

APhEx 17, 2018 (ed. Vera Tripodi)
Ricevuto il: 30/12/2017
Accettato il: 10/04/2018
Redattore: Paolo Labinaz & Francesca Ervas

APhEx
PORTALE ITALIANO DI FILOSOFIA ANALITICA
GIORNALE DI **FILOSOFIA**
NETWORK
N° 17, 2018

T E M I

La teoria dell'informazione integrata

Niccolò Negro

Il presente scritto mira a fornire un'introduzione critica a una delle principali teorie della coscienza nel panorama filosofico e neuroscientifico contemporaneo, ovvero la teoria dell'informazione integrata (IIT). Dopo una breve esposizione dei pregi della teoria, si analizzerà in dettaglio la sua formulazione e le sue principali implicazioni. In seguito, ci si concentrerà sulla relazione di IIT sia con il "problema difficile della coscienza" (Chalmers 1995), sia con certe forme di panpsichismo che sembrano essere suggerite dalla teoria stessa. In conclusione, saranno delineati criticamente problemi, obiezioni e controversie che IIT deve affrontare, offrendo così una panoramica delle future sfide, tanto filosofiche quanto neuroscientifiche, che le si pongono dinanzi.

INDICE

1. INTRODUZIONE
2. UN'ALTRA TEORIA NEUROSCIENTIFICA DELLA COSCIENZA?
3. ASSIOMI, POSTULATI E IDENTITÀ
4. RILEVANZA E PROBLEMI FILOSOFICI DI IIT

1. Introduzione

La teoria dell'informazione integrata (IIT, dall'acronimo inglese per *Integrated Information Theory*) è una recente proposta che mira a fornire una spiegazione neuroscientifica della coscienza. IIT è stata sviluppata dal neuroscienziato italiano Giulio Tononi in una serie di articoli pubblicati dalla metà del primo decennio degli anni duemila (Tononi 2004, 2005, 2008 e Balduzzi e Tononi 2008). Da allora, la teoria ha subito importanti variazioni, pur mantenendo intatte tanto le intuizioni di fondo che la caratterizzano, quanto gli aspetti epistemologici e metafisici che la sottendono (si vedano Tononi 2012 e Oizumi, Albantakis e Tononi 2014). La presente esposizione introduttiva si basa, dunque, sulla formulazione più recente, presentata nell'articolo "From The Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0" (Oizumi *et al.* 2014), e trova la sua ragion d'essere nel crescente peso che IIT sta assumendo nell'attuale dibattito neuroscientifico e filosofico sulla coscienza e sulla sua relazione con il mondo fisico. Sintomo della crescente popolarità della teoria può essere rintracciato nell'entusiasmo e nel sostegno dimostrato da importanti neuroscienziati, quali, ad esempio, Christoph Koch (2012) e Marcello Massimini¹, nonché da David Chalmers e da altri filosofi che ne hanno apprezzato soprattutto le sue (possibili) implicazioni panpsichiste (si veda, ad esempio, Hassel-Mørch 2017).

Intenzione primaria della presente esposizione è fornire un'introduzione critica e ragionata della teoria, analizzandone i principali aspetti filosofici. Nella prima sezione, dunque, si definirà lo scopo e le fondamenta epistemologiche di IIT, argomentando come sia proprio questa base a caratterizzare la forza della teoria e la sua principale differenza rispetto ad altre proposte neuroscientifiche circa il problema della coscienza. Nella seconda sezione, invece, l'attenzione si sposterà sui dettagli teorici di IIT, seguendo gli assiomi e postulati che la informano. Particolare enfasi verrà conseguentemente posta sull'identità fondamentale che caratterizza la

¹ Si veda Massimini e Tononi (2013).

teoria, cioè l'identità fra coscienza e informazione integrata. In seguito, si passerà a esaminare la rilevanza filosofica di IIT, nella misura in cui essa sembra suggerire, da un lato, una metafisica vicina a una certa forma di pansichismo, dall'altro, una problematica discontinuità concettuale fra coscienza ed azione che la avvicina all'epifenomenismo. Infine, saranno esposti i problemi, tanto filosofici quanto neuroscientifici, che la proposta in questione deve affrontare, sottolineando le controverse assunzioni che ne stanno alla base.

2. Un'altra teoria neuroscientifica della coscienza?

Nel corso degli ultimi decenni si è assistito ad un sempre crescente interesse per lo studio scientifico della coscienza, tanto che, ad oggi, può essere considerato un problema genuinamente scientifico, strappato definitivamente al campo filosofico-religioso che ha da sempre costituito il terreno naturale del dibattito sul tema.

Tale naturalizzazione del fenomeno-coscienza, oggi condivisa non solo da filosofi di estrazione analitica, ma anche da esponenti vicini alla tradizione fenomenologica (si vedano, ad esempio, Thompson 2007 e Varela 1996), si fonda sulla concezione che una spiegazione della coscienza può essere fornita dagli strumenti offerti dalle scienze naturali tradizionali, come fisica, chimica e biologia. Volendo individuare un momento chiave di questo percorso, si può considerare la teoria proposta da Francis Crick e Christoph Koch (1990) come il definitivo riconoscimento del problema della coscienza all'interno della ricerca neuroscientifica.

Altre proposte che hanno recentemente avuto buon seguito nell'ambito delle neuroscienze della coscienza sono la "Teoria dello spazio neuronale di lavoro" (*Global Neuronal Workspace Theory*) di Stanislas Dehaene, Michel Kerszberg e Jean-Pierre Changeaux (1998) e la Teoria del darwinismo neurale di Gerald Edelman (1987). Senza addentrarsi in questa sede nei dettagli di tali teorie, vale la pena evidenziare come esse condividano il metodo di base, che consiste nel rintracciare un determinato meccanismo neurale che sia correlato alla presenza di una certa esperienza cosciente. La questione, dunque, riguarda la possibilità di individuare un "ingrediente" a livello neurale capace di colmare lo scarto tra mondo fisico e mondo dell'esperienza (fenomenico). Per Crick e Koch questo ingrediente è la sincronizzazione delle onde a 40 Hz; Dehaene e Changeaux lo individuano invece nell'innesco (*ignition*) di certi pattern di attività neurale che vengono mobilitati all'interno di uno spazio accessibile ai sistemi esecutivi; per

Edelman, infine, l'esperienza fenomenica deriva dall'attività di circuiti rientranti (capaci di trasmettere informazione in modo feed-forward e feed-back) all'interno del sistema talamo-corticale, dopo una selezione dei pattern neurali rilevanti che è affine a quella attuata dalla selezione naturale.

Tutte queste concezioni condividono l'idea che la coscienza sia sostanzialmente un fenomeno biologico e che possa essere naturalizzata grazie ad una sempre maggiore conoscenza dei meccanismi cerebrali. L'approccio di fondo assegna perciò una priorità alle caratteristiche del sistema nervoso centrale.

