], cap. V, pp. 90-95; *Quaestiones super librum de causis* [= *InCausis*], q. 12, pp. 63-67.

four main eternalist arguments we will meet again and again – with more or less significant variations – in the other texts where he handles this issue.

The arguments are the following:

1. The first is an argument *ex parte dei*, which goes back to Aristotle's *Physics*³⁵ and has found in Avicenna probably its most famous version of the Medieval period: the effect X must always subsist together with its cause Y, if Y is *in actu* and lacks of nothing concerning its being-the cause-of-X, i.e. if Y is not only *causa efficiens* of X, but also its *causa sufficiens*. It is a simple *modus ponens* in an ontological meaning: $Y \rightarrow X, Y, \vdash X$. If cause Y is not only *causa sufficiens* of X but also eternal, then its effect X must also be eternal³⁶. If one considers that the Christian *causa prima*/God is by definition *in actu*, lacking in absolutely nothing and needing absolutely nothing for the causation of X (this is the meaning of Siger's insistence on the word «*immediate*», i.e. without any mediation of any *causa secunda*) and that He is eternal, then the argument seems to also be valid for the relation between God and the world (or God and the *intellectus* or God and every other being that is *immediately* caused by Him), as remarked by Siger on several occasions³⁷. Due to the *sufficientia* of the cause (God), its eternity and the *unmediated* character of the causation of the effect, the world (or the intellect) as effect seems to emerge as something eternal.

2. The second argument is also *ex parte dei* and represents – as pointed out by E. Berti, speaking about the original Aristotelian version – «a formidable objection to every form of creation in time»³⁸. It is worth directly quoting the words of the philosopher of Stagira in the *Physics* and Cicero's reporting of his opinion in the *Academica*:

[n]ow, if every one of those things which are capable of change came into existence, then the given change must be preceded by some other change – namely, the coming into existence of that which is capable of change or of that which is capable of causing change [...] As for there being rest for an infinite amount of time and then change at some time, and there being no differentiation involved in this

³⁵ Aristoteles, *Physica*, II, 3, 195b17-18.

³⁶ Cfr. Avicenna Latinus, *Liber de philosophia prima sive scientia divina*, tract. IX, c. 1, p. 435,26-27: [...] causa, quantum in se est, facit necessario esse causatum; quae, si fuerit semper, facit causatum necessario esse semper.

³⁷ Cfr. Sigerus de Brabantia, *InIIIAn*, q. 2,43-50, pp. 5-6; *Quaestiones in Metaphysicam*, ed. W. Dunphy, [= InMetD], rep. München, III, q. 16, pp. 134-136; *Quaestiones in Metaphysicam*, ed. A. Maurer, [= InMetM], rep. Paris, III, q. 6, pp. 413-414; Id., *InCausis*, q. 12,46-48, pp. 64-65.

³⁸ Berti (2007: 15).

(to explain why it should happen now rather than earlier) [...] this cannot be the work of nature³⁹.

veniet flumen orationis aureum fundens Aristoteles [...] neque enim ortum esse unquam mundum, quod nulla fuerit novo consilio inito tam praeclari operis inceptio, et ita esse eum undique aptum ut nulla vis tantos queat motus mutationem que moliri, nulla senectus diuturnitate temporum existere ut hic ornatus unquam dilapsus occidat⁴⁰.

Siger's reception of this material can be summarized through the two following arguments:

2a) In order to cause/create something new (X), cause Y must change. In order to cause/create X, Y must "come out" from its status s_1 (during which it was not creating X) and "enter" a new status s_2 (during which the causation/creation of X actually happens). However, the Christian God – in His perfection/immutability/eternal present – experiences no newness/change. Therefore, due to the impossibility of any sort of change in God, every being that is in an *unmediated* causal relation with Him (as the world or the intellect) cannot be caused as something new (*factum novum*), but rather must be something eternal (*factum aeternum*)⁴¹.

2b) Moreover, one has to indicate a cause, an explanation, a ground, a *ratio* for that change of Y from the situation s_1 (when it was not causing X and X actually did not exist) to the situation s_2 (when Y creates/begins to cause X): why does Y begin at time t_1 (which divides s_1 from s_2) to create X? Why not at t_2 ? Or today? Or one million years ago? Or tomorrow? This *ratio* can be found neither *ex parte dei* nor *ex parte effectus*: a change in the immutable first principle cannot be assumed (as explained in argument **2a**) and X cannot be the cause of its "coming to the existence", precisely due to its non-existence⁴²!

The strong connection between arguments **1)** and **2)** is repeatedly stressed by Siger, remarking that the First Principle – because of its lacking of nothing

³⁹ Aristoteles, *Physica*, VIII, 1, 251a17sqq.; 252a11sqq. [trans. by R. Waterfield, Oxford, Oxford University Press, 2008, pp. 186-189].

⁴⁰ Id., *Fragmenta selecta*, ΠΕΡΙ ΦΙΛΟΣΟΦΙΑΣ, f. 20, p. 89.

⁴¹ Cfr. Sigerus de Brabantia, *InIIIAn*, q. 2,53-63, cit., p. 6; *InMetM*, rep. Cambridge, III, q. 16,41-56, p. 112.

⁴² Cfr. Id., *AnInt*, cap. V,22-32, cit., pp. 91-92; *InMetD*, rep. München, III, q. 16,29-33, p. 135; *InCausis*, q. 12,46-59, cit., pp. 64-65.

and its being *causa sufficiens* – cannot have any reason for changing its status⁴³. For Siger, the one, who ascribes to God the will to create something new at t_1 and at the same time recognizes his status as a cause needing absolutely nothing for its action, is actually predicating two contradictory properties on the same subject. Nothing, not even Christian God, despite his omnipotence can either realize or possess them – in the background lies the idea of the principle of non-contradiction as the limit of God’s omnipotence. Indeed, on the one hand God as *causa sufficiens* must cause X as something eternal, as shown in **1**); on the other hand God’s will to create X in t_1 implies that he wants X as something non-eternal (just because it will begin to exist at t_1). The absurd conclusion is that God wants X as something eternal yet also as something non-eternal: «*Hoc autem est velle impossibilia. Deum autem velle impossibilia est impossibile*»⁴⁴.

3. Reporting argument 2a), I have affirmed that X, as something that does not exist before t_1 , cannot be the cause of its coming to existence. The third main eternalistic argument, this time *ex parte effectus*, explains and deepens this idea, marking the inner connection of all three of these arguments. Against Plato’s idea of a generated (by the Demiurge) yet nevertheless in the future incorruptible heaven/time/*kosmos*⁴⁵, Aristotle has massively underlined, especially in his *De caelo*, that the possibility/capability of being generated and the possibility/capability of being corrupted are actually the same, which a being either possesses or does not possess at all: the capacity of change (= C). A generated yet incorruptible being X must possess C, because X has changed during its passage from the status of non-being/being formless to the condition of actually being/having a form. The decisive point is that X, if truly possessing C, can actualize C again; moreover, we have to expect that it *will* actualize C sooner or later. Indeed, for Aristotle, the possession of a capacity that will be never actualized, corresponds to non-possessing such a capacity at all⁴⁶. Consequently, the incorruptibility of X in the future also implies its incorruptibility/eternity regarding the past: «*una virtus est per quam aliquid*

⁴³ Id., *InMetD*, rep. München, III, q. 16,29-56, p. 135; *InMetM*, rep. Paris, III, q. 6,1-18, p. 413.

⁴⁴ Id., *InMetM*, rep. Cambridge, III, q. 16, p. 113,85-86; cfr. *Ibidem*. II. 57-88, pp. 112-113; *InMetD*, rep. München, III, q. 16,53-56, p. 135; *Ibidem*. II. 29-55, p. 135.

⁴⁵ Plato, *Timaeus*, 37d-38c.

⁴⁶ Cfr. Aristoteles, *De caelo*, 279b4-280a11; 281a20-283b22.

est in toto praeterito et toto futuro»⁴⁷. Siger mostly reuses this argumentation in two versions:

(3a) The first version regards the *anima intellectiva*, whose eternal persistence in the future becomes the cornerstone for asserting its eternal persistence in the past – actually, one must always consider that the rational soul of a human being must eternally persist after death for a Christian, otherwise all the eschatological promises of happiness in heaven or the threats of punishment in the hell would have no sense⁴⁸;

(3b) The second concerns the notion of *creatio ex nihilo*. According to Siger, the *creatio ex nihilo* of X as a *factum novum* is something intrinsically impossible, both *ex parte effectus* (**3b₁**) and *ex parte dei* (**3b₂**).

3b₁) *Ex parte effectus*, one must never forget that the appearing *in actu* of a being X at t_1 implies the actual existence of something else with the potency-of-being-X before t_1 : the generation of X as being *in actu* is nothing other than the actualization of the potency-of-being-X (a typically Aristotelian argumentation). This cannot be the case for the creation of X *ex nihilo*, because nothing can possess the potency-of-being-X in a state of radical nothingness.

3b₂) Even *ex parte dei*, the concept of *creatio ex nihilo* is actually contradictory. Assuming that God as first efficient cause wants to create X *ex nihilo* at t_1 , we subsequently have to affirm that nothing in the non-existing universe had the potency-of-being-X before t_1 . Therefore, according to the metaphysical principle that something is ontologically impossible if nothing in the universe presents its potency-of-being, one is then allowed to affirm that X was something ontologically impossible before t_1 . Now, how could an impossibility be a possible subject of action/creation by an efficient cause, even by God⁴⁹?

4) There is fourth eternalistic argument that Siger deals with, which is already present in his *Quaestiones in tertium de anima*, before dominating the work *De aeternitate mundi*. Actually, it does not represent an argument *pro aeternitate mundi*, but rather – I would say – *in defensione aeternitatis mundi* and

⁴⁷ Sigerus de Brabantia, *InMetD*, rep. München, III, q. 15 “*Commentum*”, 79-80, p. 132; Cfr. Averroes, *Averrois Commentaria magna in Aristotelem De celo et mundo*, Textus/Commentum 121/122, pp. 230-232.

⁴⁸ Sigerus de Brabantia, *AnInt*, cap. V, 13-21; 33-53, pp. 91-93; *InCausis*, q. 12, 14-45, pp. 63-64.

⁴⁹ Id., *InMetM*, rep. Cambridge, III, q. 15, 20-58, pp. 108-109; *Ibidem*. rep. Paris, III, q. 5, 1sq, pp. 411-412.

contra novitatem mundi through the defense of the rational tenability of the thesis of the eternity of the human race. Indeed, Siger is faced with a possible argument against the eternity of the world (= AE), which identifies the eternity of the human species (a statement correlated with the claim of the eternity of the world) as the possible “weak point” for a general criticism of the eternalistic position. AE is built in the following way: if we admit, with Aristotle, that universals do not exist *extra mentem* as separated, independent, numerically different substances *in actu*, and that therefore they only exist *extra mentem* through their instances, then the human species cannot be eternal, because it would need at least an «*individuum sempiternum*» as its instance *extra mentem in actu*. In the light of the evident non-existence of eternal human beings, one has to affirm that the human species cannot be eternal and must have been created as something new at time t_1 .

It is not difficult for Siger to counter such a «*frivola ratio*»⁵⁰. The human species is not eternal because of the existence of a hyperuranic platonic idea of humanity or the existence of an *individuum sempiternum*; rather, it is eternal due to the generation *ad infinitum* of always new, numerically different yet specifically identical human beings through sexual reproduction. The species is eternal *per accidens*. According to this mode, the human species cannot be generated as a *factum novum* (in *De aeternitate mundi* Siger uses the formulation «*esse incepit cum penitus non praefuisset*»), because – in order to exist – it would need the pre-existence of other human beings as instances of the species prior to the existence of humanity at all⁵¹. Moreover, the claim that the human species cannot be a *factum aeternum* but only a *factum novum* would represent an unacceptable limitation of the power of God⁵². The Christian doctrine of the creation of the world and the human species (Adam and Eva as first human beings) is not the conclusion of a rational demonstration: it is a truth *per fidem*⁵³.

3. Truth against reason? Siger’s efforts for an “incompatibilism without conflict”

If *creatio ex nihilo* is impossible according to natural reason and a cosmology on an Aristotelian basis, if also the philosophy when discussing a central issue

⁵⁰ Id., *De aeternitate mundi* [= *Aet*], cap. II,57, p. 120.

⁵¹ *Ibidem*. cap. I, pp. 115-117.

⁵² *Ibidem*. cap. II,52-55, p. 120.

⁵³ *Ibidem*. cap. IV,82-83, p. 136.

such as the origin of the world and its ontological status, leads without committing mistakes to conclusions patently contrary to the articles of faith, to «what must be denied» as false on the basis of the *true* Christian religion⁵⁴, if reason is unable to prove the creation of the world as a *factum novum* or successfully counter and «dissolve» the eternalist arguments⁵⁵, then the traditional model of a strong compatibilism between faith and reason seems inadequate and heavily jeopardized. The rationally valid is now false, while the truth becomes the opposite of the rationally valid and demonstrable. The ideal of the *sapientia christiana* and its organic unity between the articles of faith and rational demonstrations seem to crumble and require replacement by a sharply incompatibilistic model in which truth stands *against* the conclusions of the natural reason⁵⁶.

However, Siger attempts to avoid an “incompatibilism with conflict”, first through relativizing and limiting the validity of the conclusions of natural

⁵⁴ Cfr. Id., *AnInt*, cap. V, 11-12, p. 91: Et hoc [sc. the argument (3)] etiam apparet rationibus naturalibus quas Aristoteles secutus est;

InMetD, rep. München, q. 19, 11-20, p. 144: Dico [...] circa istam rationem [sc. the connection of the arguments 1) and 2)] [...] quod ratio humana ducit in hoc quod debet negari.

InMetM, rep. Cambridge, III, q. 15, 77-79, p. 110: [...] Et credo quod, sicut ea quae fidei sunt per rationem humanam probari non possunt, ita sunt aliquae rationes humanae ad opposita eorum, quae per humanam rationem dissolvi non possunt. Propter hoc tamen in his quae fidei sunt non est errandum; sed est alia via credendi ea quam per rationem humanam [...];

Aet, “*Introductio*”, 42-43, p. 115: [...] secundum viam philosophiae dicendum erit quid istorum [sc. potentia et actus] alterum duratione praecedat, [...].

Ibidem. cap. I, 4-5, p. 115: [...] secundum philosophos species humana non est causata nisi per generationem.

Ibidem. cap. I, 36-37, p. 116: Sic igitur ex iam dictis patet qualiter species humana a philosophis ponitur sempiterna et nihilominus causata.

Ibidem. cap. I, 43-44, p. 117: [...] quod species humana secundum philosophos semper est, nec esse incepit cum penitus non praefuisset.

⁵⁵ Cfr. Id., *InMetM*, rep. Cambridge, III, q. 15, 77-79, p. 110;

ibidem. rep. Paris, III, q. 5, 41-46, p. 412: Nec debet aliquis conari per rationem inquirere quae supra rationem sunt, vel rationes in contrarium dissolvere. [...] non debet aliquis negare veritatem catholicam propter aliquam rationem philosophicam, licet illam dissolvere nesciat. *InMetD*, rep. München, q. 16, 85-89, p. 136: [...] licet oppositum huius per fidem teneatur, quamquam non possit per demonstrationes probari, et quamquam rationes Aristotelis possint impediri, non tamen videtur quod possit ad plenum satisfieri intellectui humano.

InCausis, q. 12, 89-92, p. 66: Nihilominus tamen, <quia> auctoritas fidei christianae maior <est> omni ratione humana et etiam philosophorum auctoritate, dicamus intelligentiam non esse in aeternitate, licet ad hoc demonstrationem non habeamus.

⁵⁶ Id., *InMetD*, rep. München, III, q. 15 “*Commentum*”, 87-88, p. 132: sic autem velare philosophiam non est bonum; unde non est hic intentio Aristotelis celanda, licet sit contraria veritati; cfr. *InMetD*, rep. München, q. 19, 11-20, p. 144; *InMetM*, rep. Paris, III, q. 5, 40-46, p. 412.

reason to precise boundaries, and secondly by sketching the model of a truth *beyond* – and not actually *against* – natural reason.

The path that he follows was opened by Albert the Great, who – in his epochal attempt to make the whole Aristotle intelligible to the Latins⁵⁷– already systematically separated natural philosophy and revealed theology because of their characters and main aims (contemplative for the former, practical for the latter), principles (only the *principia per se nota* for the former; the supra-natural, meta-rational articles of faith for the latter), their subjects (natural things in their normal causal relations and what is rationally/empirically accessible for the former; the “*miracula*” of God for the latter). It is worth quoting some passages from the Dominican theologian:

Non enim possumus in physicis probare, nisi quod sub principiis est physicis. Et si nos extenderemus nos ad loquendum de his quae supra physim sunt, non possemus esse physici, quia non procederemus ex probatis vel per se notis⁵⁸;

Theologica autem non convenient cum philosophicis in principiis, quia fundantur super revelationem et inspirationem, et non super rationem: et ideo de illis in philosophia non possumus disputare⁵⁹;

nihil ad me de dei miraculis cum ego de naturalibus disseram⁶⁰;

[...] naturalia non sunt a casu, nec a voluntate, sed a causa agente et terminante ea: nec nos in naturalibus habemus inquirere qualiter Deus opifex secundum sua liberam voluntatem in creatis ab ipso utatur ad miraculum quo declaret potentiam suam, sed potius quid in rebus naturalibus secundum causas naturae insitas naturaliter fieri possit⁶¹.

Inheriting Albert’s stimuli, Siger pointedly remarks upon the differences between the natural/rational theology which is part of philosophy on the one hand and revealed theology on the other: while the former rests on principles that are known «*via sensus, memoriae et experimenti et lumine naturali*», the latter rests on principles that are only known «*per divinam revelationem*»; while the former finds the boundaries of its competence-field in the set of subjects that can be known «*per rationem humanam et per creaturas*», the latter extends its competence to subjects «*quae sunt supra rationem humanam*

⁵⁷ Cfr. Albertus Magnus, *Physica*, pars I, lib. I, tract. I, c. 1, p. 1,43-49.

⁵⁸ Id., *Physica*, pars 2, lib. VIII, tract. 1, c. 2, p. 551,11-15.

⁵⁹ Id., *Metaphysica*, XI, tract. III, c. 7, p. 542,25-31.

⁶⁰ Id., *De generatione et corruptione*, lib. I, tract. 1, c. 22, p. 129,15-16.

⁶¹ Id., *De caelo et mundo*, lib. I, tr. 4, c. 10, p. 103,5-12.

et quae per creaturas tantum non possunt cognosci»; if the former has a speculative character, the latter is «practica et non tantum speculativa»⁶².

Concretely concerning the eternalist arguments, Siger mainly insists upon three strategies in order to delimit their validity in the field of the demonstrative, Aristotelian natural reason, and, in this way, to be able to guarantee a space “beyond” it for existence of the (revealed) Truth:

- A) The distinction between free, voluntary agents and involuntary, just natural agents;
- B) A call for prudence when extending the results of philosophical/scientific investigations (conducted just on the basis of creatures) to a non-creature such as God;
- C) An insistence on the meta-rationality of the articles of faith.

A) The first strategy is amply used in the q. 2 of the *Quaestiones in tertium de anima* in order to avoid the necessity of the conclusions of the arguments 1) and 2): the *causa efficiens* of the intellect/world is a “poietic” cause that creates the effect «*secundum formam voluntatis*» and can create the effect as a *factum novum* as well as a *factum aeternum*, due to the absolute freedom of its creative will. Consequently, in order to adequately know whether the intellect/world has been created as a *factum novum* or *factum aeternum*, we should be able to investigate God’s will, which by definition is inscrutable. As a possible subject of investigation, we only have the nature of the intellect/world as something *immediate* created by God, and from this perspective we are allowed to say that it is probably a *factum aeternum*⁶³.

Strategy A) was a *topos* among the theologians of the 13th century, not only among authors with a marked disposition in favor of Aristotle, such as Albert the Great⁶⁴ and Thomas Aquinas⁶⁵, but also by authors such as Bonaventura, who distinguishes between a *causa sufficiens* «*operans per naturam*» and a *causa sufficiens* «*operans per voluntatem et rationem*» in order to criticize argument 1)⁶⁶.

In the *Quaestiones in Metaphysicam* Siger is more radical and does not hesitate to remark that 1) and 2) are valid and consistent arguments as well concerning an «*agens secundum voluntatem*»:

⁶² Sigerus de Brabantia, *InMetD*, rep. Wien, VI, q. 1, “*Commentum*“ 1, pp. 359-361.

⁶³ Id., *InIIIAn*, q. 2, 64sq., pp. 6-8.

⁶⁴ Cfr. Albertus Magnus, *Summa Theologiae, Pars Secunda*, tract. 1, q. 4, a. 5, part. 1, p. 97.

⁶⁵ Cfr. Thomas de Aquino, *Quaest. disp. de potentia*, q. 3, a. 17, ad 6, p. 214.

⁶⁶ Bonaventura, *In secundum librum sententiarum*, d. 1, p. 1, a. 1, q. 2, ad 5, p. 23.

[...] quoniam ex quo existit totum illud ab aeterno unde effectus debet sortiri quod sit effectus, videtur quod in utroque agente [sc. in agente naturali et agente secundum voluntatem] sit eadem ratio⁶⁷;

[...] in utroque agente videtur necessarium quod, existente toto a quo debet effectus sortiri, existit effectus; et existente ab aeterno, et existit effectus ab aeterno⁶⁸.

Now, in order to avoid the necessity of the eternalist arguments – and accordingly an incompatibilism with conflict between faith and reason – the option of insisting upon **B**) and **C**) is necessary.

B) The second strategy can be interpreted as an invitation to avoid the case of the fallacy of hasty-generalization: the ontological abyss between the Creator and His creatures is too deep to allow a direct extension of the validity of a causal connection among creatures to the causal connection between creatures and their Creator. Indeed, it is not possible for our limited intellect – adapted for understanding the *sensibilia* – to completely understand the *modus agendi* of God, simply because every substance acts «*secundum suam substantiam*», and God’s substance – in His creatureliness and beyondness – cannot be adequately understood by us⁶⁹. In general, one should be prudent in the acceptance of a validity of the conclusions of natural reason concerning the separated substances, because they are not possible objects of empirical experience⁷⁰.

C) The discussion of **B**) also reflects the best possible introduction for strategy **C**), which actually represents the further, consequent step of the prudential remarks just described concerning the probability and limited validity – though not the necessity – of eternalist arguments: the truth of the articles of the Christian faith are meta-rational, «*supra rationem humanam*», they are not possible objects of philosophical knowledge; rather, they are “beyond” every satisfying human demonstrations/confutations⁷¹. Even in a work such as the *Quaestiones in librum de causis*, where the combative Siger of the *De anima intellectiva* and the *Quaestiones in Metaphysicam* seems to surrender

⁶⁷ Sigerus de Brabantia, *InMetD*, rep. München, III, q. 16,40-42, p. 135.

⁶⁸ Id., *InMetM*, rep. Cambridge, III, q. 16,65-68, pp. 112-113.

⁶⁹ Id., *InMetD*, III, rep. München, q. 19, cit., pp. 144-145; *InMetM*, rep. Cambridge, III, q. 16,89-104, cit., pp. 113-114.

⁷⁰ *Ibidem*. rep. Paris, III, q. 6,19-28, p. 413.

⁷¹ Cfr. *Ibidem*. q. 5,42, p. 412; Id., *InMetD*, rep. München, q. 16,85-91, p. 136; *Ibid.*, VI, q. 1, “*Commentum*” 1, pp. 359-361, especially ll. 81-84.

to the request of the traditional theologians⁷² for at least a rational refutation of the eternalist arguments and a defense of the non-impossibility *on the rational level* of the idea of a *creatio ex nihilo*⁷³, the philosopher of Brabant cannot evade affirming that we actually do not have demonstrations for what the authority of the Christian faith demands us to believe and what we should believe «*firmiter et pie sine ampliore inquisitione*»⁷⁴. Truth and rational demonstrability/validity, even if not contradicting each other on the same level, seem to have definitively divorced, representing a hard blow to the model of the *fides quaerens intellectum* and the ideal of the *sapientia christiana*.

4. Radical Aristotelianism, truthfulness and autonomy of philosophy: some conclusive reflections

Drawing my analysis to a conclusion, I would like to develop some reflections on Siger's dealing with Aristotle's eternalist arguments exposed above. First, allow me to sketch some critical remarks concerning the use of the label "radical Aristotelianism" for the thought of Siger, a widespread definition⁷⁵ defended by F. Van Steenberghen in his undoubtedly fundamental studies on him⁷⁶, but which may prove to be misleading for the interpretation of his philosophy and Aristotelianism. Adequate space for a sufficient handling of this issue would probably be a monograph rather than the following few pages; nevertheless, I believe that the exposed material concerning Siger's reception of the arguments *pro aeternitate mundi* urges a reconsideration of the adequacy of such a label. Despite such a tempting occasion as the topic of the eternity of the world for showing his presumed uncompromising, unscrupulous, bold radicalism – a consequence of his being an intransigent, fervent and slavish, yet therefore faithful, disciple of Aristotle (this is the Siger that Van Steenberghen portrays⁷⁷) –, on the contrary, the Brabantine philosopher tries to constantly relativize the validity of the conclusions of the Philosopher and natural reason in order to avoid the *necessity* of such conclusions, not only through confining their cogency within the boundaries of natural reason but also through the repeated *negation* of their truth. In this way, he prevents

⁷² Cfr. H. Denifle-A. Chatelain (edd.), "Statutum facultatis artium contra artistas tractantes quaestiones theologicas", pp. 499-500; [Pseudo-]Bonaventura, *Dominica III. Adventus, sermo II*, p. 63. Cfr. Bianchi (1984: 84-87); Putallaz & Imbach (1997: 123sqq.).

⁷³ Cfr. Sigerus de Brabantia, *InCausis*, q. 20, pp. 83-89.

⁷⁴ *Ibidem*. q. 13, 89-120, pp. 66-67.

⁷⁵ Cfr. e.g. Corti (1982: 12sqq); Lindberg (2002: 68); R. Heinzmann (³2008: 160-164).

⁷⁶ Cfr. Van Steenberghen (²1991: 351; 354-359); cfr. also Id. (1977: 397-399).

⁷⁷ Cfr. *Ibidem*. p. 223, p. 346, pp. 389sqq.

the absurd contrapositions of two truths both valid *simpliciter* (the deforming interpretation and charge of Thomas Aquinas as well as bishop Étienne Tempier against the *avveroistas*⁷⁸). In this regard, L. Bianchi rightly pointed out that Albert the Great and even Bonaventura appear more “radical” than Siger himself, because they do not hesitate to speak of «truth» concerning the eternalist argumentations of Aristotle (even if it is merely a relative truth, limited in the field of the physics and the usual physical relations between natural causes and their effects)⁷⁹. Moreover, we cannot ignore the similarity between Thomas Aquinas – the “orthodox” and critical reader of Aristotle, according to Van Steenberghen – and the “heterodox” and “slavish” Siger regarding the indemonstrability of the creation of the world as a *factum novum*, the consequent affirmation of the creation of the world as a truth only «*per revelationem divinam credibile*»⁸⁰, and also the crucial remark that eternity does not automatically exclude creatureliness, as Bonaventura claimed⁸¹: a creature is actually something which ontologically depends «*ab alio*», on another cause as its «*principium originis*», and not necessarily something that began to exist in t_1 through creation in a temporal meaning⁸². Furthermore, the (presumed) radical, intransigent, servile disciple of Aristotle explicitly charges Aristotle of *petitio principis* concerning **3b2**⁸³, and stresses without any hesitation the *insufficiency* of authorities such as Aristotle, Avicenna or Proclus in the investigation of the truth, pleading on the contrary for an autonomous use of his reason: «*Sumus autem nos homines sicut et illi. Quare igitur non debemus nos niti ad huius inquisitionem per rationem sicut et illi?*»⁸⁴

The best complement for such a statement can be found in a sentence taken from Siger’s *Quaestiones Morales*, which on the one hand illustrates his actual conception of philosophy and its aim, while on the other hand allows us to return to main topic of the relation between reason and faith and

⁷⁸ Cfr. H. Denifle-A. Chatelain (edd.), *Opiniones ducentae undeviginti [...] a Stephano episcopo Parisiensi de consilio doctorum Sacrae Scripturae condemnatae 1277 [...]*, p. 548; Cfr. Thomas de Aquino, *De unitate intellectus contra averroistas*, cap. V, ed. Leonina, p. 314,412-430.

⁷⁹ Cfr. Bianchi (1984: 53-63).

⁸⁰ Thomas de Aquino, *In II Sent.*, d. 1, q. 1, a. 5 co., p. 126; Cfr. *Summa Theologiae*, I, q. 46, a. 2 co., p. 255.

⁸¹ Bonaventura, *In secundum librum sententiarum*, d. 1, p. 1., a. 1, q. 2, Resp., pp. 22-24. Cfr. Putallaz-Imbach (1997: 80): Non seulement la foi dit que le monde a commencé, mais la raison le tient également ; il est proprement impossible que le monde soit éternel.

⁸² Cfr. Thomas de Aquino, *In I Sent.*, d. 19, q. 2, a. 1 co., p. 52; *In II Sent.*, d. 1, q. 1, a. 5, ad 2 in contr., p. 126; *Summa Theologiae*, I, q. 10, a. 5 co, p. 197. Cfr. Porro (1996: 104sqq).

⁸³ Sigerus de Brabantia, *InMetM*, rep. Paris, III, q. 5,19-21, p. 412.

⁸⁴ *Ibidem*. rep. Cambridge, q. 5,39-42, p. 84.

the model of incompatibilism without conflict that the brabantine philosopher has tried to sketch concerning the issue of the Aristotelian eternalist arguments. The sentence to which I refer is as follows:

[...] philosophus intendit finaliter cognitionem veritatis, et ille status qui minus retrahit a cognitione veritatis magis convenit et competit philosophis. [...] Unde cognitio veritatis est ultimus finis hominis; veritatis praecipue de Deo⁸⁵.

If the final aim of philosophical research and human being is the *cognitio veritatis*, then one must be cautious in the interpretation of all the preceding passages which I have quoted or referred to in the pages, whose focus was on stressing the contrariety between truth on the one hand and a reliable reconstruction of the opinion of Aristotle and the conclusions of rational investigation concerning the issue of eternalism on the other. On the face of them, the price that Siger *seems* ready to pay in order to avoid an absurd conflict between two “truths” is the renunciation of the idea of a philosophy as source of truth, as a way for discovering the Truth: not only some parts of the Truth, i.e. some true theses (the ideal of the Truth as an organic and homogenous whole remains in the background) are unachievable by human reason – and this is a first aspect of its not being a source of truth – but, concerning some fundamental topics, reason emerges as a source of falsity, of «*what must be denied*» on the basis of what we “know” *per revelationem*.

Now, in the light of the sentences quoted at the end of the preceding paragraph as well as this one, it must be specified that Siger does not go so far as to completely deny reason its truthfulness, its privilege of being a way to discover the Truth. The subject that the *lumen naturale* aims to illuminate is not different from that of faith, i.e. – in accordance with authors such as Augustine or Anselm – God/the Truth. However, in comparison to the Archbishop of Canterbury’s optimism concerning the capability of natural reason to rationally penetrate the articles of faith/the Truth, Siger defends a much humbler concept of *lumen naturale* and philosophy: firstly they are recognized as unable to enlighten important “zones” of the Truth, and secondly they are often prone to real “blackouts”, i.e. sometimes becoming a source of “darkness”, because it leads – without committing grave internal mistakes – to dissonant theses in comparison with the articles of faith regarding central issues such as *creatio ex nihilo* and the origin of the world. At this place, one can eventually discern the hearth of Siger’s strategy in the dealing with the eternalist arguments: the use of a truly more modest concept of natural reason as discoverer of the truth but strongly based upon the systematic remark of its

⁸⁵ Id., *Quaestiones morales*, q. 4,13sq., p. 102.

autonomy in its principles, in its methodology, and therefore also in its conclusions, even if false. This represents also the problematic compromise Siger attempted to achieve, trying to avoid having to sacrifice either the truthfulness of philosophical investigation in order to guarantee its autonomy, or its autonomy in the determination of what is demonstratively valid within the boundaries of a specific philosophical discipline in order to guarantee its truthfulness.

Indeed, Siger – as well as his contemporaries who, dealing with the reception of Aristotle, were particularly sensitive to the strength of the Philosopher's argument and his model of philosophy – was pressed by the following, difficultly compatible concerns/needs:

- a) The acknowledgment of faith as a fundamental, eminent source of Truth, unshakable by rational disputations;
- b) The existence of philosophy as “another” source of Truth “near” the Christian faith, now in the strong version of the Aristotelian paradigm;
- c) The autonomy of philosophy in its principles, methodology and conclusions based upon the Aristotelian paradigm;
- d) The impossibility of the coexistence of different, contrasting truths, i.e. the impossibility that this “otherness”, this autonomy of philosophy would come into direct conflict, or even negate, the articles of faith.

Reading the passages concerning the issue of the eternity of the world, one sometimes has the impression that Siger would be willing to sacrifice **b** – i.e. the truthfulness of philosophy – in order to “save” its autonomy (**c**) and – through such a castration of natural reason – avoid the problem exposed in **d**). Now we know that was not this actually Siger's project or concept of philosophy, given that he tried, on the contrary – through the compromise of a *lumen naturale* that is unreliable on crucial questions, yet at least an autonomous discoverer/enlightener of the truth – to balance all four of the demands mentioned here.

Today, after secularization, one would not have too much difficulty in querying the need for **a**) and consequently drastically reformulating **d**). Apart from the subjective feeling of a single person or private groups, faith has indeed lost its eminence as the objectively acknowledged primary source for grasping Truth, at the cultural and social level, in science, in philosophy, as well as in politics or even private life. Now, faith usually appears as *just one*

of several, competing and often contrasting voices that *claim* to have a privileged relation with the Truth. Siger – a son of his epoch – cannot and did not desire – on the basis of the texts we have – to “downgrade” the Christian faith in such a way, i.e. by casting doubt on the idea of its being the privileged source of Truth, that Truth, which is Jesus Christ, the *Verbum/Logos* and the organic whole of all particular truths.

In conclusion, Siger’s “incompatibilism without conflict” model emerges as a difficult, delicate compromise between the needs mentioned, perhaps unstable and hardly tenable from the contemporary point of view yet certainly interesting for the history of philosophy as an emblematic example of the *crisis* of the ideal of the *sapientia christiana* that Aristotle’s reception caused, i.e. as a direct witness of one of the fundamental and critical moments in the history of the European thought. Furthermore, the reconstruction of Siger’s dealing with the eternalist arguments and his attempts to find a problematic mediation and a convergence between such contrasting demands provide the reader an opportunity to reflect – for once free from contemporary polemics – upon which concerns, which needs, which risks are actually at stake when one debates about the *vexata quaestio* of the relation between reason and faith.

Acknowledgements

I would like to heartily thank Christian Rode (Rheinische Friedrich-Wilhelms-Universität Bonn) for allowing me to openly discuss the present paper with him on several occasions, including some problems I had to overcome during the writing of the text.

Bibliography

Abbreviations

AFN *Ante-Nicene Fathers*, edd. A. Roberts et. al., Buffalo, NY, Christian Literature Publishing Co., 1885sqq.

CSEL *Corpus Scriptorum Ecclesiasticorum Latinorum*, hrsg. von der Akademie der Wissenschaften zu Wien, Wien, 1866sqq.

CCCM Corpus Christianorum. Continuatio Mediaevalis, Turnhout, Brepols, 1971sqq.

CCSL *Corpus Christianorum. Series Latina*, Turnhout, Brepols, 1954sqq.

PL *Patrologia Latina*, ed. J.-P. Migne, Paris, apud J.-P. Migne Editorem, 1844sqq.

SC *Sources Chrétiennes*, Paris, Les Éditions du Cerf, 1941sqq.

Primary Sources

Albertus Magnus, *De caelo et mundo*, ed P. Hossfeld, Ed. Coloniensis V/1, Münster, Aschendorff, 1971.

Albertus Magnus, *De generatione et corruptione*, ed. P. Hossfeld, Ed. Coloniensis V/2, Münster, Aschendorff, 1980, pp. 107-213.

Albertus Magnus, *Metaphysica* VI-XIII, ed. B. Geyer, Ed. Coloniensis XVI/2, Münster, Aschendorff, 1964.

Albertus Magnus, *Physica*, pars I, ed. P. Hossfeld, Ed. Coloniensis IV/1, Münster, Aschendorff, 1987.

Albertus Magnus, *Physica*, pars 2, ed. P. Hossfeld, Ed. Coloniensis IV/2, Münster, Aschendorff, 1993.

Albertus Magnus, *Summa Theologiae, Pars Secunda*, ed. A. Borgnet, *B. Alberti Magni Opera Omnia*, vol. XXXII, Paris, L. Vivès, 1895.

Anselmus Cantuariensis, *Cur Deus Homo*, ed. F.S. Schmitt, *S. Anselmi Cantuariensis Archiepiscopi Opera Omnia*, vol. II, Edinburgh, Thomas Nelson & Son, 1946, pp. 37-133.

Anselmus Cantuariensis, *Epistola De Incarnatione Verbi*, ed. F.S. Schmitt, *S. Anselmi Cantuariensis Archiepiscopi Opera Omnia*, vol. II, Edinburgh, Thomas Nelson & Son, 1946, pp. 1-35.

Anselmus Cantuariensis, *Proslogion*, ed. F.S. Schmitt, *S. Anselmi Cantuariensis Archiepiscopi Opera Omnia*, vol. I, Edinburgh, Thomas Nelson & Son, 1946, pp. 89-139.

Aristoteles, *De caelo*, ed P. Moraux, Paris, Les Belles Lettres, 1965 [Bekker pp. 268a1-313b22].

- Aristoteles, *Physica*, ed. W.D. Ross, Oxford, Clarendon Press, 1950 [Bekker pp. 184a10-267b26].
- Aristoteles, *Fragmenta selecta*, ed. W.D. Ross, Oxford, Clarendon Press, ²1958.
- Augustinus Hipponensis, *Contra Academicos*, ed. W.M. Green, CCSL 29, Turnhout, Brepols, 1970, pp. 2-61.
- Augustinus Hipponensis, *De gratia et libero arbitrio*, ed. J.-P. Migne, PL 44, Paris, apud J.-P Migne Editorem, 1845, pp. 881-912.
- Augustinus Hipponensis, *De libero arbitrio*, ed. W.M. Green, CCSL 29, Turnhout, Brepols, 1970, pp. 207-321.
- Augustinus Hipponensis, *De vera religione*, ed. K.-D. Daur, CCSL 32, Turnhout, Brepols, 1962, pp. 171-260.
- Augustinus Hipponensis, *De Trinitate*, ed. W.J. Mountain et al., CCSL 50-50A, Turnhout, Brepols, 1968.
- Augustinus Hipponensis, *Epistulae* II, ed. A. Goldbracher, CSEL 34.2, Prag-Wien-Leipzig, Tempsky-Freytag, 1898.
- Augustinus Hipponensis, *Sermones de vetero testamento* I-L, ed. C. Lambot, CCSL 41, Turnhout, Brepols, 1961.
- Averroes, *Averrois Commentaria magna in Aristotelem De celo et mundo*, lat. transl. by Michael Scotus, edd. F. J. Carmody-R. Arnzen-G. Endress, Leuven, Peeters, 2003.
- Avicenna Latinus, *Liber de philosophia prima sive scientia divina*, V-X, ed. S. Van Riet, intr. G. Verbeke, Louvain-Leiden, Peeters-Brill, 1980.
- Bonaventura, *In secundum librum sententiarum.*, ed. PP. Collegii S. Bonaventurae, *Doctoris Seraphici S. Bonaventurae Opera Omnia*, vol. II, Ad Claras Aquas, Quaracchi, Typographia Collegii S. Bonaventurae, 1885.
- Bonaventura, *Sermones de Tempore*, ed. PP. Collegii S. Bonaventurae, *Doctoris Seraphici S. Bonaventurae Opera Omnia*, vol. IX, Ad Claras Aquas, Quaracchi, Typographia Collegii S. Bonaventurae, 1901, pp. 23-461.

- H. Denifle-A. Chatelain (edd.), *Chartularium Universitatis Parisienis*, t. I, Paris, Delalain, 1889.
- Franciscus PP., Littera Encyclica “*Lumen Fidei*” (Die XXIX m. Iunii, A.D. MMXIII), On-line: http://www.vatican.va/holy_father/francesco/encyclicals/documents/papa-francesco_20130629_enciclica-lumen-fidei_lt.html (consulted 14th October 2013).
- Hugo de Sancto Victore, *Didascalicon (De studio legendi)*, ed. C.H. Buttmer, Washington, D.C., Catholic University Press, 1939.
- Iohannes Scotus Eriugena, *De divina praedestinatione*, ed. G. Madec, CCCM 50, Turnhout, Brepols, 1978.
- Iohannes Scotus Eriugena, *Periphyseon*, lib. I, ed. E. Jeaneau, CCCM 161, Turnhout, Brepols, 1996.
- Ioannes Paulus II PP.: Littera Encyclica “*Fides et ratio*” (Die XIV m. Septembris, A.D. MCMXCVIII), On-line: http://www.vatican.va/holy_father/john_paul_ii/encyclicals/documents/hf_jp-ii_enc_14091998_fides-et-ratio_lt.html (consulted 14/10/2013).
- Lactantius, *Divinae Institutiones*, ed. S. Brandt, CSEL 19, Prag-Wien-Leipzig, Tempsky-Freytag, 1890, pp 1- 672.
- Lactantius, *Epitome Divinarum Institutionum*, ed. S. Brandt, CSEL 19, Prag-Wien-Leipzig, Tempsky-Freytag, 1890, pp. 673-761.
- Plato, *Timaeus*, ed. J. Burnet, *Platonis opera*, vol. 4, Oxford, Clarendon Press, 1902 (repr. 1968) [Stephanus pp. 17a-92c].
- Origenes, *Contra Celsum*, ed. M. Borret, SC 132, 136, 147, 150, 227, Paris, Les Éditions du Cerf, 1969-1976.
- Origenes, *De principiis*, ed. H. Crouzel/M. Simonetti, SC 252, 253, 268, 269, Paris, Les Éditions du Cerf, 1978-1980.
- Sigerus de Brabantia, *De aeternitate mundi*, ed. B. Bazán (*Philosophes Médiévaux* 13), Louvain-Paris, Publications Universitaires-B. Nauwelaerts, 1972, pp. 113-136.

