**Università degli Studi di Torino**

*Corso di Laurea in Filosofia*

**Causalità del mentale: dall’argomento di Kim ai qualia in una prospettiva fisicalista**

Tesi di Laurea

**Relatore/Relatrice**

Di Bona Elvira

**Correlatore/Correlatrice**

Barbero Carola

**Candidato/a**

**Capitaneo Leonardo**

Matricola

949776

Anno Accademico 2022/2023

*A Noemi, la prima lettrice di questo piccolo lavoro.*

Sommario

[Introduzione 5](#_Toc137316158)

[1. L’argomento di Kim 10](#_Toc137316159)

[1.1 I presupposti 10](#_Toc137316160)

[1.1.1 Sopravvenienza e irriducibilità del mentale al fisico 10](#_Toc137316161)

[1.1.2 Causalità del mentale, esclusione, e chiusura causale del mondo fisico 13](#_Toc137316162)

[1.2 Esposizione dell’argomento 14](#_Toc137316163)

[1.2.1 Completamento 1 15](#_Toc137316164)

[1.2.2 Completamento 2 16](#_Toc137316165)

[1.3 Conseguenze dell’argomento 17](#_Toc137316166)

[1.3.1 Realizzabilità multipla 18](#_Toc137316167)

[1.4 Rispondere alla realizzabilità multipla: funzionalismo 20](#_Toc137316168)

[1.4.1 Funzionalismo del ruolo 21](#_Toc137316169)

[1.4.2 Funzionalismo dell’occupante 24](#_Toc137316170)

[2. Qualia in una prospettiva fisicalista 28](#_Toc137316171)

[2.1. “Problemi facili” e il “problema difficile” 28](#_Toc137316172)

[2.1.1. *The explanatory gap* 28](#_Toc137316173)

[2.2 Esperimento mentale degli zombie 30](#_Toc137316174)

[2.2.1 Diversi sensi di possibilità 31](#_Toc137316175)

[2.2.2 Argomento della concepibilità 32](#_Toc137316176)

[2.2.3 Concepibilità 32](#_Toc137316177)

[2.2.4 Esperimento mentale contro la concepibilità di un mondo-zombie 34](#_Toc137316178)

[2.2.4.1 Precisazione su qualia e colori 38](#_Toc137316179)

[2.3 Analisi della possibilità che i qualia possano essere causa di comportamenti 39](#_Toc137316180)

[2.3.1 Spiegazione controfattuale 41](#_Toc137316181)

[2.3.2 Soluzione basata su un controfattuale 42](#_Toc137316182)

[2.3.3 Risultati teorici raggiunti 46](#_Toc137316183)

[2.4 Considerazioni finali 47](#_Toc137316184)

[Bibliografia 51](#_Toc137316185)

[Primaria: 51](#_Toc137316186)

[Secondaria: 51](#_Toc137316187)

