SOPHIA

Ricerche su i fondamenti
e la correlazione dei saperi

Semestrale

Anno XVIII n. 1
2026

SOPHIA *Rivista internazionale*
Istituto Universitario Sophia
Via San Vito 28, loc. Loppiano
50063 Figline e Incisa Valdarno (FI) – Italia
Anno XVIII – 2026/1 (Gennaio-Giugno 2026)
ISSN 2036-5047

Direttore scientifico: Marco Martino

Comitato di redazione: Giuseppe Argiolas, Antonio Maria Baggio, Luigino Bruni, Bernhard Callebaut, Piero Coda, Valentina Gaudiano, Benedetto Gui, Marco Martino, Declan O'Byrne, Paul O'Hara, Giovanna Maria Porrino, Judith Povilus, Sergio Rondinara, Daniela Ropelato, Gérard Rossé, Marco Sanchioni.

Comitato scientifico internazionale: Angela Ales Bello (*Pontificia Università Lateranense, Roma*), Kurt Appel (*Università di Vienna, Austria*), Alessandra Beccarisi (*Università del Salento*), Maria Clara Lucchetti Bingemer (*Pontificia Università Cattolica di Rio de Janeiro, Brasile*), Vincenzo Buonomo (*Pontificia Università Lateranense, Roma*), Massimo Cacciari (*Università Vita-Salute San Raffaele di Milano*), Filipe Campello (*Universidade Federal de Pernambuco, Brasile*), Giuseppe Cantillo † (*Università di Napoli "Federico II"*), Peter Casarella (*Duke University School of Divinity, U.S.A.*), Bernhard Casper † (*Università Freiburg im Breisgau, Germania*), Claudio Ciancio (*Università del Piemonte Orientale*), Giuseppe D'Anna (*Università Cattolica del Sacro Cuore*), Mario De Caro (*Università degli studi Roma Tre, Roma*), Massimo Donà (*Università Vita-Salute San Raffaele, Milano*), Adriano Fabris (*Università di Pisa*), Emmanuel Falque (*Institut Catholique de Paris, Francia*), Riccardo Ferri (*Pontificia Università Lateranense, Roma*), Lorenzo Fossati (*Università Cattolica del Sacro Cuore*), Emmanuel Gabellieri (*Université Catholique de Lyon, Francia*), Gianluca Garelli (*Università di Firenze*), Giulio Giorello † (*Università degli Studi di Milano*), André Habisch (*Catholic University Eichstaett-Ingolstadt, Germania*), Vittorio Hösle (*University of Notre Dame, U.S.A.*), Philippe Hu (*Fu Jen Catholic University, Taiwan*), Luca Illetterati (*Università degli Studi di Padova*), Marco Ivaldo (*Università degli studi "Federico II" di Napoli*), Mario Longo (*Università degli Studi di Verona*), Daniel López (*Universidad Católica de Córdoba, Argentina*), Giancarlo Magnano San Lio (*Università di Catania*), Carmelo Meazza (*Università degli Studi di Sassari*), Mauro Mantovani (*Università Pontificia Salesiana, Roma*), Massimo Marassi (*Università Cattolica del Sacro Cuore, Milano*), Massimiliano Marianelli (*Università degli Studi di Perugia*), Giulio Maspero (*Pontificia Università Santa Croce di Roma*), Edoardo Massimilla (*Università degli studi "Federico II" di Napoli*), Letterio Mauro (*Università degli Studi di Genova*), Eugenio Mazzarella (*Università degli studi "Federico II" di Napoli*), John Milbank (*University of Nottingham, Inghilterra*), Donald W. Mitchell (*Purdue University, U.S.A.*), Juan Carlos Scannone † (*Facoltà di Filosofia e Teologia di San Miguel, Buenos Aires, Argentina*), Pierangelo Sequeri (*Facoltà Teologica dell'Italia Settentrionale di Milano*), Francesco Tomatis (*Università degli Studi di Salerno*), Roberto Tomichá (*Universidad Católica Boliviana de Cochabamba, Bolivia*), Giovanni Ventimiglia (*Università di Lucerna, Svizzera*), Vincenzo Vitiello (*Università Vita-Salute San Raffaele, Milano*), Lubomir Žak (*Università Palacký di Olomouc, Repubblica Ceca*), Stefano Zamagni (*Università degli Studi di Bologna*), Gonzalo Zarazaga (*Universidad Católica de Córdoba, Argentina*).

Segreteria scientifica:
Denise Vincenti, Coordinatore
(denise.vincenti@unipg.it);
Francesco Patrone
(francesco.patrone@sophiauniversity.org);
Alessandro Borghesi
(alessandro.borghesi@sophiauniversity.org);
Francesca Annamaria Gambini
(francescaannamaria.gambini@sophiauniversity.org)
Riccardo Rezzesi
rrezzesi@univ-catholyon.fr

Editorial Director: Michele Zanzucchi
(michele.zanzucchi@sophiauniversity.org)

Rivista Semestrale Open Access
journal.sophiauniversity.org

Registrazione al Tribunale di Roma n° 405/2008
Iscrizione al R.O.C. n. 5849 del 10/12/2001

Gli scritti proposti per la pubblicazione in questa rivista sono peer reviewed

SOPHIA

* *Disegno della testata e progetto grafico di Ezio Bartocci*

SOPHIA

Ricerche su i fondamenti
e la correlazione dei saperi

SOPHIA *2026* INDICE

EDITORIALE

9 Uno sguardo sul presente in evoluzione?
Francesco Patrone

SAGGI

11 Transdisciplinarità e μεταξύ: la conoscenza nello spazio relazionale tra i saperi
Marco Martino - Marco Sanchioni

29 Phenomenological Elements in Carnap's Aufbau: The Epistemological Role of Sensible Intuition
Vincenzo Fano - Niccolò Covoni

53 Quando un'analogia guida la scienza. Il ruolo metodologico dell'analogia nella termodinamica dei buchi neri
Vincenzo Fano - Marco Sanchioni

69 Once upon a time in the Pleistocene: Integrating Paleoanthropology and Evolutionary Psychology
Matteo Bedetti

89 Beauty as Divine Order: Berkeley on Proportion, Reason, and Subordination
Aurora Yu

113 Divine Necessity and Human Freedom in Seventeenth-Century Metaphysics: Spinoza's *causa sui* and Mullā Sadrā's *wājib al-wujād*
Jamil Akhtar

127 Oltre la dimostrazione. Il limite della ragione e l'argomento ontologico in Kant
Francesco Porchia

137 Baruch Spinoza e l'idea vera. La conoscenza di Dio come conoscenza-in-Dio nel *Trattato sull'emendazione dell'intelletto* e nel *Breve trattato su Dio, l'uomo e la sua felicità*
Federico Brutti

149 Automazione e cibernazione. Hannah Arendt tra *The Human Condition* e *On the Human Condition*
Saverio Ferrieri

FORUM

173 Sulla pertinenza etica della *restorative justice*. Una discussione
Chiara Gabbolini

EDITORIALE

Uno sguardo sul presente in evoluzione

Questo primo numero dell'anno 2026 di *Sophia. Ricerche su i fondamenti e la correlazione dei saperi* incarna perfettamente la visione che sta alla base non soltanto della rivista, che arriva al suo diciannovesimo anno d'attività, ma dell'intero ateneo, dell'Istituto Universitario Sophia. Alla base di questa pubblicazione, così come della nuova offerta formativa, sta l'attenzione all'interdisciplinarità e alla transdisciplinarità, un punto di vista fecondo ed estremamente importante per la comprensione di un orizzonte, come quello contemporaneo, che sempre di più sembra sfuggirci dalle mani, evolvendosi continuamente. Questo numero, quindi, forte di approcci diversi e di spunti tematici sempre estremamente interessanti e rigorosamente trattati, può considerarsi come una sorta di bussola per chiunque voglia abitare in modo consapevole e critico il presente.

Con Marco Martino e Marco Sanchioni entriamo direttamente nel vivo delle questioni in gioco, adottando uno sguardo innovativo che ci porta alla considerazione della transdisciplinarità come spazio epistemico, uno spazio interpretato attraverso il concetto filosofico di μεταξύ. Vincenzo Fano e Niccolò Covoni offrono uno sguardo sull'*Aufbau* di Rudolf Carnap, proponendone una lettura epistemologica, in direzione di un "empirismo strutturale" e ancora Fano e Sanchioni ci invitano a riflettere sul concetto di analogia, un concetto che viene studiato nel contesto della termodinamica dei buchi neri. Matteo Bedetti, poi, integra le evidenze provenienti dalla paleoantropologia e dall'archeologia nel ricostruire le origini evoluzionistiche della cognizione e del comportamento umani. La transdisciplinarità e l'interdisciplinarità prendono il centro del campo fin da questi primi contributi, per poi emergere in maniera ancora più evidente con i seguenti.

A testimonianza del respiro internazionale della rivista, Aurora Yu, dell'Università di San Diego, sposta cronologicamente la questione, portando il *focus* sull'epoca moderna e proponendo una riflessione che cerca una ricostruzione filosofica coerente della teoria della bellezza in George Berkeley. Jamil Akhtar, dell'Università del Punjab (Pakistan), intreccia il concetto di *causa sui* di Baruch Spinoza con

quello di *wāib al-wujād*, di Mullā Sadrā, mentre Francesco Porchia ci introduce a una riflessione sui limiti della ragione e sull'argomento ontologico in Kant. Si ritorna a Spinoza e si chiude questa parentesi "modernista" con il contributo di Federico Brutti, che prende in esame la concezione spinoziana dell'idea vera, proponendo un ripensamento del processo di autorivelazione di Dio.

Saverio Ferrieri, in chiusura di numero, esamina il rapporto tra esseri umani e macchine attraverso alcuni scritti di Hannah Arendt, mentre il *Forum* è dedicato a un'interessante discussione sul tema dell'etica e della *restorative justice*, a cura di Chiara Gabbolini.

FRANCESCO PATRONE
Ricercatore in Storia della filosofia presso l'Istituto Universitario Sophia e membro della Segreteria Scientifica di *Sophia. Ricerche su i fondamenti e la correlazione dei saperi*
francesco.patrone@sophiauniversity,org

SAGGI

Transdisciplinarità e μεταξύ: la conoscenza nello spazio relazionale tra i saperi

di Marco Martino - Marco Sanchioni

If many phenomena in contemporary research possess an intrinsically relational structure, knowledge adequate to them cannot be contained within a single discipline. The literature on transdisciplinarity, however, rarely thematizes the epistemic space in which the relation between forms of knowledge becomes productive of knowledge. This contribution proposes to interpret that space through the concept of μεταξύ*, understood not as synthesis but as a relational structure that allows different perspectives to enter into relation without being reduced to one another.*

Keywords: Metaxy, In-between, Transdisciplinarity, Unity of Science, Philosophy of Science

1. Introduzione

La configurazione moderna del sapere, articolata in discipline autonome dotate di oggetti e metodi propri, rappresenta l'esito di un processo di differenziazione che ha reso possibile l'elevata specializzazione della conoscenza[1]. Tuttavia, tale articolazione presuppone implicitamente una concezione degli oggetti come determinabili indipendentemente dalle relazioni in cui sono implicati – presupposto che, nel contesto contemporaneo, mostra limiti strutturali.

Nel panorama contemporaneo questa presupposizione entra sempre più frequentemente in tensione con la natura degli oggetti studiati. In diversi ambiti – dalle scienze naturali alle scienze sociali – i fenomeni si presentano come strutture intrinsecamente relazionali, nelle quali le proprietà non possono essere attribuite a elementi isolati né ricostruite come semplice somma di contributi parziali[2]. In tali casi, l'astrazione disciplinare, pur metodologicamente necessaria, risulta epistemicamente insufficiente: essa isola dimensioni del fenomeno che non sono pienamente intelligibili se non nel contesto delle relazioni che le costituiscono.

In risposta a questa difficoltà, negli ultimi decenni si è sviluppato un crescente interesse per approcci interdisciplinari e transdisciplinari, volti a mettere in relazione prospettive differenti al fine di affrontare problemi complessi. In queste prospettive la conoscenza non è più prodotta esclusivamente all'interno delle discipline, ma emerge dall'interazione tra saperi eterogenei, spesso in contesti applicativi e socialmente distribuiti. Tuttavia, pur insistendo sul carattere relazionale della conoscenza, tali approcci tendono a rimanere sul piano metodologico o organizzativo, senza interrogare in modo esplicito le condizioni epistemiche entro cui tale relazione diventa effettivamente produttiva di conoscenza.

Il punto cruciale che si intende mettere in risalto è che la difficoltà non riguarda soltanto come mettere in relazione discipline differenti, ma che cosa significhi conoscere quando la determinazione stessa dell'oggetto dipende da relazioni che non sono secondarie rispetto ad esso. Se le proprietà dei fenomeni emergono nel-

1 - Sulla funzione epistemica delle discipline si veda J.A. Jacobs, *In Defense of Disciplines: Interdisciplinarity and Specialization in the Research University*, University of Chicago Press, Chicago 2014, che sottolinea come la specializzazione costituisca una condizione necessaria per la produzione di conoscenza rigorosa. In una prospettiva diversa ma complementare, cfr. T.S. Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago 1962, dove l'organizzazione disciplinare è interpretata come condizione della "scienza normale", intesa come pratica condivisa all'interno di comunità scientifiche strutturate.

2 - Sul carattere non riducibile dei sistemi complessi e sulla critica a una concezione analitico-riduzionista della conoscenza si veda E. Morin, *Introduction à la pensée complexe,* Seuil, Paris 2005, che insiste sulla necessità di pensare l'organizzazione dei fenomeni come irriducibile alla semplice somma delle parti e sulla centralità delle relazioni nella costituzione del reale.

le relazioni, allora anche la conoscenza non può essere concepita come composizione di descrizioni indipendenti, ma deve essere pensata a partire dallo spazio in cui tali descrizioni entrano in rapporto. In questo senso, il problema della transdisciplinarità rinvia a una questione più fondamentale: la natura dello spazio epistemico in cui la relazione tra prospettive eterogenee diventa conoscenza.

Nonostante la centralità di questo tema, tale spazio resta in larga misura implicito nella letteratura esistente. Esso viene presupposto nei processi di integrazione, descritto nei suoi effetti o fondato in termini ontologici, ma raramente tematizzato come problema epistemico autonomo. È precisamente questa lacuna che il presente lavoro intende affrontare. In particolare, si propone di interpretare lo spazio "tra" i saperi attraverso il concetto filosofico di μεταξύ, inteso non come sintesi o integrazione, ma come struttura relazionale che consente a prospettive differenti di entrare in rapporto senza essere ridotte l'una all'altra. In questa prospettiva la transdisciplinarità non viene intesa come semplice coordinamento tra discipline, ma come forma di conoscenza che emerge nello spazio intermedio tra di esse.

2. I modelli di transdisciplinarità nella letteratura

Per collocare in modo preciso il contributo qui proposto, è utile richiamare, sia pure in forma sintetica, i principali significati attribuiti alla transdisciplinarità nella letteratura contemporanea[3]. L'intento non è quello di offrire una ricostruzione esaustiva, ma di mettere in luce alcune linee interpretative rilevanti in relazione al problema della conoscenza, in particolare per quanto riguarda le condizioni entro cui essa viene prodotta e articolata. Negli ultimi decenni, infatti, il termine è stato utilizzato in contesti teorici diversi e talvolta eterogenei, dando luogo a una pluralità di interpretazioni che non sempre convergono su un nucleo concettuale comune. Senza alcuna pretesa di esaustività, ciò che segue intende semplicemente delineare alcune famiglie di posizioni rilevanti, al fine di chiarire il quadro entro cui si inserisce la proposta sviluppata in questo lavoro.

3 - La rassegna che segue non intende offrire una ricostruzione completa della letteratura sulla transdisciplinarità, ma piuttosto individuare alcune linee interpretative ricorrenti, utili a chiarire il posizionamento teorico della proposta qui sviluppata.

Tra queste, una delle interpretazioni più influenti è quella proposta nell'ambito della sociologia della scienza da Gibbons[4] e Nowotny[5], comunemente indicata come modello della *Mode 2 knowledge production*.

> (M2) Transdisciplinarità Mode 2. In questo quadro, la transdisciplinarità è intesa come una modalità di produzione della conoscenza che emerge direttamente nei contesti di applicazione, in cui problemi, metodi e criteri di validazione non sono definiti all'interno delle discipline, ma co-costruiti attraverso l'interazione tra una pluralità di attori eterogenei.

Più precisamente, (M2) descrive una trasformazione del modo stesso in cui la conoscenza viene prodotta. A differenza del paradigma disciplinare tradizionale, in cui le domande di ricerca prendono forma all'interno di comunità scientifiche relativamente autonome, qui esse nascono in ambienti aperti e distribuiti, nei quali si intrecciano esigenze scientifiche, sociali, economiche e tecnologiche. La transdisciplinarità designa così una configurazione operativa della ricerca: un processo flessibile in cui competenze diverse vengono mobilitate in funzione di problemi situati, spesso caratterizzati da un'elevata complessità e da una forte rilevanza pratica. Si può dire che la conoscenza assume un carattere intrinsecamente contestuale e orientato all'uso. Essa non è più prodotta esclusivamente all'interno delle istituzioni accademiche, ma in una pluralità di luoghi e attraverso reti di collaborazione temporanee, che includono attori non accademici e forme di sapere eterogenee. Di conseguenza, anche i criteri della sua validazione tendono a trasformarsi, includendo – accanto alla coerenza teorica – forme di legittimazione legate alla sua efficacia e alla sua "robustezza" nel contesto di applicazione[6].

Tra le interpretazioni più rilevanti, una seconda linea di sviluppo identifica la transdisciplinarità come una forma di ricerca orientata alla gestione di problemi complessi, che eccedono i confini disciplinari tradizionali e coinvolgono una pluralità di attori sociali[7].

---

4 - Cfr. M. Gibbons - C. Limoges - H. Nowotny - S. Schwartzman - P. Scott - M. Trow, *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies,* Sage, London 1994.

5 - H. Nowotny - P. Scott - M. Gibbons, *Introduction: "Mode 2" revisited: The new production of knowledge,* in «Minerva», 41 (2003) 3, pp. 179-194.

6 - Per una rassegna critica del modello della Mode 2 e delle sue diverse interpretazioni nella letteratura, cfr. L.K. Hessels - H. Van Lente, *Re-thinking new knowledge production: A literature review and a research agenda, in «Research Policy»*, 37 (2008) 4, pp. 740-760.

7 - Per una formulazione sistematica della transdisciplinarità orientata alla risoluzione di problemi, cfr. C. Pohl - G. Hirsch Hadorn, *Principles for Designing Transdisciplinary Research,* Oekom, München 2007; J.T. Klein (ed.), *Transdisciplinarity: Joint Problem Solving among Science, Technology, and Society,* Springer, Dordrecht 2001. Il riferimento alla "complessità" in questo contesto è tuttavia intrinsecamente poli-

(PS) Transdisciplinarità problem-solving. La transdisciplinarità può essere definita, in questo senso, come una modalità di produzione della conoscenza orientata alla comprensione e alla trasformazione di problemi del mondo reale caratterizzati da elevata complessità, incertezza e rilevanza sociale, attraverso l'integrazione di prospettive disciplinari e non disciplinari all'interno di un processo condiviso di definizione, analisi e soluzione del problema.

Questa impostazione, sviluppata in particolare da Pohl e Hirsch Hadorn[8], concepisce la ricerca transdisciplinare come un processo articolato in diverse fasi – identificazione e strutturazione del problema, analisi e produzione di conoscenza, applicazione e trasferimento dei risultati – nel quale discipline scientifiche e attori extra-accademici collaborano fin dall'inizio alla definizione stessa dell'oggetto di indagine. In questo quadro, la transdisciplinarità non si limita a combinare saperi esistenti, ma mira a produrre forme di conoscenza contestualizzate, orientate alla pratica e sensibili alla pluralità di prospettive coinvolte. Un elemento centrale di questo approccio è il riferimento alla complessità dei problemi affrontati[9]. Tali problemi sono caratterizzati da interdipendenze multiple, dinamiche non lineari, incertezza epistemica e rilevanza normativa, e richiedono pertanto forme di conoscenza che non possono essere prodotte all'interno di un singolo dominio disciplinare. La transdisciplinarità emerge così come una risposta metodologica ed epistemica alla necessità di trattare problemi che non sono dati come oggetti scientifici già definiti, ma devono essere costruiti attraverso un processo di interazione tra sapere scientifico e contesto sociale.

Accanto ai modelli di carattere più operativo, una terza famiglia di approcci interpreta la transdisciplinarità in senso propriamente epistemologico, come una tra-

semico e copre una gamma ampia di significati, che spaziano dalla complessità come interdipendenza sistemica fino a forme più forti di emergenza e non-riducibilità. Nel presente lavoro, tale nozione è assunta in senso operativo, rinviando a un'analisi più approfondita - in particolare rispetto ai problemi della dipendenza tra parti e totalità - a studi successivi.

8 - C. Pohl - G. Hirsch Hadorn, *Principles for Designing Transdisciplinary Research,* Oekom, München 2007.

9 - Il riferimento alla "complessità" in questo contesto è tuttavia intrinsecamente polisemico e copre una gamma ampia di significati, che spaziano dalla complessità intesa come interdipendenza sistemica e coordinamento di molteplici variabili, fino a forme più forti di emergenza e non-riducibilità, in cui il comportamento dell'insieme non è derivabile dalla somma delle parti. Questa ambiguità non è soltanto terminologica, ma segnala una differenza teoricamente rilevante tra modelli di transdisciplinarità: da un lato, approcci di carattere prevalentemente operativo – come la Mode 2 o la transdisciplinarità problem-solving – nei quali la complessità è trattata come una proprietà dei problemi da gestire; dall'altro, prospettive più radicali, come quella nicoleschiana, in cui la complessità rinvia alla struttura stessa del reale e implica una riformulazione dei presupposti epistemologici della conoscenza. Nel presente lavoro, tale nozione viene inizialmente assunta in senso operativo, rinviando a un'analisi più approfondita – in particolare rispetto ai problemi della non-riducibilità e della dipendenza tra parti e totalità – a studi successivi.

sformazione delle condizioni stesse della conoscenza. In questa prospettiva, la transdisciplinarità non si limita a coordinare discipline o a risolvere problemi complessi, ma implica una riformulazione del rapporto tra soggetto, oggetto e realtà.

> (EPI) Transdisciplinarità epistemologica: la transdisciplinarità è una forma di conoscenza che emerge dall'interazione tra diversi livelli di realtà, implicando una trasformazione del soggetto conoscente e una riorganizzazione dei presupposti epistemici del sapere.

Questa impostazione trova una delle sue formulazioni più sistematiche nell'opera di Basarab Nicolescu[10], per il quale la transdisciplinarità si fonda su tre assi portanti: l'esistenza di livelli di realtà irriducibili tra loro, la logica del terzo incluso e una nozione di complessità non riducibile alla semplice interazione tra componenti[11]. In tale quadro, la conoscenza non è più concepita come rappresentazione di un oggetto dato, ma come processo relazionale che attraversa differenti livelli del reale e coinvolge costitutivamente il soggetto. In modo convergente, anche il pensiero della complessità sviluppato da Edgar Morin insiste sul fatto che il sapere adeguato alla realtà contemporanea non può essere né riduzionista né puramente analitico, ma deve essere capace di pensare simultaneamente il tutto e le parti, l'unità e la molteplicità, riconoscendo la centralità delle relazioni e delle interdipendenze[12]. All'interno di questo quadro concettuale, la complessità non è semplicemente una proprietà dei sistemi, ma richiede una riforma del pensiero stesso. Ciò che caratterizza questo modello, rispetto agli approcci precedenti, è dunque lo spostamento dal piano metodologico a quello epistemico: la transdisciplinarità non è più intesa come una strategia di coordinamento tra saperi, ma come uno spa-

---

10 - B. Nicolescu, *La transdisciplinarité. Manifeste*, Éditions du Rocher, Monaco 1996; Id., *Methodology of Transdisciplinarity. Levels of Reality, Logic of the Included Middle and Complexity,* in «Transdisciplinary Journal of Engineering & Science», 1 (2010), pp. 19-38.

11 - Per "livelli di realtà", Nicolescu intende domini ontologicamente distinti, caratterizzati da insiemi di leggi e strutture concettuali non riducibili gli uni agli altri; il passaggio tra livelli non è descrivibile come semplice derivazione o limite, ma implica una discontinuità nei qudri teorici di riferimento. La "logica del terzo incluso" designa invece un'estensione della logica classica che caonsente di articolare la coesistenza di contraddizioni apparenti (A e non-A) attraverso un terzo termine situato a un differente livello di realtà, evitando così una loro risoluzione puramente formale o una riduzione a identità. Per una ricostruzione sistematica di questi concetti e del loro ruolo nella teoria della transdisciplinarità, si veda M. Sanchioni, *Il ruolo della transdisciplinarità nell'era dell'intelligenza artificiale (generativa),* in «Sophia» XVII (2025/1), pp. 69-94.

12 - E. Morin, *La Méthode*, 6 voll., Seuil, Paris 1977-2004; Id., *Introduzione al pensiero complesso*, Sperling & Kupfer, Milano 1993.

zio di emergenza della conoscenza, in cui si ridefiniscono le condizioni stesse del conoscere[13].

Tutti e tre i modelli esaminati, pur nelle loro differenze, condividono un presupposto fondamentale: la conoscenza transdisciplinare non nasce all'interno di una singola disciplina, ma nello spazio di relazione tra prospettive eterogenee. In (M2), tale spazio coincide con il contesto di applicazione e con la distribuzione sociale dei processi conoscitivi; in (PS), con il processo di integrazione tra saperi e attori coinvolti in un campo problematico; in (EPI), con l'attraversamento di livelli di realtà differenti. Tuttavia, proprio questo spazio relazionale, pur costitutivo di ogni forma di transdisciplinarità, resta raramente tematizzato in modo esplicito. Esso viene descritto nei suoi effetti, organizzato nei suoi dispositivi, o postulato nei suoi presupposti ontologici (il terzo incluso), ma non ancora interrogato come problema epistemico in sé. È precisamente in questo scarto che si colloca la proposta qui sviluppata: assumere lo "spazio tra" non come semplice condizione esterna del dialogo tra saperi, ma come questione filosofica fondamentale per comprendere la struttura della conoscenza transdisciplinare.

Questa lacuna si accompagna a un'ulteriore tensione, spesso implicita nelle teorie considerate. Se, da un lato, la transdisciplinarità nasce come risposta alla frammentazione del sapere, dall'altro essa espone il rischio opposto: quello di essere intesa come una forma di sapere superiore, capace di unificare dall'alto le diverse discipline. In questa prospettiva, la transdisciplinarità tenderebbe a configurarsi come una sorta di meta-scienza o di "scienza prima", incaricata di ricomporre l'unità del sapere oltre la sua articolazione disciplinare[14]. Una simile interpretazione, tuttavia, deve essere evitata. Essa presupporrebbe infatti un livello epistemico privilegiato indipendente dalle pratiche scientifiche concrete, reintroducendo una forma di fondazionalismo difficilmente sostenibile alla luce del dibattito contemporaneo in filosofia della scienza. In linea con approcci naturalizzati – come quello proposto, ad esempio, da Ladyman e Ross – la riflessione filosofica non può

13 - Per una discussione recente di questi temi nel contesto del rapporto tra transdisciplinarità e intelligenza artificiale, cfr. M. Sanchioni, *Il ruolo della transdisciplinarità nell'era dell'intelligenza artificiale (generativa)*, cit.

14 - Il rischio di interpretare la transdisciplinarità come una forma di sapere sovraordinato può essere evitato adottando una concezione della filosofia come "sapienza modesta", nel senso proposto da V. Fano, Pensare la transdisciplinarità in una prospettiva storico-filosofica, in M. Sanchioni (ed.), *Pensare il reale. Prospettive di filosofia della natura. Festschrift in onore di Sergio Rondinara,* Città Nuova, Roma 2026. In questa prospettiva, la riflessione filosofica non si configura come fondazione ultima o istanza epistemicamente privilegiata, ma come attività che si sviluppa in continuità con le pratiche scientifiche, chiarendone strutture, limiti e implicazioni senza pretendere di unificarle in un sistema superiore. La filosofia opera così nello spazio delle relazioni tra saperi, senza assumere il ruolo di una meta-scienza.

porsi come istanza esterna o superiore rispetto alle scienze, ma deve essere vincolata alle strutture teoriche e alle pratiche conoscitive effettivamente in uso[15]. Ne segue che la transdisciplinarità non può essere compresa né come una meta-scienza, né come un dispositivo di integrazione totale dei saperi. Essa non fonda le discipline, né le ricompone in un sistema unitario, ma opera nello spazio della loro relazione, mantenendo aperta la pluralità dei punti di vista. In questo quadro, la conoscenza non si configura come una sintesi definitiva, ma come una dinamica in cui prospettive eterogenee entrano in rapporto senza essere ridotte a un denominatore comune.

Proprio questa rinuncia a una fondazione unitaria rende tuttavia ancora più urgente la questione delle condizioni di possibilità della relazione tra saperi. Se la transdisciplinarità non consiste in una sintesi, ma in una messa in relazione, occorre chiarire in che modo tale relazione sia possibile e in che senso essa possa essere produttiva di conoscenza. Non basta infatti affermare che le discipline dialogano o si integrano: è necessario comprendere che cosa renda possibile questo dialogo senza che esso si risolva in una semplice giustapposizione o in una riduzione reciproca. In altri termini, la riflessione sulla transdisciplinarità rinvia inevitabilmente a uno spazio intermedio tra le discipline, che non coincide con nessuna di esse e che tuttavia rende possibile il loro incontro. È precisamente questo spazio – implicito nelle teorie esaminate, ma raramente tematizzato in quanto tale – che costituisce il problema teorico della sezione seguente.

3. Il problema del "tra": una lacuna strutturale nella teoria della transdisciplinarità

Una prima risposta alla difficoltà appena evidenziata consiste nel tentativo di fondare ontologicamente lo spazio del "tra". Questa linea è sviluppata in modo esemplare in (EPI) di matrice nicoleschiana e nelle sue elaborazioni successive, tra cui quelle proposte da Rondinara[16]. In questo quadro, la transdisciplinarità non è pensata semplicemente come coordinamento tra discipline, ma come espressione della struttura stessa del reale. Il punto di partenza è l'introduzione di una pluralità di livelli di realtà, tra loro irriducibili, la cui dinamica rende possibile la conoscenza transdisciplinare. In questo quadro, la relazione tra discipline non è primaria, ma deriva da una relazione più fondamentale tra livelli del reale. Il riferimento al cosid-

15 - J. Ladyman - D. Ross, *Every Thing Must Go: Metaphysics Naturalized,* Oxford University Press, Oxford 2007.

16 - S. Rondinara., *La metodologia transdisciplinare tra sfide e opportunità,* in «ApTh», 10 (2024) 1, pp. 15-33.

detto "terzo incluso" consente di pensare l'articolazione tra domini differenti senza ricondurli a una logica puramente duale o oppositiva. Il "tra" si configura così come una dimensione strutturata del reale, che rende possibile la connessione tra domini irriducibili senza ridurli a unità. La transdisciplinarità viene dunque interpretata come la dinamica che attraversa e connette tali livelli, rendendo possibile l'emergere di forme di conoscenza che non possono essere prodotte all'interno di un singolo dominio disciplinare.

Il tentativo di fondare ontologicamente lo spazio del "tra" rappresenta senza dubbio un avanzamento significativo rispetto agli approcci puramente metodologici. Esso consente infatti di evitare una concezione della transdisciplinarità come semplice coordinamento tra saperi e di riconoscere che la relazione tra discipline rinvia a una struttura più profonda del reale. In questa prospettiva, il "tra" non è più pensato come un vuoto da colmare, ma come una dimensione costitutiva della conoscenza. Tuttavia, proprio questa operazione introduce una difficoltà ulteriore. Nella misura in cui lo spazio intermedio viene fondato all'interno di una ontologia già articolata – ad esempio in termini di livelli di realtà e della loro dinamica – esso tende a essere interpretato come momento interno a tale struttura. Il "tra" non si presenta più come problema aperto, ma come elemento la cui funzione è già, almeno in parte, definita dal quadro ontologico di riferimento. In altre parole, (EPI) riconosce correttamente che il "tra" non può essere ridotto a interfaccia metodologica. Tuttavia, nel momento in cui lo inscrive in una ontologia dei livelli di realtà, ne definisce le condizioni di funzionamento in termini già determinati. In questo senso, la fondazione ontologica del "tra" rischia di trasformare ciò che dovrebbe restare condizione di possibilità in un elemento della struttura, ossia di reintrodurre, sotto altra forma, una determinazione di ciò che eccede ogni determinazione.

In questo modo, lo spazio intermedio non è più ciò in cui le relazioni si costituiscono, ma ciò attraverso cui esse sono già strutturate. Si può dire che la fondazione ontologica dello spazio del "tra", pur evitando una sua riduzione a semplice interfaccia metodologica, tende a saturarne la funzione teorica. Sottraendo il "tra" all'indeterminatezza, ne risulta una tensione difficilmente risolvibile: se da un lato è necessario evitare che lo spazio intermedio resti epistemicamente opaco, dall'altro la sua completa determinazione rischia di impedirne una comprensione adeguata come ambito effettivamente generativo della conoscenza.

Al termine di questa analisi, il problema dello spazio intermedio si presenta in termini più precisi. Non è sufficiente riconoscere che la conoscenza emerge dalla relazione tra prospettive eterogenee, né è sufficiente fondare tale relazione in una struttura ontologica determinata. In entrambi i casi, il "tra" rischia o di rimanere implicito, oppure di essere completamente assorbito all'interno di un quadro teorico che ne fissa in anticipo le condizioni. La difficoltà, allora, non consiste sempli-

cemente nell'individuare un concetto più adeguato per descrivere la relazione tra saperi, ma nel chiarire che tipo di concetto sia in grado di pensare tale relazione senza esaurirla. Se lo spazio intermedio viene lasciato indeterminato, esso resta epistemicamente opaco; se invece viene interamente determinato, esso perde la propria funzione generativa. Il problema consiste dunque nel pensare un "tra" che non coincida né con un vuoto metodologico né con una struttura ontologica già data. È in questo senso che si rende necessario introdurre un concetto capace di tematizzare lo spazio relazionale della conoscenza senza ridurlo a un contenuto determinato. Il riferimento al μεταξύ, nella sua elaborazione filosofica, si colloca precisamente in questa direzione.

4. Il μεταξύ: il "tra" come non-determinabile

Il concetto di μεταξύ (metaxý), che letteralmente significa "ciò che sta in mezzo", trova una delle sue formulazioni più significative nella tradizione platonica. In particolare, nel *Simposio*[17], attraverso il discorso di Diotima, esso viene impiegato per descrivere la natura di Eros, inteso non come realtà compiuta o pienamente determinata, ma come ciò che si colloca tra due poli senza coincidere con nessuno di essi. Il μεταξύ non designa semplicemente una posizione intermedia, né una via di mezzo tra opposti, ma una condizione strutturale caratterizzata da una duplice appartenenza. Eros, infatti, non è né sapiente né ignorante, né mortale né immortale, ma si situa tra questi termini, partecipando in qualche modo di entrambi senza identificarsi con essi. Il μεταξύ indica dunque una modalità dell'essere che non può essere compresa nei termini di una identità stabile, ma che si definisce attraverso la relazione che istituisce tra poli distinti. Questa concezione introduce una forma di intermedietà che non si riduce a equilibrio o mediazione nel senso comune del termine.

Per comprendere adeguatamente la portata del concetto di μεταξύ, è necessario distinguerlo con precisione da altre nozioni apparentemente affini, in particolare da quella di medietà (μέσος). Sebbene entrambi i termini possano essere tradotti, in modo generico, come forme di "intermedio", essi rinviano a strutture concettuali profondamente diverse[18]. La medietà designa infatti una posizione centrale o equilibrata tra due estremi. Essa implica una forma di stabilizzazione: ciò che

17 - Platone. *Simposio*, 204b, Bompiani, Milano 2014, p. 183.

18 - Per una ricostruzione del concetto di μεταξύ nella tradizione platonica e delle sue rielaborazioni contemporanee, si veda innanzitutto Platone, *Simposio*, 204b, in particolare il discorso di Diotima, dove il μεταξύ viene tematizzato attraverso la figura di Eros come realtà intermedia non riducibile a una sintesi tra opposti. Per una discussione sistematica del concetto nel pensiero contemporaneo, cfr. M. Martino, *Sul metaxy in Platone: Un itinerario,* Guerini Scientifica, Milano 2022; E. Gabellieri, *Le phénomène et l'en-*

è "medio" si colloca tra poli opposti secondo una logica di bilanciamento, nella quale la tensione tra i termini viene in qualche modo ricondotta a un punto di equilibrio. Il μεταξύ, al contrario, non designa né un centro né un punto di equilibrio. Esso non si configura come una posizione intermedia determinata, ma come ciò in virtù di cui la relazione tra termini distinti può darsi, senza essere ridotto né a una collocazione spaziale né a una funzione di mediazione stabilizzante. Il "tra" non è qui ciò che si situa a metà strada tra due termini, ma ciò che rende possibile la loro relazione senza annullarne la differenza. Questa distinzione è cruciale, poiché consente di evitare una lettura del μεταξύ in termini di sintesi o di composizione. Se la medietà tende a risolvere la tensione tra gli estremi in una forma unitaria, il μεταξύ mantiene tale tensione come elemento costitutivo della relazione. Esso non elimina l'opposizione, né la supera, ma la attraversa, rendendo possibile un rapporto che non si traduce in una riduzione a unità.

La distinzione rispetto alla medietà consente di mettere a fuoco un tratto ancora più fondamentale del μεταξύ: la sua irriducibilità a un contenuto determinato. Il μεταξύ non può essere inteso come un ente, né come un luogo determinato all'interno di una struttura già data. Ogni tentativo di identificarlo come un elemento positivo – come qualcosa che "sta in mezzo" nel senso di una presenza localizzabile o descrivibile indipendentemente – finisce per tradirne la natura. Il μεταξύ non è ciò che occupa uno spazio tra due termini, ma ciò in virtù di cui tale spazio si apre come relazione. La sua funzione è dunque eminentemente strutturale. Esso rende possibile la relazione tra poli distinti, ma senza annullarne la differenza né ricondurli a un'unità superiore. Il μεταξύ non può essere compreso come un momento intermedio all'interno di un processo che conduce a una sintesi, ma come ciò che mantiene aperta la tensione tra i termini, rendendo possibile il loro reciproco riferimento senza risolverlo.

Proprio per questo, il μεταξύ non si lascia determinare in modo esaustivo. Se viene trattato come un oggetto, un livello o una struttura già definita, esso cessa di svolgere la funzione che gli è propria. Il suo statuto è infatti inseparabile dalla dinamica relazionale che istituisce: esso non è qualcosa che si aggiunge ai termini, ma la condizione stessa del loro rapporto. In questo senso, ogni tentativo di oggettivarlo – di trasformarlo in un contenuto determinato – ne neutralizza la portata teorica, riducendolo a uno degli elementi che dovrebbe invece rendere in relazione.

Alla luce di quanto precede, il μεταξύ può essere caratterizzato attraverso una duplice funzione, che ne definisce la specificità: esso opera al tempo stesso come principio di separazione e come principio di connessione. Da un lato, il μεταξύ

tre-deux: Pour une métaxologie, Hermann, Paris 2019; M. Marianelli (ed.), *Entre: La relazione oltre il dualismo metafisico*, Città Nuova, Roma 2020.

mantiene la distinzione tra i termini in relazione. Esso non elimina la differenza né la attenua, ma la preserva come condizione stessa del rapporto. I poli non vengono fusi né ricondotti a un'unità indifferenziata, ma restano irriducibili l'uno all'altro. Dall'altro lato, il μεταξύ rende possibile la loro connessione. La separazione non si traduce in isolamento, ma si accompagna a una dinamica relazionale che consente ai termini di entrare in rapporto senza perdere la propria specificità. La relazione non è esterna ai termini, ma costituisce il modo stesso in cui essi si danno. Ciò che caratterizza il μεταξύ è precisamente l'inscindibilità di queste due funzioni. Connessione e separazione non rappresentano momenti successivi di un processo che conduce a una sintesi, ma coesistono come dimensioni strutturali della relazione. Il μεταξύ non risolve la tensione tra i poli, ma la mantiene operativa, rendendo possibile un rapporto che non si traduce in una riduzione a unità.

In questo senso, il μεταξύ può essere ulteriormente chiarito attraverso il riferimento alla nozione di "vuoto", a condizione però di precisarne il significato. Il "tra" non deve essere inteso come un vuoto nel senso di una mancanza o di una semplice assenza di contenuto. Il μεταξύ designa piuttosto una forma di "vuoto" nel senso di una non-saturazione della relazione: non un nulla, ma ciò che impedisce alla relazione di chiudersi in una determinazione già data. Proprio perché non è determinato come contenuto, esso non impone una forma alla relazione, ma la rende possibile mantenendola aperta. Ogni tentativo di riempire questo spazio – di determinarlo come un livello, un oggetto o una struttura già data – finisce per chiuderlo, trasformando ciò che dovrebbe rendere possibile la relazione in uno dei suoi termini. Il μεταξύ non è un vuoto da colmare, ma l'apertura stessa in cui la relazione può avere luogo. Esso non precede i termini come una struttura data, né li segue come effetto della loro interazione, ma coincide con la loro esposizione reciproca, che non si lascia saturare da alcuna determinazione.

5. Il μεταξύ e la struttura epistemica della transdisciplinarità

Il riferimento al μεταξύ non introduce una nuova ontologia della relazione, ma rende esplicita una condizione che le pratiche conoscitive contemporanee già presuppongono senza tematizzarla. Non si tratta altresì di sostenere che la realtà sia relazionale e che, per questo motivo, debba essere interpretata attraverso la nozione di μεταξύ.

L'argomento procede in una direzione diversa. Ciò che emerge, piuttosto, è una convergenza tra piani distinti di analisi. Da un lato, la riflessione filosofica sulla transdisciplinarità ha messo in evidenza i limiti di una concezione della conoscenza come composizione di descrizioni parziali, mostrando la necessità di pensare uno spazio intermedio in cui tali descrizioni possano entrare effettivamente in relazione. Dall'altro lato, diverse pratiche scientifiche e strutture teoriche sembrano in-

contrare difficoltà analoghe, nella misura in cui non possono essere comprese adeguatamente nei termini di proprietà intrinseche attribuite a entità isolate. In questo senso, il problema del "tra" non si presenta come una costruzione teorica *ad hoc,* ma come una questione che emerge da più direzioni convergenti. Il riferimento al μεταξύ consente di articolare concettualmente questa difficoltà, senza pretendere di risolverla attraverso una deduzione ontologica, ma offrendo uno schema interpretativo capace di rendere esplicita una tensione già presente nelle pratiche della conoscenza.

Alcuni ambiti della ricerca scientifica contemporanea offrono esempi significativi di questa difficoltà. Senza entrare in questioni ontologiche più ampie, è possibile osservare come, in diversi contesti, la descrizione dei fenomeni non possa essere articolata nei termini di proprietà intrinseche attribuite a entità isolate, ma richieda un riferimento essenziale alle relazioni. Un caso emblematico è rappresentato dalla meccanica quantistica nella sua interpretazione relazionale[19]. In questo contesto, le proprietà fisiche non sono intese come caratteristiche assolute dei sistemi, ma come grandezze definite relativamente a specifiche interazioni. Ciò non implica una rinuncia alla descrizione fisica, ma segnala che tale descrizione non può essere costruita a partire da proprietà assegnate indipendentemente dal contesto relazionale in cui esse emergono[20]. Un'analoga difficoltà si manifesta nello studio dei sistemi cosiddetti *complessi*. In questi ambiti, il comportamento dell'insieme non è riducibile alla somma dei comportamenti delle parti, anche quando queste siano note con precisione. Le dinamiche osservate dipendono infatti da reti di interazione che non possono essere ricondotte a una semplice aggregazione di elementi, ma richiedono una descrizione che tenga conto delle relazioni come dimensione costitutiva del fenomeno. Questi esempi, pur tra loro eterogenei, convergono su un punto: la conoscenza non può essere intesa come composizione di descrizioni relative a entità isolate. Le descrizioni risultano invece significative solo nella misura in cui articolano configurazioni relazionali, nelle quali le proprietà acquistano significato in funzione del contesto in cui si danno[21].

---

19 - C. Rovelli, *Relational quantum mechanics*, in «International Journal of Theoretical Physics», 35 (1996) 8, pp. 1637-1678.

20 - Per una discussione sistematica di questa impostazione in termini di proprietà relazionali, cfr. V. Fano - M. Sanchioni, *Relational properties and relational quantum mechanics*, in «Foundations of Physics», 55 (2025) 4, p. 52; per il quadro generale, cfr. C. Rovelli, *Relational quantum mechanics*, in «International Journal of Theoretical Physics», 35 (1996) 8, pp. 1637-1678.

21 - Una possibile linea di interpretazione di questa centralità delle relazioni è offerta dalle ontologie di processo, in cui le entità non sono concepite come sostanze dotate di proprietà intrinseche, ma come nodi di relazioni e dinamiche. In questa prospettiva, cfr. A.N. Whitehead, Process and Reality, Free Press, New York 1978 (orig. 1929). Il rapporto tra tali approcci ontologici e la nozione di spazio relazionale qui

Il riferimento a strutture relazionali nelle pratiche scientifiche può essere ulteriormente precisato considerando il ruolo dell'indeterminatezza nella meccanica quantistica. In questo contesto, le proprietà dei sistemi non risultano sempre definite indipendentemente dalle interazioni attraverso cui vengono stabilite. La loro determinazione dipende dal contesto fisico in cui il sistema è coinvolto, e non può essere semplicemente presupposta come dato iniziale della descrizione[22]. Questa caratteristica è stata interpretata in modi filosoficamente distinti. In una prima lettura, essa viene ricondotta a un limite epistemico, come se l'indeterminatezza riflettesse soltanto una restrizione nella nostra capacità di conoscere proprietà altrimenti ben definite. In una seconda prospettiva, più radicale, l'indeterminatezza viene invece intesa in senso ontologico, come indicazione del fatto che le proprietà non sono pienamente determinate in sé, indipendentemente da ogni contesto relazionale.

Tuttavia, entrambe queste interpretazioni risultano insufficienti a rendere conto del modo in cui le descrizioni quantistiche sono effettivamente costruite. Più precisamente, esse presuppongono ancora che la determinazione delle proprietà possa essere pensata in modo assoluto, sia pure come epistemicamente inaccessibile o ontologicamente incompleta. Ciò che emerge dalle pratiche teoriche è invece un limite più profondo: non è possibile articolare la conoscenza a partire da proprietà attribuite a sistemi considerati isolatamente[23]. In questa prospettiva, l'indeterminatezza non segnala un limite del conoscere, ma il venir meno del presupposto secondo cui le proprietà possano essere pensate indipendentemente dalle relazioni che le determinano. Le proprietà non risultano indeterminate simpliciter, ma determinate soltanto relativamente a specifiche interazioni. L'indeterminatezza non è quindi il contrario della determinazione, ma il segno negativo di una struttura relazionale della determinazione stessa. In questo senso, l'indeterminatezza non mette semplicemente in questione la possibilità di misurare o prevedere grandezze fisiche, ma il modello compositivo della conoscenza. Non è possibile, in generale, procedere descrivendo separatamente le proprietà dei singoli sistemi per poi integrar-

proposta richiederebbe tuttavia un'analisi specifica, che esula dagli obiettivi del presente lavoro e verrà sviluppata in contributi successivi.

22 - Questa lettura è tipica delle interpretazioni relazionali della meccanica quantistica, nelle quali la determinazione delle proprietà è intrinsecamente dipendente dal contesto di interazione. Altri approcci offrono tuttavia inquadramenti differenti del problema, interpretando l'indeterminatezza in termini di indeterminazione metafisica (cfr. C. Calosi - J. Wilson, *Quantum metaphysical indeterminacy,* in «Philosophical Studies», 176 (2019) 10, pp. 2599-2627) oppure di disposizioni o potenzialità (cfr. M. Dorato, *Bohr meets Rovelli: A dispositionalist account of the quantum limits of knowledge*, in «Quantum Studies: Mathematics and Foundations», 7 (2020) 2, pp. 233-245). Nel presente lavoro non si intende entrare nel merito di tale dibattito, poiché l'argomento sviluppato richiede soltanto di segnalare la possibilità di una lettura relazionale della determinazione delle proprietà.

23 - Sempre avendo in mente l'interpretazione relazionale della Meccanica Quantistica.

le in una descrizione complessiva del fenomeno. La determinazione delle proprietà avviene nel contesto delle relazioni, e non può essere ricostruita come risultato della semplice combinazione di descrizioni parziali.

Ne segue che la conoscenza non può essere pensata come un processo che parte da elementi isolati per giungere, attraverso successive integrazioni, a una comprensione più ampia. Al contrario, essa richiede di essere articolata fin dall'inizio nello spazio relazionale in cui le determinazioni prendono forma. In questo senso, il problema del "tra" non è un momento successivo dell'elaborazione conoscitiva, ma la condizione entro cui essa può darsi. Le considerazioni precedenti consentono di chiarire il ruolo del μεταξύ come modello dello spazio epistemico in cui la conoscenza si articola. Se, da un lato, le pratiche scientifiche mostrano i limiti di una descrizione compositiva, e, dall'altro, l'indeterminatezza mette in questione la possibilità di attribuire proprietà indipendenti a sistemi isolati, diventa necessario pensare uno spazio in cui le determinazioni possano emergere senza essere presupposte. In questo senso, il μεταξύ non può essere identificato né con una forma di indeterminatezza, intesa come assenza o mancanza di determinazione, né con una struttura già data che organizza dall'esterno la relazione tra prospettive. Esso si configura piuttosto come lo spazio in cui prospettive differenti entrano effettivamente in relazione e in cui, proprio attraverso tale relazione, si producono configurazioni determinate. Il "tra" non è quindi ciò che resta quando le determinazioni mancano, ma ciò in cui esse diventano possibili. Le descrizioni parziali non costituiscono unità isolate che vengono successivamente messe in relazione, ma si danno fin dall'inizio all'interno dello spazio relazionale. Esse non precedono il "tra", ma ne sono articolazioni: il loro significato non deriva da una integrazione successiva, ma dal rapporto che le costituisce. In questo senso, la conoscenza non procede da descrizioni isolate verso una sintesi successiva, ma si costituisce fin dall'inizio nello spazio relazionale che le connette. Il μεταξύ può così essere interpretato come il luogo epistemico in cui si dà il passaggio da descrizioni parziali a configurazioni di senso. Non si tratta di un passaggio lineare o cumulativo, ma di una dinamica relazionale in cui le determinazioni emergono come esito del loro reciproco riferimento, senza essere riducibili né ai singoli contributi né a una struttura predefinita.

Alla luce di quanto emerso, è possibile precisare in modo più rigoroso il significato della transdisciplinarità. Se, in determinati ambiti, la conoscenza non può essere articolata come composizione di descrizioni relative a entità isolate – in particolare quando le determinazioni dipendono costitutivamente da contesti di interazione – allora anche il ruolo delle discipline deve essere riconsiderato. Esse non descrivono semplicemente "parti del mondo" che possono essere successivamente integrate, ma articolano dimensioni differenti – e tra loro non riducibili – di fe-

nomeni la cui intelligibilità dipende dal contesto relazionale in cui tali dimensioni si collocano. In questo senso, le descrizioni disciplinari non costituiscono unità isolate che vengono messe in relazione in un secondo momento, ma si danno fin dall'inizio all'interno di uno spazio relazionale che ne condiziona il significato.

In questo quadro, la conoscenza non si configura come un processo additivo, nel quale contributi disciplinari distinti vengono progressivamente sommati, ma come una dinamica in cui le determinazioni emergono nel rapporto tra prospettive eterogenee. Il "tra" non è un momento successivo dell'integrazione, ma la condizione entro cui le descrizioni acquistano senso quando il modello compositivo risulta insufficiente. In tali casi, il riferimento allo spazio relazionale non costituisce una opzione interpretativa tra le altre, ma un vincolo epistemico: è ciò che rende possibile l'articolazione stessa della conoscenza.