Diversamente, l'approccio di IIT considera prioritaria la fenomenologia (ed è dunque definito *phenomenology first*): in altri termini, IIT non parte dal chiedersi quali siano le caratteristiche che certi meccanismi a livello cerebrale debbano condividere per poter generare l'esperienza conscia; all'opposto, la teoria parte dal domandarsi quali siano le caratteristiche di tale esperienza, per poi chiedersi, in subordine, quali siano le proprietà che un sostrato fisico deve manifestare per generare quelle determinate caratteristiche.

Inoltre, IIT pone l'accento su un particolare strumento che caratterizza le scienze naturali in quanto tali, e cioè la possibilità di *misurazione* dei fenomeni studiati. Si noti che quest'ultimo punto può apparire in contrasto con la natura *phenomenology first* appena menzionata: come combinare la priorità assegnata alla prospettiva in prima persona con il bisogno di una descrizione *in terza persona* della coscienza? È importante notare che tale frizione è solo apparente. La priorità in questione è, infatti, di tipo epistemologico, e non tanto logico-ontologica. Essa dipende dal fatto che l'esperienza soggettiva è esattamente ciò che una teoria della coscienza deve *spiegare*, e, dunque, ciò da cui è necessario partire. Questa concezione non implica in alcun modo un'impossibilità, per i metodi in terza persona, di dar conto dell'esperienza in prima persona, ma, al contrario, richiede che una seria indagine scientifica individui un metodo di descrizione oggettivamente fondato anche nell'ambito della ricerca sulla coscienza. Sin dalla rivoluzione scientifica del XVII secolo, infatti, le scienze naturali si sono distinte per la loro capacità di descrivere numericamente i fenomeni, e, così facendo, hanno trovato un metodo in grado di inquadrare i fenomeni analizzati entro una cornice capace di escludere la soggettività dell'osservatore, favorendo la possibilità di una validazione oggettiva (o, quanto meno, intersoggettiva) delle teorie proposte. Richiamandosi a questo fondamentale aspetto delle scienze naturali, IIT si propone di trattare *quantitativamente* la coscienza, definendo un metodo di misurazione capace di descrivere il fenomeno anche laddove l'attribuzione intersoggettiva possa

fallire. Si pensi, ad esempio, al caso studiato dal team guidato da Adrian Owen, in cui una ragazza considerata in stato vegetativo dimostrò di essere cosciente degli stimoli esterni grazie a risposte rintracciate attraverso l'attivazione neurale registrata dalla risonanza magnetica funzionale (fMRI) (si veda Owen, Coleman, Boly, Davis, Laureys e Pickard 2006). Misurare oggettivamente la coscienza significherebbe, infatti, svincolare il concetto di coscienza da quello di guida razionale dell'azione, sviluppando, almeno in linea teorica, la possibilità di fondare un utile "metro" per la coscienza applicabile in situazioni critiche e delicate come quelle in cui si trova chiunque debba definire lo stato interno di soggetti incapaci di agire o di muoversi. Tale "metro" è stato effettivamente testato grazie al lavoro del team guidato da Marcello Massimini, che ha ideato un indice di misurazione della *capacità* che un cervello ha, in potenza, di generare coscienza (Massimini, Boly, Casali, Rosanova e Tononi 2009). È interessante notare che questo indice (*Perturbational Complexity Index*) è teoreticamente fondato su IIT, nella misura in cui esso mira a considerare quanto una certa stimolazione cerebrale diffonde il proprio segnale in modo altamente differenziato e integrato.

IIT, dunque, si propone non tanto come ulteriore proposta nell'ambito della ricerca dei cosiddetti "correlati neurali della coscienza", quanto più come una proposta teorica in grado di spiegare in termini quantitativi il fenomeno-coscienza, inquadrando la mole di dati osservativi a disposizione entro una cornice esplicativa e predittiva:

Clearly, only a theory – one that says what consciousness is and how it can be generated – can hope to offer a combination of explanatory, inferential, and predictive power starting from a few basic principles, and provide a way to quantify both the level of consciousness and its content. Integrated Information Theory (IIT) is an attempt to characterize consciousness mathematically both in quantity and in quality (Oizumi *et al.* 2014, 1).

Indubabilmente, il fatto che IIT abbia offerto la base teorica per lo sviluppo di un indice dall'evidente valore pratico e funzionale come PCI testimonia del suo potere tanto esplicativo quanto predittivo.

Prima di addentrarsi nella disamina degli assiomi e dei postulati che informano IIT, è bene spendere due parole circa i concetti di *livello* e *contenuto* di coscienza, poiché la teoria afferma esplicitamente di spiegare scientificamente tali aspetti del fenomeno. Per contenuto cosciente s'intende ciò che è presente nell'esperienza conscia di un soggetto ad un dato momento, mentre il livello di coscienza indica quanto un soggetto è in grado di partecipare di una certa esperienza, indipendentemente dal suo

contenuto². IIT, allora, propone di descrivere in termini numerico-matematici sia il contenuto che il livello dell'esperienza cosciente, identificandoli rispettivamente con la sua qualità e quantità.

Nel quadro di quell'approccio *phenomenology first* menzionato in precedenza, la teoria è formulata attraverso cinque assiomi riguardanti le caratteristiche della nostra fenomenologia e cinque postulati concernenti le proprietà che un sostrato fisico deve possedere per generare le caratteristiche elencate negli assiomi. Nella prossima sezione, perciò, si procederà a esporre tale formulazione in assiomi e postulati, per poi chiarire l'identità fondamentale fra coscienza e informazione integrata. Da subito, però, è bene evidenziare quali siano i punti di forza di IIT che rendono la teoria degna di essere seriamente presa in considerazione come teoria scientifica della coscienza:

- IIT parte dal chiedersi quali siano le caratteristiche della nostra fenomenologia, e come possano essere assiomatizzate. Tale posizione riconosce dunque la nostra fenomenologia come vero *explanandum* di una spiegazione scientifica della coscienza.
- IIT propone una naturalizzazione del fenomeno coscienza attraverso la caratterizzazione numerico-matematica del fenomeno studiato, inquadrando la mole di dati empirici messi a disposizione dalla ricerca sui correlati neurali della coscienza entro una cornice teorica esplicativa e predittiva.
- IIT vuole spiegare scientificamente tanto la generazione di contenuti coscienti quanto lo sviluppo di diversi livelli di coscienza.
- IIT ha reso possibile un "metro" per individuare la coscienza in soggetti incapaci di agire e/o muoversi, attraverso lo sviluppo del PCI.

Nella prossima sezione questa forza relativa tanto agli aspetti epistemologici quanto a quelli riguardanti l'applicabilità pratica della teoria sarà analizzata dettagliatamente, facendo riferimento ai postulati, assiomi e relazioni d'identità da essa previsti.

² Per un interessante sguardo critico circa la nozione di livello di coscienza, si veda Bayne, Hohwy e Owen (2016).

3. Assiomi, Postulati e Identità

Come già suggerito nella sezione precedente, IIT si presenta lungo una serie di assiomi che riguardano le caratteristiche della nostra esperienza fenomenica, e postula successivamente le corrispondenti caratteristiche che un certo sostrato fisico deve avere per generare tale esperienza.