# Introduzione

Che la mente abbia potere causale sembra sia scontato per tutti, ognuno di noi si occupa di faccende che tutti i giorni ci fanno pensare di essere agenti cognitivi. Per avere un’idea di che cosa stiamo parlando proviamo a pensare ad una scena molto comune: immaginiamo di alzarci improvvisamente dal letto in piena notte perché abbiamo sete e inavvertitamente sbattiamo un piede sullo stipite della porta, il dolore provocato dall’urto magari ci farà gridare e molto probabilmente il focus della nostra attenzione si sposterà dal bere un bicchiere d’acqua al prendere un po’ di ghiaccio da applicare sul piede dolorante. Questa è una scena che potrebbe essere capitata a chiunque, raramente però ci interrogheremo sulle dinamiche causali che ci sono dietro ad azioni quotidiane come questa. Cerchiamo di analizzare meglio cosa è successo: lo stato mentale dell’“avere sete” ha causato la nostra azione di alzarci dal letto e lo stato mentale del “dolore” ha provocato il desiderio di andare a prendere un po’ di ghiaccio da applicare sul piede; questi due stati mentali sembrano essere cause genuine di movimenti e comportamenti differenti e nessuno di noi normalmente avrebbe dubbi nell’affermare l’effettiva efficacia causale degli stati mentali qui sopra descritti. Lo spazio del mentale sembra giocare un ruolo decisivo nella nostra vita quotidiana, ci sentiamo possessori di una vita mentale fatta di pensieri, immagini, credenze, e conoscenze e pensiamo che l’agire in modo giusto o sbagliato sia frutto di scelte calcolate in base a queste stesse credenze e conoscenze che sembrano abitare nella nostra mente. L’essenza stessa del libero arbitrio sembra essere la capacità di possedere una mente che possa avere un potere causale e che possa farci agire nel modo migliore possibile (o anche in quello peggiore); in sintesi ogni nostra azione concreta sembra necessitare il potere causale del mentale. Ci sono inoltre scienze che esistono in quanto presuppongono l’esistenza del mentale, la psicologia per esempio è una di queste. Gli psicologi sembrano infatti non avere dubbi sull’esistenza di una mente capace di agire sul fisico, nel trattare per esempio problemi di ansia o depressione, la prima cosa che farà un terapeuta sarà cercare di suggerire tecniche che possano migliorare pensieri disfunzionali che interferiscono e rendono difficile la vita quotidiana del paziente. Rendere inerte il mondo mentale significa mettere in serio pericolo i presupposti della stessa psicologia contemporanea. Nel dibattito contemporaneo il problema dell’efficacia causale del mentale nasce dal cercare di conciliare esigenze teoriche diverse, spesso derivanti dal tipo di dottrina metafisica che si sostiene. Colui che per primo ha dovuto affrontare il problema della causalità del mentale fu Cartesio nelle *Meditazioni metafisiche*. Cartesio è considerato il padre della dottrina metafisica chiamata “dualismo delle sostanze”, che consiste nel considerare la mente e il corpo come due sostanze completamente irriducibili, in quanto l’essenza stessa della materia, di cui è parte il corpo, è pura estensione, mentre l’essenza del mentale è il pensiero e quest’ultimo per definizione non è esteso[[1]](#footnote-1). Conseguenza di questa irriducibilità è la spaccatura del mondo in due tipi di enti: la *res extensa*, che consiste in tutto il mondo fisico contraddistinto dall’essere estesoe le *res cogitantes[[2]](#footnote-2)*, che consistono nelle singole menti pensanti che popolano le singole porzioni di *res extensa* che comunemente chiamiamo “corpi”. Questi due tipi di enti, poiché sono per definizione irriducibili, pongono in luce un importante problema: come può il mentale interagire causalmente con il fisico se stiamo parlando di due sostanze diverse e irriducibili? Cartesio stesso era ben consapevole del problema e ha cercato di dare una soluzione che già alla sua epoca veniva vista come insostenibile. La soluzione è quella della cosiddetta “ghiandola pineale”, una parte del cervello (ipofisi) che in virtù delle sue proprietà caratteristiche sarebbe in grado di fare da “ponte” tra le due sostanze e farle così interagire. Il problema che sorgeva evidente già al filosofo contemporaneo di Cartesio, Gassendi, e che è ancora la classica critica che continua a perdurare nei confronti del dualismo delle sostanze cartesiano, è che la ghiandola pineale non può essere utile per spiegare l’interazione tra le due sostanze, per il fatto che essa è pur sempre una parte di materia avente estensione e rimane dunque irriducibile al mentale. L’alternativa per evitare questo problema sembrerebbe quella di sostenere forme più deboli di dualismo, come quello delle proprietà o forme di identificazione tra mentale e fisico, come le teorie dell’identità. In questi ultimi decenni, chi sostiene il dualismo delle proprietà non ha potuto fare a meno che scontrarsi con un famoso argomento proposto dal filosofo Jaegwon Kim, che sembra mettere in serio pericolo sia una causalità genuina del mentale sia l’esistenza stessa di proprietà non riducibili a quelle fisiche. Nel nostro caso utilizzeremo l’argomento di Kim come una sorta di “vaglio metodologico” per cercare la teoria metafisica sul mentale più soddisfacente[[3]](#footnote-3). Compito del primo capitolo di questo elaborato sarà infatti trovare la migliore teoria in grado di resistere all’argomento di Kim ma che possa spiegare in modo soddisfacente la causalità del mentale, non riducendola ad una pura illusione. Come vedremo l’alternativa migliore che ci rimane da adottare sembra essere una forma di teoria dell’identità. Un secondo problema di grande importanza è quello dei qualia, i quali sembrano essere completamente irriducibili in un quadro fisicalista, per via di un *gap* esplicativo che sembra esserci tra le proprietà fenomeniche e quelle fisiche. Analizzerò il famoso esperimento mentale di Ned Block, volto a mostrare che il funzionalismo è falso ed esporrò un ulteriore esperimento mentale volto a mostrare l’impossibilità da parte del linguaggio funzionalista di rendere conto delle proprietà fenomeniche. Come verrà spiegato, la buona riuscita dell’esperimento mentale, a differenza di quello di Block, non dipende dallo stato di cose, ma dal semplice linguaggio funzionalista. Nel secondo capitolo ci occuperemo proprio del problema dei qualia e in particolare proveremo a conciliarli in un quadro fisicalista. Essi infatti sono sempre stati considerati come l’ultima difesa per chi vuole a tutti i costi preservare forme di dualismo, trovare una soluzione per includerli completamente in un’ontologia fisicalista sembra essere di vitale importanza per chi adotta una posizione materialista. Analizzeremo a tal proposito l’esperimento mentale degli zombie, proposto da Chalmers, che intende stabilire che i qualia non sono qualcosa di fisico. Argomenterò a favore di una posizione che vede i qualia come responsabili di alcuni comportamenti umani e con questo cercherò di dimostrare che l’esperimento mentale di Chalmers non dimostra ciò che intende dimostrare. I comportamenti di cui i qualia sembrano responsabili, sono quelli che riguardano scelte di “gusto”, che sia la scelta tra colori o che sia la scelta tra gusti di gelato diversi. Ritengo che l’”effetto particolare[[4]](#footnote-4)” che si prova quando si è davanti ad una scelta di questo tipo, sia rilevante nel determinarla. Poiché questo problema può essere formulato attraverso un condizionale controfattuale, cercherò di dimostrare questo punto attraverso la teoria di Lewis sui controfattuali che si basa su relazioni di somiglianza tra mondi possibili. La tesi centrale che intendo sostenere in questo elaborato consiste nel mostrare che si può benissimo adottare una teoria fisicalista che possa comprende ogni aspetto del mentale, inclusi i qualia. Sembra infatti assurdo, o quanto meno curioso, il fatto che quasi tutti i fenomeni mentali possano essere considerati come fenomeni fisici mentre la coscienza sembra rimanere fuori dal dominio fisico. Come cercherò di dimostrare in questo elaborato, il semplice problema del gap esplicativo non deve portarci a pensare che non stiamo parlando di qualcosa di fisico. Il vero problema che sembra esserci dietro la difficoltà di concepire il mentale come fisico sembra essere legata a dei nostri preconcetti legati spesso a certi termini che nella vita quotidiana utilizziamo. Comunemente infatti utilizziamo i termini “mentale” e “fisico” come due termini che si riferiscono a cose diverse, quando invece potrebbero riferirsi alla stessa cosa. Concepire il mentale come fisico, come vedremo, ha il vantaggio di preservarne la sua efficacia causale e quindi accordarsi con il nostro senso comune secondo cui siamo agenti cognitivi ed abbiamo interazioni causali derivanti da decisioni mentali. Come cercherò di mostrare in questo elaborato, concepire il mentale come fisico non significa privarne della sua particolarità, significa semplicemente considerare la mente da un altro punto di vista. In questo modo possiamo parlare del “mentale”, riferirci ad esso in tutte le sue sfaccettature, senza per questo uscire dal dominio fisico. Questo ha il vantaggio, non da poco, di legittimare le nostre asserzioni quotidiane sul mentale, comprendere come esso possa essere causa genuina delle nostre azioni e legittimare le scienze che si basano sulla sua esistenza, come la psicologia scientifica. In più, elaborare una posizione filosofica che possa rendere conto di tutti gli aspetti del mentale, inclusi gli aspetti qualitativi, può togliere finalmente quell’aurea di misticità che sta dietro alla coscienza, a cui tanti dualisti si aggrappano, senza in alcun modo privarne la sua particolarità. Come ho già evidenziato prima, si tratta semplicemente di cambiare i nostri preconcetti dovuti spesso a vecchie tradizioni filosofiche[[5]](#footnote-5), e provare a concepire il fatto che la materia possa essere artefice di cose meravigliose e misteriose come la coscienza.