Ne consegue che la transdisciplinarità non può essere intesa né come un semplice metodo di coordinamento tra saperi, né come una forma di integrazione funzionale orientata alla soluzione di problemi, come suggerito rispettivamente in (M2) e in (TS). Al tempo stesso, essa non può essere ricondotta a una struttura epistemica già determinata, fondata su una articolazione ontologica dei livelli di realtà, come avviene in (EPI). La transdisciplinarità designa piuttosto una condizione della conoscenza che emerge quando le determinazioni non possono essere ricondotte a descrizioni isolate, ma si costituiscono nella relazione tra di esse. In questo senso, essa non si identifica né con una funzione operativa né con una struttura ontologica predefinita, ma con lo spazio in cui tali relazioni diventano effettivamente produttive di conoscenza.

In questa prospettiva, la transdisciplinarità può essere intesa come la forma epistemica del μεταξύ: non un dispositivo di integrazione dei saperi, ma lo spazio in cui essi si espongono reciprocamente senza convergere in una sintesi, rendendo possibile una determinazione che non appartiene a nessuno di essi in quanto tale[24].

24 - La prospettiva qui delineata solleva una questione che meriterebbe di essere sviluppata ulteriormente, ovvero il suo possibile rapporto con la nozione di strong objectivity elaborata nell'ambito dell'epistemologia femminista, in particolare da S. Harding (cfr. S. Harding, *Rethinking standpoint epistemology: What is "strong objectivity"?*, in L. Alcoff - E. Potter (edd.), *Feminist epistemologies,* Routledge, New York 2013, pp. 49-82). Con tale espressione si intende una concezione dell'oggettività che, anziché prescindere dalle condizioni storiche e sociali della conoscenza, le assume esplicitamente come parte integrante del processo conoscitivo, richiedendo una riflessività critica sulle posizioni da cui si conosce. In questo senso, l'idea che la conoscenza emerga nello spazio relazionale tra prospettive differenti presenta alcune affinità con tale impostazione. Tuttavia, il presente lavoro non intende sviluppare questo confronto, limitandosi a segnalarne la rilevanza.

6. Conclusioni

Il presente lavoro ha mostrato che il problema della transdisciplinarità non riguarda soltanto il coordinamento tra saperi, ma la struttura stessa della conoscenza. Se gli oggetti della ricerca contemporanea presentano una natura intrinsecamente relazionale, allora la conoscenza non può essere intesa come composizione di descrizioni indipendenti, ma deve essere pensata nello spazio intermedio in cui tali descrizioni entrano in relazione.

In questa prospettiva, si è sostenuto che la difficoltà principale delle teorie della transdisciplinarità consiste nella mancata tematizzazione esplicita di questo spazio del "tra", che viene spesso presupposto ma non concettualizzato. Il ricorso al concetto di μεταξύ ha consentito di articolare tale spazio come una struttura relazionale non oggettivabile, capace di connettere senza sintetizzare e di rendere possibile l'emergere di determinazioni senza essere ridotta a una struttura già data. Ne è derivata una reinterpretazione della transdisciplinarità non come metodo o integrazione, ma come forma epistemica del μεταξύ: il modo in cui la conoscenza si produce quando le determinazioni emergono nella relazione tra prospettive eterogenee.

Questa impostazione ha inoltre mostrato una convergenza significativa con alcune pratiche scientifiche contemporanee, nelle quali la descrizione dei fenomeni non può essere articolata nei termini di proprietà intrinseche attribuite a entità isolate, ma richiede un riferimento costitutivo alle relazioni. In questo senso, il "tra" non appare come un vuoto da colmare né come una struttura da presupporre, ma come lo spazio in cui la conoscenza prende forma.

Le implicazioni di questa analisi non riguardano soltanto l'organizzazione della ricerca, ma la comprensione stessa della conoscenza. Se il sapere si costituisce nello spazio relazionale del μεταξύ, allora l'unità della conoscenza non può più essere pensata nei termini della sintesi o dell'integrazione, ma come articolazione di una pluralità che si mantiene tale nella relazione. Se la conoscenza contemporanea nasce sempre più spesso nello spazio relazionale tra discipline, allora anche le istituzioni della ricerca devono essere ripensate alla luce di questa trasformazione. Non si tratta semplicemente di promuovere collaborazioni interdisciplinari o di sviluppare nuovi strumenti metodologici, ma di creare contesti in cui lo spazio del "tra" possa essere effettivamente abitato come luogo epistemico. La transdisciplinarità richiede, in questo senso, ambienti istituzionali capaci di sostenere processi di conoscenza che non si sviluppano all'interno di singole discipline, ma emergono nella relazione tra di esse. Si tratta di costruire spazi in cui il dialogo tra saperi non sia un momento accessorio o successivo, ma la condizione stessa della produzione di conoscenza.

In termini di prospettive future, un possibile sviluppo di questa linea di ricerca consiste nell'approfondire il rapporto tra il concetto di μεταξύ e il dibattito contemporaneo sull'unità della scienza. In particolare, sarà necessario chiarire in che misura la nozione di "spazio relazionale" qui proposta possa offrire una alternativa alle concezioni tradizionali dell'unità del sapere, fondate sulla riduzione o sull'integrazione tra discipline, aprendo invece alla possibilità di una articolazione non compositiva della conoscenza. Parallelamente, un ulteriore sviluppo riguarderà l'analisi del rapporto tra il μεταξύ, alcune interpretazioni della meccanica quantistica – in particolare la Relational Quantum Mechanics – e una più ampia riflessione sull'ontologia relazionale.

MARCO MARTINO
Professore associato di Storia della filosofia presso il Dipartimento di Filosofia, Scienze Sociali, Umane e della Formazione (FISSUF) dell'Università degli Studi di Perugia
marco.martino@unipg.it

MARCO SANCHIONI
Professore stabile straordinario di Filosofia della Scienza presso l'Istituto Universitario Sophia
marco.sanchioni@sophiauniversity.org

Phenomenological Elements in Carnap's Aufbau: the Epistemological Role of Sensible Intuition

di Vincenzo Fano - Niccolò Covoni

The Carnap of Der logische Aufbau der Welt *was aware of the problem of "imperfect community." Nonetheless he ignores it. We argue that Carnap's tendency to downplay this difficulty is best understood in light of the enduring influence of Husserlian phenomenology on his early philosophy. On this reading, the* Aufbau *preserves an epistemological framework partially grounded in the structures of subjective experience, already at the heart of* Der Raum.

Keywords: Rudolf Carnap; Aufbau; Structural Empiricism; Sensible Intuition; Quasi-Analysis.

1. Introduction

Rudolf Carnap's first majot work is *Der Logische Aufbau der Welt* (*The Logical Structure of the World*, hereafter *Aufbau*)[1]. This seminal text has generated extensive debate and diverse interpretations. In this paper, we propose a novel reading of *Aufbau* through the lens of "structural empiricism," an epistemological position that emphasizes the formative role of experiential structures in knowledge without grounding justification solely in empirical data. Our central claim is that Carnap, in his attempt to outline a general constructive system via the method of "quasi-analysis," anticipates this position. By emphasizing the structural organization of experience, particularly through the notion of recollection of partial similarities, Carnap offers a sophisticated response to the problem of the relation between logic and experience. We define structural empiricism as a view characterized by two key principles:

- The primacy of structural features of experience in shaping epistemic justification.
- The impossibility of justifying propositions through intuition, despite its necessary role from an epistemological point of view.

A conceptual parallel to this approach is found in Angelucci,[2] who emphasizes the distinction between the state of being justified in holding a belief and the activity of *justifying a belief*. This distinction suggests that philosophical reasoning frequently draws on intuitive or experiential structures rather than formal inferential justification. Stegmüller[3] similarly highlights the epistemological role of self-evidence (*Evidenz*), treating it as a subjective but indispensable dimension of reasoning, particularly in mathematics and philosophy. This paper argues that Aufbau implicitly aligns with this epistemological stance. While drawing from empiricism, Kantianism, and Phenomenology, Carnap's project defies reduction to any single tradition. Instead, it presents an original and nuanced synthesis that privileges structured experience without appealing to metaphysical foundations. Understanding science from Carnap's perspective requires conceptualizing knowledge through auto-psychological experience, even though such experience cannot be the direct object of scientific inquiry.[4] Only through the logical construction

1 - R. Carnap, *The Logical Structure of the World*, University of California Press, Berkeley 1967.

2 - A. Angelucci, *On Justifying Case Verdicts. A Dialectical Hypothesis*, in «Inquiry», 68 (2022), n. 2, pp. 297-317.

3 - W. Stegmüller, *Metaphysik Skepsis Wissenschaft*, Springer-Verlag, Berlin 2013

4 - Although the present paper engages with the historical background of Carnap's early thought, our aim is primarily philosophical rather than strictly historiographical. We do not claim that the interpretive the-

of empirical constituents can abstract scientific concepts be rendered intelligible. The epistemological interpretation we propose also addresses some historical critiques of Carnap's project, most notably those by Nelson Goodman[5] and Quine.[6] Goodman questioned the logical adequacy of the quasi-analysis method, particularly its reliance on similarity relations, while Quine challenged the possibility of reducing physical-object language to a sensory base. Although we explore these objections in detail in Section 5, it is important to flag them early on, as they have shaped much of the secondary literature's skepticism toward the *Aufbau*. Our reinterpretation of Carnap's method as structurally empiricist offers a way to answer these critiques without abandoning the core insights of his constructional system. The structure of this paper is as follows: Section 2 introduces the framework of structural empiricism through the works of Angelucci and Stegmüller. Section 3 examines the main philosophical influences on *Aufbau*. Section 4 analyzes the internal mechanics of Carnap's quasi-analysis. Section 5 discusses the rehabilitations of *Aufbau* by Leitgeb and Mormann. Section 6 concludes by arguing that Carnap's young project is best understood through the epistemological lens of structural empiricism.

2. Structural Empiricism

Before discussing the interpretation of Carnap's work, it is necessary to define what we mean by structural empiricism. This position can be clarified by reference to Adriano Angelucci's *On Justifying Case Verdicts: A Dialectical Hypothesis.*

They develop a dialectical hypothesis analyzing the method of cases (hereafter MOC) in philosophical theorizing[7]. MOC is a technique in which philosophers examine specific scenarios to explore philosophical problems. Characteristically, this technique involves an appeal to intuition, where "a case is typically taken as decisive in supporting or undermining a given theory (thesis, or analysis) in virtue of a certain proposition purportedly holding true of it"[8]. Angelucci presents a decisive

ses advanced here apply to Carnap's work as a whole, nor that they capture his later views in logical empiricism. Rather, the claims developed in this paper are intended to apply, at most, to *Der logische Aufbau der Welt* and to the specific epistemological framework articulated therein.

5 - N. Goodman, The Structure of Appearance, Bobbs-Merrill, Indianapolis 1966.

6 - W.V.O. Quine*, Two Dogmas of Empiricism*, in «The Philosophical Review», 60 (1951), pp. 20-43.

7 - E. Machery, *Philosophy Within Its Proper Bounds*, Oxford University Press, Oxford 2017.

8 - A. Angelucci, *On Justifying Case Verdicts. A Dialectical Hypothesis*, in «Inquiry», 68 (2022), n. 2, p. 2.

challenge to the arguments of Deutsch[9] and Cappelen[10], who claim that when philosophers argue for their case verdicts, they always provide reasons for accepting them. This supports the view that philosophers do not regard intuitions as their justificatory source, a view known as the *Argument View*. Angelucci does not dispute that philosophers offer arguments. However, they argue that this alone is insufficient to conclude that these arguments are regarded as sources of justification tout court. Instead, Angelucci proposes a meaningful distinction between the activity of "justifying a belief" and the state of "being justified in holding a belief." According to the Argument View, philosophers provide arguments as reasons for their beliefs, which constitute evidence or justification. As Deutsch puts it,

> In general, if one gives an explicit argument for p and does no more by way of indicating why p ought to be accepted, then one can be taken as having provided what one takes to be the justification or evidence for p[11].

Arguing for one's beliefs, therefore, is typically interpreted as offering reasons as the main justification. However, in epistemology, a key distinction exists between being justified in holding a belief (a state) and justifying that belief (an activity). The former concerns one's current epistemic position, while the latter refers to a potentially fallible process of reason-giving. According to Angelucci, this distinction is crucial to understanding MOC because "justifying a given belief is not always perceived as a necessary condition for being justified in holding it; at times, that is, reason-giving is not seen as required for justification"[12]. This observation undermines the Argument View, suggesting that while philosophers may offer arguments, they do not necessarily regard them as the justificatory basis for their beliefs. Instead, their justification may rest on other grounds. Argumentation can serve dialectical purposes, aimed at persuasion or consensus-building, rather than providing the epistemic foundation for belief. This perspective allows for an interpretation of MOC that accommodates the idea that intuitions remain a component of justification. Thus, Angelucci contends that philosophers employing MOC may not treat their stated reasons as justificatory sources but continue to hold beliefs accompanied by intuitive states. These beliefs may align with

---

9 - M. Deutsch, *Experimental Philosophy and the Theory of Reference*, in «Mind and Language», 24 (2009), n. 4, pp. 445-466.

10 - H. Cappelen, *Philosophy Without Intuitions*, Oxford University Press, Oxford 2012.

11 - M. Deutsch, *Gettier's Method*, in J. Nado (ed.*), Advances in Experimental Philosophy and Philosophical Methodology*, Bloomsbury Academic, New York 2016, p.74.

12 - A. Angelucci, *On Justifying Case Verdicts. A Dialectical Hypothesis*, in «Inquiry», 68 (2022), n. 2, p. 7.

prevailing views[13] and need not always be explicitly justified. Philosophical justification, in this case, becomes a dialectical act aimed at fostering agreement, rather than demonstrating formal proof. It is true that a thesis cannot be justified solely by an appeal to intuition, but it is also true that intuitions cannot be eliminated from the reasons for holding a thesis. While fallible, these processes still retain cognitive value.

Although Angelucci's argument suggests that intuition plays a role as a source of justification, our claim is somewhat more moderate. We maintain that intuition cannot be employed in the process of justifying something, as it cannot be falsified; however, intuition remains essential to the process of knowledge, and for that reason, it is indispensable. Consider, for example, a simple auditory scenario with three tones: A, B, and C. Tones A and B share a similar timbre. In contrast, A and C share the same pitch, and finally, C and B have no evident equivalence. Even in this context, the human brain can immediately recognize the similarity in frequency between B and C without needing to resort to the whole equivalence class. Our auditory system is capable of capturing directly patterns and universals, often from a single situation, highlighting the immediacy of sensory recognition. This phenomenon mirrors Brentano's idea that induction can arise from a single case, which gives intuition a crucial epistemological role without relying on the necessity of multiple experiences.[14] A paradigmatic case is touching the fire with a hand; right after the first experience, we can say that fire will hurt us. Similarly, in visual perception, encountering only aa few chairs may be sufficient to apprehend the general category of "chair" or its archetypal form. These examples illustrate how justification can be grounded in immediate, single-case experiences rather than in large classes of repeated instances for recognition and full demonstration. To better understand the point raised by Angelucci, it is helpful to consider another scholar who contributes to the epistemological foundation of structural empiricism and who also grants significant importance to intuition: Stegmüller.[15] In his work, they examine the epistemological function of self-evidence (*Evidenz*), highlighting that this concept resists rigorous definition. According to Stegmüller, self-evidence cannot be captured through an exhaustive formula but must instead be illustrated with concrete examples. This suggests that evidence should be understood more as a phenomenological experience than as a formalizable concept. Despite these limitations, Stegmüller insists that self-evidence is indispensable in knowledge practices. It arises in diverse contexts, including sensory perception, mathematical

13 - With the prevailing view, we intend the most dominant position.

14 - F. Brentano, *Psychologie vom empirischen Standpunkt*, Duncker & Humblot, Leipzig 1874.

15 - W. Stegmüller, *Metaphysik Skepsis Wissenschaft*, Springer-Verlag, Berlin 2013.

proof, and philosophical argumentation. In each of these domains, we frequently rely on what strikes us as "evident." A paradigmatic example is found in mathematics. When subject A explains a proof to subject B, A cannot justify a step merely by stating, "this step is self-evident." Nevertheless, A will often depend on B's intuitions, guiding them toward a direct, immediate grasp of the reasoning's validity. The structural empiricism proposed by Angelucci and Stegmüller reframes the role of intuition not as a substitute for rigorous argumentation, but as a complementary and empirically grounded source for belief formations. Intuition, in this view, plays a valuable epistemic role in shaping our judgments and conceptual frameworks. The notion of intuition in structural empiricism is analogous to the "logic of discovery", as articulated by Lakatos in Proofs and Refutations: *The Logic of Mathematical Discovery.*[16] According to Lakatos, this process is not purely deductive, but it involves heuristic and intuitive insights that guide the development of mathematical concepts and proofs. Similarly, in structural empiricism, intuition does not serve as a basis for inferential justification but plays a role in shaping the structure of experience and enabling scientific abstraction. For this reason, we argue that the role of intuition is local rather than global: it cannot be dismissed, yet it does not function as a source of justification in the strict sense, as Angelucci rightly emphasizes. Nonetheless, it remains a crucial part of the epistemological process. What we have in mind is illustrated in cases of analogical reasoning. For instance,

> In the case of the Euler characteristic formula, experience may suggest to us that a rigorous proof for the three-dimensional case will almost certainly involve removing and adding vertices, edges, and faces. In all such cases, broadly speaking, we retain the basic intuition that the analogy exhibits reasonable potential for generalization. If we decide to persist with the analogy, then our model provides clear guidance about how it may be restricted or reformulated in order to meet our criteria of plausibility.[17]

Bartha analyzes how analogies can motivate generalizations and conjectures about explanatory principles. He highlights that the role of analogy is not only a tool for illustrating facts, but it also helps to formulate open questions and guide the direction of research. This generative aspect is captured in what Bartha calls the plausibility argument: analogies may justify tentative intuitions about the applicability of a model, even in the absence of complete theoretical or empirical

---

16 - I. Lakatos, Proofs and Refutations: The Logic of Mathematical Discovery, Cambridge University Press, Cambridge 1976.

17 - P. Bartha, By Parallel Reasoning: The Construction and Evaluation of Analogical Arguments, Oxford University Press, Oxford 2010, p.198.

support. This interplay between analogy and intuition becomes even more significant when expert knowledge is involved. In these situations, intuition plays a unique role in the reasoning process. Although expert intuition can be fallible, it helps identify significant structural similarities and modal constraints that extend beyond superficial appearances[18]. This kind of expert-guided intuition reflects the structural features of experience that structural empiricism aims to emphasize: not as a basis for formal justification, but as a necessary component in the discovery and formation of scientific concepts.

3. The influences of Carnap's philosophy

The main goal of this paper is to show that the epistemological position outlined in the previous section closely resembles Carnap's proposal of quasi-analysis. However, this perspective significantly differs from the three main philosophical positions that influenced Carnap's *Aufbau*. In this section, we briefly discuss and compare these positions with structural empiricism, highlighting also the key differences that emerge from Carnap's view. The first position to consider is empiricism. According to this view, all knowledge is derived from sensory experience. It rejects the notion of innate ideas or a priori knowledge independent of perception. Mental representations or scientific notions are considered generalizations formulated from data provided by the senses. Within this framework, any theoretical construction lacks absolute truth and is instead regarded as a useful convention for organizing and simplifying our experience of the world. This position is championed by Ernst Mach,[19] the Austrian physicist and philosopher who had a profound influence on Carnap and is frequently cited by him. Mach believed that science should remain rooted in the immediacy of sensory experience. Scientific concepts, he argued, should serve merely as tools for organizing perceptions rather than as representations of metaphysical entities. As Coffa notes in the analysis of Mach's thoughts: "The truth in science was entirely a matter of sense and sensibility, and the farther science went in the direction of insensible human postulation, the farther it departed from truth and reality".[20]

Empiricism thus aims to identify the invariants in experience, those regularities that recur and can be described by natural laws. Yet, it grants no epistemological

18 - For a detailed discussion of the role of expert intuition in analogical reasoning, see Chapters 5 and 6 of Bartha.

19 - E. Mach, *The Analysis of Sensations, and the Relation of the Physical to the Psychical*, Open Court Publishing Company, La Salle 1914.

20 - A. Coffa, *The Semantic Tradition from Kant to Carnap: To the Vienna Station*, edited by L. Wessels, Cambridge University Press, New York 1991, p.191.

weight to subjective understanding. Theories do not explain reality "in itself" but provide a pragmatic and economical means of describing observations. Empiricism denies the possibility of grounding knowledge in principles independent of observation. Science, for empiricism, does not uncover ultimate truths but constructs effective models to navigate the complexity of the phenomenal world. As Carus notes in this framework: "it was pretty much taken for granted that subjectivity is, ultimately, part of nature, of which science gives us a larger and more systematic picture than any individual's collection of subjective experiences".[21] By contrast, structural empiricism maintains that there are structures within our experience, such as intuitions and self-evidence, that play a crucial role in our theorizing. These structures constrain and, in a sense, define the theories we build. Before examining another philosophical position that shaped Carnap's thought, a key distinction must first be drawn. In the phenomenological tradition, two kinds of intuition are usually distinguished: apodictic and assertoric evidence. This classification originates in Franz Brentano's *Psychologie vom empirischen Standpunkte*[22]. The kind of intuition that concerns us in structural empiricism, and that plays a decisive role in shaping the structure of experience, is the apodictic one. Brentano illustrates what he calls *apodiktisch Urteil* with the following example:

> If someone says the impossibility of a round square exists someone who carelessly heeds only the grammatical form may think that he is affirming something, while in fact he is negating something, since he is apodictically denying the round square. In so doing he is thinking about the round and the square and in both cases the reference has a thing as its object[23].

Apodictic intuition does not merely present a proposition as evidently true; it also reveals the impossibility of its negation, in a form of phenomenological necessity. It must be regarded as a pre-theoretical certainty that informs how experience is structured and how objects are constituted at the level of experience. By contrast, intuitions are not to be understood in the sense of what Brentano called *assertorisch Urteil*:

> All men are mortal.' This proposition, therefore, seems to affirm the conjunction of man and mortal, only on the assumption that a man exists. Yet a glance at the existential proposition which is equivalent to the categorical

21 - A. W. Carus, *Carnap and Phenomenology: What Happened in 1924*?, in C. Damböck (ed.), Influences on the Aufbau, Springer, Berlin 2016, pp. 159-186.

22 - F. Brentano, *Psychologie vom empirischen Standpunkt*, Duncker & Humblot, Leipzig 1874.

23 - F. Brentano, *Psychology from an Empirical Standpoint*, edited by O. Kraus and translated by L.L. McAlister, Routledge, New York 1995, p. 252.

> proposition solves the whole difficulty. It shows that the proposition is not really an affirmation, but a denial.[24]

Assertoric evidence provides factual certainty but offers no further new information.[25] This evidence can be used in scientific reasoning, but it does not play the structural or constitutive role that the apodictic intuition has in capturing the framework of knowledge. In structural empiricism, we do not claim that intuition justifies empirical propositions inferentially; instead, apodictic evidence functions as the form of intuition that structures experience and enables scientific abstraction. As Stegmüller observes, it is a fundamental component of knowledge, even though its existence cannot be demonstrated directly.

The next step is to examine another philosophical position that influenced Carnap, namely that of Edmund Husserl.[26] Carnap explicitly acknowledges this influence in his 1922 thesis *Der Raum*,[27] particularly when defining intuitive space. He writes:

> Experience does not supply the justification for them; the axioms are independent of experience. More precisely, they are independent of the 'quantity of experience' (as Driesch puts it), i.e., knowledge of them does not, as in the case of empirical statements, become progressively more secure with repeated experience. For as Husserl has shown, this is not a question of facts in the sense of empirical reality, but rather a question of the essence (*eidos*) of certain items given to us, an essence that can be immediately grasped in its particular character even from a single instance.[28]

This deep connection between Husserl and Carnap has been highlighted by many scholars. Carus,[29] for example, characterizes the pre-1924 Carnap as a kind of "young Husserlian Carnap," emphasizing the extent to which his thinking was shaped by Husserl's phenomenology, critique of empirical psychology, and the method of constitution. A central feature of Husserl's phenomenology is the

---

24 - F. Brentano, Psychology from an Empirical Standpoint, edited by O. Kraus and translated by L.L. McAlister, Routledge, New York 1995, p.167.

25 - V. Fano, *La filosofia dell'evidenza. Saggio sull'epistemologia di Franz Brentano*, CLUEB, Bologna 1993.

26 - M. Hartimo, *Rudolf Carnap* (1891-1970), in Encyclopedia of Phenomenology, 2025.

27 - R. Carnap, *Early Writings*, edited by A. W. Carus, M. Friedman, W. Kienzler, A. Richardson and S. Schlotter, general editor R. Creath, Oxford University Press, New York 2019.

28 - *Ibid.*, 55.

29 - A. W. Carus, *Carnap and Phenomenology: What Happened in 1924*?, in C. Damböck (ed.), Influences on the Aufbau, Springer, Berlin 2016, pp. 159-186.

rehabilitation of the synthetic *a priori* in a new form, which he calls "material." This concept is introduced in the *Logische Untersuchungen*[30]:

> Each pure law, which includes material concepts, so as not to permit of a formalization of these concepts *salva veritate* - each such law, i.e., that is not analytically necessary - is a *synthetic a priori law*. Specifications of such laws are synthetic necessities: empirical specifications of course are so also, e.g. *This red is different from this green*[31].

This idea is further developed in *Erfahrung und Urteil*,[32] where Husserl argues that the material *a priori* is a general intuition intrinsic to experience. The goal of Husserl's phenomenology is to construct a foundational science based on this material *a priori*, which would underlie all other sciences. Structural empiricism, however, diverges here. It holds that experiences can never justify a sentence; at most, one may be justified in maintaining a belief because of a specific experience. The crucial point is that the process of holding a belief does not represent any inferential process but maintains its indispensability in the process of justification of a sentence. As a consequence, intuitions have an epistemological key role in the process of justification, but they cannot be taken as part of the inferential process. Therefore, intuitions play a key epistemological role in the justification process, but they cannot be regarded as part of the inferential process.

Finally, Carnap was also influenced by Neo-Kantianism.[33] Neo-Kantian thinkers such as Cassirer, Natorp, and Rickert are cited throughout the *Aufbau*, and Carnap attended the lectures of the Neo-Kantian philosopher Bruno Bauch in Jena. Both Natorp and Cassirer played crucial roles in shaping Carnap's *Der Raum*, as Fano[34] has shown. In his *Autobiography*,[35] Carnap later writes:

> Formal space is an abstract system, constructed in mathematics, and more precisely in the logic of relations; therefore, our knowledge of formal space is of a logical nature. Knowledge of intuitive space I regarded at that time, under the

30 - E. Husserl, *Logical Investigations*. Volume 2, edited by D. Moran, Routledge, New York 2001.

31 - *Ibid.*, p. 21.

32 - Id., *Erfahrung und Urteil*, Felix Meiner, Hamburg 1939.

33 - A.W. Richardson, *Carnap's Construction of the World: The Aufbau and the Emergence of Logical Empiricism*, Cambridge University Press, New York 1998.

34 - V. Fano, *L'analisi carnapiana del concetto di spazio*, in A. Pasquinelli (ed.), L'eredità di Rudolf Carnap, CLUEB, Bologna 1995, pp. 297-318.

35 - R. Carnap, *Intellectual Autobiography*, in P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, Open Court, La Salle 1963, pp. 3-84.

> influence of Kant and the neo-Kantians, especially Natorp and Cassirer, as based on 'pure intuition' and independent of contingent experience.[36]

From the Neo-Kantian perspective, reality lies beyond the reach of direct knowledge. What we perceive and understand of the world is shaped by the a priori forms of intuition (space and time) and the categories of the intellect. This mediation of subjectivity in the constitution of experience resonates in Carnap's epistemological framework. Carnap adopts and radicalizes this idea by asserting that all statements about reality are metaphysical, and, in his view, meaningless.[37] He maintains that such statements, involving ontological commitments, lack empirically verifiable content and are logically undecidable. "In other words, the 'metaphysical' concept of reality lies outside the boundary of science, not simply because it has no experiential or verifiable meaning, but because it has no 'constructional' meaning at all: that is, it has no 'logical,' 'rational,' 'non-intuitive'".[38]

Statements such as "reality exists independently of our perception" or "there is an objective external world" are problematic, as they cannot be confirmed or falsified by experience.[39] For Carnap, philosophy should abandon traditional ontological disputes and instead focus on the logical analysis of scientific language and its formal structures. He retains a firmly anti-metaphysical stance, holding that meaningful statements must be traceable to empirical data or logical relations. All others, especially those of classical metaphysics, are to be dismissed as pseudo-problems. This view is reinforced in his *Autobiography*, where he acknowledges the influence of Wittgenstein's verification principle:[40]

> The view that these sentences and questions are non-cognitive was based on Wittgenstein's verifiability principle. This principle says first, that the meaning of

---

36 - R. Carnap, *Intellectual Autobiography*, in P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, Open Court, La Salle 1963, p. 23.

37 - As Mario Alai has convincingly shown, Carnap's anti-metaphysical stance should not be interpreted as a blanket rejection of all discourse concerning reality, but rather as a critique of statements that lack empirical or logical criteria of meaning. Our formulation is therefore intended only as a simplified characterization of Carnap's position in the context of the Aufbau. For a more nuanced and definitive discussion of Carnap's relation to realism and metaphysics, see M. Alai, *How I Developed My Views on Realism* in A. Angelucci, V. Fano, G. Ferretti, G. Galli, P. Graziani, G. Tarozzi, M. Alai et al., *Realism and Antirealism in Metaphysics, Science and Language. Festschrift for Mario Alai*, FrancoAngeli, Milano 2024.

38 - M. Friedman, *Carnap's Aufbau Reconsidered*, in «Noûs», 21 (1987), n. 4, pp. 521-545.

39 - The statement of Carnap is quite similar to 6.5 in L. Wittgenstein, *Tractatus Logico-Philosophicus*, translated by C. K. Ogden, Routledge & Kegan Paul, London 1922.

40 - For a detailed analysis of the debate surrounding the verification principle, which does not yet appear in the Aufbau but begins to emerge in *Scheinprobleme*, see T. Uebel, *Verificationism and (Some of) its Discontents*, in «Journal for the History of Analytical Philosophy», 7 (2019), n. 4, pp. 1-31.

> a sentence is given by the conditions of its verification and, second, that a sentence is meaningful if and only if it is in principle verifiable, that is, if there are possible, not necessarily actual, circumstances which, if they did occur, would establish the truth of the sentence. This principle of verifiability was later replaced by the more liberal principle of confirmability.[41]

This marks a clear divergence from the Neo-Kantian tradition and moves Carnap toward a radical form of logical empiricism. In the final section, we will argue that the *Aufbau* exemplifies a form of structural empiricism distinct from this radical constructivism. First, however, the next section will analyze the central points of Carnap's quasi-analysis framework.

4. The key points of Aufbau

Carnap's *Aufbau* is a controversial work open to a variety of interpretations and it has been read as an empiricist text. This section provides an overview of the key elements in *Aufbau* that challenge the standard reading of the work and supports an interpretation aligned with the structural empiricism presented in Section 2. In his work, Carnap identifies four categories of notions in the empirical sciences: auto-psychological experiences, physical entities, hetero-psychological states, and cultural objects. Auto-psychological experiences include perceptions, emotions, and mental states. These are accessible only to the subject who experiences them and, for Carnap, form the foundational basis of empirical knowledge.

> The thus-ordered elementary experiences themselves, their constituents and components, and the more complex entities to be constructed from these, form the domain of objects of which I am conscious, or my consciousness. This domain forms the foundation of the domain of the auto-psychological.[42]

Physical entities refer to objects of the physical world, described by the natural sciences. These are constituted from classes of qualities and relations. Although natural sciences postulate them, the theory of constitution is only a rational reconstruction of already established objects. "It can furthermore be shown that physical objects are epistemologically secondary relative to auto-psychological objects, since the recognition of physical objects depends upon perception".[43]

---

41 - R. Carnap, *Intellectual Autobiography*, in P. A. Schilpp (ed.), The Philosophy of Rudolf Carnap, Open Court, La Salle 1963, p. 53.

42 - *Ibid.*, p. 203.

43 - *Ibid.*, p. 307.

Hetero-psychological states encompass the beliefs, emotions, and perceptions of others. These differ from auto-psychological experiences as they are inferred rather than directly perceived. In Carnap's system, these are essential for understanding intersubjective communication.

> The construction of the heteropsychological can be an assignment only to the body of the other, not to his mind, which, after all, cannot be constructed in any other way than through this assignment; thus, constructionally, the other mind does not even exist before this assignment is carried out. Secondly, the assigned psychological events are auto-psychological events for the very same reason: the only psychological entities which have been constructed up to this point are auto-psychological entities, and no other can be constructed prior to this assignment; there is no possibility of constructing non-autopsychological entities other than with the aid of precisely this assignment.[44]

Lastly, cultural objects exist only within a cultural or social context, such as artworks or social norms.

> The object "state" could perhaps be constructed in the following form: a relational structure of persons is called a "state" if it is characterized in such and such a way through its manifestations, namely, the psychological behaviour of these persons and the dispositions toward such behaviour, especially the disposition, on the part of some persons, to act upon the volitions of others. Sociological groups or organizations are, among others, the most important higher cultural objects.[45]

For Carnap, cultural objects are not reducible to physical objects; they have a normative rather than purely empirical dimension. Carnap's epistemological system arranges these objects according to their epistemological reducibility. The result is a step-by-step logical system in which scientific objects are formally framed through basic concepts (auto-psychological experiences). Importantly, we have direct access only to these experiences. Yet, as Carnap notes, the recognition of this epistemological primacy comes after the development of scientific theories. That is, it is not initially obvious that an experience is subjective. Carnap refrains from attributing any ontological commitment to these categories. All scientific notions are quasi-objects: they possess logical structure but lack ontological status. His primary aim is to unify the sciences—to create a general framework capable of expressing all scientific objects using a single class of quasi-object. The

44 - R. Carnap, *Intellectual Autobiography*, in P. A. Schilpp (ed.), The Philosophy of Rudolf Carnap, Open Court, La Salle 1963, p. 215.

45 - *Ibid.*, p. 231.

method of quasi-analysis is central to this project. It echoes the logic of equivalence relations developed by Frege[46] and by Russell and Whitehead [47]. The process begins with the auto-psychological experiences of an individual and attempts to construct all other scientific objects from these elementary experiences. This construction is supported by one additional two-place logical term: the recollection of similarity among observed data. Carnap's method groups elementary experiences into *quality classes* (elements sharing a property) and *sense classes* (all experiences of a particular sensation). These are then distinguished using dimensional criteria. For example, Carnap explains that the visual quality space has five dimensions: two for spatial coordinates, and three for hue, brightness, and saturation. This construction is not a proper analysis but a quasi-analysis, since lived experiences are methodologically indecomposable. The process yields quasi-components from which scientific objects are constructed. Notably, Carnap acknowledges that scientific construction has already occurred; quasi-analysis merely reconstructs it. A key insight of the *Aufbau* is the principle of epistemological primacy: A is epistemologically prior to B if B is required to know A. Carnap uses formal logic, drawing on Frege's logic courses and Russell's *Principia Mathematica*, to translate all scientific statements into statements about a single class of entity. He chooses auto-psychological experiences as the epistemological base, because physical objects are known through them, but not vice versa: "Since, according to Carnap, physical objects are known through the mediation of auto-psychological objects but not vice versa, it follows that an auto-psychological basis is most appropriate from an epistemological viewpoint"[48]. In this way, Carnap assigns epistemological primacy to auto-psychological experiences over all other scientific quasi-objects. His logical construction thus mirrors the epistemological process of acquiring scientific notions, illustrating that Carnap's goal is genuinely philosophical: to make science both logically transparent and intellectually comprehensible. It is worth noting, as Francesco Barone[49] observed, that the only realism Carnap allows is sense-data realism, inherited from Russell.[50] This is apparent in the preface to

46 - G. Frege, *Über Sinn und Bedeutung,* in «Zeitschrift für Philosophie und philosophische Kritik», 100 (1892), pp. 25-50.

47 - B. Russell - A. N. Whitehead, *Principia Mathematica*, vol. I, Cambridge University Press, New York 1910.

48 - M. Friedman, *Carnap's Aufbau Reconsidered*, in «Noûs», 21 (1987), n. 4, p. 524.

49 - F. Barone, *Il neopositivismo logico*, Laterza, Bari 1953.

50 - Although F. Barone's interpretation remains historically significant, especially within the context of twentieth-century Italian scholarship on logical empiricism, its reception has been more limited in the broader international literature on Carnap and the Vienna Circle. We cite Barone primarily because of his insightful emphasis on the epistemological dimension of Carnap's early work, rather than as a definitive

the second edition of *Aufbau*: "I should now consider for use as basic elements, not elementary experiences, but concrete sense data"[51]. This philosophical position has largely been discredited. Today, we generally regard perceptions as representations of external objects that causally interact with us. Whether these representations are accurate is a separate question, which we do not address here. One final remark: the starting point of the *Aufbau* must be experience, due to its epistemological primacy. However, this experience should not be understood as in Mach's framework[52], these experiences should not be understood as the elements of Mach's approach, in which each element is considered a unique sensory datum and analyzed independently, but rather as *Gestalten*, in the sense of Köhler[53] and Wertheimer[54]. These are experiences that gain significance only when seen holistically in the process of quasi-analysis: "the basic elements of the constructional system cannot be analyzed through construction. Thus, the elementary experiences cannot be analyzed in our system since this system takes them as basic elements"[55]. In the *Aufbau*, Carnap moves away from the material a priori as proposed – for instance by Husserl in *Erfahrung und Urteil*.[56] Instead, Carnap adopts a softer form of empiricism compared to his later works; however, the influences of Husserl remain present in the project. Although Carnap does not fully adopt a Kantian position, the process of abstraction in his system can still proceed even in the absence of transitivity, and withstand the critique later raised

account of the Aufbau. For broader historical and philosophical perspectives on the Italian reception of logical empiricism and its relation to phenomenology, see P. Parrini, *Filosofia e scienza nell'Italia del Novecento*, Le Lettere, Firenze 2004; and Id., Empirismo logico e fenomenologia, in F. Masi (ed.), *Tra Experientia ed Experimentum. Percorsi del pensiero del Novecento*, ETS, Pisa 2023.

51 - R. Carnap, *Intellectual Autobiography*, in P. A. Schilpp (ed.), The Philosophy of Rudolf Carnap, Open Court, La Salle 1963, p. 8.

52 - The present discussion is intended primarily to provide the historical and conceptual background necessary for our comparison with Carnap's Aufbau. It is therefore not meant to offer a comprehensive treatment of the contemporary debate on intuition and self-evidence. For more systematic and up-to-date analyses of these issues, see E. C. Banks, *The Realistic Empiricism of Mach, James, and Russell: Neutral Monism Reconceived*, Cambridge University Press, New York 2014, P. Gori, *Ernst Mach dal punto di vista storico-critico*, in *Ernst Mach tra scienza e filosofia*, ETS, Pisa 2018, pp. 11-31 and M. Textor, *Mach's Neutral Monism*, in «HOPOS: The Journal of the International Society for the History of Philosophy of Science», 11 (2021), n. 1, pp. 143-165.

53 - W. Köhler, *Die physischen Gestalten in Ruhe und im stationären Zustand*, Vieweg Verlag, Braunschweig 1925.

54 - M. Wertheimer, *Laws of Organization in Perceptual Forms*, in «Psychologische Forschung», 4 (1923), pp. 301-350.

55 - R. Carnap, *Intellectual Autobiography*, p. 110.

56 - E. Husserl, *Erfahrung und Urteil*, Felix Meiner, Hamburg 1939.

by Goodman, due to the firmly structured nature of experience, which manifests in Gestalt-like processes, as we will show in Section 6. A similar perspective can be found in Cassirer's *Substanzbegriff und Funktionsbegriff*,[57] particularly in his discussion of concept formation. In his view, scientific objects are conceived as idealizations of sensations[58]. These objects are not given directly, but are constructed, one could say, as a sum of intuitions. Cassirer contrasts his view with that of Scholasticism, which raises the issue of

> The necessity, to which physical theory tends, has been opposed to the evidence and truth, which comes to consciousness in the experience of individual facts. This separation, however, rest upon a false abstraction; it is based in the attempt to isolate two moments from each other, which are inseparably connected together by the very presuppositions of the formation of concepts.[59]

In this sense, one may observe a parallel between Cassirer's account and the notion of quasi-abstraction as we interpret it within the framework of structural empiricism. Like Carnap, Cassirer is not strictly Kantian: in his view, the idealization of scientific objects originates in intuition and sensation, and these constructions form the ideal objects that science investigates.

In our view, to rehabilitate Carnap's philosophy today, we should retain the core principles of the *Aufbau*. The only element we might discard is sense-data realism. In epistemology, perceptions are not considered ontologically independent; rather, they are modes of interacting with the world. Still, Carnap's emphasis on the primacy of auto-psychological experience, the parallelism between logical construction and epistemological acquisition, and the aspiration to unify the sciences remain vital. Carnap himself makes this clear:

> Construction theory agrees with transcendental idealism in the conception that all objects of cognition are constructed (in idealistic language: 'are created in thought'); in fact, the constructed objects are objects of conceptual knowledge only qua logical forms which are generated in a certain way.[60]

It is important to note that the primary role of auto-psychological experiences over all other scientific quasi-objects is a conventional stance, as it applies to all

---

57 - E. Cassirer, *Substanzbegriff und Funktionsbegriff,* Wissenschaftliche Buchgesellschaft, Darmstadt 1910.

58 - For a detailed analysis see E. Cassirer, Substance and Function and Einstein's Theory of Relativity, Dover Publications, New York 1953, pp. 148-151.

59 - *Ibid.*, p.148.

60 - R. Carnap, *Intellectual Autobiography*, in P. A. Schilpp (ed.), The Philosophy of Rudolf Carnap, Open Court, La Salle 1963, p. 285.

possible bases. However, once the choice of a basis is made, it determines the entirety of the logical structure, which in turn constitutes the method of quasi-analysis. For the Carnap of the *Aufbau*, the epistemological order is of primary importance, and this is the guiding aim of his logical reconstruction. This structure, however, must be understood in light of the "imperfect community" problem identified by Goodman: "the very problem of abstraction is to define, in terms of the primitive relation of concrete individuals, the conditions under which a set of such individuals have a common quality."[61] Even though the quasi-analysis systematically groups experiences based on similarity, these groupings may not be perfectly consistent across all observers. Nonetheless, by giving to auto-psychological experiences a primary role, Carnap establishes a stable epistemic framework that allows the logical construction to proceed, providing a methodical way to handle experiential variation while preserving the structural integrity of the system.

5. Against the critiques of the Aufbau

Before proceeding further, it is useful to briefly consider two significant rehabilitations of *Aufbau* by authoritative scholars: Thomas Mormann[62] and Hannes Leitgeb[63] . Both authors attempt to reformulate Carnap's *Aufbau* in ways that respond to the criticisms raised by Goodman and Quine. Leitgeb proposes a revised formulation of the *Aufbau* program through what they term a "new attempt at an Aufbau-like system:" "every scientific sentence can be translated to an empirically equivalent one which consists solely of logico-mathematical signs and terms that refer to experience, such that the translation image expresses a subject-invariant constraint on experience"[64]. What is relinquished from Carnap's original program is the commitment to strict logical reduction and universal definability, i.e., the explicit step-by-step logical definitions Carnap sought, along with the ambition to reconstruct science solely through logical processes. Leitgeb's proposal addresses several problems raised by Goodman.[65] One key issue concerns the possibility that the quasi-analysis method may result in groupings of experimental

61 - Goodman, N. (1966). *The structure of appearance*. Indianapolis: Bobbs-Merrill.

62 - T. Mormann, *New Work for Carnap's Quasi-Analysis*, in «Journal of Philosophical Logic», 38 (2009), n. 3, pp. 249-282.

63 - H. Leitgeb, *New Life for Carnap's Aufbau?,* in «Synthese», 180 (2011), n. 2, pp. 265-299.

64 - *Ibid*, p. 268.

65 - N. Goodman, The Significance of Der logische Aufbau der Welt, in P. A. Schilpp (ed.), The Philosophy of Rudolf Carnap, Open Court, La Salle 1963, pp. 545-558.

elements based on relational similarity for unrelated or arbitrary reasons. To counter this, Leitgeb modifies Carnap's framework in two steps:

- Introduction of the trope theory: Each object is viewed as a bundle of tropes[66] (i.e., localized properties), which helps to spatially and temporally ground experiential content.
- Replacement of the single relation with two distinct relations: one governs the constitution of basic elements, and the other governs temporal overlap. These refinements allow for a more nuanced representation of similarity and help avoid arbitrary groupings. The "gluing" of tropes into quasi-objects is governed by principled constraints derived from trope theory.

Another concern raised by Goodman is the discrepancy between Carnap's notion of finite subjective experience and the five-dimensional structure of visual quality space. Leitgeb resolves this by allowing for an infinite number of possible tropes, thereby departing from Carnap's finite conception of quality spaces:

> The similarity relation functions as a 'neighbourhood' relation on the phenomenal quality points, which is all that is needed in order to define a topology on its connectivity components. What Carnap overlooked when doing so, but what Goodman observed, was that every finite topological space is zero-dimensional.[67]

This allows for the coherent reconstruction of spaces such as the five-dimensional visual field (two spatial, three chromatic dimensions) without inconsistency. Leitgeb also addresses Quine's criticism of Carnap's treatment of physical objects:

> Carnap did not seem to recognize, however, that his treatment of physical objects fell short of reduction not merely through sketchiness, but in principle. Statements of the form 'Quality q is at point-instant x; y; z; t' were, according to his canons, to be apportioned truth values in such a way as to maximize and minimize certain over-all features, and with growth of experience the truth values were to be progressively revised in the same spirit. I think this is a good schematization (deliberately oversimplified, to be sure) of what science does; but it provides no indication, not even the sketchiest, of how a statement of the form 'Quality q is at x; y; z; t' could ever be translated into Carnap's initial language of

66 - For an overview of Tropes theories, see P. Simons, *Particulars in Particular Clothing: Three Trope Theories of Substance*, in «Philosophy and Phenomenological Research», 54 (1994), n. 3, pp. 553-575.

67 - H. Leitgeb, *New Life for Carnap's Aufbau?*, in «Synthese», 180 (2011), n. 2, p. 276.

> sense data and logic. The connective 'is at' remains an added, undefined connective; the canons counsel us in its use but not in its elimination.[68]

Leitgeb responds by abandoning Carnap's demand for definability of theoretical terms solely from experiential terms. Instead, they propose a more modest revision that focuses on empirical equivalence: a scientific statement may be translated into another with the same observable consequences, even if not logically equivalent. In this updated model, theoretical terms gain meaning only within the context of the entire theory.

The second significant reinterpretation comes from Mormann[69], who explores the broader applicability of quasi-analysis in fields such as mathematics, cognitive science, and computer science. Of particular interest for our discussion is their response to Goodman's criticisms. Mormann refers to the unpublished 1923 manuscript by Carnap[70], which provides an alternative set of conditions for quasi-analysis:

> Let M be a non-homogeneous set. A set Q provides a quasi-analysis of M iff the following requirements are satisfied for all elements of M:
>
> (C1) If two elements of M are similar, they share at least one quasi-property q Î Q.
>
> (C2) If two elements of M are not similar, they do not share any quasi-property q Î Q.
>
> (C3) If two elements a and b of M are similar to exactly the same elements, they have the same quasi-properties.
>
> (C4) There is no proper subset Q' of Q such that the elements of Q' satisfy (C1), (C2), and (C3) for all elements of M.[71]

The first two conditions, (C1) and (C2), are essential for establishing a valid connection between similarity and shared properties. These are present in *Aufbau* and are necessary for quasi-analysis to function. The additional conditions, (C3) and (C4), while not included in *Aufbau*,[72] are crucial in addressing Goodman's objections.

---

68 - *Ibid.*, pp. 267.

69 - T. Mormann, *New Work for Carnap's Quasi-Analysis*, in «Journal of Philosophical Logic», 38 (2009), n. 3, pp. 249-282.

70 - R. Carnap, *Die Quasizerlegung: Ein Verfahren zur Ordnung nichthomogener Mengen mit den Mitteln der Beziehungslehre* (unpublished manuscript RC-081-04-01), University of Pittsburgh, Pittsburgh 1923.

71 - T. Mormann, *New Work for Carnap's Quasi-Analysis*, in «Journal of Philosophical Logic», 38 (2009), n. 3, pp. 253.

72 - One interesting point to note is that Carnap was aware of the problem with the *Aufbau*. He explained it in a letter to Goodman in 1938.

- Condition (C3) ensures that if two elements share the exact similarity structure, they must share the same quasi-properties. This addresses the problem of arbitrary or inconsistent distribution of properties.
- Condition (C4) invokes a kind of *Occam's razor* by eliminating redundant quasi-properties, those that always co-occur with others, thus mitigating the problem of "constant companionship" identified by Goodman.

Mormann also demonstrates how these refinements help avoid Goodman's "imperfect community" problem, where similar objects may be described with sets of properties that lack a univocal rule for property assignment. The added conditions enable the construction of more coherent representations, thereby resolving issues like the *Goodmanian triangle*, whose vertices are similar but share no single property. This scenario violates (C3) and is thus excluded by Mormann's improved model.

It is worth noting that both Leitgeb and Mormann largely set aside the broader epistemological context we explored in Section 2. While their technical and formal contributions are invaluable, our aim is distinct: to show that Carnap's quasi-analysis, far from being a merely formal reconstruction, represents a deep philosophical response to the problem of the relation between logic and experience, a position closely aligned with what we have termed structural empiricism. It is important to keep in mind that the *Aufbau* should be read in the same context as Carnap's earlier works, rather than through the lens of his later views. The Carnap of the *Aufbau* still preserves several important elements inherited from the Husserlian phenomenological tradition, even though these are reinterpreted within a distinct logical and epistemological framework. As we presented in Section 3, although he rejects the notion of a material a priori, he posits a form of intuition that plays a key epistemological role. Carnap himself further supports this reading of the *Aufbau*: in 1962, when the work was translated into English, he reiterated his interpretation of the entire work, without introducing a neutral theory of constitutional systems of concepts applicable to all scientific bases. Indeed, the choice of epistemic basis is fundamental.[73] This point will be the focus of the concluding section.

6. A new perspective on the *Aufbau*: structural empiricism

Carnap builds his logical system on the *Gestalten* of our sensible experience. Certain relations connect these *Gestalten*, most notably through the memory of

73 - While it is true that Carnap's position changes significantly in his later works, it is necessary to attempt a reversal of the usual analysis of Carnap in order to reevaluate his work. Our view is that the *Aufbau* should be read as the culmination of German philosophy, rather than as the starting point of empiricism.

partial similarities between experiences. These experiences can be analyzed in terms of their constituents; however, it is crucial to recognize that experiences are psychological, they are "lived," whereas "constituents" belong to theoretical language and are already conceptually constructed. As Carnap states: "Elementary experiences are essentially unanalyzable units, which has, after all, led us to choose them as basic elements"[74]. They are, in other words, unanalyzable wholes. Contrary to common belief, the fundamental issue in quasi-analysis is not the problem of abstraction. Carnap is well aware that similarity is not a transitive relation. A reflexive and symmetric relation cannot generate non-overlapping similarity classes. As discussed in the previous section through Mormann's analysis, Carnap had already formulated additional axioms in 1923 to address this issue, yet he chose not to include them in the *Aufbau*. He devotes several pages to discussing the imperfect separation of similarity circles. Why did Carnap avoid adding more axioms to resolve the problem?