- Sigerus de Brabantia, *De anima intellectiva*, ed. B. Bazán (*Philosophes Médiévaux* 13), Louvain-Paris, Publications Universitaires-B. Nauwelaerts, 1972, pp. 70-112.
- Sigerus de Brabantia, *Ecrits de logique, de morale et de physique*, ed. B. Bazán (*Philosophes Médiévaux* 14), Louvain-Paris, Publications Universitaires-B. Nauwelaerts, 1974.
- Sigerus de Brabantia, *Quaestiones in Metaphysicam, Édition revue de la reportation de Munich. Texte inédit de la reportation de Vienne*, ed. W. Dunphy (*Philosophes Médiévaux* 24), Louvain-la-Neuve, Éditions de l'Institut Supérieur de Philosophie, 1981.
- Sigerus de Brabantia, *Quaestiones in Metaphysicam, Texte inédit de la reportation de Cambridge. Édition revue de la reportation de Paris*, ed. A. Maurer (*Philosophes Médiévaux* 25), Louvain-la-Neuve, Éditions de l'Institut Supérieur de Philosophie, 1983.
- Sigerus de Brabantia, *Quaestiones in tertium de anima*, ed. B. Bazán (*Philosophes Médiévaux* 13), Louvain-Paris, Publications Universitaires-B. Nauwelaerts, 1972, pp. 1-69.
- Sigerus de Brabantia, *Quaestiones super librum de causis*, ed. A. Marlasca (*Philosophes Médiévaux* 12), Louvain-Paris, Publications Universitaires-B. Nauwelaerts, 1972.
- Tertullianus, *Apologeticum*, ed. E. Dekkers, CCSL 1, Turnhout, Brepols, 1954, pp. 74-171.
- Tertullianus, *De Carne Christi*, ed. E. Kroymann, CCSL 2, Turnhout, Brepols, 1954, pp. 871-917.
- Thomas de Aquino, *De unitate intellectus contra averroistas, Opera Omnia iussu Leonis XIII P.M. edita*, tom. XLIII, Roma, Editori di San Tommaso, 1976, pp. 245-314.
- Thomas de Aquino, *In quattuor libros sententiarum*, ed. R. Busa, *S. Thomae Aquinatis Opera Omnia I*, Stuttgart-Bad Cannstatt, Frommann-Holzboog, 1980.

Thomas de Aquino, *Quaest. disp. de potentia*, ed. R. Busa, *S. Thomae Aquinatis Opera Omnia III*, Stuttgart-Bad Cannstatt, Frommann-Holzboog, 1980, pp. 186-269.

Thomas de Aquino, *Summa Theologiae*, ed. R. Busa, *S. Thomae Aquinatis Opera Omnia II*, Stuttgart-Bad Cannstatt, Frommann-Holzboog, 1980, pp.184-926.

Secondary Sources

Berti, E. (2007), *In principio era la meraviglia. Le grandi questioni della filosofia antica*, Roma-Bari, Laterza.

Bianchi, L. (1984), *L'errore di Aristotele: la polemica contro l'eternità del mondo nel XIII secolo*, Firenze, La Nuova Italia Editrice.

Bianchi, L. (1990), «Gli aristotelismi della scolastica», in: L. Bianchi - E. Randi, *Le verità dissonanti*, Roma-Bari, Laterza, pp. 3-31.

Corti, M. (1982), *Dante a un nuovo crocevia*, Firenze, Casa Ed. Le Lettere/Sansoni.

Gombocz, W. L. (1997), *Die Philosophie der ausgehenden Antike und des frühen Mittelalters*, in: *Geschichte der Philosophie*, hrsg. von W. Röd, Bd. IV, München, Beck.

Grant, E. (1996), *The Foundations of Modern Science in the Middle Ages: Their Religious, Institutional, and Intellectual Contexts*, Cambridge, Cambridge University Press.

Heinzmann, R. (2008), *Philosophie des Mittelalters*, Stuttgart, W. Kohlhammer.

Horn, C. (1995), *Augustinus*, München, Beck.

Lindberg, D. (2002), «Medieval Science and Religion», in: *Science and religion: a historical introduction*, ed. by G. B. Ferngren, Baltimore (Md.)-London, Johns Hopkins University Press, pp. 57-72.

Petagine, A. (2004), *Aristotelismo difficile: l'intelletto umano nella prospettiva di Alberto Magno, Tommaso d'Aquino e Sigieri di Brabante*, Milano, Vita e Pensiero.

- Porro, P. (1996), *Forme e modelli di durata nel pensiero medievale: l'aevum, il tempo discreto, la categoria "quando"*, Louvain, Louvain Univ. Press.
- Putallaz, F. X., Imbach, R. (1997), *Profession: philosophe. Siger de Brabant*, Paris, Éditions du Cerf.
- Swindal, J. (2013), «Faith and Reason», *Internet Encyclopedia of Philosophy*;
On-line: <http://www.iep.utm.edu/faith-re/> (consulted 14/10/2013).
- Vanni Rovighi, S. (1978), «Perché Bonaventura ha criticato Aristotele», *Studi di filosofia medievale. Vol. II.: Secoli XIII e XIV*, Milano, Vita e Pensiero, pp. 40-52.
- Van Steenberghen, F. (1977), *Maître Siger de Brabant (Philosophes médiévaux 21)*, Louvain, Publ. Univ..
- Van Steenberghen, F. (1991), *La philosophie au XIIIe siècle (Philosophes médiévaux 28)*, Louvain-La-Neuve/Louvain, Éditions de l'Institut Supérieur de Philosophie-Éditions Peters.

« Vernünftige Selbstliebe » Universale e particolare nel giudizio politico kantiano

Emanuela Macera
Università degli Studi di Urbino Carlo Bo
emanuelamacera@libero.it

1. Introduzione

Se è possibile sostenere con Pasolini che «non esiste razionalità senza senso comune e concretezza. Senza senso comune e concretezza la razionalità è fanatismo»¹; e sebbene sia anche possibile sostenere con Eraclito che «la mente è separata da tutte le cose»², possiamo, altresì, sostenere con la Arendt che:

Pensare, volere e giudicare sono le tre attività spirituali fondamentali: non si possono dedurre l'una dall'altra e sebbene posseggano certe caratteristiche comuni non si possono ridurre ad un unico denominatore. Alla domanda «che cosa ci fa pensare?» non c'è in definitiva altra risposta se non ciò che Kant chiamava il «bisogno della ragione», l'impulso interiore di tale facoltà ad attualizzare se stessa nella speculazione³.

Il presente contributo si propone di indagare la funzione della razionalità, all'interno di una selezione scelta di scritti etico-storico-politici kantiani, le possibilità conoscitive della ragione speculativa a partire dai risultati che Kant teorizza con il giudizio riflettente della terza *Critica* e che, come sostiene la Arendt, hanno eminentemente a che fare con il politico⁴.

Tali temi saranno svolti ponendo sullo sfondo il metodo critico kantiano, analizzato dal punto di vista della sua applicazione non solo all'ambito fenomenico, ma anche alla scienza etica, alla riflessione sul diritto e sulla storia.

¹ Pasolini (2013: 26).

² Eraclito, *Framm. 108*, Cfr. Giannantoni (1969: 217).

³ Arendt (1978, trad. it., 151).

⁴ Cfr. Arendt (1982, trad. it., 27).

L'analisi prenderà di mira alcuni luoghi del 'trascendentale politico' kantiano, il che comporta il riconoscimento di uno statuto gnoseologico per quei principi a priori - cardini di una *constitutio civilis* fondata sul *pactum unionis civilis*: libertà, giustizia, uguaglianza -, secondo cui opera la ragion pura a priori legislatrice.

Come è noto, è stata Hannah Arendt, nelle sue *Lectures on Kant's Political Philosophy*, a rivendicare l'esistenza di una teoria del giudizio politico in Kant⁵, muovendo dall'interpretazione del Giudizio come pensiero rappresentativo, che si combina con l'immaginazione produttiva e riproduttiva della prima *Critica*, in modo da individuare le condizioni di possibilità e i criteri di validità e di oggettività del giudizio politico. Le lezioni della Arendt acquistano un rilievo paradigmatico nell'ambito di tutta la storiografia contemporanea, per la vivacità del loro sviluppo teorico e per l'esemplarità delle tesi sostenute. Una lettura, certo, che va ben al di là della lettera del testo kantiano, ma non perciò arbitraria e che sembra in grado di giustificare la derivazione del giudizio politico dal giudizio estetico riflettente della terza *Critica*.

Ora, grazie all'uso di categorie come identità e distinzione nella ricerca dell' 'universale' e del 'particolare politico', si vuole mostrare come e perché si possa parlare di una 'critica della ragion politica' in Kant.

Il rinvio kantiano, che compare nella chiusa della terza *Critica*, all'idea d'uno scopo finale nell'uso della libertà secondo leggi morali che abbia una realtà *soggettivamente* pratica, consente di individuare nella condotta etico-politica un termine medio che giustifica e comprende tale distinzione nella circolarità: soggettività-oggettività, individuo-libertà. Il luogo preposto a mantenere e giustificare tale distinzione è rappresentato proprio dal campo dell'esistenza e non dell'essenza dell'individuo, considerato non aprioristicamente, ma a posteriori. Ciò non contraddice, come vedremo, la singolarità dell'agire morale, né l'ambito del molteplice multivario in cui si sviluppano le particolarità della volontà.

Di fatto, secondo la Arendt, Kant giungerebbe a porsi la questione del senso comune a partire dalla seguente considerazione: nonostante il pensare sia un qualcosa di privato che sorge nella sfera singola di ogni uomo, quasi un dialogo fra sé e sé, di contro, per esser tale, la verità deve possedere una generale comunicabilità e, quindi, uno 'statuto' di intersoggettività.

⁵ Interpreti di una possibile teoria del giudizio politico sono, altresì, Vollrath (1977, 140-173) e Beiner (1983), intenti, però, a garantire un ambito e una funzione conoscitiva autonoma alla filosofia politica. Si veda in particolare Vollrath (1987, 11, 13-18). Per una lettura ermeneutica del fenomeno politico a partire dalla terza *Critica* di Kant, si veda, inoltre, Henry (1992).

L'attività del pensare, sostiene Kant, non può funzionare al di fuori della società umana. La compagnia è indispensabile per il *pensatore*. Lo statuto pubblico del pensiero comporta, come principale conseguenza, l'allargamento del proprio modo di pensare, coinvolgente, dunque, il pensiero altrui, realizzantesi attraverso il paragone dei propri con gli altrui giudizi. Ne consegue un'altra considerazione che concerne il carattere disinteressato della prima operazione di astrazione dalla propria soggettività: il pensiero assume il carattere di imparzialità.

In altri termini, i temi della *Critica del Giudizio* – il particolare, sia esso un fatto della natura o un evento della storia; il giudizio, come facoltà della mente umana che lo concerne; la socievolezza dell'uomo come condizione del funzionamento di questa facoltà, cioè l'idea che gli uomini dipendono dai loro simili non soltanto perché hanno un corpo e dei bisogni fisici ma proprio in virtù delle loro facoltà mentali – questi temi, tutti di eminente significato politico, vale a dire rilevanti per il politico, avevano occupato Kant ben prima che egli, ormai vecchio, si volgesse ad essi avendo portato a compimento la sua “fatica critica”⁶.

Coerentemente alle sue precedenti ricerche, miranti alla comprensione dell'originario e autentico significato dell'agire umano e del politico, Hannah Arendt cerca qui di operare una sorta di *decostruzione* del carattere metafisico e teoreticistico del pensiero politico kantiano, così come la storiografia critica ce l'ha tramandato. È proprio questo Kant che permette alla Arendt di superare il solipsismo etico in cui Heidegger aveva rinchiuso l'agire, dischiudendo quella dimensione intersoggettiva, inter-azionale, comunicativa che era rivendicata fin da *Vita activa*⁷ come la dimensione propria dell'uomo.

Non è dunque un caso che il Kant della Arendt sia innanzi tutto il Kant della *Critica del Giudizio*. Infatti, il giudizio presuppone un pensiero rappresentativo, la capacità cioè di rappresentare la propria mente e i punti di vista altrui. Il giudizio, argomenta la Arendt, così come proviene a noi dalla tradizione filosofica, è concepito come pratico-politico, cioè è un sapere che non pretende di essere rigorosamente scientifico, né si svincola nell'arbitrarietà dell'opinione puramente soggettiva. È un giudizio che segue l'agire al fine di ampliarne i caratteri di comunicatività, quindi cresce insieme al politico. La natura di un siffatto giudizio è quella dell'opinione basata sul ‘render conto’. Quale precedente storico, quindi, poteva meglio vestire questo abito, se non la facoltà del giudizio riflettente nel suo coincidere perfettamente con il concetto di *enlarged thinking* (pensare allargato) che è fondamento per il politico?

⁶ Arendt (1982: 27).

⁷ Cfr. Arendt (1958, trad. it., 3).

2. Il ruolo dell'analogia nella tematica estetico-conoscitiva

Hannah Arendt ha individuato, dunque, i caratteri del giudizio estetico-riflettente idonei ad essere applicati all'agire politico nel modo che si è appena brevemente ricostruito. La massima della sua prescrizione, 'il pensarsi al posto di un altro', desoggettivizza la prospettiva che lega ogni singolo individuo ad una situazione di interesse particolare, orientando verso l'universalità: «Esso non indica, come l'intelletto, ciò che è vero in assoluto, bensì – come Kant dice – ciò che conta (*Worauf kommts an?*)»⁸. Viene così riconfermata l'esigenza di mantenere la diversità nell'unità: la Arendt, nell'attingere l'universale teoretico della verità nelle *doxai* particolari, non le annulla in esso, poiché l'universale politico non si raggiunge togliendo le opinioni particolari, ma solo mantenendole, garantendole nella «pluralità del politico»⁹. Merito della Arendt sarebbe stato, dunque, quello di illuminare i percorsi dei soggetti immersi nella loro dimensione storica, cogliendone, appunto, i limiti, senza annullare la finitezza della condizione umana¹⁰.

⁸ Volpi (1987: 89).

⁹ Vollrath (1987: 33-35); Cfr. inoltre, Held (1981). Della stessa opinione sembra essere anche Passerin d'Entrèves (1987: 155-169); questi, dopo aver ricostruito lo specifico del pensiero politico arendtiano, sottolinea la derivazione speculativa della sua teoria del giudizio dalla teoria estetica kantiana. Vale a dire che «il giudizio politico è libero nella stessa misura in cui per Kant lo è il giudizio estetico, in quanto i partecipanti possono decidere liberamente se una specifica proposta è valida, ragionevole, ovvero degna di essere sostenuta». Continua, infatti, sostenendo che «nel pensiero politico della Arendt la attività pubblica e l'intera sfera delle azioni umane vengono situate fermamente nell'ambito della *doxa*, sebbene non mera od occasionale, e si reggono interamente sul *giudizio*. A differenza di molti filosofi politici che hanno denigrato questa dimensione della esperienza umana e hanno cercato criteri di verità che potessero redimerla, la Arendt la accetta senza esitazione. Per lei, infatti, il costringere la prassi e la politica a rispondere a dei criteri di verità è fuorviante, perché significherebbe che i cittadini non sarebbero più tenuti a formulare dei giudizi, a coltivare una mentalità aperta e a ingaggiare le proprie idee in un ambito di discussione e deliberazione pubblica, poiché la verità è per sua natura coercitiva e lascia ben poco spazio al giudizio o all'opinione dei partecipanti», *ivi*, p. 165. In generale, sul rilievo del giudizio per la prassi cfr. Pleines (1983); mentre per una messa a fuoco dei concetti di 'regola' ed 'esempio' cfr. Buck (1979: 61-81).

¹⁰ Questo aspetto dell'operazione teoretico-speculativa compiuta dalla Arendt è stato valorizzato da Judith Shklar (Shklar (1977: 80-90) in cui riconosce alla Arendt il merito di aver individuato il legame profondo tra politica e filosofia con le facoltà del giudizio e della persuasione, (*ivi*, 89-90). Utile, a tal proposito è, inoltre, il contributo di Hutchings (1996), in cui il rapporto tra critica e politica, affrontato e discusso all'interno del pensiero di diversi teorici, quali, Habermas, Arendt, Foucault e Lyotard, si delinea come una sorta di pratica politica che avvicina ai temi teorici moderni e postmoderni.

L'ingegnosa lettura delle idee politiche di Kant svolta dalla Arendt, cerca di far derivare *una* politica dal giudizio estetico: in realtà, il suo insistere sul giudizio estetico in generale, o 'allargato', può essere anche inteso come una traccia lasciata da Kant verso la costituzione dell'ideale politico e rappresenterebbe, pertanto, un ponte che unisce il bello al pratico, rimasto a lungo insospettabile¹¹.

Nel giudizio politico, come nel giudizio estetico, insiste la Arendt, c'è uno scontro tra il principio secondo cui si agisce e il principio secondo cui si giudica, il cui riverbero è manifesto ad esempio ne *Il conflitto delle facoltà*, in cui Kant descrive il ruolo dello spettatore di fronte agli eventi del dramma dei grandi cambiamenti politici. Ma se la nozione di *sensus communis* si trasferisce dal giudizio estetico alla realtà politica, al punto che il pensiero allargato e la comprensione simpatetica del punto di vista degli altri diventano il centro della politica, c'è da chiedersi che cosa accade alla semplice verità morale, a cui Kant dà un reale e preminente peso politico, soprattutto quando afferma che «la vera politica non può fare alcun progresso, se prima non ha reso omaggio alla morale»¹².

È giusto allora collocare la filosofia politica di Kant in una regione mediana, vale a dire, dar rilievo a un giudizio intersoggettivo¹³ che si situa tra il puro soggettivismo e la verità oggettiva?

Secondo la Arendt, Kant si era reso conto che la sua filosofia morale non poteva aiutarlo nel dare forma ad un ordine politico, e attribuisce al giudizio di gusto il carattere della comunicazione in quanto accordo interpersonale, una via d'uscita puramente soggettiva che si appella ad una nozione di 'gusto morale' che agisce da tramite nel collegare soggetti giudicanti e giudizi condivisi.

Il giudizio riflettente è intrinsecamente intersoggettivo, in quanto tiene conto *a priori* del modo di rappresentare di tutti gli altri, ponendoci al posto di ciascuno di loro nella valutazione dell'oggetto¹⁴. Il punto è che lo stesso Kant non voleva 'ricostruire' l'orizzonte morale, voleva soltanto illuminarlo e far sì che la politica gli rendesse omaggio. Kant contava sulla presunzione dell'oggettività morale, nella forma dell'imperativo categorico che comanda il rispetto nelle persone come fine in sé.

¹¹ Per la seguente lettura interpretativa ci si riferisce a Riley (1992: 305-323).

¹² Kant (1795, trad. it., 329).

¹³ Particolare rilievo assume la lettura che Hans-Georg Gadamer intese attribuire al concetto kantiano di senso comune e del suo rapporto con il giudizio di gusto in *Verità e metodo*, tutta incentrata sulla tesi della soggettivizzazione kantiana dell'estetica.

¹⁴ Flynn (1988: 128-140).

La questione allora si pone in questi termini: come può un ordine politico esterno, spesso coercitivo ma legale, rendere omaggio alla morale kantiana, il cui centro è il sommo bene interiore e il rispetto per le persone?

A questo risponde chiaramente Kant che non ha trasformato la sua politica in un accordo universale (alla Rousseau): la giurisprudenza e l'etica si distinguono non tanto per i rispettivi *doveri* e *fini*, quanto per i loro differenti stimoli. Ciò comporta che ci si può sforzare di realizzare moralmente fini politici importanti – come il repubblicanesimo universale e la pace perpetua – persino in totale assenza di una volontà buona come il solo stimolo veramente morale, facendo in modo che l'amor proprio razionale dei cittadini (più precisamente l'autostima) conduca ad un ordine morale eterno in cui loro stessi devono sperare per ragioni più profonde della semplice auto-preservazione.

Se, come Kant dimostra nella terza *Critica*, la natura può essere valutata – e tuttavia mai conosciuta attraverso scopi e funzioni che la causalità meccanica trascura di spiegare –, se inoltre le persone, in quanto libere volontà, possiedono fini che esse cercano di realizzare e, necessariamente, vedono se stesse come fini ultimi della creazione e, infine, se l'arte esibisce involontarietà senza scopi che fanno di essa non la morale, ma il simbolo della moralità, è possibile, allora, che la speciale versione della teleologia del Giudizio possa soddisfare l'architettonica unità della ragione, a cui Kant spera di pervenire.

Kant assegna allo Stato lo scopo di realizzare, tramite leggi, fini morali (la pace in particolar modo) senza il vantaggio di stimoli morali. Se in Kant si sottolineano l'universalità, il pensiero allargato e il punto di vista degli altri e si negano invece fini e scopi oggettivi, come fa la Arendt, si perderà di vista ciò che caratterizza il discorso di Kant, il quale ritiene che la bellezza sia il simbolo della moralità¹⁵, e che la bellezza e la morale si riferiscano a un comune 'substrato' teleologico, senza mai pretendere di trovare una nuova dottrina morale e politica nel giudizio estetico. La Arendt, quindi, nonostante l'esattezza della sua intuizione sulla centralità della *Critica del Giudizio* nell'*opus* kantiano, opera una sorta di estetizzazione della politica, poiché considera esplicitamente solo la prima parte della terza *Critica* e solo implicitamente la seconda parte, la *Critica del giudizio teleologico*¹⁶. Al contrario,

¹⁵ Cfr. Guyer (1979).

¹⁶ Non tutti gli interpreti sono d'accordo con l'assimilazione del giudizio politico nel giudizio estetico. Richard Erstein, ad esempio, contesta alla Arendt il carattere non-cognitivo da lei attribuito alla nozione di giudizio politico, poiché appunto derivato dal giudizio riflettente kantiano, di cui ha esagerato i criteri di validità che lo sottendono. Ritiene, inoltre, che esagerando la distinzione tra verità e opinione con i loro criteri di validità, la filosofa tedesca sconfini in pseudo-argomentazioni (Erstein (1983)). Va tuttavia considerata l'opinione secondo cui la Arendt, nel suo testo, miri, kantianamente, ad individuare i criteri di validità del

tutta la terza *Critica* deve essere letta come la chiave per intendere il pensiero politico di Kant (e addirittura, qualora Goethe avesse ragione, come la chiave per l'intera filosofia critica!).

È aperta convinzione di Kant – è anzi la sua centrale convinzione politica – che moralità e politica debbano essere correlate, dal momento che la vera politica non può fare un solo passo senza prima aver reso omaggio alla morale. Allo stesso modo, comunque, egli illustra una decisa distinzione tra motivazioni morali e motivazioni che agiscono secondo la legge e insiste sul fatto che tali definizioni non devono mai essere confuse l'una con l'altra. Si pone quindi un problema: la politica, nel suo riflettere la morale, non può fare affidamento su cause morali, ma su quelle regolate dalla legge.

La morale deve essere correlata alla politica, senza che il significato di giustizia pubblica secondo la legge confluisca nel sommo bene e nel rispetto delle persone. La soluzione potrebbe essere la seguente: la morale e la giustizia pubblica secondo la legge devono essere correlate in modo tale che la morale deve modellarsi sulla politica (insistendo sulla pace perpetua e i diritti umani) senza diventare il movente della politica. Data la tensione tra morale e giustizia pubblica, tra ciò che deve essere correlato e ciò che deve rimanere distinto, si può individuare nella nozione di fine ciò che può dare soluzione al problema, costituendo un ponte tra i due piani.

La legge pubblica sostiene senz'altro alcuni fini morali; non per questo, come sottolinea Kant nella *Rechtslehre*, la giurisprudenza e l'etica si identificano; le distinguono, infatti, non tanto i loro differenti doveri, quanto i diversi stimoli. Va comunque notato che utilizzare i fini come una sorta di ponte che connette la morale al regno politico-legale non rappresenta una forzatura dello spirito della filosofia critica, dal momento che lo stesso Kant nella *Critica del Giudizio* cerca di superare l'estraneità tra regni diversi tra loro: innanzi tutto quello della natura e della libertà tramite la nozione di fine, lo scopo soggettivamente valido per l'umano giudizio riflettente, ma tra natura e libertà da un lato e l'arte dall'altro.

Kant fece questo affinché fosse più chiaro che la natura può essere giudicata (sebbene mai conosciuta) attraverso cause e funzioni che la causalità

giudizio politico, offrendo uno spazio idoneo in cui inserirli per argomentarli legittimamente. Analogamente a Ernststein, anche Habermas, in un primo intervento del 1977 (cfr. Habermas (1977: 3-24)), contesta fondatezza cognitiva alla formulazione del suo pensiero rappresentativo, in quanto ritiene che la Arendt, mantenendosi fedele alla distinzione classica tra teoria e pratica, fondi quest'ultima su opinioni e convinzioni che non possono essere rigorosamente vere o false, cadendo, pertanto, in un insormontabile abisso tra la conoscenza e l'opinione (*ivi.* 22-23); salvo poi, in un altro scritto, di poco successivo (Habermas (1980: 127-131)), ascriverle il merito di aver colmato quell'abisso e riconoscere la legittimità della sua lettura.

meccanica trascura di spiegare. Allo stesso modo, le persone ipoteticamente libere hanno fini che si sforzano di realizzare e vedono se stessi come il fine ultimo della creazione. Sicuramente, poi, se i fini (qualche volta confinati in un ruolo riflessivo e regolativo) possono essere pensati come vincoli che legano la natura, la libertà umana e l'arte, ne discende che si possono unire i due lati della libertà umana: il regno della morale e il regno della legge.

Si ammette, certamente, che la continuità non sia perfetta, poiché Kant dice che la mancanza di fini è soltanto un principio riflettente quando è usato nel giudicare natura e arte, mentre la morale, di contro, ha fini oggettivi che sono scopi della ragione e che ognuno deve avere. Ma una possibile continuità è stabilita quando Kant ricorda che ci deve essere un terreno di unità del sovrasensibile che giace come base della natura e che il concetto di libertà di fatto contiene. In un linguaggio più diretto, è come dire che un'intelligenza diversa dalla nostra potrebbe vedere nella natura fini oggettivi come quelli che la nostra ragione conosce attraverso la legge morale. Questo vuol dire che il giudizio aiuta a ristabilire quell'unità della ragione che è sempre l'interesse centrale kantiano.

Ora, se il sommo bene nel regno della morale pura è costruito per intendere un massimo d'azione mai universalizzato che dimentica di rispettare le persone come aventi fini in se stessi, allora la morale e la politica possono essere unite alla teleologia kantiana. Se tutte le persone hanno una buona volontà, si dovrebbero rispettare l'un l'altra come fini, in verità come membri di un regno dei fini, poiché una natura razionale così come una persona è non un fine da dover essere realizzato, ma un fine esistente indipendentemente.

Se il Giudizio è decisivo per la filosofia politica di Kant (non perché una 'critica della ragion politica' possa essere in qualche modo segretamente contenuta nel suo concetto di giudizio estetico, ma perché la terza *Critica* stabilisce ciò che si intende con fini), allora la lettura di tale opera può risolvere il paradosso della politica di Kant. La vera politica deve rendere omaggio alla morale, ma senza fare affidamento su uno stimolo morale, cioè il sommo bene.

Ciò, se fosse efficace realmente, trascenderebbe in senso puro la politica e, insieme, produrrebbe una comunità etica sotto le leggi della virtù. Allo scopo di dare senso al paradosso che rende la politica aderente e attenta alla morale, è necessario indicare la distinzione kantiana fra fini e scopi e dire che il repubblicanesimo e la pace perpetua potrebbero fungere bene da fini morali che sono raggiungibili legalmente (o alla fine accessibili) tramite l'amor proprio dei cittadini di una Repubblica, anche se quei cittadini fossero dei 'diavoli' privi del tutto del sommo bene.

Si potrà constatare più avanti come la lettura degli scritti politici renda ragione di questa analisi. L'attività trascendentale esposta da Kant nel corso delle sue tre *Critiche*, essendo rivolta alla comprensione dell'inesauribilità del pensiero, gli consente di aprirsi una strada tra la metafisica e il formalismo. Evitando di percorrere le due strade in modo parallelo, egli ha cercato di costruire una filosofia come scienza autonoma capace di studiare la realtà piuttosto che la possibilità.

3. La natura del giudizio politico

Si è visto fin qui come il giudizio riflettente debba essere inteso non come una specie particolare di giudizio, ma come la *sintesi* delle condizioni soggettive del giudizio in generale. Ora, secondo la *Critica della ragion pura*, l'unità puramente formale della natura è il prodotto di un giudizio determinante che applica le categorie alle pure intuizioni di spazio e tempo attraverso schemi trascendentali, mentre le specificazioni e la sistematizzazione delle leggi empiriche dipendono dall'uso regolativo delle idee della ragione. Così, sebbene nella prima *Critica* Kant non consideri esplicitamente il ruolo del giudizio riflettente nel rendere possibile il giudizio determinante, egli attribuisce alla ragione funzioni che saranno in seguito attribuite al giudizio riflettente nella produzione e sistematizzazione della conoscenza empirica. E non è tutto: nell'introdurre lo schematismo, Kant discute la necessità di un'armonia tra sensibilità e intelletto per la possibilità stessa del giudizio, e discute di come il giudizio morale sia soggettivamente determinante e al tempo stesso oggettivamente, cioè praticamente riflettente, in modo che l'individuo, agente libero, possa rendere la massima soggettiva adeguata alla razionalità dell'imperativo categorico, e insieme capire la molteplicità delle leggi empiriche della volontà.

Ma è solo nella trattazione del giudizio estetico che è venuto in primo piano il ruolo fondamentale del giudizio riflettente, non essendo occultati gli aspetti soggettivi del giudizio da fini oggettivi cognitivi e/o razionali, vale a dire la conoscenza degli oggetti o la determinazione dell'azione morale.

È dunque al giudizio riflettente che dobbiamo volgerci per cogliere, in conclusione, la natura del giudizio politico, come era apparso chiaro fin da quando erano state prese in considerazione le *Lectures* arendtiane. Come si è detto, Hannah Arendt ha individuato le condizioni di possibilità del giudizio politico ed evidenziato i criteri di validità ed oggettività, sostenendo la tesi secondo cui esso presuppone un pensiero rappresentativo, inteso come capacità di rappresentare il proprio punto di vista e insieme quelli altrui.

Il giudizio politico non pretende di produrre conoscenza scientifica rigorosa, ma non per questo cade nell'arbitrarietà dell'opinione puramente soggettiva: sua natura è quella dell''opinione basata sul render conto'; suo carattere specifico è la pubblicità come rivendicazione della socievolezza e della libertà 'pubblica'; suo protagonista è il *Weltbetrachter*, lo spettatore del mondo. Fin qui la Arendt.

Ma, come si è già potuto constatare, il discorso non può esaurirsi con le conclusioni della Arendt. Ritengo, infatti, che l'analisi della natura del giudizio politico nella filosofia trascendentale kantiana porti a porre alcune questioni che toccano la sua stessa unità sistematica e architettonica.

Generalmente il 'predicato' kantiano ha una funzione sintetizzatrice che esplica la propria attività nell'unificazione del molteplice puro, ponendosi contemporaneamente come unità interna del fenomeno sintetizzato. Avvenendo una simile applicazione automaticamente, emerge il carattere puramente meccanico del riferimento all'unità. Ma per un giudizio individuale, come quello politico (e storico) è necessaria una forma di unità che non cancelli le differenze qualitative che intercorrono tra gli elementi che costituiscono il molteplice multivario dei fatti storici e politici.

È possibile, allora, ravvisare tale funzione nel giudizio riflettente, che contribuisca a mantenere le differenze nell'unità?

Kant è esplicito nel dichiarare che l'accordo delle molteplici leggi della natura con la nostra facoltà di conoscere è un presupposto a priori del Giudizio allo scopo di riflettere su di essa secondo le sue leggi empiriche (*KU, Einleitung, § V*). La terza *Critica* è dunque il luogo in cui riconsiderare *ex novo* il problema della connessione delle percezioni empiricamente intese: il predicato del giudizio riflettente porta con sé una funzione distinguente, anche se le modalità con cui ciò avviene non sono chiarite fino in fondo da Kant. È il giudizio nella sua nuova funzione, che riflette su tale legge di specificazione, consentendo il costante riferimento alla conoscenza delle leggi empiriche della natura. Kant supera così l'omogeneità della sostanza e l'applicazione meccanica della sintesi a priori e aggiunge, altresì, la necessità del riferimento alle leggi singolarmente determinate, garantendo logicamente la possibilità della rappresentazione delle individualità, quindi anche delle individualità politiche e storiche.

La politica in Kant, dunque, può essere intesa come un'attività che realizza legalmente alcuni fini morali – caratterizzati in generale come pace e civiltà –, in un modo che è in realtà inferiore alla realizzazione di una comunità etico-universale, ma che (facendo affidamento per sua esplicita ammissione su motivi non-morali) rispetta le persone, i fini ultimi della creazione e

le preserva dall'essere usati come puri mezzi per scopi arbitrari (soprattutto la guerra).

Il giudizio politico, a sua volta, si rivela un giudizio regolativo come quello morale, ma al tempo stesso singolare, in quanto si applica nell'ambito contingente dell'individuale. In realtà, tutto questo si traduce in ciò che Kant afferma nei paragrafi 82-84 della *Kritik der Urteilskraft*, problematizzando la natura dell'uomo in quanto noumeno, poiché è fine della creazione e non solo soggetto della moralità, in cui possiamo riconoscere sia una facoltà soprasensibile, che è la libertà intesa come caratteristica sua propria, sia la legge della causalità, nonché quell'oggetto corrispondente che l'uomo si propone come fine supremo, vale a dire il sommo bene nel mondo. Ma lo scopo ultimo che la natura può ragionevolmente porre alla specie umana è la *cultura*, cioè la capacità, prodotta in un essere ragionevole, di proporsi fini arbitrari (e quindi – dice Kant – nella sua libertà).

È questa la condizione soggettiva principale della capacità di promuovere fini in generale, ma è insufficiente a promuovere la volontà nella determinazione e nella scelta dei suoi fini, che è parte essenziale della facoltà umana di proporsi degli scopi. La condizione formale sotto cui la natura può raggiungere il suo scopo finale è quella costituzione nei rapporti degli uomini tra loro, che in un tutto che si chiama società civile oppone un potere legittimo alle infrazioni reciproche della libertà, poiché solo in tale costituzione si può effettuare il massimo sviluppo delle disposizioni naturali.

Si profila così un passaggio intermedio rappresentato dal giudizio morale soggettivamente determinante e, al tempo stesso, oggettivamente riflettente. Esso ci aiuta a comprendere quella circolarità in cui, da sempre, nel pensiero morale kantiano sono incluse la legge morale (universale) e la volontà (particolare).

Il rinvio kantiano, nella chiusa della terza *Critica*, all'idea d'uno scopo finale che, nell'uso della libertà secondo leggi morali, abbia una realtà *soggettivamente* pratica, consente di individuare un termine medio che giustifica e comprende la distinzione nella circolarità: l'individuo con la sua libertà. E il luogo preposto a mantenere e giustificare tale distinzione è rappresentato proprio dal campo dell'esistenza e non dell'essenza dell'individuo stesso conosciuto, non più aprioristicamente ma, questa volta, a posteriori. Ciò non contraddice la singolarità dell'agire morale, né l'ambito del molteplice multivario in cui si sviluppano le particolarità della volontà. Allora il giudizio politico può essere così definito: *in termini kantiani chiamo giudizio politico la facoltà legislatrice a priori in grado di essere sia determinante (avendo l'universale può operare la sussunzione del particolare), sia riflettente*

(avendo il particolare può risalire di condizione in condizione fino all'universale). In altre parole: il giudizio politico è determinante nella misura in cui la ragion pura pratica, con le sue regole costituite dalle leggi morali, raggiunge validità oggettiva; è riflettente, nella misura in cui deve mediare tra volontà e legge morale, trovando il canone di validità oggettiva nelle soggettività dei singoli individui¹⁷.

Le implicazioni teoriche di tale presupposto mi spingono a non isolare icasticamente il giudizio politico, ma a verificare la sua azione nella contingenza, in quello spazio in cui si situa il discrimine tra ciò che può essere e ciò che può non essere. Il giudizio politico, di cui si è cercato di individuare le forme concettuali costitutive, rappresenta davvero la sintesi tra giudizio determinante e giudizio riflettente attraverso un *tertium* (come vuole la Arendt) inteso come simbolo, schema, esempio (come lo schema di Kant, dunque, come un prodotto dell'immaginazione) sottoposto ad uno schematismo soggettivo; è un *Faktum* della ragione riflettente inserito nell'ambito della riflessione trascendentale emersa nella *Critica del Giudizio*.

È un giudizio basato sull'esemplarità oggettiva che coniuga istanze di singolarità e universalità, strumento ermeneutico¹⁸ capace di cogliere il significato che ogni periodo storico esprime. Infatti, tutte e tre le *Critiche* di Kant partono da un preciso *Faktum* della ragione: la conoscenza matematica e la scienza della natura per la prima *Critica*, la legge morale per la seconda, la libertà della schematizzazione nella sintesi estetica e teleologica rispetto al molteplice empirico per la terza. Il giudizio politico può così essere assunto come il *Faktum* della ragione politica e si può provare a delineare la sua attività all'interno della parabola del divenire del giudizio storico kantiano.

¹⁷ Cfr. Macera (2001: 183-187).

¹⁸ Sullo statuto epistemologico della conoscenza storica cfr. Mancini (1986: 251), di cui val la pena rileggere quanto segue: «Mentre i fatti naturali sono oggetto di visualizzazione eidetica, per la conoscenza dei fatti storici si esige [...] una situazione ermeneutica, dove la posizione dell'ascolto e dell'accertamento non può essere disgiunta dalla totale precomprensione teoretica. Se le scienze mirano a un fondamento definitivo, metafisico in ultima analisi, cavando dal singolare la norma generale, astratta, l'essenza nella sua universalità, la conoscenza dei fatti storici non può ambire che ad una verità di tipo ermeneutico, sempre aperta e sempre integrabile. Il che comporta non solo uno *status* veritativo diverso da quello scientifico (esplorazione del senso *contro* visualizzazione eidetica); non solo precomprensione come filosofia *préalable*, ma anche quella precomprensione indisciungibile dalla precedente, che significa l'insufficienza dell'elemento soggettivo della scelta, non solo, ma l'impegno integrale del soggetto pensante; non solo, ma quell'anticipazione rischiosa, inesaurita, sempre rivedibile che è l'anticipazione globale del significato».

Il giudizio politico kantiano non giace in solitudine, come se fosse trasfigurato dalla roboante dialettica del linguaggio kantiano, non deriva da presupposti ontologici, ma scaturisce da presupposti storico-esistenziali.

Il problema politico, altresì, può determinarsi solo appoggiandosi su una solida e valida dottrina del diritto, sintesi dinamica di momento naturale e momento etico. È così che mi riferirò brevemente solo ai seguenti scritti, rimandando ad altro momento una più estesa trattazione di altri testi kantiani. Esaminerò criticamente: *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht* (1784), *Die Religion innerhalb der Grenzen der blossen Vernunft* (1793).

4. Identità e distinzione nella ricerca dell'universale politico: *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht* (1784)

L'importanza storica di questo scritto risiede innanzi tutto nella sua collocazione storica all'interno del *corpus* degli scritti politici. Esso segna, in un certo senso, l'affermarsi dell'interesse kantiano per una 'metafisica della storia', come afferma Cassirer¹⁹. Si evince qui un nuovo punto di vista nella considerazione del mondo di fronte al decorso dell'esistenza storico-empirica. È da questo momento in poi che in Kant si passa dal *Sein* al *Sollen*; egli guarda la storia non soltanto come successione temporale di singoli momenti che si connettono causalmente, ma come serie di eventi da ricondurre all'unità ideale di un fine immanente.

Ed è proprio nel divenire (*Werden*) che Kant proietta il suo nuovo molteplice, superando la posizione di infinità statica della molteplicità delle manifestazioni della vita umana espressa nelle *Ideen* di Herder, testo che pure aveva motivato e influito sull'elaborazione della sua *Idee*.

Da una prima lettura dello scritto si evince che il molteplice empirico, di cui si devono ricostruire i contorni nell'opera politica di Kant, consiste nei singoli individui che vanno riuniti nella totalità della specie, mantenendo le loro specifiche tendenze originarie e i loro fini particolari. L'unità della specie si realizza attraverso la sussunzione del particolare (singolo individuo) ad un principio regolatore della specie stessa. In realtà però tale molteplice è regolato dal principio della finalità della natura che, inavvertitamente, cioè quasi automaticamente, è principio regolatore e razionale dei molteplici aspetti fenomenici che caratterizzano l'agire e il vivere dei singoli individui. Nell'analizzare trascendentalmente lo scritto *Idee zu einer allgemeinen Geschichte in*

¹⁹ Cassirer (1918, trad. it. 257-275).

weltbürgerlicher Absicht, ritengo che il problema principale consista nella disamina della validità del principio analitico kantiano, che consente allo stesso Kant di scoprire quel disegno della natura che funge da filo conduttore di una «storia che si svolga secondo un piano naturale determinato»²⁰. Che potere riveste, infatti, l'analisi di Kant dell'universalità della storia?²¹

L'analisi ha il potere di mettere in luce la presenza di un processo indefinito che si articola nel mondo fenomenico, penetrandolo fin nei suoi principi o elementi indivisibili, anche se, come si sa, il procedimento analitico ha in sé il carattere di identità omogenea o formale che rende possibile la presenza del processo stesso. Ci si deve chiedere, tuttavia, se possa essere governato dalla ragion pura speculativa, poiché solo ciò che la ragione «trae interamente da se stessa non può rimaner celato, ma per opera della ragione viene alla luce, appena scoperto il principio generale che la governa»²². Bisogna ora verificare se nello scritto Kant sia giunto o no a scoprire il principio generale che governa la ragion pura speculativa, che ruolo svolga nella configurazione di una storia universale e in che consista questo procedimento analitico.

Un primo dato che emerge è costituito dal metodo analitico che Kant applica nella formulazione delle sue nove tesi. Secondo tale metodo analitico egli parte da ciò che è già dato, la storia, per poi risalire alle sue condizioni. Il procedimento critico non può operare utilizzando un'analisi disgiunta dalla sintesi o una sintesi disgiunta dall'analisi, non può produrre nulla se i due procedimenti rimangono isolati.