Volevo ringraziare la prof.ssa Elvira Di Bona per avermi seguito in questo lavoro e per i suoi preziosi consigli. Un ringraziamento è dovuto anche a tutti gli amici e colleghi, che indirettamente hanno contribuito ad arricchire la mia tesi di laurea. Devo tantissimo alla mia ragazza, che con pazienza ha letto passo dopo passo ogni riga di questo elaborato in fase di stesura. Uno speciale ringraziamento lo devo al mio professore di filosofia del liceo, che se non avesse fatto nascere in me l’interesse per questa disciplina questo mio breve lavoro non avrebbe mai visto la luce.

# L’argomento di Kim

## I presupposti

Prima di affrontare l’argomento bisogna comprendere alcuni concetti chiave utili a farci entrare nel cuore del dibattito: il primo concetto è la sopravvenienza delle proprietà mentali sulle proprietà fisiche, il secondo è l’irriducibilità delle prime a queste ultime, il terzo è il potere causale che sempre le prime hanno, il quarto è l’esclusione causale, e il quinto è la chiusura causale del mondo fisico. Essi sono le premesse dell’argomento e accettarli come tali significa accettare le loro necessarie conseguenze. Dunque qualsiasi posizione teorica che ammette la verità di queste premesse non potrà fare altro che confrontarsi con l’argomento di Kim. Per Kim le prime tre premesse congiuntamente formulano la posizione che in filosofia della mente è comunemente conosciuta come “fisicalismo non riduzionista”, posizione metafisica tipica di chi cerca di conciliare posizioni materialiste riservando però uno spazio legittimo al mentale, ritenendolo non riducibile alle proprietà fisiche. La conseguenza di queste premesse come si vedrà è che le proprietà mentali o sono solamente epifenomeniche, dunque impotenti dal punto di vista causale, oppure sono completamente riducibili a proprietà fisiche. Prima di esporre l’argomento è bene esplicitare le cinque premesse generali[[6]](#footnote-6).