For Carnap, logical and linguistic instruments must take intuition into account. While the two additional axioms, which represent an application of Occam's razor that is "is used to exclude "unfavorable circumstances" in which one quasi-property always appears accompanied by another one"[75] and a kind of indiscernibility condition (which states that two empirical experiences that are similar to the same experiences should share the same quasi-properties), are well-suited to addressing the problem of the "imperfect community," they are too stringent when considering intuition. These axioms are more appropriate within the framework of dianoia, which refers to the step-by-step, logical form of thinking that characterizes much of scientific and mathematical reasoning. Intuition, however, should be aligned more closely with nous, which refers to the immediate, non-discursive intellectual insight that grasps truths without the need for reasoning or argumentation. Unlike dianoia, which proceeds through formal logic and argumentation, nous reflects the kind of intuition that is crucial to mathematical insight and which we propose as fundamental in structural empiricism. A possible explanation for Carnap's decision not to add these axioms is that quasi-analysis must adhere to an epistemological constraint, just like the *Gestalten* and the similarity relation themselves. Reflexivity and symmetry of similarity are epistemologically reasonable, more stringent properties of the similarity relation would not be effective. Carnap was aware that Husserl believed *Gestalten* possess an internal structure that could be explained in a pure theory of experience, i.e., phenomenology. However, Carnap does not follow

74 - R. Carnap, *The Logical Structure of the World*, University of California Press, Berkeley 1967.

75 - T. Mormann, *New Work for Carnap's Quasi-Analysis*, in «Journal of Philosophical Logic», 38 (2009), n. 3, p. 253.

this idea, and his theory relies on Gestalt-like structures and similarity relations, aligning with the role assigned by Brentano to apodictic experiences. These Gestalt-like structures show that experiential configurations appear necessary or natural, even though they are not derived from formal reasoning. While auto-psychological experience holds epistemological primacy, Carnap recognized that it does not possess ontological priority. Gestalten, as we now understand them, are constructs that emerge only through the development of empirical science. Experience is indeed epistemologically prior, but it is not a pure given upon which science can be built directly. Moreover, Carnap should not be labelled an empiricist in the tradition of Mach. The quasi-analysis does not yield a conventionalist theory inspired purely by experience. Rather, it generates a framework grounded in experience without reducing theory to empirical convention, since the structure of quasi-constituents cannot be derived straightforwardly from psychological data. At the same time, Carnap does not fully align with Kantianism. He accepts the Kantian epistemological renunciation of absolute reality, a view reflected in Husserl's *epoché*[76]. The Kantian approach offers significant epistemological support to science, and Carnap clearly draws on this line of thought. For him, experience is deeply structured through logical constructs. Nevertheless, it retains an inherent and irreducible structure. As Carnap writes:

> Our problem now is how science can arrive at intersubjectively valid assertions if all its objects are to be constructed from the standpoint of the individual subject, that is, if in the final analysis all statements of science have as their object only relations between 'my' experiences? Since the stream of experience differs for each person, how can there be one objective statement of science in this sense (i.e., which holds for every individual, even though he starts from his own individual stream of experience)? The solution to this problem lies in the fact that, even though the material of the individual streams of experience is entirely different, or rather altogether incomparable, since a comparison of two sensations or two feelings of different subjects, as far as their immediately given qualities are concerned, is absurd, specific structural properties are analogous for all streams of experience. Now, if science is to be objective, then it must restrict itself to statements about such structural properties, and, as we have seen earlier, it can restrict itself to statements about structures, since all objects of knowledge are not content, but form, and since they can be represented as structural entities[77].

76 - E. Husserl, *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*. First Book, translated by F. Kersten, Nijhoff, The Hague 1982.

77 - R. Carnap, *The Logical Structure of the World*, University of California Press, Berkeley 1967, pp. 106-107

This quotation explains our use of the term structural empiricism. In this paper, we have argued that *Aufbau* can be fruitfully reinterpreted from an epistemological standpoint grounded in structural empiricism. Within this framework, the focus is not on justifying scientific concepts directly through sensory data but on emphasizing the structural role of subjective experience and self-evidence in the constitution of knowledge. We contend that Carnap's quasi-analysis is not a mere technical misstep but a profound epistemological gesture: science, no matter how abstract, must remain anchored in the structure of subjective experience. In this regard, Carnap anticipates a view similar to those of Angelucci and Stegmüller, who emphasize that experience, intuition, and self-evidence play a regulative, though not justificatory, role in epistemology. The epistemological thesis at the heart of the *Aufbau* is extremely nuanced. In our interpretation, it should not be reduced to empiricism, Kantianism, or phenomenology. Rather, Carnap critically reflects on the possibility of creating a scientific language that is both logically rigorous and grounded in lived experience. Ultimately, reading *Aufbau* through the lens of structural empiricism allows us to rehabilitate Carnap's project. His work remains a compelling attempt to reconcile the formal structures of science with the complexity of human experience. To understand science, we must view it through the lens of our auto-psychological experiences. However, we must also recognize that these experiences themselves cannot be the subject of scientific inquiry. What is possible, rather, is to construct a language of empirical quasi-constituents that renders even the most abstract concepts of science more comprehensible.

VINCENZO FANO
Professore ordinario di Logica e Filosofia della scienza presso il Dipartimento di Scienze Pure e Applicate (DiSPeA) dell'Università di Urbino Carlo Bo
vincenzo.fano@uniurb.it

NICCOLÒ COVONI
Dottorando in Filosofia della Fisica (Research Methods in Science and Technology program) presso l'Università di Urbino Carlo Bo e in Filosofia presso l'Università della Svizzera Italiana (USI)
niccolo.covoni@gmail.com

Quando un'analogia guida la scienza. Il ruolo metodologico dell'analogia nella termodinamica dei buchi neri

di Vincenzo Fano - Marco Sanchioni

Analogy is often understood as an inferential tool for justifying projections across domains, yet such a reading proves insufficient when analogy orients theoretical development over the long term. We therefore introduce the notion of analogical support – a form of methodological justification distinct from both heuristics and empirical confirmation, capable of rendering rational the continuation of research programmes even under conditions of limited empirical access – taking black hole thermodynamics as our case study.

Keywords: Analogy, Analogical Reasoning, Black Holes, Philosophy of Physics

1. Introduzione

In molte aree della scienza contemporanea, in particolare nei domini in cui l'accesso empirico è limitato, la ricerca procede attraverso una forte dipendenza da strutture teoriche indirette[1]. In tali situazioni, lo sviluppo della conoscenza scientifica non può essere descritta adeguatamente come una semplice successione di ipotesi sottoposte a verifica empirica. Piuttosto, essa assume la forma di uno sviluppo progressivo guidato da quadri concettuali che orientano la formulazione dei problemi, delimitano lo spazio delle soluzioni ammissibili e conferiscono coerenza ai programmi di ricerca nel lungo periodo[2].

Tra gli strumenti che svolgono questa funzione orientativa, l'analogia occupa un ruolo centrale. Tradizionalmente, la filosofia della scienza ha interpretato l'analogia come una forma di ragionamento che consente di trasferire proprietà da un dominio noto a uno meno noto, sulla base di una struttura di similarità. In questa prospettiva, il suo valore epistemico è valutato nei termini della plausibilità delle singole inferenze che essa rende possibili. Tuttavia, questa lettura inferenziale risulta insufficiente quando si considerano quei casi in cui l'analogia non si limita a sostenere ipotesi isolate, ma contribuisce a strutturare intere traiettorie di ricerca.

In tali contesti, il problema non è tanto stabilire se una specifica proiezione analogica sia giustificata, quanto comprendere in che senso il ricorso continuato a un'analogia possa risultare razionale, anche in presenza di tensioni teoriche persistenti e in assenza di conferme empiriche dirette. L'analogia, in altre parole, sembra operare non solo come strumento inferenziale, ma come elemento che contribuisce alla stabilità e alla continuità dei programmi di ricerca sotto condizioni di incertezza.

Questo lavoro si propone di chiarire tale funzione introducendo la nozione di *supporto analogico*. Con questa espressione intendiamo una forma di giustificazione metodologica distinta sia dall'uso euristico dell'analogia sia dal suo eventuale

1 - Un esempio emblematico è fornito dalla gravità quantistica. Gli effetti quantistici della gravità dovrebbero diventare rilevanti alla scala di Planck, caratterizzata da una lunghezza LP ~ 10-35*m* e da un'energia EP ~ 1019 GeV. Le energie attualmente accessibili negli esperimenti di fisica delle particelle, come quelli realizzati al Large Hadron Collider, raggiungono al massimo circa 104 GeV, rimanendo quindi molti ordini di grandezza al di sotto della scala in cui tali effetti dovrebbero manifestarsi direttamente. In queste condizioni, lo sviluppo teorico procede inevitabilmente in assenza di un accesso empirico diretto.

2 - Si veda, in una direzione affine, il dibattito sulle *non-empirical virtues* o *super-empirical virtues*, intese come criteri teorici – quali coerenza interna, semplicità, potere unificante o capacità esplicativa – che possono guidare la valutazione e lo sviluppo delle teorie in assenza di evidenza empirica diretta. Per una discussione sistematica, cfr. R. Dawid, *String Theory and the Scientific Method,* Cambridge University Press, Cambridge 2013; J.W. McAllister, *Beauty and Revolution in Science,* Cornell University Press, Ithaca 1996;

ruolo confermativo. Il supporto analogico, come vedremo, riguarda le condizioni che rendono razionale proseguire nello sviluppo di un programma di ricerca guidato da un'analogia, anche in presenza di difficoltà concettuali non ancora risolte. Per articolare questa proposta, faremo riferimento alla nozione di programma di ricerca, reinterpretata in modo da distinguere tra diverse forme di sviluppo teorico. In particolare, sosteniamo che, nei contesti in cui l'accesso empirico è limitato, la produzione di sviluppi teorici significativi può costituire una forma di progresso metodologicamente significativa, pur in assenza di nuove predizioni empiricamente controllabili. È proprio in questo spazio che il supporto analogico acquista rilevanza.

Il caso della termodinamica dei buchi neri fornisce un esempio particolarmente istruttivo di questa dinamica[3]. In tale contesto, un'analogia inizialmente formulata a livello formale ha progressivamente acquisito un ruolo strutturante nello sviluppo teorico, contribuendo sia alla formulazione di nuove domande sia alla riorganizzazione di quadri concettuali esistenti. Allo stesso tempo, essa ha generato tensioni profonde, che non hanno tuttavia condotto all'abbandono del programma di ricerca, ma piuttosto alla sua ristrutturazione. In questo senso, il caso dei buchi neri funge da esempio esplicativo attraverso cui articolare una tesi metodologica più generale.

L'obiettivo del lavoro è dunque duplice. Da un lato, si intende mostrare che il ruolo epistemico dell'analogia non può essere esaurito nei termini del ragionamento analogico tradizionalmente inteso. Dall'altro, si propone una caratterizzazione del supporto analogico come forma intermedia di giustificazione metodologica, distinta sia dalla generazione di ipotesi sia dalla loro conferma. Questa prospettiva consente di comprendere come, in alcune aree della scienza contemporanea, la razionalità della ricerca possa essere sostenuta nonostante l'assenza di evidenza empirica diretta e la presenza di tensioni teoriche significative.

L'articolo è organizzato come segue. Nella Sezione 2 introduciamo una tipologia delle analogie scientifiche e ne discutiamo il ruolo epistemico. Nella Sezione 3 sviluppiamo la nozione di supporto analogico e ne analizziamo le implicazioni metodologiche. Nella Sezione 4 consideriamo il caso della termodinamica dei buchi neri come studio esemplare. Infine, nella Sezione 5 discutiamo le implicazioni più generali della nostra proposta per la filosofia della scienza.

---

3 - Con "termodinamica dei buchi neri" si intende l'insieme di risultati, sviluppati a partire dagli anni Settanta, che stabiliscono una corrispondenza tra le leggi della meccanica dei buchi neri e le leggi della termodinamica. In questo quadro, grandezze geometriche come l'area dell'orizzonte e la cosiddetta *surface gravity* assumono un ruolo analogo rispettivamente all'entropia e alla temperatura. Per le formulazioni classiche di questo quadro si vedano J.D. Bekenstein, *Black Holes and Entropy*, in «Physical Review D», 7 (1973), pp. 2333-2346; J.M. Bardeen - B. Carter - S.W. Hawking, *The Four Laws of Black Hole Mechanics*, in «Communications in Mathematical Physics», 31 (1973), pp. 161-170; S.W. Hawking, *Particle Creation by Black Holes*, in «Communications in Mathematical Physics», 43 (1975), pp. 199-220.

2. Analogia e ruolo epistemico nelle scienze naturali

Per comprendere il ruolo dell'analogia nei contesti considerati, conviene anzitutto distinguere tra *analogia* e *ragionamento analogico*. Con il primo termine intendiamo una relazione di corrispondenza tra due domini, un sistema sorgente e un sistema bersaglio, stabilita attraverso una mappatura tra proprietà, relazioni o strutture. Il ragionamento analogico consiste invece nell'uso inferenziale di tale corrispondenza: a partire da similarità già riconosciute tra i due domini, si attribuiscono al dominio bersaglio proprietà ulteriori non ancora osservate o stabilite. L'analogia, dunque, è la relazione di corrispondenza; il ragionamento analogico è l'uso epistemico che se ne fa.

Una parte significativa della letteratura filosofica ha analizzato l'analogia soprattutto in questa seconda accezione, cioè come forma di ragionamento che consente di estendere il contenuto conoscitivo oltre le informazioni iniziali. In questa prospettiva, sviluppata in modi diversi da Hesse, Gentner e Bartha, l'attenzione si concentra sulle condizioni sotto le quali il trasferimento di proprietà da un dominio a un altro può dirsi epistemicamente giustificato. Nei lavori di Hesse, per esempio, l'analogia viene ricostruita attraverso la distinzione tra aspetti positivi, negativi e neutri: i primi corrispondono a somiglianze già stabilite tra i due domini, i secondi a differenze note, mentre i terzi individuano proprietà ancora non determinate[4]. Il suo interesse epistemico consiste proprio nel permettere un'estensione controllata di questi aspetti neutri, sulla base della struttura già condivisa. Gentner insiste invece sul fatto che la forza dell'analogia dipende principalmente dall'allineamento tra relazioni, piuttosto che dalla semplice somiglianza tra proprietà isolate: ciò che conta non è tanto la presenza di tratti comuni, quanto la possibilità di mettere in corrispondenza sistemi di relazioni strutturalmente analoghi[5]. Bartha, infine, ne offre una ricostruzione normativa, trattando il ragionamento analogico come una forma di inferenza ampliativa sottoposta a vincoli di coerenza globale e di rilevanza delle corrispondenze, e quindi valutabile in termini di condizioni di accettabilità razionale[6]. Nonostante queste differenze, il punto comune è chiaro: l'analogia viene valutata primariamente in funzione della sua capacità di sostenere proiezioni plausibili.

---

4 - M. Hesse, *Models and Analogies in Science*, University of Notre Dame Press, Notre Dame 1966;

5 - D. Gentner, *Structure-Mapping: A Theoretical Framework for Analogy*, in «Cognitive Science», 7 (1983), pp. 155-170;

6 - P. Bartha, *By Parallel Reasoning: The Construction and Evaluation of Analogical Arguments,* Oxford University Press, Oxford 2010.

Per chiarire meglio questo quadro, è utile distinguere tre forme di analogia, a seconda del livello al quale la corrispondenza tra i due domini si stabilisce[7].

Una prima forma è l'*analogia materiale*. In questo caso, la corrispondenza riguarda principalmente proprietà osservabili o aspetti fenomenici. Un esempio classico è l'analogia tra il flusso di calore e il flusso di un fluido. In entrambi i casi si parla di intensità del flusso, di gradienti e di equilibrio, e questa somiglianza descrittiva ha storicamente favorito il trasferimento di intuizioni da un dominio all'altro. Tuttavia, a questo livello, la corrispondenza resta relativamente superficiale: il fatto che due fenomeni presentino tratti osservabili simili non implica ancora che siano governati dalla stessa struttura teorica o che i rispettivi concetti svolgano lo stesso ruolo esplicativo. Per questo motivo, un'analogia materiale può essere utile per formulare ipotesi iniziali, ma in genere sostiene proiezioni epistemicamente più deboli.

Una seconda forma è l'*analogia formale*. Qui la corrispondenza non riguarda soltanto ciò che si osserva, ma la struttura matematica con cui i due sistemi vengono descritti. L'esempio più noto è quello dell'oscillatore armonico: la stessa equazione differenziale descrive una massa attaccata a una molla, un circuito elettrico oscillante e, in altri contesti, determinati modi di un campo. In questi casi, la similarità non è soltanto fenomenica. Essa riguarda la forma della legge che governa il sistema. Proprio per questo, l'analogia formale ha una forza epistemica maggiore di quella puramente materiale: se due sistemi condividono la stessa struttura matematica, è ragionevole aspettarsi che proprietà dimostrate in un caso possano essere trasferite, almeno in parte, anche nell'altro. Naturalmente, questa forza resta condizionata dal fatto che l'identità formale non garantisce ancora identità fisica o ontologica; ma consente proiezioni più robuste di quelle basate su una semplice somiglianza osservativa.

Una terza forma è l'*analogia funzionale*. In questo caso, la corrispondenza si estende al ruolo teorico svolto dalle grandezze o dalle strutture coinvolte. Non basta più che due sistemi si somiglino empiricamente o siano descritti da equazioni analoghe; occorre che gli elementi corrispondenti svolgano funzioni simili all'interno della spiegazione del fenomeno. Un buon esempio è offerto dall'uso della nozione di potenziale in contesti diversi: in meccanica e in elettrostatica, il potenziale

7 - La tripartizione qui proposta tra analogia materiale, formale e funzionale non pretende di offrire una classificazione normativa o esaustiva delle forme di analogia. Essa mira piuttosto a distinguere diversi livelli ai quali può stabilirsi una corrispondenza tra domini, al fine di chiarirne il diverso peso epistemico. In letteratura sono presenti sistematizzazioni alternative. In particolare, Hesse distingue tra analogie "orizzontali", che mettono in relazione sistemi diversi sulla base di somiglianze strutturali, e analogie "verticali", che collegano livelli differenti di descrizione all'interno di uno stesso dominio teorico. Cfr. M. Hesse, *Models and Analogies in Science,* University of Notre Dame Press, Notre Dame 1966.

non è soltanto una grandezza formalmente analoga, ma svolge anche un ruolo comparabile nell'organizzare la dinamica del sistema, nel vincolare il moto e nel rendere possibile una formulazione unificata delle leggi. Quando questo livello funzionale è presente, l'analogia acquista una profondità maggiore, perché la corrispondenza non riguarda soltanto il comportamento dei sistemi o la forma delle equazioni, ma anche il modo in cui certe grandezze contribuiscono alla spiegazione.

Questa distinzione tra analogia materiale, formale e funzionale permette di articolare con maggiore precisione il peso epistemico delle analogie. Non tutte le corrispondenze sostengono inferenze con la stessa forza. Una somiglianza limitata al piano osservabile fornisce in genere un appoggio modesto. Una corrispondenza formale offre invece una base più robusta, perché riguarda la struttura interna della descrizione teorica. Una corrispondenza anche funzionale è ancora più forte, perché coinvolge il ruolo esplicativo delle strutture messe in relazione. In questo senso, ciò che rafforza un'analogia non è la presenza di una similarità qualsiasi, ma la convergenza di più livelli di corrispondenza tra i due domini[8].

È però essenziale evitare un possibile equivoco. Dire che un'analogia rende una certa ipotesi più plausibile non significa dire che la confermi. La plausibilità generata dall'analogia non coincide con una forma di evidenza empirica. Un'analogia può rendere ragionevole prendere in considerazione una certa ipotesi, può orientare la sua formulazione e può motivarne lo sviluppo; ma non basta, da sola, a stabilirne la verità. La sua funzione epistemica consiste piuttosto nel selezionare e organizzare possibilità teoriche, non nel determinarne lo statuto empirico. Questa caratterizzazione è particolarmente importante nei contesti in cui l'accesso empirico è limitato. In tali casi, l'analogia non opera come sostituto della conferma, ma come criterio di orientamento teorico. Essa può contribuire a restringere lo spazio delle possibilità, a rendere alcune linee di sviluppo più promettenti di altre e a guidare l'esplorazione concettuale in assenza di verifiche dirette.

8 - La distinzione tra analogia materiale, formale e funzionale ha qui una funzione analitica e non introduce categorie rigidamente separate. In molti casi scientifici, una stessa analogia può infatti combinare simultaneamente diversi livelli di corrispondenza. Proprio per questo, il peso epistemico di un'analogia non dipende semplicemente dalla presenza di una singola similarità, ma dalla convergenza di più dimensioni analogiche. In generale, una corrispondenza soltanto materiale tende a sostenere proiezioni più deboli rispetto a una corrispondenza anche formale, mentre la presenza di un livello funzionale rafforza ulteriormente l'analogia, poiché coinvolge non solo la struttura matematica o fenomenica dei sistemi, ma anche il ruolo esplicativo svolto dalle strutture corrispondenti all'interno della teoria.

3. Dal Ragionamento Analogico al Supporto Analogico

Le considerazioni sviluppate nella sezione precedente permettono di chiarire un primo punto: l'analogia non può essere valutata soltanto nei termini della plausibilità delle singole inferenze che essa rende possibili. Questo livello di analisi è certamente importante, e descrive con precisione molti casi in cui una corrispondenza tra due domini consente di formulare nuove ipotesi o di estendere proprietà da un *sistema sorgente* a un *sistema bersaglio*. Tuttavia, esso non esaurisce il ruolo epistemico dell'analogia nella pratica scientifica. In numerosi casi, infatti, l'analogia non si limita a sostenere inferenze puntuali, ma contribuisce a organizzare un'intera linea di sviluppo teorico nel tempo. In tali contesti, il problema metodologico cambia natura. Non si tratta più soltanto di stabilire se una certa proiezione analogica sia più o meno plausibile, ma di comprendere in che senso il ricorso continuato a un'analogia possa risultare razionale anche quando esso si accompagna a tensioni, difficoltà o fallimenti parziali.

Per articolare questo secondo livello è utile ricorrere, almeno in prima approssimazione, alla nozione di programma di ricerca. Non è necessario, per il momento, assumere tutti i dettagli della ricostruzione lakatosiana[9]. Basta coglierne l'intuizione di fondo: la scienza non procede soltanto per ipotesi isolate, ma attraverso strutture relativamente stabili che orientano nel tempo la formulazione dei problemi, la selezione delle strategie teoriche e la valutazione dei risultati. Un programma di ricerca, in questo senso, non è una singola teoria, ma una struttura teorica organizzata attorno a un nucleo di assunzioni centrali, protetto da ipotesi ausiliarie e accompagnato da criteri che ne orientano lo sviluppo e la valutazione nel tempo[10].

Ora, in molti casi l'analogia può contribuire precisamente a questo livello. Essa non si limita a suggerire una proprietà nuova, ma fornisce un principio di organizzazione. Un'analogia stabile può fungere da quadro di orientamento per un programma di ricerca, perché rende intelligibile il trasferimento di intere famiglie di domande da un dominio all'altro. Una volta che due domini vengono messi in cor-

9 - I. Lakatos, *Falsification and the Methodology of Scientific Research Programmes*, in *Philosophy, Science and History*, Routledge, London 2014, pp. 89-94;

10 - La nozione di programma di ricerca è qui utilizzata in un senso lato, ispirato a Lakatos, senza assumerne integralmente l'apparato metodologico. Essa consente di rendere esplicita una dimensione già presente, seppure in forme diverse, in altri modelli della dinamica scientifica: da un lato, l'idea popperiana di sviluppo per congetture e confutazioni; dall'altro, la nozione kuhniana di paradigma come quadro relativamente stabile che orienta la ricerca. Rispetto a questi approcci, il riferimento a Lakatos permette di tematizzare in modo più esplicito la continuità e la struttura interna dei processi di sviluppo teorico. Cfr. K. R. Popper, *The Logic of Scientific Discovery*, Routledge, London 2005; T.S. Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago 1997; I. Lakatos, *Falsification and the Methodology of Scientific Research Programmes*, in *Philosophy, Science and History*, Routledge, London 2014, pp. 89-94.

rispondenza in modo sufficientemente robusto, l'analogia non suggerisce soltanto una risposta, ma un'agenda di lavoro. Essa contribuisce a stabilire che cosa debba ancora essere spiegato, quali difficoltà siano davvero decisive e quali, invece, possano essere trattate come tensioni locali.

Questa funzione organizzativa si vede bene nei casi in cui un'analogia continua a operare anche dopo l'emergere di difficoltà rilevanti. Se l'analogia fosse soltanto uno strumento per formulare inferenze puntuali, la comparsa di una disanalogia importante dovrebbe semplicemente ridurre la plausibilità delle proiezioni interessate. Ma quando l'analogia sostiene un programma di ricerca più ampio, la situazione è diversa. In tal caso, una difficoltà può essere trattata non come una confutazione immediata dell'analogia, ma come un problema interno al programma: qualcosa che richiede riformulazione, approfondimento, o revisione di alcune assunzioni. In altre parole, l'analogia entra a far parte della dinamica del programma e, proprio per questo, la sua valutazione non può essere ridotta a quella delle singole inferenze.

Un esempio storico aiuta a chiarire il punto. Nella seconda metà dell'Ottocento, l'elettromagnetismo fu spesso modellato per analogia con sistemi meccanici continui. In Maxwell, per esempio, il campo elettromagnetico viene rappresentato come uno stato di tensione e movimento di un mezzo elastico: le linee di forza sono interpretate come linee di tensione, mentre fenomeni come l'induzione vengono ricondotti a processi analoghi a vortici o a correnti in un fluido. Questa analogia suggeriva che le interazioni dovessero propagarsi localmente nel mezzo e che grandezze come il campo avessero un correlato fisico reale, non meramente formale. Tuttavia, già alla fine del secolo, questo quadro incontrava difficoltà: non era possibile fornire una descrizione coerente dell'etere come mezzo meccanico, e alcune proprietà richieste dalle equazioni di Maxwell risultavano incompatibili con i modelli elastici disponibili. Nonostante ciò, tali problemi non portarono immediatamente all'abbandono dell'analogia. Al contrario, la struttura matematica del campo, originariamente introdotta in quel contesto analogico, venne mantenuta e sviluppata, mentre il riferimento al mezzo meccanico fu progressivamente rielaborato e infine abbandonato. Ciò mostra che l'analogia poteva continuare a sostenere lo sviluppo teorico anche quando alcune sue implicazioni specifiche risultavano difficili da mantenere. Il punto non è se, a posteriori, quel programma fosse "giusto" o "sbagliato", ma che la sua razionalità metodologica non può essere compresa soltanto in termini di singole inferenze analogiche.

È precisamente questo tipo di dinamica che la nozione di *supporto analogico* intende catturare. Con l'espressione *supporto analogico* intendiamo una forma di giustificazione metodologica che riguarda non la plausibilità di una singola ipotesi, ma la razionalità del proseguimento di un programma di ricerca guidato da un'ana-

logia. Più precisamente, diremo che un'analogia fornisce supporto analogico quando contribuisce a rendere metodologicamente giustificato il proseguimento di una linea di ricerca, anche in assenza di conferma empirica diretta e anche in presenza di difficoltà teoriche ancora aperte.

Questa definizione richiede necessariamente alcuni chiarimenti. Anzitutto, il supporto analogico non coincide con il ruolo euristico dell'analogia. Un'analogia può essere euristicamente feconda nel senso minimo di suggerire nuove ipotesi o nuove domande. Ma ciò non basta ancora per parlare di supporto analogico. Il supporto analogico richiede qualcosa di più: che l'analogia non si limiti a produrre intuizioni iniziali, ma continui a sostenere lo sviluppo teorico nel tempo, contribuendo a orientare il lavoro concettuale e la gestione delle difficoltà interne al programma. In secondo luogo, il supporto analogico non coincide con una forma di conferma. Dire che un'analogia sostiene un programma di ricerca non significa dire che essa fornisca evidenza a favore della verità delle ipotesi coinvolte. Il punto è diverso. L'analogia può rendere razionale continuare a lavorare entro una certa linea teorica senza per questo aumentare direttamente il grado di conferma empirica delle sue tesi. Il supporto analogico riguarda le condizioni sotto le quali una certa traiettoria di ricerca merita di essere perseguita, non la misura in cui le sue ipotesi debbano essere considerate vere. In terzo luogo, il supporto analogico non è una proprietà puramente soggettiva. Non basta che gli scienziati trovino un'analogia "ispiratrice" o "suggestiva". Per parlare di supporto analogico deve esserci una base più oggettiva: l'analogia deve contribuire effettivamente alla produzione di sviluppo teorico, alla riorganizzazione di problemi, alla formulazione di vincoli, o alla ricomposizione di tensioni che altrimenti resterebbero opache. In questo senso, il supporto analogico non descrive semplicemente una motivazione psicologica, ma una forma di legittimazione metodologica.

Si può allora formulare il punto in modo più preciso. Un'analogia fornisce supporto analogico quando:

(i) stabilisce una corrispondenza sufficientemente robusta da rendere intelligibile il trasferimento di problemi e strutture da un dominio a un altro;
(ii) continua a orientare lo sviluppo teorico anche oltre la formulazione iniziale di singole proiezioni;
(iii) contribuisce alla produzione di sviluppi teorici significativi anziché limitarsi a sorreggere intuizioni vaghe o puramente descrittive[11];

11 - In un senso affine, come vedremo in seguito, alla nozione lakatosiana di contenuto teorico nuovo: un programma è progressivo quando le sue modifiche non si limitano a proteggere il nucleo da difficoltà già note, ma introducono elementi teorici che eccedono tali difficoltà, eventualmente anche anticipando nuove conseguenze empiriche.

(iv) mantiene questa funzione anche in presenza di difficoltà, purché tali difficoltà non distruggano il nucleo della corrispondenza.

A questo punto si vede meglio anche la differenza rispetto al ragionamento analogico. Il ragionamento analogico riguarda il passaggio da una corrispondenza riconosciuta a una proprietà ulteriore. Il supporto analogico riguarda invece la funzione che l'analogia svolge in una dinamica di ricerca più ampia. Il primo è locale; il secondo è globale. Il primo ha come oggetto una proiezione; il secondo una traiettoria teorica. Il primo si misura nei termini di plausibilità inferenziale; il secondo nei termini di razionalità metodologica del proseguimento della ricerca.

Per collocare questa nozione all'interno del paradigma lakatosiano, occorre ora precisare in che senso un'analogia possa contribuire alla progressività di un programma di ricerca. In Lakatos, la distinzione fondamentale è tra programmi progressivi e programmi regressivi. Un programma è progressivo quando le sue modifiche conducono a sviluppi teorici o empirici nuovi; è regressivo quando gli aggiustamenti servono principalmente a proteggere il programma da difficoltà già note, senza produrre avanzamenti significativi.

Se adottiamo questo schema, il punto diventa il seguente: un'analogia può fornire supporto analogico nella misura in cui contribuisce a sostenere un programma progressivo, o almeno non regressivo. Più precisamente, la funzione metodologica dell'analogia si manifesta quando essa non si limita a proteggere il programma da obiezioni, ma continua a generare risultati teorici non banali, ed eventualmente conferme empiriche. Tali risultati possono assumere forme diverse: l'introduzione di nuovi vincoli strutturali, la riorganizzazione di problemi fino ad allora separati, la formulazione di nuove connessioni concettuali, o l'integrazione coerente di principi che apparivano in tensione.

Qui però emerge una difficoltà. Nei contesti di frontiera, specialmente nella fisica teorica contemporanea, la progressività empirica è spesso difficile da stabilire. Vi sono programmi che producono un notevole sviluppo teorico senza generare, almeno nel breve periodo, nuove predizioni direttamente controllabili[12]. Se ci attenessimo rigidamente al criterio lakatosiano più forte, dovremmo forse considerarli sospesi o metodologicamente indeterminati. Ma una conclusione simile rischierebbe di essere troppo severa, e soprattutto di non descrivere adeguatamente la pratica effettiva della ricerca in questi domini. Per questo motivo conviene distin-

12 - Un esempio paradigmatico è la teoria delle stringhe, spesso discussa come un programma teoricamente fecondo nonostante l'assenza di predizioni empiriche direttamente controllabili nel breve periodo. Come argomenta Dawid, la sua razionalità metodologica è stata sostenuta da criteri non empirici, quali la coerenza interna, l'unificazione teorica e l'assenza di alternative comparabili. Cfr. R. Dawid, *String Theory and the Scientific Method*, Cambridge University Press, Cambridge 2013.

guere tra una forma forte e una forma debole di progressività. Un programma sarà *fortemente progressivo* quando i suoi sviluppi teorici si accompagnano a nuove conseguenze empiriche controllabili. Sarà invece *debolmente progressivo* quando, pur in assenza di tale crescita empirica, esso continua a produrre sviluppi teorici significativi, non meramente difensivi. La nozione di *progressività debole* non sostituisce quella lakatosiana standard, ma ne estende l'applicazione a contesti in cui l'accesso empirico è fortemente limitato.

Questa distinzione è importante per la nozione di supporto analogico. Se il programma guidato da un'analogia è fortemente progressivo, allora il ruolo metodologico dell'analogia appare relativamente chiaro. Ma anche quando il programma è soltanto debolmente progressivo, l'analogia può continuare a fornire un sostegno reale, purché essa contribuisca a generare sviluppo teorico non banale. In questo senso, il supporto analogico è compatibile con una forma di razionalità metodologica che non coincide con la conferma empirica, ma che nemmeno si riduce a mera persistenza conservativa.

Resta tuttavia da chiarire il ruolo delle disanalogie. Come visto nella sezione precedente, ogni analogia è inevitabilmente parziale. La presenza di differenze tra dominio sorgente e dominio bersaglio non elimina di per sé la funzione metodologica dell'analogia. Ciò che conta è il peso di tali differenze. Se esse riguardano aspetti secondari o locali, il programma può continuare a svilupparsi senza che il nucleo della corrispondenza venga compromesso. Se invece colpiscono proprio gli elementi che rendono l'analogia teoricamente feconda, allora la sua funzione di orientamento può indebolirsi fino a dissolversi. Ne segue che il supporto analogico non è una proprietà incondizionata. Esso dipende da un equilibrio tra due fattori: da un lato, la capacità dell'analogia di continuare a sostenere sviluppo teorico; dall'altro, il grado in cui le disanalogie lasciano intatto il nucleo della corrispondenza. Quando il primo fattore prevale e il secondo non è distruttivo, l'analogia conserva la propria funzione metodologica. Quando invece le disanalogie colpiscono il centro stesso del trasferimento analogico e il programma smette di produrre sviluppo teorico genuino, il supporto analogico viene meno.

Possiamo quindi formulare la tesi generale di questa sezione nel modo seguente. Il ruolo epistemico di un'analogia non si esaurisce nella plausibilità delle inferenze che essa giustifica. In certi casi, l'analogia contribuisce a sostenere il proseguimento razionale di un programma di ricerca nel tempo. Questa funzione, che abbiamo chiamato *supporto analogico*, dipende dal fatto che l'analogia continui a orientare sviluppo teorico non banale e che le disanalogie emergenti non distruggano il nucleo della corrispondenza che rende il programma fecondo.

È precisamente in questo senso che il caso della termodinamica dei buchi neri risulta filosoficamente istruttivo. Esso non costituisce soltanto un insieme di infe-

renze analogiche più o meno plausibili, ma un esempio in cui una analogia ha contribuito a organizzare e sostenere una linea di ricerca di lungo periodo, pur attraversata da tensioni concettuali profonde. La sezione seguente mostrerà in che modo questo caso possa essere letto alla luce della nozione di supporto analogico qui introdotta.

4. La termodinamica dei buchi neri e il paradosso dell'informazione

Il caso della termodinamica dei buchi neri offre un esempio particolarmente chiaro del ruolo che un'analogia può assumere nello sviluppo teorico. In questo contesto, un'analogia inizialmente formulata a livello formale è stata progressivamente rielaborata, messa sotto pressione da tensioni concettuali e infine trasformata attraverso nuovi sviluppi teorici.

Il punto di partenza è abbastanza noto. All'inizio degli anni Settanta, Bekenstein propose che i buchi neri possedessero un'entropia proporzionale all'area dell'orizzonte degli eventi[13]. Questa proposta si fondava su una analogia tra le leggi della meccanica dei buchi neri e le leggi della termodinamica: in particolare, la crescita dell'area dell'orizzonte ricordava il comportamento dell'entropia nei processi irreversibili. Poco dopo, Bardeen, Carter e Hawking formularono in modo sistematico le cosiddette "leggi della meccanica dei buchi neri", che riproducevano in modo sorprendente la struttura formale delle leggi termodinamiche[14].

A questo stadio, tuttavia, l'analogia era puramente formale. Le grandezze coinvolte – area, massa, superficie di gravità – occupavano posizioni strutturalmente analoghe a quelle di entropia, energia e temperatura, ma mancava un elemento cruciale: i buchi neri, nella teoria classica, non emettono radiazione. Essi si comportano come assorbitori perfetti. Di conseguenza, non esisteva alcun meccanismo fisico che giustificasse l'attribuzione di una temperatura[15].

Questa assenza costituiva una disanalogia rilevante. In un sistema termodinamico ordinario, la temperatura non è una semplice etichetta formale, ma una grandezza che governa lo scambio di energia con l'ambiente. Senza un processo di emis-

---

13 - J.D. Bekenstein, *Black Holes and Entropy*, in «Physical Review D», 7 (1973), pp. 2333-2346;

14 - J.M. Bardeen, B. Carter, S.W. Hawking, *The Four Laws of Black Hole Mechanics*, in «Communications in Mathematical Physics», 31 (1973), pp. 161-170.

15 - Per una ricostruzione filosofica della natura inizialmente formale dell'analogia tra meccanica dei buchi neri e termodinamica, cfr. D. Wallace, *The Case for Black Hole Thermodynamics, Part I: Phenomenological Thermodynamics*, in «Studies in History and Philosophy of Modern Physics», 64 (2018), pp. 52-67; Id., *The Case for Black Hole Thermodynamics, Part II: Statistical Mechanics*, in «Studies in History and Philosophy of Modern Physics», 66 (2019), pp. 103-117.

sione, l'identificazione tra superficie di gravità e temperatura rimaneva priva di fondamento fisico. L'analogia, pur suggestiva, risultava quindi incompleta. La situazione cambiò radicalmente con il lavoro di Hawking[16]. Mostrando che, in presenza di effetti quantistici, i buchi neri emettono radiazione con spettro termico, Hawking fornì il tassello mancante. I buchi neri non sono oggetti completamente "neri", ma sistemi che irradiano come corpi caldi, con una temperatura ben definita.

Questo risultato non rappresenta soltanto una scoperta indipendente, ma una trasformazione dell'analogia stessa. La corrispondenza tra meccanica dei buchi neri e termodinamica non è più soltanto formale, ma acquisisce una dimensione funzionale: i buchi neri non si limitano a soddisfare equazioni simili a quelle termodinamiche, ma si comportano effettivamente come sistemi termici. In questo senso, la radiazione di Hawking può essere vista come una realizzazione fisica di ciò che l'analogia suggeriva.

Tuttavia, proprio questa scoperta ha generato una nuova tensione, forse più profonda della precedente. La radiazione di Hawking è termica. Nella descrizione semiclassica, ogni modo della radiazione è distribuito secondo uno spettro di corpo nero, determinato unicamente da parametri macroscopici del buco nero, come massa, carica e momento angolare, e non dalle condizioni iniziali che hanno portato alla sua formazione. Di conseguenza, la radiazione appare priva delle correlazioni necessarie per codificare l'informazione microscopica dello stato iniziale. Se il buco nero evapora completamente, il processo nel suo insieme descrive una transizione da uno stato quantistico puro a uno stato misto, ossia a una distribuzione termica che conserva solo grandezze macroscopiche e non permette, neppure in linea di principio, di ricostruire lo stato iniziale[17].

Questo scenario entra in conflitto con un principio centrale della meccanica quantistica: l'unitarietà dell'evoluzione. In un sistema quantistico isolato, l'evoluzione conserva l'informazione. Anche quando un processo appare irreversibile a livello macroscopico, come nel caso dei sistemi termodinamici ordinari, l'informazione non viene distrutta, ma dispersa in correlazioni sempre più complesse. In linea di principio, uno stato iniziale può essere ricostruito a partire dallo stato finale. Nel caso dei buchi neri, la situazione sembra diversa. La radiazione termica prodotta da Hawking appare priva delle correlazioni necessarie per una tale ricostruzione. Ne segue quello che è noto come paradosso della perdita dell'informazione: o la meccanica quantistica deve essere modificata, oppure la descrizione semiclassica

16 - S.W. Hawking, *Particle Creation by Black Holes*, in «Communications in Mathematical Physics», 43 (1975), pp. 199-220.

17 - S.W. Hawking, *Breakdown of Predictability in Gravitational Collapse*, in «Physical Review D», 14 (1976), pp. 2460-2473.

dell'evaporazione dei buchi neri è incompleta[18]. Dal punto di vista della nostra analisi, è importante osservare che questa tensione può essere letta come una nuova disanalogia rispetto al modello termodinamico.

Questa disanalogia non ha condotto all'abbandono dell'analogia termodinamica. Al contrario, ha orientato lo sviluppo successivo della ricerca. Invece di rinunciare all'idea che i buchi neri siano sistemi termodinamici, molti lavori hanno cercato di comprendere in che modo l'unitarietà potesse essere restaurata all'interno di un quadro compatibile con la struttura termodinamica.

Un passo decisivo in questa direzione è stato il lavoro di Page[19], che ha riformulato il problema in termini di entanglement. Se l'evaporazione del buco nero è un processo unitario, allora l'entropia della radiazione non può crescere indefinitamente. Dopo una fase iniziale di aumento, essa deve raggiungere un massimo e poi diminuire, fino a tornare a zero quando il sistema finale è puro. Questo comportamento, noto come "curva di Page", rappresenta ciò che ci si aspetta da un sistema quantistico finito che evolve unitariamente. Per lungo tempo, tuttavia, questo risultato è rimasto in tensione con il calcolo di Hawking, che predice invece una crescita monotona dell'entropia. Negli ultimi anni, nuovi sviluppi nella gravità semiclassica hanno mostrato come questa tensione possa essere risolta. Senza entrare nei dettagli tecnici, l'idea centrale è che il calcolo originale di Hawking non include tutti i contributi rilevanti. Quando si tiene conto di effetti più sottili, la dinamica dell'entropia della radiazione risulta compatibile con la curva di Page e quindi con l'unitarietà[20]. Questo risultato può essere interpretato come una revisione interna del quadro teorico, che permette di mantenere sia la struttura termodinamica sia i principi della meccanica quantistica. In altri termini, la disanalogia introdotta dalla non-unitarietà non viene accettata come definitiva, ma utilizzata come guida per individuare i limiti della descrizione precedente e sviluppare un quadro più completo[21].

Nel suo insieme, questa sequenza di sviluppi mostra una dinamica caratteristica. L'analogia iniziale tra buchi neri e sistemi termodinamici viene prima formulata a livello formale, poi messa in difficoltà da una disanalogia rilevante (l'assenza di

---

18 - Per una discussione filosofica del paradosso dell'informazione e del ruolo dell'unitarietà, cfr. E. Cinti, M. Sanchioni, *I paradossi dei buchi neri. Una spiegazione elementare per comprendere l'universo*, Carocci, Roma 2022; Id., *The Devil in the (Implicit) Details: On the AMPS Paradox and Its Resolution*, in «International Journal of Theoretical Physics», 60 (2021), pp. 3234-3253.

19 - D.N. Page, *Average Entropy of a Subsystem*, in «Physical Review Letters», 71 (1993), pp. 1291-1294;

20 - G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, in «Journal of High Energy Physics», 2020 (9), pp. 1-45.

21 - Per un'analisi filosofica più approfondita del ruolo del supporto analogico nel contesto della termodinamica dei buchi neri e del paradosso dell'informazione, cfr. V. Fano, M. Sanchioni, *Black Hole Thermodynamics as Scientific Analogy* (manoscritto in preparazione).

radiazione), quindi rafforzata da un nuovo risultato (la radiazione di Hawking), e infine nuovamente problematizzata da una tensione più profonda (la perdita di informazione). Questa tensione, a sua volta, non porta all'abbandono dell'analogia, ma a una sua rielaborazione, che conduce a nuovi sviluppi teorici. È proprio questa dinamica – fatta di corrispondenze, tensioni e riformulazioni – che rende il caso della termodinamica dei buchi neri particolarmente adatto a chiarire il ruolo metodologico dell'analogia nei contesti in cui lo sviluppo teorico procede in condizioni di accesso empirico limitato.

5. Conclusioni

Le analisi sviluppate in questo articolo consentono di trarre una conclusione generale sul ruolo epistemico dell'analogia nei contesti di ricerca caratterizzati da accesso empirico limitato. L'analogia non può essere compresa esclusivamente come uno strumento inferenziale volto a giustificare singole proiezioni tra domini. In molti casi rilevanti, essa svolge una funzione più ampia: contribuisce a sostenere nel tempo lo sviluppo di un programma di ricerca, orientandone le direzioni, selezionando i problemi ritenuti significativi e rendendo metodologicamente razionale il proseguimento della ricerca anche in presenza di difficoltà teoriche non risolte.

È precisamente questa funzione che abbiamo cercato di articolare attraverso la nozione di *supporto analogico*. Tale nozione consente di distinguere tra due livelli che vanno distinti chiaramente: da un lato, la plausibilità delle inferenze analogiche; dall'altro, la razionalità del mantenimento di una traiettoria teorica guidata da un'analogia. Mentre il primo livello riguarda la giustificazione locale di singole ipotesi, il secondo riguarda la dinamica complessiva di un programma di ricerca. Il supporto analogico si colloca a questo secondo livello e non coincide né con il ruolo euristico dell'analogia né con una forma di conferma empirica.

Il caso della termodinamica dei buchi neri mostra con particolare chiarezza questa distinzione. In esso, l'analogia non si limita a sostenere alcune inferenze isolate, ma contribuisce a strutturare un'intera linea di sviluppo teorico. Ciò che emerge in modo particolarmente significativo è il ruolo delle disanalogie. Lungi dal costituire semplici anomalie da eliminare o segnali immediati di fallimento, le disanalogie si rivelano elementi interni alla dinamica del programma.

Questo punto merita di essere enfatizzato. La dinamica osservata non è quella di una progressiva eliminazione delle disanalogie, ma piuttosto quella di una loro trasformazione in vincoli per lo sviluppo teorico. Le disanalogie rilevanti non interrompono necessariamente il programma; al contrario, possono contribuire a ridefinirne gli obiettivi e a guidarne la ristrutturazione. In questo senso, esse svolgono una funzione metodologica positiva: segnalano i limiti del quadro corrente e, al tempo stesso, orientano la ricerca verso forme più articolate di coerenza teorica.

L'analisi proposta in questo articolo apre diverse linee di sviluppo.

Una prima direzione riguarda l'estensione della nozione di supporto analogico ad altri contesti della fisica teorica contemporanea. In particolare, negli ultimi anni si è sviluppata una seconda linea analogica, parallela a quella termodinamica, che mette in relazione i buchi neri con sistemi di informazione quantistica. In questo caso, nozioni come entanglement, codici di correzione degli errori e complessità computazionale vengono utilizzate per analizzare la struttura della gravità quantistica e della corrispondenza olografica. Anche qui, l'analogia non sembra operare soltanto come fonte di inferenze puntuali, ma come principio di organizzazione di un programma di ricerca. Un'analisi sistematica di questo secondo asse analogico potrebbe permettere di raffinare la nozione di supporto analogico proposta in questo lavoro.

Una seconda direzione riguarda il rapporto tra analogia e pluralità di programmi. Se diverse analogie possono coesistere e orientare lo sviluppo teorico in modi parzialmente indipendenti, si pone il problema di comprendere come esse interagiscano, si rafforzino o entrino in tensione tra loro. Nel caso dei buchi neri, l'analogia termodinamica e quella informazionale sembrano talvolta convergere, talvolta suggerire prospettive differenti. Comprendere questa interazione potrebbe offrire un quadro più articolato della dinamica dei programmi di ricerca in fisica teorica.

Infine, una terza direzione riguarda la portata generale della nozione di supporto analogico nella filosofia della scienza. Se l'analogia può sostenere in modo strutturato lo sviluppo teorico in condizioni di accesso empirico limitato, allora essa rappresenta un elemento costitutivo della razionalità scientifica in tali contesti, e non un semplice ausilio contingente. Questo suggerisce la necessità di integrare più esplicitamente il ruolo delle analogie all'interno delle teorie metodologiche della scienza, in particolare nei domini in cui la relazione tra teoria ed esperienza non può essere ridotta a uno schema diretto di verifica.

VINCENZO FANO
Professore ordinario di Logica e Filosofia della scienza presso il Dipartimento di Scienze Pure e Applicate (DiSPeA) dell'Università di Urbino Carlo Bo
vincenzo.fano@uniurb.it

MARCO SANCHIONI
Pofessore stabile straordinario di Filosofia della Scienza presso l'Istituto Universitario Sophia
marco.sanchioni@sophiauniversity.org

Once upon a time in the Pleistocene: integrating Paleoanthropology and evolutionary psychology

di Matteo Bedetti

This paper examines how evolutionary psychology integrates paleoanthropological evidence to reconstruct the evolution of the human mind. Using the savanna hypothesis as a case study, it introduces the concept of stylization bias and argues for a more historically grounded integration of paleoanthropological evidence into the discipline's methodology.

Keywords: evolutionary psychology, savanna hypothesis, paleoanthropology, philosophy of evolutionary biology, archaeology.

1. Introduction[1]

What role should evidence from deep human prehistory play outside the disciplines that produce it? In the philosophical debate surrounding the historical sciences, geology, paleontology and archaeology are frequently praised for their remarkable ability to draw upon and integrate methods and insights from other fields[2]. One well-known example is the adoption in archaeology of 14C dating, a method originally developed by physicists[3]. Another striking example is the more recent, and extraordinarily successful, incorporation of ancient DNA analysis into paleoanthropology[4]. While this phenomenon certainly deserves careful philosophical reflection, in this paper we will argue that attention should also be directed toward the opposite flow of information: from paleoanthropology and archaeology to other disciplines. Although paleoanthropology has traditionally borrowed more than it has lent in terms of technologies and methods, evidence concerning human evolution nonetheless plays a pivotal role in several neighboring fields. Indeed, research programs that seek to reconstruct our psychological and behavioral evolution depend (also) on data provided by paleoanthropologists and archaeologists. Human behavioral ecologists, evolutionary psychologists, cultural evolutionary theorists, and related scholars must rely on fossils, artifacts, ancient DNA, and the inferences drawn from these sources as the fundamental epistemic scaffolding for their hypotheses. For instance, human behavioral ecologists seeking to formulate adaptive hypotheses concerning the evolution of childhood cannot disregard the paleoanthropological record, which offers crucial evidence regarding when this life stage emerged, in which ecological and social contexts, and so on. To neglect such data would amount to speculation detached from historical reality[5].

In this paper, we will focus on one of the most renowned sciences concerned with the evolution of human behavior, evolutionary psychology. The aim is to discuss how evolutionary psychologists integrate (or should integrate) findings

1 - The content of this paper has been presented in condensed form in my book M. Bedetti, *La mente preistorica: analisi epistemologica della psicologia evoluzionistica,* Mimesis Edizioni, Milano 2025.

2 - See, for example, A. Currie, *Rock, bone, and ruin: An optimist's guide to the historical sciences* MIT Press, Cambridge 2024.

3 - R. Chapman - A. Wylie, *Evidential reasoning in archaeology,* Bloomsbury Publishing, London 2018, chapter 4.

4 - D. Reich, *Who we are and how we got here: Ancient DNA and the new science of the human past*, Oxford University Press, Oxford 2018. See also S. M. Downes, *The role of ancient DNA research in archaeology*, in «Topoi» 40 (2021), pp. 285-293.

5 - A. Meneganzin - A. Currie, *Not wasted on the young: Childhood, trait complexes & human behavioral ecology, in Studies in History and Philosophy of Science* 109 (2025), pp. 12-20.

from paleoanthropology and related disciplines in their methods. More specifically, the emphasis will be on paleoenvironmental reconstructions and their importance for formulating adequate hypotheses about our cognitive adaptations.

The structure of this paper is as follows. We begin with a brief overview of evolutionary psychology, outlining its core tenets. We then examine the methodological framework of the discipline, with particular attention to the so-called top-down and bottom-up approaches, both of which rely in essential ways on paleoanthropological evidence. Next, we examine a specific hypothesis in evolutionary psychology regarding the purportedly evolved human preference for certain habitats, a case that illustrates how claims about our evolutionary past are often advanced without sufficient engagement with the pertinent paleoanthropological and paleoecological record. We show how this hypothesis ought to be revised in light of such evidence and, finally, conclude with broader reflections on how the methodological foundations of evolutionary psychology might be reconfigured to better integrate paleoanthropological insights.