Il primo assioma è quello dell'esistenza: la coscienza esiste come aspetto del mondo naturale in cui viviamo, e la sua esistenza è data dal punto di vista intrinseco del soggetto. In altre parole, il fatto che la mia esperienza fenomenica esiste non necessita un osservatore esterno. Il corrispondente postulato afferma, dunque, che anche il sostrato fisico della coscienza dovrà esistere intrinsecamente, riconoscendosi come punto di vista con potere causa-effetto:

The first axiom of IIT states that experience exists intrinsically. As recognized by Descartes, my own experience is the only thing whose existence is immediately and absolutely evident, and it exists for myself, from my own intrinsic perspective. The corresponding postulate states that the PSC must also exist intrinsically. For something to exist in a physical sense, it must have cause-effect power — that is, it must be possible to make a difference to it (that is, change its state) and it must be able to make a difference to something. Moreover, the PSC must exist intrinsically — that is, it must have cause-effect power for itself, from its own intrinsic perspective (Tononi *et al.* 2016, 1).

Il secondo assioma asserisce che la nostra esperienza è strutturata, o composita. In altre parole, una certa scena cosciente si compone di diversi aspetti fenomenici, come può essere il vedere il verde di una copertina di un libro, l'ascoltare un certo suono, il distinguere diverse forme geometriche nello spazio. Analogamente, il postulato della composizione prevede che il sostrato fisico della coscienza sia dato da meccanismi con potere causa-effetto presi in combinazione.

Il terzo assioma riguarda uno dei due pilastri fondamentali che caratterizzano la teoria sin dalla sua prima formulazione, e cioè l'informazione: ogni esperienza cosciente è estremamente informativa nella misura in cui essa esclude un numero praticamente infinito di altre possibili scene coscienti. Ad esempio, richiamandosi a un esempio classico, un fotodiode capace di distinguere solo due possibili stati di cose (luce vs buio), nel momento in cui è in uno stato di attivazione, sta escludendo solo una possibile alternativa, trasmettendo 1 bit di informazione ($\log_2 2 = 1$) (si veda Tononi 2003). Un essere umano, diversamente, nel momento in cui distingue uno stato di luce da uno di buio, è capace di distinguere una luce blu da una rossa, gialla, verde, etc. È bene sottolineare che dal punto di vista

fisico l'informazione non è definita in senso shannoniano, come stringa di simboli capace di diminuire lo stato di entropia di un certo sistema, ma, piuttosto, come differenza che fa la differenza³. A livello di sostrato fisico, infatti, la teoria postula che un meccanismo in uno stato specifica un determinato repertorio causa-effetto: un certo stato costringe una determinata distribuzione probabilistica tanto sui diversi stati possibili *passati* del sistema, quanto sui possibili stati *futuri* del sistema.

Il quarto assioma verte invece sull'integrazione che caratterizza ogni esperienza cosciente. In altri termini, una scena cosciente è sempre unitaria: «the content of an experience (information) is integrated within a unitary consciousness» (Tononi *et al.* 2016, 3). Questo assioma mira ad indicare che una certa scena cosciente non è riducibile alle proprie componenti, così come non è possibile esperire separatamente il campo visivo di destra da quello di sinistra. Analogamente, a livello di sostrato fisico la teoria postula che l'informazione trasmessa da un meccanismo in uno stato è irriducibile; o meglio, il potere causa-effetto sviluppato dal meccanismo che sottende alla coscienza non può essere ridotto al potere causa-effetto dei propri componenti.

L'ultimo assioma è, infine, quello dell'esclusione, e sta a indicare come la nostra esperienza cosciente sia sempre una *ed una sola*, data una certa grana temporale, che la teoria individua in un *range* temporale che va da poche decine a poche centinaia di millisecondi⁴. Tradotto a livello di sostrato fisico, ciò significa che solo il meccanismo che genera la maggior quantità di potere causa-effetto irriducibile contribuisce al sostrato fisico della coscienza.

Ora, diventa fondamentale evidenziare il ruolo che questi assiomi e postulati hanno nell'identificare l'esperienza cosciente con il concetto di informazione integrata (denotata attraverso la lettera greca Φ). Per fare questo, è bene prendere in considerazione le parole dello studio che esprime la più recente formulazione di IIT:

according to IIT, there is an identity between phenomenological properties of experience and informational/causal properties of physical systems [...]. The central identity is the following: The maximally irreducible conceptual structure (MICS) generated by a complex of elements is identical to its experience (Oizumi *et al.* 2014, 3).

³ L'espressione è mutuata da Bateson (1972).

⁴ Vedi Holcombe (2009).

Dunque, una volta definito un *complex*, e cioè un sistema fisico che soddisfa i cinque postulati appena esposti, è possibile, secondo la teoria, costruire uno spazio multidimensionale in cui rappresentare geometricamente la struttura informazionale caratterizzata dallo stato di attivazione del complesso. Tale spazio sarà costituito da tanti assi quanti sono i possibili stati passati e futuri del sistema fisico in questione, mentre le distribuzioni probabilistiche di ciascuno stato fungeranno da “coordinate cartesiane” necessarie per rappresentare la struttura informazionale specificata dal sistema. La teoria, perciò, non identifica l’esperienza con un certo stato fisico, o, più in generale, con una certa attività cerebrale, ma, piuttosto, con la forma di una struttura informazionale in tale spazio multidimensionale (che la teoria, riprendendo un termine dal dibattito sulla coscienza in ambito analitico, definisce *qualia space*). La particolare forma geometrica di tale struttura sarebbe la *qualità* della coscienza, e cioè cosa è presente a un certo punto di vista, mentre il valore d’irriducibilità specificato dall’integrazione d’informazione fornirebbe la *quantità* della coscienza, e cioè quanto un certo concetto è presente all’esperienza:

the quantity or level of consciousness corresponds to the Φ^{Max} value of the quale; the quality or content of the experience corresponds to the particular constellation of concepts that constitutes the quale – a particular shape in qualia space (Oizumi *et al.* 2014, 15).

Nelle seguenti sezioni s’insisterà sulla rilevanza filosofica della teoria e sulle assunzioni che la sottendono. Prima di procedere, però, è bene evidenziare alcune conseguenze implicate dalla posizione qui introdotta.

Prima di tutto, si noti che la coscienza, secondo IIT, è una proprietà *intrinseca* di un sistema fisico in uno stato di attivazione. In altre parole, essa emerge perché un certo meccanismo fisico è in grado di riconoscersi come potere causa-effetto irriducibile, senza alcun bisogno di un osservatore esterno. Così, la struttura informazionale irriducibile rappresentata nel qualia space è tanto *observer-independent* quanto lo è l’esperienza cosciente⁵. L’idea che le proprietà fenomenologiche siano sostanzialmente proprietà intrinseche di un sistema fisico arrangiato in un modo tale da soddisfare i postulati di IIT porta inoltre a predire che anche un fotodiode, se opportunamente progettato, può essere cosciente, sebbene in misura minimale. IIT esplora, infatti, tale possibilità: si pensi a un fotodiode composto da un *detector* D e un predittore P, interconnessi in modo tale che se D si attiva, causa la conseguente attivazione di P. Successivamente, P

⁵ Per una critica circa questo punto, si veda Searle (2013).

fungerà da “memoria”, aumentando la sensitività di D alla luce e modificandone, dunque, le possibilità di attivazione. Secondo IIT, tale meccanismo è minimamente cosciente:

Simple as it is, the photodiode system satisfies the postulates of IIT: both of its elements specify selective causes and effects within the system (each element about the other one), their cause-effect repertoires are maximally irreducible, and the conceptual structure specified by the two elements is also maximally irreducible. Consequently, the system DP=11 forms a complex that gives rise to a MICS, albeit one having just two concepts and Φ^{Max} value of 1. DP is therefore conscious, albeit minimally so (Oizumi *et al.* 2014, 19).