Si era partiti, tuttavia, dall'affermazione per cui l'analisi è attività intrinsecamente omogenea e perciò identica. Ma, se ci si fermasse alla constatazione che Kant impiega il metodo analitico in modo omogeneo, non si riuscirebbe a spiegare come egli possa giungere alla sua idea di storia universale. In realtà l'operazione critica (ed è questo il piano ermeneutico sotteso agli scritti politici) che Kant compie, è un nuovo utilizzo del procedimento analitico che necessita, pertanto, di una riconsiderazione e rilettura.

²⁰ Kant (1784: 18, trad. it. 124).

²¹ È sintomatico, infatti, che Kant esordendo nella tesi prima dello scritto secondo cui «tutte le disposizioni naturali di una creatura in quanto destinate un giorno a svolgersi in modo completo e conforme al loro scopo», poi dica: «Ciò è confermato in tutti gli animali dall'osservazione così esterna *come interna o analitica* [cors. mio]. Un organo che non dev'essere usato, un ordinamento che non raggiunge il suo scopo sono contraddizioni nella dottrina teleologica della natura. Poiché, se noi prescindiamo da questo principio fondamentale, non abbiamo più una natura regolata da leggi, ma un giuoco senza scopo, e il caso sconcertante regnerebbe in luogo della guida della ragione» (Kant (1784: 18, trad. it. 125)).

²² Kant (1781: 13, trad. it., 11).

A seguito dei risultati speculativi ottenuti con le tre *Critiche*, il procedimento analitico subisce in Kant una revisione; non resta più fermo, quasi immobilizzato nel suo stesso processo di indagine in cui la reale multivarietà dei fenomeni irrimediabilmente si perde. Secondo l'insegnamento fondamentale della *Kritik der Urteilskraft* è possibile, tramite il giudizio riflettente, ricondurre il molteplice multivario o «terzo molteplice»²³ dell'esperienza concreta ad unità, mantenendo di quel molteplice le caratteristiche di peculiare varietà interna. Il procedimento analitico kantiano, quindi, per evitare il rischio di annullare nell'attività identicamente omogenea la varietà dei suoi oggetti, si fa a sua volta plurale e multivario: più esattamente, il carattere non identificante e omogeneizzante di tale analisi è garantito dalla capacità di cambiare all'infinito la propria struttura costitutiva.

Nell'*Idee* si può sperare, secondo Kant, di trovare un ordine ben preciso nel disarmonico gioco delle volontà umane che agiscono secondo libertà e, a volte, secondo un principio che non sembra evidente ma spesso confuso e irregolare. Nella molteplicità empirica dei singoli individui tale ordine non sarà mai evidente, ma se si prova a scoprirlo nell'unità della specie, forse la storia sarà in grado di narrare le manifestazioni dei soggetti empirici; le azioni umane considerate nella totalità della specie possono essere descritte in maniera determinata. Nell'affermare questo *Faktum* della ragione, Kant dà concretezza al processo infinito di cui il metodo analitico postulava l'esistenza. Il processo infinito, infatti, è ciò che garantisce la reale concretezza del processo storico, perché solo nelle costanti modificazioni dell'agire umano trova concretezza il reale.

Quando parla di «avanzamento»²⁴ che gli uomini promuovono con il loro agire, in realtà Kant si riferisce al cammino dell'umanità nella storia che si svolge secondo un piano naturale e determinato. Per giungere a questa soluzione egli ha necessariamente preso in considerazione l'infinità con cui il metodo analitico procede nell'indagine della storia dell'umanità. Un'analisi che mostra come la realtà sia contemporaneamente formata dalla presenza del

²³ *Terzo molteplice* è la formula usata da Luigi Scaravelli (1955) e che qui riutilizzo ermeneuticamente, per indicare il molteplice delle leggi empiriche con cui si apre la *Critica del Giudizio* e che deve essere distinto dal molteplice empirico delle sensazioni e dal molteplice puro costituito dalle determinazioni spazio-temporali che emergono dalla prima *Critica*.

²⁴ Testualmente Kant si esprime così: «Singoli individui ed anche interi popoli non pongono mente al fatto che, pur perseguendo i loro particolari fini, ognuno a suo modo e spesso in contrasto con gli altri, procedono in realtà inavvertitamente secondo il filo conduttore di un disegno della natura e promuovono quell'avanzamento che essi stessi ignorano e al quale, se anche lo conoscessero, non farebbero gran caso» (Kant (1784: 17, trad. it. 124)).

piano meta-noumenico, dalla ragione che lo pensa, dalla multivarietà dei fenomeni che si producono temporalmente, dal realizzarsi della volontà.

Nell'intitolare il proprio scritto Kant ha usato l'espressione «storia universale». Soffermiamoci un attimo sul significato di «universale» e interrogiamoci sul valore attribuito al termine. Deve essere inteso come estensione logica, come giudizio il cui soggetto è la totalità degli individui? oppure *ex causis*, aprioristicamente?

Se ci si stesse muovendo in sede di *Critica della ragion pura*, non si esiterebbe nella risposta a riconoscere l'*Allgemeinheit* come attributo della storia. In realtà esso vale come significato in un certo senso metonimico, di senso unitario degli eventi, del loro accadere in un disegno globale che non tiene conto della mera consequenzialità delle *res gestae*.

È una scienza storica che ambisce ad essere storia filosofica, che prescinde dall'analisi empirica, pur prendendo motivo da essa. È la difesa del particolare nell'universale.

Che con questa idea di una storia universale avente in certo qual modo un filo conduttore *a priori* io abbia voluto toglier di mezzo l'elaborazione della storia propriamente detta, concepita in maniera puramente *empirica*, sarebbe un fraintendere la mia intenzione. La mia è solo un'idea di ciò che una mente filosofica (che del resto dovrebbe conoscere assai bene la storia) potrebbe ancora tentare da un altro punto di vista. D'altra parte la cura dei particolari, per altro lodevole, con cui si scrive ora la storia del proprio tempo, deve portare ognuno naturalmente a chiedersi come i nostri tardi nipoti potranno abbracciare la quantità di fatti che noi tra qualche secolo avremo tramandato loro²⁵.

È in tale finalità epistemologica che Kant, in questo scritto, fa confluire tutto ciò che si è ricondotto sotto l'espressione di 'terzo molteplice', vale a dire: la libertà del volere umano (nell'espressione del suo concetto di *Antagonism* come *ungesellige Geselligkeit der Menschen*) con la *Willkür* (cioè la volontà liberamente determinata di cui sopra) e la determinazione dei fenomeni non della natura, ma della storia, secondo leggi naturali²⁶. Anche l'enfasi posta nel testo delle singole tesi (lo spaziato nella versione originale) intende far risaltare gli aspetti teorico-argomentativi dell'analisi.

Kant vuole così dar risalto non più al singolo uomo, ma alla specie umana²⁷, non solo in senso storico-naturalistico, ma anche storico-politico, in

²⁵ Kant (1784: 30-31, trad. it. 138-139).

²⁶ Galston (1975). È opinione dell'autore che il pensiero di Kant sulla storia non sia laterale alla sua filosofia, ma si situi al suo stesso centro, ecco perché si evidenzerebbe tutto il suo sforzo nel far convergere il dualismo di moralità e scienza naturale.

²⁷ Si rinvia a tal proposito a Kain (1989: 325-345), secondo cui l'importanza del pensiero politico di Kant può essere meglio compresa se facciamo due cose: se lo paragoniamo alla

quanto investe la sfera morale fatta di doveri universali della socialità e del politico, per il pieno sviluppo della *Selbstschätzung*²⁸.

Solo in quella società in cui si attui, da un lato, la massima libertà (e quindi un generale antagonismo dei suoi membri) e, dall'altro lato, la più rigorosa determinazione e sicurezza dei limiti di tale libertà (affinché essa possa coesistere con la libertà degli altri), il supremo fine della natura, lo sviluppo di tutte le facoltà, possono essere raggiunti nell'umanità.

La natura vuole ancora che l'umanità debba attuare da se stessa questi così come tutti gli altri fini della sua destinazione. Perciò una società, in cui la libertà *sotto leggi esterne* vada congiunta nel più alto grado possibile con un potere irresistibile, cioè una *costituzione civile* perfettamente giusta, è il compito supremo della natura nei riguardi della specie umana; poiché solo assolvendo e attuando tale compito la natura può raggiungere tutti gli altri suoi fini nei riguardi della nostra specie²⁹.

È in questo scritto, e pertanto nel campo della storia umana, che Kant, come osserva Vlachos³⁰, abbandona le sue riserve critiche per un uso 'costitutivo' della teleologia, dando spiegazione dell'assunto esposto nell'introduzione alle nove tesi in cui, appunto, si propone di «scoprire nel gioco della libertà umana, considerato in grandi proporzioni, un ordine per cui ciò che nei singoli individui si rivela confuso e irregolare, nella totalità della specie possa riconoscersi come sviluppo continuato e costante, anche se lento, delle sue tendenze originarie»³¹. Per Vlachos, dunque, l'idealismo trascendentale kantiano, eliminando chiaramente la teleologia negativa della storia di Rousseau³², non guarda all'incondizionato, all'astratta e inaccessibile totalità per la comprensione umana delle condizioni dell'intera conoscenza.

Queste stesse condizioni, teleologicamente articolate ed ordinate, servono per tracciare allo stesso tempo, attraverso i fatti della storia, il percorso

teoria politica come essa esisteva prima di Kant e se vediamo come essa dipenda, fondamentalmente, dalla sua filosofia della storia. È la filosofia della storia di Kant, afferma l'autore, che gli consente di fare un passo in avanti dai precedenti pensatori politici. Kant metterebbe insieme, per la prima volta, due progetti che sono rimasti tradizionalmente separati, sviluppando una teoria dello stato ideale e, inoltre, una filosofia della storia, cioè una teoria su come, attualmente, realizzare lo stato ideale nella storia.

²⁸ Kant (1784: 20, trad. it. 126). A tal proposito segnalo l'interessante lettura di Onora O'Neill che spiega come per Kant la storia volga lo sguardo al passato mentre la politica apra le porte per il futuro, poiché solo la politica può portare a termine ciò la ragione speculativa ha cercato di rischiarare in modo incompleto. (Cfr. O'Neill (1992: 71-72).

²⁹ Kant (1784: 22, trad. it. 128-129).

³⁰ Vlachos (1962 : 181-192).

³¹ Kant (1784: 17, trad. it. 123).

³² Tenenbaum (1974: 343-392).

di un'evoluzione ascendente empiricamente verificabile. Kant sarebbe indotto a considerare il naturalismo come la base della teleologia critica, che finisce per invadere progressivamente la filosofia della storia, poiché, argomenta Vlachos, la concezione del progresso che è propria dell'idealismo kantiano non porta avanti una selezione razionale dei fatti della storia, ma piuttosto una concezione antinomica dell'esperienza storica.³³ In virtù di questa presa di posizione l'«histoire systématique» di Vlachos, meta essenziale della teleologia critica, diviene propriamente una storia naturale dell'umanità e il progresso, come Kant lo concepisce, è un «devenir au sein du quel le réel est impliqué inexorablement avec tout son être»³⁴.

L'idealismo si afferma così come un realismo assoluto: partendo dal *Leitfeiden* della *Naturabsicht*³⁵, si amplia e si sviluppa in una dialettica filosofica della storia e raggiunge infine l'universale politico, che era il fine perseguito nel nostro assunto iniziale nella determinazione delle *anomalie razionali*.

5. Identità e distinzione nella ricerca del particolare politico: *Die Religion innerhalb der Grenzen der blossen Vernunft* (1793)

Riflettendo sul rapporto tra *Die Religion* e il sistema critico kantiano, si vuole riflettere sul ruolo del particolare politico quale momento costitutivo del giudizio politico. Lo scritto non rappresenta esclusivamente la «dottrina filosofica pura della religione»³⁶, come Kant afferma nella *Prefazione alla prima edizione* (1793) costituisce piuttosto la prosecuzione del *Leitfaden* programmatico esposto nella *Logik* e poi nella lettera del 4 maggio 1793 a Carl Friedrich Stäudlin. In questa lettera Kant pone le famose quattro domande fondamentali, di cui la terza: «Che cosa mi è lecito sperare?», rappresenta il suo tentativo di non chiudere il programma critico, anzi di continuarlo sulla

³³ Vlachos (1962 : 192).

³⁴ *Ibidem.* p. 194.

³⁵ Nella tesi nona, infatti, emerge il tentativo di elaborare una storia universale secondo ciò che Kant stesso chiama intenzione o disegno della natura, il cui scopo sia, per l'appunto, un'unione civile del genere umano, tentativo da ritenersi non solo possibile, ma tale, anzi, da promuovere e sostenere l'intenzione primaria della natura. È l'idea, tutta kantiana, della possibilità di una filosofia della storia che individua il fine della natura unitamente a quello della storia nel genere umano inteso come *unionis civilis*.

³⁶ Kant (1793: 10, trad. it. 331).

scia di quel percorso etico-teologico³⁷ che la «omnitude collective»³⁸ (*allgemeine Vereinigung*) emersa con la terza *Critica*, gli aveva offerto. Considero, pertanto, lo scritto a partire dalla sua funzione metodologica all'interno del *corpus* delle opere kantiane, vale a dire leggendolo nella sua posizione simmetrica rispetto alla *Critica del Giudizio*, con cui si lega, a mio avviso, sia per intenti programmatici sia, come si mostrerà, per finalità intrinseche legate al rilievo da attribuire all'attività innovativa della riflessione trascendentale, così come è venuta emergendo nelle pagine precedenti nella ricerca in questo caso non dell'universale, ma del 'particolare' politico in Kant.

L'interrogativo: «Che cosa mi è lecito sperare?», relativo alla religione, racchiude in sé intrinsecamente il quesito su come ciò che è oggetto del dovere possa diventare realtà; così «si pone, dal punto di vista della prassi, il compito di mediazione tra natura ("realtà") e morale ("ciò che è dovuto")». Kant tratta questo compito di mediazione ancora una volta, dal punto di vista del conoscere, nella *Critica del Giudizio*³⁹.

Tralasciamo adesso la prima delle due possibili declinazioni della speranza, cioè la libertà esterna, oggetto della filosofia della storia, e concentriamoci sulla libertà interna, oggetto della filosofia della religione. Più esattamente, ciò che interessa non è l'oggetto particolare della filosofia della religione, ma il fatto che la filosofia della religione si venga a trovare in un ambito contingente, poiché il suo oggetto non ha realtà oggettiva: la religione, considerata soggettivamente, non è oggetto del sapere, ma piuttosto di una speranza fondata filosoficamente.

È importante tener fermo che Dio è un postulato della ragion pura pratica, poiché, «in accordo con una convinzione di fondo dell'illuminismo, Kant prosegue la trattazione della filosofia morale nella propria filosofia della religione. Dopo l'estinzione della cosmoteologia metafisica, il concetto di Dio appartiene eminentemente all'etica»⁴⁰. Pertanto, per Kant l'anima (immortale) e Dio sono oggetti reali non del mondo empirico, ma di quello morale. Questo significa sottolineare la 'deficienza ontologica' della natura umana, cogliere l'uomo nella sua fenomenicità, quale l'esperienza ce lo fa conoscere⁴¹, mettendo a nudo la lotta, congenita alla sua natura finita, tra bene e

³⁷ Kant (1790, trad. it. § 86).

³⁸ Kant (1793: 157, trad. it. 486).

³⁹ Höffe (1983, trad. it. 225).

⁴⁰ Höffe (1983, trad. it. 231-232).

⁴¹ Kant nel criticare la risposta negativamente rigorosa degli antichi filosofi moralisti rispetto agli interrogativi per cui sia necessario insegnare la virtù e se esista più di una virtù, afferma in nota che: «[...] se si vuol dare un giudizio morale su quell'essere morale, che è l'uomo nel

male, e, quindi, l'irresistibile amore del proprio io empirico. Aspetti fondamentali per la ricerca del particolare politico.

«L'ontologico e l'empirico si congiungono così, pur opponendosi, per completarsi»⁴². In quest'ottica, si può facilmente vedere come, nonostante la tesi più forte dello scritto sia l'accentuazione della lotta fra il bene e il male, con il prevalere dell'uno sull'altro, motivo incessantemente propulsivo dell'ascesa alla moralità, nella vita etica di ogni uomo torna ripetutamente nell'opera un altro aspetto su cui credo valga la pena riflettere.

Qui Kant sembra compiere un salto logico-metodologico, passando da una dimensione antropologico-individualistica a una dimensione comunitaria. Tale passaggio si può cogliere soprattutto nell'impostazione argomentativa data da Kant alle *Annotazioni generali*, intese dal filosofo come parti confinanti alla dottrina che sta esponendo e che definisce, difatti, «*parerga della ragione nei limiti della ragion pura*». Nell'intento di rafforzare le proprie tesi, egli attribuisce carattere storico alle dottrine religiose, ricorrendo al piano storico-empirico per mostrare gli elementi irrazionali delle religioni storiche. In questo modo Kant può mantenere la molteplicità non indifferenziata nell'unità, può cioè ricondurre la molteplicità ad unità nei termini di una vicenda storico-universale, di un progresso verso l'instaurazione del regno di Dio, fatto concreto dall'istituzione di una *chiesa universale*, che «riveste la forma di uno Stato etico retto da Dio»⁴³ e che altro non è se non la *comunità etica* che dovrebbe compiere la storia.

Uso il condizionale poiché, come ben si sa, questo resterà per Kant un puro ideale che non verrà mai pienamente realizzato⁴⁴.

Kant, ponendo la questione della differenza fra fede razionale e fede storica, con il richiamo alla voce interiore⁴⁵, esorta a saldare idealmente ciò che,

fenomeno [cors. mio], ossia come ci è fatto conoscere dall'esperienza, si può rispondere affermativamente a tutt'e due le questioni; perché in tal caso l'uomo non è giudicato sulla bilancia della ragion pura (in un tribunale divino), ma con una misura empirica (da un giudice umano)» (Kant (1793: 25, trad. it. 344)).

⁴² Weil (1962: 1-32). Più precisamente: «Ontologicamente l'uomo è moralmente insufficiente; empiricamente si dimostra criminale proprio perché, nel suo fondo, rimane libero e non può perdere la coscienza della legge morale, legge della propria libertà. Ontologicamente è chiamato al Bene; empiricamente si presenta come un uomo che ha scelto» (p. 165).

⁴³ Kant, (1793: 124, trad. it. 452).

⁴⁴ L'universalità è paradossalmente 'circostritta' al cristianesimo, mentre l'unità, all'interno della storia del cristianesimo viene data, ancora paradossalmente, dalla constatazione del continuo differire e del continuo allontanarsi della religione storica da quella razionale, (cfr. Olivetti (1973: 103-133, xxx).

⁴⁵ Mi riferisco alla chiusa della *Partizione seconda*, «Rappresentazione storica della graduale fondazione del dominio del buon principio sulla terra», in cui Kant nella sua attesa millenaristica alla domanda: «Quando arriverà dunque il regno di Dio?», risponde seraficamente,

nel corso di tutta l'opera critico-trascendentale, ha tenuto tenacemente distanti: *phaenomenon* e *noumenon*, attraverso il ricorso alla dimensione comunitaria, di cui *Die Religion* si fa promotrice.

Va avanzata, comunque, un'importante precisazione: dalla constatazione che le religioni abbiano un carattere storico non segue affatto che della religione sulla terra possa farsi una *storia universale*, poiché essa, «in quanto fondata sulla fede morale pura», non è un'istituzione pubblica, ma ciascuno può comprendere solo da sé i progressi compiuti in nome in questa fede.

Storia ci potrà essere solo del rapporto fra le diverse e mutevoli forme della «fede di chiesa» con la «pura fede religiosa, unica e immutabile»; una storia che assumerà la forma della narrazione «della lotta costante fra la fede culturale e la fede religiosa morale»⁴⁶. L'unità di principio invocata da Kant viene da lui ascritta alla storia di quella Chiesa che, fin dai suoi inizi, ha portato con sé i germi e i principi per l'unità oggettiva della fede religiosa vera e universale, alle origini del cristianesimo, depurato però della componente della fede giudaica, in quanto insieme di leggi semplicemente statutarie, su cui si basava una costituzione statale. Al di là delle considerazioni critiche sull'attribuzione di tale ruolo al cristianesimo⁴⁷, emergono a questo punto alcune questioni di grande rilievo. Innanzi tutto, qual è l'operazione ermeneutica che fa da sfondo ad un tale procedimento, che ha consentito a Kant di trattare un tema così problematico in uno scritto che, a ben vedere, non solo non è estraneo al sistema critico, ma anzi lo completa, specie con l'aggiunta alla filosofia pratica della teoria del male radicale? E poi, da quali premesse speculative è partito Kant per sollevare il problema circa i limiti della religione entro la ragione umana, e cioè il problema del confine tra religione storica e religione razionale?

La soluzione a questi interrogativi si trova, a mio avviso, di nuovo nella *Kritik der Urteilskraft*, e più precisamente nella dottrina del giudizio riflettente, relativamente allo schematismo senza oggetto (§ 38); ma si trova anche

adagiandosi sulle parole di Luca, che: «Il regno di Dio non giunge in forma visibile. Non è possibile dire: esso è qui o là. *Perché, vedete, il regno di Dio è dentro di voi* [cors. mio] (*Luca, XVII, 21-22*)» (Kant (1793: 136, trad. it. 465)).

⁴⁶ *Ibidem.* p. 124-125, trad. it. 452-53.

⁴⁷ Per lo sviluppo di questo tema si potrebbe partire dalla considerazione secondo cui il privilegiamento del cristianesimo avverrebbe entro gli stessi limiti della ragione già individuati nelle opere della trilogia; una religione, quindi, non dedotta a partire dalla ragione, ma da quella religione che si concretizza proprio nella scelta di rappresentare il cristianesimo come una sorta di schematizzazione della religione razionale. Si veda a tale proposito la lunga nota relativa allo schematismo dell'analogia, che Kant formula in relazione alla realtà oggettiva dell'idea personificata del principio buono. (Kant (1793: 64-66, trad. it. 385-87)).

in un'Annotazione generale a *Die Religion*, laddove si parla di una «fede riflettente».

Nella nota posta a conclusione del primo capitolo Kant scrive:

Consapevole della sua impotenza nella soddisfazione delle sue esigenze morali, la ragione si estende fino a certe idee trascendenti che potrebbero compensare questa sua manchevolezza, senza tuttavia ricavarne un'estensione del proprio dominio. Essa non contesta né la possibilità né la realtà degli oggetti di queste idee, ma soltanto non può ammetterli nelle massime del proprio pensiero e del proprio agire. Essa ritiene anzi che, se nell'impenetrabile campo del soprannaturale c'è ancora qualcos'altro, oltre a ciò che essa è in grado di comprendere; che se esso è necessario per supplire alla propria impotenza morale, questo qualcosa, pur essendo ignoto, sarà di giovamento alla sua buona volontà; e ritiene ciò con una fede che potrebbe esser detta *r i f l e t t e n t e* (sulla possibilità della cosa) perché la fede *d o g m a t i c a*, che pretende valere come *s c i e n z a*, appare alla ragione insincera e presuntuosa⁴⁸.

Nella *Kritik der Urteilskraft* e in *Die Religion* manca una facoltà conoscitiva capace di estendersi ad un campo esterno alla ragione. In tal modo, sia l'attività postulativa della ragion pura sia quella della ragion pratica mostrano la loro insufficienza nel dar luogo alla conoscenza; esse sollecitano però un altro aspetto di particolare importanza: la necessità di schematizzare ciò che, seppur postulato moralmente, è teoreticamente trascendente.

Kant esplicita questa esigenza con l'utilizzo dello schematismo dell'analogia contenuto in *Die Religion*, Sezione Prima: Della giusta pretesa del principio buono al dominio sull'uomo, relativamente all'Idea personificata del principio buono e alla Realtà oggettiva di questa idea.

Kant, nel definire lo schematismo dell'analogia in relazione alla necessità di ricorrere sempre a certe analogie per rendere comprensibili istanze so-prasensibili, afferma:

La stessa Scrittura si adatta a questo tipo di rappresentazione perché possiamo comprendere l'alto grado che l'amore di Dio per il genere umano raggiunge quando si attribuisce a Dio il massimo sacrificio che un essere che ama possa fare per rendere felici anche creature indegne [...], anche se non ci è possibile, mediante la ragione, intendere in qual modo un essere, del tutto autosufficiente, possa sacrificare qualcosa che rientra nella sua felicità e spogliarsi di ciò che possiede. Questo è appunto lo *schematismo dell'analogia* [cors. mio] (che serve alla spiegazione) di cui non possiamo fare a meno⁴⁹.

⁴⁸ Kant (1793: 52, trad. it. 375).

⁴⁹ Kant (1793: 65, trad. it. 386).

E poco più oltre, avvertendo del rischio infausto di trasformare tale schematismo in uno «schematismo della determinazione oggettiva», che si tradurrebbe in un antropomorfismo assolutamente inaccettabile e dalle più funeste conseguenze dal punto di vista della religione, amplia e traduce a chiare lettere la natura e la funzione dello schema:

[...] voglio osservare che, nel risalire dal sensibile al soprasensibile, si può certo *schematizzare* [cors. mio] (rendere comprensibile un concetto attraverso l'analogia con qualcosa di sensibile) ma, mediante l'analogia, non si può assolutamente *c o n c l u d e r e* che una cosa propria dell'uno debba essere necessariamente attribuita anche all'altro (*e s t e n d e n d o* così il concetto); e ciò perché una conclusione del genere contrasterebbe con ogni analogia in quanto questa, dalla necessità in cui ci troviamo di introdurre uno schema per renderci comprensibile un concetto (facendolo poggiare su un esempio) pretenderebbe ricavare la conseguenza che tale schema spetti necessariamente anche all'oggetto in questione, come predicato di esso⁵⁰.

È questa, kantianamente, la deduzione del giudizio riflettente, la sussunzione del particolare che permette di giungere al concetto di esso. Non essendo un semplice giudizio di sensazione esso riveste, per così dire, la forma di un giudizio formale di riflessione che ha bisogno della deduzione con la quale si spieghi la sua pretesa alla necessità, «senza concludere» che sia da attribuire al soprasensibile, valutazione che spetta solo al giudizio di conoscenza. In questo caso si potrebbe soltanto affermare l'esigenza ineliminabile di cui 'non possiamo fare a meno' e che serva alla spiegazione dello schematismo dell'analogia che accompagna la rappresentazione dell'oggetto.

A ben vedere, anche nella *Kritik der Urteilskraft* i giudizi così costituiti sono 'sintetici', in quanto riflettenti, proprio perché «oltrepassano il concetto ed anche l'intuizione dell'oggetto ed aggiungono a questa, come predicato, qualcosa che non è conoscenza, cioè il sentimento di piacere (o dispiacere)»⁵¹. Ma, diversamente dal giudizio riflettente della terza *Critica*, che è singolare perché lega il suo predicato (il piacere) non con un concetto, ma con una rappresentazione empirica singolare, qui, in assenza dell'operazione dell'intelletto, la rappresentazione si lega a se stessa.

È il problema dell'unità di se stesso, vale a dire trovare una forma (o un concetto) la cui struttura sia tale che contenga come sua nota costitutiva il problema della propria possibilità.

⁵⁰ *Ibidem.*

⁵¹ Kant (1790: 288, trad. it. 115).

Il ricorso all'analogia da parte di Kant significa cogliere l'inferenza per analogia nel proprio significato specifico, in quanto – come si legge nella *Logica*, «estende le *proprietà date* di una cosa a più proprietà della *stessa medesima cosa. Uno in molti*, dunque, in tutti»⁵². Nell'inferire per analogia, pertanto, non è richiesta l'*identità del fondamento (per ratio)* in quanto ciò potrebbe valere solo per gli abitanti razionali della Luna, non per gli uomini. Né si può concludere per analogia, al di là del *tertium comparationis*.

Ogni inferenza della ragione deve dare necessità e così *induzione e analogia* non sono perciò inferenze della ragione, ma solo *presunzioni* logiche, ovvero inferenze empiriche; e attraverso l'induzione si ottengono sì proposizioni generali, ma non propriamente universali.

Segnalo, di seguito, alcuni passi della *Kritik der Urteilskraft*, per la straordinaria rilevanza gnoseologica che riveste il giudizio riflettente nel pensiero kantiano⁵³. Per ciò che attiene il rapporto etico-teologico, Kant in una lunga nota al paragrafo 87, puntualizza quanto segue:

Non l'uomo che vive s e c o n d o le leggi morali, cioè tale che agisce conformemente ad esse, è lo scopo finale della creazione. Perché, esprimendoci in tal modo, diremmo più di quello che sappiamo, cioè che stia in potere di un autore del mondo di far sì che l'uomo si conduca sempre conformemente alle leggi morali, il che presuppone un concetto della libertà e della natura (per la quale ultima non si può pensare che un autore esterno), concetto che dovrebbe implicare una cognizione del sostrato soprasensibile della natura e della sua identità con ciò che è reso possibile nel mondo della causalità libera; e ciò supera di molto la portata della nostra ragione. Solo dell'uomo s o t t o p o s t o alle leggi morali possiamo dire, senza oltrepassare i limiti della nostra conoscenza, che la sua esistenza costituisce lo scopo finale del mondo. Ciò si accorda perfettamente col giudizio della ragione umana, che riflette moralmente sull'andamento del mondo⁵⁴.

Ma c'è una limitazione relativa alla validità della prova morale (§ 88) ed è il ricorso opportuno all'inferenza per analogia (§ 90), quindi ad uno schematismo dell'analogia, il cui risultato Kant chiarisce in questo modo:

Il risultato di tutto ciò è, che per la ragione umana, relativamente all'esistenza di un essere originario in quanto Dio, e all'anima in quanto spirito immortale, non è

⁵² Kant (1800: 133, trad. it. 127).

⁵³ Cfr. il giudizio espresso da E. Weil sull'importanza della terza *Critica* per gli scritti posteriori alla trilogia: «il mondo esprime la ragione come è in sé. Occorre sottolinearlo, poiché niente dimostra più chiaramente ciò che di rivoluzionario si è prodotto nel pensiero kantiano con la *Critica del giudizio*, anteriore, e non per caso, alla *Religione nei limiti della semplice ragione*, alla *Metafisica dei costumi*, ai grandi saggi sulla politica e sulla storia» (Weil (1962: 174)).

⁵⁴ Kant (1790: 448-449, trad. it. 265, n. 1).

possibile assolutamente alcuna prova dal punto di vista teoretico, sia pure per produrre il minimo grado di adesione; e ciò per la ragione evidentissima, che non abbiamo alcuna materia per la determinazione delle idee del soprasensibile, e dovremmo trarla dalle cose del mondo sensibile, in modo che poi non riuscirebbe adeguata all'oggetto che si vuole applicare; e che, per conseguenza, tolta ogni determinazione, non resta altro che il concetto di qualcosa di non sensibile, che contiene il fondamento ultimo del mondo sensibile, ma non costituisce alcuna conoscenza (in quanto estensione del concetto) della sua interna costituzione⁵⁵.

Ma affinché si abbia ben presente la natura fondativa di tale discorso, Kant avverte la necessità di ammettere 'qualche cosa' che contenga il fondamento della possibilità e della realtà pratica, cioè dell'attuabilità di uno scopo finale moralmente necessario. In tal senso, dichiara:

Ma noi possiamo pensare questo qualche cosa, secondo la natura dell'effetto che ne aspettiamo, come un essere saggio e che governa il mondo secondo leggi morali, e, conformemente alla costituzione delle nostre facoltà conoscitive, dobbiamo pensarlo come una causa delle cose distinta dalla natura, soltanto per esprimere il rapporto di questo essere, che trascende tutte le nostre facoltà di conoscere, all'oggetto della nostra ragion pratica: senza pretendere però di attribuirgli teoreticamente l'unica causalità di questa specie che conosciamo, cioè un'intelligenza ed una volontà, e nemmeno di distinguere oggettivamente, in questo essere stesso, la causalità che pensiamo in lui relativamente a ciò che è scopo finale per noi, da quella che è relativa alla natura (e alla sua finalità in generale), mentre possiamo ammettere tale distinzione solo come soggettivamente necessaria, per la costituzione della nostra facoltà di conoscere, e come valida pel Giudizio riflettente, non pel Giudizio determinante⁵⁶.

La *Kritik der Urteilskraft* contiene, dunque, lo sviluppo della prova morale, che è ciò che è sempre esistito nella ragione umana, anteriormente al suo primo sviluppo, «e si è sviluppata sempre più con la progressiva coltura della ragione medesima»⁵⁷. Ecco perché tale motivo ritorna costitutivamente in *Die Religion*.

La pretesa di necessità che sul piano metasensibile porta allo schematicismo dell'analogia, risiede, dunque, nella «fede riflettente».

Questa sola è capace di supplire all'impotenza della morale ed essere di aiuto sia al fine umano da raggiungere (per ciò che attiene la buona volontà degli essere umani come specie e come singolo) sia a Kant, per ciò che era la

⁵⁵ Kant (1790: 466, trad. it. 281).

⁵⁶ Kant (1790: 457, trad. it. 273).

⁵⁷ Kant (1790: 458, trad. it. 273).

sua intenzione: restituire la dogmatica teologica al piano storico-critico, connettendola in un certo qual modo al piano dell'etica e a quello della *Publizität*, tanto caro alla Arendt e da lei stessa considerato un principio trascendentale⁵⁸.

Quando Kant, oltretutto, parla di fede, non intende una mera adesione a un credo, bensì una fede che pensa, quindi la ragion pratica, portatrice dell'unica fede vera. Prendiamo in considerazione, per concludere, alcuni passi kantiani, indicativi dei risultati da lui raggiunti nello scritto sulla religione:

Che la legge morale chiami l'uomo a una buona condotta di vita, che egli scopra dentro di sé, per il rispetto incancellabile che sente per essa, la promessa che si può nutrir fiducia in questo spirito buono e si può sperare di poterlo in qualche modo soddisfare, che, infine, mediante la conciliazione di questa speranza coi comandi rigorosi della legge, egli debba costantemente prendersi in esame come si richiede a chi dovrà presentarsi a un giudice, tutto ciò ci insegnano sia il cuore sia la coscienza, spingendoci ad agire in questo senso. Sarebbe presunzione pretendere che ci fosse rivelato qualcosa di più; se questo dovesse accadere, non potrebbe esser considerato come un bisogno universale dell'uomo⁵⁹.

E poco dopo, parlando dell'opera dello Spirito Santo:

Ma benché quel grande mistero, che raccoglie in un'unica formula tutti quelli finora nominati, possa riuscire comprensibile alla ragione di ogni uomo quale *idea religiosa praticamente necessaria* [cors. mio], si può affermare che esso, per divenire il fondamento morale di una religione, specialmente se pubblica, fu rivelato per la prima volta quando fu insegnato p u b b l i c a m e n t e, elevandolo a simbolo di un'epoca religiosa del tutto nuova⁶⁰.

Oltre ad un'intenzione critico-dottrinale, Kant, con questo suo scritto, assolve anche il compito di chiarire il rapporto che si istituisce, dopo lo scoppio della Rivoluzione francese, tra riflessione politica e riflessione religiosa. La religione, infatti, unifica i principi dell'organizzazione del popolo etico di Dio con i principi della sua costituzione, dove i primi hanno un valore empirico perché costituiscono solo il mezzo per l'esecuzione dei secondi. In questo senso ha ragione Losurdo quando scrive:

[...] la religione si risolve sì nella morale, ma nella morale non dell'uomo privato, bensì del *citoyen*: è vero che «la religione (considerata soggettivamente) è la conoscenza di tutti i nostri doveri come comandamenti divini», ma i doveri di cui qui

⁵⁸ Arendt (1982: 75).

⁵⁹ Kant (1793: 144-145, trad. it. 475).

⁶⁰ *Ibidem*.

si tratta sono quelli «etico-civili» (*ethisch-bürgerlich*), quelli «degli uomini verso gli uomini»; al di fuori di questi, non ci sono «doveri speciali, riferentisi immediatamente a Dio». La religione si è risolta fino a tal punto nella consacrazione della comunità politica e degli obblighi (e dei diritti) del *citoyen*, che essa non richiede «alcun sapere assertorio» (nemmeno quello dell'esistenza di Dio)⁶¹.

La non assertorietà dell'esistenza di Dio è possibile anche per la presenza nel giudizio riflettente di leggi che non sono costitutive della natura, ma sono leggi della facoltà stessa del giudizio.

6. Se il giudizio politico puro sia facoltà legislativa a priori

Con l'ideale e futuristica previsione del progresso verso il meglio Kant fa valere la necessità di mantenere l'unità nella molteplicità senza annichilire il particolare nell'universale, mantenendo le differenze. Essa viene considerata, quindi, non nel concetto della specie (*singulorum*) ma nel concetto della totalità degli uomini, affinché tale *ideale* venga fatto valere nella continuità immanente alla sua stessa formulazione come un fattore efficace sul piano politico.

Da questo punto di vista Kant non poteva fondare il presupposto epistemologico della progressione verso il meglio se non nell'«esposizione a priori possibile dei fatti che devono succedere»⁶². Una storia a priori sarà possibile se «chi predice fa e organizza ciò che egli stesso ha predetto», la condizione consisterà tutta nell'ordine non solo meccanico, ma soprattutto morale: così solo infondendo principi morali che portino al meglio si potranno evitare deviazioni nello sviluppo lineare della storia dell'umanità.

Nella capacità dell'uomo di operare scelte ragionevoli risiede la possibilità del suo costruirsi storico, poiché per prevedere «l'uomo ha bisogno del nesso secondo le leggi naturali, mentre in ordine alle *libere* azioni future deve fare a meno di questa guida o indicazione»⁶³.

All'idea regolativa, meccanicamente necessitante, dell'incidenza della Provvidenza sulle umane vicende, Kant contrappone con forza l'intervento attivo dell'uomo in continuo progresso verso il meglio, in relazione alla sua destinazione morale, che si pone come un dovere praxeologico.

L'idea di un ordinamento cosmopolitico che l'uomo può e deve realizzare, in quanto è in suo dovere e in suo potere farlo, è l'ipotesi che Kant vede come in grado di dare corpo all'idea di una federazione di popoli che tragga

⁶¹ Losurdo (2007: 150).

⁶² Kant (1798: 79, trad. it. 213).

⁶³ *Ibidem.* p. 84; trad. it. 217.

forza dalla collettività che delibera secondo leggi della volontà comune. È una concezione non necessitante, ma solo possibile, che è di sprone agli uomini che entrano nella società civile e che possono avere la tendenza alla depoliticizzazione, vale a dire allo scarso interesse personale. Mi riferisco ancora una volta alle osservazioni presenti nella *Kritik der Urteilskraft* in cui si sottolinea il significato che assume un 'essere organizzato' in natura, il quale non può essere «semplice macchina» dotata di un'unica forza motrice, ma presuppone uno scambievole rapporto tra le parti affinché nessuna di essa venga relegata a un mero ruolo passivo⁶⁴.

In analogia coi fini della natura – avverte Kant nella nota al paragrafo 65 – si può parlare «dell'impresa di una totale trasformazione di un grande popolo in uno Stato»: anche qui abbiamo un essere organizzato (non a caso si parla di «o r g a n i z z a z i o n e» per designare l'assetto delle magistrature ecc.) che non può spiegarsi con un cieco meccanismo della natura, perché «in un tutto come questo ogni membro dev'essere non soltanto mezzo, ma anche scopo; e, mentre concorre alla possibilità del tutto, è determinato a sua volta dall'idea del tutto, relativamente al suo posto e alla sua funzione»⁶⁵.

La possibilità di prevedere la storia futura nel suo svolgimento verso il meglio può risiedere, dunque, in una 'legislazione pura a priori del giudizio politico', sia per la sua natura intrinsecamente determinante (porta con sé il dovere da parte dell'uomo di agire nella serie indefinita delle generazioni) sia per quella meramente riflettente (porta con sé la dimensione della socialità che converge su un altro principio garante del diritto che è la pubblicità). E come si esibisce?

Attraverso la concatenazione delle cause e degli effetti della natura che si manifestano storicamente: esempi sono le rivoluzioni che devono essere interpretate come evoluzioni della storia verso il suo fine.

Il giudizio riflettente porta sempre con sé la capacità della rappresentazione dovuta al libero e scambievole gioco armonico di intelletto e immaginazione, stimolando analogamente l'agevolazione e l'intensificazione della vita di ognuno, per «il fiorire della propria identità e l'esaltazione della propria autenticità»⁶⁶, grazie ad una sua nota distintiva data dalla capacità intrinsecamente da esso posseduta: l'orientamento. L'attenzione kantiana rivolta allo stimolo a progredire viene ulteriormente caricato di significato attraverso l'attribuzione al giudizio riflettente della capacità di legiferare nelle azioni

⁶⁴ Kant (1790: 347, trad. it. 196).

⁶⁵ Kant (1790: 375, trad. it. 197, n. 1).

⁶⁶ Sull'aspetto qui ripreso, si veda Makkreel (1990), per aver individuato il carattere peculiare del giudizio operando la distinzione tra giudizi riflettenti puri e giudizi riflettenti orientati, all'interno del saggio kantiano del 1786 *Was heißt: sich im Denken orientieren?*

umane nell'ordine, anche e soprattutto, temporale, il che lo carica del valore politico dell'universale comunicabilità nella società (*sensus communis*). La sensazione mentale di un'esaltazione della vita insita nel piacere estetico e quella di un suo essere sminuita nella sgradevolezza estetica comportano senz'altro un'anomalia razionale.

Eppure il discernimento estetico si rapporta a giudizi riflettenti intorno al mondo, né più né meno come la distinzione spaziale fra destra e sinistra si rapporta a giudizi determinanti intorno alla natura. Entrambi sono sentimenti soggettivi ma costanti, che ci forniscono il necessario orientamento nel momento in cui la nostra immaginazione si muove da ciò che è dato in modo diretto a ciò che è dato in maniera solo indiretta.

È l'identità come autorealizzazione di sé, vale a dire relazione autentica del Sé con se stesso, o meglio, traducendo tutto questo in termini kantiani: è la difesa della tendenza a fare di sé il motivo determinante della volontà, quindi, rispetto e cura dell'*amor razionale di sé* (*vernünftige Selbstliebe*)⁶⁷.