2. Evolutionary psychology: a primer

Evolutionary Psychology (EP) explains human cognition and behavior as products of evolution, arguing that the mind is composed of specialized adaptations evolved through natural and sexual selection to solve recurrent ancestral problems such as mate choice, kin recognition, social exchange or danger avoidance. The field emerged between the 1980s and the 1990s, especially through the work of Cosmides and Tooby integrating evolutionary biology with information-processing models of the mind borrowed from cognitive sciences[6]. Nowadays, it remains a vibrant research program, with a large community of scholars identifying with its framework and a notable degree of public visibility.

One of the main tenets of evolutionary psychology is that psychological mechanisms are designed by natural selection to process information in ways that (pre)historically increased fitness. These mechanisms are thought to be predominantly domain-specific, modular, universal, and largely invariant across cultures, and the core goal of the discipline is to uncover the structure and function of these evolved modules and to understand how they manifest in modern behavior, even when modern environments differ markedly from ancestral ones. Closely

6 - L. Cosmides - J. Tooby, *From Evolution to Behavior: Evolutionary Psychology as the Missing Link*, in *The Latest on the Best*, edited by J. Dupre, MIT Press, Cambridge 1987, pp. 277-306; J. Tooby L. Cosmides, *The psychological foundations of culture*, in *The Adapted Mind*, edited by J. Barkow L. Cosmides J. Tooby, Oxford University Press, Oxford 1992, pp. 19-136.

related to this aspect is the concept of the Environment of Evolutionary Adaptedness (EEA), the statistical composite of selection pressures that shaped human cognition. Although in principle each cognitive mechanism (adaptation) may have its own specific EEA, reflecting different selective contexts, many evolutionary psychologists usually associate our general EEA with the African savanna of the Pleistocene (from 2.58 million years ago to 11.700 years ago), where the members of our genus (Homo) evolved and lived as hunter-gatherers. What both conceptions have in common is the belief that our cognitive evolution stopped with the invention of agriculture, at the beginning of the Holocene[7]. According to evolutionary psychologists, evolutionary processes are slow, especially when they involve complex organs such as the brain, so it is a priori very unlikely that new psychological adaptations have evolved in such a short time. As a result, many of our evolved mechanisms are mismatched with modern environments, as illustrated by our preferences (e.g., for high-fat or sugary foods that were once adaptive but can be maladaptive today[8]).

3. The methods of evolutionary psychology

We briefly presented evolutionary psychology and some of its core tenets. Now we will introduce EP's two main heuristics for hypotheses generation: the top-down approach and the bottom-up approach. The top-down method, also called functional analysis[9], moves from general evolutionary principles to specific hypotheses about the structure and function of cognitive mechanisms. It is usually presented in six steps[10] (Tooby & Cosmides, 1989; Lieberman, Cosmides & Tooby, 2007). We will focus on the first two steps, since it is primarily at this stage that insufficiently constrained reconstructions of ancestral conditions risk generating serious epistemic distortions:

Step 1: Identification of a recurrent adaptive problem: a problem that repeatedly affected ancestral survival or reproduction and was relevant to fitness.

---

7 - J. Tooby L. Cosmides, *The Theoretical Foundations of Evolutionary Psychology*, in *The Handbook of Evolutionary Psychology*, edited by D.M. Buss, John Wiley & Sons Hoboken 20162, pp. 3-87.

8 - D. Buss, *Evolutionary Psychology: The New Science of the Mind*, Routledge London7 2024.

9 - D.J. Buller, *Adapting minds: Evolutionary psychology and the persistent quest for human nature*, MIT Press Cambridge 2005.

10 - J. Tooby - L. Cosmides, *Evolutionary psychology and the generation of culture, part I: Theoretical considerations*, in «Ethology and sociobiology», 10 (1989), pp. 2949; D. Lieberman J. Tooby L. Cosmides, *The architecture of human kin detection*, in «Nature», 445 (2007), pp. 727-731.

Step 2: Determining how that adaptive problem appeared in the evolutionary past: once the problem is defined abstractly, it must be reconstructed in its ancestral context.

The remaining stages involve developing a computational model of the adaptive challenge, inferring the minimal structure of the putative cognitive mechanism, empirically testing alternative hypotheses through a variety of methods, and finally assessing the degree of fit between the proposed model and observed human behavior. A classic example is Cosmides' hypothesis of a cheater-detection module[11], which starts from the adaptive problem of social exchange, reconstructs its Pleistocene context, proposes a domain-specific cognitive adaptation, and tests its predictions empirically, including through cognitive experiments and neuroscientific lesion studies[12].

The bottom-up method works in the opposite direction. It begins with an observed psychological trait, behavior, or preference in the present and tries to infer the adaptive function it may have had in the past. This is why it is often called *reverse engineering*[13]. For example, the widespread fear of snakes has been interpreted as evidence for a threat-detection mechanism shaped by recurrent danger in ancestral environments[14]. Bottom-up reasoning is especially vulnerable to producing "just-so stories"[15], narratives that explain the presence of a trait by appealing to a hypothetical adaptive history, without seeking independent evidence beyond the story's apparent plausibility. On the other hand, the top-down approach would not be subject to this kind of criticism. Indeed, while reverse engineering *assumes* that a trait is an adaptation and proposes an adaptive explanation for it, functional analysis is predictive, since it is possible to test hypotheses derived from it against the empirical evidence:

---

11 - L. Cosmides, *Deduction or Darwinian Algorithms?*, Doctoral dissertation, Harvard University, Cambridge 1985; L. Cosmides J. Tooby, *Evolutionary psychology and the generation of cultur*e, part II: Case study: A computational theory of social exchange, in «Ethology and sociobiology», 10 (1989), pp. 51-97; L. Cosmides - J. Tooby, *Cognitive adaptations for social exchange*, in *The Adapted Mind*, edited by J.H. Barkow - L. Cosmides - J. Tooby, Oxford University Press Oxford 1992, pp. 163-228;

12 - L. Cosmides - J. Tooby, *Neurocognitive adaptations designed for social exchange*, in *The handbook of evolutionary psychology*, edited by D.M. Buss, John Wiley & Sons, Hoboken 2016.

13 - D. Dennett, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Simon and Shuster, New York 1995; S. Pinker, *How the Mind Works*, Penguin Books, London 1997.

14 - A. Öhman - S. Mineka*, Fear, phobias and preparedness: Toward an evolved module of fear and fear learning*, in «Psychological Review», 108 (2001), pp. 483-522.

15 - With reference to Rudyard Kipling's 1902 collection of stories for children, in which the English writer invents amusing and imaginative explanations for how animals acquired certain characteristics.

> Our hominid ancestors had to be able to solve a large number of complex adaptive problems, and do so with special efficiency. By combining data from paleontology and hunter-gatherer studies with principles drawn from evolutionary biology, one can develop a task analysis that defines the nature of the adaptive information-processing problem to be solved […]. Once one understands the nature of the problem, one can then generate very specific, empirically testable hypotheses about the structure of the information processing mechanisms that evolved to solve it […]. One virtue of this approach is that it is immune to the usual (but often vacuous) accusation of post hoc storytelling: The researcher has predicted in advance the properties of the mechanism[16].

As the quotation above shows, Cosmides, Tooby, and Barkow adopt strikingly strong language, suggesting that functional analysis is entirely immune to charges of storytelling. But as we will show shortly, this is not the case. Indeed, both top-down and bottom-up approaches often share a deeper methodological limitation: a tendency to operate with minimal integration of paleoanthropological and archaeological evidence. When the Environment of Evolutionary Adaptedness is invoked without careful attention to the actual ecological, technological, and social conditions of ancestral populations, as reconstructed through the historical sciences, evolutionary explanations risk becoming disconnected from plausible scenarios of human evolution and, consequently, being empirically inadequate. The next sections will address this specific issue.

4. The stylization bias

Evolutionary psychology has produced many results but faces key criticisms: adaptationist bias[17], challenges to massive modularity and universality from evidence of plasticity and cross-cultural variation[18] and pressures from niche

16 - L. Cosmides - J. Tooby - J.H. Barkow, *Introduction: Evolutionary psychology and conceptual integration*, in *The Adapted Mind*, Oxford University Press, Oxford 1992, p. 11.

17 - S.J. Gould - R.C. Lewontin, *The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme*, in «Proc. R. Soc. Lond. B.», 205 (1979), pp. 581-598; E. Lloyd, *Adaptation*, Cambridge, University Press Cambridge 2021.

18 - F. d'Errico - I. Colagè, *Cultural exaptation and cultural neural reuse: A mechanism for the emergence of modern culture and behavior*, in «Biological Theory», 13 (2018), pp. 213-227; D.J. Buller - V. Hardcastle, *Evolutionary psychology, meet developmental neurobiology: Against promiscuous modularity*, in «Brain and Mind», 1 (2000), pp. 307-325.

construction and gene–culture coevolution, which undermine a strictly Pleistocene-fixed view and point to continued evolution into the Holocene[19] and many more.

In this section, we address a final critical point that has received relatively little attention in existing literature. As we will argue, the reification of the Environment of Evolutionary Adaptedness, combined with the limited use of paleoanthropological evidence in top-down and bottom-up approaches, has led to the formulation of empirically inadequate hypotheses. This discussion aims to contribute to a more historically grounded approach within evolutionary psychology.

We have already noted that there are two distinct ways of conceptualizing the Environment of Evolutionary Adaptedness (EEA) of the human mind. The first tends to equate it with the Pleistocene African savanna, whereas the second defines it as the set of selection pressures that shaped a given adaptation. Tooby and Cosmides claim that they only ever advocated the latter conception, and only appealed to the former in the early stages of their thinking, to make their ideas more understandable even to those without prior knowledge of evolutionary biology or those who considered all human behavior adaptive in the modern world[20]. Despite this, a stereotypical image of the EEA – as a Pleistocene savanna inhabited by small bands of hunter-gatherers – appears to have crystallized within certain research traditions of EP. Combined with the tendency to rely on superficial descriptions of the EEA while largely disregarding recent findings from paleoanthropology and prehistoric archaeology, this has the effect of diverting their heuristic strategies and fostering the development of inadequate hypotheses. We can call this phenomenon "*stylization bias*", because it involves replacing complex and up-to-date reconstructions of our evolutionary history with a crystallized and stereotypical view of human evolution. This bias does not stem from a lack of evidence about our deep past. Also, stylization bias should not be understood as a criticism of abstraction or idealization per se. Some degree of abstraction is unavoidable in any scientific attempt to reconstruct long-term evolutionary processes, and idealized models often play an important heuristic role in scientific explanation and theory construction[21]. Rather, the problem arises when idealized representations of the EEA become reified and disconnected from the

19 - K.N. Laland - J. Odling-Smee - S. Myles, *How culture shaped the human genome: bringing genetics and the human sciences together*, in «Nature Reviews Genetics», 11 (2010), pp. 137-148; R. Powell, *The Future of Human Evolution*, in «The British Journal for the Philosophy of Science», 63 (2012), pp. 145-175.

20 - K.N. Laland - G.R. Brown, *Sense and nonsense: Evolutionary perspectives on human behaviour*, Oxford University Press, Oxford 2011, p. 124.

21 - *E.g.* Michael Weisberg, *Simulation and Similarity: Using Models to Understand the World,* Oxford University Press, Oxford 2013; Angela Potochnik, *Idealization and the Aims of Science,* University of Chicago Press, Chicago 2017.

increasingly fine-grained evidence provided by paleoanthropology and archaeology. In this sense, stylization bias can be interpreted as a form of epistemically problematic idealization: a simplification that ceases to function as a useful heuristic and instead distorts the historical complexity of human evolution.

As we will see when examining the environments inhabited by our ancestors, human evolution has been (and continues to be) a complex phenomenon that resists simplistic generalizations. Over time, both our ancestors and the contexts where they lived underwent profound transformations: their habitats, diets, social practices, and much else changed dramatically. For this reason, attempts to reduce such complexity to a single statement– for instance, that "our ancestors lived in the African savanna"– are inevitably inadequate. In the next section we provide an example of this bias.

5. The savanna hypothesis

Deciding where to move, live and camp is an extremely important adaptive problem for every mobile living organism. Selecting a suitable location – one that offers a reasonable balance between shelter and camping opportunities, access to essential resources (water, food, tool-making materials, and so on), and minimal exposure to predators or other dangers – could mean the difference between life and death. This line of reasoning (step 1 of the top-down approach) led evolutionary psychologists to hypothesize, on the basis that much of our evolution took place in the African savanna (step 2), that modern humans possess a predilection for environments reminiscent of the latter. In other words, a cognitive adaptation expresses an aesthetic preference for savanna-like landscapes where our species evolved. This inclination was then tested through psychological experiments. We are thus faced with a paradigmatic case of functional analysis, where evidence about our evolutionary past serves to provide constraints during the phase of hypothesis generation. Aptly named the *savanna hypothesis*[22], this idea was famously proposed by Gordon Howard Orians in a series of articles and book chapters[23]. For Orians, the African savanna was an optimal ancestral environment

22 - This should not be confused with the other "savanna hypothesis", widely debated in paleoanthropology, which argues that the African savanna played a crucial role in shaping the evolution of bipedality in our ancestors (see, for example, M. Domínguez-Rodrigo, *Is the "savanna hypothesis" a dead concept for explaining the emergence of the earliest hominins?*, in «Current Anthropology», 55 (2014), pp. 59-81, and comments therein). Both versions of the hypothesis, however, have been challenged by recent evidence concerning the paleoecologies of our hominin ancestors.

23 - G. Orians, *Habitat selection: General theory and applications to human behavior*, in *The evolution of human social behavior*, edited by J.S. Lockard, Elsevier, Chicago 1980, pp. 49-66; G. Orians, *An ecological*

because its scattered acacia trees offered food, shelter, and vantage points, while the open landscape promoted grazing herbivores and wide visibility, aiding both resource acquisition and danger detection. Most importantly, under the "classical" conception of the EEA, the African savanna is portrayed as the environment in which "99 percent of our evolution" took place. According to this view, our lineage evolved and remained in this setting for most of its history until, between roughly 100,000 and 70,000 years ago, *Homo sapiens* left Africa and began to disperse across much of the world. Because our ancestors are thought to have lived in the savanna for the majority of their evolutionary past, it is assumed that humans developed an enduring preference for this type of environment.

Orians' hypothesis was initially supported by studies showing cross-cultural preferences for savanna-like features, such as specific acacia tree forms[24], and for savanna landscapes overall, especially among younger individuals, suggesting a possible innate bias[25]. However, subsequent research produced mixed and contradictory results with participants not consistently preferring savannas and sometimes favoring other environments like rainforests, even without direct experience, a result that cannot be fully explained by cultural learning[26]. Despite these issues, the savanna hypothesis remains influential in evolutionary psychology and is still cited as an explanation for landscape preferences in textbooks and evolutionary scholarly works[27]. At this point, it is legitimate to ask why this case

and evolutionary approach to landscape aesthetics, in *Landscape meaning and values*, edited by E.C. Penning-Rowsell - D. Lowenthal, Allen & Unwin, London 1986, pp. 3-25; G. Orians - J.H. Heerwagen, *Evolved responses to landscapes*, in *The adapted mind: Evolutionary psychology and the generation of culture*, edited by J.H. Barkow - L. Cosmides - J. Tooby, University Press, New York 1992. The latter contribution contained in *The Adapted Mind*, the manifesto of evolutionary psychology.

24 - *Ibidem.*

25 - J.D. Balling - J.H. Falk, *Development of visual preference for natural environments*, in «Environment and behavior», 14 (1982), pp. 5-28.

26 - R. Sommer, *Further cross-national studies of tree form preference*, in «Ecological Psychology», 9 (1997), pp. 153-160; D. Dutton, *Aesthetics and evolutionary psychology*, in *The Oxford handbook for aesthetics*, edited by J. Levison, Oxford University Press, New York 2003; J.H. Falk - J.D. Balling, *Evolutionary influence on human landscape preference*, in «Environment and behavior», 42 (2010), pp. 479-493; K.T. Han, *Responses to six major terrestrial biomes in terms of scenic beauty, preference, and restorativeness*, in «Environment and Behavior», 39 (2007), pp. 529-556; P. Hartmann - V. Apaolaza-Ibáñez, *Beyond savanna: An evolutionary and environmental psychology approach to behavioral effects of nature scenery in green advertising*, in «Journal of Environmental Psychology», 30 (2010), pp. 119-128; J.M. Moura - W.S. Ferreira Júnior - T.C. Silva - U.P. Albuquerque, *The influence of the evolutionary past on the mind: An analysis of the preference for landscapes in the human species*, in «Frontiers in Psychology», 9 (2018), p. 2485.

27 - I. Silverman - J. Choi, *Spatial Navigation and Landscape Preferences*, in *The Handbook of Evolutionary Psychology*, edited by D.M. Buss, John Wiley & Sons, Hoboken 2015; D. Buss, *Evolutionary Psychology: The New Science of the Mind*, cit.

study has been chosen for closer examination. After all, isn't it generally accepted that our species evolved primarily on the African savanna? Most depictions of early humans, whether in museum displays, textbooks, or documentaries, tend to situate them within environments evocative of that biome. The following section examines what paleoanthropological and paleoecological research reveal about this.

6. The paleoecology of our ancestors

Human evolution is sometimes portrayed in popular representations as an "escape" from closed wooded or forested environments, where the common ancestor between humans and chimpanzees probably lived (and where the latter still live), to the more open ones of African savannas. More open environments are thought to have favored the adoption of an upright posture, which was crucial for freeing the hands to manipulate and make tools. In turn, these tools enabled access to higher-energy resources, thereby supporting the evolution of our brain, a metabolically expensive organ. This process likely created a self-reinforcing, virtuous cycle in which technological, ecological, and cognitive changes mutually amplified one another[28].

While not entirely mistaken, some aspects of this narrative have been increasingly challenged over the past decades. In particular, the savanna does not appear to have consistently played a decisive role in human evolution. Evidence suggests that some of the earliest hominins lived in closed, wooded, or at least "mosaic" environments, where more closed and more open habitats coexisted, rather than in open savannas. The paleoecology of *Orrorin tugenensis* (~6 Mya), one of the earliest putative hominins, points in this direction[29]. The case is even clearer with *Ardipithecus ramidus*, whose fossils were discovered in Ethiopia and date to about 4.4 million years ago. Paleoenvironmental reconstructions based on its morphology, carbon isotope analyses of dental enamel, associated fauna and flora, and soil isotopic composition all point to a predominantly wooded habitat[30].

---

28 - Richard Potts dubbed these reconstructions of human evolution "intrinsic explanations", see for example R. Potts, *Environmental hypotheses of hominin evolution*, in «American Journal of Physical Anthropology», 107 (1998), pp. 93-136, for references to several of these reconstructions.

29 - M. Pickford - B. Senut, *The geological and faunal context of Late Miocene hominid remains from Lukeino, Kenya*, in «Comptes Rendus de l'Académie des Sciences-Series IIA-Earth and Planetary Science», 332 (2001), pp. 145-152.

30 - T.D. White - B. Asfaw - Y. Beyene - Y. Haile-Selassie - C.O. Lovejoy - G. Suwa - G. WoldeGabriel, *Ardipithecus ramidus and the paleobiology of early hominids*, in «Science», 326 (2009), pp. 64-86; T.D. White - C.O. Lovejoy - B. Asfaw - J.P. Carlson - G. Suwa, *Neither chimpanzee nor human, Ardipithecus reveals the surprising ancestry of both*, in «Proceedings of the National Academy of Sciences», 112 (2015), pp. 4877-

Moreover, although *A. ramidus* retained many adaptations for arboreal life, it also displayed traits specialized for bipedal locomotion, such as a shortened upper pelvis. This is particularly significant because it shows that around 4 million years ago our ancestors lived in wooded environments and were at least partly arboreal, yet were already exhibiting adaptations for bipedal walking.

Subsequently, a series of climatic changes that promoted the spread of open grasslands across Africa, particularly in East Africa, a region central to several major developments in human evolution[31]. Accordingly, members of the genus *Australopithecus* appeared to be better adapted to arid environments than *Ardipithecus*. But even during this period, evidence indicates that hominins inhabited mosaic environments combining features of both forests and savannas. Moreover, these long-term aridification trends were punctuated by alternating wet and dry phases, which caused significant fluctuations in vegetation patterns[32].

However, this seemingly privileged link with the African savanna is disrupted by *Homo ergaster/Homo erectus*, the first hominin to leave Africa and colonize parts of Asia[33]. The reasons behind these migrations are still not entirely clear, but during its dispersals *H. ergaster* encountered a wide range of environments beyond the savanna, from tropical forests[34] to highly arid deserts[35]. This demonstrates that *H. erectus* already possessed considerable ecological flexibility that, until recently, was thought to characterize only later members of our genus, such as the *Neanderthals* and *Homo sapiens*. Subsequently, *Homo heidelbergensis*, which emerged during the Middle Pleistocene[36] is thought to have been the main actor in a second major dispersal out of Africa, spreading into Europe and Asia.

4884; but see T.E. Cerling - J.G. Wynn - S.A. Andanje - M.I. Bird - D.K. Korir - N.E. Levin - C.H. Remien, *Woody cover and hominin environments in the past 6 million years*, in «Nature», 476 (2011), pp. 51-56.

31 - T.E. Cerling, *Development of grasslands and savannas in East Africa during the Neogene*, in «Palaeogeography, Palaeoclimatology, Palaeoecology», 97 (1992), pp. 241-247; T.E. Cerling - J.M. Harris - B.J. MacFadden - M.G. Leakey - J. Quade - V. Eisenmann - J.R. Ehleringer, *Global vegetation change through the Miocene/Pliocene boundary*, in «Nature», 389 (1997), pp. 153-158.

32 - N.E. Levin, *Environment and climate of early human evolution*, in «Annual Review of Earth and Planetary Sciences», 43 (2015), pp. 405-429.

33 - G.P. Rightmire, *Patterns of hominid evolution and dispersal in the Middle Pleistocene*, in «Quaternary International», 75 (2001), pp. 77-84.

34 - P. Roberts - N. Boivin - J. Lee-Thorp - M. Petraglia - J. Stock, *Tropical forests and the genus Homo*, in «Evolutionary Anthropology: Issues, News, and Reviews», 25 (2016), pp. 306-317.

35 - J. Mercader - P. Akuku - N. Boivin - A. Camacho - T. Carter - S. Clarke - P. Durkin, *Homo erectus adapted to steppe-desert climate extremes one million years ago*, in «Communications Earth & Environment», 6 (2025), p. 1.

36 - A. Profico - F. Di Vincenzo - G. Lorenza - P. Marcello - G. Manzi, *Filling the gap. Human cranial remains from Gombore II (Melka Kunture, Ethiopia; ca. 850 ka) and the origin of Homo heidelbergensis*, in

But it was primarily our species that inhabited and exploited a wide range of biomes – from open savannas and temperate woodlands to coastal areas and high-altitude environments – including tropical forests, which were also widespread in Africa, and remain poorly studied[37]. During the Pleistocene and the Holocene, African forests went through periods of expansion and contraction driven by climatic fluctuations. In many regions, this resulted in mosaic landscapes where forests alternated with savannas and grasslands. These mixed environments offered unique opportunities for hominins, allowing them to exploit the resources of both ecosystems. Far from being hostile places to avoid, as sometimes still assumed, these environments instead represented resource-rich contexts[38]. For example, the site of Panga ya Saidi (Kenya) shows that humans produced symbolic tools, practiced burials and exploited mixed forest environments 78 ka[39]. With the outflow of sapiens from Africa, we witness within not even 100,000 years the colonization of almost every biome on Earth, confirming our incredible adaptive abilities[40].

Finally, the great climatic variability of the Pleistocene should also be considered. Indeed, throughout the Pleistocene there was an alternation of cold and dry phases and warm and wet phases, which caused a series of profound environmental changes[41]. Climatic fluctuations probably favored the selection of some behavioral flexibility, rather than adaptation to a specific biome. Adaptability to different

«Journal of Anthropological Sciences», 94 (2016), pp. 41-63.

37 - E.M. Scerri - P. Roberts - S. Yoshi Maezumi - Y. Malhi, *Tropical forests in the deep human past*, in «Philosophical Transactions of the Royal Society B», 377 (2022), 20200500. There are several reasons for this. First, fossils are less likely to be preserved in tropical environments, where heat, humidity, and acidic soils accelerate decomposition. Second, Western researchers and colonizers have long regarded tropical forests as hostile, untouched wilderness, which has contributed to a lasting bias against studying them (Patrick Roberts, personal communication).

38 - P. Roberts - M. Petraglia, *Pleistocene rainforests: barriers or attractive environments for early human foragers?*, in «World Archaeology», 47 (2015), pp. 718-739.

39 - C. Shipton - C. Clarkson - R. Cobden, *Were Acheulean bifaces deliberately made symmetrical? Archaeological and experimental evidence*, in «Cambridge Archaeological Journal», 29 (2018), pp. 65-79. M. Martinón-Torres - F. d'Errico - E. Santos - A. Álvaro Gallo - N. Amano - W. Archer - M.D. Petraglia, *Earliest known human burial in Africa*, in «Nature», 593 (2021), pp. 95-100.

40 - K.E. Westaway - J. Louys - R.D. Awe - M.J. Morwood - G.J. Price - J.X. Zhao - B. Sulistyanto, *An early modern human presence in Sumatra 73,000-63,000 years ago*, in «Nature», 548 (2017), pp. 322-325; reports one of the oldest non-African sites where *sapiens* are attested in a tropical forest, while we refer to the article by Scerri et al. 2022 for an overview of the most recent findings on the occupation of tropical African and non-tropical forests during prehistoric times).

41 - L.E. Lisiecki - M.E. Raymo, *A Pliocene Pleistocene stack of 57 globally distributed benthic δ18O records*, in «Paleoceanography», 20 (2005); T.D. Herbert, *The mid-Pleistocene climate transition*, in «Annual Review of Earth and Planetary Sciences», 51 (2023), pp. 389-418.

environmental conditions is also highlighted by the various waves of migrations that our ancestors underwent. Building on this, several theoretical models try to explain the relationship between the evolution of our ancestors and the environment. In recent years, the "*variability selection hypothesis*", proposed by Richard Potts, has had particular resonance[42]. According to Potts, fundamental changes in human evolution occurred concurrently with these phases of increased instability and phase reversal and have increased the ecological flexibility and adaptability of our ancestors to a range of different environments[43].

In summary, human evolution did not unfold exclusively in open environments such as the savanna; rather, our ancestors were capable of dispersing into and colonizing a wide range of habitats. Therefore, while the savanna undoubtedly played an important role in our evolutionary history, there is no compelling reason to regard it as our preferred biome or as the one in which we felt safest. At the present state of knowledge, it does not appear that most human adaptations have evolved to respond to the demands of a particular environment, so much as to respond flexibly and adaptively to a very variable ecological context[44].

In light of the evidence presented, the savanna hypothesis, as intended by evolutionary psychologists, is clearly inadequate. The heuristic process (the functional analysis) that generated it is based on a stereotypical view of human evolution that equates our evolutionary past with life in African savannas, and remains largely uninformed by paleoecological data. Recall that functional analysis starts by identifying an adaptive problem – in this case, determining the ideal habitat for human survival (step 1) – and then reconstructing how that problem manifested in our evolutionary past (step 2). It is precisely at this stage that the main difficulties arise when it comes to integrating paleoanthropological and paleoenvironmental evidence. As we have seen, most evolutionary psychologists argue that the African savanna provided the primary stage for human evolution.

42 - R. Potts, *Environmental hypotheses of hominin evolution*, cit; R. Potts, *Evolution and environmental change in early human prehistory*, in «Annual Review of Anthropology», 41 (2012), pp. 151-167, see E.S. Vrba, *Role of environmental stimuli in hominid origins*, in *Handbook of paleoanthropology*, edited by W. Henke - I. Tattersall, Springer, Berlin, Heidelberg 2015, pp. 1837-1886, for a complementary account of the relationships between human evolution and the unstable climatic and ecological conditions.

43 - For an overview of the role of ecological patterns in human evolution see also A. Parravicini - T. Pievani, *Multi-level human evolution: ecological patterns in hominin phylogeny*, in «Journal of Anthropological Sciences», 94 (2016), pp. 167-182.

44 - To further explore the wide range of environments in which our ancestors lived, see T.I. Foister - I. Žliobait - O.E. Wilson - M. Fortelius - M. Tallavaara, *Homo heterogenus: Variability in early Pleistocene Homo environments*, in «Evolutionary Anthropology: Issues, News, and Reviews», 32 (2023), pp. 373-385; E. Zeller - A. Timmermann - K.S. Yun - P. Raia - K. Stein - J. Ruan, *Human adaptation to diverse biomes over the past 3 million years*, in «Science», 380 (2023), pp. 604-608.

From this assumption, they infer that we should have developed an enduring preference for such environments, a preference that should still be observable in humans today. While this reconstruction may have seemed plausible in the '80s when the hypothesis was first proposed, it is no longer tenable in light of paleoanthropological evidence, which now points to a far more complex evolutionary scenario. Unfortunately, the tendency of evolutionary psychologists to underutilize paleoanthropological evidence and to rely on overly simplified reconstructions of our past, what we deemed the stylization bias, has allowed the savanna hypothesis to persist, albeit with mixed results. The inconsistent results of landscape-preference experiments can be better understood by recognizing that human evolution occurred in diverse habitats, shaped by adaptability and ecological versatility rather than specialization in a single biome.

But how should we revise the functional analysis underlying the savanna hypothesis in light of these remarks? Step 1 is not controversial in itself; choosing the right place to live is crucial for survival. The problem arises with step 2, which should draw on all the relevant evidence provided by paleoanthropology and, in this case, paleoenvironmental reconstructions. Rather than relying on a stereotyped account of human evolution, we should acknowledge its complexity and the environmental diversity in which our ancestors lived. If we were to preserve the same inferential structure as the original model, we might instead expect humans to show a preference for mosaic environments, a conclusion also reached by Rathmann and colleagues[45] or, more generally, for natural settings over heavily anthropized ones such as large cities. Alternatively, this preference could reflect a broader adaptive strategy for identifying suitable habitats without privileging any specific environment.

7. How to revise the methods of evolutionary psychology

In section 3, when presenting the different heuristics of hypothesis generation, we showed that the founders of evolutionary psychology regard the top-down method as immune to accusations of producing just-so stories. This supposed immunity stems from the method's predictive nature: it begins with the formulation of a hypothesis about a cognitive adaptation, which is then tested against empirical evidence, thereby avoiding the construction of *ad hoc* explanations. In the last sections, we have shown that this immunity is far from guaranteed: if evidence from our past is not properly integrated into the process of hypothesis generation,

45 - J. Rathmann - K.M. Korpela - P. Stojakowits, *Pleistocene hypothesis-moving savanna perceptual preference hypothesis beyond savanna*, in «Frontiers in Psychology», 13 (2022), 901799.

there is a risk of producing empirically inadequate hypotheses. One might object that, in functional analysis, hypotheses are in any case tested against empirical evidence, and therefore can be refined or corrected through this process. This objection would be valid if evolutionary psychologists were willing to revise their hypotheses in light of empirical testing. But this is not what happened with the savanna hypothesis. Despite mixed results, the hypothesis continues to thrive in evolutionary psychologists' works.

Because of the stylization bias, EP hypotheses, compared to those of paleoanthropology, have typically a low *spatial*, *temporal* and *evolutionary* "resolution": they often treat the entire evolutionary history of humankind as a single, homogeneous block, overlooking the numerous transformations our lineage has undergone across space, time and phylogeny – from the succession of different species, to shifting environmental and ecological conditions, to the growing diversity and complexity of human lifeways. Paleoanthropology and archaeology, on the other hand, operate at a much finer resolution. Although EP must to some extent abstract from certain details to keep the generation process viable (it would not be feasible to consider all the evidence about our past), the level of abstraction caused by stylization bias leads to a hypothesis generation process that is inadequate.

But how should evolutionary psychologists revise functional analysis? We have seen that Tooby and Cosmides propose a process with six steps. In this partition of functional analysis, historical evidence is relevant to the second step: identification of a recurrent adaptive problem and characterize how the adaptive problem manifested itself in our evolutionary past. For these steps to be successfully completed, it is necessary to adequately reconstruct the EEA of an adaptation, that is, the set of selective pressures that shaped it. To do so, we should ideally do two things:

> Outline the period of interest. As already mentioned, not all adaptations will have the same EEA. To identify the relevant period, an initial review of paleoanthropological literature is already necessary. For example, meat consumption plays a fundamental role in hypotheses about the evolution of disgust: the evolution of this primordial emotion would be linked to proximity to carcasses, as scavengers and/or predators[46]. But when did the systematic consumption of meat begin? In other words, what is the EEA of disgust?

46 - D.R. Kelly, *Yuck! The nature and moral significance of disgust*, MIT Press, Cambridge (MA) 2011.

> Evolutionary psychology alone cannot answer this question; it needs paleoanthropological evidence[47].

Select all available relevant evidence. This step clearly depends on the adaptive problem we are investigating. If, for example, we are investigating the assumed differences in partner choice strategies between men and women, any data about the possible social structure of our ancestors and their evolution over time is pertinent. Number of individuals per group, sexual dimorphism, ratio between males and females, the presence of different roles, group mobility and intra-group migration. But also patterns of development and other data that can give us clues as to how the burden of childcare was distributed, and many more. In the case of the savanna hypothesis, paleoecological evidence was pertinent. (Historical evidence is then relevant for the following steps as well, as they require details of what informational resources were available to our ancestors, to infer what specifics of the cognitive mechanism evolved).

Of course, this process is not straightforward. Our knowledge of the past is extremely fragmentary, but it is rapidly increasing and being revised, which makes it difficult to stay up to date when conducting research in another discipline. At the same time, understanding what is relevant for reconstructing the EEA of a particular adaptation is not an easy task. In this regard, greater interdisciplinarity and collaboration with paleoanthropologists and archaeologists could be beneficial. Evolutionary psychologists often claim that the overarching aim of their discipline is the unification of all the human evolutionary sciences. Regardless of whether this undertaking is ultimately feasible, incorporating evidence from these disciplines appears to be one of the minimum requirements for pursuing this goal.

Historical evidence plays a similar role in reverse engineering. In this case, we start with the observation of a behavioral or psychological pattern in modern humans, but knowledge of the past is still required to work out the adaptive function of that particular trait. What is the function of jealousy (if any)? We cannot know unless we reason in terms of the specifics of the context in which it would have evolved. Moreover, paleoanthropological evidence is relevant for another reason. We mentioned that, according to Gould and Lewontin, adaptationist bias improperly leads researchers to divide organisms into independent traits. This issue also arises in evolutionary psychology, where the choice of a particular character as a trait with a separate evolutionary history is rarely justified. Considering the archaeological and fossil record can help evolutionary psychologists

47 - If the EEA of a particular adaptation for example calls into question primate evolution more generally (think of the attachment relationship between mother and child) then evidence from primate paleontology and primatology also become relevant.

reconstruct the history of certain characters and determine which are genuine individual traits and which are not. For example, Hauser and colleagues proposed that the language faculty should not be regarded as a single character but as a mosaic of different traits with distinct evolutionary histories[48] – a view consistent with the fossil and archaeological record[49]. For this reason, an evolutionary psychologist applying reverse engineering to the language faculty may actually be analyzing a composite of different characters, and thus working at the wrong level of analysis.

One final consideration is in order: fully integrating the evidence about our past risks being deeply disruptive to the methods of evolutionary psychology. The increasing complexity of reconstructions of our evolution seems to indicate a certain instability in the selective pressures to which we have been subjected, especially with regard to certain aspects of group life and adaptation to the environment. Identifying a cheater, collaborating with one's companions, finding the best place to set up camp, recognizing environmental dangers, etc. are problems that have changed continuously throughout our evolution. If selective environments changed repeatedly throughout the Pleistocene, then selection may have favored behavioral flexibility, ecological versatility, and broad learning capacities. On the other hand, evolutionary psychologists traditionally assumes that many human cognitive mechanisms evolved as domain-specific solutions to recurrent and relatively stable adaptive problems[50]. The inclusion of paleoanthropological evidence, therefore, risks being incompatible with evolutionary psychology, at least in its more traditional formulations. However, this need not necessarily be the case. In recent years, some evolutionary psychologists have become increasingly open to the possibility that cognitive modules may be more flexible and context-sensitive than originally assumed[51]. Should evolutionary psychology move in this

48 - M.D. Hauser - N. Chomsky - W.T. Fitch, *The faculty of language: what is it, who has it, and how did it evolve?*, in «Science», 298 (2002), pp. 1569-1579.

49 - A. Parravicini - T. Pievani, *Continuity and discontinuity in human language evolution: putting an old-fashioned debate in its historical perspective*, in «Topoi», 37 (2018), pp. 279-287; A. Parravicini - T. Pievani, *Mosaic evolution in hominin phylogeny: meanings, implications, and explanations*, in «Journal of Anthropological Sciences», 97 (2019), pp. 45-68.

50 - A perspective shared by several critics, see for example K. Sterelny - P.E. Griffiths, *Sex and death: An introduction to philosophy of biology*, University of Chicago Press, Chicago 1999.

51 - Consider, for example, how recent research in evolutionary psychology on mate-choice mechanisms has placed far greater emphasis on variability and sensitivity to a wide range of contextual factors. David Buss, *The Sexual Selection of Human Mating Strategies: Mate Preferences and Competition Tactics*, in *The Oxford Handbook of Evolutionary Psychology and Romantic Relationships*, a cura di Justin Mogilski e Todd Shackelford, Oxford University Press, Oxford 2023.

direction, its hypotheses may ultimately prove more compatible than previously thought with the growing body of evidence concerning human evolutionary history.

8. Conclusions

We have seen that, despite its relevance, evidence from paleoanthropology and archaeology is not always integrated into disciplines studying human cognitive and behavioral evolution. Focusing on evolutionary psychology and the savanna hypothesis, we have shown that its formulation overlooks paleoenvironmental evidence about the contexts of human evolution. Consequently, evolutionary psychologists may be testing empirically inadequate hypotheses[52]. Although there appear to be tentative signs of renewal within EP[53], this remains an issue that evolutionary psychologists must address. Indeed, some evolutionary psychologists seem to have recently taken a self-critical stance, admitting that the discipline should take archaeological and paleoanthropological evidence more seriously[54]. And ultimately, evolutionary psychologists cannot get away from doing this if they really want to achieve that unification of multiple fields of knowledge that they so yearn for.

The same concerns also apply to other related disciplines. As already mentioned, Meneganzin and Currie recently emphasized that human behavioral ecology needs to integrate evidence from paleoanthropology in order for its hypotheses to be historically sound[55]. Others have expressed the importance that models of cultural evolutionary theory be grounded in and tested against the archaeological record[56].

---

52 - To be clear, we are not claiming that every hypothesis in evolutionary psychology suffers from this problem; but rather that if historical evidence is not taken into account during hypothesis generation, then we risk formulating inadequate hypotheses.

53 - H.C. Barrett, *The Shape of Thought: How mental adaptations evolve*, Oxford University Press, Oxford 2015; see also G.R. Brown - K.N. Laland, *Sense and nonsense: Evolutionary perspectives on human behaviour*, Oxford University Press, Oxford3 2024.

54 - D.M.T. Fessler - J.A. Clark - E.K. Clint, *Evolutionary psychology and evolutionary anthropology*, in *The handbook of evolutionary psychology: Integrations*, 2nd edition, edited by D.M. Buss, John Wiley & Sons, Inc., Hoboken 2016, pp. 1029-1046; M.A. Abramiuk, *Evolutionary Psychology and Archaeology*, in *The SAGE Handbook of Evolutionary Psychology. Integration of Evolutionary Psychology with other Disciplines*, edited by T.K. Shackelford, SAGE Publications, London 2021, pp. 257-282.

55 - A. Meneganzin, Andra - A. Currie, *Not wasted on the young: Childhood, trait complexes & human behavioral ecology*, cit.

56 - See for example R. Garvey, *Current and potential roles of archaeology in the development of cultural evolutionary theory*, in «Philosophical Transactions of the Royal Society B: Biological Sciences», 373 (2018), 20170057; G. Tostevin, *Content Matters: the materiality of cultural transmission and the intersection of Paleolithic archaeology with cultural evolutionary theory*, in *Beyond the Meme: Articulating*

In any case, disciplines concerned with the evolution of the human mind cannot afford to treat paleoanthropology and archaeology as mere background sources of illustrative examples. If hypotheses about human cognition are meant to be genuinely evolutionary and historically grounded, they must be constrained by the best available evidence about the environments, ecologies, and lifeways of past hominins. Otherwise, the risk is that stylized narratives of human evolution will continue to replace the far more complex history revealed by the historical sciences. These fields have long drawn on insights and methods from other sciences, yet they also have much to contribute in return.

MATTEO BEDETTI
PhD in Research Methods in Science and Technology from the University of Urbino Carlo Bo. His research interests include the methods of historical sciences, and in particular, how paleoanthropology and archaeology deal with epistemically unfavorable situations.
m.bedetti1@campus.uniurb.it

Dynamic Structures in Evolution, edited by A.C. Love - W.C. Wimsatt, University of Minnesota Press, Minneapolis 2019, pp. 311-364.

Beauty as Divine Order: Berkeley on Proportion, Reason, and Subordination

di Aurora Yu

This paper reconstructs George Berkeley's theory of beauty as a coherent philosophical account grounded in relations of proportion, subordination, and teleological fittingness. It argues that beauty, for Berkeley, spans both natural and moral domains under a single unifying principle – the rational ordering of parts within a hierarchy of ends – and that this commitment leads to a qualified but significant rejection of the moral sense tradition in favor of a view on which beauty is an object of rational apprehension, ultimately grounded in divine reason.

Keywords: Divine Order, Berkeley, Proportion, Reason

1. Introduzione

Early modern philosophy is centrally concerned with the problem of order: how a world composed of diverse and heterogeneous elements can nevertheless exhibit unity, intelligibility, and normativity. In this context, the role of God is typically understood in terms of causation, epistemic guarantee, or metaphysical grounding. Yet an equally important, though often neglected, dimension of this problem concerns how such order becomes perceptible to finite minds. This paper argues that, for Berkeley, beauty is precisely the mode under which the rational structure of a divinely ordered world becomes accessible to human cognition.

Berkeley's scattered remarks on beauty – primarily in *Alciphron*, but also in the *Three Dialogues*, *Passive Obedience*, the *Principles*, and *Siris* – have often been treated as marginal or unsystematic. When discussed, they are typically assimilated either to early modern moral sense theories or to broader empiricist accounts of aesthetic response[1]. Such readings, however, obscure a central feature of Berkeley's view: beauty is not grounded in feeling, but in intelligible relations of order, proportion, and subordination, relations that ultimately reflect the teleological structure of creation.

In *Alciphron*, the character Alciphron defends a "moral sense" theory, according to which beauty is apprehended by a special faculty analogous to the senses: On this view, beauty is to be felt rather than understood. It attracts us without reason.

Berkeley, through Euphranor, rejects this account. The problem with the moral sense view is that it cannot distinguish between apparent and real beauty. If beauty is simply what pleases, then there is no standard for correcting mistaken judgments. Berkeley insists instead on an analogy between truth and beauty:

> They [truth and beauty] proceed from nature, and are conformable to right reason? Truth and beauty are in this alike, that the strictest survey sets them both off to advantage while the false lustre of error and disguise cannot endure being reviewed, or too nearly inspected[2].

This analogy implies that beauty, like truth, is objective and intelligible. It can withstand scrutiny because it is grounded in real relations. The distinction between apparent and true beauty is crucial. What pleases us is not always truly beautiful. Human beings are often guided by self-interest and immediate gratification, which can obscure the perception of real order. True beauty, by contrast, is what is genuinely good within the system as a whole.

---

1 - See, e.g., discussions of Berkeley in relation to Shaftesbury and Hutcheson in Jerome Stolnitz, *Aesthetics and the Philosophy of Art Criticism* (1960); also Guyer (2005).

2 - *Dialogue* II.

This paper reconstructs Berkeley's theory of beauty as a unified account spanning both natural and moral domains. I argue for three claims. First, Berkeley distinguishes between proper beauty (the beauty of visible and natural objects) and moral beauty (the beauty of actions and virtues), while insisting on a deep structural parallel between them. Second, in both domains, beauty consists in relations of proportion relative to ends, where parts are subordinated to wholes and coordinated within a system. Third, contrary to the moral sense theorists he engages, Berkeley holds that beauty is not immediately felt but rationally apprehended, and that this rational apprehension reveals the order of a world structured by divine reason.

On this interpretation, beauty plays a distinctive epistemic role. It is not merely an object of perception, but a way of perceiving order as such. To judge something beautiful is to recognize its place within a system of ends and its contribution to the harmony of the whole. In this way, aesthetic judgment becomes a site at which metaphysics, epistemology, and theology converge: it is through beauty that finite minds apprehend the intelligibility of divine creation.

This reading both builds on and departs from existing scholarship. Commentators such as Winkler and Stoneham have emphasized the role of divine language and signification in Berkeley's metaphysics, while Downing has stressed the importance of relational properties in his ontology[3]. Yet the aesthetic dimension of these relations – their appearance as beauty – has received comparatively little attention. Recent scholarship has begun to register precisely this gap. As Endre Szécsényi observes, Berkeley is "usually not discussed in the canonical histories of modern aesthetics," and even Berkeley specialists "do not seem to have paid attention to his possible contribution to modern aesthetics"[4]. This neglect is not accidental. Berkeley does not present a formal aesthetic treatise; rather, his reflections on beauty are dispersed across dialogues, sermons, and practical writings, and are "profoundly interwoven with his philosophy of language and his (popular) theology"[5]. As a result, the aesthetic dimension of his thought has often remained implicit, overshadowed by epistemological and metaphysical concerns. The present paper builds on this recent reassessment, but departs from it in one crucial respect. Whereas Szécsényi emphasizes Berkeley's contribution to an emerging "theologico-aesthetic language," the argument here is that Berkeley offers not merely a

3 - Cf. K. P. Winkler, Berkeley: An Interpretation, Clarendon Press, Oxford 1989; T. Stoneham, Berkeley's World: An Examination of the Three Dialogues, Oxford University Press, Oxford 2002; L. Downing, George Berkeley, Oxford University Press, Oxford 2005.

4 - E. Szécsényi, "The Aesthetics of the Invisible: George Berkeley and the Modern Aesthetics," in «History of European Ideas», 48 (2022) n. 6, pp. 731–743, here p. 731.

5 - E. Szécsényi, "The Aesthetics of the Invisible," cit., p. 732.

language or mode of discourse, but a coherent, if dispersed, account of beauty grounded in relations of proportion, subordination, and rational order. Beauty, on this reading, is not an ancillary feature of Berkeley's philosophy, but a central category through which the intelligibility of divine order becomes accessible to finite minds. By foregrounding this dimension, the present paper shows that Berkeley's aesthetics is not an ancillary doctrine, but an integral part of his broader philosophical system.

The argument proceeds by developing a unified account of beauty across both natural and moral domains. Section 2 begins by clarifying Berkeley's distinction between proper and moral beauty, showing that this distinction does not mark two fundamentally different kinds of phenomena, but two domains structured by a single principle: proportion relative to ends. Section 3 then turns to natural beauty, arguing that Berkeley conceives it as the perception of a system of mutually dependent parts whose coordination reveals an underlying order. Section 4 extends this framework to the moral domain, where beauty consists in the rational ordering of action within a system governed by stable and intelligible rules. Section 5 draws out the philosophical stakes of this account by examining Berkeley's critique of the moral sense tradition and his insistence that beauty, like truth, is an object of rational apprehension rather than immediate feeling. Section 6 brings these strands together by articulating the concept of subordination as the unifying structure of beauty, and by showing how apparent imperfections can be understood as contributing to the harmony of a larger whole.

2. Proper and Moral Beauties: a Single Structure

Berkeley's account of beauty begins from a distinction that is at once intuitive and theoretically loaded: the distinction between proper beauty and moral beauty. In Alciphron, this distinction is introduced in a deceptively simple formulation:

> As the eye perceiveth the one, so the mind doth by a certain interior sense perceive the other[6].

At first glance, this passage appears to align Berkeley with the moral sense tradition. Proper beauty – the beauty of visible, corporeal objects – is perceived by the eye, while moral beauty – the beauty of virtue – is perceived by an "interior sense". Read in isolation, this suggests a familiar early modern picture: just as colors are given to the external senses, moral qualities are given to an internal faculty of feeling.

---

6 - *Alciphron* 3.3.

Yet such a reading is ultimately misleading. As we shall see, Berkeley introduces this formulation only to reject the moral sense model it seems to support. The distinction between proper and moral beauty is real, but it does not correspond to a division between sensation and feeling. Rather, it marks two domains in which a single underlying structure – proportion relative to ends – is instantiated.

To see this, we must first clarify what is meant by proper beauty. Berkeley's discussion makes clear that proper beauty is not restricted to works of art. On the contrary, its primary domain is nature itself. As the materials rightly emphasize, natural beings are not merely included among beautiful objects; they are paradigmatic instances of beauty. The opening scenes of Berkeley's dialogues, with their detailed descriptions of landscapes, are not ornamental but philosophically significant. They direct our attention to a world in which beauty is pervasive and structurally grounded.

Proper beauty, then, is not merely visual in a superficial sense. It is not reducible to color, shape, or immediate sensory appeal. Rather, it concerns the way in which objects are situated within a system of relations. A tree is beautiful in a forest not simply because of its form, but because of its role within a network of mutual dependencies: it shelters birds, supports growth at its roots, and contributes to the overall harmony of its environment. The same object, removed from this context, loses its beauty – not because its intrinsic features change, but because its relations do.

Moral beauty, by contrast, concerns the domain of action and character. Yet Berkeley insists that it is structurally analogous to proper beauty. As Euphranor puts it, moral beauty consists in:

«A beauty of the mind, a charm in virtue, a symmetry and proportion in the moral world»[7].

The key terms – symmetry, proportion, order – are the same. What differs is not the structure, but the domain of application. In the natural world, beauty arises from the coordination of physical parts; in the moral world, it arises from the coordination of rational actions. In both cases, beauty consists in fittingness within a system.

This parallelism is central to Berkeley's view. It allows him to unify aesthetics across domains while preserving the distinctiveness of each. Proper beauty and moral beauty are not two fundamentally different kinds of phenomena, but two expressions of a single principle: the principle that beauty consists in the ordered relation of parts to ends and to wholes.

7 - *Alciphron* 3.4

At this point, however, Berkeley introduces a competing account, represented by the character Alciphron. According to this view, moral beauty is grounded in a special faculty: "there is an idea of beauty natural to the mind of man," and "men's first thoughts and natural notions are the best in moral matters"[8]. On this account, moral beauty is not the object of reason, but of immediate intuition or feeling. It is apprehended through a moral sense that operates independently of deliberation. This position places Berkeley in dialogue with figures such as Shaftesbury and Hutcheson, for whom moral and aesthetic judgments are grounded in a distinctive affective faculty. It is worth noting that Berkeley does not simply dismiss this tradition. He shares with Shaftesbury the conviction that beauty is objective and not merely reducible to private whim, and with Hutcheson the recognition that aesthetic response can be immediate and not always the product of explicit inference. His disagreement concerns the ground of this objectivity: whether it lies in a divinely calibrated affective faculty, or in intelligible relations that are, at least in principle, accessible to reason.

The force of Berkeley's disagreement with the moral sense tradition becomes clearer when situated against Hutcheson's own account. Hutcheson famously grounds beauty in a special internal sense, such that objects are called beautiful "in virtue of their causing us to have a certain kind of pleasurable feeling"[9], and this connection between form and pleasure is itself contingent upon "some Constitution of the AUTHOR of our Nature"[10]. On this view, beauty is not intelligible through reason alone, but depends on a divinely instituted correspondence between our sensibility and the world.