### Sopravvenienza e irriducibilità del mentale al fisico

Esistono diversi tipi e modi di intendere la sopravvenienza, McLaughlin caratterizza l’idea centrale della sopravvenienza come quella relazione particolare in cui dati due tipi di oggetti, x e y, nessun cambiamento avviene in x se non avviene anche in y[[7]](#footnote-7); Kim suggerisce due versioni di sopravvenienza, quella “debole” e quella “forte”, la sopravvenienza debole si presenta quando in ogni mondo *W*, se due enti *x* e *y* non differiscono rispetto a proprietà-A allora non differiscono neanche rispetto a proprietà-B, dunque due oggetti in un mondo con le stesse proprietà microscopiche non differiranno nelle proprietà macroscopiche. Si parla di sopravvenienza forte quando per ogni mondo *W1* e *W2*, se le proprietà-A di *x* in *W1* non differiscono rispetto alle proprietà-A di *y* in *W2,* allora anche le proprietà-B di *x* non differiranno rispetto alle proprietà-B in *y*. Dunque per esempio si dirà che le proprietà mentali sopravvengono in modo “forte” alle proprietà fisiche se e solo se dati due oggetti in due mondi possibili con le stesse proprietà fisiche, manifestano esattamente le stesse proprietà mentali. La sopravvenienza forte è esattamente quella che Kim ha in mente nel suo argomento e perciò ci concentreremo su questa. Il secondo presupposto da considerare rappresenta il nucleo centrale di ogni posizione fisicalista non riduzionista, un fisicalismo dunque che accetta l’esistenza di proprietà che non possono essere ridotte a quelle fisiche pur mantenendo uno stretto legame ontologico con queste ultime. Per mantenere una metafisica fisicalista con la presenza di proprietà non riducibili a proprietà fisiche, in questo caso quelle mentali, c’è bisogno che queste siano in qualche modo dipendenti ontologicamente da quelle fisiche in modo tale che se non ci fossero quelle fisiche non ci sarebbero neppure quelle mentali. Questo tipo di dipendenza giustifica la sopravvenienza delle proprietà mentali su quelle alle fisiche. Questo presupposto esclude per principio il dualismo delle sostanze, posizione che in filosofia della mente quasi nessuno sostiene[[8]](#footnote-8) per via dei grossi problemi che esso sembra implicare. Uno dei classici problemi è il come spiegare l’interazione causale tra sostanze completamente diverse, che è lo stesso problema che ha avuto Cartesio; un altro problema attualmente discusso oggi è quello dell’ “accoppiamento”, ovvero ponendo che esistano due sostanze pensanti M1 ed M2 esse dovranno essere “accoppiate” rispettivamente a due corpi estesi C1 e C2 per avere interazioni causali con essi, ma dato che le relazioni causali si presentano in spazi estesi e due sostanze pensanti sono per definizione non estese, allora è difficile concepire come questo possa essere possibile. Anche se l’accoppiamento fosse possibile, rimarrebbe un problema irrisolvibile stabilire in che parte del corpo ha sede la sostanza pensante, non essendo infatti per definizione estesa, non può essere localizzata in nessun luogo[[9]](#footnote-9). Per non andare incontro a questo tipo di problemi, la soluzione migliore sembra essere l’adozione di posizioni in cui pur salvando la specificità del mentale lo si identifica con proprietà di sostanze fisiche. Tali posizioni, definite “dualismi delle proprietà” sostengono l’esistenza di un’unica sostanza, ovvero la materia, che presenta però proprietà non riducibili a quest’ultima. Il vantaggio di adottare questo tipo di posizioni consiste nel preservare il legame tra mente e cervello, senza che ciò implichi un’identità stretta tra i due. Anche grazie a dimostrazioni sperimentali nelle neuroscienze nessuno più ha dubbi riguardanti le correlazioni del mentale con il cervello. Danni cerebrali infatti determinano modificazioni nelle manifestazioni mentali, dunque anche la sperimentazione empirica sembra confermare la sopravvenienza del mentale sul fisico. Kim assume l’irriducibilità nel suo argomento soprattutto perché gli sembra che gli argomenti proposti da altri autori che supportano invece una completa riduzione del mentale al fisico non siano completamente conclusivi. Sempre in “*Physicalism or Something Near Enough*” Kim fa notare che tutti coloro che sostengono forme di riduzionismo come le teorie di identità di tipo, giustificano la loro posizione ritenendo che la loro conclusione sia un’inferenza alla spiegazione migliore, questo tipo di inferenza non essendo deduttiva non implica la necessità della conclusione[[10]](#footnote-10). Conclusioni come quelle offerte da Smart[[11]](#footnote-11), appellandosi ai rasoi di Occam, sembrano funzionare solo se si assume che il mondo sia per natura semplice, dunque in questo caso la teoria più semplice potrebbe rivelarsi quella corretta, ma poiché non possiamo escludere che in realtà il mondo sia complesso, le teorie più semplici potrebbero non essere corrette. Altre posizioni non riduzioniste sono l’emergentismo e l’epifenomenismo, scegliere però tra queste due posizioni significa portarsi dietro i connessi problemi teorici. L’emergentismo afferma che le proprietà mentali “emergono” dalle proprietà fisiche[[12]](#footnote-12) e le prime sono causalmente efficaci (l’argomento stesso di Kim escluderà questa possibilità), l’epifenomenismo invece riduce le proprietà mentali a semplici epifenomeni, rendendole impotenti dal punto di vista causale. Per evitare di escludere ogni tipo di possibilità Kim utilizza questo secondo presupposto per affermare che l’argomento funziona per quasi[[13]](#footnote-13) ogni tipo di posizione che ammette l’irriducibilità del mentale al fisico.