Una siffatta esperienza cosciente, ovviamente, sarebbe limitata al fatto di essere sempre “così piuttosto che così”, ma ciò non toglie che anche una minima quantità di coscienza non sia legata a un certo aspetto fenomenico. Ancora una volta, è possibile considerare la possibilità di progettare un fotodiode conscio solo se si considera la coscienza come proprietà intrinseca di un sistema fisico. Nella sezione seguente, ci si soffermerà sulla rilevanza filosofica di tale posizione, e sulla sua vicinanza a certe forme di pansichismo.

Ora, un'altra importante conseguenza di IIT è che anche meccanismi silenti possono essere parte del sostrato fisico che genera l'esperienza cosciente. Infatti, anche elementi inattivi contribuiscono a selezionare i possibili stati passati e futuri del sistema, implementandone, dunque, la struttura informazionale da esso specificata. Tale idea è ben distante da forme di teorie della coscienza affini al Global Workspace suggerito da Bernard Baars (1998) prima, e sviluppato da Dehaene e Changeaux poi⁶. Secondo tali punti di vista, infatti, l'esperienza cosciente emerge laddove una competizione di processi cerebrali di computazione dell'informazione avvengono in parallelo, e solo alcuni riescono a segnalare, o trasmettere, l'informazione convogliata da tale attività ai networks fronto-parietali. Al contrario, «in IIT, information is not in the message that is broadcasted by an element, but in the shape of the MICS that is specified by a complex» (Oizumi *et al.* 2014, 18).

⁶ Si noti che anche Dennett (1991, 2001) ha sviluppato una teoria che può essere considerata affine alla *Global Workspace Theory* di Baars. La celebre espressione «consciousness as fame in the brain» (Dennett 1996a) può interpretarsi, infatti, lungo quella linea che vede la coscienza come fenomeno emergente a livello cerebrale a seconda dei processi sub-personali che competono tra loro; e che, infine, hanno la capacità di mostrarsi, ed essere visti, dagli altri.

Infine, un importante risultato di IIT è che la coscienza non è associata ad alcuna funzione, o scopo, particolare. Di fatto, la teoria afferma che è plausibile considerare la coscienza come proprietà interamente solipsistica, del tutto slegata dall'ambiente che la circonda (si veda Tononi 2008, 239). Si pensi, ancora, al caso dei pazienti locked-in, o in stato di coscienza minimo (MCS), laddove la ricchezza dell'esperienza percepita non può essere espressa attraverso l'azione, o, più in generale, il comportamento. Inoltre, la teoria prevede la possibilità di costruire dei sistemi interamente feed-forward che abbiano la capacità di eseguire le stesse funzioni di sistemi integrati che generano una certa quantità di Φ . Prima di tutto, è bene sottolineare qui un dettaglio neuroscientifico: anche IIT, in accordo con altre proposte (si vedano Edelman 1989 e Hofstadter 2007), considera che l'idea di "rientro", o connessione "feed-back", sia un elemento essenziale alla coscienza, in quanto starebbe alla base della possibilità di generare un punto di vista intrinseco del sistema, consentendo una trasmissione reciproca e circolare dell'informazione. Al contrario, un sistema puramente feed-forward trasmetterebbe l'informazione solo a livelli successivi, fino ad arrivare a sistemi esecutivi capaci di trasformare un certo input in output. Tali sistemi, dal punto di vista del comportamento esterno, sarebbero essenzialmente indistinguibili da meccanismi consci, caratterizzati da connessioni rientranti:

IIT claims [...] that input-output behavior is not always a reliable guide: one needs to investigate not just "what" functions are being performed by a system, but also "how" they are performed within the system. Thus, IIT admits the possibility of true zombies, which may behave more and more like us while lacking subjective experience (Oizumi *et al.* 2014, 21).

Questo passaggio ci costringe a prendere in considerazione la relazione della teoria con uno dei soggetti filosofici che ha popolato i dibattiti sulla coscienza negli ultimi vent'anni: lo *zombie* filosofico. Considerata l'importanza di tale concetto e la sua rilevanza per una teoria neuroscientifica della coscienza, è bene riservare una sezione di questa introduzione alle conseguenze filosofiche implicate da IIT.

4. Rilevanza e problemi filosofici di IIT

Dopo la pubblicazione di *The Conscious Mind* (1996) da parte del filosofo australiano David J. Chalmers, il dibattito filosofico e neuroscientifico sulla coscienza è stato scosso da una fortunata nozione che mira alle basi di ogni

spiegazione neuroscientifica della coscienza: quella di zombie filosofico. Come suggerito nella precedente sezione, lo *zombie* è un'entità esattamente uguale a un essere umano vivente dal punto di vista fisico, e indistinguibile dal punto di vista comportamentale. Se punto con uno spillo, lo *zombie* griderà dal dolore; se posto di fronte a un pericoloso serpente, lo *zombie* si dimostrerà atterrito. Ma, pur avendo un cervello e un corpo identico al nostro, lo *zombie* non prova né dolore né paura: qualsiasi computazione dell'informazione che avviene dentro di lui, anche se si manifesta tramite outputs esterni, non avviene all'interno di una sfera di coscienza, ma completamente "al buio".

Ora, è bene evidenziare che Chalmers non nega che la coscienza sia effettivamente causata da un certo substrato fisico. Ciò che l'esperimento mentale dello *zombie* vuole dimostrare, piuttosto, è che il mentale non è in alcun modo *implicato necessariamente* dal fisico, essendo possibile, dal punto di vista logico, pensare senza contraddizione a delle proprietà fisiche non congiunte a proprietà fenomenologiche. Dunque, il fisico non implica *necessariamente* il mentale. Si noti che questo è esattamente ciò che una spiegazione scientifica vuole dimostrare, nel momento in cui individuare una relazione d'identità fra un certo stato fisico e un certo stato mentale vuol dire implicare che quando c'è uno, allora c'è anche l'altro.