Berkeley's objection can thus be seen as targeting precisely this contingency. If beauty depends on a merely given alignment between human feeling and external forms, then its normative force becomes unstable: what appears beautiful would lack any independent standard. By contrast, Berkeley insists that beauty is grounded in relations that can withstand rational scrutiny. The disagreement is therefore not merely about the faculty by which beauty is perceived, but about whether beauty itself is fundamentally affective or intelligible.

Berkeley's strategy is not to deny that beauty is sometimes felt, but to argue that this account fails to distinguish between apparent and real beauty. If beauty is identified with what pleases, then any pleasurable object would count as beautiful, and there would be no standard for correcting mistaken judgments. Yet Berkeley clearly wants to preserve such a standard. Not everything that appears beautiful is

8 - *Alciphron* 3.3.

9 - Kivy 2003, p. 112.

10 - Kivy 2003, p. 113.

truly so. This point marks the beginning of a decisive shift. While Alciphron treats the distinction between proper and moral beauty as a distinction between kinds of perception (external vs internal), Berkeley reconceives it as a distinction between domains governed by the same rational structure. Both kinds of beauty involve relations of proportion and order. Both require the apprehension of how parts are fitted to ends. And both ultimately depend on a system whose intelligibility exceeds immediate perception.

The distinction between proper and moral beauty, therefore, is not foundational but derivative. It reflects two ways in which a single underlying order is manifested. The eye and the mind do not disclose fundamentally different kinds of beauty; rather, they provide different modes of access to the same structure. This reinterpretation allows Berkeley to position himself against the moral sense tradition while preserving its insights. He agrees that beauty is not reducible to explicit reasoning in every case; we often recognize beauty immediately. But he denies that this immediacy is primitive. It is, instead, the result of a tacit grasp of relations that can, in principle, be made explicit. Beauty is not a brute feeling, but an intelligible relation that may be more or less clearly apprehended.

In this way, the distinction between proper and moral beauty becomes the starting point for a more ambitious claim: that beauty, in all its forms, is grounded in the rational structure of a world ordered by divine reason. The next section develops this claim by examining Berkeley's account of natural beauty in greater detail.

3. Natural Beauty and the System of the World

If Berkeley's distinction between proper and moral beauty establishes a shared structure, it is in his treatment of natural beauty that this structure becomes fully visible. For Berkeley does not begin with abstract definitions, but with a striking methodological move: he begins with experience itself, and in particular with the experience of natural beauty.

His philosophical dialogues famously open not with arguments, but with scenes. In the *Three Dialogues,* Philonous begins:

> Can there be a pleasanter time of the day, or a more delightful season of the year? That purple sky, those wild but sweet notes of birds, the fragrant bloom upon the trees and flowers, the gentle influence of the rising sun, these and a thousand nameless beauties of nature inspire the soul with secret transports[11].

11 - Dialogue I.

Similarly, The Minute Philosopher "began upon the beauty of this rural scene, the fine season of the year, and some late improvements which had been made the adjacent country by new methods of agriculture"[12]. These openings are not merely rhetorical or decorative. They perform an essential philosophical function. They direct our attention to a world in which beauty is not rare or exceptional, but pervasive and structured. What is at issue is not whether nature is beautiful, but how its beauty is to be understood.

This question is immediately deepened by Philonous' later remarks:

> Is there not something in the woods and groves, in the rivers and clear springs, that soothes, that delights, that transports the soul? At the prospect of the wide and deep ocean, or some huge mountain whose top is lost in the clouds, or of an old gloomy forest, are not our minds filled with a pleasing horror? Even in rocks and deserts, is there not an agreeable wildness? How sincere a pleasure is it to behold the natural beauties of the earth [...]. What delicacy, what beauty, what contrivance, in animal and vegetable bodies! How exquisitely are all things suited, as well to their particular ends, as to constitute opposite parts of the whole! And while they mutually aid and support, do they not also set off and illustrate each other? [...]. The motion and situation of the planets, are they not admirable for use and order? [13].

This passage reveals that Berkeley's account of natural beauty has a twofold structure: each thing is suited to its own end and each thing contributes to a larger whole. Beauty, therefore, is not merely local but systemic. It does not reside in isolated objects, but in the way those objects are integrated within a network of relations. The phrase "mutually aid and support" is especially important. It indicates that natural beauty is not additive, but integrative. The elements of nature are not simply juxtaposed; they are interdependent. Their relations form a structured system in which each part contributes to the functioning of the whole. This point is reinforced in *Alciphron* 1.16:

> Even the inanimate unorganized elements seem to have an excellence relative to each other. Where was the excellency of water, if it did not cause herbs and vegetables to spring from the earth, and put forth flowers and fruits? And what would become of the beauty of the earth, if it was not warmed by the sun, moistened by water and fanned by air? Throughout the whole system of the visible and natural world, do you not perceive a mutual connexion and correspondence

12 - *Alciphron* 1.1.

13 - *Dialogue* II.

> of parts? And is it not from hence that you frame an idea of the perfection, and order, and beauty of nature? [14].

Here Berkeley explicitly names what was implicit before: "the whole system of the visible and natural world." Beauty arises from the "connexion and correspondence of parts" within this system. It is not an intrinsic feature of individual things, but a relational feature of the whole.

At this point, however, Berkeley introduces a crucial philosophical clarification. The system described above is not directly given in sensation. While the senses present colors, sounds, and textures, they do not present relations of order as such. The recognition of harmony requires something more: "beauty is an object, not of the eye, but of the mind"[15]. To perceive beauty is to grasp relations – relations of dependence, coordination, and proportion. This requires intellectual activity. The senses provide the material, but the mind discerns the structure. This insight is central to Berkeley's view. It establishes that aesthetic experience is not merely sensory, but cognitive. Beauty is not simply felt; it is understood. This intellectual dimension also explains why Berkeley repeatedly links beauty with truth. Both involve conformity to a structure that is not of our making. Just as truth consists in the correspondence of ideas with reality, beauty consists in the conformity of parts to an order.

From this point, the theological implication follows naturally. If beauty consists in the perception of systemic order, and if that order is not produced by human minds, then it must be grounded in a higher source of intelligibility. The harmony of nature is not accidental; it is the result of a purposive structure instituted by God. In this respect, Berkeley's account stands in a complex relation to contemporaneous theological aesthetics. Hutcheson, for example, treats aesthetic experience as evidence of divine design: the striking correspondence between our sense of beauty and the order of the world would be "extremely unlikely" if it arose from chance, and is therefore best explained by a rational creator[16]. Yet for Hutcheson, this argument presupposes that beauty is first and foremost a matter of pleasurable response. Berkeley reverses this order of explanation. Beauty is not evidence for divine design because it pleases us; rather, it pleases us because it reflects an intelligible structure grounded in divine reason. The perception of beauty is thus not merely a sign of design, but a partial apprehension of it. What is revealed in aesthetic experience is not simply that the world is well suited to us, but that it is ordered according to principles that can be grasped, however imperfectly, by the

14 - *Alciphron* 1.16.

15 - *Alciphron* 3.8.

16 - Kivy 2003, p. 113.

intellect. Beauty, therefore, is not merely an aesthetic phenomenon. It is a mode of access to divine order. Through the perception of natural beauty, finite minds apprehend – however partially – the intelligibility of creation.

4. Moral Beauty, Interest, and Rational Order

The transition from natural to moral beauty does not introduce a new structure, but a new standpoint. In the natural world, we perceive order; in the moral world, we are called upon to participate in and realize it.

Berkeley makes this parallel explicit by extending the language of proportion and system to the moral domain. Just as natural beauty arises from the coordination of parts within a whole, so moral beauty arises from the coordination of actions within a system of rational beings. As he writes in *Passive Obedience:*

> As without such a steadiness in nature, we should soon… see nothing but a disorderly and confused chaos; so if once it become current, that the moral actions of men are not to be guided by certain definite inviolable rules, there will be no longer found that beauty, order, and agreement, in the system of rational beings… but… darkness and violence[17].

The analogy is exact. Both the natural and moral worlds are systems, and both derive their beauty from the maintenance of order within those systems.

Yet the moral domain introduces a complication: the role of agency. Unlike natural objects, rational beings are not automatically coordinated. Their actions must be guided. This raises the question: how is moral beauty to be apprehended and realized?

It is here that Berkeley confronts the moral sense theory. According to Alciphron, moral beauty is "an object, not of the discursive faculty, but of ... the moral sense, being adapted to the perception of moral beauty as the eye to colors, or the ear to sounds"[18]. On this view, beauty is immediate and non-rational. It is something we feel rather than understand. This position is reinforced by the claim that: "interest is a mean ungenerous thing, destroying the merit of virtue"[19], and that beauty depends on a certain "je ne sais quoi"[20].

Berkeley's response is to show that this account cannot distinguish between apparent and real beauty. If beauty is simply what pleases, then there is no standard for correction. Yet such a standard is clearly required. He therefore

---

17 - *Discourse of Passive Obedience*, Chapter XXVIII.

18 - *Alciphron* 3.5.

19 - *Alciphron* 3.3.

20 - *Alciphron* 3.5.

replaces the moral sense view with a rational account of moral beauty, grounded in judgment and interest properly understood:

> duty and virtue are in a fairer way of being practiced, if men are led by reason and judgment, balancing low and sensual pleasures with those of a higher kind, comparing present losses with future gains, and the uneasiness and disgust of every vice with the delightful practice of the opposite virtue[21].

Here, interest is not rejected, but transformed. It becomes a rational evaluation of goods across time and across levels of value. Moral beauty consists in the correct ordering of these goods – in aligning one's actions with what is truly best within the system as a whole. This account is reinforced in *Principles §CIX:*

> we should propose to ourselves nobler views, such as to recreate and exalt the mind, with a prospect of the beauty, order, extent, and variety, of natural things, hence, by proper inferences… to make several parts of the creation, so far as in us lies, subservient to the ends they were designed for[22].

This passage reveals the full structure of Berkeley's view. Moral action is not merely rule-following; it is the active realization of teleological order. To act beautifully is to make oneself and one's actions subordinate to the ends embedded in creation. Thus, moral beauty is not a matter of feeling, but of rational participation in divine order.

Berkeley's most explicit account of beauty appears in Alciphron, where Euphranor states:

> There is no beauty without proportion, so proportions are to be esteemed just and true only as they are relative to some certain use or end, their aptitude and subordination to which end is[23].

This passage establishes the central features of Berkeley's theory. Beauty is neither intrinsic nor merely formal, but essentially relational and teleological. Proportion is not simply a matter of symmetry or balance; it is a relation of fitness to an end. The crucial terms – *use, end, aptitude, and subordination* – all point to a normative structure in which objects are evaluated according to how well they fulfill a purpose.

It is worth pausing here to note the historical-conceptual weight these terms carry. The category of "teleology" – the orientation of things toward ends – has its roots in Aristotle's natural philosophy, where it characterizes the internal

---

21 - *Alciphron* 3.5.

22 - *Of the Principles of Human Knowledge*, CIX.

23 - *Alciphron* 3.9.

directedness of natural substances toward their characteristic forms of flourishing. By the early modern period, however, teleological explanation had become contested: Descartes and his followers largely expelled final causes from natural philosophy, while thinkers such as Leibniz sought to rehabilitate them within a metaphysical framework. Berkeley's appeal to teleology thus represents a deliberate counter-movement within early modern thought, retrieving the Aristotelian insight that things are intelligible in light of their ends, while grounding those ends in divine rather than immanent nature. Similarly, the term "subordination" carries a specific conceptual history: in scholastic and early modern discussions of order, subordination typically denotes the hierarchical dependence of lower on higher principles – instrumental causes subordinated to final causes, parts to wholes. Berkeley inherits this hierarchical vocabulary but deploys it aesthetically, making subordination the very form in which beauty is apprehended. The term "system," finally, reflects the influence of seventeenth – and eighteenth-century natural philosophy and theology, where it denoted an interconnected whole governed by law – as in Newton's "system of the world." For Berkeley, the aesthetic and theological dimensions of this term converge: the world is a system in the scientific sense – ordered, lawlike, intelligible – and its beauty consists precisely in the perceptibility of that order to finite minds.

This account immediately rules out a purely subjectivist understanding of beauty. Berkeley's position can be further clarified by situating it within the broader eighteenth-century debate concerning the relation between beauty and utility. As Robert Clewis notes, this period is marked by competing models: some treat beauty and utility as independent, while others regard beauty as a "form or mode of perfection or goodness"[24]. Berkeley's view aligns most closely with this latter tradition. For Berkeley, beauty is not identical with utility in a crude instrumental sense, but neither is it independent of it. Rather, beauty arises from the fittingness of parts to ends – from what he calls their "aptitude and subordination." In this respect, his view anticipates later accounts of "purpose-based" or "adherent" beauty, in which aesthetic evaluation depends on an object's relation to a concept of its function or perfection. Yet Berkeley's account remains more deeply theological: the system of ends relative to which beauty is defined is ultimately grounded in divine reason. If beauty consists in relations of fitness to ends, then it cannot be reduced to the feelings of observers. At the same time, Berkeley's view is not a crude objectivism. Beauty is not a fixed property independent of context, since the

24 - R. R. Clewis, "Beauty and Utility in Kant's Aesthetics: The Origins of Adherent Beauty," in «Journal of the History of Philosophy», 56 (2018) n. 2, pp. 305–335, here p. 306.

same object may be beautiful in one setting and not in another. What matters is the relation between the object and a system of ends.

Consider, for instance, a tree. In a forest, the tree contributes to a harmonious environment: it provides shelter for birds, participates in ecological cycles, and stands in proportion to surrounding elements. Yet the same tree, if placed in the middle of a lecture room, would appear out of place. Its beauty depends not on its intrinsic features alone, but on its relation to an appropriate context. This example illustrates a crucial point: beauty is inherently systemic. It arises from the integration of parts within a whole. The notion of proportion presupposes a broader framework in which objects can be evaluated relative to ends. Beauty, therefore, is not a local property, but a feature of relational structure.

The teleological character of this structure raises a further question: what determines the system of ends? Berkeley's answer, I suggest, is theological. The order of nature is not self-explanatory; it reflects the rational activity of a divine mind. Proportion, in this sense, is an expression of divine order, and beauty is the way in which this order becomes intelligible to human cognition.

5. Subordination, Imperfection, and Divine Harmony

We are now in a position to state Berkeley's central claim in its most general form: beauty consists in relations of subordination within a system of ends. Parts are subordinated to wholes; wholes are ordered toward ends. This structure provides the standard for evaluating both natural objects and moral actions. A thing is beautiful when it fits – when it is appropriately related to its function and to the system of which it is a part. This explains Berkeley's emphasis on fittingness and use, as well as his claim that beauty is inherently relational. It also explains why beauty depends on thought: without the recognition of ends, there can be no basis for evaluating proportion.

Yet Berkeley's account goes further. It also addresses a problem that arises for any teleological theory of beauty: the problem of imperfection. If beauty consists in order, how are we to account for apparent disorder?

Berkeley's answer is: "blemishes and defects of nature are the not without their use," because "they make an agreeable sort of variety, and augment the beauty of the rest of the creation, as shades in a picture serve to set off the and more enlightened parts"[25]. Imperfections are not exceptions to order, but components of it. They contribute to the overall harmony of the system, much as contrast contributes to the beauty of a painting. This idea is developed more fully in *Siris:*

25 - Of the Principles of Human Knowledge, Chapter CLII.

> As for the blots and defects which appear in the course of this world, which some have thought to proceed from a fatality or necessity in nature, and others from an evil principle, that same philosopher observes, that it may be the governing reason produceth and ordaineth all those things; and, not intending that all parts should be equally good, maketh some worse than others by design, as all parts in an animal are not eyes: and in a city, comedy, or picture, all ranks, characters, and colours, are not equal or alike; even so excesses, defects, and contrary qualities, conspire to the beauty and harmony of the world[26].

Here Berkeley explicitly adopts a holistic perspective. Beauty cannot be assessed at the level of individual parts; it must be understood in relation to the whole. What appears defective locally may be necessary globally. This view has clear theological implications. It suggests that apparent disorder is compatible with, and even required by, a divinely ordered system. Human judgments of beauty are therefore limited by perspective. What seems imperfect to us may be perfectly ordered within the divine plan.

Beauty, in this sense, reveals not only the structure of the world, but also the limits of our understanding of it. Berkeley's texts suggest that beauty is not merely associated with order, but grounded in a specific kind of ordering – one that can be understood in terms of subordination, fittingness, and thought.

At the most general level, Berkeley's account of beauty is governed by a principle of functional coordination. As Euphranor explains:

> The parts, therefore, in true proportions must be so related, and adjusted to one another, as that they may best conspire to the use and operation of the whole[27].

This passage makes explicit what earlier discussions implied: beauty consists in the coordination of parts relative to a whole. The notion of "conspiring" to a use emphasizes that the relation is not merely structural, but teleological. Parts are not simply arranged; they are arranged *for the sake of an operation.* This allows us to distinguish two dimensions of aesthetic structure.

First, there is a vertical relation between an object and its end. As Berkeley states:

> A thing [is] perfect in its kind when it answers the end for which it was made[28].

An object's beauty depends on how well it fulfills its purpose. This relation establishes a standard of evaluation: the end determines what counts as proper proportion.

---

26 - *Siris* 262.

27 - *Alciphron* 3.8.

28 - *Alciphron* 3.8.

Second, there is a horizontal relation among parts within a system. The parts must be "so related, and adjusted to one another" as to contribute to the functioning of the whole. This relation concerns coordination among elements that share a common end.

These two dimensions are inseparable. The suitability of parts to each other presupposes a shared orientation toward an end, and the realization of an end requires the coordination of parts. Together, they form a unified structure in which subordination is the organizing principle: parts are subordinated to wholes, and wholes are ordered toward ends. This framework allows us to understand why beauty is inherently relational. It is not located in isolated features, but in the way those features are integrated within a hierarchy of purposes. Beauty, in this sense, is the perception of ordered subordination.

Berkeley explicitly characterizes beauty in terms of fittingness: "the subordinate relative nature of beauty"[29]. This phrase captures the essence of his view. Beauty is both relative – depending on relations among parts and ends – and subordinate – structured by hierarchical ordering.

The notion of fittingness can be clarified through Berkeley's own examples. To recognize the beauty of an object, we must understand its use: recognizing beauty requires "knowing its use" and "comparing its figure with that use"[30]. This comparison is essential. An object's form is evaluated in light of its function. A chair, for instance, is beautiful not simply because of its shape, but because its shape is appropriate to its use. A door properly installed in a house is beautiful; the same door lying unused is not. The difference lies in its relation to an end. This example reveals that beauty is not a perceptual given, but a relational judgment. It requires the recognition of how an object fits within a system of purposes. Without such recognition, the object's beauty cannot be fully apprehended.

At this point, Berkeley introduces what may be the most philosophically significant element of his account: the role of thought. As he writes:

> without thought there can be no end or design, and without an end there can be no use, and without use there is no aptitude or fitness of proportion, from whence beauty springs[31].

This passage establishes a hierarchy of dependence: beauty depends on proportion; proportion depends on use; use depends on ends; ends depend on thought. The conclusion is that beauty ultimately depends on thought. Without

29 - Alciphron 3.9.

30 - Alciphron 3.8.

31 - Alciphron 3.10.

thought, there would be no ends, and therefore no basis for evaluating proportion or fitness. This emphasis on thought is consistent with Berkeley's broader philosophical commitments. As Stephen Clark (2005) notes, Berkeley repeatedly insists that "intellect and reason are alone the sure guides to truth"[32], even while recognizing the limits of demonstrative knowledge. The same structure appears in his aesthetics. Beauty is not reducible to sensation, but requires the activity of a rational faculty capable of discerning relations of order and purpose. At the same time, this rationalism is not purely secular. The "ends" on which beauty depends cannot be generated by finite minds alone; they presuppose a system of purposes grounded in divine thought. Berkeley's account of beauty thus mirrors his account of knowledge more generally: both involve the apprehension of an order that is not constructed by the human mind, but discovered within a world sustained by a divine intellect.

This claim has two important implications. First, it confirms that beauty is not a purely sensory phenomenon. While perception provides the materials of aesthetic experience, the recognition of beauty requires an intellectual grasp of relations. Beauty is, in this sense, an object of the mind. Second, it points toward a theological foundation. If ends depend on thought, and the system of ends that structures the world is not produced by human minds, then it must ultimately depend on a divine mind. The teleological order that grounds beauty is, therefore, an expression of divine thought. This move transforms Berkeley's aesthetics into a form of theological metaphysics. Beauty is not merely a feature of the world as perceived, but a manifestation of the rational structure imposed by God.

The emphasis on fittingness and thought allows Berkeley to avoid two opposing errors. On the one hand, he rejects formalism, the view that beauty consists in abstract properties such as symmetry or regularity. Such properties are insufficient because they do not take into account the relation between form and function. A perfectly symmetrical object may fail to be beautiful if it is not suited to its purpose. On the other hand, he rejects subjectivism, the view that beauty consists in the pleasure experienced by the observer. Pleasure may accompany the perception of beauty, but it does not constitute it. Without a standard grounded in relations of fittingness, there would be no way to distinguish between true and apparent beauty. By grounding beauty in thought and teleology, Berkeley provides an account that is both objective and intelligible. Beauty is not a matter of arbitrary preference, but of rational judgment.

32 - S.R.L. Clark, *Berkeley on Religion,* in K.P. Winkler (ed.), *The Cambridge Companion to Berkeley*, Cambridge University Press, Cambridge 2005, pp. 369-404, here p. 374.

We can now see that the concept of subordination is not merely one element of Berkeley's account, but its unifying principle. This hierarchical structure is what gives beauty its intelligibility. To perceive beauty is to recognize this structure – to see how something fits within a larger system. From a theological perspective, this structure reflects the organization of creation itself. The world is not a collection of independent entities, but a system in which each part is ordered toward the good of the whole. Beauty is the mode under which this order becomes perceptible.

Berkeley's account of beauty is thus neither purely aesthetic nor merely descriptive. It is a theory of ordered relations grounded in thought and structured by subordination. Beauty arises from the fittingness of parts to ends and to wholes, and this fittingness ultimately depends on a system of purposes grounded in divine reason. By making thought essential to beauty, Berkeley transforms aesthetic judgment into a form of rational insight. To judge something beautiful is to understand its place within a hierarchy of ends. In this way, beauty becomes a bridge between perception and metaphysics, linking the experience of the world to its underlying rational structure.

The next section situates this account within a broader philosophical tradition by examining Berkeley's relation to ancient theories of beauty and his engagement with their methodological assumptions.

6. Berkeley, the Ancients, and the Ascent to Beauty

Berkeley's relation to the ancient tradition is further illuminated when placed within the emerging eighteenth-century discourse of the aesthetic. As Szécsényi emphasizes, early modern aesthetic experience was characterized by a renewed attempt to connect "the sensible and the transcendent," such that sensory experience becomes "the indispensable and constitutive framework" for apprehending higher realities[33]. Berkeley's account can be seen as a particularly sophisticated development of this tendency.

For while he begins from sensory experience – natural scenes, ordinary objects, and everyday practices – he treats these not as ends in themselves, but as entry points into a deeper intelligible order. The ascent from sensible beauty to intelligible beauty is thus not a departure from experience, but its fulfillment. In this respect, Berkeley both inherits and transforms the Platonic tradition: beauty remains a guide to intelligibility, but its object is no longer a transcendent Form, but the structured order of a divinely created world.

33 - E. Szécsényi, "The Aesthetics of the Invisible," cit., p. 732.

The account of beauty developed thus far – centered on proportion, order, subordination, and thought – places Berkeley within a broader philosophical tradition that extends back to ancient Greek philosophy. While Berkeley does not explicitly present himself as a systematic inheritor of this tradition, *Alciphron* contains clear indications that his view is both informed by and responsive to it. By examining these connections, we can better understand both the originality and the limits of Berkeley's position.

In ancient theories of beauty, Aristotelian τὸ καλόν is ἐπαινετόν (laudable), and Platonic τὸ καλόν is ἡδύ (pleasant) or ὠφέλιμον (profitable), and that beauty is perceived by "a reasonable mind" in relation to "true interest"[34]. This discussion situates Berkeley within a long-standing debate concerning the nature of beauty. In Aristotle, the beautiful is closely tied to what is worthy of praise, and thus to ethical evaluation. In Plato, it is associated with pleasure and benefit, but also with a higher intelligible order that transcends immediate perception. Both traditions agree, however, that beauty is not merely sensory, but involves a relation to reason and to value. Berkeley's position both appropriates and transforms these views. Like the ancients, he rejects the idea that beauty is purely subjective. Beauty is not simply what pleases, but what is worthy of rational approval. Yet he also recognizes a limitation in the ancient accounts. As Euphranor notes, the criteria of pleasure, profit, or praise are insufficient because they can conflict: different goods may be "canceled out" by greater goods, and thus cannot provide a stable foundation for judgment.

This problem leads Berkeley to a more robust conception of beauty, grounded not in isolated criteria but in the systemic relation of parts to ends. What ultimately matters is not whether something is pleasant or useful in a local sense, but whether it fits within the overall structure of a divinely ordered world. In this respect, Berkeley radicalizes the ancient insight that beauty is intelligible, by embedding it within a comprehensive teleological framework.

Berkeley's engagement with the ancients is not limited to substantive claims about beauty; it also extends to method. Euphranor explicitly expresses a preference for proceeding from particular cases rather than abstract generalities: he prefers particulars to generalities in discussing philosophical matters[35]. This methodological commitment is reflected throughout Berkeley's discussions of beauty. Rather than beginning with a definition, he invites us to consider concrete examples: landscapes, natural scenes, objects in use. These examples are not illustrative but constitutive of the argument. It is through them that we come to recognize the relations of

34 - *Alciphron* 3.4.

35 - *Alciphron* 3.8.

proportion and order that define beauty. This approach parallels Aristotle's method in the *Nicomachean Ethics,* where ethical concepts are clarified through reflection on particular cases, and Plato's use of examples in dialogues such as the *Symposium.* Berkeley thus situates himself within a tradition that treats philosophical understanding as emerging from engagement with concrete phenomena.

At the same time, Berkeley departs from the ancients by emphasizing the systemic integration of these particulars. The examples he offers – trees in forests, elements in nature, parts of a whole – are not isolated instances, but components of a larger structure. The movement from particular to general is therefore not merely inductive, but systemic: it reveals an underlying order that connects diverse phenomena.

Berkeley's engagement with Platonic themes is most evident in his reference to a "scale or ladder" by which we ascend to higher forms of beauty: a "scale or ladder" by which we "ascend to moral and intelligible beauty"[36] beginning from sensory experiences such as "the smell of a rose, or the taste of an apple"[37]. This passage unmistakably echoes Plato's *Symposium*, where the ascent from physical beauty to the contemplation of the Form of Beauty is described as a progression from particular instances to universal intelligibility. Berkeley appropriates this structure but reinterprets it within his own framework.

For Berkeley, the ascent is not from illusion to reality, but from partial to more complete apprehension of order. Sensory experiences provide the starting point, but they do not exhaust the content of beauty. Through reflection, we come to recognize the relations that underlie these experiences – relations of proportion, fitness, and subordination. The endpoint of this ascent is not an abstract Form, but the intelligible structure of the world as a whole. Beauty, at its highest level, is the perception of this structure. It is the recognition that the diversity of phenomena is unified by a system of relations grounded in divine reason.

By integrating these elements, Berkeley both continues and transforms the ancient tradition. Like Plato, he treats beauty as a guide to intelligibility. Like Aristotle, he emphasizes the role of rational judgment in aesthetic evaluation. But he departs from both by situating beauty within a theologically grounded system of teleological relations.

This transformation has two key features. First, Berkeley rejects the idea that beauty can be understood independently of function. While ancient accounts often emphasize proportion or harmony in abstract terms, Berkeley insists that these

36 - Alciphron 3.8.

37 - Alciphron 3.8.

must be understood relative to ends. Beauty is not merely formal order, but functional order. Second, Berkeley embeds this functional order within a comprehensive system governed by divine reason. The hierarchy of ends that structures beauty is not self-grounding; it reflects the purposes of a divine mind. In this respect, Berkeley's account is more explicitly theological than its ancient predecessors.

Berkeley's engagement with the ancient tradition reveals both continuity and innovation. By drawing on Platonic and Aristotelian insights, he situates his account of beauty within a long-standing philosophical framework. Yet by emphasizing teleology, subordination, and divine order, he transforms this framework into a distinctly early modern – and deeply theological – theory. Beauty, on this view, is not merely an object of perception or contemplation. It is a pathway to understanding the structure of reality itself. Through the ascent from sensory particulars to intelligible relations, we come to apprehend a world ordered by divine reason.

7. Conclusion

Berkeley's remarks on beauty, read in isolation, can appear occasional, rhetorical, or subordinate to larger concerns. Read together, however, they disclose a philosophically coherent orientation, one that gains in systematic appearance when read across the corpus as a whole. Beauty is not, for Berkeley, a peripheral topic appended to his metaphysics or moral theory. It is a privileged site at which those theories become visible. The central claim of this paper has been that Berkeley understands beauty as the intelligible manifestation of divine order: beauty is the mode under which finite minds apprehend the rational, teleological, and hierarchical structure of creation.

This conclusion follows from the convergence of several lines of argument. First, Berkeley's distinction between proper beauty and moral beauty does not divide two fundamentally different kinds of phenomenon. It identifies two domains in which the same basic structure is realized: in both cases, beauty consists in proportion, order, and fittingness within a system. Berkeley's aesthetics is therefore unified at a deeper level than the distinction between natural and moral beauty might initially suggest.

Second, this unity depends on Berkeley's qualified rejection of the moral sense theory. Alciphron's appeal to an interior sense, to "first thoughts and natural notions," to disinterestedness, and to a certain je ne sais quoi threatens to collapse beauty into mere feeling. Berkeley's resistance to this view reflects a more basic commitment: beauty must be capable of withstanding rational scrutiny. Just as truth survives the "strictest survey", so too does beauty. What this means,

importantly, is that Berkeley's disagreement with the moral sense tradition is not primarily one about faculty psychology – about which inner organ perceives beauty – but about the very nature of the aesthetic object. For Berkeley, the object of aesthetic judgment is not a quale of pleasurable feeling but a structure of relations, and it is that structure which makes the distinction between true and apparent beauty possible.

Third, the objective content of beauty is teleological. Berkeley's most explicit statements insist that there is no beauty without proportion, and that true proportion must always be assessed relative to some use or end. Beauty is therefore not a merely formal matter of symmetry or balance abstracted from function. Nor is it a fixed intrinsic property inhering in individual things regardless of context. It is relational through and through. A thing is beautiful insofar as it answers the end for which it was made and insofar as its parts are adjusted so as to conspire to the use and operation of the whole. This is why Berkeley's examples are so philosophically important. A tree in a forest, a door in a house, the elements within the visible world, and the various ranks and characters within a larger totality all illustrate the same point: beauty is not simply in the object, but in the object's fittingness within an ordered arrangement.

This brings us to the fourth and deepest feature of Berkeley's account: beauty is structured by subordination. The language of subordination is not ornamental in Berkeley; it names the basic form of intelligible order. Parts are subordinated to wholes; wholes are ordered toward ends; ends themselves depend upon thought. Once this sequence is made explicit, the metaphysical depth of Berkeley's aesthetics comes into view. Beauty presupposes proportion, proportion presupposes use, use presupposes end, and end presupposes mind. The upshot is that beauty is ultimately rooted in thought. Since the relevant system of ends is not constituted by human thought, it must finally depend upon divine thought. Beauty is thus not merely perceived by minds; it is grounded in the prior rationality of a world ordered by God.

From this perspective, Berkeley's descriptions of natural beauty take on their full significance. The opening scenes of the *Three Dialogues and Alciphron* are not simply literary flourishes. They are philosophical demonstrations of a world whose beauty emerges from "mutual aid and support," from "connexion and correspondence of parts," and from the way things are suited both to their own ends and to the whole. Natural beauty thus reveals the world not as an aggregate of independent items, but as a coordinated system. Because this coordination is not directly given in sensation, beauty must be "an object, not of the eye, but of the mind." Aesthetic experience is therefore cognitive as well as affective. It is a way of grasping the order that underlies the sensible world.

The same structure reappears, with an important modification, in the moral domain. Here the system is no longer one merely contemplated, but one in which rational agents must participate. Moral beauty lies in the rule-governed coordination of actions within "the system of rational beings." Its opposite is chaos, violence, and the breakdown of order. Berkeley's rehabilitation of interest is crucial here. What Alciphron dismisses as incompatible with virtue Berkeley reinterprets through reason and judgment. To balance lower and higher pleasures, present losses and future gains, is not to debase virtue, but to understand the hierarchy of goods in which beautiful action consists. Moral beauty is therefore not a spontaneous feeling of approval, but rational participation in the teleological order of creation.

Berkeley's treatment of imperfection confirms the same point from another angle. If beauty consisted only in local smoothness, symmetry, or immediate pleasure, then blemishes, defects, and contrary qualities would count straightforwardly against it. But Berkeley insists that such features may contribute to the beauty of the whole, just as shades heighten the force of light in a picture. This holistic view reveals both the strength and the modesty of human aesthetic judgment. We can genuinely perceive order, but only partially. We can recognize beauty, but not from the standpoint of the whole. What appears disordered to us may have its place in a larger harmony whose full rationale is visible only to the divine mind. In this respect, Berkeley's aesthetics approaches a theodicy of beauty: local imperfection need not compromise universal order, and may indeed be required by it.

Finally, Berkeley's engagement with Plato and Aristotle helps clarify the broader philosophical significance of this view. Like the ancients, he treats beauty as something accessible to a reasonable mind and as bound up with value rather than mere sensation. Yet he transforms that inheritance by giving beauty an explicitly teleological and theological foundation. The ascent from sensible particulars to intelligible beauty is, for Berkeley, not an escape from the world, but a deepened apprehension of the world as a divinely ordered whole. Beauty becomes a path to metaphysical insight, because it discloses the structure of reality itself.

Berkeley's theory of beauty, then, deserves to be understood as an integral part of his philosophy rather than an incidental aside. It gathers together themes that often appear separately in Berkeley scholarship: divine reason, the intelligibility of nature, the normativity of moral order, the relational character of qualities, and the cognitive role of perception. Beauty is the point at which these themes converge. To perceive beauty is to perceive more than pleasing surfaces; it is to discern the fittingness of things within a hierarchy of ends. To judge beauty truly is to recognize

that hierarchy as rational. And to recognize it as rational is already to glimpse the divine order that grounds it.

This account carries implications that extend beyond the interpretation of Berkeley himself. If beauty is conceived as the mode in which intelligible order becomes perceptible, then aesthetic experience is not a secondary or merely hedonic phenomenon – something that might be eliminated from a fully rational account of the world – but a genuine epistemic achievement. It reveals what argument and analysis can only approximate: the felt intelligibility of a structured whole. This may help explain why Berkeley so frequently begins his philosophical dialogues with aesthetic openings rather than definitions. The experience of beauty, for Berkeley, is not an ornament to philosophical inquiry but its natural threshold – the point at which finite minds first encounter, however partially, the order they will later seek to understand.

If this reading is right, then Berkeley offers more than an eighteenth-century contribution to aesthetic theory. He offers a philosophical account of how order becomes perceptible to finite minds. Beauty is not an embellishment added to an otherwise intelligible world. It is the sensible-intelligible form in which the world's intelligibility appears. In that sense, Berkeley's aesthetics is inseparable from his theology: beauty is the appearance of divine order to creatures whose understanding is finite, yet genuinely capable of truth.

AURORA YU
Assistant Professor of Philosophy at the University of San Diego
aurorayu@alumni.unc.edu

Divine Necessity and Human Freedom in Seventeenth-Century Metaphysics: Spinoza's *causa sui* and Mullā Sadrā's *wājib al-wujād*

di Jamil Akhtar

This article compares Spinoza's causa sui with Mullā Sadrā's wājib al-wujād to examine how divine necessity coexists with human freedom. Through a non-genealogical analysis, it distinguishes modal necessity from graded ontological dependence. The result is two distinct compatibilist models of agency, each grounded in a different metaphysical architecture of divine causality.

Keywords: Freedom, Metaphysical, Mullā Sadrā, Necessity, Spinoza

1. Introduction

The seventeenth century refines one of the core questions that early modernity raises: if God is necessarily existing and founding the intelligibility of everything that is, what understanding of human freedom remains coherent? This question is often posed in the European canon as a conflict between rational grounds and moral agency within a non-contingent order. Spinoza pushes the problem to its end by identifying God as *causa sui* (self-caused) and holds that whatever exists is "in God" and a derivative of the necessity of the divine nature[1]. On such a framework, contingency is not a metaphysical condition of reality but a correlate of insufficient cognition, and freedom will have to be redefined in the absence of an undetermined power of choice[2].

Yet a supernatural project of the seventeenth century – Mullā Sadrā's systematic formulation of God as *wājib al-wujād* (the Necessary Existent) – structures another design of dependence, which cannot be reduced either to voluntarist creationism or to a thesis of "God-or-Nature" identity. Sadrā's existence-based metaphysics (as commonly formulated through the priority of *wujād* and hierarchical dependence of finite realities) provides the means to answer how divine necessity can be maximal and also how finite agency is meaningful as a graded and dependent power and not an illusion[3].

Although chronologically contemporaneous and thematically overlapping, the early modern discussions of "God, necessity, and freedom" remain principally intra-European, and how particular models of necessity in the seventeenth century should be used to hone the conceptual resources one may bring to bear on interpreting Spinoza's compatibilism remains largely unexplored. The article fills that gap using a non-genealogical, controlled comparison: following a methodological clarification, it reconstructs (i) the account of freedom based on necessity proposed by Spinoza and (ii) the account of dependent agency based on necessity proposed by Sadrā and then compares the divergent explanatory logic of each and its implications for the early modern metaphysics of God.

---

1 - B. de Spinoza, *Ethics*, public domain text (Latin/English), Ethics I, D1; P11; P15-P18; P29. https://www.gutenberg.org/ebooks/3800

2 - M. Della Rocca, *Spinoza*, Routledge, London/New York 2008, pp. 75-76; 187-92; Y.Y. Melamed, *Spinoza's Metaphysics*, in M. Della Rocca (ed.), *The Oxford Handbook of Spinoza*, Oxford University Press, Oxford 2017, pp. xx-xx, here pp. 139-1ee52.

3 - S.H. Rizvi, *Mullā Sadrāδand Metaphysics: Modulation of Being*, Routledge, London 2009, pp. 15-9; S.H. Rizvi, *Mullā Sadrā*, in «The Stanford Encyclopedia of Philosophy», rev. 2021.

2. Methodological note

The comparison suggested here is correlational, not genealogical. It does not purport to draw historical lines of transmission between Spinoza and Mullā Sadrā; instead, it treats "divine necessity" as an identical problem-space in the metaphysics of the seventeenth century and inquires how two quite different conceptual systems respond to the same pressure: intervening an ultimately necessary divisional divine basis with an intelligible sense of human freedom4. To prevent anachronism and false equivalence, I apply two controls highly suggested in contextual and comparative methodology. First, I bracket explanatory narratives of "influence" and reconstruct each stance from its own argumentative environment – a methodology that resonates with the warning that introducing alien questions can corrupt the author's work5. Second, I do indicate a tertium comparationis on the level of functionality: what "necessity" does in either system (grounding, dependence, agency), rather than presuming that there is semantic identity between *causa sui* and *wājib al-wujād.*

Operationally, the paper will take the following steps: (i) text-anchored implementation of each account by each framework of necessity being; (ii) extraction of each framework's freedom concept as a reaction to necessity; (iii) synthesis through a systematic comparison (tables and figure) that brings out explicitly where the two architectures of dependence are convergent and divergent. This provides a disciplined "correlation of knowledge" whilst rendering the comparison philosophically fruitful and historically accountable.

3. Spinoza: divine necessity as *causa sui* and freedom as rational self-determination

Spinoza approaches the problem of God and of freedom by metaphysical necessity intelligibility: that which is intelligible must be explainable in itself and in terms of what it relies upon. In *Ethics* I, it acquires a strictly anti-contingent form. God is characterized as *causa sui* – "that whose essence involves existence"[6] – and as absolutely infinite substance[7]. It is not that God creates himself in time, but rather that the existence of God is rooted in (identical with) the very nature that the definition describes. This is why, according to numerous commentators, Spinoza's

4 - B. de Spinoza, *Ethics*, cit., E I D1, D7; I P11, P29.

5 - Q. Skinner, *Meaning and Understanding in the History of Ideas*, in «History and Theory», 8 (1969), pp. 3-53, here pp. 23-5.

6 - B. de Spinoza, *Ethics*, cit., E I D1.

7 - *Ibid.* E I D6.

causa sui is not a story of individual activity but rather a thesis concerning the priority of explanation and ontological autonomy[8].

Based on this, Spinoza asserts that God is necessarily existent[9] *and that nothing* other than God can exist or be imagined[10]. Finite beings are therefore not substances but *modes* – ways in which the attributes of God are manifested[11]. The "necessitarianism" of Spinoza is a consequence: out of the necessity of the divine nature, there are, in infinitely many ways, infinitely many things[12], and nothing in nature is contingent[13]. The seeming contingency is insoluble; however, we know nothing of the overarching causality and its sensible dependence[14]*;* Divine necessity in this sense is not the assertion that God exists but the assertion that the order of nature is the effect of the existence of God.

This metaphysical image appears at first to eradicate freedom. Instead, Spinoza aims to reformulate freedom so that it becomes consistent with universal necessity. A thing is "free," he says, because it "exists solely from the necessity of its own nature and is determined to act by itself alone"[15]. No finite thing could be free in the sense of being not dependent on external causes since only God is absolutely self-determined. They may be free in another, to be more exact, ethical and intellectual sense: to be free is to be *active* and not *passive* – to act out of proper knowledge of oneself and of one's causes, and not to be forced by external influences poorly understood. That is why Spinoza sees human bondage as a state of thought-action: while we are controlled by passions, which are the results of insufficient ideas, we are determined; when our ideas become sufficient, we are the sufficient causes of our deeds[16].

The dynamics of this liberty are Spinoza's psychology of conatus: everything struggles to continue its existence[17], and the human agency is a systematic

---

8 - E. Curley, *Behind the Geometrical Method: A Reading of Spinoza's Ethics*, Princeton University Press, Princeton 1988, pp. 51-86; B. Lord, *Spinoza's Ethics: An Edinburgh Philosophical Guide*, Edinburgh University Press, Edinburgh 2010, p. 49-82.

9 - B. de Spinoza, *Ethics*, cit., E I P11.

10 - *Ibid.* E I P14.

11 - *Ibid.* E I D5; I P15.

12 - *Ibid.* E I P16.

13 - *Ibid.* E I P29.

14 - *Ibid.* E I P33; B.D. Dutton, *Benedict De Spinoza (1632-1677)*, in «Internet Encyclopedia of Philosophy», University of Tennessee at Martin, accessed 18 February 2026; S. Nadler, *Baruch Spinoza*, in «The Stanford Encyclopedia of Philosophy», edd. E.N. Zalta - U. Nodelman, Fall 2025 Edition..

15 - *Ibid.* E I D7.

16 - *Ibid.* E III D1-D3; E III P1-P3; E IV Preface.

17 - *Ibid.* E III P6-P7.

manifestation of that struggle with thought and extension as its qualities. The increase of rationality does not change whether we are in the causal order or not (we always are), but it does change the way in which the order appears to us. Sufficient ideas make us more powerful to act; insufficient ideas make us more vulnerable to outer determination. Therefore, the distinctive Spinozist postulate is that true autonomy is rational self-determination in necessity – the accomplishment of understanding, not a metaphysical capacity to disrupt the nexus of causality[18].

Ethics V completes this account by connecting freedom with a stable form of self-governance based on reason's grasp of necessity. The effectiveness of the mind over the affects rests on the establishment of pleasant ideas and the arranging of affects by the understanding[19]. The ultimate manifestation of freedom is not arbitrariness but the "intellectual love of God," a rational joy following adequate cognition of God/Nature in the mind and hence in the necessity of things[20]. In modern terminology, Spinoza presents a comprehensive compatibilism: determinism does not oppose freedom since freedom is not the ability to act in one or two more different ways; freedom is the ability to act by one's own nature under the supervision of sufficient cognition[21].

This reconstruction is critical to the comparative purpose of this article. Spinoza offers a maximalist version of divine necessity where dependence is manifested through the modal framework of substance-attribute-mode. Human freedom is thus redefined as rational self-determination, a progressive accomplishment of agency that becomes more sufficient with higher cognition that is adequate to the desired order. The following section examines the conceptual approach to this redefinition by juxtaposing it with one of the contemporaneous accounts of necessary being in the seventeenth century – Mullā Sadrā's *wājib al-wujād* – whose structure of dependence exists in other metaphysical words.

18 - M.J. Kisner, *Spinoza on Human Freedom: Reason, Autonomy and the Good Life*, Cambridge University Press, Cambridge 2011, pp. 17-45; 46-56; 57-71.

19 - *Ibid.* E V P3-P10.

20 - *Ibid.* E V P32-P36.

21 - B. Lord, *Spinoza's Ethics*, cit., pp. 136-58; D. Garrett, *Spinoza's Ethical Theory*, in *Necessity and Nature in Spinoza's Philosophy*, Oxford University Press, Oxford 2018, pp. 312-15.

Table 1 - Spinoza (*Ethics*): primary-text anchors for divine necessity and rational freedom

Ethics locus (Curley tr.)	Key term(s)	Minimal textual claim	Function in this article's argument	Interpretive note (for precision)
I D1	*causa sui*	God's essence involves existence (self-grounding).	Fixes "divine necessity" as ontological/explanatory independence, not a contingent fact.	Not temporal self-production; it marks non-derivative being and ultimate explanation.
I P11	necessary existence	God necessarily exists.	Establishes necessity as the starting-point of the metaphysical system, not a conclusion from cosmological premises.	The proof relies on Spinoza's substance/essence framework; avoid importing a Thomist-style "real distinction" reading.
I P16	"follows from the necessity of the divine nature"	Infinitely many things follow from God's nature in infinitely many ways.	Grounds Spinoza's necessitarian order: nature's intelligibility is an expression of divine nature.	"Following" indicates internal intelligibility (expression), not an external decree or purposive choice.
I P18	immanent cause	God is the immanent (not transient) cause of all things.	Specifies the dependence relation: finite beings depend on God without God standing outside nature as a separate efficient cause.	Immanence here is a causal-ontological thesis; it does not by itself settle debates about "pantheism" unless further argued.
I P29	no contingency in nature	Nothing in nature is contingent; all is determined to exist and act.	Provides the determinist backdrop that motivates a revisionary account of freedom.	Contingency becomes epistemic (a feature of inadequate knowledge), not metaphysical chance.

I D7	freedom	Free = determined to act solely from the necessity of one's nature.	Supplies Spinoza's compatibilist definition of freedom: autonomy as self-determination within necessity.	Freedom ≠ alternative possibilities; it is graded in finite beings via the active/passive distinction developed in Ethics III–V.

Table 1 summarizes the precise *Ethics loci* that are doing the principal philosophical job in §3 so that the prose can be interpretative. It also demonstrates how Spinoza presents his assertions about necessity[22] and freedom[23] as one argument structure and not two distinct subjects. The interpretative notes anticipate typical errors of category (e.g., temporal self-causal; freedom as indeterminism) and do not make any new assertions that would need more space.

4. Mullā Sadrā: *wājib al-wujād* and freedom under ontological dependence

If Spinoza rebuilds freedom on the basis of a modal metaphysics of substance, attributes, and modes, Mullā Sadrā (Sadr al-Dīn Shīrāzī, d. 1640) treats the same problem – how a necessary divine ground can be compatible with finite agency – through an ontology of dependence, expressed in terms of existence (*wujūd*). The central point is God as *wājib al-wujād* (the Necessary Existent): something necessary "in itself," without another. Although Sadrā adopts the w*ā*jib/mumkin (necessary/possible) distinction of the Avicennan tradition, his unique point of contribution is to radicalize the sense of what it means for non-divine beings to be a mere "possible": their contingency is not just an epistemic predicate but a structural reliance of their very existence on the Necessary[24].

One of the key Sadrian gestures (as it appears in modern scholarly work as the "primacy of existence") is that the explanatory burden lies on *existence*, as opposed to quiddities/essences (*māhiyyāt*). According to such a perspective, finite objects are not autonomous objects that then get existence; their reality is wholly bound to a given and received act of being. This gives a metaphysical grammar for divine

22 - B. de Spinoza, *Ethics*, cit., E I D1, P11, P16, P18, P29.

23 - *Ibid.*, E I D7.

24 - I. Kalin, *Knowledge in Later Islamic Philosophy: Mulla Sadra on Existence, Intellect, and Intuition*, Oxford University Press, Oxford 2010, pp. 45-55; 122-25; M. Kamal, *Mulla Sadra's Transcendent Philosophy*, Routledge, London 2006, pp. 24-42.

necessity: the God's necessity is not just one property among others, but it is the absolutely non-borrowed reality, upon which all other realities owe their ontological status[25].

This dependency is usually described as the univocation (*tashkīk*) that Sadrā describes as existence: existence is not lax in all beings in the same manner; rather, the existence comes in degrees or intensities. What now becomes related to the God-world relation is of philosophical importance to the current comparison: the dependence of finite beings may be extensive without identity collapsing. In other words, it is possible to hold to a maximal divine necessity and have some sort of intimate divine "presence" (in the sense of ontological dependence) and remain opposed to the equation "God = Nature" as the mere identity statement. The reality of the world is real-yet-derived: it is neither an independent world beyond the Necessary nor a mere subjective illusion[26].

Then, how will such ontological dependence survive the fact of "freedom"? Sadrā 's response to this question does not lie in the adding of pockets of metaphysical inconclusivity but in the re-conceptualization of agency as a gradual perfection of a dependent entity. Human action is not defined as an independent origination; it is a construed power that is exercised in its own ontological position and which can be enhanced through knowledge and self-transformation. That is, dependence does not eliminate agency; it identifies in which terms agency will be intelligible. This helps to understand why Sadrā's debates on moral psychology and epistemology often associate human well-being with the augmenting actuality – being more of what one is as an intellectual and spiritual being – rather than with an abstract power of opposite determination[27].

Another doctrine, frequently highlighted in the secondary literature, renders this compatibilist course particularly salient, that is, substantial motion (al-Saraka al-jawhariyya*).* Though the development of its complete statement would overstep the bounds of this article, its applicability can be concisely expressed thus: assuming that the soul (and more generally substances) is not given in a static manner in any case but is dynamically perfected, then insofar as human agency can be conceived as self-formation within an ordered metaphysical dependence even on the Necessary. Within this paradigm, freedom is not the absence of causality or

---

25 - I. Kalin, *Knowledge in Later Islamic Philosophy*, cit., pp. 65-68; M. Kamal, *Mulla Sadra's Transcendent Philosophy*, cit., pp. 24-42; K. El-Rouayheb - S. Schmidtke (edd.), *The Oxford Handbook of Islamic Philosophy*, Oxford University Press, Oxford 2016, pp. 465-488.

26 - S.H. Rizvi, *Mullā Sadrā*, cit.; I. Kalin, *Knowledge in Later Islamic Philosophy*, cit., pp. 305-314.

27 - M. Kamal, *Mulla Sadra's Transcendent Philosophy*, cit., pp. 64-87.

metaphysical autonomy but a self-actualization position, progressively more "one's own" as the cognitive and existential capacities of the agent become stronger[28].

This Sadrian model thus establishes a definite comparison with Spinoza. Both philosophers criticize freedom as a mindless contingency and treat freedom as a refined, transformed agency within necessity. Yet they locate the corresponding dependence relationship differently: Spinoza defines dependence by articulating the modal framework of a single substance (modes being derivatives of the divine nature), whereas Sadrā defines dependence as ontological derivation in transcendences of existence through the Necessary. The subsequent blurring uses this juxtaposition to describe two "architectures of dependence" and to explain how different metaphysical accounts of necessary being yield distinct reconstructions of human freedom.

Table 2 - Mullā Sadrā: core doctrinal elements used in the comparison (necessity dependence agency)

Doctrinal element	Technical term(s)	Minimal claim (as used here)	Role in grounding divine necessity	Direct relevance to human freedom (in this paper's sense)
Necessary Existent	*wājib al-wujād*	God is necessary "in itself," non-derivative, and the ultimate ontological ground.	Fixes the absolute pole of being: only God is necessary without dependence.	Rules out freedom as absolute independence; motivates a reconstructed, compatibilist freedom.
Dependent/ possible being	*mumkin al-wujād*	Finite beings are possible "in themselves" and exist through dependence on the Necessary.	Explains why created reality is not self-explanatory; being is received/ derivative.	Human agency can be real yet non-independent: actions are "one's own" within dependence.