### Causalità del mentale, esclusione, e chiusura causale del mondo fisico

Il terzo presupposto restringe il campo delle posizioni in filosofia della mente che vengono attaccate dall’argomento, l’epifenomenismo infatti sembra escluso, quindi più in generale l’argomento è proposto per dualisti delle proprietà che accettano l’efficacia causale del mentale. Accettare la causalità del mentale sembra infatti scontato alla maggior parte delle persone, privare il mentale di efficacia causale rischierebbe di mettere in discussione la gran parte della nostra quotidianità, le nostre azioni come agenti cognitivi sembrerebbero venire meno, i nostri ricordi e la nostra conoscenza, che sembrano condizionarci così tanto nella nostra vita, in realtà non lo farebbero e dunque il loro potere causale sarebbe un’illusione. Accettare questo presupposto ha dunque un obiettivo ben preciso, ovvero mettere alla prova un’intuizione che accetteremmo tutti e che potrebbe mettere a rischio una delle nostre convinzioni più scontate. Il principio dell’esclusione causale nega per principio casi di sovradeterminazione, se si ammette la sovradeterminazione infatti, due cause differenti *x* e *y* possono essere entrambe cause genuine di un effetto *e*. Per spiegare come funziona un caso di sovradeterminazione causale viene spesso fornito un esempio del tipo seguente: un uomo viene colpito nello stesso momento da una pietra e da un proiettile e l’uomo muore. Se accettiamo la sovradeterminazione, le cause della sua morte sono da imputare sia alla pietra sia al proiettile. Se ammettiamo l’esclusione causale, casi come questo non possono darsi; riporterò durante l’esposizione dell’argomento la formulazione dell’esclusione causale proposta da Kim. L’ultimo presupposto, il principio di chiusura causale del mondo fisico, può essere formulato in diversi modi, Kim stesso ne dà due diverse formulazioni, una più “debole” ed una più “forte. Quest’ultima afferma che: “ogni causa di un evento fisico è essa stessa un evento fisico, […], nessun evento che non è fisico può essere una causa di un evento fisico”[[14]](#footnote-14), con questo tipo di formulazione se è presente un effetto fisico si escluderà da sé la possibilità di sovradeterminazione da parte di cause non fisiche. Kim preferisce mantenere una versione più “debole” nel suo argomento poiché crede sia più forte filosoficamente arrivare alla stessa conclusione con premesse più deboli. Questo gli permette di utilizzare nell’argomento anche il principio di esclusione causale, poiché se si accetta la versione “forte” della chiusura si potrebbe anche evitare di usare, dato che esclude per principio ogni tipo di causa che non sia fisica[[15]](#footnote-15). Darò la formulazione “debole” della chiusura causale offerta da Kim durante l’esposizione dell’argomento.

## Esposizione dell’argomento[[16]](#footnote-16)

Supponiamo che ci sia una relazione causale tra due proprietà mentali M e M1, in particolare M causa M1. Kim sottolinea che non sono le proprietà mentali in quanto tali ad entrare in relazione causale, ma sono le istanze M ed M1, in particolare l’istanziazione di M causa l’istanziazione di M1. In virtù della sopravvenienza, per ogni cambiamento che avviene nel dominio mentale ci sarà un cambiamento che avviene nel dominio fisico. Kim inoltre, interpreta la sopravvenienza come un rapporto di dipendenza ontologica, tale per cui se le proprietà fisiche non ci fossero allora non ci sarebbero neanche le proprietà mentali. Dunque se c’è M1 , ci deve anche essere una base di sopravvenienza fisica che permetta la sua istanziazione, ovvero P1, o detto in altri termini, ogni qual volta vi è una proprietà fisica P1 si necessita per sopravvenienza l’istanziazione di una proprietà mentale M1. Questo sembra rendere completamente inutile M come causa di M1 tranne nel caso in cui M è in qualche modo responsabile dell’istanziazione di P1. Da questo momento in avanti l’argomento si può completare in due modi diversi, dedicheremo due paragrafi a queste due possibilità per rendere più chiare possibili le differenze che ci sono tra una e l’altra.