Non è questa la sede per disaminare l'enorme dibattito scatenato da siffatta linea di pensiero (conosciuta come *conceivability argument*), e nemmeno per analizzarla nei dettagli semantici e metafisici necessari per capire il passaggio che rende possibile passare dalla concepibilità logica alla possibilità metafisica. Ciò che è necessario qui considerare è la rilevanza che questa argomentazione assume in relazione a una teoria come IIT, che postula una relazione di identità fra struttura informazionale irriducibile ed esperienza fenomenica. Chalmers, infatti, argomenta che una spiegazione scientifica tradizionalmente intesa prevede di ridurre un certo fenomeno alle funzioni e al potere causale di determinati meccanismi fisici. Ad esempio, è possibile spiegare la trasmissione dell'informazione genetica attraverso il funzionamento del DNA, o la vita attraverso i meccanismi che regolano la riproduzione, l'equilibrio e le funzioni di un organismo in un ambiente. Grazie alla forza del *conceivability argument*, dunque, Chalmers (1995) sostiene che una spiegazione di questo tipo non può essere sufficiente a spiegare la coscienza fenomenica, visto che, anche una volta individuati i processi che integrano informazione, memoria, attenzione ed esecuzione dei compiti motori (tutti quei problemi che l'autore etichetta come *easy problems of consciousness*), rimane ancora la domanda del *perché* tali processi avvengano entro una sfera esperienziale dotata di una certa qualità,

invece che “al buio”, così come accade per gli *zombies* filosofici. Tale domanda circa il perché questa esperienza conscia accompagni i nostri processi mentali è invece considerata il vero problema difficile della coscienza (HPC, dall'inglese *hard problem of consciousness*). Secondo Chalmers, ogni tentativo di spiegare questo *perché* attraverso una visione fiscalista che riduca il fenomeno ai meccanismi funzionali e alla struttura causale di un fenomeno fisico di più basso livello è destinato a cadere in quel «gap esplicativo» (Levine 1983) derivante dal *knowledge argument* presente in pensatori come Thomas Nagel (1974) e Frank Jackson (1986)⁷.

Riesce dunque IIT a superare quest'obiezione? La teoria, dopotutto, non traccia un'identità fra stato fisico (ad esempio, una certa attivazione di networks cerebrali) e stato mentale, ma, piuttosto, come visto in precedenza, postula un'identità fra scena cosciente e struttura informazionale irriducibile in un *qualia space*. Lo stesso Chalmers ammette che IIT può essere in grado di risolvere un problema della coscienza che egli stesso definisce “pretty hard”, e cioè il problema di distinguere quando una certa entità è cosciente o meno⁸. Ad ogni modo, tale possibile passo in avanti nello studio scientifico della coscienza non sembra bastare a spiegare *perché* la nostra vita mentale sia associata all'esperienza conscia. Come scrive Garrett Mindt, prendendo le mosse da una posizione vicina a quella di Chalmers,

If we are to understand function as ‘causal roles in the production of a system’s behaviour’, as Chalmers suggests, then I take it that ‘intrinsic cause–effect structures of certain mechanisms in a state’ (as explained by IIT) satisfy the relevant causal role in producing the behaviour of a system, as it is the intrinsic cause–effect structures that constrain the possible states mechanisms within which a system can instantiate, i.e. neurons in the brain. [...] If information according to IIT is about how a mechanism through its ‘cause–effect power’ and ‘space of possibilities’ is nothing over and above structure and function, then IIT is committed to being a physicalist account of consciousness. If this is so, IIT succumbs to the same explanatory gap argument against physicalism. (Mindt 2017, 149).

⁷ L'idea di base del *knowledge argument* è che per quanto si possa sapere tutto della descrizione fisica di un evento cosciente (ad esempio, come funziona il vedere rosso), tale descrizione non esaurisce tutti gli aspetti del fenomeno che la fisica crede spiegato (l'esperienza in sé del vedere rosso sarebbe qualcosa di nuovo anche per un neuroscienziato che sapesse tutto del meccanismo fisico della visione dei colori, ma che non avesse mai visto il rosso).

⁸ Si veda la discussione nata dal post pubblicato nel blog di Scott Aaronson nel Maggio del 2014: <https://www.scottaaronson.com/blog/?p=1799>.

Il punto, perciò, sta nel capire se la nozione di informazione presente in IIT sia coerente con il fisicalismo classico o meno. Nel caso essa fosse riducibile alla struttura causale e al comportamento funzionale dei meccanismi fisici di base di un sistema, come sembra suggerire Mindt, l'identità fra struttura informazionale irriducibile ed esperienza cosciente non sarebbe essenzialmente diversa dalle teorie fisicaliste tradizionali che tracciano un'identità fra stato mentale e stato fisico (si vedano Place 1956 e Smart 1959). Nella sua incapacità di superare il sopracitato *explanatory gap*, dunque, la teoria perderebbe gran parte del suo potere esplicativo, e rimarrebbe distante dallo spiegare il problema difficile della coscienza.

D'altra parte, la teoria sostiene di superare il problema degli *zombies* attraverso la possibilità di progettare dei sistemi non integrati funzionalmente identici a sistemi coscienti, come già evidenziato in conclusione della precedente sezione. Il punto è che la coscienza non sarebbe essenziale all'espletamento delle funzioni che vengono comunemente associate all'esperienza fenomenica. In uno studio successivo alla formulazione 3.0 di IIT, infatti, si dimostra come dei semplici networks adattivi (*animats*) riescano a incrementare il livello d'informazione integrata una volta posti di fronte a compiti che implicano una complessità sempre maggiore (Albantakis, Hintze, Koch, Adami e Tononi 2014). Il risultato cruciale suggerisce che a un certo livello di complessità del compito, gli *animats* sviluppano spontaneamente un'interconnessione integrata, che, secondo IIT, genera una certa quantità di Φ . Si noti, però, che tale generazione non è in alcun modo *necessaria*, ma solo *preferibile*, come gli autori dello studio sottolineano: «Note also, however, that any task could, in principle, be solved by a modular brain with $\Phi=0$ given an arbitrary number of elements and time-steps» (Albantakis *et al.* 2014, 13). Il problema per IIT risiede nel fatto che essa afferma che «a complex of a physical mechanisms in a certain state is necessarily associated with its quale» (Albantakis *et al.* 2014, 13). Com'è possibile tenere assieme questa associazione di necessità fra aspetto fisico ed aspetto fenomenico, anche se considerata *a posteriori*, con l'idea che sia l'evoluzione per selezione naturale a rendere preferibile lo sviluppo di un'interconnessione a livello di sostrato fisico capace di generare una certa esperienza cosciente? Non sembra consigliabile interpretare tale associazione come fenomeno contingente derivante da quel processo contingente che è l'evoluzione per selezione naturale? Il principale problema per la teoria sembra stare nel fatto che la coscienza non è tanto interpretata unicamente come fenomeno biologico emergente entro un determinato contesto comprendente la relazione fra un organismo vivente e il suo ambiente d'azione; piuttosto, essa è considerata una proprietà

intrinseca di un certo stato della materia. Di conseguenza, IIT afferma che la coscienza è una proprietà fondamentale dell'universo fisico, presente laddove la materia che costituisce l'universo, così come conosciuto dalla fisica, sia disposta in modo tale da soddisfare i postulati della teoria. Individuare la coscienza come proprietà intrinseca della materia, dunque, da un lato porta la teoria a postulare una problematica identità fra struttura informazionale ed esperienza fenomenica; dall'altro, apre la teoria a delle posizioni che possono essere definite come panpsichiste.