28 - K. El-Rouayheb - S. Schmidtke, *Oxford Handbook of Islamic Philosophy*, cit., pp. 472-475.

Primacy/ explanatory priority of existence	(commonly) *alat al-wujūd*	Existence, not quiddity, carries explanatory weight in metaphysics.	Makes dependence ontological (about *being*), not merely mechanical or epistemic.	Freedom can be framed as increased actuality/power through knowledge (not as chance).
Gradation/ modulation of being	*tashkīk al-wujūd*	Existence is manifested in degrees/intensities rather than a flat univocal distribution.	Allows pervasive dependence without identity: creatures participate derivatively without collapsing into "God = world".	Agency becomes a graded capacity: more "free" as one's existential-cognitive power intensifies.
Essence/ quiddity and existence	*māhiyya* / *wujūd*	What a thing is does not itself secure that it is; existence is not self-contained in quiddity.	Blocks treating finite entities as ontologically autonomous "units" alongside God.	Undercuts libertarian "uncaused choice" models; supports freedom as perfected dependent agency.

This table includes only the Sadrian doctrines that are strictly required to compare "necessary being" and "freedom". It maintains the non-redundancy of the comparison by still connecting each doctrine to a given argumentative position (how necessity grounds dependence and how dependence allows agency). It also clarifies which component (substantial motion) is optional.

5. Comparative synthesis: two architectures of necessity, two reconstructions of freedom

Sections 3 and 4 reconstruct two seventeenth-century strategies for holding together: the maximal divine necessity and a non-illusory account of human agency. In the synthesis, it is now possible to put it more accurately: Spinoza and Mullā Sadrā meet on the denial of contingency as an essential metaphysical characteristic, yet they diverge from each other in the structure of reliance within which the necessity is expressed – and this deviation produces distinct reconstructions of "freedom".

In the case of Spinoza, necessity presents itself in the modal form of a single substance; God as *causa sui*[29] exists necessarily[30], and the whole order of nature is a result of the necessity of the divine nature[31]. Dependent is thus mainly modal: finite states are modes the intelligibility of which belongs to the immanent order of causation[32]. It is against the usual assumption that determinism is merely a denial of freedom to discover that the freedom is to be found in the opposition between passivity and activity: we are "free" insofar as we act under the necessity of our nature[33] – insofar as we are the adequate causes of our actions[34]. It is rational self-determination within necessity – under a graded agent dependence, that is, in agency via understanding that renders freedom[35].

Sadrā also confirms that the divine ground is necessary in itself (*wājib al-wujūd),* but he expresses reliance within an existence-oriented ontology. The non-divine realities are possible by themselves (*mumkin*) and even dependent on their own existence; dependence is therefore not only epistemic but also ontological. Moreover, gradational descriptions of being (*tashkīk*) make possible a pervasive dependency without needing to reduce to a univocated identity between God and the world: creatures are real as derivative, ranked manifestations of being and not independent entities along with the Necessary[36]. In this architecture, freedom is unrealistically understood as independence; rather, it can be intelligible as a refined working of dependent agency – self-formation, and self-governance in the measure of enhanced actuality and knowledge[37].

The key compensation is that compatibilism with a necessary God is not one-pointed. Spinoza's compatibilism lies in explanatory closure within a single substance and describes liberation as a rational adequacy. Sadrā's compatibilism is based on graded ontological dependence and models liberation as enhanced perfection of a dependent agent. These are not simply variant vocabularies for the

29 - B. de Spinoza, *Ethics*, cit., E I D1.

30 - *Ibid.*, E I P11.

31 - *Ibid.*, E I P16, P29.

32 - *Ibid*, E I P18.

33 - *Ibid*, E I D7.

34 - *Ibid*, E III D1-D3; E V.

35 - M.J. Kisner, *Spinoza on Human Freedom*, cit., pp. 17-45; 46-56; 57-71; O. Koistinen (ed.), *The Cambridge Companion to Spinoza's Ethics*, Cambridge University Press, Cambridge 2009, 167-87.

36 - I. Kalin, *Knowledge in Later Islamic Philosophy*, cit., pp. 223-9; 301-5; F. Rahman, *The Philosophy of Mullā Sadrā (Sadr al-Dīn al-Shīrāzī)*, SUNY Press, Albany 1975, pp. 27-40; 47-58.

37 - M. Kamal, *Mulla Sadra's Transcendent Philosophy*, cit., pp. 64-87; 88-105; J.W. Morris (trans.), *The Wisdom of the Throne: An Introduction to the Philosophy of Mullā Sadrā*, Princeton University Press, Princeton 2003, pp. 60-64.

same view: they represent variant dependence relations (modal expression vs. graded derivation) and consequently variant meanings in which an action may be seen as "one's own". This contrast becomes expressible (and graphically expressible in *Figure 1*) to make the conceptual map of early modern deliberations of God and freedom clearer: that the necessity of divine origin can give rise to greater than a single consistent reconstruction of human agency.

Table 3 - Comparative synthesis matrix

Dimension	Spinoza (modal architecture)	Mullā Sadrā (ontological architecture)	Point of convergence	Divergence / payoff for "freedom under necessity"
What "divine necessity" primarily means	God is *causa sui*, existence follows from essence[a]	*God is wājib al-wujūd; uniquely necessary "in itself"*	God is not contingent; divine necessity is foundational	Spinoza: necessity is built into one-substance explanatory closure; Sadrā: necessity grounds a hierarchy of dependent being
Structure of dependence	Finite things are modes of the one substance[b]	Finite things are possible/dependent beings (*mumkin*) receiving being from the Necessary	Dependence is not optional or merely moral; it is metaphysical	Spinoza: dependence is modal-expression within one substance; Sadrā: dependence is ontological derivation across degrees of being
Status of contingency	Nothing is contingent in nature[c]; contingency = epistemic (linked to inadequate cognition)	Contingency is real-as-dependent: beings are possible in themselves, dependent in existence	Both reject brute chance as ultimate	Different diagnoses of "could have been otherwise": Spinoza explains via ignorance; Sadrā via genuine creaturely possibility-with-dependence

Definition of freedom	Free = determined by necessity of one's nature alone[d]; becomes *active* via adequate ideas[e]	Freedom = meaningful agency as perfected operation within dependence (often tied to intensification/ perfection)	Freedom ≠ uncaused choice	Spinoza: freedom is rational self-determination (adequacy); Sadrā: freedom is graded perfection/ actualization within ontological hierarchy
Ethical "mechanism" of liberation	Liberation through understanding necessity; re-ordering affects via reason[f]	Liberation through ascent/ intensification of being/knowledge (perfectionist trajectory)	Both yield a transformative ethics grounded in metaphysics	Two different compatibilisms: "cognitive-therapeutic" (Spinoza) vs "ontological-perfectionist" (Sadrā)

Table 3 compresses the argument into five comparison-dimensions.

a - *B. de Spinoza, Ethics, cit., E I D1; I P11.*
b - *Ibid., E I D5; I P15.*
c - *Ibid., E I P29.*
d - *Ibid., E I D7.*
e - *Ibid., E III-V.*
f - *Ibid., E IV-V.*

These sets of nodes/edges capture precisely the synthesis-claim: Spinoza models dependence as modal expression in one substance only; Sadrā models dependence as graded ontological derivation from the Necessary. The compatibilist difference could be seen in the fact that the nodes of "Freedom" are the results of various mechanisms (adequacy vs. perfected dependence).

6. Conclusion

This article claimed that "divine necessity" in seventeenth-century metaphysics does not entail a standardized, constant outcome for human freedom. By matching the doctrine of God as *causa sui* with Mullā Sadrā's account of God as *wājib al-wujūd* – without any presumption of historical influence – the paper explains two different orders of dependence and the different reconstructions of agency they authorize. In Spinoza, necessity is expressed in the modal form of one substance

and its immanent inner order: freedom is redefined as self-determination of rationality, obtained as a result of progressing adequacy and activity within a necessary nexus of causation. In Sadrā, necessity can be described by ontological dependence, which must be expressed in terms of existence and gradation: freedom is understandable as a perfected dependent agency, a graded power of self-formation that does not need the metaphysical independence.

Such comparative payoff is conceptual and not ecumenical: it perfects what is meant by "compatibilism under a necessary God" by indicating that one dependence-relation (modal expression versus graded ontological derivation) yields different criteria for what counts as an action "one's own". More broadly, the outcome confuses linear narratives that correlate early modern necessity with the decline of agency. The extension of the framework to providence and theodicy in future work, or to the theological-political applications of necessity to the issue of authority and secularization in early modern thought, is possible.

JAMIL AKHTAR
Assistant Professor of Islamic Studies in the Faculty of Islamic Studies at the University of the Punjab, Lahore, Pakistan
jamillahori@gmail.com

Oltre la dimostrazione. Il limite della ragione e l'argomento ontologico in Kant

di Francesco Porchia

This paper examines Kant's critique of the ontological argument, arguing that the limit of reason does not lead to the dissolution of metaphysics, but rather opens the possibility for its reconsideration. The impossibility of logically demonstrating the existence of God points toward an excess that cannot be reduced to conceptual determination. From this perspective, the dialogue between Kant and the tradition of the ontological argument allows for a rethinking of both the subject and the meaning of speculative thought.

Keywords: Kant, Ontological Argument, Limit of Reason, Metaphysics.

Introduzione

Il tema teologico ricompreso all'interno del pensiero kantiano costituisce una sfida, se non proprio la sfida per eccellenza, che la tradizione filosofica occidentale ha ereditato e sulla quale si è dovuta confrontare. Da un lato, la scissione tra fenomeno e noumeno è stata interpretata da una certa parte della tradizione filosofica come pretesto per negare il problema Dio, presentando così inevitabilmente una certa forma di nichilismo, un mondo che era stato scisso da Dio[1]. Parallelo è stato invece il tentativo di coloro che hanno cercato di confrontarsi con la scissione, quella *Zerrissenheit* tra ideale e reale, tra finito e infinito, che ha caratterizzato il filone dei post-kantiani e in particolare l'itinerario tracciato da Reinhold, dal romanticismo e dall'idealismo tedesco[2]. Quest'ultimo filone si è posto in stretta continuità col pensiero kantiano, tanto che i differenti sentieri tracciati dagli esponenti dell'idealismo tedesco, trovano la loro unità proprio nel pensiero del filosofo di Königsberg nel tentativo di rendere chiaro il suo pensiero e di ricercare, proprio dalle tematiche kantiane, un punto unitario, che potesse superare quella scissione apparente tra fenomeno e noumeno[3]. Tornare a indagare

1 - Cfr. M.M. Malimpensa, *La scienza inquieta. Sistema e nichilismo nella Wissenschaftslehre di Fichte*, InSchibboleth, Roma 2020, p. 113; F. Vercellone, *Introduzione a il nichilismo*, Laterza, Roma-Bari 1992, pp. 3-30; F. Volpi, *Il nichilismo*, Laterza, Roma-Bari 1996, pp. 11-19.

2 - Per un itinerario che evidenzia un particolare sviluppo a partire dal piano della ragione pratica si rimanda a: M. Ivaldo, *Ragione pratica. Kant, Reinhold, Fichte*, Edizioni ETS, Pisa 2012. Per uno studio focalizzato sul romanticismo si veda: M. Frank, *"Unendliche Annäherung". Die Anfänge der philosophischen Frühromantik*, Suhrkamp, Frankfurt am Main 1977.

3 - Per l'importanza relativa all'eredità kantiana, si rimanda alle stesse parole di Fichte il quale indica, all'inizio della propria carriera filosofica, la necessità di rendere noto ciò che Kant aveva già detto nella propria filosofia: «L'autore sa di non poter mai dire qualcosa cui Kant, in modo diretto o indiretto, chiaro od oscuro, non abbia già accennato», in J.G. Fichte, *Gesamtausgabe der Bayerischen Akademie der Wissenschaften,* a cura della Fichte-Kommission der Bayerischen Akademie der Wissenschaften, Frommann-Holzboog, Stuttgart-Bad Cannstatt 1962-2012: I,2, p. 110; trad. it., A. Tilgher, *Dottrina della scienza*, Laterza, Roma-Bari 1971, p. 6 (D'ora in poi citato come "GA"). Si pensi all'importanza cruciale che Kant ha poi avuto per gli studenti dello *Stift* di Tubinga (cfr., E. Mirri, *Gli anni della formazione Stoccarda e Tubinga. 1770-1793*, in G.W.F. Hegel, *Scritti giovanili*, a cura di E. Mirri, Orthotes, Napoli-Salerno 2015 pp. 21-41. Qui p. 34; W.G. Jacobs, *Zwischen Revolution und Orthodoxie? Schelling und seine Freunde im Stift und an der Universität Tübingen*, Frommann-Holzboog, Stuttgart-Bad Cannstatt 1989). Hegel testimonia fortemente nei suoi *Scritti giovanili* di Berna, specie nella *Vita di Gesù*, in cui la preghiera riportata rimanda chiaramente ad un'adesione a Kant e ai principi dell'illuminismo (cfr., G.W.F. Hegel, *Gesammelte Werke*, a cura della Rheinisch-Westfälischen Akademie der Wissenschaften, Felix Meiner Verlag, Hamburg 1968-: 1, pp. 218-219; trad. it., *Scritti giovanili*, cit., p. 250-251: d'ora in poi citato come "GW"), una adesione forte al pensiero kantiano. Lo stesso Schelling lo testimonia sin dai primi studi col suo commentario al *Timeo* (F.W.J. Schelling, *Historisch-Kritische Ausgabe*, a cura della Schelling-Kommission der Bayerischen Akademie der Wissenschaften, Frommann-Holzboog, Stuttgart-Bad Cannstatt 1976-: II,2; trad. it. e curatela di L. Follesa,

il problema teologico nella filosofia kantiana significa fare i conti col tema del limite, con quello del rapporto tra libertà e necessità, cosa che è restituita chiaramente dalla posizione che Kant ha assunto verso l'argomento ontologico.

Ovviamente, il tema sopracitato costituisce un tema molto vasto e complesso all'interno della produzione kantiana, tanto da non poter essere ripercorso per intero in questo breve contributo. Lo scopo di questo scritto è pertanto limitato a indagare la riflessione kantiana condotta sull'argomento ontologico e come il relativo tema del limite, da esso scaturito, possa essere traccia di un'apertura ad un pensiero metafisico: l'intento è pertanto quello di mostrare come l'esito delle argomentazioni kantiane non conduca ad una mera rinuncia del filosofare o alla distruzione della metafisica. Ciò che si intende mettere in luce è come, al contrario, essa conduca a quella medesima tensione che è stata tramandata alla tradizione filosofica successiva che ha fatto del problema intorno alla ricerca di Dio la domanda filosofica per eccellenza[4].

Sulle prove dell'esistenza di Dio in Kant.

Già nel periodo degli *Scritti precritici* Kant si era confrontato con la valenza assunta dal tentativo di provare l'esistenza di Dio. Cruciale risulta essere, sin qui, l'argomentazione secondo la quale l'esistenza non può essere intesa né come un

Sullo spirito della filosofia platonica. Il Timeo di Platone: commento manoscritto (1794), Mimesis, Milano-Udine 2022. D'ora in poi l'edizione critica verrà citata come "HKA"), in cui il confronto con Kant culmina nel tema dell'abisso della ragione in riferimento alla prova cosmologica della *Critica della ragion pura* (*Kants Gesammelte Schriften* a cura della Königlich Preußische Akademie der Wissenschaften, Reimer, Berlin 1904-1911: III-IV, A 613/B 642; trad. it. di G. Gentile e G. Lombardo-Radice, *Critica della ragion pura*, Laterza, Roma-Bari 2005, p. 390: d'ora in poi l'edizione critica sarà indicata con "AA"). Quel medesimo tema sarà cruciale proprio per l'impianto più maturo del filosofo di Leonberg, nella sua suddivisione tra filosofia negativa e positiva (si pensi in merito alle *Lezioni di Erlangen* e alle riflessioni sull'estasi della ragione: cfr., HKA II, p. 10; trad. it., L. Pareyson, *Conferenze di Erlangen*, in *Scritti sulla filosofia, la religione, la libertà*, a cura di L. Pareyson, Mursia editore, Milano 1974, pp. 197-226. Qui p. 212).

4 - Anche da questo punto di vista, il tema squisitamente teologico verrà ereditato dell'idealismo tedesco come nucleo centrale di ricerca. Si pensi in tal senso alle affermazioni significative di Fichte: «Io esigo qualcosa che sia al di là della mera rappresentazione, che esista, sia esistito ed esisterà, anche se non ci fosse la rappresentazione», in; GA I,6, p. 253; trad. it., R. Cantoni, *La missione dell'uomo*, Laterza, Roma-Bari, 1944 p. 105. Come non rimandare alle stesse parole di apertura della *Enciclopedia* di Hegel: «*Dio*, e *Dio solo*, è la verità», in GW 20, p. 39; trad. it., B. Croce, *Enciclopedia delle scienze filosofiche in compendio*, Laterza, Roma-Bari 1983, p. 3; o, infine, alla stessa formulazione della *Grundfrage* di Schelling che ricerca proprio nell'evento storico della rivelazione la risposta alla «domanda della disperazione»: cfr., F.W.J. Schelling, *Sämmtliche Werke*, 14 vol. (II sez.), a cura di K.F.A. Schelling, J.G. Cotta, Stuttgart-Augsburg 1856-1861: vol. XIII, p. 7; trad. it., A. Bausola, *Filosofia della rivelazione*, Bompiani, Milano 2002, p. 11: l'edizione critica verrà d'ora in poi citata come "SW".

predicato né come la totalità dei predicati di una cosa, poiché non aggiunge alcuna determinazione al contenuto del concetto. Si rimanda alla prima proposizione della *Considerazione prima:*

> Pure noi ci serviamo dell'espressione di esistenza come di un predicato; e questo si può fare sicuramente e senza inquietanti errori, sempreché non ci sia pericolo di voler dedurre l'esistenza da concetti semplicemente possibili, come si suol fare, quando si vuol provare l'esistenza assolutamente necessaria[5].

L'esistenza non costituisce un predicato che noi possiamo dedurre da un concetto, facendo così dipendere in un certo qual modo la condizione di possibilità di tale oggetto dalla deduzione stessa dell'esistenza. L'esistenza non può essere intesa come la totalità dei predicati che costituiscono una cosa, ma come il porsi di qualcosa cui spettano quei predicati che ne costituiscono il concetto, senza che per questo venga aggiunta alcuna nuova determinazione al contenuto concettuale. La differenza è data dal modo in cui si intende il concetto di essere, ovvero se lo si considera una cosa come posta in relazione a qualcos'altro o se lo si consideri, in senso più marcato, in sé e per sé[6]. Da questo punto di vista, viene tracciata una differenza fondamentale tra ciò che costituisce la copula della proposizione, in funzione di un giudizio logico, e l'attestazione dell'esistenza di qualcosa che prescinde dalla dimostrazione logica e la cui esistenza non è espressa allora da una determinazione ulteriore, ma consiste nel porre come esistente qualcosa cui spettano quelle determinazioni:

> "Dio è onnipotente" è una proposizione che deve rimanere vera anche nel giudizio di colui che non ne riconosce l'esistenza, quand'egli mi intenda nel mio concepire Dio […]. Quando io dico: Dio è una cosa esistente, pare ch'io esprima la relazione di un predicato con un soggetto. Ma in verità c'è una inesattezza in questa espressione. A parlare esattamente, si dovrebbe dire: qualcosa di esistente è Dio; cioè: ad una cosa esistente spettano quei predicati che presi insieme noi contrassegniamo con la espressione Dio [7].

Quest'ultima affermazione conduce ad una riflessione cruciale: se effettivamente l'esistenza di Dio non può esser determinata logicamente mediante il giudizio del soggetto razionale, non dipendendo così dalle operazioni della ragione, ci si dovrebbe chiedere infatti se l'esistenza di Dio possa essere intesa quale mera somma dei predicati, o se, con tale espressione, non ci sia spazio per

---

5 - AA II, p. 72; trad it., *L'unico argomento possibile per una dimostrazione dell'esistenza di Dio*, in I. Kant, *Scritti precritici*, a cura di R. Assunto - R. Hohenemser, Laterza, Roma-Bari 1982, pp. 103-213. Qui p. 113.

6 - Cfr., AA II, p. 73; trad. it., *Ibid.*, p. 114.

7 - *Ibid.*, pp. 115-116.

una riflessione ulteriore: il fatto stesso che l'esistenza di Dio non sia concepibile come risultato di una dimostrazione, suggerisce che essa non possa essere definita, determinata. Ciò implica che l'esistenza non possa essere ricondotta a una somma definibile di determinazioni. In quanto non determinabile mediante il concetto, essa non è esauribile dal pensiero e può essere indicata solo come ciò che eccede ogni determinazione concettuale, senza tuttavia potersi costituire come una determinazione ulteriore[8]. È a partire da questa riflessione che si comprende il richiamo al tema dell'abisso, molto presente in Kant tanto che in apertura di questo scritto precritico egli afferma:

> Tuttavia dal ricercare tal dimostrazione non ci si può astenere, finché essa non si faccia, dovecchessia, presente. Poiché, senza accennare al legittimo desiderio, di cui non può spogliarsi un intelletto abituato all'indagine, di raggiungere qualcosa di compiuto e di chiaramente concepito in così importante conoscenza, è da sperare ancora che, una volta padroni di una cognizione cosiffatta, il suo oggetto possa esserne molto più chiarito. Ma a raggiungere questo scopo bisogna avventurarsi entro l'abisso senza fondo che è la metafisica. Oceano tenebroso, senza sponde e senza fari, in cui bisogna condursi come chi, navigando in mare non ancora solcato, non appena metta piede su una qualche terra, esamina il suo cammino, e cerca se mai delle inavvertite correnti marine non abbian deviato il suo corso, non ostante ogni precauzione che possa mai prescrivere l'arte di navigare. / Tuttavia tal dimostrazione, come altri han già notato, non è stata ancora trovata[9].

Il tema dell'abisso della ragione torna in modo significativo nella *Critica della ragion pura* nella trattazione dedicata alla prova Cosmologica: il disgregarsi di ciò che costituisce il fondamento, in quanto appare alla ragione come mera determinazione, come ente e non come Dio in sé e per sé, spalanca l'abisso. Con le stesse parole di Kant:

> La necessità incondizionata, di cui abbiamo bisogno in maniera così indispensabile, come dell'ultimo sostegno di tutte le cose, è il vero baratro della ragione uma-

8 - Si fa riferimento con questo al modo in cui è stata intesa la monade di Leibniz: la ragion sufficiente della monade non risponde al solo principio di causalità, ma presuppone un'eccedenza, ovvero una ulteriorità incalcolabile che rompe la linearità della causalità: cfr., K. Appel, *Tempo e Dio. Aperture contemporanee a partire da Hegel e Schelling*, Queriniana, Brescia 2018, p. 48. In questo senso, tale eccedenza non può essere intesa in senso costitutivo né come oggetto di conoscenza, ma come ciò che, secondo l'uso regolativo delle idee di ragione, orienta il pensiero verso una unità che resta indeterminabile; proprio per questo, essa indica una dimensione che, dal punto di vista ontologico, rimane problematica e sarà oggetto di successive elaborazioni. Sull'uso regolativo a partire da Kant, in merito alla comprensione di questa eccedenza: cfr., M. Trudu, *The Regulatory Use of the Ideas of Reason in Kant and Husserl*, in *Kant and the Space of Feeling*, a cura di F. Fabbianelli - A. Falduto, De Gruyter, Berlino 2026, pp.221-233.

9 - AA II, pp. 65-66; trad. it., *L'unico argomento possibile per una dimostrazione dell'esistenza di Dio*, cit., pp. 105-106.

> na [...]. Qui tutto si sprofonda sotto di noi, e la massima come la minima percezione pende nel vuoto senza sostegno innanzi alla ragione speculativa, alla quale non costa nulla far disparire l'una come l'altra senza il più piccolo impedimento[10].

È significativo che il tema dell'abisso ricompaia proprio nella *Critica del giudizio* in relazione al sublime, in quanto ciò che è abisso può essere considerato come quella eccedenza che svela ad un tempo la piccolezza dell'umano e delle proprie facoltà ma che, parimenti, ne suggerisce una partecipazione[11]. Tale eccedenza, in relazione al tema dell'abisso della ragione, pone nuovamente la necessità di indagare quelli che sono i limiti del soggetto dinnanzi alla possibilità del determinare, di dimostrare.

Ecco che allora la questione richiama direttamente la formulazione dell'argomento ontologico, nella sua differenza, non esaminata direttamente da Kant ma presente in base a ciò che è stato citato sopra, con ciò che vien definito prova. Diviene cioè necessario, a partire da quanto sostenuto sopra con Kant, fare un passo indietro sino alla radice stessa della formulazione dell'argomento ontologico e nella sua distinzione dalla successiva ricezione che lo ha ricondotto ad una prova.

Da argomento a prova. La differenza della posizione del soggetto nello stabilirsi dei limiti della ragione.

Kant definisce la prova ontologica come quella più rigorosa. Scrive:

> Tuttavia, se si domandasse quale dunque delle due prove sia la migliore in generale, si risponderebbe: quando si abbia riguardo alla esattezza logica ed alla completezza, migliore è la ontologica; ma se si desidera comprensibilità per il retto intendimento comune, vivacità dell'impressione, bellezza e forza motrice per i movimenti morali dell'umana natura, la preferenza è da concedersi alla prova cosmologica[12].

In questo passaggio Kant sta tracciando una differenziazione interessante tra ciò che è rigorosamente certo e ciò che fa maggiormente leva a quel che potremmo definire come "sentimento" degli altri. La ricerca della certezza costituisce proprio ciò che, a partire da Descartes, ha sempre più preso peso in filosofia, tanto che ci si è quasi dimenticati purtroppo che ciò che viene ricercato non è più il Vero,

---

10 - AA IV, p. 613 (A); AA III, p. 642 (B); trad. it., *Critica della ragion pura*, cit., p. 390.

11 - Cfr., AA V, p. 265; trad. it., A. Gargiulo e V. Verra, *Critica del giudizio*, Laterza, Roma-Bari, 1997, pp. 201-203.

12 - AA II, p. 161; trad. it., *L'unico argomento possibile per una dimostrazione dell'esistenza di Dio*, cit., p. 207. Anche nella *Critica della ragion pura* la prova ontologica verrà intesa come quella che, infine, sorregge tutte le altre: cfr., AA IV, p. 631 (A); AA III, p. 658 (B); trad. it., *Critica della ragion pura*, cit., p. 399.

ma il certo[13]. Kant si è difatti confrontato con l'esposizione cartesiana della prova ontologica, avanzando una critica sulla base di quanto già menzionato, ovvero a partire dalle considerazioni secondo le quali non risulta possibile intendere l'esistenza come un predicato della perfezione, altrimenti:

> si potrebbe dal concetto di ogni cosa, purché rappresentata anche come la più perfetta della sua specie, concludere alla sua esistenza; per esempio, concludere alla esistenza di un mondo perfettissimo, già per il solo fatto che può essere pensato[14].

Questa argomentazione richiama quella più celebre, scritta successivamente nella *Critica della ragion pura*[15], risulta però altrettanto immediato il richiamo a quella critica che Gaunilone aveva mosso ad Anselmo d'Aosta, rispetto a quella che veniva indicata come "prova ontologica"[16]. L'argomento, perché in tal guisa lo si dovrebbe piuttosto richiamare, è il seguente: «Dunque, Signore, non solo sei ciò di cui non si può pensare il maggiore, ma sei anche qualcosa di maggiore

---

13 - La verità è per il Kant della prima *Critica* un «nome allettatore» la cui ricerca rischia di traviare condurre all'errore: cfr., AA IV, p. 236 (A); AA III, p. 295 (B); trad. it., *Critica della ragion pura*, cit., p. 199. Quando nella *Critica della ragion pura* si usa la parola "verità": «vi si intende la verità del "giudizio": o piuttosto la sua "veridicità" chi sa mai come dichiarabile. […] In una tale esaustione della verità nell'ambito del giudizio, è chiaro che la regola fondamentale del giudicare, il principio di non contraddizione, diventi il criterio stesso e la "pietra di paragone" della verità», in E. Mirri, *Sull'uso della parola "verità" nella* Critica della ragion pura, in Id., *Pensare il Medesimo. Raccolta di saggi*, a cura di F. Valori-M. Moschini, Edizioni Scientifiche Italiane, Napoli 2006, pp. 287-295. qui p. 290. Si comprendono le critiche di Hegel, che nella sua *Fenomenologia dello spirito* scrive dalle paura dell'errore, come paura della verità: «Frattanto, se la tema di cadere in errore insinua sfiducia nella scienza che senza preoccupazioni di questo genere si accinge da sé all'opera e realmente conosce, bisogna vedere come mai, per contro, in questa diffidenza non si debba insinuare una diffidenza e non s'abbia a temere che una tale paura di errare non sia già stessa l'errore», GW 9, p. 54; trad it. E. De Negri, *Fenomenologia dello spirito*, 2 voll., La Nuova Italia, Firenze 1973: vol. 1, p. 67.

14 - AA II, p. 156; trad. it., *L'unico argomento possibile per una dimostrazione dell'esistenza di Dio*, cit., p. 202.

15 - Cfr., AA IV, p. 599 (A); AA III, p. 627 (B); trad. it., *Critica della ragion pura*, cit., pp. 382-383.

16 - Si pensi ad esempio al celebre esempio dell'isola: «Qualcuno dice che in qualche parte dell'Oceano vi è un'isola che, a causa della difficoltà o piuttosto della impossibilità di trovare ciò che non esiste, alcuni chiamano "perduta". […] Ma se poi aggiungesse, come se fosse una conseguenza: non puoi dubitare che questa isola migliore di tutte le terre esista veramente da qualche parte in realtà, più di quanto non dubiti che esista nel tuo intelletto; e poiché è migliore essere non solo nell'intelletto, ma anche nella realtà, è necessario che quell'isola esista nella realtà, perché se non esistesse nella realtà qualsiasi altra terra esistente nella realtà sarebbe migliore di essa, e così l'isola già da te intesa come superiore non sarebbe superiore», in Gaunilone, *Difesa dell'insipiente*, in Anselmo d'Aosta, *Proslogio*, p. 341, in Id., *Monologio e Proslogio*, trad. it. I. Sciuto, Bompiani, Milano 2002, pp. 363-379. Qui p. 375.

di quanto si possa pensare»[17]. Una prima cosa subitaneamente emerge con chiarezza: nell'affermazione citata manca uno sviluppo di una qualche forma di dimostrazione. L'argomento non sta presentando una definizione o una prova, ma attesta un'esperienza estatica di un soggetto che, d'innanzi a ciò che risulta così grande da esperire, dichiara l'insufficienza stessa dell'attività razionale, giacché ciò che viene pensato non è semplicemente un oggetto, ma è ciò che sta al di là dello stesso pensiero. Cosa allora potrebbe restituire il senso di questa attività pensante, in una posizione che sta al di là della prospettiva kantiana, se non il richiamare il genitivo soggettivo di tale pensare dell'essere, dove a pensare non è allora il soggetto che, nell'attività di giudicare, determina l'idea di Dio come oggetto, ma a pensare è lo stesso Dio e a tale attività il soggetto si abbandona. Il limite dell'attività del soggetto non può essere considerata altrimenti che nell'insufficienza di determinare ciò che per sua stessa natura non può essere assoggettato a determinazione alcuna. Gaunilone critica Anselmo perché ha concepito l'argomentazione come una prova, ovvero come deduzione del predicato di esistenza a partire dalla perfezione del concetto Dio[18]. L'argomento di Anselmo non voleva però provare nulla. In Kant, nel suo segnare i limiti della ragione, non si nega allora il problema Dio, si confuta la volontà di rendere l'idea di Dio un oggetto. Altrimenti, non sarebbe stata giustificata la successiva postulazione dell'esistenza di Dio nella *Critica della ragion pratica*[19]. Una questione risulta di particolare rilevanza, ovvero la consapevolezza di Kant che la riduzione di Dio ad oggetto della dimostrazione rischia di portare ad un ateismo, come nel caso della argomentazione kantiana sulla prova fisico-teologica negli *Scritti precritici*:

> Questo metodo può servire soltanto a provare un autore dei nessi e dei congegni artificiali del mondo, ma non a provare l'autore della materia stessa e l'origine degli elementi dell'universo. Questa notevole deficienza necessariamente lascia tutti quelli che si servono soltanto di esso metodo, in pericolo dell'errore detto ateismo raffinato, pel quale Dio, propriamente inteso, vien considerato come direttore e non come creatore del mondo, direttore che certo ha ordinata e formata la materia, ma che non l'ha prodotta, non l'ha creata[20].

---

17 - Anselmo d'Aosta, *Proslogio*, in Id., *Monologio e Proslogio*, cit., pp. 305-361. Qui p. 341.

18 - Per un'analisi dedicata all'argomento ontologico, in particolare per la distinzione tra argomento e prova si rimanda a: M. Casucci, *Deus summe sensibilis. Argomento ontologico e resurrezione estetica del pensare in S. Anselmo d'Aosta*, in «La Nottola di Minerva. Journal of Philosophy and Culture», XX/1, 2022, pp. 32-40. Per l'evoluzione dell'argomento ontologico si rimanda a: F. Tomatis, *L'argomento ontologico. L'esistenza di Dio da Anselmo a Schelling*, Città Nuova, Roma 1997.

19 - Cfr., AA V, p. 124; trad. it., F. Capra, *Critica della ragion pratica*, Laterza, Roma-Bari 1997, p. 273.

20 - AA II, pp. 122-123; trad. it., *L'unico argomento possibile per una dimostrazione dell'esistenza di Dio, cit., pp.* 166-167.

È questo il punto cruciale che verrà poi rielaborato successivamente nell'idealismo tedesco: a cosa serve un Dio distante, un «direttore», che nulla ha da dire rispetto a quelle che costituiscono le sfide della vita? Insistere su un Dio razionalmente dimostrato, non potrà che condurre all'ateismo, giacché ciò che si ottiene dalla dimostrazione non potrà che essere un oggetto indifferente rispetto all'umano e alla creazione tutta[21].

Nonostante il pensiero kantiano sia squisitamente impregnato di elementi illuministici, tali da portarlo in seguito a negare la validità stessa della rivelazione, relegata a quelli che sono solamente da considerare quali meri fini educativi in vista del perseguimento di un mito di una ragione autosufficiente[22], resta il fatto che il problema era stato lanciato. La *Critica della ragion pura* ha definitivamente aperto il problema ed il rimando alla *Critica della ragion pratica* ha mostrato come in verità il problema Dio non sia da risolvere per mezzo di una dimostrazione razionale, ma attraverso un libero esercizio della libertà che ha da esser costantemente attuata per mezzo dell'uso della ragione: questo allora costituisce il mito della ragione che, lungi dal porre una determinazione dell'idea di Dio, sempre ricerca una possibile approssimazione. Resta da chiedersi, e ciò Kant lo ha fatto, nella *Critica del giudizio,* come sia possibile riconciliare un atteggiamento determinante del soggetto con un libero agire dello stesso. Quello di cui si necessita è un'argomentazione fondata proprio sul silenzio di questa ragione che, partendo dai limiti costitutivi che la caratterizzano, volge lo sguardo verso una metafisica che fa dell'ascolto verso un contenuto rivelativo che non può esser determinato, ma di cui purtuttavia si partecipa. Questo uno degli argomenti ereditati dal pensiero kantiano che ha contribuito a ritornare a quel pensare dell'essere che, all'interno di un mondo

21 - Come sosterrà Schelling, in merito, proprio la non distinzione tra una prospettiva razionale, ed una che ricomprende a posteriori ciò che costituisce il contenuto teologico, è ciò che conduce all'ateismo, al nichilismo: cfr., SW XIII, p. 80; trad. it, *Filosofia della rivelazione*, cit., pp. 131-133. L'idea di Dio non restituisce la relazione genuina che la persona ricerca di intrattenere col fondamento. Il personale, sostiene Schelling, cerca ciò che parimenti è personale: Cfr., SW XI, p. 566-567; trad. it. L. Lotito, *Introduzione filosofica alla filosofia della mitologia*, Bompiani, Milano 2002, pp. 611-613. Non è un caso che nell'avvenuta incomprensione di questa differenziazione di paradigma, e nel permanere nella scissione tra finito e infinito, verrà decretata la morte di Dio da Friedrich Nietzsche: cfr., F. Nietzsche, *Sämtliche Werke. Kritische Studienausgabe in 15 Einzelbänden*, a cura di G. Colli e M. Montinari, Deutscher Taschenbuch Verlag-de Gruyter, München-Berlin-New York 1988: vol. 3, pp. 480-481; trad. it., *La gaia scienza*, p. 129, in *Opere*, vol. V, Tomo II, versioni di F. Masini e M. Montinari, Adelphi, Milano 1965.

22 - Si rimanda alle parole dello stesso Kant: «La fede storica allora può bastare, sì, per la fede di chiesa (di cui si possono avere molte forme); ma solo la fede religiosa pura, interamente fondata sulla ragione, può essere riconosciuta come necessaria e, per conseguenza, come la sola che caratterizzi la chiesa *vera*», in AA VI, p. 115; trad it., A. Poggi, *La religione entro i limiti della sola ragione*, Laterza, Roma-Bari 1980, p. 125.

costituito dalla perenne volontà di determinare l'alterità, si ripropone come cifra inattuale dell'umano.

Conclusioni: il limite della ragione come apertura per un ripensamento della metafisica

In questo studio si sono ripercorse alcune tematiche fondamentali del pensiero kantiano che hanno condotto la modernità ad affrontare la sfida relativa alla necessità di confrontarsi col tema metafisico. L'impossibilità di sviluppare coerentemente una dimostrazione dell'esistenza di Dio ha rafforzato ciò che costituisce il limite della ragione del soggetto, rimarcando la distinzione, operata da Kant, tra un mondo fenomenico ed uno noumenico. La sfida principale scaturisce proprio dalla necessità di doversi misurare con la frattura. È possibile pensare ad una metafisica post-kantiana o la definizione dei limiti della ragione hanno posto sotto scacco ogni possibile ripensamento di un pensiero speculativo che non sia concepito come scienza? Forse occorre proprio riflettere sul significato che tale limite assume alla luce dell'impossibilità di porre una dimostrazione dell'esistenza di Dio. Riflettere sul fatto che il soggetto abbia un limite non comporta inevitabilmente assumere una posizione di chiusura: la kantiana distinzione tra fenomeno e noumeno pone i limiti sul piano della conoscibilità, non una differenza che si consuma sul piano squisitamente ontologico. Ciò che è possibile sostenere a partire da queste riflessioni, è una maggiore consapevolezza verso la posizione che assume il soggetto: se esso viene considerato solamente come soggetto del giudizio esso assume la forma di un giudice che è relegato al solo piano fenomenico. La possibilità di porre uno sguardo con rinnovato senso all'esistenza è data dalla possibilità di affiancare, a questo paradigma del soggetto razionale, una dimensione di apertura che ricomprende a posteriori un contenuto che eccede il piano del giudizio del soggetto, come mostra la riflessione post-kantiana.

Questo è il risultato di alcuni sviluppi che, come si è voluto mostrare in questo lavoro, sono stati possibili a partire dal confronto con il pensiero kantiano. Da quanto emerso, è possibile sostenere che ritornare all'umano, indagandone la modalità di stare al mondo può essere considerata una delle sfide ereditate da Kant che ancora risultano preziose per la nostra contemporaneità.

FRANCESCO PORCHIA
Dottorando in *Advanced Theological Studies* presso l'Università di Vienna ed è Postdoctoral Fellow presso la Facoltà di Teologia Protestante della stessa università.
francesco.porchia@collaboratori.unipg.it

Baruch Spinoza e l'idea vera. La conoscenza di Dio come conoscenza-in-Dio nel *Trattato sull'emendazione dell'intelletto* e nel *Breve trattato su Dio, l'uomo e la sua felicità*

di Federico Brutti

The paper aims to examine Spinoza's vision of the «true idea», highlighting how this concept enables the Dutch philosopher to radically rethink the process of God's self-revelation and the traditional theological constructs of Providence, freedom, and salvation.

Keywords: Spinoza; Ontology; Theology; True Idea; Beatitude.

1. Il trattato sull'emendazione dell'intelletto

Il presente contributo intende ricostruire la genesi della dottrina spinoziana concernente «l'idea vera», così come essa viene a configurarsi nelle meditazioni contenute nel *Trattato sull'emendazione dell'intelletto* e nel *Breve trattato su Dio, l'uomo e la sua felicità,* feconde anticipazioni teoretiche del più maturo frutto dell'*Etica* e del suo rigoroso sistema[1]. Non vi è modo migliore, per introdursi al pensiero spinoziano, che volgersi alla lettura del celebre racconto di apertura che segna l'inizio del *Trattato sull'emendazione*. Si tratta di un documento di inestimabile valore, non soltanto per le chiare indicazioni teoretiche in esso comunicate – le quali testimoniano la sostanziale e profonda continuità delle posizioni dell'autore –, ma anche in quanto vi trovano espressione, con eccezionale vividezza, lo slancio esistenziale e la passione dell'anima informanti l'indagine speculativa spinoziana. Si esprime qui, in altri termini, un carattere essenziale della dottrina di Spinoza, così riassumibile: si dà autentico pensare soltanto laddove esso diventi capace di risultare salvifico, introducendo al godimento del sommo bene. Vi si legge:

> Dopo che l'esperienza mi ebbe insegnato come fossero vane e futili tutte quelle cose che capitano così frequentemente nella vita quotidiana; e vedendo che ciò che mi atterriva o che temevo, in sé non aveva niente né di buono né di cattivo se non in quanto l'animo ne veniva scosso, decisi infine di cercare se esistesse qualcosa di veramente buono e che fosse di per sé accessibile, e da cui solo, abbandonati tutti gli altri, l'animo potesse venire affetto; meglio ancora, se esistesse qualcosa che, una volta trovato e acquisito, mi facesse godere in eterno di una continua e somma letizia[2].

---

1 - Per ciò che concerne la letteratura secondaria riguardante il *Trattato sull'emendazione*, si rimanda – senza pretesa di esaustività – ai seguenti studi fondamentali: F. Mignini, *Per la datazione e l'interpretazione del Tractatus de intellectus emendatione di B. Spinoza*, «La Cultura», 17 (1-2) 1979, pp. 87-160; G. Boss, *Méthode et doctrine dans le Traité de la Réforme de l'Entendement*, «Studia Spinoziana», 2, 1986, pp. 93-108; D. Garrett, *Truth and Ideas of Imagination in the Tractatus de Intellectus Emendatione*, «Studia Spinoziana», 2, 1986, pp. 61-92; F. Mignini, *Per una storia dell'argomento ontologico in Spinoza. Le "prove" del "Tractatus de intellectus emendatione"*, in *L'argomento ontologico*, a cura di M. M. Olivetti, Cedam, Padova 1990, pp. 203-221; L. Vinciguerra, *Iniziare con Spinoza. Errore e metodo nel Tractatus de Intellectus Emendatione*, «Rivista di storia della filosofia», 4, 1994, pp. 665-687. In riferimento, invece, al *Breve trattato*, si vedano tra gli altri: G. Boss, *L'enseignement de Spinoza: Commentaire du Court Traité,* Éditions du Grand Midi, Zürich 1982; H. Ravven, *Spinoza's individualism reconsidered: some lessons from the Short Treatise on God, man and his well-being,* «Iyyun», 47, 1998, pp. 265-292; F. Mignini, *Il lessico teologico nella Korte Verhandeling di Spinoza,* in *Studi in onore di G. Sasso,* Bibliopolis, Napoli 1999, pp. 373-401; trad. ted., *Die theologische terminologie in Spinozas Korte Verhandeling*, «Studia Spinoziana», 14, pp. 137-157.

2 - B. Spinoza, *Trattato sull'emendazione dell'intelletto, e sulla via per dirigerlo nel modo migliore alla vera conoscenza delle cose* (1660), in Id., *Tutte le opere*, trad. it. a cura di A. Sangiacomo, Bompiani Edito-

La narrazione autobiografica si fa portavoce di una crisi intellettuale che, ben più radicalmente del ridursi a mera questione di conoscenza astratta, investe invece la concretezza dell'esistenza nella sua integralità, nascendo dall'insopprimibile bisogno di felicità che alberga nel cuore di ogni essere umano. È qui soltanto che l'autentica interrogazione filosofica conosce la luce: essa prende forma dall'instancabile desiderio di trovare una risposta pura e definitiva alla domanda sul senso del vivere, domanda che risulta costantemente tradita e disattesa dalla vanità delle esperienze che affollano il quotidiano. La filosofia – questo pare dirci Spinoza – comincia laddove si sperimenti la totale relatività dei valori transeunti e dei beni effimeri, per i quali l'uomo comune sempre si affanna. Perciò, la ricerca del sommo bene richiede una scelta estrema, che non prevede sconti, giacché ogni concessione al desiderio di piacere, onore e ricchezza – considerati come beni in sé – non fa che allontanare, fatalmente, dal raggiungimento della meta. Ne consegue che il cammino della conoscenza comporti, almeno agli inizi, un'inevitabile sofferenza: si affaccia il tarlo del pessimismo, la tendenza alla malinconia. Compito del filosofo è saper abitare e quindi trascendere tale fase, non fuggendo dinanzi alla gravità della riflessione, in modo da aprirsi, in ultimo, alla gioia coincidente con l'esperienza del vero bene. È possibile riconoscere qui, implicitamente, le tracce di quella che diverrà una tesi cardine dell'*Etica*: il primato accordato alla dimensione dell'affettività, della passione – il *conatus* –, da cui discende l'impossibilità di sanare un affetto negativo attraverso la fuga nell'intellettualità astratta; una passione negativa può venire rovesciata soltanto da una passione di segno contrario, più forte, più stabile, che accresca la propria potenza al sommo grado. Non si dà passione più salda – qui il punto decisivo – di quella che sappia rivolgersi all'essere eterno e perfetto, in quanto il suo oggetto incarna ciò che, in assoluto, mai può venir meno: il destino della nostra felicità o infelicità, cioè, «risiede unicamente nella qualità dell'oggetto a cui l'amore ci unisce»[3]. Risulta già evidente, dalle brevi notazioni considerate, come il sommo bene spinoziano sia assolutamente irriducibile a un giudizio qualitativo operato dall'essere umano, in accordo ai criteri attraverso i quali egli comunemente si orienta nel vivere quotidiano. Se così fosse, esso finirebbe per venire subordinato alla passionalità particolare e alla limitata prospettiva intellettuale che definisce la condizione del singolo. L'unico bene davvero autentico, la pura perfezione in sé, non può che identificarsi con la realtà stessa: la piena coincidenza tra essere – Dio –, verità e bene è considerabile a tutti gli effetti la tesi capitale del filosofare spinoziano. Il pensatore olandese addita, qui, l'inconsistenza e dunque la relatività

re, Milano 2021, pp. 103-184, qui p. 111.

3 - *Ibid.*, pp. 113-115.

del bene e del male empirici, in quanto valori originantesi dallo sguardo utilitaristico ed egocentrico dell'umano:

> Occorre notare che bene e male si dicono solo in senso relativo, sicché una sola medesima cosa può essere detta buona o cattiva a seconda dei diversi punti di vista, e lo stesso vale per perfetto e imperfetto. Niente, infatti, considerato nella sua natura, si dirà perfetto o imperfetto, soprattutto dopo che avremo saputo che tutto ciò che accade, accade secondo un ordine eterno e secondo determinate leggi di Natura[4].

All'occhio limitato di chi identifica il proprio interesse particolare – quindi ciò che risulta favorevole o spiacevole – con la manifestazione reale del bene e del male, Spinoza contrappone l'ideale del sapiente capace di innalzarsi alla prospettiva dell'eternità e del suo inscalfibile decreto. In ciò consta il sommo bene: nella conoscenza e nell'amore di ciò che, semplicemente e unicamente, è; esso corrisponde cioè alla «conoscenza dell'unione che la mente ha con tutta la Natura»[5]. Si tratta di un'impresa immediatamente configurantesi nella sua portata etico-politica, in quanto l'innalzamento spirituale che conduce al godimento del primo principio, nella consapevolezza dell'unione che a lui ci lega, insieme rivela – a un tempo – la strutturale condivisione di tale appartenenza con l'intera comunità degli essenti, suscitando il vivo desiderio che la medesima esperienza indiante possa essere goduta da altri e con altri:

> Questo è dunque il fine al quale tendo: acquisire una tale natura e cercare che molti insieme a me l'acquisiscano; cioè fa parte della mia felicità anche l'impegnarmi perché molti altri comprendano ciò che io ho compreso e perché il loro intelletto e i loro desideri si accordino con i miei[6].

Il progetto del filosofare spinoziano appare così, fin dalla sua prima ed embrionale esposizione, lucidamente orientato all'unica ed essenziale meta: pervenire alla certa identificazione con il principio, nell'esperienza comunitariamente condivisa del sapersi in-Dio.

1.1 I modi della conoscenza

Indicata la direzione speculativa della ricerca, non resta che procedere alla disamina metodologica delle differenti modalità di conoscenza caratterizzanti la riflessione umana, cosicché il processo di emendazione dell'intelletto possa

---

4 - *Ibid.*, p. 115.

5 - *Ibid.*, p. 117.

6 - *Ibid.*

concretamente prendere forma e volgersi all'intendimento dell'idea vera, scartando gli errori e le rigide astrazioni. Scrive, a tal proposito, Spinoza:

> Se faccio un'accurata disamina, tutti questi modi si possono ridurre senza problemi a quattro:
>
> 1. C'è una percezione che abbiamo per sentito dire o da qualche altro segno, che viene detto arbitrario.
>
> 2. C'è una percezione che abbiamo per esperienza vaga, cioè tramite un'esperienza non determinata dall'intelletto; viene chiamata così solo perché, essendocisi offerta occasionalmente e non avendo noi nessun altro esperimento che la contrasti, essa continua a restare valida come non smentita.
>
> 3. C'è una percezione dove l'essenza della cosa si inferisce da un'altra cosa, ma non adeguatamente. Ciò accade o quando da una qualche effetto risaliamo alla sua causa, oppure quando si conclude da un qualche universale che è sempre accompagnato da una qualche proprietà.
>
> 4. Infine, c'è una percezione nella quale la cosa è percepita mediante la sua sola essenza, o tramite la percezione della sua causa prossima[7].

Laddove nei primi due livelli – quelli del "sentito dire" e dell'esperienza vaga – la conoscenza ci si offre in maniera del tutto casuale, e perciò arbitraria, in quanto non generantesi per mezzo del contributo attivo dell'intelletto, il terzo grado della percezione identifica l'ingresso nella dimensione della *ratio*: qui, il reale risulta comprensibile – cioè concatenabile – mediante l'applicazione del principio di causa, e attraverso la riconduzione dell'essente particolare a un universale in grado di esplicarlo e quindi di renderne ragione. Spinoza riconosce, in una simile forma di conoscenza, un sapere soltanto parzialmente adeguato, in quanto strutturalmente costretto a ricavare da altro, cioè in maniera indiretta, l'essenza di ciò che si indaga. Al contrario, la conoscenza autentica, quella esprimentesi nel quarto modo, si concretizza in una esperienza diretta dell'ente, colto immediatamente nella sua essenza o in quella della sua causa prossima. In virtù di una tale capacità intuitiva di comprensione dell'essenza, il quarto livello della conoscenza è eletto dall'autore come modalità preferenziale di indagine speculativa, in quanto attraverso di esso è possibile trascendere assolutamente il pericolo dell'errore, nella misura in cui l'errore stesso può nascere soltanto laddove si dia scollamento tra l'essenza della cosa pensata e la ragione che tenti di rivolgersi a essa. Nel caso della percezione immediata, evidentemente, si assiste invece a una totale coincidenza tra pensiero e realtà, giacché l'essenza giunge qui ad auto-manifestarsi, comunicando direttamente se stessa nell'atto intuitivo che di essa risulta informato. Perciò Spinoza può affermare, più avanti, che «non può sapere cosa sia la somma certezza

7 - *Ibid.*, p. 119.

se non colui che abbia l'idea adeguata o l'essenza oggettiva di qualcosa, e ciò perché la certezza e l'essenza oggettiva sono il medesimo»[8].