### 1.2.1 Completamento 1

Il primo modo consiste nell’affermare che per sopravvenienza anche M ha una base di sopravvenienza fisica P, così facendo sembra che siano due le cause dell’istanziazione di P1, ovvero M e P; è bene precisare che queste due cause siano distinte perché stiamo assumendo che mentale e fisico siano irriducibili, ed poiché stiamo anche assumendo l’efficacia causale del mentale dobbiamo poter ipotizzare che sia P che M possano essere cause genuine dell’istanziazione di P1. Arrivati a questo punto Kim applica il principio di esclusione causale da lui stesso formulato, secondo cui: “se un evento *e* ha una sufficiente causa *c* in *t*, nessun evento in *t* distinto da *c* può essere causa di *e*”[[17]](#footnote-17). Dunque P1 non potrà essere causalmente sovradeterminato, ma dovrà avere o P o M come causa della sua istanziazione. Ora che si pone il problema di quale causa scegliere, e Kim gioca la carta del principio di chiusura causale del mondo fisico così da lui formulato nella versione “debole”: “se un evento fisico ha una causa in *t*, allora ha una causa fisica in *t*”[[18]](#footnote-18), quindi dato che P1 è un evento fisico, esso avrà una causa fisica, dunque in questo caso P. Il principio di chiusura causale insieme a quello dell’esclusione ci porterà a scegliere come unica causa P. L’istanziazione della proprietà fisica P quindi è l’unica causa genuina dell’istanziazione della proprietà fisica P1. Schematicamente, Kim formula l’argomento come segue:

1. M causa M1.
2. Per ogni proprietà fisica P1, M1 ha P1 come sua base di sopravvenienza.
3. M causa M1 causando la sua base di sopravvenienza P1.
4. M ha una base di sopravvenienza fisica P.
5. M causa P1, e P causa P1. [da (3) e (4)]
6. M≠P. [dal principio di irriducibilità del mentale al fisico]
7. P1 non è causalmente sovradeterminato da M e P. [dal principio di esclusione causale]
8. La presunta causa mentale M, è esclusa dalla causa fisica P; dunque P non M è una causa di P1. [dal principio di chiusura causale del mondo fisico]

### 1.2.2 Completamento 2

Il secondo modo viene ritenuto più semplice, in quanto dopo che M causa l’istanziazione di M1 per aver causato l’istanziazione di P1 come sua base di sopravvenienza e che dunque M sembra causa di P1, si procede con l’applicare direttamente il principio di chiusura causale esposto sopra, così facendo si ammette già che P1 deve avere una causa fisica, ovvero P. Allo stesso tempo in cui c’è P occorre M, ma in quanto irriducibili non si esclude la possibilità che P1 abbia due cause distinte M e P, non essendo però un caso di sovradeterminazione, per il principio di esclusione causale solo uno tra M e P rimarrà come causa genuina dell’istanziazione di P1. Applicando sia il principio di esclusione causale sia quello di chiusura avremo che P rimane l’unica causa genuina dell’istanziazione di P1. Come formulato schematicamente da Kim avremo:

1. M causa M1.
2. Per ogni proprietà fisica P1; M1 ha P1 come sua base di sopravvenienza.
3. M causa M1 causando la sua base di sopravvenienza P1.
4. M è una causa di P1.
5. P1 ha una causa fisica P, al cui tempo occorre M.
6. M≠P [per il principio di irriducibilità del mentale al fisico].
7. Quindi P1 ha due cause distinte, M e P, e questo non è un caso di sovradeterminazione.
8. Quindi per l’esclusione causale, rimarrà solo M o P come causa genuina di P1.
9. Per chiusura causale ed esclusione causale, resta solo P.

I prossimi paragrafi saranno dedicati a valutare quali alterative rimangono una volta accettata la validità dell’argomento di Kim.