Bibliografia

- Arendt, H. (1982), *Lectures on Kant's Political Philosophy* (*Teoria del Giudizio politico*, trad. it. di P.P. Portinaro e Ronald Beiner (a cura di), Genova, Il Melangolo, 1990).
- Arendt, H. (1958), *The Human Condition* (*Vita activa. La condizione umana*, trad. it. di S. Finzi, Milano, Bompiani, 1991).
- Arendt, H. (1978), *The Life of the Mind*, New York-London, Harcourt Brace Jovanovich (*La vita della mente*, introduzione all'ed. it. di A. Dal Lago, pref. all'ed. americana di M. McCarthy, Bologna, il Mulino, 1987).
- Beiner, R. (1983), *Political Judgment*, London, Methuen.
- Buck, G. (1979), *Über die Identifizierung von Beispielen – Bemerkungen zur 'Theorie der Praxis'*, in AA.VV., *Identität*, hrsg. von O. Marquard und K.H. Stierle, München, Fink.
- Cassirer, E. (1918), *Kants Leben und Lehre* (*Vita e dottrina di Kant*, trad. it. di G. A. De Toni, Firenze, La Nuova Italia, 1984).

⁶⁷ Kant (1788: 73, trad. it., 1970: 214).

- Eraclito, Framm. 108, trad. it a cura di G. Giannantoni, in *I presocratici*, 2 voll., Bari, Laterza, 1969.
- Ernstein, R. (1983), *Beyond Objectivism and Relativism*, Philadelphia, University of Pennsylvania Press.
- Flynn, B. (1988), "Arendt's appropriation of Kant's Theory of Judgment", *Journal Brit. Soc. Phenomenol*, pp. 128-140.
- Galston, A. W. (1975), *Kant and the problem of History*, Chicago, University of Chicago Press.
- Guyer, P (1979), *Kant and the Claims of Teste*, Cambridge (Mass), Cambridge University Press.
- Habermas, J. (1977), "Hannah Arendt's Communications Concept of Power", *Social Research*, 44, 1, pp. 3-24.
- Habermas, J. (1980), On the German-Jewish Heritage, *Telos*, 44, pp. 127-131.
- Held, K. (1981), Stato, interessi e mondi vitali. Per una fenomenologia della politica, di A. Ponsetto (a cura di), Brescia, Morcelliana.
- Henry, B. (1992), Il problema del giudizio politico fra criticismo ed ermeneutica, Napoli, Morano.
- Höffe, O. (1983), *Immanuel Kant (Immanuel Kant*, trad. it. di S. Carboncini e V. Verra (a cura di), Bologna, Il Mulino, 1986).
- Hutchings, K. (1996), *Kant, Critique and Politics*, New York, Routledge.
- Kain, J. P.(1989), "Kant's Political Theory and Philosophy of History", in *Clio*, 18, (4), pp. 325-345.
- Kant, I. (1795), *Zum ewigen Frieden. Ein philosophischer Entwurf*, in *Gesammelte Schriften*, Herausgegeben von der Königlich Preußischen Akademie der Wissenschaften (und Nachfolgern), Berlin-Leipzig, 1900 sgg., (da ora in poi KGS), vol. VIII (*Per la pace perpetua*, trad. it. di G. Solari, in *Scritti politici e di filosofia della storia e del diritto*, con un saggio di Christian Garve, a cura di Norberto Bobbio, Luigi Firpo e Vittorio Mathieu, Torino, UTET, 1965²).

- Kant, I. (1781), *Kritik der reinen Vernunft*, in KGS, IV (*Critica della ragion pura*, trad. it. di G. Gentile e G. L. Radice, con introduzione di Vittorio Mathieu (a cura di), Roma-Bari, Laterza, 1993).
- Kant, I. (1784), Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht, in KGS, VIII (Idea per una storia universale dal punto di vista cosmopolitico, in *Scritti politici...*, cit.
- Kant, I. (1790), *Kritik der Urteilskraft* (1790), in KGS, V (*Critica del Giudizio*, trad. it. di A. Gargiulo, con glossario e indice dei nomi a cura di V. Verra, Roma-Bari, Laterza, 1994).
- Kant, I. (1793), Die Religion innerhalb der Grenzen der bloßen Vernunft, in KGS, VI, 10 (La religione entro i limiti della sola ragione, trad. it. di P. Chiodi, in *Scritti morali*, Torino, UTET, 1970 (ristampa del 1986)).
- Kant, I. (1798), Der Streit der Fakultäten in drei Abschnitten, in KGS, vol. VII, Il conflitto delle facoltà, in *Scritti politici...*, cit.
- Kant, I. (1799), *Kritik der praktischen Vernunft*, in KGS, V (*Critica della ragion pratica*, di P. Chiodi (a cura di), in *Scritti Morali*, Torino, UTET, 1970).
- Kant, I. (1800), *Logik*, di G. B. Jäsche (a cura di), in KGS, IX (*Logica*, di L. Amoroso (a cura di), Roma-Bari, Laterza, 1984).
- Losurdo, D. (2007), Autocensura e compromesso nel pensiero politico di Kant, Napoli, Bibliopolis.
- Macera, E. (2001), “Kant e la «critica della ragion politica»”, in *La filosofia politica di Kant*, di G.M. Chiodi (a cura di), G. Marini, R. Gatti, Milano, Franco Angeli.
- Makkreel, R. (1990), Imagination and Interpretation in Kant. The Hermeneutical Import of the «Critique of Judgment», Chicago, University of Chicago Press.
- Mancini, I. (1986), *Filosofia della religione*, Genova, Marietti.

- O'Neill, O. (1992), *Reason and Politics in the Kantian Enterprise*, in *Essays on Kant's Political Philosophy*, di H. Williams (ed.), Cardiff, University of Wales Press, pp.50-80.
- Olivetti, M. M. (1973), "Chiesa e ideologia. Un problema di filosofia della storia, in Demitizzazione e ideologia", in *Archivio di filosofia*.
- Passerin d'Entrèves, R. (1987), "Il concetto di giudizio politico nella filosofia di Hannah Arendt", in *La pluralità irrepresentabile*, di R. Esposito (a cura di), Urbino, Quattro venti, pp. 155-169.
- Pasolini, P. P. (2013), *Scritti corsari*, Milano, Garzanti.
- Pleines, J. E., (1983), *Praxis und Vernunft. Zum Begriff praktischer Urteilskraft*, Würzburg-Amsterdam, Königshausen & Neumann-Rodopi.
- Riley, P. (1992), "Hannah Arendt on Kant, Truth and Politics", in *Essays on Kant's Political Philosophy*, di H. Williams (ed.), Cardiff, University of Wales Press
- Scaravelli, L. (1955), "Osservazioni sulla «Critica del giudizio»", in *Opere*, II, *Scritti kantiani*, Firenze, La Nuova Italia, 1968.
- Shklar, J. N. (1977), "Rethinking the Past", *Social Research*, 44, n. 1, pp. 80-90.
- Tenenbaum, K. (1974), "Il pensiero politico di Immanuel Kant e l'influsso di Jean-Jacques Rousseau", *Giornale critico della filosofia italiana*, 53, pp. 343-392.
- Vlachos, G. (1962), *La pensée politique de Kant. Métaphysique de l'ordre et dialectique du progrès*, Paris, Presses Universitaires de France.
- Vollrath, E. (1977), *Zur Rekonstruktion der politischen Urteilskraft*, Stuttgart, Klett-Cotta.
- Vollrath, E. (1987), *Grundlegung einer philosophischen Theorie des Politischen*, Würzburg-Amsterdam, Königshausen & Neumann.

- Volpi, F. (1987), “Il pensiero politico di Hannah Arendt e la riabilitazione della filosofia pratica”, in *La pluralità irrepresentabile, Il pensiero politico di Hannah Arendt*, di R. Esposito (a cura di) Urbino, Quattroventi, pp. 73-92.
- Weil, E. (1962), “Kant et le problème de la politique”, in *La Philosophie politique de Kant (Annales de Philosophie politique, IV)*, Paris, Puf, pp. 1-32; poi col titolo “Histoire et politique” in Id., *Problèmes kantians*, Paris, J. Vrin, 1963 (“Storia e politica”, trad. it. di P. Salvucci. in Id., *Problemi kantiani*, Urbino, Quattroventi, 1980).

Mythos e logos.

Alle origini del pensiero occidentale

Giorgio Grimaldi
Università degli Studi “G. d’Annunzio” di Chieti-Pescara
giorgio.grimaldi@yahoo.it

1. *Mythos e logos* come semplice opposizione o come passaggio liquidatorio in senso assoluto?

Mythos e logos rimandano immediatamente, nella rappresentazione comune, a due tipi di atteggiamento e di approccio riguardo il loro significato e la loro relazione: o i due termini appaiono in reciproca opposizione, insuperabile perché costitutiva, oppure essi vengono concepiti l’uno come un diverso, e superiore, grado dell’altro, quasi che a un tratto la storia abbia compiuto una sorta di definitivo salto qualitativo dal primo al secondo e sia ormai approdata dalle acque turbinate e agitate, eppure estremamente affascinose, del *mythos* alla saldezza di un *logos* come *terra firma*, *luogo* che finalmente concede uno sguardo (e un pensiero) quieto, sereno e che può osservare le cose in un ambito in cui vige la *chiarezza*. Fuor di metafora, il passaggio dal *mythos* al *logos* costituirebbe il *momento* in cui il pensiero non si muove più attraverso immagini, figure, rappresentazioni che cercano di offrire risposte secondo forme e modalità ancora confuse, ma riesce invece a elevarsi man mano al livello di razionalità compiuta attraverso modelli argomentativi sempre più complessi, nonché precisi, rispetto alla conoscenza dell’oggetto particolare, e all’elaborazione di concetti che presuppongono una capacità piuttosto elevata di astrazione. Questa rende possibile il superamento di spiegazioni che scambiano le immagini attraverso cui si muovono per la verità ultima dei contenuti da esprimere: così il *logos* smantella sistematicamente concezioni del mondo che implicano l’accettazione di presupposti irrazionali (ma è proprio il *logos* a mostrarli come tali), e in questo senso uno dei suoi risultati più recenti e decisivi è lo sviluppo delle scienze sperimentali.

Senonché, mentre l'approccio che vede *mythos* e *logos* come semplicemente in opposizione mostra la sua fragilità piuttosto presto, perché ignora la derivazione del secondo dal primo, il secondo approccio appare sicuramente più fruttuoso e più rispondente allo sviluppo storico concreto, ma apre una serie di questioni che non riesce a risolvere, proprio perché tende a ignorare l'*intreccio* di entrambi come *intreccio nel tempo*.

La "svista" fondamentale di questa impostazione è il considerare il *passaggio* dal *mythos* al *logos* come definitivo; si tratta sì di un *momento*, e di un *momento cronologico*, ma non solo: esso è un *momento* dello svolgersi stesso del pensiero e perciò non relegabile a un determinato periodo del mondo antico e, per essere più precisi, del pensiero greco. Se esso non è un momento solo cronologico, non è fissato in un punto e definitivo, ma è *ripetibile, reiterabile* e forse *costitutivo* del pensiero stesso.

È questo *movimento* in sé unidirezionale? Si dà quindi solo un movimento *progressivo dal mythos al logos* oppure è possibile un movimento in senso contrario, che va dal *logos* al *mythos*? E questo secondo moto è da considerarsi come sempre, e meramente, regressivo?

Per cercare di dare una risposta a queste domande dobbiamo prima interrogare le parole stesse.

2. Interrogare le parole stesse

Ma cosa significa *interrogare le parole stesse*? Ricostruirne con precisione filologica la storia del loro significato? Cercare di mirare alla comprensione del loro concetto, di ciò che, al di là del loro significato in senso stretto (e comune), esse dicono?

Nelle lezioni su Parmenide del semestre invernale del 1942/43, Heidegger espone una serie di considerazioni in merito che sono estremamente indicative:

La conoscenza dei vocaboli non garantisce ancora la comprensione delle parole. Esse dicono ciò che propriamente c'è da dire: la parola, il detto. Se quindi seguiamo la letteralità in modo tale che già da prima, e perciò costantemente, pensiamo alla parola e in base a essa, allora, ma soltanto allora, la nostra grande stima per la 'letteralità' appare giustificata. Il termine preso alla lettera dobbiamo udirlo in modo da prestare ascolto alle indicazioni che ci indirizzano verso la parola. In questo stare in ascolto il nostro percepire ubbidisce quindi a ciò che la parola dice. Esso esercita così l'attenzione. Diventa un pensare¹.

¹ Heidegger (1982b, trad. it. 52).

Heidegger si esprime attraverso un linguaggio adeguato alla sua concezione dell'essere come *evento*, il quale si dà anche nella modalità di un *dire* di cui l'uomo, e solo l'uomo, può porsi in ascolto. Tale ascolto può accogliere anche il silenzio, che non è semplice assenza di contenuto, ma esso stesso *segno* di un essere che, come *evento*, si dà nella modalità del darsi/sottrarsi, velamento/disvelamento. Di qui l'importanza della modalità dell'*ascolto*. Ma al di là del significato dell'intero passaggio all'interno dell'impianto heideggeriano, i due punti che ci interessano maggiormente sono posti all'inizio e alla fine del passo stesso.

Dapprima vi è una constatazione, un punto di partenza problematico: «La conoscenza dei vocaboli non garantisce ancora la comprensione delle parole». Alla fine vi è il passaggio per cui lo «stare in ascolto» diviene «un pensare». Questo *pensare* mira «a ciò che la parola dice», ma non (solo) in base al risultato che può comportare «la conoscenza dei vocaboli», in quanto occorre andare oltre, più in profondità, verso «la comprensione delle parole». Questo *in-oltrarsi* lo può il pensiero.

Heidegger, però, pensa la verità delle parole come custodita nel loro venire a espressione originario: questa verità, quasi subito, si occulta, si vela, si sottrae, riluce nel pensiero greco presocratico per poi ritrarsi. Ascendenza nietzscheana, il pensiero greco delle origini è per Heidegger un *inizio* che, non a opera dell'uomo ma dell'essere stesso, entra presto nella modalità del velamento, un velamento che, proprio in quanto tale, lo *salva*. Tale *inizio*, non come semplice *ripetizione*, può portare a compimento ciò che in quel primo inizio era stato pensato per poi cadere nell'oblio: il pensiero dell'essere, e non dell'ente come se fosse l'essere, un pensiero che non entizzi l'essere, con il conseguente superamento (o meglio, oltrepassamento) del pensiero rappresentativo-metafisico.

Perciò la verità delle parole sta nel loro venire a espressione originario, è custodita nella loro forma originaria, iniziale, così come il pensiero presocratico le ha pensate. Si scatenano, allora, in Heidegger, un vero e proprio *pathos* dell'originario, una ricerca e un'analisi minuziose volte a raggiungere il più possibile il sostrato originario delle parole, al di là dei significati comuni e sedimentati, come se in quell'originario risiedesse la verità. Poter attingere nuovamente a quella verità potrebbe essere il segno della fine di una (rovinosa, in fondo, per Heidegger) parabola costituita dal pensiero metafisico occidentale, da Socrate e Platone in poi.

Ma, possiamo chiederci, la verità risiede veramente nell'*origine*? L'essere si è veramente velato a partire dall'impostazione della metafisica occidentale che ha entizzato l'essere? Oppure tale entizzazione è inevitabile, in quanto pensare significa e comporta sempre pensare *qualcosa*? L'essere,

però, non è *qualcosa*, e Heidegger, magistralmente, ne parla come *l'aperto*: «L'aperto [*Das Offene*], in cui ogni ente è liberato come nel suo libero, *questo aperto è l'essere stesso* [*das Offene ist das Sein selbst*]. Ogni svelato, in quanto tale, è salvato nell'aperto dell'essere, cioè nel senza-fondo [*im Boden-Losen*]»². Cercare di rispondere a questi interrogativi ci porterebbe del tutto, qui, fuori strada³; invece l'indicazione heideggeriana ci orienta verso una «comprensione delle parole» che è un *pensarle*.

3. Essere e verità

La lezione di Heidegger riguardo il *mythos* è presente nel corso su Parmenide: si tratta di una parola-chiave per cercare di comprendere la portata dei frammenti parmenidei. Com'è noto, Parmenide espone il suo pensiero come rivelatogli da una dea. Ella non ha *un* nome, ma diversi, segno di una molteplicità che risulta essere una sua modalità caratterizzante. Luigi Ruggiu ha analizzato questa modalità di manifestazione della dea:

In Parmenide, l'unica Grande Dea si rifrange nei suoi molteplici aspetti, ciascuno dei quali si mostra necessariamente connesso con l'altro e tutti insieme costituiscono la forma totale del divino. Senza la polinomia, la potenza del divino rimarrebbe in qualche modo muta e inespressa, indistinta. Ogni nome non solo esprime un momento o aspetto essenziale del divino, ma anche necessariamente rinvia all'altro, evidenziando un momento dell'operare della Dea e insieme un aspetto del reale⁴.

Tra i modi in cui la dea viene nominata incontriamo alcuni termini essenziali del pensiero occidentale: giustizia, sorte, legge divina, verità, necessità. È pur vero, come sottolinea Giovanni Reale riportando un rilievo di Mario Untersteiner, che vi è un «nesso culturale ben preciso fra la Dea e la figura mitica della Grande Madre propria della religiosità mediterranea preellenica, cui Parmenide deve essersi ispirato»⁵, eppure tale elemento religioso, mitico, diviene in Parmenide la chiave per il passaggio al *logos*. Ma andiamo con ordine e torniamo a Heidegger per seguirlo nella sua interpretazione del *mythos*.

² *Ibidem*. p. 268; Id. (1982a: 224).

³ Per l'analisi di questi problemi relativi all'impostazione heideggeriana, vedi Grimaldi (2012: 185-203).

⁴ In Reale (2001: 20).

⁵ *Ibidem*. p. 20.

Essa, nel corso del 1942/43, si sviluppa a partire dalla lettura dei frammenti parmenidei. In quella modalità del velarsi/disvelarsi, a cui abbiamo accennato prima, dell'essere stesso come *evento*, la parola *verità* assume una posizione eminente. Heidegger mostra come nel passaggio al latino *veritas* il carattere greco dell'*alétheia* sia andato perduto, proprio a causa di un velarsi dell'essere stesso. Egli mostra come l'alfa privativo indichi che il termine greco per *verità* parli invece di una vera e propria *dis-velatezza* anziché rimandare a una corrispondenza con la realtà, con il *verum*, il *certum*, di un soggetto che si autoassicura così un rapporto di controllo, di dominio, nei confronti degli oggetti e del mondo circostante. Invece, pensata in modo greco (ecco il pensare le parole senza "accontentarsi" di averne appurato il mero significato, di averle "tradotte"), la parola *alétheia* ci indica che quella che noi chiamiamo *verità* dice della modalità del disvelamento e quindi rinvia a un velamento. Il discorso diviene più chiaro se noi cerchiamo di specificare cosa l'essere sia, un'impresa che il linguaggio tenta ma che forse non può (ancora) interamente esprimere: «essere significa: presenza di ciò che è presente»⁶. Prezioso e illuminante è in questo senso un passo heideggeriano del *Parmenide*, un'indicazione che è anche una sorta di "metodo":

Per pensare l'essere non occorrono né la solenne pomposità che fa uso di una complicata erudizione, né gli stati singolari ed eccezionali quali gli sprofondamenti mistici e le estasi in un'assorta pensosità. C'è bisogno, invece, solo del semplice risvegliarsi nella vicinanza di un qualsiasi ente inappariscnte, un risveglio che vede improvvisamente che l'ente 'è' [das Seiende 'ist']⁷.

E, in uno scritto sul *logos* del 1951, Heidegger chiarisce ulteriormente: «Prendiamo in considerazione anzitutto che 'essere' significa originariamente 'esser presente'; e che 'essere presente' vuol dire pro-dursi e durare (*her-vor-währen*) nella disvelatezza»⁸. Quest'ultima parola ci mostra quanto l'*alétheia* sia centrale nella riflessione heideggeriana. È proprio per questo venire-alla-presenza nella disvelatezza che l'*aperto* è il termine con cui Heidegger vuole indicare, vuole cercare di pensare l'ambito a partire dal quale si dà il venire-alla-presenza: «l'aperto, in cui ogni ente è liberato come nel suo libero, *questo aperto è l'essere stesso*». Per pensare l'essere non dobbiamo concentrare l'attenzione sull'*ente* che è, ma sull'*è* dell'ente (ecco perché abbiamo visto Heidegger porre l'*è* tra virgolette).

⁶ Heidegger (1954, trad. it. 172).

⁷ Heidegger (1982b: 266).

⁸ Id. (1951, trad. it. 157).

È a partire da questo complesso di indicazioni sull'essere e sull'*alétheia*, tenendo presente il loro rapporto che è instaurato dalla modalità di disvelamento, e che rinvia a un velamento, che è possibile accingersi alla lettura dell'interpretazione heideggeriana del *mythos*.

4. *Mythos*

Mythos originariamente è *ciò che viene detto oralmente*, ed è a partire dal significato che, al di là dei vari significati che si sono aggiunti nel tempo, lo lega al *dire*, che Heidegger lo inserisce all'interno del corso su Parmenide:

I termini greci κρύπτειν e κρύπτεσθαι (da cui derivano il latino *cripta* e il tedesco *Gruft*, *cripta*) significano il velare che cela. Κρύπτειν viene detto soprattutto della νύξ, della notte. Ma giorno e notte, in generale, indicano per l'appunto l'evento del disvelamento e del velamento. Ora, è proprio perché per i Greci tutto l'essenziale sorge fondamentalmente dall'essenza del velamento e della svelatezza che essi, quando esprimono l'inizialità (*Anfängnis*) del tutto, parlano per l'appunto di νύξ e ουρανός, di notte e giorno luminoso. Ciò che è in tal modo detto è il 'da dire' (*das zu Sagende*) in termini iniziali. È la saga (*die Sage*) in senso proprio, la parola iniziale. Il termine greco che sta per 'parola', e in cui il 'da dire' anzitutto si dice, è μῦθος. Anche l'essenza del μῦθος si determina in base all'ἀλήθεια. Μῦθος è ciò che schiude, disvela e fa vedere: vale a dire ciò che dal principio si mostra ovunque come ciò che è presente in tutto quanto è 'presente'⁹.

Il *mythos* è proprio dei Greci perché essi pensano l'essere come *evento* e in quanto esso *dice* dello svelarsi e del velarsi dell'essere stesso: in quest'ambito fanno cenno gli dèi. Ed è questo l'ambito nel quale la dea parla a Parmenide e gli rivela che «è necessario il dire e il pensare che l'essere sia: infatti l'essere è, il nulla non è»¹⁰. Da qui parte l'ontologia occidentale e da qui il pensiero filosofico greco, e poi occidentale, prende le mosse non solo per pensare l'essere ma, contrariamente a quanto richiesto da Parmenide, per pensare il nulla, o meglio, il ni-ente, come cercherà di fare lo stesso Heidegger. Tra le innumerevoli eredità della riflessione parmenidea tramandateci dai frammenti vi è anche il nucleo importante di caratteristiche che, nel corso dei secoli, e con declinazioni diverse, giungeranno dall'essere astratto al Dio della teologia cristiana. In altri termini, i frammenti di Parmenide inaugurano per molti versi il pensiero occidentale, almeno per quanto riguarda alcuni suoi caratteri fondamentali.

⁹ Id. (1982b: 124).

¹⁰ Parmenide, *Sulla natura*, p. 49.

Heidegger, non senza toni neopaganeggianti, inserisce il *mythos* in una interpretazione molto articolata del pensiero greco presocratico e del suo rapporto con l'essere. Ma, a parte i toni, il *mythos* qui è restituito a una sua dignità originaria e sottratto all'essere considerato mero "racconto" dal sapore ormai quasi favolistico. Ciò che attraverso il *mythos* si compie è il *dire* in senso eminente, è il *dire* degli dèi che appaiono nella modalità del disvelamento:

Il da-dire e il detto della saga sono il divino in quanto apparire e in quanto percepito nel suo apparire; perciò il divino è 'mitico'; perciò la saga degli dèi è 'mito'; perciò, infine, solamente l'uomo esperito in modo greco, e solamente esso, nella sua essenza e conformemente all'essenza dell'ἀλήθεια, è 'colui che dice Dio' (*Gottsager*)¹¹.

Questo vero e proprio *pathos* della greicità originaria conduce Heidegger a constatazioni piuttosto impegnative: «Se la divinità iniziale sorge dall'essenza dell'essere, la dimenticanza dell'essere non dovrebbe forse costituire la ragione del fatto che, da quando l'inizio della verità dell'essere si è ritratto nella velatezza, non è più potuto apparire alcun Dio sorgente dall'essere stesso?»¹² Questa sorta di "fuga degli dèi", questo loro ritrarsi, ne spiega il dileguare e ne può annunciare il ritorno?

Il fatto è che il *pathos* della greicità originaria, così declinato nella fase successiva a *Essere e tempo*, se ci offre una interpretazione del *mythos* non appiattita su atteggiamenti che lo sviliscono, lo sottovalutano e non ne comprendono il senso e la portata, diviene però più che potenzialmente regressivo. Se altissima è la vetta teoretica di concepire l'essere come *l'aperto*, nella tensione a pensarlo senza entizzarlo e a indicarne l'ambito, il riferimento puntuale agli dèi e al loro apparire nella disposizione dell'essere come *evento*, pur fascinoso e costituente una magistrale interpretazione del pensare e del sentire greci, rischia di essere, a sua volta, e in senso diverso da quello della riflessione heideggeriana, mito regressivo. Chi sono gli dèi? Anche a partire da questo punto ne va dell'essenza del *mythos*.

5. Storia dell'essere e processo storico

Ne va, a ben vedere, anche dell'essenza del *logos*, se, proprio in Parmenide, il *mythos* in senso heideggeriano e il *logos* rimandano l'uno all'altro. Ma, al contrario della rilettura di Heidegger del *mythos*, che ne dischiude nuova-

¹¹ Heidegger (1982b: 206).

¹² *Ibidem*.

mente il senso intendendo sottrarlo a successivi sviamenti di senso, l'interpretazione di questi del *logos* risente eccessivamente dell'intero impianto di riferimento, cui diventa quasi del tutto subalterna. Beninteso, grandezza e problematicità di Heidegger è il fatto che, attraverso il confronto con altri autori e l'ascolto del loro pensiero, parla lo stesso Heidegger. Ciò accade in particolar modo per Hölderlin, ma non ne sono immuni né Parmenide né Eraclito. Quando si dispiega il pensiero heideggeriano si può essere d'accordo o dissentire da esso, ma nel momento in cui questo rilegge il pensiero altrui fino a travolgerlo e a stravolgerlo, maggiore cautela è d'obbligo. Se quindi, pur non esente da rischi, come vedremo, la rilettura del *mythos* è assai fruttuosa, lo è meno quella del *logos*, inclinata non senza eccessiva pressione perché risulti organica alla riflessione heideggeriana nella sua interezza.

Ma leggiamo alcuni passaggi dell'interpretazione di Heidegger del *logos*:

Ciò che sia il λόγος lo possiamo desumere dal λέγειν. Che significa λέγειν? Chiunque conosca la lingua greca sa che λέγειν significa 'dire' e 'parlare' [...]. Chi può negare che nella lingua dei greci, fin dalle origini, λέγειν significhi parlare, dire, narrare? Però, in un tempo ugualmente antico e in modo ancora più originario, e quindi già sempre e perciò anche nel significato ora menzionato, λέγειν significa anche ciò che è espresso nella analoga parola tedesca 'legen': posare, metter dinnanzi (*nieder- und vorlegen*). [...] Λέγειν significa propriamente il posare e metter dinnanzi raccogliente se stesso e altro (*das sich und anderei sammelnde Nieder- und Vorlegen*). [...] Il dire e il discorrere dei mortali accade fin dalle origini come λέγειν, come 'posare'. Dire e parlare sono (*wesen*) come il 'lasciar-stare-insieme-dinnanzi' tutto ciò che, posato nella disvelatezza, è presente. [...] La parola ó λόγος designa quello che riunisce ogni cosa presente nella presenza e in essa la lascia-stare. [...] ó λόγος, pensato come il posare raccogliente, sarebbe anche l'essenza, pensata in modo greco, del Dire originario (*Sage*). Il linguaggio sarebbe Dire originario. Linguaggio sarebbe: riunente lasciar-stare-dinnanzi ciò che è presente nella sua presenza. Di fatto, i greci *abitavano* in questo modo di essere (*Wesen*) del linguaggio. Solo che essi non hanno mai *pensato* questa essenza del linguaggio, [...]¹³.

Come è evidente, la riflessione heideggeriana non manca di coerenza, ma comprime i termini-chiave del pensiero perché essi siano interni all'intero impianto. A prescindere dal fatto che quest'ultimo convinca o meno, si può rilevare un dato: quella sorta di forma particolare di "filosofia della storia" su cui si impernia il lavoro concettuale di Heidegger (una "filosofia della storia" senza *ratio*, cioè non retta da questa ma da un essere come *evento* che può diventare facilmente potenza arbitraria, cogliendo nel pensiero greco delle

¹³ Heidegger (1951: 142, 144, 155-156).

origini un bagliore che subito si ritrae) legge l'intero movimento storico successivo come "sviamento". Tale "sviamento" non è opera dell'uomo ma dell'essere stesso che si vela e si disvela, nella sua modalità di *evento*. Il passaggio, proprio della romanità, dall'*alétheia* alla *veritas* e dal *logos* alla *ratio*, non costituisce una mera *tra-duzione* ma consolida e sancisce un vero e proprio cambiamento di senso rispecchiando una fase dell'essere in cui esso è ritratto. È vero, ogni tradurre non è un mero trasportare un termine da una lingua all'altra: ogni termine porta con sé, e può altresì inaugurare, un orizzonte di senso di un'intera cultura, e perciò, letteralmente, è intraducibile. Inoltre, la storia dei termini conserva e mostra, rispecchiandola, tradizione, mentalità, e la storia stessa del pensiero. E, in questo senso, la lezione heideggeriana è imprescindibile e illuminante.

Ma la storia dell'essere come *evento*, con il suo *pathos* dell'originario, riesce veramente a comprendere in profondità il processo storico nella sua interezza? Qui si diverge dall'impostazione heideggeriana, che in questo sguardo storico, "storiografico", rinviene proprio un effetto dello "sviamento". In realtà la storia dell'essere così impostata risulta funzionale alla critica dell'idea di "progresso" e in particolar modo della modernità nel suo complesso. Non a caso la filosofia di Heidegger è intrisa di figure, di immagini che rimandano a un mondo pre-industriale e potenzialmente pre-moderno. La posizione di Heidegger in merito non è affatto semplicisticamente "anti-industriale", e tuttavia il nuovo *inizio* da questi auspicato e atteso, non senza toni messianici, comporta l'oltrepassamento della metafisica di cui la tecnica è il culmine, proprio perché questa rappresenta lo sbocco definitivo dell'entizzazione dell'essere e quindi la manifestazione suprema di un ente considerato come totalmente manipolabile. Riflessioni del genere hanno indubbiamente una portata, diremmo oggi, "ecologica", e rappresentano un'indicazione preziosa per contrastare e superare il dominio, l'*Herrschaft*, la signoria sulle cose e sull'uomo da parte dell'uomo: esse però mostrano, rispetto alla lettera heideggeriana, che Heidegger di certo non è così grossolano da rifiutare la modernità in quanto tale, ma che non l'accetta in realtà mai pienamente, e ne prefigura in questo senso un oltrepassamento che non è mero "ritorno" al passato, bensì, potremmo dire, *inizio* nel segno, nel solco del «primo inizio»¹⁴.

¹⁴ Su questo punto Heidegger stesso è molto chiaro: «l'inizio è iniziale solo se il pensiero stesso e l'uomo nella sua essenza pensano in modo iniziale. Ciò non significa l'impossibile, cioè una ripetizione del primo inizio e una sua traduzione in termini attuali concepita come rinnovamento della grecità, bensì questo: pensando in termini iniziali, entrare in discussione e in dialogo con l'inizio al fine di percepire la voce della disposizione e destinazione future» (Heidegger (1982b: 292)).

Ma la storia si sviluppa effettivamente attraverso uno svelamento e poi un velamento iniziali atti a mostrare e poi custodire la *verità*, oppure questa si dà *attraverso* un processo storico complesso, come *risultato* e non all'*origine*, e anzi, sorge e si manifesta, *diviene* attraverso il processo storico stesso?

Se osserviamo l'intero movimento storico da questa seconda prospettiva approdiamo a un senso storico robusto. Quest'ultimo non è un "punto di vista", così come non lo è l'impostazione heideggeriana. Questa rischia di essere chiusa in sé, di costituirsi come pensiero che liquida l'altro-da-sé come *ratio* calcolante e così si esaurisce in se stesso, divenendo autoreferenziale. Così però si impoverisce, e ripete se stesso. La critica della *ratio* calcolante, che ha al suo interno punti di estremo interesse e validità, tende ad essere in Heidegger la liquidazione della *ratio* stessa. Liquidata con essa, e per mezzo di questa operazione, è la storia concreta, alla quale si sostituisce una storia dell'essere dai tratti quasi insondabili, che solo una sorta di "iniziati" può forse cogliere. Il *logos* universale (e universalistico) si polverizza dinanzi a un intero impianto di tipo religioso-sacrale, esclusivistico.

Se invece cerchiamo, con sguardo storico-dialettico, di osservare il processo storico, le cose di certo non diverranno meno complesse, ma forse più chiare e più rispondenti a una realtà concreta e non mistificata da considerazioni di sapore quasi oracolare, in attesa di cenni di dèi, possibilità e appannaggio di pochi privilegiati.

Che la realtà sia problematica e complessa è un dato di fatto, ma l'itinerario per cercare di giungere alla sua comprensione non è fondato su attese messianiche e ascolto di messaggi rivelati a quei pochi che, quasi per motivi insondabili, sarebbero da loro stessi capaci di ascoltare. Esso si fonda sul duro lavoro del concetto.

Abbiamo visto un passaggio di Heidegger in cui egli esclude che «per pensare l'essere» servano «complicata erudizione», «sprofondamenti mistici» e «estasi in un'assorta pensosità». L'accorgersi che «l'ente 'è'» è, potenzialmente, alla portata di tutti. Questo è un primo passo, e, per inoltrarsi nella questione dell'essere, Heidegger richiede un pensiero che, oltrepassando la ragione calcolante, il pensiero rappresentativo-metafisico, si avvicini a un pensare-poetante, simile a quello espresso non a caso da Parmenide.

Tale pensare-poetante, come indicazione, è del tutto salutare e necessario in un tempo come il nostro in cui la filosofia si attegga ad *ancilla scientiae* per non "perdere terreno" rispetto al felice momento delle scienze sperimentali, per essere *accettata* e *accreditata* da una comunità scientifica internazionale che (non nella sua totalità) ha ossificato il *dato* in feticcio, a volte guardiano e sanzionatore dello *status quo*, "mito" in un senso che vedremo

fra poco, come regressione e fissazione. Ma il pensare-poetante non può sostituire/liquidare la *ratio*: si invertirebbe così lo stesso lavoro per mezzo del quale proprio Parmenide perviene al passaggio dal *mythos* al *logos*, segnalandone allo stesso tempo una loro particolare forma di coappartenenza.

6. Automovimento del concetto

Il contenuto del *logos*, della *ratio*, è universale non perché disponibile *immediatamente*, ma perché *universalmente* valido e rinvenibile. Ma, appunto, non come *immediato* bensì come risultato di un *lavoro*, di un processo, di un movimento. Tale movimento è interno, immanente alla storia stessa: la storia dei termini vi partecipa e lo rappresenta, ne mostra i passaggi e gli approdi. E allora, anziché a un alquanto difficilmente sondabile darsi e ritrarsi dell'essere che comporterebbe vicinanza all'autenticità e sviamenti rinvenibili nel linguaggio, si può concentrare lo sguardo su ciò che i termini indicano di volta in volta, ma non in base a una sorta di *volontà dell'essere* bensì alla logica storica a loro immanente. E così, se l'*alétheia* come dis-velamento è diventata *veritas* come certezza e corrispondenza di un enunciato con una realtà oggettiva, e se il *mythos* diviene *logos*, e il *logos* acquista un suo significato addirittura opposto al *mythos*, e man mano assume la forma sempre più *calcolante*, ciò non è dovuto a un essere sfuggente, astratto perché vuoto di contenuto in base alla sua potenziale arbitrarietà, e, allo stesso modo, per ciò stesso potenzialmente autoritario, ma al movimento storico stesso, ricostruibile razionalmente, anzi, potremmo spingerci a dire, *mosso* dalla *ratio* stessa, sebbene non ogni suo momento o risultato possano essere considerati razionali. E però lo è il suo contenuto, in una lunga provenienza e durata, come intero.

La storia dell'essere è allora la storia del pensiero che si sviluppa attraverso il lavoro del concetto, attraverso l'automovimento dei concetti che divengono altro-da-sé. Questo divenire altro-da-sé indica che l'automovimento, l'immanenza, non è autoreferenziale, ma implica la presenza e l'irrompere del non-identico. Occorre pensare l'identità come al contempo identica e non-identica a se stessa, e pur sempre identità. *Pensare* comporta il pensare questa contraddizione. In altri termini, per movimento *immanente* un concetto *trascende* se stesso.

Ciò si mostra nel linguaggio: per movimento immanente il *mythos* diviene *logos*, per movimento immanente l'*alétheia* diviene *veritas*, per movimento immanente il *logos* diviene *ratio*. Ciò non significa che ogni passaggio sia necessariamente, e in termini assoluti, un progresso, ma che ciò avviene per una logica che, *immanente*, *trascende*.

7. Parmenide

Questo movimento è presente e osservabile nei frammenti parmenidei. Risalenti al VI secolo a.C., i frammenti *Sulla natura* possono essere considerati come la fondazione dell'ontologia occidentale e come uno dei passaggi decisivi per l'identità del pensiero occidentale in quanto tale. Questa fondazione, questo *inizio*, avviene in una zona centrale del Mar Mediterraneo, a Elea, nella Magna Grecia, l'odierna, campana, Velia. Qui ci concentreremo non sui contenuti del pensiero tramandatici da Parmenide, ma sul metodo, sulla *via* attraverso cui il filosofo giunge al proprio pensiero. In questi frammenti è visibile il passaggio dal *mythos* al *logos*.

L'opera inizia con un Proemio nel quale l'autore *racconta* in che modo sia venuto a conoscenza delle verità che verranno di lì a poco esposte. Sarebbe meglio dire *rivelate*, in quanto, come sappiamo, è una dea a *rivelarle* a Parmenide. Ma chi è questa dea?

Sappiamo già che ella viene nominata in modi diversi, che la sua origine affonda molto probabilmente nel culto della Grande Madre, ma qui la sua figura è divenuta più complessa, e di difficile interpretazione. Sta di fatto che si è in presenza di un *mythos*, da declinare secondo l'indicazione heideggeriana. La dea *parla, dice, rivela*. Parmenide viene portato dinanzi la dea: «là mi portarono accorte cavalle / tirando il mio carro, e fanciulle indicavano la via»¹⁵. Giunto alla «porta dei sentieri della Notte e del Giorno»¹⁶, egli assiste all'apertura della porta da parte della Giustizia per intercessione delle «fanciulle Figlie del Sole»¹⁷. Quindi,

di là, subito, attraverso la porta, / diritto per la strada maestra le fanciulle guidarono carro e cavalle. / E la Dea di buon animo mi accolse, e con la sua mano la mia mano destra / prese, e incominciò a parlare così e mi disse: / 'O giovane [...] Bisogna che tu tutto apprenda: / e il solido cuore della Verità ben rotonda / e le opinioni dei mortali, nelle quali non c'è una vera certezza [...]»¹⁸.

L'opera vera e propria è il *dire* della dea:

Ora, io ti dirò – e tu ascolta e ricevi la mia parola - / quali sono le vie di ricerca che sole si possono pensare: / l'una che 'è', e che non è possibile che non sia / - è il sentiero della Persuasione, perché tien dietro alla Verità - / l'altra che 'non è', e che è necessario che non sia / E io ti dico che questo è un sentiero su cui nulla si

¹⁵ Parmenide, *Sulla natura*, p. 41.

¹⁶ *Ibidem*.

¹⁷ *Ibidem*.

¹⁸ *Ibidem*. pp. 41-43.

apprende. / Infatti, non potresti conoscere ciò che non è, perché non è cosa fattibile,
/ né potresti esprimerlo¹⁹.

La dea enuncia, rivela, delle proposizioni caratterizzate da una logica di assoluto rigore rispetto all'essere e al non-essere. Il *mythos* della dea *dice* il *logos*. Ancora, lo *nomina* espressamente:

È necessario il dire e il pensare che l'essere sia: infatti l'essere è, / il nulla non è: queste cose ti esorto a considerare. / [...] questo non potrà mai imporsi: che siano le cose che non sono! / Ma tu da questa via di ricerca allontana il pensiero, / né l'abitudine, nata da numerose esperienze, su questa via ti forzi / a muovere l'occhio che non vede, l'orecchio che rimbomba / e la lingua, ma con la ragione giudica la prova (*κρίναι δέ λόγοι [...] ἐλεγκόν*) molto discussa / che da me ti è stata fornita²⁰.

L'invito è quello di ricercare la verità non affidandosi ai sensi, che possono essere ingannevoli, ma alla ragione; in altre parole la dea *dice* che la verità (ed ella stessa è figura della Verità) non è verità dei sensi ma di *ragione*. La verità, per bocca della dea stessa, non è rivelazione fideistica, ma risultato del vaglio di una *prova*, vaglio effettuato non dai sensi o dal sentimento, dalla volizione, dal desiderio, ma dal rigore del *logos*, della *ragione*.

Gli sviluppi successivi, che si riflettono e si danno essi stessi nei termini nelle *traduzioni* delle diverse fasi storiche, sono frutto di un velamento dell'essere che ne custodisce la verità iniziale, oppure sono sviluppi, necessari o potenziali, ma ricostruibili razionalmente secondo quella dialettica dell'immanente e del trascendente a cui abbiamo fatto riferimento in precedenza?

È vero, la *ratio*, in particolar modo quella meramente calcolante, liquida il *mythos* come fantasia irrazionale o inutile residuo del passato, ma nel fare questo si irretisce in se stessa, diviene essa stessa "mito".

8. Figure del *mythos* e *logos*

Il senso in cui abbiamo appena utilizzato la parola "mito" diverge dal *mythos* della dea che si rivolge a Parmenide, eppure appartiene alla medesima storia, solo che, a causa della storia stessa e di quell'automovimento del concetto di cui abbiamo parlato, la *stessa* parola indica qualcos'altro. Adorno, assai brillantemente, definisce il «mito» come «concetto immobilizzato o ipostatizzato»²¹. Come lo stesso Adorno insegna, dobbiamo cercare di guardare ai

¹⁹ *Ibidem*. p. 45.