1.2 L'idea vera

Riconosciuta la modalità atta a giungere alla conoscenza autentica – quella fondantesi sulla percezione intuitiva capace di cogliere immediatamente l'essenza della cosa –, Spinoza si accinge a definire più nel dettaglio quale sia l'oggetto supremo della conoscenza stessa, la realtà cui essa è chiamata a rivolgersi. Si affaccia qui, inevitabilmente, lo spauracchio del *regressus in infinitum*: come è possibile trascendere il pericolo che la ricerca della verità, e quindi del metodo per giungervi, non si infranga in un interminabile vicolo cieco? Così – lapidariamente – l'autore fuga le incertezze:

«Occorre in primo luogo considerare che non incorreremo in un'indagine all'infinito: per giungere al miglior metodo per ricercare il vero, non è necessario un altro metodo per ricercare il metodo per la ricerca del vero, e per cercare il secondo metodo non ne serve un terzo, e così via all'infinito. Infatti, in questo modo non si giungerebbe mai alla conoscenza del vero, anzi non si arriverebbe a nessuna conoscenza»[9].

Spinoza è, fin dall'inizio, saldo nel negare consistenza all'ipotesi per la quale la ricerca della verità possa condurre a un nulla di fatto. La ragione della sua certezza è la seguente: l'idea vera, proprio poiché sussistente in quanto tale, non può in alcun modo venire ridotta alla conseguenza di un ragionamento corretto; piuttosto, essa è da riconoscersi come il necessario fondamento di ogni credenza veritiera. La verità – detto altrimenti – non dipende dal metodo, ma ne è l'assoluto presupposto: la sua ricerca, da parte dell'intelletto, non può che concludersi nell'attestazione di ciò che, già da principio, si dà come presenza criteriante. Ne discende che, per sapere, non è in alcun modo necessario sapere di sapere – ciò comporterebbe il paradosso del regresso all'infinito –, poiché il possesso dell'idea vera è già, in sé, norma indubitabile del sapere stesso: «Alla certezza della verità non è necessario disporre di nessun altro segno che dell'idea vera»[10]. È evidente che, nel sostenere una tale tesi, Spinoza abbia in mente non un'idea vera qualsiasi, bensì l'idea assolutamente vera, in quanto concepibile come ragione e criterio della verità di tutte le altre: si tratta dell'idea dell'ente perfettissimo, ovvero dell'essere eterno e infinito. In quanto fondamento dell'essente tutto, infatti, l'idea dell'essere non può

8 - *Ibid.*, p. 129.

9 - *Ibid.*, p. 125.

10 - *Ibid.*, pp. 127.

che riconoscersi come la fonte e l'origine di ogni altra idea possibile: ciò significa che l'atto del pensare si configura immediatamente come pensiero dell'essere, che l'idea di ogni ente particolare richiede e presuppone il porsi dell'idea dell'essere come suo principio di esistenza e conoscenza. Prende forma, così, la suprema intuizione spinoziana per la quale «*Veritas se ipsam patefacit*»:

«Se qualcuno, nelle sue indagini sulla Natura, avesse proceduto per caso in questo modo, ossia acquisendo nell'ordine dovuto altre idee secondo la norma di un'idea vera data, non dubiterebbe mai della sua verità, giacché la verità, come si è mostrato, manifesta se stessa, e tutte le cose gli giungerebbero in modo addirittura spontaneo»[11].

Si profila però, a tale stadio della trattazione, il rischio di un ulteriore inciampo, riconosciuto dallo stesso Spinoza. Se davvero la verità fosse tale da manifestarsi autonomamente in quanto norma del pensare, per quale ragione – di fatto – si rende necessaria un'argomentazione discorsiva atta a giustificare una verità così intesa? Scrive l'autore:

> A questo punto ci si meraviglierà forse che noi, dopo aver detto che buon metodo è quello che mostra come sia da dirigere la mente secondo la norma dell'idea vera data, proviamo questa dottrina con un ragionamento, il che sembrerebbe mostrare che questa non sia per sé nota, e si può persino dubitare se ragioniamo bene. Se ragioniamo bene dobbiamo cominciare da una data idea, e poiché l'iniziare da una data idea necessita di dimostrazione, dovremmo di nuovo provare il nostro ragionamento, e poi di nuovo quell'altro e così all'infinito[12].

La risposta spinoziana alle obiezioni presentate si fonda, significativamente, sul carattere "pedagogico" riconosciuto al discorso intorno all'idea vera e al metodo per giungervi. È necessario, in tale direzione, distinguere due livelli di realtà e di pensiero: da un lato si dà l'intuizione dell'essere perfetto che, in sé, non richiede alcun metodo, essendo norma del metodo stesso; dall'altro, è compito del filosofo prendere coscienza della condizione deficitaria dell'uomo comune, perennemente travolto, sovrastato dalle passioni e dai pregiudizi che ne discendono. È a quest'ultimo – all'uomo nel suo *status* di naturale insipienza – che il discorso spinoziano sul metodo si rivolge, ed è soltanto nell'orizzonte di un simile livello della riflessione che la paradossalità ravvisata, come tale, può prendere forma. In altre parole: il sapiente, alla luce della superiore esperienza di verità alla quale si è innalzato, giunge ad abitare consapevolmente la dimensione del paradosso nell'intento di ricostruire per via discorsiva, e dunque mostrando progressivamente le tappe del proprio percorso intellettuale, il cammino di elevazione della coscienza

11 - *Ibid.*, p. 131.

12 - *Ibid.*

di cui si è reso protagonista e a cui invita il discepolo-lettore:

> Se per caso qualcuno mi chiedesse perché, subito e prima di tutto, abbia esposto proprio con quest'ordine le verità della Natura, dacché la verità manifesta se stessa, a costui rispondo e contemporaneamente lo esorto a non voler respingere quelle verità come false a causa dei paradossi che per caso si incontrino qua e là, ma prima ci si degni di esaminare l'ordine nel quale le proviamo e allora si sarà certi che abbiamo raggiunto la verità[13].

È particolarmente rilevante notare come, in questo contesto, Spinoza faccia riferimento a una Filosofia[14] sistematica – di cui, probabilmente, l'*Etica* diventerà l'effettiva incarnazione – rispetto alla quale il *Trattato sull'emendazione* non costituisce che una tappa propedeutica, una sorta di iniziazione speculativa[15].

2. Il Breve trattato su Dio, l'uomo e la sua felicità

È ora opportuno considerare altresì il contributo speculativo offerto dal *Breve trattato su Dio, l'uomo e la sua felicità*. In tale scritto giovanile – tornato alla luce soltanto posteriormente, nel XIX secolo – è già possibile, infatti, riconoscere i capisaldi teoretici della dottrina spinoziana, gli stessi che, nella forma dell'ordine geometrico, troveranno espressione nell'*Etica*. Significativamente, l'esordio dell'opera è già informato della prospettiva teologica sopra considerata, per la quale il discorso su Dio non può che prendere avvio da Dio stesso. Il capitolo primo – dal titolo emblematico *Che Dio esiste* – è così introdotto:

> Esiste un Dio? Affermiamo che lo si può dimostrare:
> *A priori* così:
> 1. Tutto quello che intendiamo chiaramente e distintamente appartenere alla natura di una cosa, possiamo affermarlo secondo verità di quella cosa; ma possiamo chiaramente e distintamente intendere che l'esistenza appartiene alla natura di Dio; dunque (Dio esiste, ndr).

---

13 - *Ibid.*, p. 133.

14 - *Ibidem*: «Il motivo per cui capiti raramente che le indagini sulla natura si facciano nel debito ordine, è dato dai pregiudizi, le cui cause spiegheremo poi nella nostra Filosofia.»

15 - Obiettivo del Trattato spinoziano è quello cioè, di pervenire alla chiarificazione dell'idea vera, distinguendola «da tutte le altre percezioni, e impedire alla mente di confondere le idee false, le idee finte e le idee dubbie con le idee vere» (*Ibid.*, p. 135). Nel caso dell'idea finta, «si finge la sola esistenza e si conosce, o si suppone di conoscere, la cosa che in tale atto si finge» (*Ibid.*); nell'idea falsa si aggiunge, alle caratteristiche sopra elencate, l'assenso soggettivo; infine, si dà l'idea dubbia, che coinvolge una relazione tra idee, nella misura in cui il dubbio dipende dall'influenza di un'idea non chiara né distinta sulle altre che richiedono il nostro assenso. Le idee chiare e distinte sono quelle che, sottrattisi all'arbitrarietà dei sensi e delle opinioni, risultano «prodotte dalla pura mente» (*Ibid.*, p. 161).

Oppure così:

2. Le essenze delle cose sono da tutta l'eternità e per tutta l'eternità devono restare immutabili; l'esistenza è essenza di Dio; dunque (Dio esiste, ndr).

A posteriori così:

se l'uomo ha un'idea di Dio, Dio deve esistere formalmente; ma l'uomo ha un'idea di Dio; dunque (Dio esiste, ndr)[16].

Traspare in maniera evidente, in queste righe, la ripresa spinoziana dell'argomento ontologico di Anselmo: l'unico perno attorno al quale ruotano le varie formulazioni espresse da Spinoza, infatti, è rintracciabile nell'intuizione fondamentale per la quale, in Dio, l'essenza è esistenza e l'esistenza è essenza. In altre parole: non può darsi l'idea di Dio senza che, in essa, non si dia anche la sua realtà, giacché l'idea è da riconoscersi come la presenza immediata della realtà stessa. L'uomo si scopre, così, abitato dalla presenza di Dio, in quanto la manifestazione dell'idea rivela la realtà di chi, lungi dal potersi ridurre a oggetto della ragione discorsiva, ci rende in lui essenti e pensanti: «Che l'uomo abbia l'idea di Dio risulta chiaramente dal fatto che egli ne conosce gli attributi, il quali attributi non possono essere da lui prodotti, in quanto lui è imperfetto»[17]. Il finito, essendo tale, può accedere all'esperienza intellettuale dell'infinito[18] soltanto laddove sia l'infinito stesso a comunicarsi in esso. Il favore accordato da Spinoza all'argomento ontologico assume, peraltro, esplicita espressione:

> Da tutto questo (dagli argomenti attraverso cui si mostra la necessità dell'essere-Dio, ndr), dunque, segue chiaramente che l'esistenza di Dio può essere dimostrata tanto a priori quanto a posteriori. Anzi, molto meglio a priori: infatti, le cose che non si dimostrano in tal modo, devono essere dimostrate ricorrendo alle cause esterne, il che implica un'evidente imperfezione, perché esse non possono farsi conoscere da se stesse, ma solamente in forza di cause esterne. Dio, tuttavia, prima causa di tutte le cose, è anche causa di se stesso, si fa conoscere per se stesso[19].

La dimostrazione *a posteriori* dell'esistenza di Dio, rileva Spinoza, finisce per

---

16 - B. Spinoza, *Breve trattato su Dio, l'uomo e la sua felicità* (1662), in Id., *Tutte le opere*, cit., pp. 185-362, qui p. 193.

17 - *Ibid.*, p. 197.

18 - *Ibidem*: «Egli sa, per esempio, che l'infinito non può essere formato da diverse parti finite; che non vi possono essere due infiniti, bensì uno solo; che esso è perfetto e immutabile; sa anche che nessuna cosa ricerca da sé stessa la propria distruzione, e poi che l'infinito non si può mutare in qualcosa di meglio, poiché è perfetto, e non sarebbe tale se cambiasse; e ancora che non può essere subordinato a qualche altra cosa, poiché è onnipotente etc.»

19 - *Ibid.*, p. 199.

ridurre il principio a ente tra i molti enti, a causa tra le molte cause – per quanto prima e perfetta –, giacché scade nella conseguenza paradossale di far dipendere la conoscenza di Dio dalla dimensione mondana, e quindi dalle categorie che la definiscono. Il principio è così, a un tempo, ridotto ai criteri della sfera empirico-rappresentativa e collocato in un'inattingibile ulteriorità che la trascende, rendendone strutturalmente impossibile l'esperienza. Invece «egli si fa conoscere per se stesso»[20]: in ciò si esprime l'inscalfibile certezza di cui l'argomento ontologico è portavoce.

2.1 L'essenza di Dio e la salvezza dell'uomo

Avendo chiarito il carattere di assoluta incontrovertibilità dell'idea di Dio – il quale, quindi, non può non essere –, Spinoza si volge all'indagine intorno alla sua essenza. La prima definizione fornita dall'autore – anticipando quella che, quasi con le medesime parole, si troverà nell'*Etica*[21] – è la seguente: «Diciamo (di Dio, ndr) che è un ente di cui si può affermare tutto, cioè un numero infinito di attributi, ciascuno dei quali è infinitamente perfetto nel suo genere»[22]. In quanto essere assoluto e, perciò, onnicomprensivo, Dio deve necessariamente possedere un'infinità di attributi, – ovvero di proprietà essenziali – la cui natura è, per ognuno, eternamente e perfettamente compiuta, con ciò escludendo ogni mancanza reale o finanche virtuale. Prosegue, poi, il filosofo olandese:

> Per esprimere chiaramente il nostro pensiero, porremo le seguenti quattro proposizioni:
>
> 1. Non esiste una sostanza finita, ma ogni sostanza deve essere infinitamente perfetta nel suo genere: cioè nell'intelletto infinito di Dio non vi può essere alcuna sostanza più perfetta di quella che nella natura già esiste.
>
> 2. Non esistono due sostanze uguali.
>
> 3. Una sostanza non può produrne un'altra.
>
> 4. Nell'intelletto infinito di Dio non vi è alcuna sostanza che non esista formalmente nella natura[23].

---

20 - «Invero, poiché Dio è causa di sé stesso, basta che lo proviamo per sé stesso, e una tale prova è molto più rigorosa della prova *a posteriori*, che non ha luogo comunemente che per mezzo delle cause esterne.» *Ibid.*, p. 241.

21 - B. Spinoza, *Etica dimostrata secondo l'ordine geometrico* (1677), in Id., *Tutte le opere*, cit., pp. 1143-1623, qui p. 1147: "Intendo per Dio un ente assolutamente infinito, cioè, una sostanza costituita da un'infinità di attributi, ciascuno dei quali esprime un'essenza eterna ed infinita".

22 - B. Spinoza, *Breve trattato*, cit., p. 199.

23 - *Ibid.*, pp. 199-201.

Le proposizioni qui elencate concorrono – nel loro insieme – a giustificare la tesi più estrema e radicale del filosofare spinoziano, quella per la quale si dia un'unica e infinita sostanza, coincidente con Dio stesso. Alla luce dell'intelletto emendato, ovvero innalzatosi alla pura contemplazione dell'idea vera, tale posizione appare di per sé evidente, tautologica. Come potrebbe darsi, infatti, una molteplicità di sostanze, laddove si sia pervenuti all'intendimento dell'essere come realtà che assolutamente esclude da sé ogni negazione? Significherebbe, in tutta evidenza, negare Dio nell'atto verbale che, contraddittoriamente, lo afferma; giacché una simile ammissione equivale a sancirne la limitazione, il confinamento nell'orizzonte del finito. In altri termini: Dio o è infinito, o non è; ma, in quanto sappiamo di lui alla luce dell'idea vera, non possiamo che riconoscerlo come l'unica sostanza, che in quanto tale si rivela. Ne discende che Spinoza non possa accettare, in forza della dottrina ontologica che informa il suo pensiero, alcuna riduzione di Dio alle categorie antropomorfiche dell'individualità personale, del libero arbitrio, della sospensione miracolosa delle leggi di natura; in tali concezioni egli non ravvede, infatti, nient'altro che la traccia del limite che l'inferiore prospettiva umana impone al principio, nel rappresentarsene un'immagine ricalcante le proprie virtù e i propri vizi. Attribuire all'unica sostanza il libero arbitrio, la facoltà di modificare e rivedere i propri decreti, significherebbe nientemeno che ammetterne la passività, la dipendenza dall'altro, cioè l'imperfezione. In Dio, tutto è a-temporalmente e assolutamente compiuto, nulla di altro può accadere se non il suo eterno e perfetto esplicarsi negli altrettanto infiniti attributi di cui essenzialmente consta. In quanto sostanza infinita, il principio deve dunque considerarsi, propriamente, come l'unica e assoluta causa, giacché tutto ciò che accade, accade in lui. Egli è causa immanente e non transeunte di ogni cosa: direttamente, e perciò immediatamente, degli enti infiniti e immutabili; indirettamente, e quindi in quanto causa ultima, di ogni affezione e modo particolare. La libertà di Dio – l'unica, autentica libertà – coincide così con la necessità della sua infinita ed eterna potenza. Scrive l'autore:

> Alcuni considerano ciò (il fatto che Dio non possa agire in modo diverso da come, eternamente, agisce, ndr) come un'imperfezione e un'ingiustizia contro Dio; errore che nasce soltanto dal fatto che non concepiscono rettamente in che cosa consista la vera libertà, la quale non può in alcun modo identificarsi, come s'immagina, nel poter fare o omettere il bene o il male. La vera libertà non è altro che la causa prima, la quale da null'altro sollecitata o costretta, solo in virtù della sua perfezione è causa di ogni perfezione[24].

24 - *Ibid.*, pp. 225-227.

Ne consegue evidentemente che la tesi teologica di Spinoza, con la sua correzione del vizio antropomorfico nell'affermazione dell'unica ed eterna sostanza, comporti una sensibile rivisitazione della tradizionale dottrina della provvidenza. Quest'ultima, infatti, per le ragioni sopra discusse, non può ridursi all'idea di un particolare e straordinario intervento di Dio nelle vicende storiche umane. Nella prospettiva spinoziana, la provvidenza divina non può che coincidere con il suo perfetto ed eternamente compiuto atto generativo, informante la realtà tutta. Essa «non è altra cosa che lo sforzo che noi scopriamo in tutta la natura e in tutte le cose particolari, il quale tende al mantenimento e alla conservazione del loro essere»[25]. Ne discende una fondamentale conseguenza soteriologica: per l'essere umano, la beatitudine consiste nella capacità di far coincidere, coscientemente, il raggio della propria libertà con l'inscalfibile necessità dell'essere-Dio[26]. Essa definisce, quindi, il raggiungimento dell'ultimo livello della conoscenza, che Spinoza nomina qui, emblematicamente, «rigenerazione». Nuovamente, tale stadio del cammino coscienziale «non deriva da altro, ma scaturisce dalla manifestazione immediata dell'oggetto all'intelletto»[27]. Nell'atto della rigenerazione, la mente trascende la finitudine connessa alla prospettiva dell'individuo somatico – carnale – e si innalza fino a sapersi uno con l'essere eterno, accedendo così alla consapevolezza della propria immortalità. Salvezza e conoscenza, qui, pervengono a perfetta identificazione, dal momento che la conoscenza stessa risulta tutt'altro che intendibile come un mero processo di astrazione razionale, significando piuttosto il raggiungimento di un più elevato *status* di realtà, quello dell'essere-Dio. In consonanza con il radicale esito che prenderà forma nell'*Etica*, Spinoza può concludere: «L'uomo, come tutto ciò che esiste, è in Dio, di modo che Dio sta in tutte le cose e, a parlare con proprietà, non vi può essere in lui amore per nessun'altra cosa che per se stesso, poiché tutto è in lui»[28].

FEDERICO BRUTTI
Dottorando di ricerca in Etica della comunicazione, della ricerca scientifica e dell'innovazione tecnologica presso l'Università degli Studi di Perugia
federico.brutti@dottorandi.unipg.it

25 - *Ibid.*, p. 229.

26 - Più avanti, si legge ancora (*ibid.*, pp. 333-335): «Dopo averlo riconosciuto (Dio, ndr) come il migliore di tutti i beni, siamo obbligati ad arrestarci e riposare in lui, al di fuori del quale nulla può darci la pace: la nostra vera libertà è che noi siamo e rimaniamo incatenati dalle amabili catene del suo amore».

27 - *Ibid.*, p. 321.

28 - *Ibid.*, p. 327.

Automazione e cibernazione. Hannah Arendt tra *The Human Condition* e *On the Human Condition*

di Saverio Ferrieri

The essay examines the relationship between humans and machines in Hannah Arendt's thought through a comparison of The Human Condition, On the Human Condition*, and the* Cybernetics typescript. *The analysis highlights textual differences and continuities, emphasizing the importance of Arendt's reflection on technology and its connection with themes later developed in* The Life of the Mind.

Keywords: automation, cybernation, human condition, techno-scientific progress, IA.

1. Partecipare a una conferenza sessant'anni dopo

Nel giugno del 1964 un gruppo eterogeneo di studiosi, intellettuali e personalità di spicco[1] si riunisce presso l'Hotel Madison di New York per la *First Annual Conference on the Cybercultural Revolution – Cybernetics and Automation*. Principale promotrice della conferenza è Alice Mary Hilton[2], che aveva già riflettuto a lungo sulle idee elaborate dall'*Ad Hoc Committee on the Triple Revolution*, un gruppo di intellettuali e attivisti statunitensi che vedevano nelle attività delle moderne macchine computazionali – e in particolar modo nella *cibernazione*[3] – la premessa di un cambiamento sociale profondo e radicale.

Non a caso, proprio nel marzo del 1964, questo comitato aveva redatto una famosa lettera aperta, considerata il manifesto del movimento, indirizzata al presidente degli Stati Uniti Lyndon B. Johnson (1908-1973) e ad altri leader politici. Il testo, intitolato *The Triple Revolution*[4], individuava tre grandi trasformazioni simultanee che iniziavano a rimodellare profondamente la società americana e mondiale.

(1) La prima di queste trasformazioni viene lì definita "rivoluzione della *cibernazione*" (*The Cybernation Revolution*[5]). Essa segna l'inizio di una nuova era della produzione, basata su principi organizzativi tanto diversi da quelli dell'epoca industriale quanto l'era industriale lo era stata rispetto a quella agricola. La novità consiste nella combinazione tra macchine computazionali e macchine automatizzate autoregolatrici, dalla quale scaturisce un sistema con capacità produttive quasi illimitate, ma che necessita di una quantità sempre minore di manodopera. In questo senso, l'emergente *cibernazione* non si limita a introdurre nuovi strumenti di lavoro, ma sta già riorganizzando l'intero sistema economico e sociale in funzione delle proprie logiche.

---

1 - C. Bassett, *Anti-computing. Dissent and the machine*, Manchester University Press, Manchester 2021, p. 105.

2 - Alice Mary Hilton (1919-2011), matematica e attivista – da alcuni soprannominata "socratica petulante" (Socratic Gadfly), per via della facile inclinazione alla discussione, o indicata come la voce vellutata dell'automazione (Velvet Voice of Automation) –, insistette per avere la presenza di Hannah Arendt alla conferenza.

3 - Il termine *cybernation* viene coniato da Donald N. Michael (cfr. il suo *Cybernation: The Silent Conquest. A Report of the Center for the Study of Democratic Institution*, Center for the Study of Democratic Institution, Santa Barbara [CA] 1962) e con esso si intende esprimere la convergenza e fusione dell'automazione con le attività delle moderne macchine computazionali.

4 - Ad Hoc Committee on the Triple Revolution, *The Triple Revolution*, lettera aperta, Santa Barbara (CA) 1964, in [https://scarc.library.oregonstate.edu/coll/pauling/peace/papers/1964p.7-01.html].

5 - *Ibid.*, p. 5.

(2) La seconda trasformazione è quella legata alle armi (*The Weaponry Revolution*[6]). Le nuove tecnologie belliche avevano ormai reso evidente l'impossibilità di vincere una guerra attraverso i mezzi tradizionali: se da un lato esse erano capaci di annientare la civiltà, dall'altro si immaginava che avrebbero reso la guerra stessa uno strumento obsoleto per la risoluzione dei conflitti internazionali. La minaccia costante della distruzione reciproca era accompagnata dalla crescente consapevolezza dell'inutilità della guerra e dalla conseguente necessità di pensare un futuro senza guerre, pur nella difficoltà e nella lentezza del processo che questo obiettivo implicava.

(3) Infine, la terza trasformazione è rappresentata dalla rivoluzione dei diritti umani (*The Human Rights Revolution*[7]). Sempre più evidente appariva la richiesta, ormai universale, di un pieno riconoscimento della dignità e dei diritti fondamentali di ogni individuo. Negli Stati Uniti questa esigenza si manifestava nel movimento per i diritti civili, ma esso non era che un riflesso di una più ampia mobilitazione globale volta alla costruzione di realtà sociali e politiche in cui ciascuno si sentisse valorizzato e nessuno venisse escluso o discriminato per la propria appartenenza razziale.

Riprendendo le riflessioni elaborate dall'*Ad Hoc Committee*, la *First Annual Conference on the Cybercultural Revolution* si svolge in un contesto geopolitico, politico e sociale complesso e in rapido mutamento, segnato dalla Guerra Fredda e dalla minaccia della bomba atomica, dalle politiche dell'amministrazione Johnson e dal programma della *Great Society*[8], dai primi successi dei movimenti per i diritti civili e dalle "lunghe e calde estati" di rivolte urbane nei primi anni Sessanta[9]. I partecipanti sono chiamati a discutere questioni centrali, come la sostituzione del lavoro umano, la gestione della transizione dal lavoro manuale a quello meccanico e le prospettive future di una "società dell'indolenza o dell'ozio" (*society of idleness or leisure*), destinata a essere inaugurata dalle macchine computazionali[10].

---

6 - *Ibidem*.

7 - *Ibidem*.

8 - La *Great Society* prevedeva anche investimenti in cultura, ambiente e sviluppo urbano. Il manifesto *The Triple Revolution* si rivolgeva a Johnson perché vedeva in questo contesto politico un'opportunità per affrontare le sfide poste dall'automazione, dalla tecnologia e dalle trasformazioni sociali. In questo senso, la *Great Society* rappresentava il quadro istituzionale in cui il comitato auspicava riforme innovative e progressive.

9 - Si rammenti che il picco delle rivolte si registrò nell'estate del 1967, nota come la "Long, Hot Summer", quando oltre 150 città furono teatro di violenze, saccheggi e incendi, con i casi più gravi a Detroit e Newark. Tali disordini misero in luce le profonde disuguaglianze sociali ed economiche nelle comunità afroamericane e influenzarono in modo significativo il movimento per i diritti civili.

10 - C. Bassett, *Anti-computing*, cit., p. 106.

Alla conferenza partecipa anche Hannah Arendt, la cui presenza era fortemente desiderata da Hilton, tanto che il pensiero dell'Autrice influenzò l'articolazione dei temi trattati alla conferenza[11]. Il dattiloscritto originale, intitolato *Cybenetics*[12], riporta la trascrizione dei contenuti esposti da Arendt nel suo intervento alla conferenza; esso è oggi conservato presso la *Library of Congress* di Washington insieme a uno scambio epistolare pre- e post-conferenza tra Hilton e l'Autrice[13]. Il saggio *On the Human Condition* (*Sulla condizione umana*)[14] viene così incluso in *The Evolving Society*, curata dalla stessa Hilton, che raccoglieva gli atti della conferenza (comprese trascrizioni di altri interventi, tavole rotonde e contributi dei partecipanti)[15].

Il presente saggio si propone di esaminare il rapporto tra esseri umani e macchine attraverso lo studio di *The Human Condition*, in confronto con il successivo *On the Human Condition* – peraltro considerato alla luce del dattiloscritto *Cybernetics*. Proprio l'analisi del rapporto tra questi ultimi due scritti appare particolarmente rilevante, in quanto – a nostra conoscenza – non risulta ancora affrontata e consente di mettere in evidenza alcune significative differenze testuali. Non solo, la riflessione arendtiana sulla tecnologia costituisce un aspetto non eccessivamente studiato in letteratura, rispetto ad altri temi maggiormente approfonditi. Inoltre, va considerata l'intertestualità di *On the Human Condition*: in esso emergono questioni che si pongono in continuità sia con alcuni temi già trattati in *The Human Condition*, sia con altri che saranno approfonditi in *The Life of the Mind*. Tali comunanze tematiche consentono di interpretare *On the Human Condition* come uno scritto "cerniera", una sorta di ponte che illumina il passaggio teorico dal quadro

---

11 - *Ibid.*, p. 107.

12 - Biblioteca del Congresso di Washington, Hannah Arendt Papers: Speeches and Writings File, 1923-1975; Essays and lectures; *Cybernetics*, lecture, 1964, in [https://www.loc.gov/item/mss1105601187/; d'ora in avanti: CYB].

13 - In una lettera post-conferenza, Hilton sollecita la revisione del testo della conferenza, affinché la raccolta degli atti, intitolata *The Evolving Society*, rispettasse la prevista data di stampa di dicembre 1964, poi posticipata al 1966 (cfr. CYB, img. 1).

14 - Per la prima edizione, pubblicata dall'Autrice, cfr. H. Arendt, *On the Human Condition*, in A.M. Hilton (a cura di), *The Evolving Society. First Annual Conference on the Cybercultural Revolution – Cybernetics and Automation*, Institute for Cybercultural Research, New York 1966, pp. 213-220. Per la seconda edizione, cfr. H. Arendt, *On the Human Condition*, in Id., *Thinking Without a Banister: Essays in Understanding*, 1953-1975, J. Kohn (a cura di), Schocken Books, New York 2018, pp. 322-327 (d'ora in avanti: OHC). Per la traduzione italiana, cfr. H. Arendt, *Sulla condizione umana*, trad. it. E. Pili, in M. Falcone - E. Pili (a cura di), *L'irriducibile umano. Etica e diritto delle intelligenze artificiali*, il Mulino, Bologna 2025, pp. 211-219 (d'ora in avanti: SCU).

15 - Cfr. A.M. Hilton (ed.), *The Evolving Society. First Annual Conference on the Cybercultural Revolution – Cybernetics and Automation*, Institute for Cybercultural Research, New York 1966.

della *vita activa* delineato in *The Human Condition* alla successiva indagine di *The Life of the Mind*, non sempre compresa in continuità con quel percorso.

A tal fine, inizierò lo studio analizzando il tema dell'automazione in *The Human Condition*. Svilupperò poi il percorso spostandomi su *On the Human Condition*, dove l'automazione viene declinata nei termini della *cibernazione*, tenendo anche in considerazione *Cybernetics*, che permette di mettere in evidenza il differente utilizzo dei termini *automazione* e *cibernazione*. Infine, a partire da tale differenza terminologica, concluderò rilevando come i contenuti di *On the Human Condition* possano essere considerati in continuità con alcune tematiche di *The Human Condition* e come, nella necessità di *ripensare il pensare*, si apra – per certi versi – la possibilità di considerare *The Life of the Mind* in continuità con *On the Human Condition*.

2. La fine del lavoro?

Sia nell'opera *The Human Condition* che nel saggio *On the Human Condition*, Hannah Arendt esprime le sue preoccupazioni e incertezze riguardo alla possibile fine del lavoro. Più di sessant'anni dopo, però, possiamo affermare con certezza che non si è assistito alla *scomparsa* del lavoro, ma piuttosto a una sua *trasformazione* profonda, resa possibile in larga misura dall'*automazione*[16]. Nell'opera *The Human Condition*, l'Autrice interpreta l'automazione come l'ultimo stadio della rivoluzione industriale[17], un cambiamento significativo che l'umanità ha affrontato e al quale si è adattata, come ha sempre fatto di fronte alle sfide evolutive. Il fatto che «gli uomini sono esseri condizionati perché ogni cosa con cui vengono in contatto diventa immediatamente una condizione della loro esistenza»[18] implica che «tutto ciò che è in relazione prolungata con la vita dell'uomo assume immediatamente il carattere di una condizione dell'esistenza umana»[19]. La verità è che, quando si vivono determinati cambiamenti in prima persona – non come semplici osservatori, ma come soggetti coinvolti – il loro impatto può risultare improvviso e disorientante, come un fulmine a ciel sereno. Le trasformazioni in atto portano con sé conseguenze che non sono sempre immediatamente evidenti e che richiedono tempo e riflessione per essere comprese pienamente. La velocità con cui questi cambiamenti si manifestano spesso supera la nostra capacità di comprenderne

16 - Per una lettura originale di tale "profonda trasformazione del lavoro", cfr. R.A. Waelen, *Rethinking Automation and the Future of Work with Hannah Arendt*, in «Journal of Business Ethics», (2025), dove l'autrice riflette sull'emergente realtà del "lavoro immateriale".

17 - Cfr. H. Arendt, *The Human Condition*, University of Chicago Press, Chicago 1958, ed. it. *Vita activa. La condizione umana*, Giunti, Milano 2017, pp. 166-167 (d'ora in avanti: THC).

18 - THC, p. 42.

19 - *Ibidem*.

le cause, suscitando timori e inadattabilità. Tuttavia, come sottolinea Arendt nel saggio *On the Human Condition*, «l'essere umano si è sempre adattato alle nuove condizioni più rapidamente di quanto egli si ritenesse capace di fare»[20] e tale prospettiva invita a una riflessione più profonda sulla nostra capacità di adattamento e sulla natura stessa dei cambiamenti in corso.

L'avvento dell'automazione ha profondamente trasformato la natura del lavoro, pur lasciando intatto il "ciclo di vita" – definito in *The Human Condition* come quel ritmo biologico proprio dell'*animal laborans* che ha finito per prevalere nella modernità[21]. L'esito è quello di ritrovarsi «di fronte alla prospettiva di una società di lavoratori senza lavoro, privati cioè della sola attività rimasta loro»[22] – una condizione che, come suggerito nel saggio *On the Human Condition*, impone la necessità di ripensare i precetti morali più antichi, ereditati dalla tradizione giudaico-cristiana[23], poiché si rivelano ormai inadatti a interpretare un presente in cui il lavoro ha cessato di rappresentare il fondamento dell'identità individuale e della coesione sociale. «Forse – suggerisce Arendt – il nostro precetto dovrebbe cambiare solo da "Chi non lavora non mangia" a "Chi non spende energie non riuscirà a dormire"»[24].

3. Il rapporto tra la vita umana e il progresso tecnico-scientifico

Sin dalle prime pagine di *The Human Condition*, Arendt ricorda che la condizione umana è indissolubilmente connessa al pianeta Terra[25]. In effetti, solo l'ambiente naturale terrestre garantisce agli esseri umani un *habitat* nel quale muoversi e respirare in modo spontaneo, senza l'intervento di artifici[26]. Al contrario, il "mondo costruito" – come spiegherà Arendt, opera dell'*homo faber*[27] – tende a separare

20 - SCU, p. 218.

21 - Cfr. THC, pp. 336-341. Infatti, come emerge nell'ultimo capitolo dell'opera, liberato dalle fatiche del lavoro grazie al progresso tecnico-scientifico, l'*animal laborans* non ha realmente affrancato l'umano dalla necessità, ma ha piuttosto esteso il "carattere del consumo" anche agli oggetti d'uso, sottomettendo ogni ambito dell'esistenza al ciclo incessante del bisogno e alla logica biologica della mera sopravvivenza: «rimase solo una "forza naturale", la forza del processo vitale, alla quale tutti gli uomini e tutte le attività umane erano egualmente sottomesse […] e il cui scopo, se mai aveva uno scopo, era la sopravvivenza della specie dell'animale umano» (THC, p. 338).

22 - THC, p. 38.

23 - Cfr. SCU, p. 218. Per quanto concerne tali precetti, cfr. *2Tess* 3,10, e *Gen* 1,28.

24 - SCU, p. 219.

25 - Cfr. THC, p. 34.

26 - Cfr. *Ibidem*.

27 - Cfr. THC, pp. 155-162.

«l'esistenza umana dall'ambiente meramente animale»[28], benché la vita stessa, pur appartenendo ad un ambito estraneo al mondo artificiale, mantenga inalterato il legame con gli altri organismi viventi[29]. Parallelamente, numerosi sforzi scientifici mirano a rendere la vita stessa artificiale, nel tentativo di recidere l'ultimo legame che ancora conferma l'essere umano come figlio della natura[30].

Mentre la scienza compie enormi progressi nella comprensione e nella spiegazione dei fenomeni naturali, grazie a sistemi matematici e strumenti tecnologici sempre più sofisticati, le spiegazioni di queste moderne "verità" scientifiche non possono più essere espresse attraverso il linguaggio e il modo di pensare tradizionali, che finora si sono dimostrati efficaci[31]. Gli esseri umani arriveranno mai a «comprendere, cioè a pensare e a esprimere»[32] fino in fondo, le cose che sono capaci di fare? Forse è proprio questo il rischio: il cervello umano potrebbe davvero divenire «incapace di seguirci in ciò che facciamo, tanto da rendere necessario in futuro il ricorso a macchine artificiali per produrre i nostri pensieri e le nostre parole»[33], una prospettiva che richiama scenari da fantascienza inquietante, in cui l'essere umano rischia di non riuscire più a pensare o a parlare in maniera autonoma, dipendendo sempre più dalle machine per esprimere i propri pensieri e discorsi. Il tutto suona come una *profezia* amara, ma può anche essere letta come una *critica* provocatoria sul rapporto che intercorre tra intelligenza umana e artificiale. Separare la competenza dal pensiero, infatti, comporterebbe la condanna dell'umanità. Come scrive Arendt:

> Se la conoscenza (nel senso moderno di *know-how*, di competenza tecnica) si separasse irreparabilmente dal pensiero, allora diverremmo esseri senza speranza, schiavi non tanto delle nostre macchine quanto della nostra competenza, creature prive di pensiero alla mercé di ogni dispositivo tecnicamente possibile, per quanto micidiale[34].

Tale monito arendtiano acquista una rilevanza particolare nel presente, soprattutto in relazione allo sviluppo delle *intelligenze artificiali*. Mai come ora, la prospettiva di un possibile distacco della "competenza tecnica" dal pensiero solleva interrogativi cruciali. L'interazione tra esseri umani e macchina, un tempo limitata a compiti pratici o ripetitivi, si sta progressivamente estendendo a sfere legate

28 - THC, p. 34.
29 - Cfr. *ibidem*.
30 - Cfr. *ibidem*.
31 - Cfr. THC, p. 35.
32 - THC, p. 36.
33 - *Ibidem*.
34 - *Ibidem*.

all'attività della mente. Tale evidenza viene sottolineata dalla stessa Arendt nel saggio *On the Human Condition*, laddove si afferma: «oggi, le macchine possono sostituire una certa quantità di attività umane che noi abbiamo sempre identificato come attività della mente *umana*»[35].

Più che meri strumenti, le *tecnologie emergenti* assumono il ruolo di mediatori nei processi cognitivi, suggerendo percorsi, soluzioni, immagini o strutture che, pur essendo accessibili alla mente umana, potrebbero non emergere spontaneamente (basti pensare ai moderni *software di scrittura assistita* che suggeriscono contenuti e stili, ai *sistemi generativi* capaci di creare musica o opere visive, o alle *piattaforme educative* che adattano il percorso formativo alle esigenze cognitive del singolo[36]). Di conseguenza, viene a crearsi una *zona grigia* in cui il confine tra ausilio tecnico e attività intellettuale si fa sempre più labile.

4. Le premesse di *The Human Condition*

L'intervento di Hannah Arendt alla *First Annual Conference on the Cybercultural Revolution* viene accolto con grande interesse dai partecipanti. La presenza dell'Autrice alla conferenza richiama l'attenzione su un aspetto non eccessivamente studiato in letteratura: la riflessione arendtiana sulla tecnologia. Se infatti gli studi dedicati ad Arendt tendono a concentrarsi prevalentemente sulle sue analisi politiche, etiche e storiche, non è troppo frequentato l'indagare il modo in cui la filosofa affronta i concetti legati alla tecnologia "in sé"[37].

---

35 - SCU, p. 212.

36 - Cfr. P.L. Pisa, *Come generare testi, immagini e molto altro. In modo naturale*, in F. Ferrara - P.L. Pisa (a cura di), *Intelligenza artificiale*, Gedi News Network, Torino 2025. Inoltre, cfr. R.A. Waelen, *Rethinking Automation and the Future of Work with Hannah Arendt*, cit., che costituisce un contributo teoricamente solido alla riflessione sull'automazione e la crisi del lavoro nel contesto tardo-capitalistico, attraverso una rilettura critica di *The Human Condition*, sviluppata in continuità con le letture più consolidate del pensiero arendtiano.

Le riflessioni di Waelen si inseriscono in un dibattito interdisciplinare ampio, volto a indagare la progressiva perdita di significato, centralità e valore simbolico del lavoro nella contemporaneità tecnologica e postindustriale. L'autrice rielabora le distinzioni arendtiane tra attività lavorativa (*labor*), operare (*work*) e agire (*action*), interrogandosi sulle trasformazioni prodotte dall'automazione avanzata e dalla crescente smaterializzazione dei processi produttivi. In particolare, problematizza la possibilità, oggi sempre più complessa, di "rivalutare" l'*operare* e l'*agire* come attività autenticamente umane, capaci di generare esperienze di senso, autonomia e riconoscimento. Tali interrogativi risultano più che mai urgenti in un contesto in cui le tecnologie emergenti – come ChatGPT, Sora o Gemini – mostrano una crescente capacità di sostituire l'intervento umano anche in ambiti tradizionalmente legati alla creatività.

37 - C. Bassett, *Anti-computing*, cit., p. 108.

Nell'opera *The Human Condition*, i molteplici nuclei tematici affrontati da Hannah Arendt ruotano attorno a un tema centrale di maggiore rilevanza. L'Autrice conduce un'indagine di carattere storico-filosofico, proponendo una dettagliata «analisi di quelle generali capacità umane che nascono dalla condizione umana e che sono permanenti, che cioè non possono andare irreparabilmente perdute finché la stessa condizione umana non sia cambiata»[38]. Attraverso questa prospettiva, Arendt giunge a discutere anche le problematiche che hanno investito la modernità e quelle che da essa ha ereditato la postmodernità. Il principale intento di questa analisi, a detta dell'Autrice, è quello di «ricondurre l'alienazione del mondo moderno – la sua duplice fuga dalla terra all'universo e dal mondo all'io – alle sue origini, e ciò al fine di arrivare a comprendere la natura della società, quale si è sviluppata e mostrata al momento in cui fu sopraffatta dall'avvento di un'era nuova e ancora sconosciuta»[39]. La "fuga dalla terra all'universo" sta a rappresentare il distacco degli esseri umani dal proprio ambiente naturale, reso possibile dal processo tecnico-scientifico che, pur ampliando le possibilità umane, spinge a supere i confini del pianeta senza considerare le conseguenze per la vita terrestre. La "fuga dal mondo all'io", invece, riguarda il ripiegamento del soggetto nella propria interiorità: un processo che porta all'isolamento dalla sfera pubblica e dal confronto con gli altri, privilegiando una dimensione privata e individuale.

Quando, con l'inaugurazione dell'età moderna, si compie quel «rovesciamento del rapporto di contemplazione e fabbricazione»[40] articolato in due momenti – dapprima il capovolgimento del rapporto tra contemplazione e azione, poi quello tra azione e fabbricazione[41] – sembra che *homo faber* sia destinato a uscirne vincitore. Ciò che tuttavia gli impedisce di mantenere la posizione più elevata «nell'ordine gerarchico della *vita activa*»[42] è il ruolo centrale assunto, in età moderna, dal concetto di *processo* («lo spostamento moderno […] dalla cosa stessa al suo processo di fabbricazione»[43]).

Come sottolinea Arendt, «nessuna delle attività della *vita activa* aveva tanto da perdere dall'eliminazione della contemplazione dal campo delle capacità umane significative, come la fabbricazione»[44]. Questo perché, a differenza dell'azione e dell'attività lavorativa, in cui i processi hanno un ruolo primario, «la fabbricazione

38 - THC, p. 39.
39 - *Ibidem*.
40 - THC, p. 321.
41 - THC, p. 323.
42 - *Ibidem*.
43 - *Ibidem*.
44 - THC, p. 324.

dà luogo a dei processi considerandoli semplicemente come mezzi verso un fine, cioè come qualcosa di secondario e derivato»[45].

Il sistema di valori che definiva l'orizzonte di *homo faber* crolla nel momento stesso in cui egli «cessa di definirsi […] *produttore di oggetti* e *costruttore del mondo tecnico umano* che solo accidentalmente inventa arnesi […], e si considera prima di tutto un *produttore di arnesi* […] che solo incidentalmente produce anche delle cose»[46]. Il principio di utilità viene così soppiantato dal principio di soddisfazione. Poiché il principio di utilità richiede un mondo di oggetti per potersi applicare, in un contesto in cui è il processo a occupare il centro – come accade nelle logiche moderne di tipo scientista e produttivistico, che tendono a porre il processo al centro, trasformandolo in fine a sé stesso – l'utilità si trasferisce al processo di produzione. Ne consegue che «ciò che contribuisce a stimolare la produttività e riduce la pena è utile»[47], e dunque la vera utilità si identifica ormai con la soddisfazione, ossia con la riduzione della fatica del lavoro. È questo passaggio che segna la vittoria dell'*animal laborans*, la sua emancipazione e, conseguentemente, il passaggio da una società di lavoratori a una società di consumatori.

5. Uno sguardo all' automazione «da un punto di vista non scientifico»[48]

Sebbene parlare di "automazione" nei termini usati da Arendt possa sembrare anacronistico, le sue riflessioni si rivelano molto attuali. La *cibernazione*[49] «deve essere distinta dalla rivoluzione industriale del passato, la quale consisteva nella

---

45 - *Ibidem*.

46 - THC, p. 325 (corsivo mio).

47 - *Ibidem*.

48 - In CYB, l'Autrice sottolinea la volontà di parlare «da un punto di vista non scientifico [*nonscientific standpoint*]» (CYB, img. 2). Nel saggio revisionato e pubblicato, questo concetto diventa un «punto di vista generale» (SCU, p. 211), definito come «dalla prospettiva di un cittadino medio degli Stati Uniti, non come un membro di qualche ceto specifico della popolazione» (*Ibidem.*). In questo modo, Arendt critica implicitamente la parzialità di alcuni partecipanti alla conferenza. Ad esempio James Boggs (1919-1993), impegnato nella lotta per i diritti della popolazione afroamericana (cfr. J. Boggs, *The Negro and Cybernation*, in A.M. Hilton [ed.], *The Evolving Society*, cit., pp. 167-172), e Betty Friedan (1921-2006), impegnata nella lotta per i diritti delle donne, che partecipò alla conferenza senza tenere un vero e proprio intervento (cfr. A.M. Hilton [ed.], *The Evolving Society*, cit., p. 392). Per una descrizione dei contenuti degli interventi, cfr. E. Pili, *Agli albori dell'IA. Automazione e condizione umana in Hannah Arendt*, in M. Falcone - E. Pili (a cura di), *L'irriducibile umano. Etica e diritto delle intelligenze artificiali*, cit., pp. 160-161.

49 - Come osserva anche Pili, *cibernazione* è il termine con cui con cui viene declinata l'*automazione* nel saggio *Sulla condizione umana*, in linea con l'impostazione generale della *First Annual Conference on the Cybercultural Revolution – Cybernetics and Automation* (*ibid.*, p. 162).

sostituzione della forza muscolare, non quella cerebrale»[50]. Se in passato il processo meccanico si sostituiva al ritmo del corpo – è il caso della "rivoluzione industriale del passato" – oggi il cambiamento coinvolge direttamente le *facoltà cognitive*: le intelligenze artificiali, infatti, non si limitano più a supportare l'agire umano, ma intervengono nei processi mentali stessi, ridefinendo in profondità il rapporto tra gli esseri umani e la tecnica. La *cibernazione* rappresenta l'ennesima avanguardia cui gli esseri umani sono chiamati ad adattarsi. Non si tratta più, infatti, di apparati tecnici modellati sulle esigenze umane; dopo l'avvento dell'automazione – e ancor più con quello della *cibernazione* – sono gli esseri umani che devono rimodellare le proprie esigenze in funzione delle macchine.

In questa prospettiva, Arendt riconosce nel progresso tecnico-scientifico la capacità di realizzare ciò che per secoli era rimasto un sogno irrealizzato: la *liberazione dal lavoro*. Riflettendo in *The Human Condition* sugli sviluppi già visibili nel suo tempo, ella scrive che

> l'avvento dell'automazione, […] in pochi decenni vuoterà probabilmente le fabbriche e libererà il genere umano dal suo più antico e più naturale fardello, il giogo del lavoro e la schiavitù della necessità. […] la ribellione contro di esso e il desiderio di essere liberati dalla "fatica e dall'affanno" del lavoro non sono moderni ma vecchi come la storia che c'è stata tramandata. La libertà dal lavoro in se stessa non è nuova; un tempo era uno dei privilegi più radicati di pochi individui[51].

Questo perché, come ella stessa osserva, «il progresso scientifico e l'evoluzione della tecnica» possono essere funzionalmente «impiegati […] per conseguire ciò che tutte le generazioni passate avevano sognato senza poterlo realizzare»[52].

6. Le tappe dello sviluppo tecnologico

Il progresso tecnico-scientifico si compone di tutti quei miglioramenti che hanno portato, nel corso dei secoli, a trasformazioni profonde nella vita quotidiana e nei processi produttivi. Seguendo la trattazione di *The Human Condition*, è nostro interesse analizzare quelle che sono le tappe principali dello sviluppo tecnologico dagli inizi dell'età moderna fino ad oggi, concentrandoci in particolare sul processo che ha portato alla graduale sostituzione di attrezzi e utensili con macchine sempre più sofisticate e automatizzate[53].

---

50 - SCU, p. 212.

51 - THC, p. 37.

52 - *Ibidem.*

53 - Cfr. THC, p. 165.

Nel ricostruire le tappe che hanno condotto alla moderna automazione, Arendt si confronta con alcune delle figure chiave del dibattito americano degli anni Cinquanta incentrato proprio su questo tema, tra cui spiccano Norbert Wiener (1894-1964), padre della cibernetica, e John Diebold (1926-2005), tra i primi a diffondere e sistematizzare il concetto di automazione[54]. Non è un caso che la definizione proposta da Diebold si ispiri direttamente a Wiener, così come Arendt adotta a sua volta la prospettiva di Diebold[55], lasciando trasparire una certa adesione alla visione secondo cui «l'automazione indica lo sviluppo in senso *autonomo* dei processi meccanici ed elettronici volti alla progressiva sostituzione dello sforzo fisico e (in prospettiva) mentale dell'essere umano nello svolgimento dei suoi lavori»[56].

Come suggerisce Arendt, l'inizio della rivoluzione industriale è segnato dall'avvento della *macchina a vapore*, pur rappresentando ancora una tecnologia che imita la natura, in quanto sfrutta forze naturali per scopi umani. In questa prima tappa, le macchine replicano e potenziano capacità fisiche proprie dell'essere umano[57].

«Lo stadio successivo è principalmente caratterizzato dall'uso dell'elettricità ed è certo l'elettricità che determina la fase attuale dello sviluppo tecnico»[58]. Nel passaggio alla fase dominata dall'*elettricità*, Arendt individua una cesura rispetto alla tradizione dell'*homo faber*, «per il quale ogni strumento era il mezzo per raggiungere un fine prescritto»[59]. Attraverso la sua attività tecnica, l'essere umano non si limita più a utilizzare i materiali così come vengono forniti dalla natura, interrompendone la ciclicità o imitandone i processi, ma inizia a generare da sé e a introdurre nel *mondo umano* – un mondo artificiale durevole, costruito dagli esseri umani e separato dalla natura – le forze naturali, un tempo esterne ad esso. Ne deriva una profonda trasformazione del concetto di *fabbricazione*: dalla produzione artigianale, intesa come una sequenza di azioni separate, si passa a un "processo continuo": la *catena di montaggio*[60].

---

54 - B. Simbirski, *Cybernetic Muse: Hannah Arendt on Automation, 1951–1958*, in «Journal of the History of Ideas», 77, (2016/4), p. 602.

55 - Cfr. THC, pp. 166-169.