## Conseguenze dell’argomento

Come si può vedere da entrambe le formulazioni fornite da Kim, il risultato dell’argomento è che il mentale risulta causalmente impotente, sia per quanto riguarda la causalità tra il mentale e il fisico, sia per quanto riguarda la causalità tra due proprietà mentali. Kim ci mette davanti ad una scelta imprescindibile tra il salvare le proprietà mentali come un qualcosa di non riducibile oppure adottare una forma di riduzionismo in cui le proprietà mentali sono in realtà stati cerebrali. Dunque se proprio vogliamo salvare qualche forma del mentale dal riduzionismo l’unica scelta che abbiamo è adottare un approccio epifenomenista, che però non è una posizione molto allettante dato che sembra andare contro le nostre principali intuizioni. L’alternativa a questa posizione è adottare una forma di teoria dell’identità, ma anche questo comporta diversi problemi[[19]](#footnote-19). Tra le due alternative però la posizione migliore sembra essere la seconda, poiché solo con essa si preserva la causalità, come afferma Kim: “il miglior risultato successivo, in effetti la nostra unica speranza a questo punto se vogliamo salvare la causalità del mentale, è il riduzionismo fisico”.[[20]](#footnote-20) Come si è visto infatti sia il dualismo delle sostanze sia un dualismo delle proprietà sembrano, dal punto di vista causale, totalmente inefficaci: il primo perché le relazioni causali si devono dare in un reticolo spazio-temporale e tutto ciò che non è esteso non può parteciparvi, il secondo deve affrontare l’argomento di Kim e a meno che non si neghi la chiusura causale o si affermi la sovradeterminazione, le proprietà mentali sono impotenti. Se si ammettesse però che le proprietà mentali non sono altro che la stimolazione di alcune fibre cerebrali, non sembrano esserci i problemi dei dualisti. Il mentale rientrerebbe infatti totalmente nel dominio delle cose fisiche, dunque la chiusura causale non è in pericolo, non c’è pericolo di sovradeterminazione causale dal momento che le uniche cause sono fisiche, ma soprattutto il mentale preserverebbe la sua efficacia ritenuta imprescindibile da molti. Dunque, visto che ritengo importante preservare la causalità del mentale, mi concentrerò su queste forme di riduzionismo, ma come abbiamo già accennato adottare una teoria dell’identità ha i suoi problemi, quindi in ogni caso se dobbiamo adottare questo tipo di teorie, dobbiamo capire come risolverli. I prossimi paragrafi saranno dedicati alla realizzabilità multipla e al trovare una teoria che possa sia resistere a quest’ultima, sia all’argomento di Kim.