²⁰ *Ibidem*. pp. 48-49.

²¹ Adorno (1973, trad. it. 481).

concetti senza a nostra volta renderli *immobili* in senso assoluto. Circa vent'anni prima di questa "definizione", Horkheimer e Adorno avevano parlato del mito in termini diversi, in una complessa connessione con l'illuminismo:

I miti che cadono sotto i colpi dell'illuminismo erano già il prodotto dell'illuminismo stesso. [...] Il mito voleva raccontare, nominare, dire l'origine: e quindi anche esporre, fissare, spiegare. Questa tendenza si è rafforzata con la stesura e la raccolta dei miti, che diventarono presto, da racconto di cose avvenute, dottrina. [...] Le divinità olimpiche non sono più direttamente identiche agli elementi, ma li significano. In Omero Zeus presiede al cielo diurno, Apollo guida il sole, Helios ed Eos sono già ai limiti dell'allegoria. Gli dèi si separano dagli elementi come essenze dei medesimi²².

Di «introduzione allegorica al poema περί φύσεως»²³ parmenideo parla Hegel nelle sue *Lezioni sulla storia della filosofia*. E in effetti già Sesto Empirico, grazie al quale abbiamo una versione completa del Proemio, interpreta tutte le figure presenti in questi versi in quanto allegorie (dei sensi, dei desideri, della ragione²⁴). Possiamo parlare, allora, già per Parmenide, di pensiero allegorico? E nel senso che questo potrebbe avere ai nostri giorni?

Torniamo a leggere Adorno, e facciamo un passo in avanti nella storia del pensiero greco, prendendo in considerazione Platone, che, com'è noto, fa uso del *mythos* con valore esplicativo:

I miti non sono delle semplici aggiunte, delle allegorie o delle escursioni nelle regioni superne che hanno luogo quando capita l'occasione; questi miti sono invece un fattore costitutivo del pensiero platonico. I miti platonici [...] Potremmo dire che sono il tentativo, nel movimento dialettico che termina con essi, di esprimere, come momento dialettico, proprio ciò che sfugge alla formazione soggettiva del concetto. [...] Ogniqualvolta compaiono questi miti, l'incapacità del soggetto di cogliere interamente la realtà in sé a cui essi si riferiscono viene espressa rappresentando lo stato del soggetto che parla – uno stato di ispirazione, che si tratti della veggente o di colui che si trova in un luogo sacro ed è colto dalla divina mania²⁵.

Potremmo fare delle considerazioni (anche se non del tutto) analoghe riguardo Parmenide. Werner Jaeger parla del Proemio parmenideo come del frutto di «un'autentica esperienza religiosa»²⁶, e a tal proposito aggiunge

²² Horkheimer & Adorno (1944, trad. it. 16).

²³ Hegel (1840-1844, trad. it. 136).

²⁴ Cfr. Parmenide, *Sulla natura*, p. 129 nota 2.

²⁵ Adorno (1973, trad. it. 478-479).

²⁶ Jaeger (1947, trad. it. 154); Reale (2001: 19).

Reale: «Si tratta di una religione che, ovviamente, non rientra nell'ambito della religione statale, ma piuttosto in quel tipo di religiosità connessa alle iniziazioni e ai misteri, diffuse nell'Italia meridionale»²⁷. Da questa culturalità mediterranea sorgono le figure simboliche che esprimono il passaggio dal *mythos* al *logos*.

Il *mythos* deve esprimere l'inesprimibile, o meglio, deve cercare di rappresentare un inespresso, tale perché difficile da *dire*, e non solo perché non sia ancora pervenuto al suo concetto. In Parmenide il *mythos* dice del principio, della fondazione, della fonte del proprio pensiero. Qual è l'origine, come fondamento, del contenuto del pensare parmenideo? La dea, che è Verità, *Alétheia*, ma anche «Divinità che tutto governa (*δαίμων ἢ πάντα κυβερνᾷ*)»²⁸ e «legge divina (*θέμις*)»²⁹.

Proprio perché della questione del fondamento si tratta, Heidegger vede la dea (e gli dèi) come appartenenti all'ambito che *l'aperto* apre, ma *l'aperto dell'essere* è *das Boden-lose*. Qui si giunge ai confini del *dicibile* (problema da cui nasce forse il *mythos* stesso): la soluzione heideggeriana può non convincere, ma *pone* il problema.

Dall'*aperto* la dea, come *mythos*, dice il *logos*: questo *inizio* non è il *vero come assoluto*, perché il *vero* perviene a se stesso come risultato di un *lavoro del concetto*. Il *luogo* di questo *lavoro* è la *storia*. Qui il *mythos* diviene *logos*. Quest'ultimo, se irretito in se stesso, diviene a sua volta di nuovo *mythos*, ma non in senso originario, perché il tempo fa il suo corso: come forma originaria il *mythos* è passato. Se torna come *mythos* o come *logos* cui il movimento stesso viene pietrificato, torna come «concetto immobilizzato o ipostatizzato». Così il «mito» è regressivo. Esso può invece *tornare* come *pensiero allegorico* che, attraverso le sue figure, può *trascendere* il *logos*, liberandone il movimento. Tale *pensiero*, però, non è *assolutamente-altro-dal logos*, né è una forma, come il *mythos*, di cui conserva le tracce come *ricordo*. Le sue figure non sono *vere* in se stesse, ma rimandano oltre-se-stesse, nel movimento verso il concetto.

Nelle acque del Mediterraneo, Odisseo compì il suo viaggio, le cui figure sono originarie e decisive dell'Occidente. Horkheimer e Adorno si chiedono quale sorte sia toccata alle Sirene dopo il passaggio di Odisseo, che ne udì il canto, e proseguì, grazie a uno stratagemma:

L'epos non dice che cosa accade alle cantatrici dopo che la nave di Odisseo è scomparsa. Ma nella tragedia sarebbe stata certo la loro ultima ora, come per la

²⁷ Reale (2001: 19).

²⁸ Parmenide, *Sulla natura*, pp. 58-59.

²⁹ *Ibidem*. pp. 42-43.

Sfinge quando Edipo risolve l'indovinello, eseguendo il suo ordine e così rovesciandola. Poiché il diritto delle figure mitiche, che è il diritto del più forte, vive solo dell'ineseguibilità delle loro norme. Se esse vengono soddisfatte, i miti si dissolvono fino alla più lontana posterità³⁰.

Mentre in Parmenide il *mythos* diviene *logos* quasi da se stesso, qui il superamento del *mythos* implica un sacrificio, una sofferenza, una perdita. In entrambi i casi, in una lunga durata, le figure del *mythos* si dissolvono come valide in se stesse. Ne restano il ricordo, la potenza. Odisseo oltrepassa le Sirene, ed esse svaniscono, ma il loro canto è stato udito. Oltre la porta dei sentieri della Notte e del Giorno la dea rivela a Parmenide, dice il *logos*, e nel dirlo si dissolve, perché ha impresso un movimento che la trascende. Ma la sua parola, il suo *dire*, sono stati uditi.

Delle Sirene e della dea restano le figure, con il loro canto e il suo dire: esse indicano la possibilità, la verità di una *ragione* liberante, di un *logos* emancipatore, di una conciliazione del *mythos* con il *logos* nell'unità del concetto.

Bibliografia

Adorno, Th. W. (1973), *Philosophische Terminologie (Terminologia filosofica)*, trad. it. di A. Solmi, Torino, Einaudi, 2007).

Grimaldi, G. (2012), "Essere, ente e verità. Alcune annotazioni sul Parmenide di Heidegger", in *Prospettive filosofiche, Ontologia*, di C. Tatasciore, P. Graziani, G. Grimaldi (a cura di), Acireale-Roma, Bonanno Editore, pp. 185-203.

Hegel, G. W. F. (1840-1844), *Vorlesungen über die Geschichte der Philosophie (Lezioni sulla storia della filosofia)*, trad. it. di R. Bordoli, Roma-Bari, Laterza, 2009).

Heidegger, M. (1951), "Logos", in *Saggi e discorsi*, trad. it. di G. Vattimo, Milano, Mursia, 1976, pp. 141-157).

Heidegger, M. (1957), "Moira", in *Saggi e discorsi*, trad. it. di G. Vattimo, Milano, Mursia, 1976, pp. 158-175.

³⁰ Horkheimer & Adorno (1944, trad. it. 66-67).

Heidegger, M. (1982a), *Gesamtausgabe. II. Abteilung: Vorlesungen 1923-1944. Band 54. Parmenides*, Frankfurt am Main, Vittorio Klostermann.

Heidegger, M. (1982b) *Parmenides (Parmenide*, trad. it. di G. Gurisatti, Milano, Adelphi, 1999).

Horkheimer, M., Adorno, Th. W. (1944), *Dialektik der Aufklärung. Philosophische Fragmente (Dialettica dell'illuminismo*, trad. it. di R. Solmi, Torino, Einaudi, 1997).

Jaeger, W. (1947), *The Theology of the Early Greek Philosophers (La Teologia dei primi pensatori greci*, trad. it. di E. Pocar, Firenze, La nuova Italia, 1961).

Parmenide, *Sulla natura*, a cura di G. Reale, Milano, Bompiani, 2001.

Cosa causa l'irrazionalità motivata

Patrizia Pedrini
Università di Firenze
patpedrini@gmail.com

1. Il fenomeno dell'irrazionalità motivata e le sue cause

Una venerabile tradizione filosofica ci ha invitati a pensare alla ragione e alle passioni come a due fenomeni in competizione tra di loro per l'indirizzo dell'esistenza umana. Secondo questa immagine generale della mente umana, la ragione può determinare la nostra condotta e il nostro pensiero in molte occasioni rilevanti della nostra vita quotidiana. Tuttavia, ci può accadere di essere sopravanzati da una forte passione e quindi agire, valutare situazioni e trarre conclusioni sulla scorta del potente movente di ciò che proviamo, trascurando in parte o del tutto quello che la ragione consiglia. Quando questo accade, il fenomeno che si verifica è una forma di irrazionalità precipua, definita "irrazionalità motivata".¹ Essa viene generalmente intesa come un evento mentale che non è riconducibile né a un limite qualitativo delle nostre facoltà razionali, come potrebbe esserlo la difficoltà a ragionare in termini probabilistici, né a un limite quantitativo, come ad esempio l'incapacità di processare tutti i dati che si offrono alla nostra cognizione cosciente. Al contrario, colui che cade nell'irrazionalità motivata sarebbe perfettamente in grado di trarre la conclusione pratica, doxastica o inferenziale corretta, se solo non fosse in preda a una passione che – per così dire – lo acceca. In altre parole, la causa dell'irrazionalità motivata sarebbe la passione, che travolge la ragione e la sua capacità di determinare l'individuo, assumendo il comando dell'azione o del pensiero. Tra le passioni che nell'economia psicologica della nostra esistenza avrebbero questo potente ruolo causale vengono generalmente elencate le emozioni, i desideri e i sentimenti.

¹ Pears (1984).

Nel dibattito contemporaneo dei paesi di lingua inglese, vengono studiati diversi fenomeni che appartengono alla famiglia dell'irrazionalità motivata. Tra questi, figurano centralmente le credenze precipitosamente formate sulla base di un'emozione, il cosiddetto *wishful thinking* (letteralmente "pensiero che è l'espressione di un desiderio") e l'autoinganno.

Nel caso delle credenze formate precipitosamente sulla base di un'emozione, il soggetto prova una forte emozione, come ad esempio la paura, e sulla scorta di essa viene spinto a credere una certa proposizione. Poniamo che qualcuno tema fortemente i serpenti e stia camminando per un sentiero di montagna. A un certo punto avvista in lontananza qualcosa per terra le cui forme somigliano vagamente a quelle di un serpente. Preso dal timore di stare per imbattersi in un serpente, conclude precipitosamente che quello che sta vedendo è un serpente².

Il caso del *wishful thinking* ha una struttura simile, anche se tipicamente si svolge in un arco di tempo di durata variabile, ma non necessariamente brevissima, come nel caso delle credenze formate precipitosamente. Il soggetto indulge per un certo tempo nel pensiero di ciò che desidera, ad esempio di essere oggetto dell'attenzione speciale di qualcuno, e questo pensiero, sostenuto dal forte desiderio che le cose stiano così come lui vorrebbe, finisce per installarsi sotto forma di qualcosa di molto simile a una credenza. Ad esempio, alcune azioni del soggetto saranno condizionate dall'intensa attività di pensiero sostenuta da questo desiderio.

I confini tra il *wishful thinking* e l'autoinganno sono tradizionalmente discussi³, ma in generale nell'autoinganno si ammette che una credenza *tout court* viene raggiunta sulla scorta della forte motivazione a crederla e nonostante la presenza, o perlomeno la disponibilità, di una forte evidenza contraria, che nel *wishful thinking* non sembra necessariamente presente. Vedremo meglio tra poco attraverso quali meccanismi si è a lungo ritenuto che questo possa accadere.

La spiegazione causale di massima per questi fenomeni è dunque quella secondo cui la motivazione a pensare o credere una certa proposizione opera in maniera tale da produrre tale pensiero o credenza, nonostante il soggetto sia perfettamente in grado, dal punto di vista delle sue facoltà razionali, di raggiungere la conclusione doxastica o il pensiero corretto. L'immagine della mente sottesa a tale spiegazione è quella della tradizione: quando le passioni ardono, la ragione viene vinta.

² Cfr. Scott-Kakures (2002).

³ Cfr. Pedrini (2013, cap. 1),

Questo quadro ha tuttavia cominciato a complicarsi. Alcune raffinate posizioni elaborate di recente suggeriscono di pensare al rapporto tra ragioni e passioni in maniera non più dicotomica e maggiormente composita. Alcuni⁴ sostengono che le passioni, quantunque antagoniste alla ragione, siano tuttavia immerse nel cosiddetto “spazio delle ragioni”.⁵ Altri pensano che passioni e ragioni siano moventi di natura distinta, ma che alla base delle passioni ci sarebbero ragioni⁶.

Tutto questo intreccio fa pensare che un problema causale di fondo affligga la nostra comprensione del fenomeno dell'irrazionalità motivata. Se le passioni dipendono da ragioni o ne sono composte, l'individuazione nelle passioni del *primum* causale dell'irrazionalità motivata sembra una via meno promettente di quanto non appaia a prima vista. Lontana dal voler aggiudicare la disputa generale del rapporto causale tra ragioni e passioni, mi concentrerò su uno specifico fenomeno di irrazionalità motivata, l'autoinganno, e discuterò in che modo il problema causale si declina in relazione ad esso. Nel § 2 esporrò in cosa consista il fenomeno dell'autoinganno e quale sia la spiegazione causale al momento prevalente. Nel § 3 metterò in luce una potente obiezione alla spiegazione prevalente, nota come *problema della selettività*, la quale genera un problema causale per l'autoinganno, che potrebbe essere esteso anche alle altre forme di irrazionalità motivata. Nel § 4 cercherò quindi di tracciare una soluzione al problema causale dell'autoinganno, offrendo i lineamenti di una teoria che renda contemporaneamente conto del problema della selettività. Nel § 5, infine, trarrò da questa teoria alcune implicazioni che reputo rilevanti per ulteriori sviluppi della ricerca.

2. Il caso dell'autoinganno

La vita quotidiana, nelle sue varie dimensioni (privata, sociale, politica), ci presenta una grande quantità di casi di persone che manifestano verbalmente o semplicemente attraverso il loro comportamento di essere convinte (o di cercare di convincersi) della verità di proposizioni che sono improbabili o scarsamente supportate dall'evidenza. Tuttavia, la verità di tali proposizioni è l'oggetto di un interesse speciale da parte dei soggetti coinvolti e sembra condurli a finire per crederle a tutti gli effetti. Un controfattuale sembra valere: se non fosse per questo interesse che nutrono, i soggetti si accorgerebbero che tali proposizioni non sono adeguatamente supportate dall'evidenza

⁴ Scanlon (1998).

⁵ McDowell (1994).

⁶ Wells (1997).

e non finirebbero per crederle. La formulazione dell'autoinganno appena delineata indica quindi che il meccanismo causale dell'autoinganno è da ricondursi al lavoro sotterraneo della motivazione a credere una certa proposizione p . Tale motivazione, che può essere sostanziata dal desiderio di credere che p , o da forti emozioni associate a p , opera surrettiziamente sulla cognizione del soggetto, fino a indurlo a formare la credenza che p . La teoria dell'autoinganno divenuta popolare negli ultimi 15 anni circa spiega l'autoinganno in questo modo. È nota come *motivazionalismo*⁷, ma è stata per la prima volta formulata da Al Mele⁸ con il nome di *anti-intenzionalismo*, al fine di enfatizzare il carattere non intenzionale e non volontario del fenomeno dell'autoinganno, di contro alla spiegazione offerta da Donald Davidson negli anni Ottanta⁹. Secondo Mele, l'autoinganno non è altro che il prodotto di un desiderio che p , capace di distorcere la cognizione rilevante per la formazione della credenza che p . Il soggetto che nutre un desiderio che p viene indotto a credere che p attraverso l'attivazione, da parte del desiderio che p , di una serie di *bias* che gli impediscono la raccolta equilibrata e la valutazione corretta dell'evidenza a favore e a sfavore di p . Nonostante l'evidenza disponibile mostri chiaramente che p non è supportata, o non sufficientemente, il soggetto che desidera che p si trova a raccogliere, selezionare e interpretare tale evidenza in maniera fortemente pregiudiziale, cosicché diventa facile abbracciare la conclusione doxastica che p . Tra gli esempi di autoinganno riportati in letteratura vi sono persone che credono che il partner sia fedele, nonostante le evidenze di una probabile relazione extraconiugale; medici che credono di essere in buona salute, nonostante i segnali di malattia che sono tipicamente in grado di interpretare correttamente nei loro pazienti; studiosi che credono che il loro lavoro sia stato ingiustamente respinto, sebbene le giustificazioni per il rifiuto siano valide e tali che loro stessi le considererebbero valide, se riguardassero un'altra persona¹⁰. Secondo Mele, queste persone riescono a formare la credenza autoingannevole grazie a una serie di manovre pregiudiziali, o *bias*, di raccolta e interpretazione dell'evidenza. Tra questi, Mele elenca la raccolta selettiva dell'evidenza a favore della credenza che è oggetto di preferenza, che viene attivamente cercata e valorizzata dal soggetto a discapito di quella a sfavore, l'interpretazione pregiudizialmente favorevole dell'evidenza che sostiene la proposizione autoingannevole, e l'interpretazione pregiudizialmente sfavorevole di quella che la sconfessa. Mele usa il

⁷ Funkhouser (2005).

⁸ Mele (1997, 2001).

⁹ Cfr. Davidson (1982, 1985).

¹⁰ Cfr. Pedrini (2013, cap. 1).

modello del test delle ipotesi ordinarie di Trope e Liberman¹¹ per spiegare come il desiderio che p sia vera abbassi le soglie di accettazione dell'ipotesi che p desiderata, rendendola per il soggetto meno costosa, in termini di sforzo cognitivo ed emotivo necessario per formarla e intrattenerla, della verità contraria. Tale ipotesi viene quindi abbracciata dal soggetto più facilmente di quanto non lo sarebbe se il soggetto non nutrisse quel desiderio.

Insomma, secondo il motivazionalismo, uno stato motivazionale, come ad esempio un desiderio che p , è sufficiente a causare la credenza autoingannevole che p , attraverso l'attivazione dei *bias* elencati e la distorsione della cognizione rilevante alla formazione di una credenza circa l'ipotesi che p .

3. Il problema della selettività per l'autoinganno

Nonostante l'*appeal* di questa spiegazione, il motivazionalismo è stato posto sotto attacco da una potente obiezione, mossa da Bermúdez nel 2000, e nota come il *problema della selettività*. L'obiezione è la seguente: sebbene la spiegazione motivazionalista abbia una sua presa intuitiva, a ben vedere non offre una spiegazione davvero convincente dell'autoinganno, a misura che vi sono moltissime persone che, pur in preda a un forte desiderio che p , non finiscono affatto per credere autoingannevolmente che p . In altre parole, i *bias* che secondo il motivazionalismo verrebbero attivati dal desiderio che p , non si attivano in tutti i soggetti. Molti mostrano di essere di fatto protetti dall'effetto distorsivo dei desideri, o di altri stati motivazionali, sulla cognizione. In questo modo, l'autoinganno appare come *selettivo* rispetto alla totalità dei soggetti che nutrono un desiderio che p . L'autoinganno è, cioè, più selettivo di quanto la mera presenza del desiderio che p consente di prevedere. Da questo sembra conseguire che occorrono altri elementi psicologici oltre ai desideri perché l'autoinganno si verifichi.

In tre tentativi precedenti di rispondere al problema della selettività¹² ho argomentato che il motivazionalismo dovrebbe essere incrementato attraverso l'introduzione di una spiegazione causale dell'autoinganno che attinga alla struttura valutativa dei soggetti che si autoingannano per spiegare perché l'autoinganno si verifichi selettivamente. Ho sostenuto che coloro che si autoingannano non possiedono in maniera sicura valori epistemici come il valore della verità, il valore dell'accuratezza e quello dell'imparzialità e della completezza nella raccolta e nella valutazione dell'evidenza. Il sicuro possesso di

¹¹ Trope & Liberman (1996).

¹² Pedrini (2010, 2013, 2014).

valori epistemici consente infatti di esercitare un effetto di controbilanciamento rispetto ai *bias* attivati dai desideri. Coloro che resistono all'autoinganno di fatto li possiedono e attraverso la loro azione è possibile evitare l'autoinganno.

Dopo questi primi tentativi, ho perseguito ulteriormente questa linea di ricerca e ho cominciato ad elaborare i lineamenti di quella che nelle mie intenzioni vuole essere una teoria più ampia del set causale dell'autoinganno. Si tratta di una teoria che tenta di fare appello non solo e non tanto ai valori epistemici del soggetto che si autoinganna, ma più in generale alla sua specifica psicologia individuale, con i suoi ingredienti distintivi, che ne fanno la persona che è.

4. Il problema causale dell'autoinganno: lineamenti di una soluzione

Come abbiamo visto, il problema della selettività mostra che il motivazionalismo come spiegazione dell'autoinganno soffre di un importante *problema causale* ancora irrisolto. La mia ipotesi di lavoro è che quando un desiderio che p sia vera distorce la cognizione del soggetto così da indurlo a credere autoingannevolmente che p , questo succede perché il desiderio che p sia vera non è funzionalmente equivalente al desiderio che p sia vera che opera nel soggetto che non finisce per autoingannarsi. Al contrario, si tratta di un desiderio che la psicologia generale del soggetto in questione rende funzionalmente adatto a consentirgli di raggiungere la credenza autoingannevole. I desideri sono infatti intrecciati in maniera significativa con tutta una serie di altri elementi della nostra psicologia individuale, che li plasma funzionalmente in maniera da renderli causalmente operativi in modi che sono diversi da individuo a individuo. È quindi necessario fare luce su questi altri elementi della nostra psicologia individuale al fine di identificare correttamente il set causale dell'autoinganno. In questo modo soltanto possiamo arrivare a una comprensione attendibile dei differenti esiti cognitivi in soggetti che nutrono parimenti un desiderio che p sia vera. Per raggiungere questo obiettivo esplicativo, introduco il concetto dell'*abitare un punto di vista*. Questo concetto servirà a individuare le caratteristiche dell'esistenza individuale dell'autoingannato che sono cruciali nel determinare il ruolo funzionale specifico del desiderio che p sia vera all'interno della sua psicologia.

La vita di un soggetto è caratterizzata dal fatto che tale soggetto inevitabilmente abita un punto di vista su questioni e situazioni che si offrono al suo giudizio e al suo sentire. Tale punto di vista sul mondo, o su porzioni del mondo, è definito da un complesso insieme di credenze, opinioni, valori,

emozioni, desideri, aspettative, impegni (moralì, sociali, legali) e sentimenti. Le interconnessioni tra tutti questi elementi sono immensamente variabili da soggetto a soggetto e plasmano la specifica personalità e il carattere del soggetto, con il suo specifico punto di vista sul mondo. È cruciale il fatto che, abitando un punto di vista, il soggetto è sottoposto anche a tutta una serie di condizionamenti circa le verità che gli è possibile tollerare senza smettere di occupare quel punto di vista, cioè senza smettere di essere la persona che è. Un esempio può essere utile a illustrare questa idea. Roberto ha un disturbo paranoide di personalità. La sua personalità e la sua vita ruotano attorno a un certo numero di cardini psicologici che definiscono la sua condizione psicologica generale e fanno di lui la persona che di fatto è. Roberto crede (non solo inconsciamente) che le relazioni affettive intime, di natura sia amorosa sia amicale, siano delle minacce cospicue al suo benessere. Il suo comportamento generale durante le relazioni affettive è fortemente condizionato da questa convinzione. Per esempio, non è incline a intraprendere conversazioni in cui l'interlocutore si attende che egli riveli dettagli anche abbastanza ordinari della sua vita privata, come quello che ha fatto ieri, o chi sono le persone che frequenta. Fa parte del suo quadro di personalità credere che se rivela tali dettagli, questi potranno essere usati contro di lui in vari modi. Nel momento stesso in cui questi contesti conversazionali accennano a manifestarsi, Roberto diviene fortemente ansioso e potenzialmente aggressivo. Se costretto ad affrontare le lamentele di coloro che sono entrati in relazione con lui e che notano che è reticente e disturbato in maniera anomala dalle loro domande innocenti, Roberto tende a reagire in maniera spropositata. Se l'interlocutore insiste, lo afferrano reazioni incontrollate di rabbia. La relazione che Roberto ha avuto nell'infanzia con sua madre è stata segnata da una serie di violenze emotive piuttosto significative. Sua madre era infatti solita tenere sotto controllo i vissuti emotivi e cognitivi del figlio, per poi infliggergli umiliazioni. Così Roberto ha imparato che non deve fidarsi delle persone che ama. Sulla base di questa convinzione, ha costruito la sua personalità e la sua intera vita. Se qualcuno chiede a Roberto perché si comporta in questo modo nelle relazioni affettive intime, o se si rende conto che la sua ansia e la sua aggressività sono sproporzionate in maniera anomala rispetto al contesto affettivo e conversazionale nel quale si manifestano (le persone che intrattengono relazioni affettive intime generalmente si aprono volentieri le une con le altre), egli risponde presentando una *storia di copertura* autoingannevole: è solo una persona estremamente riservata, semplicemente non ama quel genere di domande sulla sua vita e i suoi sentimenti, ed è per giunta banalmente intollerante verso coloro che tentano di *violarlo* con domande di quel genere. Da risposte come questa, si evince che Roberto non riesce a intrattenere l'ipotesi

che c'è molto di più di questo da dire sulla sua reale condizione e sui suoi comportamenti, che c'è perlomeno una possibilità che tali comportamenti siano il segnale di un disordine di personalità abbastanza serio. Se per caso questo pensiero lo sfiora, Roberto giudica questa possibilità improbabile, visto il carattere che ha (riservato e intollerante) e vista l'*oggettiva* mancanza di rispetto mostrata da tutti coloro che fanno domande di questo genere alle persone. Gli è capitato di avere molte donne con questo *difetto*, purtroppo, ed è stato *costretto* a interrompere queste relazioni non appena le sue partner mostravano di volerne sapere di più di lui. Ha così trascorso la gran parte della sua vita da solo, dicendo a se stesso che se lo è, è perché le donne che ha incontrato gli hanno sempre mancato di rispetto.

Come abbiamo visto, qualche volta Roberto arriva a pensare per un attimo che in lui c'è qualcosa che non va, ma è estremamente costoso testare questa ipotesi, per lui. La domanda che ci dobbiamo porre, quindi, è perché per Roberto è così costoso testare questa ipotesi. Il motivazionalismo risponde che è costoso testarla perché desidera che non ci sia niente che non vada in lui. Ora, però, la maggior parte di noi nutre questo desiderio. Nondimeno, molti di noi non rigetterebbero questa ipotesi, anche se costosa, nel momento in cui si presenta l'evidenza che qualcosa non va. Per molti di noi, insomma, il costo cognitivo ed emotivo necessario per testare questa ipotesi non è tale da sopravanzare il bisogno di sapere come stanno le cose, per poi intervenire, dove possibile. Roberto, però, in questo non ci somiglia. Nella sua psicologia, sono presenti ingredienti che non hanno un corrispettivo nella psicologia di altri e che gli impediscono di prendere seriamente in considerazione questa ipotesi e procedere a testarla correttamente e in maniera spassionata. Innanzitutto, la storia delle violenze emotive alle quali è stato sottoposto nella sua infanzia e adolescenza gli hanno fatto sviluppare un'impressionante volontà di sopravvivere al regime emotivo che vigeva allora, e questa volontà si è manifestata per mezzo della forte affermazione di se stesso rispetto a un mondo affettivo minaccioso. Ha elaborato obiettivi forti da raggiungere in maniera solitaria, senza l'assistenza di nessuno, contando solo su se stesso. Ha imparato a non fidarsi di nessuno, come abbiamo visto. Quindi è abbastanza normale per lui non prendere in considerazione il consiglio e il parere degli altri. Se qualcuno gli fa notare che qualcosa in lui non va o lo mette di fronte a questa evidenza, dirà a se stesso che non c'è da fidarsi di questa opinione, che la persona ha intenzioni malevole, e via dicendo. In secondo luogo, ha sviluppato il forte timore di essere stato seriamente danneggiato dalla madre, che nel suo scrutinio umiliante, spesso lo tacciava di essere pazzo. Per lui, salvarsi da questo regime ha significato anche avere la forza di dire da solo a se stesso che non era veramente pazzo. Insomma, intrattenere e testare

l'ipotesi che qualcosa in lui non va è qualcosa che metterebbe seriamente a repentaglio i fondamenti stessi della persona che di fatto è, il punto di vista che abita. Per farlo, dovrebbe essere non Roberto, ma qualcun altro, con un altro punto di vista sul mondo.

A questo punto, siamo in una posizione più favorevole per comprendere perché testare tale ipotesi è per lui costoso in maniera neppure paragonabile al costo che il test di questa ipotesi avrebbe per molti di noi, che pure desiderano che tutto vada bene. L'ipotesi costituisce una minaccia diretta per la vita che vive e la persona che è. Intrattenerla significherebbe cessare di essere ciò che è. In altre parole, il desiderio che non ci sia nulla che non va in lui non può essere funzionalmente equivalente all'apparentemente analogo desiderio di stare bene che può avere anche chi non è affetto da un disturbo paranoide di personalità, con le specificazioni del caso. Per Roberto testare questa ipotesi è intollerabile in una maniera che non è tanto spiegata dal desiderio di stare bene, comune a molti di noi, quanto dal tipo di vita che vive, dal tipo di punto di vista che abita, con i suoi vincoli e le sue possibilità. È questo punto di vista che rende il desiderio di stare bene funzionalmente adatto a produrre l'autoinganno specifico di cui soffre quest'uomo.

Il problema della selettività crea un problema causale per l'autoinganno nella misura in cui assume che il desiderio che p sia vera sia funzionalmente equivalente in tutti i soggetti. Questa assunzione è errata, dal momento che i desideri sono invece funzionalmente differenziati dal particolare ruolo che svolgono all'interno di una psicologia specifica, all'interno di uno specifico punto di vista sul mondo. Attraverso questo supplemento, possiamo cominciare a delineare il set causale dell'autoinganno in maniera tale da non esporlo all'obiezione della selettività.

5. Ulteriori implicazioni della teoria

La teoria causale i cui lineamenti ho qui sopra delineato si rivela utile anche alla risoluzione di una serie di urgenti problemi presenti nel dibattito su altri aspetti dell'autoinganno. Oltre a un problema causale, infatti, un esame attento della discussione fa registrare quello che chiamo il *problema dell'atteggiamento*, parimenti generato dal motivazionalismo senza il supplemento che offro. Il motivazionalismo, infatti, si concentra essenzialmente sui desideri come motore motivazionale dell'autoinganno, ma da tempo viene riconosciuto che altri fattori motivazionali, come le emozioni, possono generarlo. In particolare, la paura genera una forma particolarmente spinosa di autoin-

ganno, noto come *autoinganno invertito*¹³, a cui viene tradizionalmente opposto l'*autoinganno diretto*. Mentre nell'*autoinganno diretto* il soggetto desidera che *p* e finisce per credere che *p*, nell'*autoinganno invertito* il soggetto teme che *p* e finisce per credere che *p*. Le ossessioni e i disturbi d'ansia offrono una vasta gamma di esempi in cui il soggetto finisce per credere quel che teme, e che quindi non desidera, anziché quel che desidera. La sfida del teorico è quindi spiegare in maniera unificata la scaturigine di queste due forme di autoinganno le cui cause motivazionali appaiono radicalmente distinte. Per tentare l'unificazione, il motivazionalismo può postulare la presenza di un perverso desiderio che *p* nonostante la paura che *p*, ma la soluzione sembra *ad hoc* e non in linea con i dati fenomenologici. Illuminando il set causale dell'*autoinganno*, abbiamo invece le risorse per spiegare entrambe queste forme di autoinganno. Se la causa dell'*autoinganno* deve essere ricondotta al punto di vista abitato dal soggetto che si autoinganna, siamo in grado di scoprire perché per un soggetto è importante credere quel che teme, anziché quel che desidera.

Ugualmente, il motivazionalismo non sembra avere adeguate risorse per spiegare il *problema dell'uscita* dall'*autoinganno*. Il problema in questione riguarda le risorse che sono necessarie per redimersi dall'*autoinganno*, dopo esserci caduti. Quali mutamenti sono necessari affinché un soggetto che si autoinganna fuoriesca da uno stato di autoinganno? Se la teoria causale dell'*autoinganno* che sottoscriviamo è il motivazionalismo classico, sembrerà che il cambiamento richiesto al soggetto per uscire dall'*autoinganno* sia il mutamento dei suoi desideri. Ora, però, come abbiamo visto questi desideri sono forgiati dal punto di vista che il soggetto abita, che li rende funzionalmente adatti a provocare autoinganno. Affinché tali desideri si estinguano, o mutino, è necessario che il soggetto riveda il punto di vista che abita. Ad esempio, nel caso analizzato sopra, Roberto dovrà fare assai di più che mutare il suo desiderio che in lui non ci sia niente che non va. Dovrà rivedere perlomeno una parte della sua stessa personalità, se vuole evitare di cadere in autoinganno. Così, facendo luce sul set causale dell'*autoinganno*, possiamo avere più chiaro in che modo il soggetto esce dall'*autoinganno* in cui è incorso. Per uscire dall'*autoinganno*, il soggetto di fatto muta parte del suo punto di vista sul mondo.

Un ulteriore problema che ha speranza di essere risolto attraverso questa teoria è il *problema della valutazione* dell'*autoinganno*. Molte teorie sono in concorso per spiegare quale sia la valutazione più adeguata che dovremmo dare del fenomeno dell'*autoinganno* in generale. Alcuni autori sostengono

¹³ Cfr. Pedrini (2013, cap. 2).

che sia un fenomeno fortemente positivo per il benessere del soggetto, altri che sia dannoso¹⁴. Una teoria che si avvale del concetto di *punto di vista* del soggetto, secondo le linee accennate, è in grado di inquadrare l'intera disputa come una disputa tra i valori che si assegnano a diversi tipi di punti di vista sul mondo, più che al fenomeno dell'autoinganno come tale. Saremo inclini a giudicare l'autoinganno positivamente o negativamente a seconda del tipo di vita da cui tale autoinganno discende, e dal valore che assegniamo a questo tipo di vita. Se abbiamo ragione di valutare negativamente una vita in cui, come effetto collaterale, per così dire, si genera un certo tipo di autoinganno, allora giudicheremo negativamente anche l'autoinganno che ne discende. Se invece il tipo di vita da cui l'autoinganno proviene non incontra la nostra condanna (morale o di altro tipo), anche l'autoinganno che ne scaturisce potrà essere guardato con maggiore simpatia. In altre parole, il nostro atteggiamento valutativo nei confronti dell'autoinganno non è monolitico e risente della valutazione che diamo del quadro causale generale che lo provoca.

Bibliografia

- Bermúdez, L. 2000, "Self-Deception, Intentions and Contradictory Beliefs", *Analysis*, 60, n. 4, pp. 309-319.
- Davidson, D. 1982, "Paradoxes of Irrationality", in *Philosophical Essays on Freud*, ed. by R. Wollheim & J. Hopkins, Cambridge, Cambridge University Press, pp. 289-305.
- Davidson, D. 1985, "Deception and Division", in *Actions and Events*, ed. by E. LePore & B. P. McLaughlin, Oxford, Basil Blackwell, pp. 138-148.
- Funkhouser, E. 2005, "Do the Self-Deceived Get What They Want?", in *Pacific Philosophical Quarterly*, 86, n. 3, pp. 295-312.
- McDowell, J. 1994, *Mind and World*, Cambridge (MA), Harvard University Press (*Mente e mondo*, trad. it. di C. Nizzo, Torino, Einaudi, 1999).
- Mele, A. 1997, "Understanding and Explaining Real Self-Deception", in *Behavioral and Brain Sciences*, 20, pp. 127-134.

¹⁴ Cfr. Pedrini (2013, cap. 5).

- Mele, A. 2001, *Self-Deception Unmasked*, Princeton, Princeton University Press.
- Pears, D. 1984, *Motivated Irrationality*, New York, Oxford University Press.
- Pedrini, P. 2010, "Cognition and Desires: How to Solve the 'Selectivity Problem' for Self-Deception", in *Pratiche della cognizione, Proceedings of the AISC2010*, Italian Society for Cognitive Sciences, pp. 131-135.
- Pedrini, P. 2013, *L'autoinganno. Che cos'è e come funziona*, Roma-Bari, Laterza.
- Pedrini, P. 2014, "Autoinganno selettivo", in *Sistemi intelligenti*, XVII, n. 2, pp. 247-268.
- Scanlon, M. 1998, *What We Owe to Each Other*, Cambridge (MA), Harvard University Press.
- Scott-Kakures, D. 2002, "At 'Permanent Risk': Reasoning and Self-Knowledge in Self-Deception", in *Philosophy and Phenomenological Research*, vol. LXV, n. 3, pp. 576- 603.
- Trope, Y. & Liberman, A. 1996, Social Hypothesis Testing: Cognitive and Motivational Mechanisms, in *Social Psychology, Handbook of Basic Principles*, ed. by E. Higgins & A. Kruglanski, , New York, Guilford Press, pp. 239-270.
- Wells, A. 1997, *Cognitive Therapy of Anxiety Disorders*, Chichester, John Wiley & Sons Ltd.

Razionalità ed emotività

Il ruolo svolto dalle emozioni in epistemologia

Vera Tripodi
Università degli studi di Torino, LabOnt
vera.tripodi@unito.it

«Abbiamo, dunque, innata in noi una così grande passione
per la conoscenza e per il sapere»
Cicerone, *Il sommo bene e il sommo male*, V, 18

«La ragione è, e può solo essere, schiava delle passioni
e non può rivendicare in nessun caso una funzione diversa
da quella di servire e obbedire a esse»
David Hume, *Trattato sulla natura umana*, II, 3, 3

«ma voi con legami strettissimi dovete
legarmi, perché io resti fermo, in piedi sulla
scarpa dell'albero: a questo le corde mi
attacchino. E se vi pregassi, se v'ordinassi di
sciogliermi, voi con nodi più numerosi stringetemi»
(Omero, *Odissea*, XII, 160-4)

1. Prologo

È comune pensare le emozioni come opposte alla ragione e filosoficamente concepirle in antitesi al ragionamento logico. Secondo la nota definizione aristotelica (*Politica*, I, 5, 1254b), l'uomo è per natura un animale razionale. A differenza degli altri animali, la sua peculiarità è quella di essere dotato di capacità logico-matematiche. Da Aristotele in poi, la tradizione filosofica occidentale ha per lo più concepito la razionalità come la capacità di usare la ragione, intesa come la facoltà di calcolare e di argomentare correttamente, ossia di produrre inferenze logiche valide¹. All'interno di questa lunga tradizione, la razionalità è stata intesa anche come la capacità di essere in grado di

¹ Cfr. Vassallo (2009: 162-178).

individuare gli strumenti necessari per raggiungere un dato scopo, condurre indagini, risolvere problemi e decidere come agire in una determinata situazione. La gran parte della cultura filosofica occidentale considera infatti la ragione come una facoltà indispensabile per valutare, criticare e acquisire conoscenza². All'interno di questa prospettiva, con alcune eccezioni³, la ragione è stata spesso associata all'universale, al culturale, al pubblico. Diversamente, tutto ciò che è considerato in opposizione alla ragione – vale a dire la sfera dell'irrazionale, del mondo fisico, della realtà naturale, il particolare, il privato – è stato culturalmente collegato all'ambito delle emozioni⁴. Passioni ed emozioni spesso ci conducono – molti sostengono – a compiere errori di valutazione. Pertanto, per evitare forme di autoinganno e di scivolare in posizioni erranee, passioni ed emozioni dovrebbero – si ritiene – essere in qualche modo moderate o *temperate*. Più precisamente, passioni, sentimenti ed emozioni dovrebbero essere guidate (talvolta anche dominate) dalla ragione non solo perché si possa dare davvero conoscenza, ma anche per raggiungere una comprensione profonda di noi stessi, degli altri e della realtà che ci circonda.

È altresì comune, in filosofia e nella nostra cultura ordinaria, associare l'irrazionalità al femminile e conseguentemente alla donna. La razionalità, vale a dire, è in genere considerata come una caratteristica maschile⁵. Tuttavia, come è stato osservato⁶, la tesi aristotelica che vede la razionalità come puramente maschile ha conseguenze alquanto discutibili. Se in termini aristotelici ciò che contraddistingue gli esseri umani dagli altri animali è la loro natura razionale di cui le donne sarebbero però prive, segue che quest'ultime non dovrebbero essere considerate in senso proprio esseri umani non essendo dotate di una facoltà *specificatamente* umana. Sebbene non ci siano dati scientifici a sostegno della tesi che le donne siano irrazionali o parzialmente razionali, il *cliché* vuole che le donne siano – proprio perché dotate di una razionalità “dimezzata” – meno portate per lo studio di certe discipline quali

² Lakoff & Johnson (1999: 3). Per un punto di vista femminista sulla storia della filosofia come essenzialmente razionalista, si veda: Lloyd (1984); Lloyd (2000a: 245-64); Lloyd (2002).