56 - E. Pili, *Agli albori dell'IA. Automazione e condizione umana in Hannah Arendt*, cit., p. 154 (corsivo mio). Inoltre, cfr. J. Diebold, *Automation. The Advent of the Automatic Factory*, D. Van Nostrand, New York 1952, ed. it. *Oltre l'automazione. Il mondo industriale e l'attuale espansione tecnologica*, ETAS Kompass, Milano 1966 (d'ora in avanti: AUT). Su Wiener, si veda il recente studio di M. Ferrari, *La cibernetica prima della cibernetica. Filosofia, scienza e tecnica in Norbert Wiener (1914-1943)*, in «Philosophy Kitchen», 18 (2023/1), pp. 37-55.

57 - THC, pp. 165-166.

58 - THC, p. 166. Inoltre, cfr. AUT, p. 69.

59 - THC, p. 166.

60 - *Ibidem*.

Ripercorrendo le principali tappe del processo tecnico-scientifico alla luce dell'opera di Diebold, e condividendo la visione del sociologo francese Georges Friedmann[61], Arendt giunge alla conclusione che dall'analisi di Diebold si può dedurre come l'*automazione* – lo stadio più recente dello sviluppo tecnologico al momento della pubblicazione di *The Human Condition* – sia il risultato della meccanizzazione della catena di montaggio.

Arendt è consapevole della ripresa di Diebold delle previsioni di Norbert Wiener sull'impatto sociale dell'automazione. A differenza di Wiener, però, Diebold ne mise in risalto quasi solamente gli aspetti positivi: insisteva sul fatto che l'automazione avrebbe ridotto la fatica senza generare disoccupazione, emancipando le masse grazie a un maggiore tempo libero[62]. Egli rifiutava quindi le apprensioni di Wiener, secondo cui l'introduzione dei sistemi automatici – seppur più efficienti – avrebbe richiesto una riprogettazione del lavoro attorno ai sistemi stessi, minacciando altrimenti disoccupazione di massa, nuove forme di oppressione e profondi stravolgimenti sociali[63]. Inoltre, a differenza di Wiener, che sottolineava le analogie tra essere umano e macchina, Diebold era convinto che i lavoratori si sarebbero emancipati proprio in virtù della qualità umana del *libero arbitrio*, non presente nelle macchine automatiche. Egli riteneva impossibile attribuire loro una reale capacità di pensiero o qualsiasi forma di antropomorfismo[64].

Arendt, da parte sua, prende in considerazione entrambe le prospettive: da un lato riconosce, come Wiener, l'interdipendenza e la coevoluzione tra condizione umana e sviluppo tecnico-scientifico; dall'altro condivide con Diebold l'idea che le macchine automatiche non possano possedere libero arbitrio, convinzione che sembra influenzare anche la sua concezione dell'azione politica[65]. Pochi anni più tardi, l'avvento della cibernetica smentirà la posizione di Diebold. Oggi, a meno di un secolo di distanza, i sistemi di intelligenza artificiale rimettono in discussione la centralità dell'essere umano e del suo impianto valoriale a fondamento della relazione con le tecnologie emergenti[66].

---

61 - Cfr. G. Friedmann, *Problèmes humains du machinisme industriel*, Gallimard, Parigi 1946, ed. it. *Problemi umani del macchinismo industriale*, Einaudi, Torino 1970.

62 - B. Simbirski, *Cybernetic Muse: Hannah Arendt on Automation, 1951-1958*, cit., p. 603.

63 - *Ibid.*, pp. 603-604. Inoltre, per le previsioni di Norbert Wiener riguardo le conseguenze dell'automazione sui lavoratori, cfr. N. Wiener, *The Human Use of Human Beings: Cybernetics and Society*, Houghton Mifflin, Boston 1950, ed. it. *Introduzione alla cibernetica. L'uso umano degli esseri umani*, Bollati Boringhieri, Torino 1966.

64 - *Ibid.*, p. 603.

65 - *Ibid.*, p. 604.

66 - Cfr. M. Falcone - E. Pili (a cura di), *L'irriducibile umano. Etica e diritto delle intelligenze artificiali*, cit., p. 7.

6.1 Dall'automazione alla cibernazione: un salto qualitativo?

Come emerge dal confronto tra il dattiloscritto *Cybernetics* e la sua versione finale, pubblicata con il titolo *On the Human Condition* (*Sulla condizione umana*), nel dattiloscritto originale non è presente il termine *cybernation* (*cibernazione*), che invece compare nel saggio revisionato e pubblicato. Al suo posto, si trova il termine *automation* (automazione), definito come *a new revolution* (una nuova rivoluzione), mentre nel saggio revisionato e approvato da Arendt è definito *a new phenomenon* (un fenomeno nuovo)[67]. Tale differenza terminologica ha una sua connotazione teorica?

Come accennato (cfr. *infra*, nt. 3), il termine *cybernation* compare per la prima volta nel rapporto sociopolitico *The Silent Conquest*, redatto dal politologo Donald N. Michael su commissione del *Centre for the Study of Democratic Institutions* negli Stati Uniti, e indica l'evoluzione del rapporto tra automazione e computer. La tesi centrale del rapporto è che l'avvento della *cibernazione* costituisce un fenomeno subdolo e silenzioso, che penetra gradualmente nella società. L'avanzamento – quasi inevitabile – del *processo tecnico-scientifico* influisce sull'economia, sul lavoro e sulla vita quotidiana senza che la maggior parte delle persone se ne renda pienamente conto. La *cibernazione*, come determinazione di tale processo, comporta anche questi cambiamenti[68].

Secondo Michael, la *cibernazione* rappresenta un *cambiamento qualitativo della società* che impone un ripensamento della vita politica e democratica[69] – in maniera in parte analoga a quanto osservato da Wiener in *The Human Use of Human Beings* riguardo alla necessità di ripensare le attività umane in relazione ai sistemi automatici[70].

Alcuni dei temi affrontati da Michael in *The Silent Conquest* (come il carattere subdolo della *cibernazione* e le problematiche imminenti che essa avrebbe prodotto) vengono ripresi anche da Hannah Arendt nel saggio *On the Human Condition*. Come Michael, Arendt sottolinea la necessità di distinguere la *cibernazione* dal primo stadio della rivoluzione industriale, che attraverso la «sostituzione della forza muscolare»[71] aveva alleviato lo sforzo lavorativo umano. A differenza delle *macchine pesanti* dell'era industriale, che erano visibili, tangibili e associate a una liberazione dalla fatica fisica, la *cibernazione* è un fenomeno meno evidente e di

---

67 - Cfr. CYB, img. 2, e OHC, p. 322.

68 - Cfr. D.N. Michael, *Cybernation: The Silent Conquest. A Report of the Center for the Study of Democratic Institution*, cit.

69 - Cfr. *ibid*em.

70 - Cfr. N. Wiener, *The Human Use of Human Beings*, cit.

71 - SCU, p. 212.

cui non tutti sono consapevoli[72]. La *cibernazione* allevia lo sforzo mentale, dato che la maggior parte dei compiti di controllo oggi portati a termine dagli esseri umani possono essere facilmente delegati alle macchine («oggi, le macchine possono sostituire una certa quantità di […] attività della mente *umana*. Questo richiede […] una rivalutazione dell'attività mentale»[73]).

6.2 La continuità tra «The Human Condition» e «On the Human Condition»

Il saggio *On the Human Condition* esplora alcuni temi (come il ciclo della vita, la fine del lavoro, la possibilità del tempo libero) già affrontati da Arendt in *The Human Condition*, ma da una prospettiva diversa, rivolta al futuro e, in particolare, ai problemi imminenti che la *cibernazione* avrebbe potuto comportare.

La differenza nell'utilizzo dei termini *automation* e *cybernation*, evidenziata in precedenza, ha un forte significato: nell'opera *The Human Condition* (1958) Arendt aveva trattato ampiamente dell'automazione, ma non della *cibernazione*, forse perché all'epoca non vi era ancora piena consapevolezza del fenomeno. Diversa era la situazione nel 1964, anno della *Conferenza*, quando erano già stati pubblicati il rapporto di Michael *The Silent Conquest* e la lettera aperta dell'*Ad Hoc Committee on the Triple Revolution* al presidente Johnson.

In un recente studio, Pili ritiene che lo scambio del termine *automation* con il termine *cybernation* sarebbe effettuato da Arendt per assicurare una maggiore aderenza ai temi della conferenza, dato che il rapporto tra i due sarebbe come quello rispettivamente tra il genere e la specie[74]. È possibile però aggiungere che la continuità tra i due termini rinvii alla più generale continuità tra l'opera *The Human Condition* e il saggio *On the Human Condition*, garantita sia dalla ripresa di molte tematiche già discusse in *The Human Condition*, sia dal diverso uso dei termini: *automation* nel parlato e *cybernation* nello scritto.

Tanto l'automazione quanto la *cibernazione* (forme di progresso tecnologico) alleviano il lavoro, ma non l'operare in quanto tale. In questo contesto, la figura dell'*artigiano* – emblema dell'*homo faber* arendtiano, che dà forma al mondo attraverso la produzione di oggetti durevoli – ha perso ogni attualità, spazzata via dall'avvento delle fabbriche ben prima che si cominciasse a parlare di automazione. Secondo Arendt, almeno un effetto positivo dell'automazione sarà quello di

72 - C. Bassett, *Anti-computing*, cit., pp. 109-110.

73 - Ibidem.

74 - Cfr. E. Pili, *Agli albori dell'IA. Automazione e condizione umana in Hannah Arendt*, cit., p. 162.

mettere in discussione l'assurdità di tutti gli "umanesimi del lavoro"[75]. Storicamente, l'umanesimo ha posto al centro l'essere umano, la sua libertà e dignità, e tali visioni umanistiche del lavoro miravano a valorizzare l'attività lavorativa come un'esperienza capace di dare significato alla libertà dell'individuo. Se le macchine possono svolgere determinate attività al posto degli esseri umani, facendo venir meno l'idea stessa del lavoro come esperienza umana, allora perde di significato il concetto stesso di "umanesimo del lavoro". Con l'avvento dell'automazione, molte attività lavorative sono state sostituite dalle macchine, riducendo drasticamente il ruolo umano nel processo produttivo e facendo venire meno la *necessità del lavoro* – l'obbligo di dover lavorare per "procurarsi il pane" –, ma non la *possibilità di lavorare*, che rimane aperta per scelta o per passione. Un tempo si imparavano mestieri come quello del falegname o della sarta perché, oltre ad essere redditizi, rispondevano a un bisogno reale: quando un mobile si rompeva o un vestito si logorava, lo si riparava. Con l'avvento della produzione industriale, invece, è diventato più conveniente acquistare un prodotto nuovo piuttosto che riparare quello vecchio, e così queste professioni hanno progressivamente perso importanza. Oggi il falegname e la sarta esistono ancora, ma sono figure rare e spesso alcuni di loro svolgono l'attività non per necessità, bensì per passione o per *hobby*; non a caso, il termine inglese *hobby* viene comunemente tradotto in italiano con *passatempo* (*passa-tempo*): esso rappresenta in maniera concreta la possibilità di lavorare scevra da ogni necessità di lavorare.

Nella visione arendtiana, l'*ultima tappa* del processo tecnico-scientifico è rappresentata dall'*utopia* – non del tutto irreale – secondo cui l'automazione

> certamente rimarrà il punto culminante dello sviluppo moderno, anche se l'età atomica e una tecnologia basata sulle scoperte nucleari le potranno imporre una fine piuttosto rapida […]. Se la tecnologia attuale consiste nell'immettere delle forze naturali nel mondo degli artefatti umani, la tecnologia futura consisterà ancora nell'immettere le forze universali del cosmo che ci circonda nella natura della terra[76].

75 - Si veda, a tal proposito, D. Bell, *Work and Its Discontents*, Beacon Press, Boston 1956, e R.P. Gemelli, *Fattore umano o fattore sociale del lavoro?*, in «Rivista internazionale di scienze sociali», V (1950), pp. 423-441, dove è possibile trovare anche una denuncia molto dettagliata della terribile "illusione" della "gioia" del lavoro.

76 - THC, p. 167.

7. La fine dell'"artificio umano"?

«L'immissione [da parte dell'umano] delle forze naturali nel mondo umano ha mandato in rovina la finalità del mondo, [ovvero] il fatto che gli oggetti sono il fine a cui sono destinati attrezzi e strumenti»[77].

Arendt ricorda che gli oggetti prodotti manualmente nascono da una sequenza ordinata di azioni, eseguite passo dopo passo, e che il processo che li genera (il *processo di fabbricazione*) resta chiaramente distinto dall'esistenza dell'oggetto finito; al contrario, le cose naturali esistono in continuità con il processo che le origina, in quanto sono un tutt'uno con il loro stesso divenire[78]. Se guardati attraverso la lente dei fini umani, che presuppongono un'intenzione e una direzione, tutti i processi biologici appaiono come automatici[79]. Arendt definisce «automatici tutti i processi di movimento spontanei e quindi fuori dalla portata della volontà umana o di interferenze deliberate»[80]. È questo il motivo per cui, in tutte le «forme di produzione contraddistinte dall'automazione, la distinzione tra operazione e prodotto […] non ha più senso ed è già superata»[81].

Come l'Autrice sostiene in *The Human Condition*, la riflessione sul «problema globale della tecnologia, cioè della trasformazione della vita e del mondo mediante l'introduzione della macchina»[82], è stata distolta da un'analisi eccessivamente concentrata sui vantaggi e sugli effetti negativi che la tecnica produce sugli esseri umani[83].

In realtà, come suggerisce Arendt, non si tratta di «vedere se siamo i padroni o gli schiavi delle nostre macchine, ma se le macchine servono ancora il mondo e le sue cose, o se, al contrario, con l'automatismo dei loro processi abbiano comin-

---

77 - *Ibidem.*

78 - THC, pp. 167-168.

79 - THC, p. 168.

80 - *Ibidem.*

81 - *Ibidem.* Inoltre, cfr. AUT, pp. 59-60.

82 - THC, p. 168.

83 - I contributi di Hans Jonas si inseriscono in questo filone teorico che interpreta il carattere strumentale della tecnica in senso prevalentemente antropocentrico. A tal proposito, cfr. H. Jonas, *Das Prinzip Verantwortung*, Suhrkamp, Francoforte sul Meno 1979, ed. it. *Il principio responsabilità. Un'etica per la civiltà tecnologica*, Einaudi, Torino 2009. Inoltre, per quanto riguarda il significato della tecnica in Arendt, è rilevante considerare il dialogo implicito con il pensiero heideggeriano e le influenze che Heidegger ha esercitato sia sulla sua riflessione, sia su quella di Hans Jonas; a tal proposito, cfr. A. Michelis, *La questione della tecnica: evoluzioni di matrici heideggeriane nel pensiero di Hannah Arendt e Hans Jonas*, in «Problemata», 2 (1/2011), pp. 27-51. Inoltre, per l'interpretazione heideggeriana della tecnica, cfr. M. Heidegger, *Die Frage nach der Technik*, Neske, Pfullingen 1954, ed. it. *La questione della tecnica*, in *Saggi e discorsi*, Mursia, Milano 1976.

ciato a dominare e anche a distruggere il mondo e le cose»[84]. Perfino quello che un tempo costituiva «il fine più generale, la liberazione della forza-lavoro umana, che era solitamente attribuita alle macchine, è oggi ritenuto come uno scopo secondario e obsoleto, inadeguato e perfino di ostacolo»[85] al continuo perfezionamento tecnico e all'incremento dell'efficienza delle macchine, che ormai operano in maniera autonoma rispetto ai fini umani, poiché è proprio la liberazione dell'umano dall'attività lavorativa a essere considerata il fattore limitante rispetto alla piena realizzazione di tali sviluppi. Per questo, diviene «privo di senso descrivere questo mondo di machine in termini di mezzi e fini»[86].

D'altra parte, come osserva l'Autrice in un passo del *Prologo*, «l'età moderna ha comportato anche una glorificazione teoretica del lavoro, e di fatto è sfociata in una trasformazione dell'intera società in una società di lavoro»[87]. È proprio questa «*società di lavoratori* quella che sta per essere liberata dalle pastoie del lavoro, ed è una società che non conosce più quelle attività superiori e più significative in nome delle quali tale libertà meriterebbe di essere conquistata»[88].

È in questo contesto che acquista significato la riflessione sul *tempo vuoto* (*vacant time*) portata avanti da Arendt nel saggio *On the Human Condition*. La posizione arendtiana – a differenza di quella assunta da altri partecipanti alla *Conferenza*[89] – assume toni più critici circa gli esiti di una spinta eccessiva verso l'automazione. Il «tempo al di fuori del lavoro»[90], se non impiegato in maniera feconda, assume i caratteri dell'indolenza, e, come sottolinea l'Autrice, «l'indolenza, o il tempo vuoto, è tragica»[91]. Nel suo saggio, Arendt mette in evidenza come «il problema del tempo vuoto, dell'indolenza, ha solo un precedente storico, che risale al declino dell'Impero romano»[92]. Infatti, «la plebe romana aveva tempo vuoto»[93], ma «con il suo tempo vuoto […] non ha portato ad alcuna fioritura della cultura»[94]. Il vero problema da fronteggiare sarà appunto quello del tempo vuoto causato dall'automazione e dalla *cibernazione*, che libereranno la "società di lavoratori" dal lavoro. Ciò nonostante, questi lavoratori non conoscono altro che l'attività la-

84 - THC, pp. 168-169.
85 - THC, p. 169. Inoltre, cfr. AUT, pp. 110 e 157.
86 - THC, p. 169.
87 - THC, p. 37.
88 - *Ibidem*. (corsivo mio).
89 - E. Pili, *Agli albori dell'IA. Automazione e condizione umana in Hannah Arendt*, cit., pp. 160-161.
90 - SCU, p. 216.
91 - *Ibidem.*
92 - SCU, p. 217.
93 - *Ibidem.*
94 - *Ibidem.*

vorativa («ci troviamo di fronte alla prospettiva di una società di lavoratori senza lavoro, privati cioè della sola attività rimasta loro»[95]), ignorando il vero significato – noto anche agli antichi – della liberazione dal lavoro: l'ozio, la *scholé* greca. Nell'antica Grecia, infatti, «avere l'ozio significava astenersi da alcune attività – in questo caso, da quelle lavorative – per essere liberi di svolgerne altre. L'ozio […] è quindi l'antitesi dell'indolenza»[96].

Se da un lato in *The Human Condition* la riflessione arendtiana sul "problema globale della tecnologia" si concentra prevalentemente sulla «trasformazione […] del mondo mediante l'introduzione della macchina»[97], dall'altro, nel saggio *On the Human Condition* l'attenzione sembra quasi spostarsi sulla «trasformazione della vita […] mediante l'introduzione della macchina»[98]. Secondo Bassett, la *cibernazione* potrebbe segnare il superamento del primato dell'attività lavorativa sull'operare, della vittoria dell'*animal laborans* sull'*homo faber*[99].

Nonostante l'immissione dei processi naturali nel mondo umano minacci di trasformarlo radicalmente, se non addirittura di distruggerlo, gli esseri umani sembrano conservare intatta la propria "capacità di condizionamento". Che la ciclicità appartenga alla natura o invada il mondo artificiale, ciò che realmente conta è la possibilità di continuare a soddisfare le "necessità vitali del genere umano". Infatti, come osserva Arendt,

> è molto probabile che la continua immissione dei processi naturali nel mondo umano – anche se potrà distruggere il mondo in quanto tale e gli artefatti umani – riuscirà a soddisfare le necessità vitali del genere umano nello stesso modo continuo e illimitato in cui riuscì la natura prima che gli uomini edificassero la loro dimora artificiale sulla terra e frapponessero una barriera tra loro e la natura[100].

Tant'è vero che «per una società di lavoratori, il mondo delle macchine è diventato un sostituto del mondo reale»[101], sebbene il carattere *ciclico* di tale "mondo

95 - THC, p. 38.

96 - SCU, p. 216.

97 - THC, p. 168.

98 - *Ibidem*. Inoltre, come suggerisce Pili, nella conferenza del 1964 «Arendt ha l'opportunità di tornare a riflettere sui temi affrontati, ma con un approccio maggiormente incarnato e dedicato al pensare la tecnologia in sé, senza considerarla immediatamente nel più ampio contesto della parabola della modernità, come in *The Human Condition*» (E. Pili, *Agli albori dell'IA. Automazione e condizione umana in Hannah Arendt*, cit., p. 160).

99 - Cfr. C. Bassett, *Anti-computing*, cit., p. 108.

100 - THC, pp. 169-170.

101 - THC, p. 170.

tecnico" non assolva «al compito più importante dell'artificio umano»[102]: «offrire ai mortali una dimora più permanente e più stabile di quanto essi stessi non siano»[103]. Non solo, «nel processo operativo, questo mondo di macchine sta anche perdendo quel carattere indipendente che gli attrezzi, gli utensili e i primi macchinari dell'età moderna – interni all'"artificio umano" – possedevano in misura così eminente»[104]; infatti, «i processi naturali da cui quel mondo si alimenta lo assimilano progressivamente al processo biologico»[105], tanto da far assumere alla tecnica i caratteri di «uno sviluppo biologico dell'umanità in cui le strutture innate dell'organismo umano sono trapiantate in misura sempre crescente nell'ambiente circostante»[106].

8. La condizione umana dentro e oltre il processo di tecnicizzazione

Poiché lavoro e consumo costituiscono due fasi di un processo imposto agli esseri umani dalla necessità della vita, dire che viviamo in una società di lavoratori equivale a dire che viviamo in una società di consumatori[107].

Il lavoro è parte di un "ciclo di vita" chiuso, legato alla sopravvivenza biologica e privo di uno scopo ulteriore[108]. In questo contesto, è inutile domandarsi se il lavoro sia un mezzo o un fine: lavorare e consumare non conducono a un "fine superiore", ma servono solo a mantenere attivo lo stesso ciclo, che nel saggio *On the Human Condition* Arendt definisce «ciclo di vita e di lavoro, di sfinimento e di recupero»[109]. È come se si lavorasse solo per avere la forza di lavorare ancora, in un circolo senza fine né trascendenza. Sempre nel saggio *On the Human Condition*, l'Autrice osserva che «nell'Antico Testamento, un uomo doveva lavorare per il suo pane quotidiano, per portarlo a casa alla sua famiglia, e recuperare le forze per essere di nuovo in grado di guadagnarsi il medesimo pane»[110].

Il principale motivo per cui la "meccanizzazione" riesce a inserirsi in maniera così efficace nel mondo del lavoro è che «nulla può essere meccanizzato più agevolmente e meno artificialmente del ritmo del processo lavorativo, che a sua volta

---

102 - *Ibidem.*

103 - *Ibidem.*

104 - *Ibidem.*

105 - *Ibidem.*

106 - Crf. W. Heisenberg, *Das Naturbild der heutigen Physik*, Rowohlt, Hamburg 1955, ed. it. *Natura e fisica moderna*, Milano, Garzanti, Milano 1957.

107 - THC, p. 147.

108 - Cfr. THC, p. 163.

109 - SCU, p. 214.

110 - *Ibidem.*

corrisponde al ritmo ricorrente altrettanto automatico del processo vitale e del suo metabolismo con la natura»[111].

Riprendendo le parole di Karl Marx, Arendt suggerisce che «l'emancipazione dal lavoro […] è emancipazione dalla necessità, che significherebbe, in definitiva, emancipazione dal consumo, dal metabolismo con la natura che è la condizione effettiva della vita umana»[112]. Alla luce degli ultimi sviluppi del processo tecnico-scientifico e delle «possibilità aperte dall'ulteriore incremento dell'automazione»[113], diventa urgente riflettere su come l'emancipazione dal lavoro, un tempo considerata un'*utopia*, possa oggi sembrare una realtà sempre più *concreta*. Inoltre, l'avvento della *cibernazione*, lo sviluppo delle *intelligenze artificiali* e l'emergere delle più recenti innovazioni (non da ultimo i sistemi computazionali quantistici), che superano l'idea della mente umana come sede esclusiva di una *razionalità strumentale*, richiedono un profondo ripensamento del concetto stesso di pensiero. Una necessità che trova risonanza già all'inizio del saggio *On the Human Condition*, dove Arendt afferma con acume che è necessaria una «rivalutazione dell'attività mentale. Dobbiamo domandare: che cos'è l'attività intellettuale in quanto tale?»[114]. Questa è la "sfida" che ci pone Hannah Arendt: *ripensare la condizione e le attività dell'umano* in un periodo di *svolta* come quello presente, nel tentativo di trovare una risposta attuale a un problema di lunga data.

Come osservato, la progressiva automazione libererà sì gli esseri umani dal lavoro – una delle due fasi del processo imposto dalla "necessità della vita" – ma li lascerà esposti al consumo. Infatti, secondo Arendt:

> Il pericolo della futura automazione non è tanto la deplorata meccanizzazione e artificializzazione della vita naturale, quanto il fatto che, nonostante la sua artificialità, ogni produttività umana sarebbe racchiusa in un processo vitale enormemente intensificato, e seguirebbe automaticamente, senza pena o sforzo, il suo sempre ricorrente ciclo naturale. Il ritmo delle macchine intensificherebbe a dismisura il ritmo naturale della vita, ma non modificherebbe, rendendola solo più micidiale, la funzione principale della vita rispetto al mondo, che consiste nel consumare ciò che è durevole[115].

---

111 - THC, p. 164.

112 - THC, p. 150. Inoltre, cfr. S. Weil, *La condizione operaia* [1951], Edizioni di Comunità, Milano 1952, opera in cui Weil mette in luce come la speranza di essere liberati dal lavoro e dalla necessità della vita è l'unico elemento utopistico del marxismo, ma allo stesso tempo è ciò che alimenta i movimenti dei lavoratori che si ispirano ad esso. Questa speranza funge da "oppio del popolo", proprio come la religione per Marx: riesce a consolare e motivare, anche se basata su un'illusione.

113 - THC, p. 150.

114 - SCU, p. 212.

115 - THC, p. 151.

Una volta che le macchine ci avranno liberato dal lavoro, a blindare gli esseri umani nella necessità rimarrà dunque il consumo: «alla fine, solo lo sforzo del consumo sarà l'ultimo elemento rimasto di quella fatica e di quella pena connaturate al ciclo biologico al cui motore è legata la vita umana»[116]. Nulla sfuggirà alla voracità consumistica dell'*animal laborans*.

Una cosa è certa: Arendt, sia nell'opera *The Human Condition*[117] sia nel saggio *On the Human Condition*[118], riconosce la straordinaria capacità degli esseri umani di "adattarsi" alle nuove condizioni. Anche quando l'ambiente e le circostanze cambiano in modo radicale, gli esseri umani hanno sempre dimostrato una sorprendente resilienza. Infatti, come osserva l'Autrice:

> se la condizione umana consiste nell'essere-condizionato dell'uomo, per il quale ogni cosa, data all'uomo o fatta da lui, diventa immediatamente una condizione della sua ulteriore esistenza, l'uomo si "adatta" a un ambiente di macchine nel momento stesso in cui le progetta. Esse sono certamente diventate una condizione tanto inalienabile della nostra esistenza quanto gli attrezzi e gli utensili lo furono per tutte le età precedenti. […] Nessuno ha mai dubitato che l'uomo si adattasse agli strumenti che usava, o avesse bisogno di un particolare adattamento; si può anche concepire che egli si sia adattato all'uso delle sue stesse mani[119].

Tuttavia, nel passaggio da un mondo di attrezzi a un mondo di macchine, proprio questa "capacità di condizionamento" subisce una trasformazione significativa: mentre gli attrezzi «in qualsiasi momento del processo dell'operare rimangono sempre al servizio delle mani»[120] che li guidano, rimanendo sotto il controllo umano, «le macchine esigono che il lavoratore le serva»[121], imponendo i propri ritmi e le proprie logiche. In questo rovesciamento, l'essere umano passa da padrone a servo, costretto ad adattare i ritmi naturali del proprio corpo al movimento meccanico. Arendt in *The Human Condition* lo spiega così: «il processo meccanico ha sostituito il ritmo del corpo umano. Anche l'attrezzo più elaborato rimane un servo, incapace

---

116 - THC, p. 150.

117 - THC, p. 42: «Gli uomini sono esseri condizionati perché ogni cosa con cui vengono in contatto diventa immediatamente una condizione della loro esistenza. […] Tutto ciò che è in relazione prolungata con la vita dell'uomo assume immediatamente il carattere di una condizione dell'esistenza umana».

118 - SCU, p. 218: «L'essere umano si è sempre adattato alle nuove condizioni più di quanto egli si ritenesse capace di fare. Quando l'ambiente si è effettivamente trasformato, noi ne rimaniamo condizionati – anche se non sappiamo di esserlo, anche se possiamo conoscere molto poco delle condizioni che ci hanno plasmato».

119 - THC, p. 165.

120 - *Ibidem.*

121 - *Ibidem.*

di guidare o di sostituire la mano. Anche la macchina più primitiva guida il lavoro del corpo e può eventualmente sostituirlo del tutto»[122].

Come suggerisce l'Autrice, tali «implicazioni pratiche della tecnologia, [ovvero] la sostituzione […] di attrezzi e utensili con i macchinari, [come accade spesso con i processi storici, sembra che] siano venute alla luce solo nel loro ultimo stadio, con l'avvento dell'automazione»[123]. Ed è proprio con l'automazione – il cui processo continuo è assimilabile ai processi naturali – che ad andare perduta non sarà la sopravvivenza biologica degli esseri umani, bensì la permanenza e il significato del *mondo* da loro stessi costruito[124]. È in quest'ottica che Arendt invita a ripensare le condizioni e le attività che rendono possibile una vita *autenticamente umana*, anche in un contesto profondamente trasformato dalla tecnologia.

In ogni caso, va rilevato che Hilton non rimane completamente soddisfatta dell'intervento di Arendt alla *First Annual Conference on the Cybercultural Revolution*[125], ma, d'altra parte, se si guarda agli scambi epistolari post-conferenza, la stessa Arendt ammette di non aver apprezzato particolarmente la conferenza[126]. Con il presente saggio ho però voluto mostrare la continuità tra l'opera *The Human Condition* e il saggio *On the Human Condition*; inoltre, i riferimenti presenti in quest'ultimo riguardo la necessità di ripensare il pensare mettono in luce anche un legame con *The Life of the Mind*, opera dell'Autrice pubblicata postuma[127]. In questo senso, il saggio *On the Human Condition* può essere considerato una vera e propria cerniera tra l'attività nel mondo e l'attività intellettuale.

SAVERIO FERRIERI
Laureato in Filosofia e Scienze e tecniche psicologiche presso l'Università degli Studi di Perugia. Attualmente frequenta il Corso di Laurea magistrale in Filosofia ed etica delle relazioni presso lo stesso ateneo.
saverio.ferrieri@studenti.unipg.it

122 - *Ibidem.*

123 - *Ibidem.*

124 - THC, pp. 169-170.

125 - Cfr. C. Bassett, *Anti-computing*, cit., pp. 106-108.

126 - Cfr. Biblioteca del Congresso di Washington, Hannah Arendt Papers: Correspondence, 1938-1976; Organizations, 1943-1976; *Conference on Cybernetics*, New York, N.Y., 1964-1966, in [https://www.loc.gov/item/mss1105600339/], imgs. 18-19.

127 - Cfr. H. Arendt, *The Life of the Mind*, 2 voll., ed. by M. McCarthy, Harcourt and Brace, New York 1978, ed. it. *La vita della mente*, il Mulino, Bologna 2009.

FORUM

Sulla pertinenza etica della *restorative justice*. Una discussione

di Chiara Gabbolini

This article engages with Serena Meattini's recent book, Etica e Restorative Justice – Definizioni, concetti e problemi. *The text highlights the centrality of harm as a relational category oriented toward individual and community needs, thereby bringing to light accountability as a dynamic process of recognition, responsibility, and commitment to repair. By examining major critiques, including feminist objections, a conception of Restorative Justice emerges as a moral practice aimed at repairing relational fractures before and outside of legal contexts.*

Keywords: giustizia restorativa, etica, accountability, danno, relazione

Sorta nel corso degli anni Settanta sul terreno delle pratiche prima ancora che in quello teorico, la *restorative justice* (anche RJ) continua a costituire un fecondo tema di riflessione interdisciplinare, nonostante le difficoltà definitorie e i numerosi elementi che restano oggetti di dibattito[1]. In linea generale, si tratta di un approccio *incentrato sulle relazioni tra le persone coinvolte nel danno*, nato in risposta alla crisi della giustizia tradizionale di matrice retributivista, che pertanto risulta passibile di considerazione in ambito filosofico-morale e non esclusivamente giuridico. È in questa direzione che si colloca il volume *Etica e restorative justice. Definizioni, concetti e problemi*, che tenta di sottolineare la pertinenza etica di tale approccio alla giustizia, come "cambio di lenti" sul "danno" in senso morale e non solamente legale[2]. Una scelta affine a quella teorica operata dallo European Forum for Restorative Justice (EFRJ) nel quadro della definizione di linee guida e di standards di riferimento[3].

La *restorative justice* si realizza innanzitutto come una rivoluzione sul piano terminologico, causante una rottura in primo luogo semantica con la giustizia tradizionale. Le categorie con le quali leggere il fatto cambiano: come accennato si passa dal reato al danno; non si parla di vittime, ma di persone danneggiate, e neanche di criminali, ma di persone responsabili del danno. Seguendo lo sguardo di Zehr, l'autrice sottolinea che «il sistema giuridico […] si concentra su trasgressori e pene, assicurandosi che i trasgressori ricevano ciò che meritano; la RJ invece su bisogni di vittime, comunità, e autori di reato»[4]. Dunque, se la giustizia tradizionale è incentrata su ciò che le persone coinvolte nel reato si "meritano", la priorità della giustizia restorativa si focalizza sulla riparazione del danno e sull'ascolto dei bisogni che ne derivano. Il danno causa una rottura che implica un obbligo a "mettere le cose a posto", per quanto possibile, nella convinzione che il biasimo rivolto all'autore di reato non sia la soluzione appropriata al danno. Infatti esso, messo in atto a qualsiasi livello, porta all'insorgenza di nuovi bisogni, prima assenti: bisogno di riconoscimento, per chi ha subito il danno; di *accountability*, per chi lo ha commesso; e di sicurezza per la comunità di riferimento. Necessità che, in quest'ottica, si rivelano interconnesse e non contrapposte. Coloro ai quali il danno viene indirizzato necessitano innanzitutto di riconoscimento: riconoscere

1 - Cfr. F. Reggio, *Giustizia dialogica. Luci e ombre della Restorative Justice*, FrancoAngeli, Milano 2010.

2 - Cfr. P. Patrizi, *La giustizia riparativa. Psicologia e diritto per il benessere di persone e comunità*, Carocci, Roma 2024.

3 - European Forum for Restorative Justice, *Manuale sui valori della restorative justice e sugli standard della pratica*, Carocci, Roma 2021.

4 - H. Zehr, *The little book of Restorative Justice*, Good Books, Intercourse 2002, p. 18.

il danno causato ed assumersi la responsabilità e le conseguenze di esso, da parte di chi lo causa, è un primo passo verso la riparazione. Infatti, il riconoscimento dei danni e della natura illecita delle azioni da parte dell'autore

> rappresenta generalmente il nucleo del processo di *accountability* negli incontri RJ, in cui le vittime e i loro sostenitori vogliono vedere che le persone che le hanno danneggiate 'capiscono [*get it*]' e dimostrano un certo grado di comprensione e rimorso per le proprie azioni[5].

Ciò è in grado di apportare numerosi benefici per chi subisce il danno, tra cui «la validazione delle proprie esperienze ed emozioni, la conferma della veridicità dei propri racconti, la riduzione dell'ansia e della paura verso l'autore»[6]. Al contempo, la relazione non si esaurisce in quella duale tra chi causa il danno e chi lo subisce, ma prende in considerazione l'intera comunità di riferimento, al fine di "rimettere le cose a posto nel migliore dei modi possibili". Così dunque la comunità, la collettività intera entrano nell'evento insieme alle parti direttamente coinvolte, ribadendo l'interconnessione tra le persone nell'occasione aperta delle *conferences* e dei *circles*, in quanto «il danno è un evento che coinvolge le storie delle persone, sebbene con ruoli diversi»[7].

In questa prospettiva una precisazione preliminare si rivela fondante: considerando esclusivamente la mediazione come strumento di riparazione, il rischio è di giungere a una sorta di "sineddoche". La pratica della mediazione è richiamata sulla scorta della nota teorizzazione di Bonafé-Schmitt, per cui essa si delinea come «un processo […] con il quale un terzo neutrale tenta, mediante scambi tra le parti, di permettere a queste ultime di confrontare i loro punti di vista e di cercare, con il suo aiuto, una soluzione al conflitto che le oppone»[8]. Il mediatore «ha il ruolo di ricostruire fra le parti lo spazio comunicativo inter-soggettivo e trovare un linguaggio comune che possa condurre al superamento del conflitto»[9]. L'autrice prende le distanze dalla considerazione della mediazione come pratica di base nella RJ poiché essa risulta inapplicabile ad una pluralità di fattispecie connesse alle molteplici declinazioni pragmatiche del danno. Ciò poiché l'essenza

---

5 - K. Daly, *Sexual violence and victims' justice interests*, in M. Keenan (a cura di), *Sexual Violence and Restorative Justice: Legal, Social and Therapeutic Dimensions*, Routledge, London 2017, p. 1408.

6 - S. Meattini, *Etica e Restorative Justice. Definizioni, concetti e problemi*, Mimesis, Milano 2025, p. 140.

7 - *Ibid.*, p. 30.

8 - G. Mannozzi, *Giustizia riparativa* (voce), in *Annali Enciclopedia del Diritto*, Giuffrè, Milano 2017, p. 480. Cfr. J.P. Bonafé-Schmitt, *La mèdiation: une justice douce*, Syros, Paris 1992.

9 - *Ibidem*.

della mediazione risiede nel fatto che essa «colloca in un medesimo piano morale le parti in causa, rispetto alle quali il mediatore assume il ruolo di figura terza, equiprossima, orientata a ristabilire l'equilibrio dialogico»[10], mentre in taluni casi la «mediazione evoca ricomposizione di relazioni non volute, non cercate. [...] Dietro a un'ingiustizia e a un danno non sempre c'è un conflitto da mediare»[11]. Alla luce di questo, la mediazione è un assetto che «in taluni casi sembra impossibile adottare e che, forse, è anche all'origine di alcune profonde resistenze verso la giustizia riparativa *tout court*, anche in questo caso indentificata con la mediazione penale e schiacciata sui suoi limiti»[12]. Il riferimento dell'autrice è da rintracciare in alcune posizioni femministe che muovono delle obiezioni al ricorso della RJ nei casi di reati di genere, nei quali emerge con forza il problema dello squilibrio di potere tra le parti implicate nella violenza. Il cosiddetto *carceral feminism* considera il paradigma retributivista «tra le principali e più auspicabili risposte ai reati di genere e alla necessità di tutela delle donne»[13]. La preferenza per il modello retributivo si fonda sull'idea che il sistema penale, con le sue garanzie procedurali, offra una tutela più adeguata rispetto a pratiche dialogiche proprie della RJ, che rischiano di esporre la vittima a ulteriori forme di pressione o vittimizzazione secondaria. In quest'ottica, la *restorative justice* viene ritenuta, in primo luogo, incapace di eludere il rischio di privatizzazione dell'illecito attraverso la mediazione legale; in secondo luogo, vi è la messa in discussione della possibilità di un reale processo di *accountability* da parte dell'autore di violenza, e, in terzo luogo, la garanzia della tutela della libertà e dell'incolumità psicologico-emotiva delle donne implicate nel processo viene ritenuta dubbia[14]. In merito a ciò viene ipotizzato che «tali critiche possano essere calzanti per la "giustizia riparativa", così come si è configurata in Italia, ma che non siano pienamente pertinenti per la "giustizia restorativa" in senso ampio»[15]. Quest'ultima recupera sul piano terminologico una scelta in materia di traduzione fatta da altri interpreti italiani[16], e su quello contenutistico mutua l'immagine proposta da Johnstone e Van Ness di «una casa a tre piani»

10 - S. Meattini, *Etica e Restorative Justice. Definizioni, concetti e problemi*, cit., p. 96.

11 - Cfr. P. Patrizi, *La giustizia riparativa. Psicologia e diritto per il benessere di persone e comunità*, cit.; cfr. H. Zehr, *The little book of Restorative Justice*, cit.

12 - S. Meattini, *Etica e Restorative Justice. Definizioni, concetti e problemi*, cit., p. 96.

13 - *Ibid.*, p. 101, cfr. E. Bernstein, *The Sexual Politics of the "New Abolitionism*", in "Differences", 18, 2007, pp. 128-151.

14 - Cfr. L. Goodmark, *Restorative justice as feminist practice*, in "The International Journal of Restorative Justice", November 2018, pp. 372-384; pp. 373-374.

15 - S. Meattini, *Etica e Restorative Justice. Definizioni, concetti e problemi*, cit., p. 103.

16 - Cfr. B. Dighera - I. Lizzola, *Comunità territoriali e relazioni restorative*, in P. Patrizi (a cura di), La giustizia riparativa, pp. 189-204.

ciascuno dei quali caratterizzato da una concezione specifica (trasformativa, dell'incontro, riparativa) della giustizia restorativa[17]. Una casa con numerosi punti di accesso e passaggio ai differenti piani, segno che non vi è alcuna esclusione o chiusura tra le differenti concezioni, solo specificità. Le critiche femministe in questo senso sembrano indirizzate a taluni aspetti della giustizia riparativa italiana, che a dispetto del nome si caratterizza come una concezione dell'incontro, ma non alla RJ in senso lato. E dunque, lungi dal costituire un rigetto del paradigma restorativo, tali obiezioni vengono concepite come momenti di problematizzazione interna, che affermano l'esigenza di un ampliamento in senso trasformativo. In linea con una concezione della giustizia orientata non solo alla riparazione del danno, ma alla trasformazione delle condizioni sociali che lo rendono possibile, attraverso processi partecipativi e relazionali capaci di incidere tanto sul piano individuale quanto su quello collettivo. Pertanto, la restorative justice può essere concepita come un "problema" etico poiché viene pensata in quanto pratica che è necessario coltivare in primo luogo come esercizio di responsabilità e riparazione su un piano morale prima di ogni aula di tribunale.

L'ultimo capitolo è dedicato al concetto di *accountability*, centrale nella RJ. Esso viene inteso non solo in quanto «capacità di rispondere delle proprie azioni e delle conseguenze che queste hanno manifestato su altre vite e sul tessuto di relazioni nel quale esse sono inserite»[18] ma anche come «assumersi la responsabilità [*taking responsibility*] del proprio comportamento e agire [*taking action*] per riparare il danno»[19]. Tale definizione si colloca esplicitamente all'interno di un filone teorico che trova ampio consenso all'interno della *restorative justice*, che privilegia una concezione relazionale e normativa della responsabilità rispetto a una descrittiva *tout court*. In questa prospettiva, l'*accountability* si declina come dovere di risposta, come obbligo verso terzi, suscettibile di valutazione morale. Diversa, perciò, tanto dal concetto di *attributability*, che rappresenta «il "volto aretaico" della responsabilità»[20] quanto da quello di *answerability*, poiché delinea la responsabilità in termini di non-rendicontabilità. L'*accountability* si declina come capacità che esonda l'ambito della giustizia restorativa investendo l'ambito morale *tout court*, in quanto «la rendicontabilità si configura come una specifica forma di

17 - Cfr. G. Johnstone - D.W. Van Ness, *The Idea of Restorative Justice*, in Id. (a cura di), *Handbook of Restorative Justice*, Willan Publishing, Cullompton 2007.

18 - F. Reggio, *Giustizia dialogica. Luci e ombre della Restorative Justice*, cit., p. 130.

19 - Office of Juvenile Justice and Delinquency Prevention, *Guide for Implementing the Balanced and Restorative Justice Model*, Washington, DC 1998, p. 9.

20 - S. Meattini, *Etica e Restorative Justice. Definizioni, concetti e problemi*, cit., p.119; cfr. G. Watson, *Two Faces of Responsibility*, in "Philosophical Topics", 24, 1996, pp. 227-248.

relazione sociale»[21], diversa dalla responsabilità legale e non riducibile ad essa. Essa risulta intrinsecamente connessa ad un insieme di pratiche sociali, in quanto porta con sé una «dimensione "processuale", dinamica [...] (una natura) relazionale implicite alla chiamata a "rendere conto", la quale presuppone l'esistenza di più soggetti»[22].

Su questa base teorica si innesta la rielaborazione propria del paradigma restorativo, nella quale, seguendo l'impostazione di Howard Zehr, l'*accountability* non si esaurisce nella sanzione, ma si declina come processo che comprende il riconoscimento del danno, l'assunzione di responsabilità e l'impegno alla riparazione. Ne emerge una concezione della responsabilità che, pur mantenendo una dimensione retrospettiva, si orienta verso il futuro, assumendo una funzione trasformativa delle relazioni. È questo il punto di vista adottato dal Dipartimento di Giustizia canadese, che configura l'*accountability* come «assistere coloro che hanno causato un danno a riconoscerlo e ad assumersi la responsabilità [*take responsibility*] per il danno e per la riparazione»[23]. Se tale concezione manca di chiarezza dal punto di vista semantico, poiché utilizza in maniera interscambiabile *accountability* e *responsibility*, essa si rivela di cruciale importanza in una prospettiva costruttiva in ambito morale, in quanto «quest'ultima (la responsabilità personale) non è solamente quella "retrospettiva", nel riferimento all'assunzione di responsabilità attraverso il riconoscimento del proprio ruolo nel danno compiuto e a cui può conseguire biasimo, ma anche "in prospettiva", nell'impegno insito alla responsabilità di riparare o, più ampiamente, "restorare" le persone e le relazioni danneggiate»[24].

L'origine del danno non è da identificare solamente per infliggere una punizione, identificando dunque la responsabilità come correlato della sanzione come vorrebbe l'ottica retributiva, ma per offrire l'opportunità di «rimettere le cose a posto»[25] come imperativo relazionale restorativo. Un duplice livello, quindi, articolato tanto sul piano del riconoscimento del danno compiuto quanto su quello dell'impegno a mettere le cose a posto.

---

21 - *Ibid*, p. 118.

22 - *Ibidem*.

23 - Cfr. UNODC, *Handbook on Restorative Justice Programmes*, p. 6; cfr. Department of Justice Canada, *Principles and Guidelines for Restorative Justice Practice in Criminal Matters*, Ottawa 2018.

24 - S. Meattini, *Etica e Restorative Justice. Definizioni, concetti e problemi*, cit., p. 124; cfr. M. De Caro - S.F. Magni - M.S. Vaccarezza, *Le sfide dell'etica*, Mondadori, Milano 2021, pp. 34-35; cfr. S. Bonicalzi - M. De Caro, *In medio stat virtus. The Justification of Punishment between Deontology and Utilitarianism*, in "B@belonline", 10, 2023, pp. 15-24.

25 - *Ibid.*, p. 125.

Con Zehr è possibile affermare che

> l'autore deve essere "held accountable per il suo comportamento" […] in una triplice direzione: della comprensione del danno arrecato, inteso come lesione a persone e relazioni, dell'assunzione di responsabilità a partire dal riconoscimento del proprio ruolo in esso e della partecipazione proattiva alla riparazione[26].

Al centro, dunque, rimane l'alterità: assumersi la responsabilità si configura in primo luogo come impegno relazionale nei confronti dell'altro che dal danno viene annullato, per ricomporre i pezzi partendo dal riconoscimento del proprio ruolo in esso. L'alterità non si esaurisce nella vittima ma coinvolge, come anticipato, la comunità intera. A questo proposito, Zehr sottolinea che i livelli di obbligazione sono tre: vittima, comunità e autore.

> Gli autori devono essere ritenuti rendicontabili [held accountable], ma anche la società. La società deve essere rendicontabile [accountable] nei confronti delle vittime, aiutandole a identificare e soddisfare i loro bisogni. Allo stesso modo, l'intera comunità deve prestare attenzione ai bisogni degli autori, cercando non solo di punire ma di trasformare. La rendicontabilità [accountability] è multidimensionale e trasformativa[27].

Pavelka, invece, considera la *restorative justice* come una necessità di riparare la trama delle relazioni e rispondere ai bisogni di tutte le parti coinvolte in maniera costruttiva per la comunità intera piuttosto che distruttiva per l'autore del danno:

> Nel modello di giustizia penale, l'accountability si raggiunge quando il trasgressore viene punito. Nella giustizia restorativa, l'accountability si raggiunge quando il trasgressore accetta la responsabilità [responsibility] e intraprende azioni per riparare il danno. La responsabilità per il crimine è sia individuale che sociale: le vittime sono centrali nel processo di risoluzione del crimine, i trasgressori sono chiamati a fare riparazioni, e c'è particolare enfasi sulla creazione e il mantenimento del dialogo tra vittima, trasgressore e comunità[28].

L'autrice, inoltre, include nella sua riflessione l'analisi di Kathleen Daly sulle diverse forme di responsabilità[29]. Ella distingue i concetti di *vindication* e *accountability*, attribuendo alla prima la funzione di riaffermazione della giustizia attraverso la pena – intesa dunque come responsabilità passiva imposta

---

26 - *Ibid*, p. 130; cfr. H. Zehr, *Changing Lenses*, cit., p. 46.

27 - H. Zehr, *Changing Lenses*, cit., p.202.

28 - S. Pavelka - D. Thomas, *The Evolution of Balanced and Restorative Justice*, in "Juvenile and Family Court Journal", 70, 2019, pp. 37-48, pp. 39-40.

29 - K. Daly, *Sexual violence and victims' justice interests*, cit., p. 118.

dall'autorità – e riservando alla seconda una dimensione distinta e non riducibile alla sanzione. Essa viene infatti intesa come il processo mediante il quale l'autore del danno assume su di sé la responsabilità delle azioni commesse in maniera attiva, mettendo in atto comportamenti finalizzati alla riparazione (quali il riconoscimento e l'ammissione del torto, l'espressione di rimorso, le scuse, la partecipazione a programmi di contrasto alle dipendenze o a lavori di pubblica utilità). Tale analisi però non rimane esente da critiche e da una problematizzazione: il presupposto spesso implicito secondo cui l'autore del danno possa essere considerato un agente morale pienamente capace di rispondere delle proprie azioni viene messo in discussione. La teoria restorativa così intesa può risultare di difficile applicazioni in contesti minorili o in situazioni di disagi psico-sociali[30]. Alla luce di tale criticità, emerge la necessità di un lavoro sull'approccio restorativo che includa una riflessione più ampia sulle implicazioni ad esso collegate, volgendo lo sguardo verso quelle problematicità che diventano motivo di trasformazioni, revisioni e ampliamenti all'interno del paradigma.

In conclusione, in questa sede il tema viene considerato primariamente all'interno dell'ambito morale, a partire dal quale, solamente, è possibile promuovere una "trasformazione".

> Il perseguimento della giustizia non può che passare per la ricerca di strumenti per trasformare l'esistente, per cambiare quelle condizioni iniziali cosicché il danno non si verifichi più. Non si tratta in questo caso di trovare un finale alternativo, ma di riscrivere l'intera trama[31].

Altrimenti, per una cultura sostanzialmente retributivista risulta estremamente complesso accogliere qualsiasi tentativo di cambiamento di prospettiva sulla giustizia in sede penale[32]. È necessario perciò mutare le lenti morali sul danno *tout court*, non soltanto in ambito penale, e con la partecipazione di tutte le persone coinvolte, per poter accogliere la rivoluzione della *restorative justice*.

CHIARA GABBOLINI
Laureata in Filosofia ed etica delle relazioni presso l'Università degli Studi di Perugia
chiara.gabbolini@studenti.unipg.it

30 - S. Meattini, *Etica e Restorative Justice. Definizioni, concetti e problemi*, cit., p. 137; cfr. W. R. Wood, *The incident demands of accountability: trauma, marginalisation, and moral agency in youth restorative conferencing*, in "The International Journal of Restorative Justice", 2 (2020), pp. 178-183.

31 - G. Palomba, *La trama alternativa*, Minimum fax, Roma 2023, p. 204.

32 - Cfr. La complessa ricezione della Disciplina organica della giustizia riparativa (d.lgs. n. 150/2022).



Sapienza di Dio che scende dal Cielo
e scienze dell'uomo che salgono dalla Terra
diventano "uno" al centro:
la mente di Gesù (1 Cor 2,16)
a tutti partecipata nel dono dello Spirito Santo


Istituto Universitario
SOPHIA
University Institute