Prima di considerare il problema del panpsichismo, è bene chiarire un punto fondamentale riguardante la relazione di IIT con l'hard problem: spiegare quale sia la differenza fra due sistemi funzionalmente identici, uno cosciente e l'altro *zombie*, non basta a spiegare perché certe funzioni siano associate a determinate aspetti qualitativi, così come non basta calare la teoria entro il contesto della selezione naturale per poter superare HPC. Spiegare perché quel particolare tratto si sia evoluto non spiega direttamente perché tale tratto avesse certe qualità in prima istanza. Questo punto sembra essere particolarmente rilevante perché IIT, a differenza di altre teorie che negano esplicitamente HPC sin dalla sua formulazione, considerandolo fondato su false assunzioni⁹, è pensata per affrontare e superare tale questione. Non è un caso che IIT, nel momento in cui condivide le stesse assunzioni che stanno alla base della formulazione dell'hard problem, condivide anche un certo entusiasmo per posizioni affini al panpsichismo che lo stesso Chalmers ha abbracciato in alcuni suoi lavori¹⁰. Queste assunzioni possono rintracciarsi nell'idea che la coscienza sia una proprietà di un certo stato fisico in quanto tale, e che sia possibile concepire, per dirla con Evan Thompson (2011, 15), «a physical *living body* (Körper) that is not a *lived body* (Leib)». È proprio quest'idea, che mira a interpretare la coscienza come una proprietà intrinseca di un certo sistema fisico in uno stato, che porta Tononi e Koch ad affermare che essa è caratteristica fondamentale dell'universo, esattamente come la massa di un corpo:

An analogy is mass, which can be defined by how it curves space-time around it – except that in the case of experience the entities having the property are not elementary particles but complexes of elements, and experience comes not in two but in a trillion of varieties (Tononi e Koch 2015, 11).

Diversamente dal panpsichismo tradizionale, ad ogni modo, IIT non attribuisce una mente indiscriminatamente a tutte le entità presenti

⁹ Dennett (1996b), O'Regan (2016), e Thompson (2011).

¹⁰ Chalmers (2015), (2016).

nell'universo, e, tramite l'individuazione dei cinque postulati che un sistema fisico deve soddisfare per generare la coscienza, offre una possibile soluzione al cosiddetto "combination problem": come possono delle microesperienze dare vita a macroesperienze? Diversamente da alcune forme di panpsichismo, o di protopanpsichismo¹¹, IIT non attribuisce forme di coscienza agli elementi fondamentali della materia, ma solo a *complessi* di entità fisiche tradizionalmente intese, i quali generano una certa quantità di informazione integrata. Sembrerebbe dunque possibile interpretare IIT coerentemente con il fisicalismo tradizionale, secondo cui la realtà fisica deve e può essere descritta interamente dalla fisica, senza il bisogno dell'intromissione di scomode entità mentali al livello fondamentale. D'altra parte, però, alcuni autori hanno evidenziato come la fisica tradizionale non si esprima tanto su *cos'è* la materia, quanto invece su *come funziona* (si veda Goff 2017). Altri, come Galen Strawson (2006), ritengono che quel tipo di fisicalismo che non rinunci a classificare la coscienza come un fenomeno realmente esistente nel nostro universo implichi una certa forma di panpsichismo che interpreti la materia come essenzialmente e fondamentalmente disposta per la coscienza.

Una forma di fisicalismo come quello di IIT potrebbe essere interpretato coerentemente con questa proposta. Inoltre, l'idea di quantificare l'esperienza cosciente così come proposta da IIT permette di trattare la coscienza come proprietà graduata, piuttosto che come fenomeno tutto-o-nulla: la coscienza sarebbe così distribuita in diversi gradi attraverso tutto il regno animale, così come si presenterebbe in diverse forme entro la vita stessa di un organismo vivente: «in us it becomes richer as we grow from a baby to an adult whose brain has fully matured and becomes more functionally specialized» (Tononi e Koch 2015, 13). La problematicità di questa proposta sta nel definire i limiti entro cui è possibile individuare l'esperienza cosciente: c'è davvero coscienza in un fotodiodo integrato così come postulato da IIT? È dunque possibile, almeno in linea teorica, costruire macchine coscienti?

Diventa necessario notare, a questo punto, un ulteriore punto controverso implicato da IIT: la totale separazione dei concetti di vita e coscienza. Tale separazione è evidente nell'esempio del fotodiodo conscio, ed è confermata nell'idea che l'esperienza è una proprietà intrinseca di un qualsiasi stato fisico, indipendentemente dal contesto in cui esso è istanziato. Come nota Thompson (2011, 8), l'idea per cui la coscienza non abbia nulla a che fare con le funzioni vitali del soggetto cosciente è

¹¹ Per un'esauriente discussione sul tema, si veda Chalmers (2015).

l'assunzione che sta alla base dell'esperimento mentale dello *zombie*, e porta con sé la possibilità che il fenomeno-coscienza non abbia alcuna funzione nel guidare il nostro comportamento. Come si è visto in precedenza, IIT condivide queste assunzioni, assumendo su di sé tutte le problematicità e le controversie di natura filosofica che le accompagnano. Non si vuole, in questa sede, sostenere una posizione piuttosto di un'altra, ma solo evidenziare come le assunzioni filosofiche che stanno alla base di IIT non siano del tutto stabilmente condivise all'interno della comunità filosofica e scientifica.

Infine, è bene notare che scindere il concetto di coscienza dal contesto biologico in cui si presenta, al fine di generalizzarlo e considerarlo come proprietà intrinseca di un sistema fisico, vuol dire scindere anche il concetto di coscienza da quello di azione. Come si è già visto, IIT non definisce alcuna particolare funzione per l'esperienza cosciente, e lascia aperta la possibilità che essa non abbia alcun ruolo nel guidare le nostre azioni. Tale concezione può essere vista come apertura all'epifenomenismo, quella dottrina per cui l'esperienza cosciente sia generata da determinati processi fisici, ma che non abbia alcun potere causale entro il dominio fisico¹². IIT può obiettare che la teoria non prevede alcun collegamento *necessario* fra coscienza e azione, pur non negando che tale collegamento sia possibile attraverso una spiegazione evoluzionistica. In altri termini, le funzioni che compiamo coscientemente non sarebbero necessariamente connesse con l'aspetto fenomenico che assume la nostra esperienza, nondimeno, le compiamo coscientemente perché così ha "voluto" l'evoluzione per selezione naturale. Ancora una volta, però, diventa un problema, per IIT, rendere coerente tale spiegazione con l'idea che la coscienza non sia un fenomeno biologico, concettualmente slegato da quello di vita, ma una proprietà fondamentale dell'universo fisico.

5. Conclusioni

Nel corso di questa esposizione, si è più volte rimarcato che il principale punto di forza di IIT risiede nel fatto che si è dimostrata capace di caratterizzare teoricamente un indice in grado di distinguere quando un soggetto è cosciente o meno, indipendentemente dal dato comportamentale. Da un lato, infatti, l'idea di descrivere numericamente il fenomeno-coscienza sembra essere un passaggio fondamentale in quel processo di

¹² Per una recente difesa di questa posizione, si veda Wegner (2002).

naturalizzazione della coscienza che le neuroscienze stanno attuando negli ultimi decenni; dall'altro, la forza pratica dimostrata da IIT sembra essere fondamentale, poiché ogni teoria scientifica è, in conclusione, testata nel mondo attraverso la sua applicabilità.