³ Tra i filosofi non in linea con questa tradizione Susan Bordo inserisce Hume, Nietzsche, Dewey e James. Cfr. Bordo (1987: 114-118).

⁴ Sul rapporto tra razionalità ed emozioni, si veda anche: Calabi (2002: 51-92).

⁵ «Il maschio è per natura migliore, la femmina peggiore l'uno atto al comando, l'altra ad obbedire», Aristotele (2006: 72, testo originale *Politica*, Libro I, 5, 1254b). Cfr. De Beauvoir (1949, trad. it. 21-22, 27, 37); Haslanger (1993). Per un'analisi dettagliata e analitica dei diversi paradigmi di razionalità, rimando a: Lloyd (1984); cit. Witt, (2006).

⁶ C. Witt (1998).

ad esempio le scienze matematiche⁷. Oltre all'irrazionalità, ciò che caratterizzerebbe la donna sarebbe infatti l'emotività e la passività, qualità per lo più considerate superflue, svantaggiose, deleterie per chi aspira a realizzare un progetto di conoscenza che vuole definirsi scientifico. Alle donne spetterebbe – a torto si è affermato – un altro tipo di conoscenza, quella emotiva. Storicamente, la convinzione che l'irrazionalità fosse una prerogativa femminile ha avuto come esito l'esclusione delle donne dalle pratiche scientifiche e dalla sfera filosofica⁸. Com'è noto, questa esclusione ha spinto molte filosofe femministe a parlare di "mascolinizzazione" della conoscenza⁹ e ad affermare la necessità di mettere in discussione il concetto stesso di ragione così come di altre nozioni fondamentali del canone filosofico¹⁰.

In quest'articolo, l'obiettivo non è quello di mettere in luce come il concetto di ragione e quello di razionalità possano o debbano essere rigettati¹¹ né di argomentare a favore di una prospettiva di genere. Piuttosto è quello di mostrare come, alla luce di alcune obiezioni che l'epistemologia femminista¹² ha sollevato, il nostro concetto di conoscenza e la dicotomia tra emotività e razionalità possano essere in parte riconsiderati. In modo particolare, l'attenzione sarà qui posta sul ruolo attribuibile alle emozioni e su una loro possibile riabilitazione in campo epistemologico¹³. È bene precisare che l'intento di questo articolo non è quello di offrire una ricostruzione cronologica delle diverse teorie o tesi elaborate su razionalità, emotività e irrazionalità. Nei paragrafi successivi, la discussione verterà dunque principalmente sulla questione: davvero le emozioni non hanno alcun peso o valore epistemico nel processo conoscitivo? Come vedremo, alcuni studi sembrano smentire questa tesi.

2. Emozione *versus* ragione

Perché considerare la razionalità umana – anch'essa imperfetta, limitata e fallibile – una facoltà di maggior affidabilità in termini epistemologici rispetto alle emozioni?

⁷ Dweck (2006); Dweck (2008).

⁸ Tuana (1992).

⁹ Fox Keller (1985).

¹⁰ Witt (2006); Lloyd (1984).

¹¹ MacKinnon (1999); Tuana (1989).

¹² Per un approfondimento di cosa si intenda per epistemologia femminista, rimando a: Tanesini (1999); Anderson (2011); Garavaso & Vassallo (2007); Tanasini (2015).

¹³ Molti scienziati e filosofi, non solo l'epistemologia femminista, criticano oggi l'idea della contrapposizione netta tra emozione e ragione. Cfr. Damasio (1994).

A differenza della ragione, le emozioni sono state considerate per lo più come fonte d'incoerenza e di perturbazioni interne, come un ostacolo a una vita pratica ordinata poiché si è ritenuto che abbiano conseguenze talvolta distruttive e ci facciano spesso cadere in tentazioni che paiono essere dannose. Noto è lo stratagemma adottato da Ulisse per difendersi da esse. Ben consapevole della debolezza del proprio desiderio e delle proprie passioni, Ulisse fa ricorso all'espedito di farsi legare dai suoi compagni all'albero della nave per poter udire il canto delle sirene senza tuttavia cadere vittima di quella seducente tentazione. Lo stratagemma di Ulisse ha come scopo di salvaguardare se stesso e i suoi compagni dallo scatenarsi delle sue e delle loro reazioni emotive. A ben guardare, la stessa passione per la conoscenza può nel caso di Ulisse essere intesa come pericolosa. Ulisse non chiede infatti, alcuni ritengono, di farsi legare perché non vuole conoscere quanto le sirene avrebbero da dire, tantomeno perché crede che la conoscenza promessa da queste si sarebbe rilevata poi pericolosa. A differenza dei suoi compagni, come è noto, Ulisse non si tappa le orecchie. Piuttosto, egli non vuole cedere assecondando il proprio *desiderio* di conoscere e la propria *passione* per il sapere. Le sirene cercano infatti di attirare Ulisse e i suoi compagni non semplicemente attraverso la dolcezza del loro canto ma offrendo al re di Itaca e ai suoi uomini una prospettiva di conoscenza. È proclamando «di sapere molte cose» che le sirene speravano che «gli uomini, trascinati dal desiderio di sapere, restavano confitti ai loro scogli».¹⁴

È durante l'età moderna che il contrasto tra emozioni e ragione diventa netto, che a quest'ultima viene conferito un primato e viene riconosciuta come facoltà conoscitiva per eccellenza. La tradizione cartesiana ci ha abituati infatti a vedere la separazione tra sfera emotiva e sfera razionale come basata sulla convinzione che la ragione sia lo strumento migliore per indagare la realtà e su cui – nonostante le sue limitatezze – la nostra fiducia andrebbe riposta poiché meno esposta a possibili contaminazioni quali inclinazioni o passioni individuali. Più precisamente, è proprio a partire da questa tradizione che viene rafforzata l'assimilazione della ragione alla capacità di fare inferenze da premesse valide e all'abilità di eseguire calcoli. L'oggettività e l'universalità della ragione risiederebbe precisamente nel fatto che la validità delle inferenze logiche, che grazie ad essa possiamo produrre, non avrebbe nulla a che fare con le nostre personali attitudini e sarebbe del tutto indipendente dalle nostre preferenze individuali¹⁵.

¹⁴ Cicerone (1992: 309).

¹⁵ Nagel (1968).

L'età moderna ha definito anche il modo in cui comunemente intendiamo oggi le emozioni, ossia come atti involontari o fenomeni accidentali. Secondo questa prospettiva, le emozioni possiedono solo un significato motivazionale o un valore cognitivo di tipo valutativo; sarebbero prive di una realtà oggettiva poiché sono fortemente connesse a quei processi mentali che consentono a un singolo individuo di formarle e svilupparle¹⁶. Esse sono intese dunque come fenomeni individuali e non attributi universali della natura umana in quanto strettamente legate al particolare modo di sentire, pensare o agire del singolo. Non esiste infatti una singola esperienza o azione umana *affettivamente* neutra. Le emozioni sarebbero dunque stati interni di cui facciamo esperienza e che “proviamo” senza alcuna mediazione della ragione. Le emozioni avrebbero inoltre una struttura temporale, modale e individuale specifica: esse si riferiscono alle nostre esperienze presenti, non a quelle future o passate; hanno a che fare con l'esistenza affettiva di un preciso momento, non con quella che potremmo o che avremmo dovuto avere. Siamo soliti, infatti, pensare a certi stati d'animo (quali rabbia, ansia, vergogna, gioia) come a qualcosa che percepiamo in un determinato momento o a qualcosa da cui veniamo schiacciati. Non scegliamo di provare certi stati emotivi – comunemente si ritiene – ne veniamo semplicemente sopraffatti. Le emozioni possono – spesso affermiamo – “toglierci il fiato”, esaltarci, farci perdere lucidità e talvolta anche paralizzarci.

È opportuno qui precisare che, sebbene molto diffusa, l'idea che le emozioni possano e debbano essere eliminate del tutto dalle nostre pratiche conoscitive non è stata (e non è) unanimemente condivisa in filosofia. Nella Grecia antica, ad esempio, la distinzione tra emozioni e ragione non era espressa sempre in maniera così netta e contrapposta. Nel mito del carro e dell'auriga raccontato nel *Fedro*¹⁷, Platone rappresenta l'anima razionale come un auriga e le passioni come due cavalli, uno bianco (che personifica la parte dell'anima dotata di sentimenti spirituali che si dirige verso l'iperuranio) e l'altro nero (che invece raffigura l'anima concupiscibile diretta verso la realtà sensibile). Come l'auriga deve guidare il carro e decidere dove questo deve andare, così le passioni (i sentimenti e le emozioni) devono essere guidate dalla ragione che le dirige verso la giusta direzione. Tuttavia, come il mito platonico evidenzia, il carro sarebbe fermo e non andrebbe da nessuna parte senza i cavalli. Analogamente, senza passioni, la ragione non potrebbe muoversi in maniera autonoma. Secondo il mito platonico, il compito della parte razionale dell'anima umana (dell'auriga) sarebbe dunque quello di fare da guida alle

¹⁶ Cfr. D'urso & Trentin (1998); Davidson (1992).

¹⁷ Platone, *Fedro*, (246 a-d), pp. 64-71.

altre sue due parti (i cavalli). Le passioni non devono, dunque, essere completamente represses o estirpate. Piuttosto, esse devono essere condotte e accompagnate nei loro rispettivi percorsi dalla parte razionale dell'anima.

Il rifiuto di pensare il rapporto tra sfera razionale e sfera emotiva in termini di un netto contrasto verrà poi accolto e radicalizzato – sebbene in una prospettiva completamente diversa – per esempio da David Hume, secondo il quale «non parliamo né con rigore né filosoficamente quando parliamo di lotta tra la passione e la ragione»¹⁸. Nella prospettiva humeana, le emozioni vengono dunque riabilitate (diventando il fondamento della etica) e la ragione deve mettersi al servizio di desideri e passioni se vuole raggiungere ciò cui aspira. Dall'altra parte, anche il senso comune è disposto a riconoscere una certa utilità alle emozioni nell'agire quotidiano. Ad esempio, possiamo considerare irrazionale e imprudente in una situazione minacciosa non provare un sentimento di paura o di spavento che potrebbe rilevarsi cruciale per la nostra incolumità. Dunque, come questo esempio evidenzia, le emozioni possono rilevarsi vantaggiose e favorevoli in certi casi. Torneremo su questo punto più avanti.

Riassumendo quanto detto fin qui, data la loro natura, le emozioni sono state per lo più ignorate o bollate come irrilevanti in epistemologia. Tuttavia, come vedremo di seguito, l'interesse delle emozioni non può essere ridotto semplicemente all'aspetto volitivo e motivazionale. Prima di esaminare però le ragioni del perché le emozioni non dovrebbero essere epistemologicamente sottovalutate, cercheremo di analizzare nel paragrafo seguente se vi siano delle differenze tra le emozioni e altri stati d'animo che siamo soliti considerare dello stesso tipo o associare a esse. Abbiamo finora parlato di emozioni, sentimenti, passioni come se fossero sinonimi e si trattasse di fenomeni dello stesso tipo. Eppure, in molti hanno espresso perplessità rispetto all'idea di considerare le emozioni alla stregua di altri stati d'animo. Come vedremo, si è da tempo messo in discussione l'idea che le emozioni, al pari delle sensazioni, si diano per caso o non possano essere frutto di una scelta individuale o sociale.

3. Che cosa sono le emozioni?

Che cosa dunque sono propriamente le emozioni? Per rispondere a questa domanda proviamo innanzitutto a vedere più da vicino il modo ordinario in cui parliamo delle nostre emozioni e valutiamo l'adeguatezza dei modi utiliz-

¹⁸ Hume (1739-40, trad. it. II, 3, 3).

zati per esprimere le nostre esperienze emotive. Quando cerchiamo di rispondere alla questione posta sopra è difficile non imbattersi in una serie di difficoltà, una delle quali è rappresentata dal nostro disaccordo. Vale a dire, quando parliamo di emozioni, una delle maggiori difficoltà con cui dobbiamo confrontarci sta proprio nella varietà e incoerenza dei modi in cui ne parliamo, le pensiamo o concepiamo. Parliamo delle nostre emozioni in modi così diversi che, se considerassimo esclusivamente il modo in cui lo facciamo, non sarebbe possibile offrire una trattazione sistematica. Altro inconveniente, quando si ha a che fare con le emozioni e si cerca di trattarle *rigorosamente*, è legato al fatto che con il termine emozione intendiamo riferirci a fenomeni mentali o fisici spesso molto distanti tra loro. Siamo disposti ad affermare che le emozioni hanno a che fare con il modo in cui entriamo in contatto con la realtà e la percepiamo ma anche con la maniera in cui riconosciamo il valore (maggiore o minore) di una determinata cosa o situazione. Altresì, nel linguaggio ordinario, l'espressione emozione è impiegata per riferirsi al *provare* dolore (valore minore) o al *sentire* piacere (valore maggiore). Come si è accennato sopra, il termine "emozione" è utilizzato nel linguaggio ordinario come sinonimo di "sentimento", "sensazione", "passione", "volontà". Per questa ragione, sembra dunque arduo poter fornire una definizione precisa di cosa propriamente intendiamo per emozione che possa risultare poi valida e applicabile alla diversità dei casi e dei modi in cui di fatto siamo disposti a *usare* quest'espressione. Consapevoli di questa complessità, si potrebbe chiedere se vi sia comunque un criterio più affidabile di altri che possa aiutarci a fornire quantomeno una definizione¹⁹ anche approssimativa, se non di tutti i nostri usi, almeno di cosa un'emozione sia o per definizione dovrebbe essere. Tuttavia, un altro ostacolo si presenta al nostro tentativo di trattare sistematicamente le emozioni. Questo risiede nel fatto che la nozione stessa di emozione varia non solo da un individuo all'altro, ma anche da un contesto sociale a un altro, da una cultura a un'altra. Ciò ha spinto alcuni a considerare il concetto stesso di emozione, al pari di quello di "mente" o "intelligenza", un'invenzione storica²⁰.

Tuttavia, le difficoltà non sono finite qui. In filosofia non vi è neppure un grande accordo sulla natura estremamente varia e complessa delle emozioni; e sulla distinzione ontologica tra emozione e sentimento²¹. Alcuni ritengono

¹⁹ Non si passerà qui in rassegna la molteplicità delle definizioni proposte. Per un quadro generale rimando a: Scherer (2005: 693-727); Frijda (1993); Frijda (1986); Ekman & Davidson (1994: 25-31); Rorty (1980); Moravcsik (1982: 207-224).

²⁰ Jaggard (1989: 147); Rorty (1979).

²¹ Cfr. Gozzano (1999: 13-4).

che le emozioni siano stati affettivi caratterizzati da una breve durata, determinati da una causa interna o esterna, la cui funzione è quella di orientare o ri-orientare la nostra attenzione. Altro aspetto peculiare delle emozioni sarebbe quello di essere stati affettivi che interrompono i nostri processi di pensiero. Diversamente, alcuni sostengono, il sentimento può essere inteso come la fonte delle emozioni. A differenza di quest'ultime, pur avendo un oggetto e un carattere disposizionale, il sentimento non avrebbe una vera e propria causa: la disposizionalità di cui godono i sentimenti, vale a dire, non è transitoria come quella delle emozioni o passioni. Secondo questa prospettiva, i sentimenti sarebbero infatti caratterizzati dalla difficoltà di capire e individuare un loro preciso inizio. Alcuni autori – come William James²² ad esempio – ritengono che l'emozione debba essere distinta dal sentimento (*feeling*), vale a dire debba essere separata da quella sensazione soggettiva qualitativa che accompagna ogni emozione. Per James, l'emozione è un sentimento cosciente, soggettivo e qualitativo. C'è chi poi sottolinea invece che le emozioni non dovrebbero essere spiegate in stretta connessione con le nostre motivazioni e pulsioni.

Nel tentativo di fare un po' di chiarezza sul possibile ruolo svolto dalle emozioni in epistemologia, proviamo a fare allora un breve *excursus* (che non vuole essere esaustivo) su alcune teorie filosofiche delle emozioni.

3.1. Un breve excursus filosofico

Sulla natura delle emozioni e la loro funzione, fin dagli albori, la filosofia ha elaborato diverse e complesse teorie. Anche in psicologia e sociologia²³ la letteratura sulle emozioni – di cui però qui non ci occuperemo – è molto ricca. Per citarne solo alcuni, basti pensare ad autori come Sigmund Freud, “la teoria somatica delle emozioni” di James e di Carl Lange, Pierre Janet e la concezione della emozione come reazione dello scacco, la psicologia della Gestalt e Kurt Lewin. A interessarsi alle emozioni fu anche Charles Darwin, il quale elaborò una teoria evolucionistico-funzionalista secondo cui le emozioni servono per adattarsi all'ambiente²⁴. Limitando qui la nostra attenzione al solo contributo filosofico, è possibile schematicamente individuare (almeno) due diversi filoni²⁵. Il primo è quello che generalmente si fa risalire agli stoici,

²² James (1884: 188-205).

²³ Kemper (1978: 30-41).

²⁴ Darwin (1872).

²⁵ Per una ricostruzione storica della filosofia delle emozioni rimando ad Abbagnano (2002: 34ss).

secondo cui le emozioni sono paragonabili a turbamenti interiori privi di qualsiasi motivazione razionale, il secondo filone invece è quello che attribuisce alle emozioni un valore e un significato oltre che una funzione adattativa. Alla prima concezione, oltre agli stoici, si fanno aderire tradizionalmente filosofi come Spinoza, Leibniz e Hegel: l'idea condivisa da tutti questi autori è che le emozioni debbano essere vinte e dominate. Diversamente, in base alla seconda concezione, le emozioni sono utili poiché permettono agli esseri umani di orientarsi nel proprio ambiente e rispondere *in primis* ai bisogni del corpo. Tra i sostenitori di quest'approccio si inseriscono Aristotele, Hobbes, Pascal, Rousseau, Kant. A metà tra le due prospettive, si pone un filosofo come Descartes che, se da un lato è d'accordo con i stoici sul fatto che non dobbiamo farci sopraffare dalle emozioni, dall'altro riconosce a esse la funzione di contribuire a quelle azioni necessarie che permettono di preservare il corpo.

Nella filosofia contemporanea, su cui vogliamo porre qui principalmente la nostra attenzione, è molto condivisa la convinzione secondo cui le emozioni siano dotate di una razionalità diversa da quella del pensiero razionale. Come si è accennato, molte teorie contemporanee hanno infatti tracciato una netta distinzione tra il tipo di conoscenza prodotta dall'emozione e quella prodotta dal pensiero razionale. Gilbert Ryle²⁶ per esempio elabora una teoria semantica delle emozioni. Secondo Ryle, le emozioni sono azioni passive che possono essere ridotte a riflessi passivi: causate da fattori esterni, la loro funzione non è tanto quella di guidare l'agire umano nel contesto in cui viene a trovarsi, quanto quella di permettere a un individuo di adattarsi automaticamente all'ambiente attraverso schemi d'azione predeterminati. Secondo il positivismo invece, per spiegare cosa siano le emozioni e offrirne una trattazione rigorosa dobbiamo in primo luogo separarle da altre facoltà umane. Più precisamente, le emozioni sono misurabili, prevedibili, non *riguardano* qualcosa, non vertono *su* qualcosa. Piuttosto, vanno intese come fenomeni di disturbo (effettivo o potenziale) di altri fenomeni (quali il giudizio razionale, il pensiero o l'osservazione) che invece vertono su un oggetto. Dunque, le emozioni vanno separate non solo dalla ragione ma anche dalla percezione sensoriale. Secondo la prospettiva positivista, le emozioni sono identificate con le sensazioni fisiche e i movimenti involontari che spesso accompagnano questo tipo di sensazioni o con sensazioni fisiologiche come tristezza, depressione o noia. Pertanto, in un contesto come questo, il termine sensazione è spesso usato in modo del tutto corrispondente ad emozione. Questa distinzione tra emozione e ragione rimanda a quella tra mente e corpo, tra cognizione come facoltà razionale e quegli elementi della vita umana che sono

²⁶ Ryle (2013: 79-111).

meno coscienti, che riguardano il corpo e la sfera affettiva. Le emozioni sarebbero, secondo i positivisti, come la sensazione della stanchezza o gli spasmi di dolore poiché non implicano alcuno stato cognitivo.

La concezione positivista è – sostengono Elizabeth Spelman²⁷ e Alison Jaggar²⁸ – del tutto implausibile. Recentemente, Spelman ha definito tale concezione sulle emozioni *dumb*. Tale approccio può essere sinteticamente riassunto così: le emozioni semplicemente mancano di qualsiasi significato cognitivo e sarebbero dunque letteralmente fenomeni *dumb* (muti, stupidi). Contrariamente a quanto sostenuto dai positivisti, secondo Spelman e Jaggar, le emozioni differiscono da sensazioni, sentimenti e risposte fisiologiche: la loro natura non è episodica, piuttosto è disposizionale. Ad esempio, possiamo affermare con veridicità di sentirci offesi da una certa situazione o fieri di qualcuno oppure rattristati da un fatto anche se in quel preciso momento non siamo né agitati né spaventati tanto meno eccitati. In più, aspetto su cui i positivisti non pongono l'attenzione, le emozioni sono dipendenti dal contesto in cui un soggetto si trova. L'approccio positivista e la teoria *dumb* delle emozioni sono stati criticati e in parte sostituiti da approcci cognitivi e intensionali.

Diverse delle teorie filosofiche più recenti delle emozioni adottano un approccio cognitivista che tenta di elaborare una spiegazione funzionale per rispondere alla domanda “a che cosa servono le emozioni?”. Molti filosofi cognitivisti considerano le emozioni come dotate di due componenti: una affettiva e una cognitiva. La funzione dell'aspetto cognitivo delle emozioni sarebbe quello di interpretare o identificare l'emozione in questione. Più specificatamente, secondo la prospettiva cognitivista, le emozioni non sono più intese semplicemente come sentimenti soggettivi, ma piuttosto come processi di valutazione degli stimoli e di elaborazione delle informazioni. O meglio, i processi cognitivi alla base delle emozioni sono considerati fondamentali nei processi che consentono lo sviluppo e il manifestarsi di un'emozione. Anche quest'approccio è stato fortemente criticato perché, si ritiene, eccessivamente razionalista. È infatti inapplicabile a (e non rende conto di) una serie di fenomeni che inseriamo tra le emozioni, come ad esempio stati d'animo quali ansia, contentezza, angoscia, estasi o terrore.

L'aspetto interessante dell'approccio cognitivista nel dibattito filosofico in relazione alla questione che stiamo qui trattando, ossia la rilevanza o meno delle emozioni in epistemologia, è al ruolo riconosciuto ai processi di apprendimento in grado di influenzare la nostra esperienza emotiva individuale.

²⁷ Spelman (1989: 263-73).

²⁸ Jaggar (1989: 145-171).

Vale a dire, secondo i cognitivisti, ogni individuo sviluppa e matura le proprie emozioni: ciascuno può incrementare il dominio e la varietà delle proprie risposte emotive in relazione alla propria esperienza di vita e di quella degli altri. Tuttavia, l'approccio cognitivista finisce per rinforzare la distinzione che i positivisti facevano valere tra una realtà oggettiva, pubblica, condivisibile che può essere sottoposta alla verifica, al calcolo, all'osservazione e la sfera individuale dei sentimenti e delle emozioni che invece è privata, non condivisibile, soggettiva. Così anche l'approccio cognitivista finisce per rompere qualsiasi nesso concettuale tra le nostre emozioni e il mondo esterno: le nostre emozioni restano cieche, rozze e indifferenziate. Né tantomeno viene sufficientemente spiegato la relazione tra l'aspetto cognitivo e quello affettivo delle emozioni. Dando comunque una priorità all'aspetto cognitivo rispetto a quello affettivo, l'approccio cognitivista ribadisce e rinforza dunque il primato che da un punto di vista epistemologico dovremmo concedere alla mente sul corpo. Rispetto alla teoria *dumb* però, i cognitivisti considerano un aspetto rilevante delle emozioni, vale a dire l'intenzionalità.

Diversamente, all'interno della cornice del pensiero femminista, Jaggar ha posto l'attenzione su alcuni aspetti generalmente poco considerati delle emozioni in epistemologia. L'esame di questa spiegazione sarà l'oggetto dei paragrafi successivi. Come vedremo, è possibile attribuire un ruolo primario all'aspetto relazionale e comunicativo delle emozioni, alla loro variabilità culturale e dipendenza dal contesto sociale. Si esaminerà dunque in che senso le emozioni siano in parte socialmente costruite e spiegabili come fenomeni sociali senza la necessità di riferirsi ad aspetti biologico-genetici o istintivi. Si mostrerà, vale a dire, come le emozioni abbiano un significato e valore solo nell'ambito di un contesto socio-culturale.

4. Emozioni come costrutti sociali

Come si è visto, è comune considerare le emozioni come risposta irriducibilmente soggettiva e privata a una data situazione. Che cosa si intende propriamente per "soggettiva" e "privata"? Si intende banalmente che le mie emozioni sono soggettivamente percepibili in maniera diretta e immediata solo da me e non da altri soggetti: un'emozione è mia e non è di nessun'altro; gli altri non possono percepirle come se fossero loro stessi i soggetti di quell'esperienza sensibile. Le emozioni dunque sono private nel senso che non possono essere sottoposte a una comparazione fra soggetti: non siamo in grado di stabilire se una mia emozione sia uguale o simile a quella di un altro.

L'aspetto individuale delle emozioni sarebbe dunque alla base dal loro carattere pre-sociale, istintuale, innato, primitivo, corporeo e biologicamente determinato.

Questa immagine delle emozioni è stata criticata e considerata da Jaggar fuorviante. Le emozioni non sono irriducibilmente spontanee, fisiche, istintuali o innate. Piuttosto, le emozioni sono in parte socialmente costruite²⁹. Vale a dire, le emozioni possono essere intese come costrutti sociali e hanno a che fare con il modo in cui veniamo educati e cresciuti in uno specifico contesto sociale. Siamo infatti educati a rispondere *emotivamente* nel modo in cui *culturalmente* ci viene indicato come il più appropriato. A un livello inconscio, fin da bambini apprendono ciò che la nostra cultura di appartenenza riconosce come il modo adeguato (e quello meno adeguato) di esprimere emozioni e sentimenti. Ad esempio, pensiamo alla paura per gli stranieri, al piacere per un certo tipo di cibo o a quello di nuotare nelle acque fredde³⁰.

L'idea che le emozioni avessero un valore era emersa già, negli anni Sessanta, con gli studi di Paul Ekman e Wallace V. Friesen³¹ i quali mostrarono che culture diverse esteriorizzano le emozioni diversamente e che a determinare queste differenze fosse il contesto sociale di riferimento. Per spiegare questo fenomeno, Ekman e Friesen utilizzarono il concetto di *display rules*: norme, regole e aspettative sociali fanno sì che gli individui imparino a gestire e modificare l'espressione delle loro risposte emotive in relazione non solo al contesto in cui si trovano, ma anche ai soggetti con cui interagiscono. Un esempio di come le *display rules* agiscono è il seguente. Una persona potrebbe manifestare senza troppe riluttanze, tra le mura domestiche, il proprio risentimento nei confronti della propria compagna o del proprio compagno per qualche cosa che ha fatto. Al contrario, sul luogo di lavoro, la stessa persona potrebbe essere più esitante a fare lo stesso verso il proprio datore di lavoro. La diversità di questo comportamento sta nel fatto che egli (ella) ha imparato delle regole secondo cui non è appropriato esternare un certo tipo di emozioni, specie se aggressive, verso persone con le quali vi è un rapporto di lavoro o di dipendenza. Secondo Ekman e Friesen, sebbene vi sia una componente innata nelle emozioni, regole culturali determinano il modo in cui possiamo modularle e controllarle. Dunque, possiamo esercitare un controllo nella nostra esperienza soggettiva – intensificare, dominare, modificare, an-

²⁹ Jaggar (1989: 150-153).

³⁰ Gli esempi citati sono di Jaggar. Cfr. Jaggar (1989: 150).

³¹ Ekman & Friesen (1969: 49-98).

nullare le nostre emozioni – secondo modalità puramente culturali. A radicalizzare questa posizione sarà Rom Harré³², secondo il quale però le emozioni non sono dotate di una componente biologica o innata né sono fenomeni individuali. Piuttosto, le emozioni hanno una natura intersoggettiva e interpersonale. Inoltre, per Harré, sebbene abbiano una durata relativamente breve, i nostri stati emotivi possono rimanere – attraverso la condivisione della nostra esperienza – nella nostra memoria e in quella di chi ha condiviso con noi quei ricordi emotivi.

Come diverse ricerche antropologiche evidenziano³³, le emozioni sono dispositivi culturali³⁴, hanno una dimensione pubblica e una valenza sociale. Alcune espressioni emotive sono comunicabili e comprensibili solo all'interno di certe comunità sociali³⁵. Ci sono divergenze culturali rispetto al modo in cui esprimiamo o riconosciamo l'adeguatezza o meno di espressioni di disprezzo, dolore, rabbia. O meglio, comunità sociali differenti sviluppano il loro concetto di emozione e la propria teoria delle emozioni. In tal senso, questi concetti e queste teorie sono *locali*. Comunità e culture diverse elaborano diverse concezioni di cosa le emozioni siano e debbano essere, sulla distinzione tra ciò che è pubblico e privato, ciò che può essere esternato in pubblico e non deve essere pubblicamente condiviso. Da ciò segue che l'esistenza e l'esperienza stessa almeno di un certo tipo di emozioni sia determinata da una serie di complesse precondizioni linguistiche e sociali.

Se adottiamo dunque la concezione delle emozioni come in parte socialmente costruite, quale sarebbe propriamente la loro struttura? Le emozioni includono, secondo Jaggar, necessariamente un sistema di valori³⁶. Se così, allora le emozioni presuppongono anche concetti, vale a dire ciò attraverso cui socialmente costruiamo e diamo senso alla realtà intorno a noi. Pertanto, le emozioni sono allo stesso tempo rese possibili (e delimitate) dalle risorse linguistiche e concettuali di una comunità d'appartenenza. Le emozioni di cui facciamo esperienza, dunque, riflettono le forme e le norme sociali in cui siamo immersi. Sebbene ci siano naturalmente somiglianze nel modo in cui culture diverse identificano una stessa emozione, come si accennava sopra, molti studi mettono in evidenza come ad esempio rabbia e amore possano variare da un contesto culturale a un altro. Prendiamo in considerazione il caso dell'amore romantico. Tale concetto, "inventato" nel medioevo e originariamente pensato come confinato alla classe nobile, ha subito radicali e

³² Harré (1986).

³³ Rosaldo (1984: 137-157).

³⁴ Lutz & Abu-Lughod (1990).

³⁵ Rosaldo (1980).

³⁶ Jaggar (1989: 153).

continue modifiche che hanno condizionato il nostro modo di sentirci oggi *romanticamente* innamorati. Un altro caso interessante è quello della fedeltà: ci si sente infedeli o ci si comporta da infedeli sono in un contesto in cui si condividono (o meno) delle norme sociali sulla infedeltà. C'è un senso in cui il nostro sentirci fedeli o infedeli dipende dunque dal fatto che ci troviamo all'interno di un gruppo sociale capace di farci sentire fedeli o meno. Ciò significa che l'esperienza emotiva di questo tipo per un soggetto è allo stesso tempo un'esperienza individuale e sociale.

Questa visione non sembra però, si potrebbe osservare, corrispondere al modo in cui spesso interpretiamo o parliamo delle nostre emozioni, vale a dire come qualcosa da cui siamo sopraffatti o come risposte che non scegliamo di avere in maniera cosciente. Tuttavia, da una riflessione più attenta, emerge che – anche nella nostra vita ordinaria – molte delle nostre emozioni non sono del tutto atti involontari. A ben guardare, infatti, riusciamo spesso a esercitare su di loro un controllo apprendendo e mettendo in atto tecniche di comportamento che ci permettano di “tenerle a bada” oppure di reagire emotivamente in maniera diversa in situazioni analoghe che si presenteranno in futuro. Pensiamo ad esempio alle tecniche per evitare attacchi di panico.

Riepilogando quando emerso in questo paragrafo, possiamo affermare che le emozioni sono a torto intese come necessariamente risposte passive o involontarie determinate da cause esterne. Al contrario, le emozioni sono modi attraverso i quali entriamo *attivamente* in contatto con il mondo. Sono caratterizzate sia da aspetti psichici sia da aspetti che hanno a che fare con la fisiologia del corpo. Le emozioni sono solo in parte risposte passive giacché possiamo esercitare un controllo su di esse o sopprimerle, presuppongono un linguaggio così come un apparato concettuale e un contesto sociale. Chiarito questo, quali sarebbero dunque le implicazioni epistemologiche della comprensione sociale di un'emozione? Prima di esaminare quest'aspetto consideriamo il nesso tra emozioni e valori.

4.1. Emozioni, valori e osservazione

Nella sua analisi, Jaggar pone l'accento su due aspetti peculiari delle emozioni, il loro nesso con la sfera di valori e con l'osservazione. Procediamo con ordine e vediamo innanzitutto il primo di questi nessi.

Emozioni e valori sono, afferma Jaggar, intimamente correlati³⁷. Il possesso di una sfera di valori è strettamente connesso al possesso o all'espressione di certe attitudini e tendenze emotive. Non distante da quanto sostenuto

³⁷ Jaggar (1989: 153-154).

dall'emotivismo etico, per Jagggar, i valori e giudizi di valore presupporrebbero la capacità di reagire emotivamente in certe circostanze. Se non avessimo delle risposte emotive nei confronti della realtà che ci circonda, non sarebbe possibile giungere a valutare certi stati di cose e attribuire ad alcuni di questi una considerazione maggiore (o minore) rispetto ad altri. La nostra sfera di valori di riferimento presuppone dunque la nostra capacità di avere risposte emotive di un certo tipo. Questo vuol dire che vale anche il contrario? Ossia, le emozioni presuppongono a loro volta giudizi di valore? L'oggetto di un'emozione (un oggetto di paura, di orgoglio, di dolore) è un complesso stato di cose. Per esempio, il mio sentirmi orgogliosa della mia amica Belinda perché ha ottenuto un certo risultato presuppone un giudizio di valore, vale a dire presuppone che io sia in grado di attribuire un valore e un'importanza a ciò che Belinda ha ottenuto. Per sentirmi orgogliosa devo pertanto riconoscere ciò che Belinda ha raggiunto come degno di ammirazione. Emozioni e il modo in cui valutiamo certi eventi o stati di cose sono, per Jagggar, concettualmente e logicamente connesse.

C'è un altro aspetto da evidenziare, afferma Jagggar, rispetto al nesso tra emozioni e valori. Il fatto che percepiamo e valutiamo qualcosa (o il verificarsi di un evento) come ad esempio desiderabile o temibile non implica che quella cosa (o il verificarsi di quell'evento) sia *di fatto* da tutti percepita(o) allo stesso modo. Ciò che noi desideriamo o temiamo potrebbe non essere desiderato o temuto da altri. Allo stesso modo, è legittimo che una persona possa non sentirsi spaventata in una certa circostanza quando invece la maggior parte delle persone lo è. Quando infatti qualcuno esprime un sentimento di paura perché ritiene che una situazione sia pericolosa quando invece per la maggior parte degli altri così non è, quello che accade più comunemente è che si sollevino perplessità sulla validità (ma anche sulla autenticità) dello stato d'animo della persona in questione e avanzano richieste di spiegazione. Lo stesso avviene quando qualcuno non prova paura in un contesto in cui per lo più si ritiene che invece dovrebbe averne. Analogamente, anche questo tipo di risposta emotiva suscita perplessità. Si aggiunga che, in alcuni contesti sociali, la mancanza di paura è considerata come irrazionale o talvolta patologica. Segue da questa breve disamina dunque che ogni nostra emozione presuppone una valutazione da parte nostra di qualche aspetto del contesto o realtà in cui questa si manifesta e viene espressa; e che ogni valutazione di una certa situazione comporta che, chi condivide quella valutazione, condivida spesso anche una certa (simile e prevedibile) risposta emotiva a quella situazione. Dunque, le emozioni non sono – ribadisce Jagggar – semplici risposte istintive a circostanze o eventi. Piuttosto, esse dipendono anche dal

modo in cui percepiamo quelle circostanze o quegli eventi e il modo in cui siamo stati portati a rispondere emotivamente a essi.

Come si è detto, Jaggar stabilisce anche un nesso molto stretto tra osservazione ed emozione³⁸: l'osservazione dirige, modella e definisce l'emozione; l'emozione indirizza, modella e definisce l'osservazione. L'osservazione non è semplicemente un processo passivo attraverso cui si registrano stimoli e informazioni. Ciò che viene selezionato quando osserviamo e il come viene interpretato può essere fortemente influenzato anche dalla nostra propensione ad avere certe emozioni, le quali – come abbiamo visto – sono anche in parte pubblicamente e socialmente condivise o condivisibili. Anche nell'osservazione scientifica, ciò è valutato come evidente o chiaro può essere almeno in parte frutto di una costruzione sociale. Ossia, il ritenere uno stato di cose certo e oggettivo può essere basato su un accordo implicito che consiste in una condivisione di assunzioni che coinvolgono e hanno a che fare anche con quella che è ritenuta la risposta emotiva più appropriata in quel contesto. Dunque, l'attitudine verso un certo tipo di risposta emotiva è coinvolto a un livello molto profondo e talvolta anche inconsciamente in ogni nostra osservazione. Altresì, quest'attitudine è implicitamente coinvolta anche nei modi attraverso cui il contenuto delle nostre osservazioni sono *intersoggettivamente* esaminate.

5. Alcune considerazioni sul ruolo delle emozioni in epistemologia

Come si è visto, una buona parte dell'epistemologia contemporanea tende a considerare con sospetto e ostilità le emozioni. Eppure, il valore intrinsecamente strumentale delle emozioni non può essere certamente negato. Le emozioni hanno un'importanza vitale per la nostra stessa sopravvivenza e per il nostro agire sociale. È il nostro provare un certo tipo di emozione che di fatto ci conduce ad agire nel modo di appropriato, ad avvicinare persone nella maniera corretta, ad evitare di imbatterci in certe situazioni e a preferirne altre, a prenderci cura degli altri o di certe situazioni, a fuggire o a difenderci. Inoltre, sebbene non tutte le emozioni siano piacevoli o vantaggiose, è plausibile che senza emozioni avremmo difficoltà a orientarci e a muoversi come agenti conoscitivi nella realtà intono a noi. Tuttavia, specie nel modo occidentale, abbiamo coltivato con grande convinzione l'idea che si debba esercitare un controllo sulle nostre emozioni o sopprimerle con la conseguenza che molti individui non hanno sviluppato una profonda consapevolezza della

³⁸ Jaggar (1989: 154).

propria emotività, neghino di averne o abbiano difficoltà ad attribuirne ad altri. Questa mancanza di consapevolezza da un lato e la diffusa concezione positivista delle emozioni dall'altro, secondo Jaggar, hanno rinforzato il *mito* di una forma d'indagine scientifica libera da condizionamenti, "fredda" e oggettiva perché appunto scevra della componente emotiva³⁹.

In un approccio come quello positivista, come si è visto, l'influenza delle emozioni è pensata come qualche cosa che distorce l'osservazione scientifica, impedisce l'indagine rigorosa e la conoscenza. Secondo Jaggar, non c'è dubbio che le emozioni che hanno a che fare per esempio con il disprezzo, il disgusto, la vergogna, la repulsione e la paura possono falsare, rallentare o travisare l'esame di certi fenomeni o stati di cose. Molti stati emotivi (come l'innamoramento o la rabbia) ci portano ad esempio a una perdita di lucidità, a essere ignari e inconsapevoli di alcune cose o a "non vederne" altre. Tuttavia, l'epistemologia positivista afferma che il ruolo delle emozioni nei processi di conoscenza sia sempre svantaggioso o che non possa essere talvolta un contributo prezioso. Il ruolo riconosciuto alle emozioni nei processi conoscitivi, in un contesto come quello positivista, è ridotto e limitato alla loro capacità di suggerirci ipotesi sulla realtà. Tuttavia, secondo questa prospettiva, per ottenere o raggiungere conoscenza in merito a uno stato di cose, dobbiamo comunque affidarci ad altro. In epistemologia, può considerarsi lecito non porre limiti rispetto ai metodi per generare ipotesi, le quali devono essere sottoposte al criterio della falsificazione per essere eventualmente respinte e confutate. Quando però si tratta di controllare quelle ipotesi, l'epistemologia pone molte restrizioni. La ragione di ciò sta comprensibilmente nel fatto che uno dei criteri basilari della scienza sia la *riproducibilità e replicabilità* sperimentale, ossia la possibilità che un dato fenomeno possa essere riproposto e osservato più volte; criterio questo ritenuto un strumento affidabile per eliminare gli aspetti emotivi o i preconcetti individuali ascrivibili a chi conduce un'indagine. I risultati (tesi e teorie) a cui la scienza giungerebbe sarebbero così – si presume – "oggettive" precisamente nel senso in cui non sono contaminate da ciò che considerato per definizione soggetto a un sistema di valori ed emozioni individuali. La scienza sarebbe in grado di assicurarci con un certo grado di approssimazione di non essere *pregiudizievole* condizionata.

Tuttavia, come è stato osservato, l'indagine scientifica non sempre è immune da questo tipo di condizionamenti. Si potrebbe sostenere, infatti, che se da un lato sia possibile filtrare e con successo neutralizzare gli elementi emotivi puramente soggettivi e individuali, non sempre è possibile fare altrettanto

³⁹ *Ibidem.*

per quegli aspetti emotivi che hanno fortemente a che fare con la sfera di norme e valori (che, come abbiamo visto, sono fortemente legati alle nostre risposte emotive) socialmente accettati e riconosciuti come validi. Pertanto, anche le indagini scientifiche possono essere influenzate da norme sociali e sfere emotive di valori di riferimento⁴⁰. Per esempio le nostre norme sociali orientano la nostra pratica scientifica nella scelta di quali questioni possono essere riconosciute e considerate degne di un'indagine scientifica, nella selezione di quali siano le ipotesi meritevoli di considerazione o nell'accettazione di alcune conclusioni e soluzioni di problemi. Inoltre, anche la scienza contemporanea offre diversi esempi di come questa influenza possa essere dannosa: il sapere e la conoscenza scientifica non sono, infatti, privi di pregiudizi, preconcetti e stereotipi (consoci e inconsci)⁴¹. Certamente, lo schema di valori condivisi all'interno di una comunità anche scientifica è rivedibile. È chiaro altresì come il disprezzo per le persone appartenenti a etnie diverse o gli stereotipi di genere siano *implicitamente* presenti in molte delle interpretazioni antropologiche e storiche del diciannovesimo secolo. Tuttavia, non sempre è facile avere una piena consapevolezza di come pregiudizi e risposte emotive (intese come socialmente costruite) agiscano in maniera anche inconsapevole e condizionino la nostra indagine scientifica o i nostri processi conoscitivi. Secondo Jaggard, ad esempio, norme sociali gerarchiche rispetto al ruolo socialmente attribuito ai due generi e la tendenza a dominare i soggetti appartenenti a minoranze etniche influenzano anche le nostre scienze politico-economiche. I condizionamenti emotivi-culturali agiscono non solo sulla pratica scientifica ma anche a livello meta-scientifico, ad esempio sul tipo di risposta che siamo disposti a dare a questioni come: cosa è la scienza? Come dovrebbe essere praticata? Cos'è un'indagine scientifica? In cosa consiste esattamente la differenza tra metodi scientifici e non scientifici? È stato inoltre affermato che la moderna concezione della scienza, che identifica la conoscenza con strumento per dominare il mondo naturale, riflette spesso una certa concezione politico-economica, talvolta anche preconcetti razzisti e misogini. Molte teorie femministe hanno sostenuto che l'epistemologia moderna e contemporanea possa essere vista come espressione di risposte emotive e di sensibilità associate a certi elementi considerati *stereotipicamente* tipici

⁴⁰ Per un'analisi di come questo tipo di condizionamento funzioni, per esempio, in biologia rispetto alla questione delle distinzioni e classificazioni sessuali, rimando a Fausto Sterling (2000: 113, 147, 170-190). Secondo Fausto Sterling, la distinzione tra sessi non è puramente biologica ma dipende in gran parte da come scegliamo di formulare convenzionalmente il confine tra maschi e femmine.

⁴¹ Hyde (2005: 581-592); McGlone & Aronson (2006: 486-493); Steele & Aronson (1995: 797-811); Ferguson (1998: 318-374).

dell'universo maschile, quali l'ansia di separazione, l'ossessione del controllo e la paura della contaminazione.

5.1. Il mito dell'oggettività scientifica

Tradizionalmente, gran parte delle teorie femministe sollevano perplessità e si mostrano sospettose nei riguardi di generalizzazioni intorno a certe categorie sociali e agli individui che appartengono ad esse. Per esempio, parlare di esseri umani in maniera indifferenziata è considerato rischioso e forviante. Discutere *in generale* di uomini infatti non evidenzia il fatto che un individuo non sia semplicemente un essere umano. Non si può infatti definire una persona indipendentemente dalla sua appartenenza a un sesso, a una razza o un'etnia, a una cultura o a una classe sociale⁴². La categoria di genere, al pari di quella di razza, è socialmente costruita. Dunque, non si può parlare semplicemente di donne e uomini come soggetti epistemici. Piuttosto di *particolari* donne e di *particolari* uomini con tutte le loro specifiche differenze e similarità. Il motivo di ciò sta nel fatto che – come è stato mostrato – razza, genere, sesso e classe sociale plasmano e modellano le nostre vite sociali, la nostra esperienza personale e la nostra sfera emotiva. Una delle critiche femministe mosse al pensiero occidentale è quella di aver negato o di non aver dato la giusta considerazione a queste diversità.

Storicamente, la ragione è sempre stata associata a gruppi di potere, gruppi sociali e culturali dominanti e le emozioni invece relegate alle categorie sociali di minoranza (alcuni gruppi etnici e le donne). Per chiarire ulteriormente questo punto, consideriamo il caso delle donne in relazione anche a quanto detto all'inizio di questo saggio. Come si è detto, l'emotività attribuita alle donne è uno stereotipo culturale molto diffuso. Le donne appaiono essere più emotive degli uomini, o vengono percepite e si auto-percepiscono come tali, perché questo è ciò che gli viene richiesto e che ci si aspetta da loro, vale a dire di esprimere le proprie emozioni in maniera più disinvolta e palese. Questi stereotipi continuano a essere molto forti perché vengono alimentati nella vita di tutti i giorni. Socialmente è infatti *mediamente* accettato che una donna pianga in pubblico. Diverso è per un uomo. Conseguentemente, le donne che non esprimono le proprie emozioni in pubblico, in alcuni contesti, non sono considerate propriamente delle donne; al contrario uomini che lo fanno non vengono considerati *davvero* tali perché si distaccano da quello che è lo stereotipo dell'ideale virile. C'è molta letteratura in merito alle conseguenze di questo comportamento: se educati alla repressione delle proprie

⁴² Haslanger (2000: 31-55).

emozioni, si può perdere la capacità di essere empatici nei confronti degli altri, di fare esperienza di certi stati d'animo e di maturare emotivamente, di essere molto rigidi sul piano morale e di diventare insensibili agli altri non riuscendo a immedesimarsi negli altri. Tuttavia, molti studi hanno smentito che ci sia una vera differenza tra uomini e donne in merito a ciò: gli uomini possono essere tanto emotivi quanto lo sono le donne. Un discorso analogo può essere fatto anche per esempio per i meridionali, i neri o latini rispetto al gesticolare mentre si parla, considerata come una loro caratteristica peculiare.

Da un punto di vista epistemologico, il mito dell'indagine scevra da emozioni ha promosso una concezione della giustificazione epistemologica che ha avuto come conseguenza di escludere alcuni soggetti (*in primis* le donne o i membri di minoranze etniche) e di ridurli al silenzio. Viste come soggetti irrazionali perché culturalmente pensate *per natura* più emotive, le donne vengono identificate come soggetti non epistemicici. Più propriamente, le donne (così come gli individui che appartengono ad alcune categorie sociali emarginate) dunque sono oggetto di quella che viene definita *ingiustizia epistemica*⁴³ poiché non viene riconosciuta loro una autorità in campo conoscitivo⁴⁴. Tuttavia, come si è visto, molte delle affermazioni della scienza sono tendenziose perché influenzate da fattori come il genere d'appartenenza. Dunque, nella misura in cui la nostra concezione del genere (ad esempio il fatto di considerare le donne non epistemicamente autorevoli perché più emotive) influenza le relazioni sociali, abbiamo ragione di credere che anche l'indagine teoretica e scientifica non sia immune da questo tipo di condizionamenti.

La nostra conoscenza è, molte filosofe femministe affermano, *situata* e le nostre emozioni non sono né completamente istintive né interamente determinate da un punto di vista biologico. Cosa vuol dire che le emozioni siano socialmente costruite e che la conoscenza sia *situata*⁴⁵? Vuol dire che: i) al pari di altri prodotti sociali o culturali, le emozioni sono prodotti storici legati al tempo e alla società in cui si trovano immersi; ii) la comprensione che i soggetti hanno rispetto allo stesso stato di cose o fenomeno può differire perché questa riflette sempre le distinte relazioni e posizioni in cui quei soggetti socialmente si trovano e collocano. All'interno di una società gerarchica, le norme e valori tendenzialmente fanno sì che emergano e vengano privilegiate le convinzioni di chi gode di una posizione culturale e sociale più vantaggiosa. Se così, allora possiamo affermare che anche la conoscenza intorno a

⁴³ Fricker (2007).

⁴⁴ Fricker (1998: 159-177).

⁴⁵ Per un approfondimento su cosa si intenda per conoscenza "situata", rimando a Anderson (2011).

uno stato di cose o un fenomeno ritenuta “oggettiva” sia in realtà sempre parziale giacché relativa a un certo modo di costruire *socialmente* la realtà che esclude o marginalizza quello degli appartenenti ad altre categorie o gruppi sociali.

Secondo Jagggar, ogni individuo può essere soggetto a pregiudizi cognitivi inestirpabili e inconsci o a limiti o a imparzialità dovuti al genere di appartenenza o altri tipi di influenze. Questo vuol dire che non è possibile aspirare a una conoscenza oggettiva e neutrale? La risposta a questa domanda di Jagggar è strettamente legata a come si concepisce la conoscenza. In una visione della conoscenza situata e come pratica sociale comune, se le relazioni tra i soggetti ammessi ai processi conoscitivi sono socialmente egualitarie e non discriminanti, allora i pregiudizi individuali o influenze di altri tipo possono essere riconosciuti e corretti anche dagli altri agenti che in egual misura partecipano a quella stessa impresa epistemica. Rispetto alle nostre risposte emotive, la questione della loro appropriatezza o meno dipende allora dal fatto che una comunità conceda a tutti gli individui in maniera indiscriminata di trovarsi in una posizione sociale che permetta di perfezionarsi e migliorare le proprie conoscenze. In questo modo, razionalità e obiettività dunque possono essere espressi da un'intera comunità di agenti epistemici anche quando il processo di elaborazione di un pensiero da parte in un singolo non sia del tutto imparziale, obiettivo e valido.

Si può obiettare che molte delle considerazioni di Jagggar qui presentate siano estremamente vaghe. Naturalmente, la difendibilità di questa proposta sul ruolo attribuibile alle emozioni in campo epistemologico deve essere vagliata. Restano da spiegare molte questioni problematiche come per esempio: come facciamo a valutare quali risposte emotive siano più appropriate di altre? perché dovremmo fidarci della risposta emotiva di alcuni soggetti? Quali tipi di risposta emotiva dovrebbe essere incoraggiata e quale invece respinta? Che ragioni abbiamo per affermare che certi modi alternativi di percepire la realtà, fondati su emozioni socialmente bandite, siano da preferire a quelli invece fondati da emozioni convenzionali? Lo scopo di questo articolo era solo quello di mettere in luce (in maniera non troppo sistematica) alcune caratteristiche delle emozioni su cui occorrerebbe maggiormente riflettere e che andrebbero prese in seria considerazione nel dibattito intorno alla necessità e possibilità di offrire una comprensione più esaustiva di come la conoscenza è, o dovrebbero essere, concepita. Pertanto, piuttosto che vedere emozioni e valori come elementi che devono necessariamente rimossi attraverso un'applicazione rigorosa di un metodo scientifico – come afferma Jagggar – occorrerebbe ripensare la relazione tra conoscenza ed emozioni, ed elaborare modelli concettuali non in termini di una rigida opposizione tra ragione ed

emozione. Quella di una scienza scevra da emozioni è un mito irrealizzabile – come Jaggar sottolinea – in gran parte legato a un certo modo *pregiudizievole* e *stereotipicamente* di vedere il mondo e di pensare la dimensione politico-sociale.

Bibliografia

Abbagnano, N. (2002), “Storia filosofica delle emozioni”, in *Prospettive sulle emozioni e teorie del soggetto*, a cura di D. Galati, Torino, Bollati Boringhieri, pp. 34ss.

Aristotele, (2006), *Politica e Costituzione di Atene*, a cura di C. A. Viano, Torino, UTET Libreria.

Anderson, E. (2011), “Feminist Epistemology and Philosophy of Science”, in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta (ed.), On-line: <http://plato.stanford.edu/entries/feminism-epistemology/>

Bordo, W. R. (1987), *Le Deuxième Sexe*, Albany (NY), SUNY Press.

Calabi, C. (2002), “Che cosa hanno in comune l’amore, il disprezzo e l’assassinio premeditato? Emozioni, basi cognitive e razionalità”, in *Filosofia ed emozioni*, a cura di T. Magri, Milano, Feltrinelli, pp. 51-92.

Cicerone, M. T. (1992), *Il sommo bene e il sommo male*, a cura di F. Demolli, Milano, Bompiani.

Damasio, A. (1994), *Descartes’error*, New York, Putnam (*L’errore di Cartesio*, trad. it., Milano, Adelphi, 1995).

Darwin, C. (1872), *L’espressione delle emozioni nell’uomo e negli altri animali*, con introduzione, postfazione e commenti di Paul Ekman, Torino, Bollati Boringhieri, 1998.

Davidson, D. (1992), *La psicologia come filosofia in Azione ed eventi*, Bologna, il Mulino.

De Beauvoir, S. (1949), *Le Deuxième Sexe (Il secondo sesso)*, trad. it. di M. Adreose e R. Cantini, Milano, Il Saggiatore, 2008).

- D'Urso, V., Trentin, R. (a cura di), (1998), *Introduzione alla psicologia delle emozioni*, Roma-Bari, Laterza.
- Dweck, D. C. (2006), "Is Math a Gift? Beliefs that Put Females at Risk", in *Why Aren't More Women in Science? Top Researchers Debate the Evidence*, di S. Ceci e W. Williams (eds.), Washington, American Psychological Association, 2006.
- Dweck, D. C. (2008), *Mindset: The New Psychology of Success*, New York, Ballantine Books.
- Ekman, P., Friesen, W. V. (1969), "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding", *Semiotica*, I (I), pp. 49-98.
- Ekman, P., Davidson R. J. (a cura di), (1984), *The Nature of Emotion: Fundamental Questions*, New York and Oxford, Oxford University Press.
- Fausto Sterling, A. (2000), *Sexing the body: Gender Politics and the Construction of Sexuality*, New York, Basic Book.
- Ferguson, R. F. (1998), "Can schools narrow the Black-White test-score gap?", in *The Black-White test score gap* di C. Jencks, M. Phillips (editors), Washington (DC), Brookings Institutions Press, pp. 318-374.
- Fox Keller, E. (1985), *Reflections on Gender and Science*, New Haven and London, Yale University Press.
- Freeland, C. (1998), *Feminist Interpretations of Aristotle*, Pennsylvania State, University Press.
- Fricke, M. (2007), *Epistemic Injustice. Power & the Ethics of Knowing*, Oxford, Oxford University Press.
- Fricke, M. (1998), "Rational Authority and Social Power: Towards a Truly Social Epistemology", *Proceedings of the Aristotelian Society*, 98 (2), pp. 159-177.
- Frijda, N. H. (1986), *The emotions*, Cambridge, Cambridge University Press.
- Frijda, N. H. (1993), "Emozioni e sentimenti", *Enciclopedia delle scienze sociali*, Roma, Treccani..

- Galati, D. (a cura di), (2002), *Prospettive sulle emozioni e teorie del soggetto*, Torino, Bollati Boringhieri.
- Garavaso, P., Vassallo, N. (2007), *La filosofia delle donne*, Roma-Bari, Laterza.
- Gozzano, S. (1999), "Ipotesi sulla metafisica delle emozioni", in *Filosofia ed emozioni*, Milano, Feltrinelli, pp. 13-47.
- Harré, R. (ed. by), (1986), *The Social Construction of Emotions*, Oxford, Blackwell (*La costruzione sociale delle emozioni*, trad. it. di O. Pavarini Arcangeli, Milano, Giuffrè, 1992).
- Haslanger, S. (1993), "On Being Objective and Being Objectified" in *A Mind of One's Own*, di L. Antony, C. Witt (ed.), Boulder (CO), Westview.
- Haslanger, S. (2000), "Gender and Race: (What) Are They? (What) Do We Want Them To Be?", *Noûs*, 34 (1), pp. 31-55.
- Hume, D. (1739-1740), *Trattato sulla natura umana*, Roma-Bari, Laterza, 1993, II, 3, 3.
- Hyde, J. S. (2005), 'The Gender Similarities Hypothesis', *American Psychologist*, 60, 6, pp. 581-592
- Jaggar, A. M. (1989), "Love and Knowledge. Emotion in Feminist Epistemology", in A. M. Jaggar, S. Bordo (eds), *Gender/Body/Knowledge: Feminist Reconstructions of Being and Knowing*, New Brunswick, Rutgers University Press.
- James, W. (1884), "What Is an Emotion?", *Mind*, 9, pp. 188-205.
- Kemper, Th. D. (1978), "Toward a Sociology of Emotions: Some Problems and Some Solutions", *The American Sociologist*, 13, pp. 30-41.
- Lakoff, G., Johnson, M. (1999), *Philosophy In The Flesh: the Embodied Mind and its Challenge to Western Thought*, New York, Basic Books.
- Lloyd, G. (1984), *The Man of Reason: "Male" and "Female" in Western Philosophy*, London, Routledge, 1993.

- Lloyd, G. (2000a), "Feminism in History of Philosophy: Appropriating the Past", in *The Cambridge Companion to Feminism in Philosophy*, M. Fricker (ed.), J. Hornsby, Cambridge, Cambridge University Press, pp. 245-64.
- Lloyd, G. (2000b), *Feminism and History of Philosophy*, Oxford, Oxford University Press.
- Lutz, C., Abu-Lughod (eds) (1990), *Language and the politics of emotion*, New York, Cambridge University Press.
- McGlone, M. S., Aronson, J. (2006), "Stereotype threat, identity salience, and spatial reasoning", *Journal of Applied Development Psychology*, 27, pp. 486-493.
- MacKinnon, C. (1999), *Toward a Feminist Theory of the State*, Cambridge (Mass.), Harvard University Press.
- Magri, T. (a cura di) (2002), *Filosofia ed emozioni*, Milano, Feltrinelli.
- Moravcsik, J. M. E. (1982), *Understanding and the Emotions*, *Dialectica*, 36, 2-3: pp. 207-224.
- Nagel, E. (1968), "The Subjective Nature of Social Subject Matter", *Readings in the Philosophy of the Social Sciences*, M. Brodbeck (ed.), New York, Macmillan.
- Platone, (1998), *Fedro*, a cura di G. Reale (a cura di), Milano, Mondadori.
- Rorty, R. (1979), *Philosophy and the Mirror of Nature*, Princeton (NJ), Princeton University Press.
- Rorty, A. O. (1980), *Explaining emotions*, Berkeley, University of California Press.
- Rosaldo, M. Z. (1980), *Knowledge and Passion: Ilongot Notion of Self and Social Life*, Cambridge (UK), Cambridge University Press.
- Rosaldo, M. Z. (1984), "Towards an Anthropology of Self and Feeling", in R. A. Shweder, R. A. LeVine (eds), *Culture Theory: Essays on Mind, Self, and Emotion*, Cambridge (UK), Cambridge University Press, pp. 137-157.

- Ryle, G. (1949), *The Concept of Mind*, London, Hutchinson & Co (*Il concetto di mente*, trad. it. di G. Pellegrino, Roma-Bari, Laterza 2013).
- Scherer, K. R. (2005), "What are emotions? And how can they be measured?", *Social Science Information*, 44(4), pp. 693-727.
- Spelman, E. V. (1989), "Anger and Insubordination", in A. Garry, M. Pearsall (eds), *Women, Knowledge and Reality: Explorations in Feminist*, London, Philosophy, Unwin-Hyman, pp. 263-73).
- Steele, C. M., Aronson, J. (1995), "Stereotype threat and the intellectual test performance of African Americans", *Journal of Personality and Social Psychology*, 69, pp. 797-811.
- Tanesini, A. (1999), *An Introduction to Feminist Epistemologies*, Oxford, Blackwell.
- Tanesini, A. (2015), *Epistemologie e filosofie femministe della scienza*, in APhEx 11. On-line: <http://www.aphex.it/index.php?Temi=557D03012202740321070F04777327>
- Tuana, N. (ed.) (1989), *Feminism and Science*, Bloomington, Indiana University Press.
- Tuana, N. (1992), *Women and the History of Philosophy*, New York, Paragon House.
- Vassallo, N. (2009), "Su naturalismi e filosofie femministe in relazione a cognizione e conoscenza", *Etica & Politica*, XI, 2, pp. 162-178.
- Witt, C. (1996), "How Feminism Is Re-writing the Philosophical Canon", *The Alfred P. Stiernotte Memorial Lecture in Philosophy at Quinnipiac College*, October 2, Printed and distributed by Quinnipiac College On-line: <http://www.uh.edu/~cfreelan/SWIP/Witt.html>
- Witt, C. (1998), "Form, Normativity and Gender in Aristotle: A Feminist Perspective" in C. Freeland, *Feminist Interpretations of Aristotle*, The Pennsylvania State, University Press.

Witt, C. (2000), "Feminist History of Philosophy", *Stanford Encyclopedia of Philosophy* (11/2000) On-line: <http://plato.stanford.edu>; updated 11/2006.

Witt, C. (2006), "Feminist Interpretations of the Philosophical Canon", *Signs: Journal of Women in Culture and Society*, 31, 2.

Alcune riflessioni sul fenomeno dell'*Uncanny Valley*

Adriano Angelucci
Università degli Studi di Urbino Carlo Bo
adriano.angelucci@uniurb.it

Manuel Bastioni
Make Human Project
mbastioni@gmail.com

Pierluigi Graziani
Università degli Studi "G. d'Annunzio" di Chieti-Pescara
pierluigi.graziani@uniurb.it

Maria Grazia Rossi
Università Cattolica di Milano
mariagrazia.rossi@unicatt.it

1. Una strana sensazione

Nel 1906, lo psicologo tedesco Ernst Jentsch utilizzò il termine *unheimlich*¹ in un primo tentativo di definire la strana sensazione di 'innaturale' che viene percepita dagli esseri umani in presenza di situazioni o immagini particolari.

Dodici anni più tardi Sigmund Freud dedicò un lungo saggio² a questa complessa emozione, cercando di identificarne le cause, ma rimanendo comunque ben distante da un completo discernimento e ammettendo di non avere esperienza personale sull'argomento.

¹ Jentsch (1906: 195-198).

² Freud (1953: 219-252).

In questi primi lavori la definizione stessa del soggetto era vaga: si trattava di una inquietante sensazione di stranezza, non necessariamente di repulsione, che si può provare, ad esempio, in occasione di un *dejavu*, o di un attimo in cui la realtà sembra trasfigurare in un sogno.³

La *unheimlich* assunse un significato molto più preciso e tecnico solo nel 1970, quando Masahiro Mori⁴, utilizzò l'espressione "Bukimi no Tani Genshō" (tradotto con *uncanny*⁵, in Inglese) per descrivere la sensazione sgradevole, spesso di rifiuto, che si prova nell'osservare protesi artificiali o più in generale *robot* che imitano gli uomini. Mori osserva che guardando un robot fermo si ha spesso l'impressione di guardare un cadavere⁶, mentre guardandolo in azione, con i movimenti convulsi e irregolari, si prova una sensazione simile a quella che si ottiene osservando qualcuno con gravi problemi motori⁷. Si tratta di un effetto ampiamente utilizzato, in maniera più o meno conscia, dalla cinematografia *horror*, dove conservando la fisionomia base vengono alterati dei dettagli in modo da ottenere quel qualcosa di 'non-normale' in grado di generare nello spettatore la paura-repulsione voluta.

Quello dell'*uncanny* è un problema, dunque, per chiunque tenti di creare interfacce umanoidi perché questo sentimento impedisce la nascita di quell'empatia necessaria per alcune interazioni ottimali tra uomo (e forse animali⁸) e interfaccia umanoide.

Così, questo 'strano effetto', benché conosciuto solo dagli esperti, nei trent'anni successivi allo studio di Mori ha impegnato numerosi studiosi e ha letteralmente plasmato il *design* delle interfacce con cui siamo a contatto nella vita quotidiana: dalle bambole ai grandi film, dai videogiochi ai primi prototipi di androidi. Basti osservare, ad esempio, come le più famose realizzazioni robotiche giapponesi abbiano sempre utilizzato forme molto stilizzate e *hi-tech*, e solo di recente, con gli studi di Hiroshi Ishiguro, si stia tentando la strada del realismo (senza però sfuggire al fenomeno dell'*uncanny*).

La strana sensazione, dunque, sebbene spesso poco nota, condiziona, come vedremo, in maniera più o meno conscia, molti aspetti della nostra vita moderna.

³ Per un'analisi accurata dei testi di Jentsch e Freud vedi Redstone (2013).

⁴ Mori (1970: 33-3).

⁵ Mario Alai ci ha offerto una bella traduzione italiana: "sottilmente e misteriosamente inquietante".

⁶ Una impressione simile può darla anche una statua di cera riproducente un essere umano.

⁷ Sensazione probabilmente legata anche a comportamenti innati che tendono a farci allontanare da situazioni pericolose (vd. infra, § 5).

⁸ Steckenfinger & Ghazanfar (2009: 18362-18466).

2. Valle, muro o crinale?

Sino ad ora abbiamo utilizzato il termine *uncanny*, ma l'espressione con cui, dopo il lavoro di Mori, si fa riferimento al fenomeno è quella di *uncanny valley*. Il termine 'valle' trae origine dal famosissimo grafico tracciato dallo stesso Mori (vd. Fig. 1). Si tratta di una rappresentazione che pone la *familiarity*, con il personaggio osservato, sull'asse delle ordinate e il livello di *human likeness* sull'asse delle ascisse. Osservando la funzione, si vede come nel caso di un personaggio stilizzato la risposta sia quasi lineare: incrementi di *human likeness* fanno aumentare in modo proporzionale anche la *familiarity*. Tuttavia, superato un certo livello di *human likeness*, ecco comparire la 'valle': la percezione della familiarità crolla, il personaggio appare terribilmente fastidioso e poco realistico, l'interazione diventa problematica, e si arriva persino alla totale riluttanza nel guardarlo. Solo arrivando a livelli di *human likeness* altissimi, anche la *familiarity* improvvisamente risale, descrivendo così il secondo, scosceso versante della valle.

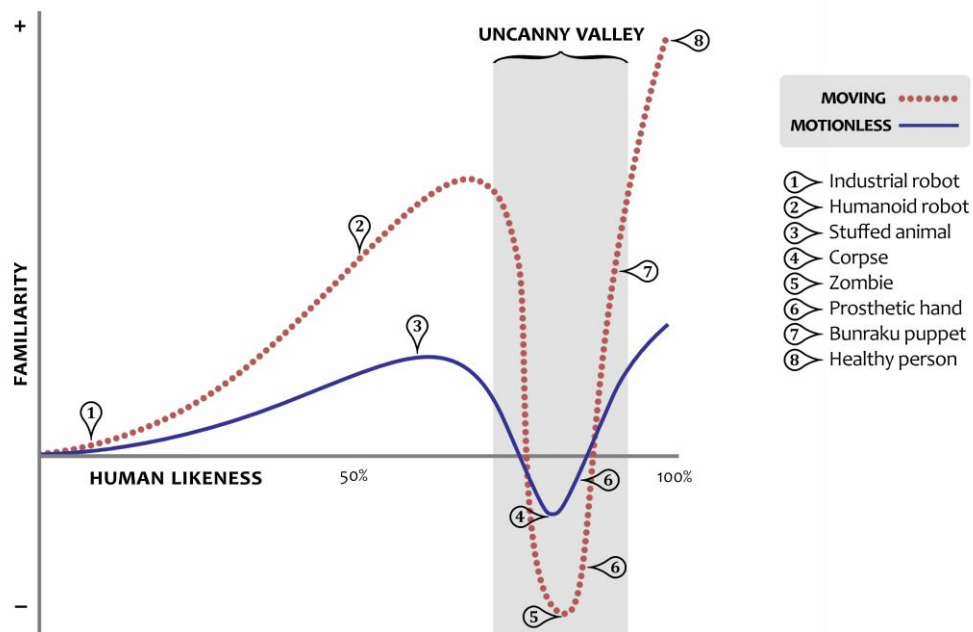


Fig. 1 Grafico che riproduce quello tracciato da Mori (1970/2012).

Il grafico di Mori ci fornisce anche altre informazioni molto preziose. Si vede infatti che le immagini statiche non possono mai raggiungere altissimi

livelli di *familiarity*, ma si vede anche che la loro valle è meno profonda, arrivando, come massimo disgusto, alla sensazione *uncanny* che si ottiene osservando, ad esempio, un cadavere. Al contrario, quando il personaggio si muove, è possibile raggiungere una credibilità molto elevata, ma per le stesse ragioni, la valle è molto più profonda, e giunge al disgusto che si prova nell'osservare, ad esempio, le ultime convulsioni di un corpo morente.

Il grafico di Mori spiega in maniera immediata e intuitiva il fenomeno di cui ci stiamo occupando. Tuttavia va notato che la curva in questione è stata da studi successivi ridisegnata con particolare riferimento alla sua fase di risalita. In particolare Bartneck et al. (2007) e Tinwell et al. (2011a) hanno elaborato due nuovi grafici parlando invece di *uncanny valley* rispettivamente di *uncanny cliff* e di *uncanny wall*. Le tre espressioni si riferiscono a tre diversi andamenti del noto grafico di Mori che ha sulle ascisse la *human likeness* e sulle ordinate la *familiarity*. La *valley*, come abbiamo detto, è l'andamento che si evince nel grafico tracciato da Mori nel 1970, in cui la curva, raggiunta una certa soglia di *human likeness*, decresce come a formare una valle, raggiungendo un punto minimo nella *familiarity* che rappresenta una forte sensazione di *uncanny*, per poi risalire e raggiungere valori di superiori a tutti quelli già raggiunti. Il *cliff*, diversamente, è l'andamento che si ottiene tracciando la curva col metodo di Bartneck ed altri (2007). In questo caso il valore di familiarità massimo si ottiene solo prima della discesa *nell'uncanny*, con valori di *human likeness* non troppo alti. Un robot stilizzato sarebbe quindi sempre più familiare di un robot molto realistico, e la curva risultante somiglia ad una scogliera, un dirupo in cui dopo la caduta c'è una lieve risalita, che tuttavia non arriva mai tanto in alto quanto il punto iniziale. Il *wall*, infine, è l'andamento che si ottiene tracciando il grafico con il metodo di Tinwell ed altri (2011a), in cui viene considerato anche il progresso tecnologico. Con l'introduzione di questa variabile, accade che il livello di *familiarity* del soggetto decresce fortemente nel momento dell'introduzione di nuove tecnologie (e quindi nuove versioni del soggetto stesso). Ad esempio il personaggio di un videogioco che oggi viene comunemente inteso come poco *uncanny*, diverrà irrimediabilmente molto *uncanny* tra qualche anno, quando il pubblico si sarà abituato a nuove tecnologie e quindi a realismi maggiori. Secondo la teoria dell'*uncanny wall* dunque, la curva non è una valle che è possibile risalire, ma un muro forse invalicabile.

Questa situazione, in mancanza di altri dati sperimentali, ci porta a riflettere essenzialmente sulla prima parte della curva, tralasciando al momento il problema del se la seconda parte sia un muro o meno.

3. Perché capire l'*Uncanny*?

La sensazione *dell'uncanny* è collegata, secondo alcuni studiosi⁹, a quell'atavica repulsione dalla morte posseduta da ciascun essere umano e in tal senso essa non è certo un elemento nuovo per gli uomini, eppure solo in epoca recente essa sta prepotentemente attirando l'interesse degli studiosi e coinvolgendo una vasta gamma di discipline, che vanno dall'informatica alla neuroscienza, dalle produzioni cinematografiche alla robotica. Perché?

Appena un secolo fa, termini come 'uomo virtuale', 'androide da compagnia', 'attore digitale', 'avatar', 'realtà sintetiche', avevano ben poco significato per l'uomo comune. Oggi le cose stanno cambiando e lo stanno facendo velocemente. Quotidianamente avvengono miliardi di interazioni tra uomini e macchine: di tipo diretto, come l'utilizzo di uno *smartphone* o la partecipazione ad un videogioco, oppure di tipo indiretto, come guardare un film elaborato al computer oppure un cartellone pubblicitario realizzato con mezzi digitali. La maggior parte di queste interazioni avviene per mezzo di strumenti grafici che cercano di rendere la fruizione piacevole, facendo uso di immagini convincenti, magari personaggi verosimili, in grado di mettere gli utenti a proprio agio, di creare empatia e sicurezza. Ci troviamo, dunque, nella situazione in cui miliardi di persone ogni giorno vengono a contatto con decine di interfacce sottoposte all'*uncanny valley*. Tutto questo fornisce il motivo per cui tale sensazione è oggi ritenuta un fattore determinante per la riuscita o il fallimento di grandi operazioni di massa e persino di tentativi di progresso tecnologico.

Una dimostrazione pratica e statisticamente valida dell'importanza del fenomeno è data da quello che può accadere nelle produzioni cinematografiche che decidono di utilizzare attori virtuali credibili. Ci sono molti altri ambiti in cui l'abbattimento della sensazione *uncanny* si rende assolutamente necessario. La creazione di mondi virtuali, ad esempio, in cui persone tra loro molto distanti possano incontrarsi e interagire in modo realistico è già iniziata e, come spesso accade in questo ambito, l'innovazione è partita dal mondo videoludico. Uffici virtuali, supporto psicologico a distanza, lezioni universitarie in aule virtuali sono altri ambiti. Tutto però è sottoposto al problema *uncanny*: non si deve avere l'impressione di interagire con delle marionette di cera, o peggio ancora, con cadaveri animati. Un'altro campo di applicazione, forse più distante, è la creazione di androidi, macchine umanoidi simili agli uomini che entreranno sempre di più nella vita comune e si suppone che

⁹ Vedi, ad esempio, MacDorman e Ischiguro (2006).

rappresenteranno un aiuto concreto in molti casi “reali”. Un androide può essere, infatti, definito come “an artificial system designed with the ultimate goal of being indistinguishable from humans in its external appearance and behavior”¹⁰. Questo lo differenzia moltissimo dal semplice robot, in cui il raggiungimento della somiglianza umana non è fondamentale. Ed è proprio il raggiungimento di questa empatia il fine ultimo che anima i tanti ricercatori, ponendoli ovviamente in scontro diretto con il problema dell’*uncanny valley*. Sappiamo, infatti, che la realizzazione di un robot che non sembri un ‘cadavere animato’ è ancora molto lontana. Certamente, l’utilizzo di materiali simili alla pelle, di capelli non sintetici e di tutte le altre accortezze ereditate dall’*animatronica* degli anni ‘80 dello scorso secolo, consentono di dare all’automa un aspetto molto realistico, ma solo finché resta immobile. Non appena tenta di muoversi, la mancanza degli oltre 600 muscoli di un essere umano, della loro fluidità e armonia, è immediatamente evidente e l’effetto *uncanny* è peggiore che nella computer grafica e proprio per questa grande difficoltà gli esperti di robotica sono tra i maggiori ricercatori sull’*uncanny*.

Abbiamo già visto come il termine stesso, *uncanny valley*, sia stato coniato dal ricercatore Masahiro Mori inaugurando un intero ambito di ricerca: inizialmente si sono studiate le scienze sociali e cognitive per capire come costruire i *robot*, ora si stanno costruendo *androidi* per capire le scienze cognitive e sociali stesse. Questo, ad esempio, è quanto afferma MacDorman in *The uncanny advantage of using androids in cognitive science research*¹¹:

An experimental apparatus that is indistinguishable from a human being, at least superficially, has the potential to contribute greatly to an understanding of face-to-face interaction in the social, cognitive, and neurosciences. It would be able to elicit the sorts of responses, including nonverbal and subconscious responses, that people typically direct toward each other. Such a device could be a perfect actor in controlled experiments, permitting scientists to vary precisely the parameters under study. The device would also have the advantage of physical presence, which simulated characters lack.

Se così fosse, e utilizziamo il condizionale perché si tratta di ricerche ancora in corso, si innescherebbe un circolo virtuoso: non si tratta infatti di capire soltanto da quali parametri dipenda un effetto di questo tipo, poiché le ricerche in questo campo potrebbero avere ricadute nella comprensione degli umani e della loro natura. La possibilità di alterare singoli parametri in questo

¹⁰ MacDorman (2005: 106-118).

¹¹ MacDorman & Ishiguro (2006: 297-337).

ambito di indagine permetterebbe un controllo sperimentale altrimenti impensabile, quel controllo sperimentale che spesso manca quando si studiano direttamente le interazioni tra gli umani.

4. *Uncanny Valley*: una sfida tecnologica

Abbiamo citato l'importanza del fattore *uncanny* nell'industria dell'intrattenimento. Questi 'esperimenti' sono stati ovviamente osservati anche dalla comunità scientifica, e commentati dagli studiosi dell'*uncanny*, tra cui occorre citare senz'altro Karl MacDorman:

A common feature of Zemeckis' films is a mismatch between the characters' physical appearance and movement owing to the misuse of motion capture technology," MacDorman said. For example, in "A Christmas Carol," we see an old man thrust into the heavens, and yet his movements are that of a young acrobat. This mismatch between appearance and behavior breaks the illusion of being transported into another world. Audiences lose their identification with and empathy for the characters.¹²

Se ne deduce che persino quando il modello è perfettamente realistico, e i movimenti anche, il risultato può essere comunque fastidioso per via di un'associazione errata dei due. I fattori che determinano questa sensazione sono quindi sempre più difficili da individuare man mano che se ne approfondisce lo studio.

La complessità del fenomeno suggerisce che forse, per indagarlo a fondo, la robotica sia qualcosa di insufficiente. Occorrono strumenti più sofisticati per individuare il segreto dell'apparire umano. Un primo aiuto può provenire dalla *computer graphics* che sembra essere il campo migliore per sperimentare e ricercare l'*uncanny*, in quanto permette di simulare umani sintetici che, se fossero costruiti in animatronica (cosa comunque non possibile con i mezzi attuali), richiederebbero grandi investimenti. Basti pensare che il *robotclone* di Hiroshi Ishiguro ha un costo che si aggira, nonostante utilizzi appena 50 attuatori (pistoni idraulici miniaturizzati), intorno al milione di dollari. Un personaggio al computer, invece, può tranquillamente utilizzare centinaia di attuatori, o meglio, dei veri e propri muscoli simulati. Gli umanoidi digitali hanno altri molteplici vantaggi rispetto alla robotica: non hanno nessun problema *hardware*, possono essere facilmente condivisi da ricercatori a grandissime distanze e possono essere migliorati in maniera cumulativa all'infinito (al contrario di un robot che va ricostruito ogni volta) come qualsiasi altro

¹² Wolchover (2011).

software. Sono molto più semplici da interfacciare con sistemi di *motion capture*, e non devono tener conto di problemi reali come la gravità, l'equilibrio ecc.. Inoltre, gli umanoidi digitali sono già correntemente utilizzati nella vita comune, poiché come abbiamo detto, sono largamente impiegati nel campo del cinema, dei videogiochi e della progettazione e quindi c'è un'industria pronta a investire per il loro perfezionamento.

Tuttavia, anche nel campo della computer grafica esistono delle difficoltà e dei problemi, di natura molto differente da quelli posti dalla robotica. Per giungere al massimo realismo esistono sostanzialmente due metodologie. Da un lato c'è chi tenta di simulare tutto con modelli matematici e anatomici sempre più accurati, riproducendo ossa, muscoli, tendini, comportamento dell'epidermide, delle masse molli e così via. Dall'altro c'è chi invece preferisce acquisire tutto questo direttamente da soggetti esistenti, duplicando un umano reale e catturando in un enorme database tutte le sue espressioni, le sue pose i colori della pelle e altro ancora¹³. In teoria, la strada più promettente appare la prima. Se si riuscissero a mettere insieme tutti i modelli matematici che definiscono l'apparenza esteriore di un essere umano, si potrebbe senz'altro giungere ad una simulazione talmente realistica da poter persino evitare l'effetto *uncanny*. Ma i problemi di questo approccio sono a dir poco enormi e probabilmente ancora ben lontani dall'essere risolti.

5. Simulare la realtà

Vorremmo a questo punto provare a fare un esempio utile per comprendere la sfida *dell'uncanny valley*. Pensiamo alla difficoltà di far fare all'umanoide digitale un'azione apparentemente semplice come quella di sorridere.

Nel volto ci sono oltre 30 muscoli, e di questi, almeno una dozzina viene utilizzato dal sorriso. Ogni muscolo si contrae in modo differente, non lineare, e con intensità e direzione propria. Allo stesso tempo però, ogni muscolo trascina con se i muscoli ad esso vicini, alterandone la direzione, subendo un'azione frenante e agendo contemporaneamente anche sulle cartilagini. Infine, tutto il complesso intreccio di movimenti si trasmette alla pelle, ma l'intensità del movimento ottenuto dipende in maniera non trascurabile dalla quantità di grasso che si interpone tra muscolo e pelle.

Mentre tutto questo accade a ridosso delle ossa, in superficie osserviamo un altro incredibile spettacolo di complessità. La pelle, compressa in certi punti e distesa in altri, presenterà anche un colore diverso. In alcune zone la

¹³Bastioni & Graziani (2011: 26-37).

compressione determinerà un maggior irroramento dei vasi sanguigni, dandole un colorito più roseo, mentre in altri, l'allungamento agirà in senso opposto, rendendola più pallida ed evidenziando le vene, nel caso in cui la pelle sia abbastanza sottile. Poi, a seconda dell'elasticità, della percentuale di collagene e di altri fattori, la pelle compressa tende a mostrare più o meno rughe, determinate anche da fattori soggettivi, come l'abitudine del soggetto a fare quel movimento (le cosiddette rughe di espressione), la presenza di cicatrici, l'esposizione al sole, e altro ancora. Al contrario, la pelle tesa tende a mostrarsi più lucida, levigando tutte le asperità e mostrando spesso il volume dell'osso sottostante.

Cosa dire infine del comportamento della luce? Essa non rimbalza semplicemente sulla pelle, ma viene in gran parte assorbita. Percorre qualche millimetro e poi quando giunge ai muscoli, più densi e rossi, si riflette in maniera disordinata uscendo di nuovo dalla pelle e giungendo a noi come e dandoci il senso di vivo e organico, in un effetto definito tecnicamente *subsurface scattering*. Ogni millimetro quadrato di pelle subisce delle variazioni durante le espressioni e quindi cambia in modo radicale il comportamento della luce. Ma la riflessione varia anche con altri fattori: si pensi all'età, alla presenza di barba, alla presenza di trucco, a quella della pioggia o del sudore.

Incredibile, fantastica complessità, catturata dai nostri occhi ogni attimo, in un quadro olistico in cui anche la perdita di uno solo di questi dettagli determina l'effetto *uncanny*. Il nostro cervello 'si attende' che la pelle si colori in quel certo modo, che il sudore scorra su di essa in una certa maniera, fermandosi dove è più ruvida e accelerando dove è più liscia, frenando sulle sopracciglia o sulla barba ecc. Se questo non accade, se anche solo un elemento viene trascurato, il cervello entra in una sorta di modalità interrogativa e determina l'*uncanny*.

È tutta qui la complessità del sorriso? No, c'è altro.