Anche per questo motivo, è bene considerare un problema di natura matematica, ma che si riflette nella ricerca neuroscientifica che mira a misurare la coscienza. Tale problema riguarda la capacità di trovare una procedura che sia *computabile*. IIT, infatti, prevede, che Φ sia calcolata attraverso l'individuazione della partizione che porta la minor perdita d'informazione possibile. Così, diventa necessario bi-partizionare il sistema lungo ogni possibile partizione, così da trovare in che modo gli elementi del sistema sono effettivamente connessi tra loro. La nozione di *effective connectivity* garantisce che la relazione fra due elementi (ad esempio, due neuroni) sia effettivamente causale e non solo correlazionale, ma porta con sé il problema di complessità computazionale qui descritto:

the evaluation of Φ requires the calculation of effective information across each bipartition of a system, and there is a factorial growth in the number of partitions that must be examined as the size of the network increases (Seth *et al.* 2006, 10801).

Inoltre, un metodo di misurazione della coscienza deve affrontare il problema che pone l'aspetto dinamico e interazionale della coscienza. Per quanto IIT si sia posta questa questione, e abbia cercato di darvi risposta attraverso la nozione di *matching* (cioè, quanto la struttura causale-informazionale generata dal complesso si distanzi dalla struttura causale dell'ambiente esterno), sembra che questo strumento, più che catturare un aspetto essenziale del fenomeno-coscienza, serva invece ad accomodare la teoria all'unico contesto in cui, ad oggi, la coscienza è presente, e cioè quello biologico. Ipotesi di possibili misurazioni dell'aspetto dinamico della coscienza sono state prese in considerazione dalla ricerca neuroscientifica negli ultimi anni, ed è verisimile pensare che diversi metodi di misurazione possano essere integrati per catturare diversi aspetti della coscienza (si vedano Seth *et al.* 2006 e Rudrauf *et al.* 2017) .

IIT, ad ogni modo, non vuole proporsi solo come teoria di un aspetto della coscienza, ma, piuttosto, come teoria comprensiva della coscienza, che sia capace di caratterizzare cosa effettivamente è la coscienza in quanto tale. Si può obiettare, però, o sostenendo che tale progetto non è auspicabile per una scienza della coscienza, oppure affermando che, anche nel caso l'idea fosse possibile da realizzare, nondimeno IIT fallirebbe nel suo intento. Alla prima classe appartiene l'idea che

Given that consciousness is a rich biological phenomenon, a satisfactory neural theory of consciousness must avoid reductionistic excess. Excessive reductionism can be revealed by improper reification, for example by converting a dynamic process into a static entity. [...] any theory that identifies consciousness with a single measure is likely to be excessively reductionistic and, as a result, limited in its scope (Seth *et al.* 2006, 10803).

Alla seconda appartiene invece Michael Cerullo (2015, 10), il quale critica IIT sostenendo come la teoria non sia una teoria della coscienza, ma, piuttosto, che «IIT is a theory of partial-panexperientialism that, even if correct, does not help us to understand or predict the kind of consciousness that is relevant to our subjective experience».

Senza entrare qui nei dettagli del dibattito, sembra evidente che ogni teoria neuroscientifica della coscienza che proponga un metodo di misurazione dell'esperienza debba porsi il problema di capire se e come tale descrizione riesca a catturare il fenomeno-coscienza in tutta la sua complessità, sia relativa alla sua generazione cerebrale, sia nel suo aspetto dinamico-interazionale.

Dal punto di vista filosofico, è ragionevole pensare che le future sfide riguardanti il problema di trattare scientificamente la coscienza saranno dirette ad analizzare la *possibilità* che ciascun metodo di misurazione garantisce al fine di spiegare l'esperienza fenomenica. Inoltre, sembra necessario esaminare le conseguenze metafisiche implicate dalle più diverse teorie che vogliono misurare la coscienza come proprietà intrinseca di un sistema fisico. È possibile che parecchio lavoro filosofico dovrà essere fatto per delineare una teoria della coscienza che parta da presupposti epistemologici vicini a quelli di IIT ma che eviti, da un lato, pesanti carichi metafisici come il panpsichismo, dall'altro, proposte estremamente distanti dalla nostra fenomenologia e dal nostro mondo dell'esperienza come l'epifenomenismo. D'altronde, è bene rimarcare che la nostra fenomenologia, con le sue caratteristiche, è esattamente ciò che una teoria della coscienza si ritiene debba spiegare.

Bibliografia

- Albantakis L., Hintze A., Koch C., Adami C., Tononi G., 2014, «Evolution of Integrated Causal Structures in Animats Exposed to Environments of Increasing Complexity», *PLoS Comput. Biol.* 10, 12.
- Baars B, 1988, *A Cognitive Theory of Consciousness*, New York, Cambridge University Press.

- Balduzzi D., Tononi G., 2008, «Integrate Information in Discrete Dynamical Systems: Motivation and Theoretical Framework», *PLoS Comput. Biol.*, 4, 6.
- Bateson G., 1972, *Steps to an Ecology of Mind*, Chicago, Chicago University Press.
- Bayne T., Hohwy J., Owen A., 2016, «Are There Levels of Consciousness?», *Trends in Cognitive Sciences*, 20, 6, pp. 405-413.
- Cerullo M.A., 2015, «Problems with Phi: A Critique of Integrated Information Theory», *PLoS Comput. Biol.*, 11, 9.
- Chalmers D.J., 1995, «Facing Up to the Problem of Consciousness», *Journal of Consciousness Studies*, 2, 3, pp. 200-219.
- Chalmers D.J., 1996, *The Conscious Mind: in Search of a Fundamental Theory*, New York, Oxford University Press.
- Chalmers, D.J., 2015, «Panpsychism and Panprotopsychism», in Alter T. e Nagasawa Y. (eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism*, New York, Oxford University Press, pp. 246-276.
- Chalmers, D.J., 2016, «The Combination Problem for Panpsychism», in Brüntrup G. e Jaskolla L. (eds.), *Panpsychism*, New York, Oxford University Press, pp. 179-214.
- Crick F. e Koch C., 1990, «Towards a Neurobiological Theory of Consciousness», *Seminars in the Neuroscience*, 2, pp. 263-275.
- Dehaene S., Changeaux J.P., Kerszberg M., (1998) «A Neuronal Model of a Global Workspace in Effortful Cognitive Tasks», *Proc. Natl. Acad. Sci. USA*, 95, pp. 14529-14534.
- Dennett D.C., 1991, *Consciousness Explained*, London, Penguin.
- Dennett D.C., 1996a, «Consciousness: More Like Fame than Television», in Maar C., Pöppel E., Christaller, T. (eds.), *Die Technik auf dem Weg zur Seele*. Munich, Rowohlt.
- Dennett D.C., 1996b, «Facing Backwards on the Problem of Consciousness», *Journal of Consciousness Studies*, 3, 1, pp. 4-6.
- Dennett D.C., 2001, «Are We Explaining Consciousness Yet?», *Cognition*, 79, pp. 221-237.
- Edelman G. M., 1987, *Neural Darwinism. The Theory of Neuronal Group Selection*, New York, Basic Books.
- Edelman G.M., 1989, *The Remembered Present: A Biological Theory of Consciousness*, New York, Basic Books.
- Goff P., 2017, *Consciousness and Fundamental Reality*, Oxford, Oxford University Press.
- Hassel-Mørch H., 2017, «Is Matter Conscious? Why the Central Problem in Neuroscience is Mirrored in Physics», *Nautilus Magazine*, 47, On-line: <http://nautil.us/issue/47/consciousness/is-matter-conscious>.
- Hofstadter D., 2007, *I Am a Strange Loop*, New York, Basic Books.
- Holcombe A.O., 2009, «Seeing Slow and Seeing Fast: Two Limits of Perception», *Trends in Cognitive Sciences*, 13, 5, pp. 216-221.