Tutto quello appena descritto riguarda una posa statica, un sorriso immobile. Ma sappiamo che la realtà è ben diversa. La contrazione muscolare cambia di continuo, i muscoli vibrano un poco, l'espressione viva si modifica attimo per attimo, in modo fluido e non meccanico. Ogni istante di esitazione, ogni ritardo o accelerazione dell'attivazione delle fibre si traduce in una sfumatura dell'emozione, qualcosa di difficile persino da descrivere, ma che ci è ben noto nella vita quotidiana, laddove vediamo un 'bel sorriso', 'un sorriso ipocrita', 'un sorriso triste', ecc.

Quanti algoritmi e che incredibile computazione sarà necessaria per simulare tutto questo? Difficile da dire. Al momento le simulazioni si rivolgono soprattutto alla riproduzione del comportamento dei muscoli, che sono senz'altro il *layer* più visibile tra quelli descritti sopra, e anche quello più

semplice. Altri ricercatori stanno lavorando sulla simulazione dell'elasticità della pelle, e altri ancora sul comportamento delle masse morbide. Ogni tanto ci sono dei tentativi di mettere insieme tutte le caratteristiche, ma è chiaro che mancano tanti e tanti dettagli, per cui l'effetto *uncanny* è ancora assolutamente inevitabile, almeno per il volto.

Per quando riguarda quali muscoli contrarre, di quanto, con che velocità e in che sequenza, gli studi sono ancora agli inizi. L'orientamento è quello di considerare la costruzione di espressioni di base utilizzando combinazioni di contrazioni generiche, ma i risultati sono ben distanti dalla naturalezza della realtà, per cui l'industria preferisce affidarsi al *motion capture*, una tecnica in cui un soggetto reale viene cosparso di sensori la cui posizione viene memorizzata durante il movimento, istante dopo istante. Si passa così dalla simulazione alla semplice memorizzazione/riproduzione del comportamento, e questo ci introduce al secondo approccio al problema.

6. Catturare la realtà

Considerata l'incredibile complessità dei fenomeni da simulare, e considerato anche il rapidissimo sviluppo della capacità di archiviazione dei dati nei computer, una soluzione per evitare l'effetto *uncanny* può essere quella di 'catturare' la realtà. In questa prospettiva, tornando al nostro sorriso, anziché utilizzare migliaia di formule e modelli matematici, si può semplicemente prendere un attore e scansionare una sessione dei suoi sorrisi. Lo scanner catturerà, fotogramma per fotogramma, ogni micro espressione, ogni ruga, ogni vibrazione, i dati saranno *gigabyte* di coordinate tridimensionali, ognuna delle quali identificherà la posizione di un punto del volto. Allo stesso modo, un altro scanner speciale catturerà il colore della pelle, e ogni istante il volto dell'attore verrà illuminato in decine di modi differenti, in modo da memorizzare le risposte della pelle alla luce e questa volta il risultato verrà memorizzato in *gigabyte* di immagini. A questo punto, sarà stata effettuata una vera e propria 'clonazione' digitale del soggetto. Il suo modo di sorridere, le sue smorfie e atteggiamenti memorizzati in un enorme archivio di dati. La mossa successiva sarà quella di gestire questa libreria in modo che i vari segmenti possano essere utilizzati e combinati a richiesta, ad esempio durante una conversazione con il personaggio virtuale.

Un primo tentativo di utilizzo di questa tecnica fu fatto da *Image Metrics* e l'*USC Institute for Creative Technologies of University of Southern California*, che nel 2008, con la *computer graphic* Emily¹⁴, realizzò un'impresionante 'intervista' in cui il personaggio digitale stesso spiegava la tecnologia utilizzata per la sua realizzazione. Questo esperimento fu un grande successo e finalmente consentì di far vacillare il regno dell'*uncanny valley*. Inoltre, nella realizzazione di Emily venne utilizzata anche un'altra tecnologia, precedente di qualche anno¹⁵, utile per catturare il comportamento della luce sulla pelle. Si tratta del *Light Stage Systems*, presentato già nel 2000 dal Paul Debevec nell'Università di Berkely e pubblicato al *Siggraph* dello stesso anno, in cui un attore reale entra in una futuristica 'sfera di luce' e viene illuminato in modo da catturare i principali tipi di illuminazione. Anche in questo caso, come nella scansione 3D, l'acquisizione tiene conto di una grande quantità di fattori, difficilissimi da simulare e persino da individuare, ma catturati comunque dalla macchina di Debevec.

Con la realizzazione di Emily, dunque, è stata inaugurata una tecnologia che ha trovato da subito applicazioni pratiche nel cinema, e che sembra stia finalmente superando la valle dell'*uncanny*. Inoltre il passaggio in atto di tale tecnologia dal cinema al mondo dei videogiochi è già stato compiuto e non è stato un passaggio banale come può apparire. Il cinema ha a disposizione tempi molto lunghi per i *rendering*, e una scena di pochi minuti può richiedere persino settimane di calcolo per essere realizzata. Nel mondo dei videogiochi invece, tutto deve essere calcolato in pochi centesimi di secondo, perché il personaggio deve reagire in *real time* agli *input* del giocatore. Altra differenza fondamentale risiede nel fatto che le produzioni cinematografiche hanno a disposizione *hardware* molto costosi, mentre il videogioco deve essere in grado di funzionare su *console* popolari. Si capisce, dunque, l'enorme difficoltà di portare una tecnologia cinematografica in un ambiente videoludico. Eppure sempre l'*USC*, in collaborazione con *Activision* e *Nvidia*, è riuscita a realizzare un prodotto *real time*. Il nuovo progetto si chiama *Digital Ira* e usa lo stesso sistema di Emily, con la differenza che il comportamento della pelle è stato ottimizzato talmente tanto da poter essere inserito all'interno di uno *shader* per le schede grafiche (anche se ancora fuori dalla portata del grande pubblico), per cui è possibile interagire con *Ira* in *real time*.

Possiamo, dunque, almeno per quel che riguarda la sua prospettiva tecnologica, considerare il problema *uncanny* come risolto?

¹⁴Alexander, Rogers, Lambeth, Chiang & Debevec (2009).

¹⁵Ghosh, Fyffe, Tunwattanapong, Busch, Yu & Debevec (2011: 129).

Non ne siamo certi.

Ricordiamo, innanzitutto, il già citato problema dell'ingombro dei dati. Solo a titolo di esempio, basti sapere che la cosiddetta *demo* di *Ira*, un volto con poche variazioni, richiede 32 *gigabyte* di dati, che anche se mirabilmente compressi, grazie ad una nuova tecnologia, in soli 400 MB, rappresentano pur sempre un dato di grandi dimensioni. Un *set* che possa coprire una buona parte delle espressioni umane utilizzate correntemente durante le conversazioni della vita quotidiana richiederà molti giga di dati. Occorre dunque scegliere con grande attenzioni cosa 'catturare', e in particolar modo, quali espressioni utilizzare. In teoria, visto che non si tratta di una simulazione, sarebbe necessario collezionarne centinaia, ma in pratica i ricercatori pensano che sia sufficiente limitarsi solo ad un sottoinsieme, e fanno molto spesso riferimento allo psicologo americano Paul Ekman, che dopo una pluriennale ricerca è riuscito a creare una codifica delle componenti principali¹⁶ classificando una serie di elementi da lui chiamati *facial actions*, che possono essere combinati secondo una sorta di algebra in modo da ottenere tutte le possibili varianti di espressioni facciali. Purtroppo però le cose nella *computer graphics* sono un poco differenti. Le *action* di Ekman funzionano con i volti reali, in cui vengono miscelate con movimenti complessi dei muscoli, molto lontani dalla semplice addizione algebrica. Nei volti digitali, invece, dove la somma è puramente matematica, le combinazioni creano svariati problemi, che richiedono soluzioni complesse non ancora perfezionate. Quindi per essere davvero certi di uscire dall'*uncanny valley*, occorrerebbe un database 'full', che abbia acquisito tutto 'dal vero', con conseguente enorme ingombro di memoria. Ad ogni modo, vista appunto la crescita esponenziale della tecnologia di archiviazione ed elaborazione di dati, questo aspetto può essere definito secondario, come anche l'incredibile complessità delle 8000 operazioni necessarie per il *rendering* di ogni singolo *pixel* per ogni istante del movimento.

Un ulteriore problema della tecnica di cattura, in contrapposizione a quella di simulazione, è la mancanza di flessibilità. Ogni personaggio avrà bisogno del suo database, e i dati validi per un volto non saranno validi per un altro. I dettagli, proprio quelli che fanno la differenza quando si parla di *uncanny*, sono strettamente personali e adatti a ogni specifico volto. Anche la reazione della pelle alla luce è del tutto personale, dipendendo, come abbiamo osservato, da innumerevoli fattori legati alla specifica anatomia. Lo stesso vale per quanto riguarda le rughe, il modo di corrugare la fronte, la gestualità stessa. Volendo semplificare il discorso, potremmo dire che la tecnologia

¹⁶ Ekman & Friesen (1978).

della ‘cattura’ del personaggio altro non è che un nuovo tipo di ripresa televisiva. Laddove la videocamera tradizionale si limitava a riprendere sequenze di immagini bidimensionali, la nuova tecnologia riprende immagini tridimensionali. Laddove la vecchia tecnologia montava le sequenze manualmente, la nuova tecnologia le monta in *real time*, a seconda della reazione da mostrare all’interlocutore. In questo modo *l’uncanny* è superata, ma come comportarsi se si vuol creare un attore che non esiste? Forse si giungerà al modo di miscelare accuratamente riprese di attori base, giocando con l’interpolazione di grandi quantità di dati. Al momento siamo ancora solo dei pionieri ed è quindi difficile poterlo dire, ma certamente possiamo essere ottimisti sul superamento della valle dell’*uncanny*.

7. *Uncanny Valley*: una sfida filosofica

Le differenti ricerche discusse nei paragrafi precedenti condividono il medesimo obiettivo: fronteggiare la sfida tecnologica posta dall’*uncanny valley* ottenendo il massimo del realismo, tanto sul piano dell’aspetto dei corpi, quanto sul piano del loro movimento.

Ora, pur accogliendo questa sfida con motivato ottimismo, la nostra convinzione è che il fenomeno dell’*uncanny valley* offra una più generale sfida teoretica o meglio filosofica. La presenza stessa della valle, infatti, ci pone davanti il problema di capire le ragioni profonde del modo in cui gli esseri umani si riconoscono o non si riconoscono in quanto umani. Ovviamente la sfida tecnologica e filosofica si muovono su due piani differenti, ma pensiamo che le ricerche su tali piani possano proficuamente arricchire entrambi gli ambiti di ricerca proprio nello spirito dell’*android science* proposta da MacDormann e Ischiguro (2006).

Dal punto di vista teoretico, il problema nell’*uncanny valley* ci sembra essere in primo luogo un problema di riconoscimento intimamente connesso a quello della *disponibilità empatica all’interazione*¹⁷. Lo scarto tra la crescita di *human likeness* e la paradossale decrescita di *familiarity* – la valle *uncanny* – si traduce in sensazioni di allarme, di repulsione, di riluttanza, di malessere, di disgusto. Il fatto che queste sensazioni non siano presenti in tutte le situazioni di disponibilità empatiche porta a pensare che nel fenomeno *uncanny* essa sia avvertita come più importante di altre, probabilmente in forza del fatto che nel caso *uncanny* sono coinvolti i nostri processi di attribuzione di umanità agli altri e, dunque, riflessivamente il nostro riconoscimento come esseri umani. Nella disponibilità empatica all’interazione i fattori in gioco

¹⁷ Vedi anche Misselhorn (2009); Redstone (2013).

sono molteplici¹⁸. Con riferimento *all'uncanny*, come evidenziato da MacDorman e collaboratori¹⁹, si possono considerare almeno due ampie categorie di spiegazioni: la prima categoria tenta di ricondurre l'effetto *uncanny* al funzionamento di processi percettivi specializzati (ad esempio quelli del riconoscimento dei volti); la seconda categoria tenta di ricondurre l'effetto *uncanny* al funzionamento di processi cognitivi di più alto livello (ad esempio quelli legati alla costruzione dell'identità personale, messi a dura prova dalla presenza di robot che, grazie al loro funzionamento meccanico, sembrano in grado di sfidare i confini tra umano e non umano mettendo in discussione la nostra unicità di esseri mentali). La quantità di dettagli da prendere in considerazione all'interno di ogni categoria è molto grande. Per fare un esempio, con riferimento alla prima categoria, Tinwell e collaboratori²⁰ hanno mostrato che la configurazione realistica dei movimenti della parte superiore del volto è essenziale per evitare l'effetto *uncanny* quando l'androide simula emozioni di base come la paura o la tristezza, ma non quando simula emozioni come il disgusto o la sorpresa.

La nostra proposta di ricerca, che qui ci limitiamo solo a tratteggiare²¹, intende evidenziare come la dinamica tipica dei processi di *deumanizzazione*²² possa avere un ruolo interessante nella spiegazione del fenomeno dell'*uncanny valley*. In particolare la nostra idea è che le sensazioni legate alla problematicità dell'interazione tipiche dell'*uncanny valley* siano riconducibili all'emozione del *disgusto* (ampiamente intesa) e alle sue dinamiche deumanizzanti, emozione generata da una sorta di corto circuito categoriale riguardante il riconoscimento di esseri conspecifici.

Cerchiamo di tratteggiare meglio la nostra proposta partendo dall'analisi dell'emozione del disgusto. Nel 1987 Paul Rozin e April E. Fallon inaugurano una nuova stagione di ricerca sull'emozione di disgusto che ha portato, a partire dagli anni Novanta del secolo scorso, a una più attenta valutazione di questa emozione e a una maggiore consapevolezza sul ruolo del disgusto nel contesto delle relazioni sociali e morali degli umani. La funzione evolutiva originaria del disgusto è connessa con la salvaguardia del corpo degli organismi: le reazioni di nausea e di repulsione tipiche di questa emozione proteggono gli organismi dall'azione dannosa di agenti patogeni. A questo propo-

¹⁸Misselhorn (2009: 345-359); Tinwell, Grimshaw & Williams (2011: 326-341); Tinwell (2014).

¹⁹MacDorman, Green, Ho, & Koch (2009: 695-710).

²⁰Tinwell, Grimshaw, Nabi & Williams (2011: 741-749).

²¹Nel presente lavoro ci limitiamo solo a una descrizione qualitativa di un progetto di ricerca in fase di sviluppo mirante a testare sperimentalmente tali nostre proposte.

²²Cfr. Haslam (2006).

sito, Rozin e Fallon definiscono il disgusto primario o di base (*core disgust*) nei termini di una «food-rejection emotion»²³. Tuttavia, come notano Rozin e collaboratori, nel corso dell'evoluzione umana il dominio del disgusto ha subito una estensione inglobando il riferimento a minacce di tipo morale e sociale. In un quadro teorico di questo tipo, la reazione di repulsione sembra pertinente non soltanto quando a essere in pericolo è il nostro corpo, ma anche quando a essere in pericolo sono alcune nostre categorizzazioni. Tramite il disgusto rifugiamo spesso da tutto ciò che ci obbligherebbe a contaminare ad esempio la 'purezza' del nostro mondo interiore con 'l'impurità' proveniente dalla nostra natura animale; siamo così portati a considerare repellenti le interiora, gli escrementi, la morte, finanche alcune pratiche sessuali²⁴.

Riconoscendo come propri del disgusto anche i domini estesi, quello sociale e quello morale, la letteratura contemporanea ci costringe a riflettere sul ruolo (più o meno positivo, più o meno razionale) di questa emozione all'interno di una dimensione squisitamente interpersonale. È questa dimensione che consideriamo rilevante ai fini del nostro discorso sull'*uncanny valley*. Più nello specifico la proposta che vogliamo avanzare è che, all'interno della dimensione interpersonale, il disgusto possa agire come una forza che disgrega, stigmatizza, causa pregiudizio, genera sospetto e che, quindi, agisce in direzione opposta rispetto alla relazione di contatto che sta alla base del riconoscimento empatico dell'interazione sociale come interazione tra umani²⁵. L'azione stigmatizzante del disgusto potrebbe così rappresentare un utile elemento teorico per riflettere sulla natura dei *pregiudizi* ampiamente intesi. In effetti, varie tipologie di pregiudizi sociali possono portare a valutazioni e/o a comportamenti di disapprovazione, di condanna, di persecuzione nei confronti di gruppi esterni ed estranei rispetto al gruppo di riferimento. Sulla base di alcuni studi sulla predisposizione a valutare automaticamente in maniera negativa persone disabili, stranieri, gruppi minoritari²⁶, Hodson e Costello hanno indagato più specificamente la relazione tra disgusto e pregiudizio enfatizzando la presenza di una forte correlazione tra la sensibilità al disgusto

²³ Rozin & Fallon (1987: 23-41).

²⁴ Rozin, Lowery, Imada & Haidt (1999: 574-586); Rozin, Haidt & McCauley (2008: 9-29); Rozin, Haidt & Fincher (2009: 1179-1180); Moll, de Oliveira-Sousa, Moll, Ignacio, Bramati & Caparelli (2005: 68-78); Olatunji, Williams, Tolin, Abramowitz, Sawchuk, Lohr, et al. (2007: 281-297); Philips, Young, Senior, Brammer, Andrews, Calder et al. (1997: 495-498).

²⁵ Per posizioni simili cfr. Pizarro, Detweiler-Bedell & Bloom (2006: 81-98); Rossi (2013); Sherman & Haidt (2011: 245-251).

²⁶ Faulkner, Schaller, Park & Duncan (2004: 333-353); Fiske, Cuddy, Glick, & Xu (2002: 878-902); Navarrete & Fessler (2006: 270-282); Park, Faulkner & Schaller (2003: 65-87).

(alta o bassa) dei soggetti e il grado di stigmatizzazione dei gruppi non appartenenti al gruppo di riferimento²⁷. Martha Nussbaum ha, inoltre, insistito a lungo su posizioni simili: occupandosi però più direttamente del ruolo del disgusto all'interno della legislazione giuridica americana²⁸.

Valutati complessivamente questi studi imputano al disgusto un'azione di deumanizzazione che ci appare compatibile con quella presente nel fenomeno *uncanny valley*. La presenza di questa emozione, anche quando è indotta sperimentalmente²⁹, può essere interpretata come una risposta a una minaccia di 'contaminazione' che, di conseguenza, comporta che al gruppo oggetto della stigmatizzazione non venga riconosciuta l'attribuzione di umanità. In questo senso, il riconoscimento e l'attribuzione di umanità vengono continuamente contrattati anche nelle interazioni tra umani e tra gruppi di umani. A dispetto del fatto che queste interazioni siano caratterizzate da un buon grado di realismo, la sensazione di repulsione può dominare e avere il sopravvento; così la disponibilità empatica può essere messa (temporaneamente o talvolta, purtroppo, definitivamente) in discussione attraverso la dinamica tipica dei processi di deumanizzazione.

Usualmente, la nozione di deumanizzazione è utilizzata in riferimento a quella che viene definita «deumanizzazione animale» o a quella che viene definita «deumanizzazione meccanica o meccanicistica»³⁰: nel primo senso, saremmo portati a considerare l'altro con cui interagiamo più come un animale che come un umano; nel secondo senso, saremmo portati a considerare l'altro con cui interagiamo più come una macchina che come un umano. Haslam³¹ collega questi due differenti sensi di deumanizzazione a due concezioni differenti di umanità (vd. Fig. 2).

²⁷ Hodson & Costello (2007: 691-698).

²⁸ Nussbaum (2004); Nussbaum (2010).

²⁹ Buckels & Trapnel (2013: 771-780).

³⁰ Haslam (2006: 252-264); Haslam, Kashima, Loughnan, Shi & Suitner (2008:248-258).

³¹ Haslam(2006).

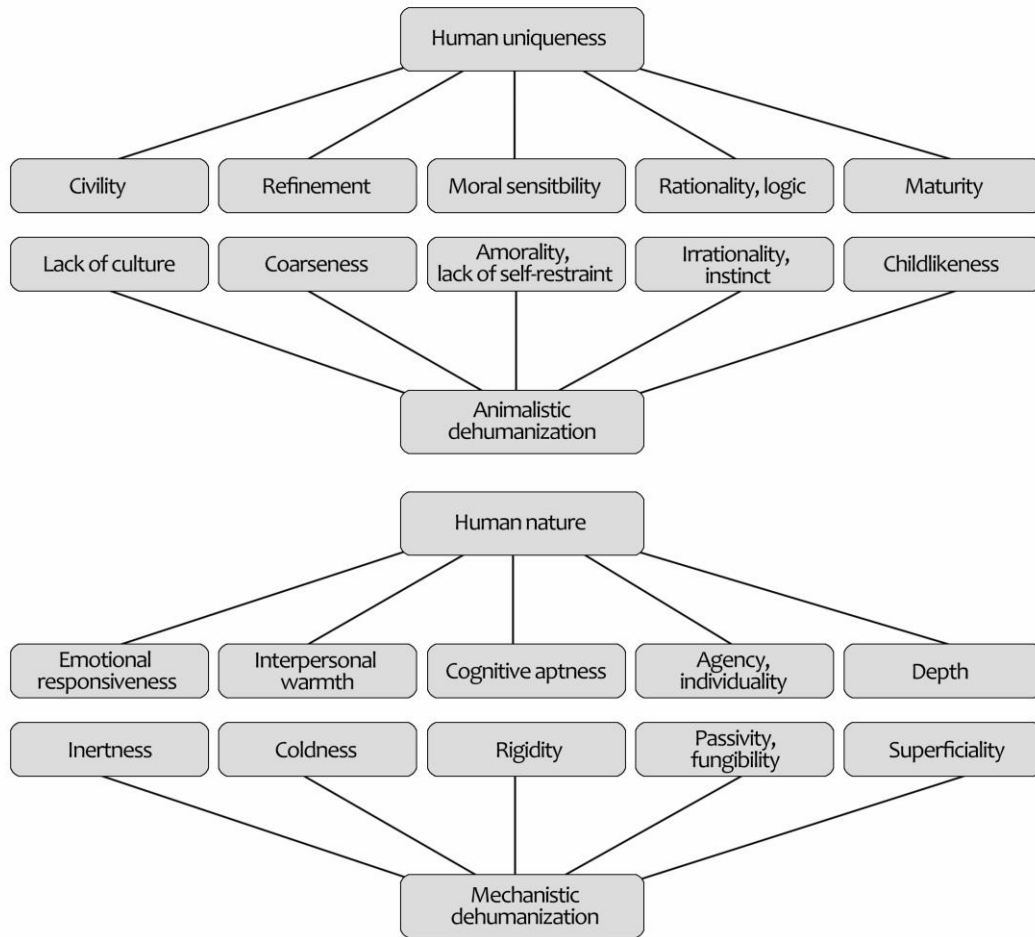


Fig. 2: La figura rappresenta le connessioni tra due differenti concezioni di umanità e due differenti modalità di de-umanizzazione come rappresentate da Haslam, 2006.

Ciò che è rilevante in questo contesto è che le due tipologie di deumanizzazione abbiano valenze emotive tra loro molto diverse. Il disgusto è considerato l'emozione preminente della forma di deumanizzazione animale. Scrive Haslam³²:

³² Haslam (2006).

Disgust and revulsion feature prominently in images of animalistically dehumanized others: Represented as apes with bestial appetites or filthy vermin who contaminate and corrupt, they are often viscerally despised»³³.

Mentre non ha molto a che vedere con la forma meccanica di deumanizzazione; continua Haslam:

The mechanistic form of dehumanization has a quite different emotional signature. As it involves emotional distancing and represents the other as cold, robotic, passive, and lacking in depth, it implies indifference rather than disgust. Typically, mechanistically dehumanized others are seen as lacking the sort of autonomous agency that provokes strong emotion and are more likely to be seen as emotionally inert³⁴.

Mantenendo ferma la distinzione di Haslam, la nostra proposta di ricerca considera che per quel che concerne l'*uncanny* ci troviamo di fronte a un caso molto particolare di *deumanizzazione appartenente allo spazio di intersezione tra la deumanizzazione meccanica e quella animale*. Questa particolare posizione emerge a nostro avviso dalle particolari cause alla base del fenomeno in questione. In un certo senso i robot con cui interagiamo sono abbastanza umani e in effetti, almeno fino a un certo livello di *human likeness*, il grado di *familiarity* e la conseguente disponibilità empatica all'interazione continuano a crescere; in un altro senso, quando il livello di *human likeness* supera quella soglia responsabile dell'*uncanny valley*, i robot con cui interagiamo sono 'pericolosamente/stranamente' umani. Si viene così a creare una sorta di cortocircuito che auto riflessivamente investe il nostro sentirci *animali razionali* e la possibilità di essere *meccanismi razionali* al pari dei corpi robotici con cui stiamo interagendo. La 'contaminazione' che temiamo, che ci disgusta e che repelle, è la contaminazione del meccanico e del materiale che un'attribuzione piena di animalità razionale, o meglio di umanità, ci appare portare con sé. Dobbiamo non poter percepire lo scarto tra ciò che vorremmo/pensiamo di essere e i corpi meccanici che potremmo essere: se lo percepiamo seppur minimamente, questo basta per innescare il corto circuito e fare emergere l'emozione del disgusto e connesse dinamiche deumanizzanti. La nostra analisi di dettagli aumenta, il nostro trattare l'umanità reciproca si problematizza complessivamente, iniziamo a dismettere di umanità molti dettagli, anche quelli che precedentemente, con livelli di *human likeness* minori, avevano garantito un livello minimale di riconoscimento e che avevamo posto a fondamento della nostra disponibilità empatica all'interazione.

³³ *Ibidem.* p. 258.

³⁴ *Ibidem.*

Proprio questo particolare corto circuito categoriale³⁵ è, a nostro avviso, alla base del senso di disgusto che induce l'attuazione delle dinamiche deumanizzanti in azione nell'*uncanny valley*, dinamiche che hanno sia i tratti della deumanizzazione animale che meccanica.

Guadagnare, dunque, il massimo del realismo potrà aiutarci a risolvere la sfida tecnologica posta dall'*uncanny valley*; per quel che invece concerne la sfida filosofica è bene essere più cauti. L'*uncanny valley* ci appare un problema ancora aperto, un problema che a nostro giudizio è profondamente connesso con i nostri processi di categorizzazione e in particolare con il nostro categorizzarci come animali razionali e il riconoscere altri individui come animali razionali.

Ringraziamenti

Ringraziamo Mario Alai, Alessandro Aldini, Vincenzo Fano, Matteo Morganti e Alfredo Paternoster per aver letto le prime bozze di questo lavoro e averle discusse con noi. Questi confronti ci hanno certamente permesso di migliorare il nostro testo ponendo le basi per ulteriori ricerche.

Bibliografia

- Alexander, O., Rogers, M., Lambeth, W., Chiang M., P. Debevec, P. (2009), "The digital Emily project: photoreal facial modeling and animation", in *ACM SIGGRAPH 2009 Courses*, ACM, New York, NY, USA, SIGGRAPH.
- Alexander, O., Fyffe, G., Busch, J., Yu, X., Ichikari, R., Jones, A., et al. (2013). *Digital Ira: Creating a Real-Time Photoreal Digital Actor*. Paper presented at the ACM SIGGRAPH 2013 Posters, Anaheim, California.
- Angelucci A., Bastioni M., Graziani P., and Rossi M.G., 2014. A Philosophical Look at the
 Uncanny Valley. In: J. Seibt, R. Hakli, and M. Nørskov eds. *Sociable Robots and the Future of Social Relations*. Amsterdam - Berlin - Tokyo - Washington DC: IOS Press, pp. 165-171.

³⁵ Vedi Redstone (2013) sul problema della categorizzazione.

- Asimov, I. (1978), *The Bicentennial Man* (*L'uomo bicentenario*, trad. it., Milano, Il Saggiatore, 2005).
- Bartneck C., Kanda T., Ishiguro H., and Hagita N., (2007). *Is the Uncanny Valley an Uncanny Cliff?. The 16th IEEE International Symposium on Robot and Human interactive Communication, RO-MAN 2007*, pp.368-373.
- Bastioni, M., Graziani, P. (2011), "Quanta matematica serve per costruire un essere umano?", *Lettera Matematica PRISTEM*, 78, pp. 26-37.
- Brenton, H., Gillies D., M., Ballin D., Chatting, D. (2005), *The Uncanny Valley: Does it exist?*, Paper presented at the HCI 2005, Animated Characters Interaction Workshop, Napier University, Edinburgh, 2005.
- Buckels E. E., Trapnel, P. D. (2013), "Disgust Facilitates Outgroup Dehumanization", *Group Processes & Intergroup Relations*, 16, 6, pp.771-780.
- Ekman, P., Friesen, W. (1978), *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologists Press, Palo Alto.
- Faulkner, J., Schaller, M., Park, J. H., Duncan, L. A. (2004), "Evolved disease-avoidance mechanisms and contemporary xenophobic attitudes", *Group Processes and Intergroup Relations*, 7, pp. 333-353.
- Fiske, S. T., Cuddy, A. J., Glick, P., Xu, J. (2002), "A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition", *Journal of Personality and Social Psychology*, 82, pp. 878-902.
- Freud, S. (1953), "The uncanny [das Unheimliche]", *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, ed. & trs. James Strachey, XVII, London, Hogarth, pp. 219-252.
- Ghosh, A., Fyffe, G., Tunwattanapong, B., Busch, J., Yu, X., Debevec, P. (2011), "Multiview face capture using polarized spherical gradient illumination", in *ACM Transactions on Graphics*, 30, 6, 129.

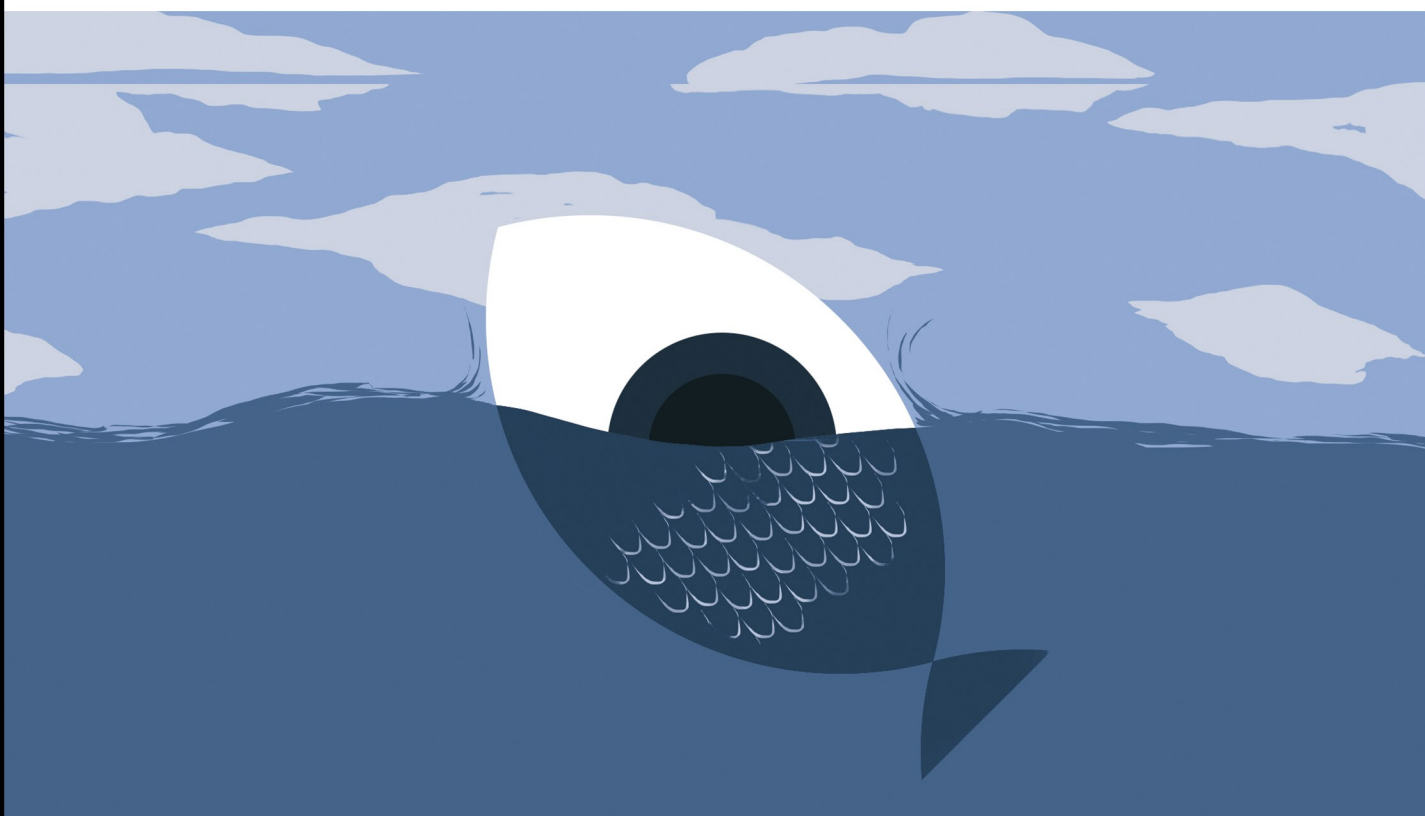
- Götz, J., Kiesler, S., Powers, A. (2003), "Matching robot appearance and behaviour to tasks to improve human-robot cooperation", *Robot and Human Interactive Communication*, Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 03), Oct. 31-Nov. 2, Millbrae, Calif., pp. 55-60.
- Gouskos, C. (2006), *The Depths of the Uncanny Valley* Online: <http://uk.gamespot.com/features/6153667/index.html>
- Haslam, N. (2006), "Dehumanization: An integrative review", *Personality and Social Psychology Review*, 10, pp. 252-264.
- Haslam, N., Kashima, Y., Loughnan, S., Shi, J., Suitner, C (2008), "Subhuman, inhuman, and superhuman: contrasting humans with nonhumans in three cultures", *Social Cognition*, 26, No. 2, pp. 248-258.
- Hodson, G., Costello, K. (2007), "Interpersonal Disgust, Ideological Orientations, and Dehumanization as Predictors of Intergroup Attitudes", *Psychological Science*, 8, 8, pp. 691-698.
- Jentsch, E. (1906), "Zur Psychologie des Unheimlichen", *Psychiatrisch-Neurologische Wochenschrift*, 8(22), pp. 195-198.
- Lewkowicz, D. J., Ghazanfar, A. A. (2012), "The development of the uncanny valley in infants", *Developmental Psychobiology*, 54, pp. 124-132.
- MacDorman, K. F. (2005). *Androids as Experimental Apparatus: Why Is There an Uncanny Valley and Can We Exploit It?*. CogSci-2005 Workshop, Stresa, Italy, July 25-26, pp. 108-118.
- MacDorman, K. F., Ishiguro, H. (2006), "The uncanny advantage of using androids in cognitive science research", *Interaction Studies*, 7(3), pp. 297-337.
- MacDorman, K. F., Vasudevan, S. K., Ho, C.-C.(2009), "Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures", *AI & Society*, 23(4), pp. 485-510.
- MacDorman, K. F., Green, R. D., Ho, C.-C., Koch, C. (2009), "Too real for comfort: Uncanny responses to computer generated faces", *Computers in Human Behavior*, 25, pp. 695-710.

- Misselhorn, C. (2009), "Empathy with Inanimate Objects and the Uncanny Valley", *Mind & Machines*, 19, pp. 345-359.
- Moll, J., Oliveira-Sousa, R. de, Moll, F. T., Ignacio, F. A., Bramati, I. E., D.E.M. Caparelli, (2005), "The moral affiliations of disgust: A functional MRI study", *Cognitive and Behavioral Neurology*, 8, pp. 68-78.
- Mori M., 1970. The Uncanny Valley. *Energy*, vol. 7, no.4, pp. 33-35 (in Giapponese). Traduzione in Inglese di MacDormann K. F. and Kageki N., The Uncanny Valley, in *IEEE Robotics & Automation Magazine*, June, 2, 2012, pp. 98-100, anche reperibile presso *IEEE Spectrum*, posted 12 June 2012:
<http://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley>
- Navarrete, C. D., Fessler, D. M. T. (2006), "Disease avoidance and ethnocentrism: The effects of disease vulnerability and disgust sensitivity on intergroup attitudes", *Evolution and Human Behavior*, 27, pp. 270–282.
- Nussbaum, M. C. (2004), *Hiding from Humanity. Disgust, Shame, and the Law*, Princeton, Princeton University Press (*Nascondere l'umanità. Il disgusto, la vergogna, la legge*, trad. it. di C. Corradi, Roma, Carocci, 2005).
- Nussbaum, M. C. (2010), *From Disgust to Humanity*, New York, Oxford University Press (*Disgusto e umanità. L'orientamento sessuale di fronte alla legge*, trad. it. di S. De Petris, Milano, Il Saggiatore, 2011).
- Olatunji, B. O., Williams, N. L., Tolin, D. F., Abramowitz, C. N., Sawchuk, C. N., Lohr, C.J.M., J.M. et al. (2007), "The Disgust Scale: Item analysis, factor structure, and suggestions for refinement", *Psychological Assessment*, 19, pp. 281-297.
- Park, J. H., Faulkner, J., Schaller, M. (2003), "Evolved disease avoidance processes and contemporary anti-social behavior: Prejudicial attitudes and avoidance of people with physical disabilities", *Journal of Nonverbal Behavior*, 27, pp. 65–87.
- Philips, M. L., Young, A. W., Senior, C., Brammer, M., Andrews, C., Calder A. J. et al. (1997), "A specific neural substrate for perceiving facial expression of disgust", *Nature*, 389, pp. 495-498.

- Pizarro, D. A., Detweiler-Bedell, B., Bloom, P. (2006), "The Creativity of Everyday Moral Reasoning. Empathy, Disgust, and Moral Persuasion", in J. C. Kaufman e J. Baer (a cura di), *Creativity and Reason in Cognitive Development*, Cambridge, Cambridge University Press, pp. 81-98.
- Ramey, C. H. (2005), "The uncanny valley of similarities concerning abortion, baldness, heaps of sand, and humanlike robots", *Proceedings of the Views of the Uncanny Valley workshop, EEE-RAS International Conference on Humanoid Robots*, Japan, Tsukuba.
- Redstone, J. D. B. (2013), "Beyond the Uncanny Valley: A Theory of Eeriness for Android science Research", MA Thesis, Ottawa, Carleton University.
- Rossi, M. G. (2013), *Il giudizio del sentimento. Emozioni, giudizi morali, natura umana*, Roma, Editori Riuniti University Press.
- Rozin, P., Fallon, A. E. (1987), *A perspective on disgust*, in *Psychological Review*, 94, pp. 23-41.
- Rozin, P., Haidt, J., McCauley, C. R. (2008), "Disgust: The body and soul emotion in the 21st century", in D. McKay e B.O. Olatunji (a cura di), *Disgust and its disorders*, Washington DC, American Psychological Association, pp. 9-29.
- Rozin, P., Haidt, J., Fincher, K. (2009), "From oral to moral", *Science*, 323, pp. 1179-1180,.
- Rozin, P., Lowery L., Imada S., and Haidt J.. (1999). "The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity).", *Journal of Personality and Social Psychology*, 76, 4, pp. 574-586.
- Sherman, G. D., Haidt, J. (2011), "Cuteness and Disgust: The Humanizing and Dehumanizing Effect of Emotion", *Emotion Review*, 3, 2, pp. 245-251.
- Steckenfinger, S. A., Ghazanfar, A. A. (2009), "Monkey visual behavior falls into the uncanny valley", *Proceedings of the National Academy of Sciences*, 106, pp. 18362-18466.

- Tinwell, A., Grimshaw, M., Williams, A. (2011a), "The Uncanny Wall", *International Journal of Arts and Technology*, 4, 3, pp. 326-341.
- Tinwell, A., Grimshaw, M., Nabi, D. A., Williams, A. (2011b), Facial expression of emotion and perception of the Uncanny Valley in virtual characters, 27, pp. 741-749.
- Tinwell, A. (2014), "Appllying psychological plausibility to the Uncanny Valley phenomenon", in M. Grimshaw (Ed.), *Oxford Handbook of Virtuality*, Oxford, Oxford University Press.
- Wolchover, N. (2011), *What scientists are doing about creepy CGI humans*, 2011, today.com On-line: www.today.com/id/44953405/ns/today-to-day_tech/t/what-scientists-are-doing-about-creepy-cgi-humans/

Il concetto di razionalità è stato qui affrontato con analisi di tipo storico-filosofico, psicologico-cognitivo, logico-matematico e sociologico, avendo di mira i limiti e le possibilità del ragionamento umano. Oltre a indagare il modo in cui il concetto di razionalità è stato differentemente declinato in diversi periodi storici, i contributi di questo volume interrogano i confini della razionalità stessa mettendo in evidenza le dinamiche “irrazionali” operanti all’interno di numerosi contesti onde emergono questioni di tipo etico, politico, economico, estetico, ecc. Il lettore è così guidato in una riflessione sulla categoria di razionalità che ne fa emergere lo statuto rispetto alle sue modalità e specificità, e viene contemporaneamente dotato di strumenti, messi a punto nell’ambito della ricerca filosofica, logica, economica, giuridica e nelle scienze cognitive, atti a fortificare le sue capacità di lettura critica di situazioni che si presentano a diversi livelli dell’interazione sociale.



Il concetto di razionalità è stato qui affrontato con analisi di tipo storico-filosofico, psicologico-cognitivo, logico-matematico e sociologico, avendo di mira i limiti e le possibilità del ragionamento umano. Oltre a indagare il modo in cui il concetto di razionalità è stato differentemente declinato in diversi periodi storici, i contributi di questo volume interrogano i confini della razionalità stessa mettendo in evidenza le dinamiche “irrazionali” operanti all’interno di numerosi contesti onde emergono questioni di tipo etico, politico, economico, estetico, ecc. Il lettore è così guidato in una riflessione sulla categoria di razionalità che ne fa emergere lo statuto rispetto alle sue modalità e specificità, e viene contemporaneamente dotato di strumenti, messi a punto nell’ambito della ricerca filosofica, logica, economica, giuridica e nelle scienze cognitive, atti a fortificare le sue capacità di lettura critica di situazioni che si presentano a diversi livelli dell’interazione sociale.



ISSN 2037-4348

Pierluigi Graziani, Giorgio Grimaldi, Massimo Sangoi, ANIMALI RAZIONALI

ANIMALI RAZIONALI

STUDI SUI CONFINI E SULLE POSSIBILITÀ DELLA RAZIONALITÀ

a cura di

Pierluigi Graziani
Giorgio Grimaldi
Massimo Sangoi

Isonomia *Epistemologica*