- Jackson F., 1986, «What Mary Didn't Know», *The Journal of Philosophy*, 83, 5, pp. 291-295.
- Koch C., 2012, *Consciousness: Confession of a Romantic Reductionist*, Cambridge, MIT Press.
- Levine J., 1983, «Materialism and Qualia: the Explanatory Gap», *Pacific philosophical quarterly*, 64, pp. 354-361.
- Massimini M., Boly M., Casali A., Rosanova M., Tononi G., 2009, «A Perturbational Approach for Evaluating the Brain's Capacity for Consciousness», *Prog. Brain Res.*, 177, pp. 201-214.
- Massimini M., Tononi G., 2013, *Nulla di più grande*, Milano, Baldini & Castoldi.
- Mindt G., 2017, «The Problem with the "Information" in Integrated Information Theory», *Journal of Consciousness Studies*, 24, 7-8, pp. 130-154.
- Nagel T., 1974, «What Is It Like to Be a Bat?», *The Philosophical Review*, 83, 4, pp. 435-450.
- Oizumi M., Albantakis L., Tononi G., 2014, «From Phenomenology to the Mechanisms of Consciousness: Integrated information Theory 3.0», *PloS Comput. Biol.*, 10, 5.
- O'Regan J.K., 2016, «How the Sensorimotor Approach Bridges Both Absolute and Comparative Explanatory Gaps», *Journal of Consciousness Studies*, 23, 5-6, pp. 39-65.
- Owen A.M., Coleman M.R., Boly M., Davis M.H., Laureys S., Pickard J.D., 2006, «Detecting Awareness in the Vegetative State», *Science*, 313, 5792, p. 1402.
- Place U.T., 1956, «Is Consciousness a Brain Process?», *British journal of psychology*, 47, 1, pp. 44-50.
- Rudrauf D., Bennequin D., Granic I., Landini G., Friston K., Williford K., 2017, «A Mathematical Model for Embodied Consciousness», *Journal of Theoretical Biology*, 428, pp. 106-131.
- Searle J.R., 2013, «Can Information Theory Explain Consciousness?» *The New York Review of Books*, 60, 1. On-line: <http://www.nybooks.com/articles/2013/01/10/can-information-theory-explain-consciousness/>.
- Seth A.K., Izhikevich E., Reeke G., Edelman G.M., 2006, «Theories and Measures of Consciousness: and Extended Framework», *Proceedings of the National Academy of Sciences*, 103, 28, pp. 10799-10804.
- Smart J.J.C., 1959, «Sensations and Brain Processes», *The Philosophical Review*, 68, 2, pp. 141-156.
- Strawson G., 2006, «Realistic monism: Why Physicalism Entails Panpsychism», *Journal of Consciousness Studies*, 13, 10-11, pp. 3-31.
- Thompson E.T., 2007, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Cambridge, MA, Harvard University Press.
- Thompson E.T., 2011, «Précis of Mind in Life», *Journal of Consciousness Studies*, 18, 5-6, pp. 10-22.
- Tononi G., 2003, *Galileo e il fotodiodo. Cervello, complessità e coscienza*, Roma-Bari, Laterza.

- Tononi G., 2004, «An Information Integration Theory of Consciousness», *BMC neuroscience*, 5, 42.
- Tononi G., 2005, «Consciousness, Information Integration, and the Brain», *Progress in Brain Research*, 150, pp. 109-126.
- Tononi G., 2008, «Consciousness as Integrated Information: a Provisional Manifesto», *The Biological Bulletin*, 215, 3, pp. 216-242.
- Tononi G., 2012, «Integrated Information Theory of Consciousness: an Updated Account», *Arch Ital Biol*, 150, 2-3, pp. 56-90.
- Tononi G., Koch C., 2015, «Consciousness: Here, There and Everywhere?», *Phil. Trans. R. Soc. B*, 370.
- Tononi G., Boly M., Massimini M., Koch C., 2016, «Integrated Information Theory: From Consciousness to Its Physical Substrate», *Nature Reviews Neuroscience*, 17, 7, pp. 450-461.
- Varela F.J., 1996, «Neurophenomenology: a Methodological Remedy to the Hard Problem», *Journal of Consciousness Studies*, 3, 4, pp. 330-349.
- Wegner D.M., 2002, *The Illusion of Conscious Will*, Cambridge, MIT Press.

AphEx.it è un periodico elettronico, registrazione n° ISSN 2036-9972. Il copyright degli articoli è libero. Chiunque può riprodurli. Unica condizione: mettere in evidenza che il testo riprodotto è tratto da www.aphex.it

Condizioni per riprodurre i materiali --> Tutti i materiali, i dati e le informazioni pubblicati all'interno di questo sito web sono "no copyright", nel senso che possono essere riprodotti, modificati, distribuiti, trasmessi, ripubblicati o in altro modo utilizzati, in tutto o in parte, senza il preventivo consenso di AphEx.it, a condizione che tali utilizzazioni avvengano per finalità di uso personale, studio, ricerca o comunque non commerciali e che sia citata la fonte attraverso la seguente dicitura, impressa in caratteri ben visibili: "www.aphex.it". Ove i materiali, dati o informazioni siano utilizzati in forma digitale, la citazione della fonte dovrà essere effettuata in modo da consentire un collegamento ipertestuale (link) alla home page www.aphex.it o alla pagina dalla quale i materiali, dati o informazioni sono tratti. In ogni caso, dell'avvenuta riproduzione, in forma analogica o digitale, dei materiali tratti da www.aphex.it dovrà essere data tempestiva comunicazione al seguente indirizzo (redazione@aphex.it), allegando, laddove possibile, copia elettronica dell'articolo in cui i materiali sono stati riprodotti.

In caso di citazione su materiale cartaceo è possibile citare il materiale pubblicato su AphEx.it come una rivista cartacea, indicando il numero in cui è stato pubblicato l'articolo e l'anno di pubblicazione riportato anche nell'intestazione del pdf. Esempio: Autore, *Titolo*, <<www.aphex.it>>, 1 (2010).