### Realizzabilità multipla

Ormai famosissimo in letteratura, l’argomento della realizzabilità multipla è stato utilizzato e formulato per la prima volta da Hilary Putnam nel suo saggio “*The Nature of Mental States*”[[21]](#footnote-21), tale saggio è considerato una delle prime formulazioni del funzionalismo, posizione che consiste nel considerare gli stati mentali in base al loro ruolo nell’organizzazione funzionale dell’organismo. Uno stato mentale verrà definito tale se e solo se svolge una certa funzione, per esempio lo stato mentale “dolore” potrà essere definito tale se e solo se dati certi input, come un taglio sul braccio, provocherà un certo stato di agitazione, dato dalla rilevazione di un danno fisico, e sarà responsabile di una serie di output, come l’andare a medicarsi. In questo modo, indipendentemente dal tipo di statuto metafisico del realizzatore del dolore, possiamo dare una spiegazione di che cos’è il dolore, come anche Putnam afferma, infatti, il funzionalismo è perfettamente compatibile con un dualismo. Lo scopo di Putnam era di trovare una valida alternativa alle teorie dell’identità di tipo che falliscono davanti all’ipotesi della realizzabilità multipla: secondo le teorie dell’identità di tipo, una determinata proprietà mentale corrisponde esattamente alla stimolazione di una certa “zona cerebrale”, di solito in letteratura per riferirsi a questo tipo di zone si usa definirle “fibre-C”, dunque una determinata proprietà mentale come il dolore, sarà presente se e solo se verrà stimolata quella determinata fibra-C[[22]](#footnote-22) responsabile della sensazione di dolore. Quella fibra-C è strettamente legata a quel determinato tipo di proprietà mentale, anzi per dirlo in una maniera migliore, la sua stimolazione è esattamente la proprietà mentale. Ma ogni proprietà mentale è strettamente legata solo ad un particolare tipo di fibra-C[[23]](#footnote-23), per esempio il dolore[[24]](#footnote-24) è legato alla stimolazione di una particolare fibra-D, il piacere alla stimolazione di una particolare fibra-E e la rabbia alla stimolazione di una particolare fibra-F. Se e solo se in un organismo sono presenti le fibre-D, quell’organismo potrà provare la sensazione del dolore e se e solo se sono presenti le fibre-E e le fibre-F, in quell’organismo saranno presenti le sensazioni di piacere e di rabbia; ovviamente in tutti questi casi gli organismi sentiranno rispettivamente dolore, piacere e rabbia se e solo se le fibre-C responsabili di quel determinato tipo di attivazione saranno stimolate correttamente[[25]](#footnote-25). Dati questi presupposti, due organismi viventi potranno sperimentare dolore se e solo se possiedono entrambi fibre-D, dunque poiché io e il lettore sentiamo entrambi dolore dopo essere stati vittima di traumi fisici[[26]](#footnote-26), possiamo inferire che entrambi possediamo fibre-D. Possiamo fare lo stesso ragionamento per analogia rispetto ad altri mammiferi simili all’uomo come i cani o i gatti, infatti abbiamo buone ragioni per pensare che anche loro provino dolore se esposti a traumi fisici, dunque anche loro sembrano possedere fibre-D. Il problema inizia a sorgere quando anche altri esseri viventi filogeneticamente lontani da noi sembrano avere sensazioni di dolore. Per esempio i polpi, che sono molluschi, sembrano averlo, ma se il dolore lo si può provare se e solo se si possiedono determinate fibre-D, allora significa che anche i polpi devono possedere fibre-D. Per alcuni può ancora sembrare plausibile che certe strutture cerebrali umane primitive siano uguali a certe strutture cerebrali dei polpi, ma il vero problema delle teorie dell’identità di tipo emerge con l’ipotesi dell’esistenza di esseri che magari esistono in pianeti lontani molto diversi dal nostro. Questi esseri infatti potrebbero avere una “struttura nervosa” di un materiale completamente diverso da quello della nostra struttura cerebrale, ma che svolge le stesse sue funzioni. Questi tipi di creature non potrebbero provare dolore in quanto non possiedono le fibre-D, ma dunque un teorico dell’identità di tipo deve poter ammettere che in tutto l’universo, se qualsiasi essere vivente ha la capacità di provare dolore, allora deve possedere esattamente le stesse fibre-D che possediamo noi; questo sembra abbastanza implausibile. Ma senza prendere in esame ipotesi troppo lontane, anche il semplice fatto che due specie viventi molto diverse una dall’altra possiedano lo stesso tipo di fibre-C sembra implausibile. Il funzionalismo permette di evitare tutto questo, dati infatti due sistemi fisici diversi, la proprietà “dolore” può essere esemplificata da realizzatori diversi. Che siano dunque neuroni o sistemi al silicio, l’importante è che questi realizzatori possano permettere di svolgere il “ruolo del dolore”, ovvero se inseriti in una rete causale, dati certi input risulteranno certi output compatibili con una descrizione che diamo di che cosa è il dolore. Potremmo dire che un robot è in uno stato funzionale del dolore se e solo se possiederà un certo realizzatore fisico, magari in silicio, che dati certi input, come un danno fisico, lo metterà in uno stato di allerta tale per cui procederà con degli output, che potrebbero essere la fuga da quella determinata situazione.

## Rispondere alla realizzabilità multipla: funzionalismo

Visto che per i motivi visti sopra le teorie dell’identità di tipo sembrano portare a conseguenze difficili da accettare e che l’epifenomenismo è completamente insostenibile se vogliamo mantenere l’efficacia causale del mentale, visti i contributi teorici che può dare il funzionalismo, potremmo pensare di adottare questo tipo di approccio. Adottare il funzionalismo però non è privo di problematiche, prima tra tutte è la sua incapacità di spiegare i qualia; prima di vedere quest’ultimo punto, è opportuno vedere quali tipi di funzionalismo esistono. Principalmente esistono due tipi di funzionalismo, tradizionalmente chiamati “funzionalismo del ruolo” e “funzionalismo dell’occupante”, compito di queste due sezioni successive sarà di mostrare quale tra questi due tipi di funzionalismo potremmo accettare per i nostri scopi, tenendo presente che quello che cerchiamo è una teoria che possa spiegare nel miglior modo possibile la causalità del mentale.

### Funzionalismo del ruolo

La prima forma di funzionalismo che analizzerò è quella del ruolo. Questo tipo di funzionalismo è esattamente quello adottato da Kim nell’argomento sopra esposto ed è la forma di funzionalismo più usata in letteratura. Esso consiste nell’ammettere che i realizzatori fisici possiedono delle proprietà di ordine superiore di essere in un determinato stato, come quello del dolore, il quale svolge il ruolo che deve svolgere nell’organizzazione funzionale complessiva[[27]](#footnote-27), dunque in questo caso il ruolo del dolore. Quindi a svolgere il ruolo del dolore negli esseri umani è una certa proprietà di ordine superiore che sopravviene al realizzatore fisico. Tale proprietà non è altro che uno stato funzionale di quell’organismo che possiede il realizzatore ed esso non è altro che una proprietà di livello inferiore. Il problema principale di questo tipo di funzionalismo è quello di restare vittima dell’argomento di Kim, se infatti le proprietà mentali non sono altro che proprietà di ordine superiore M che possiedono proprietà di ordine inferiore P, le quali non sono altro che i realizzatori fisici di queste proprietà, l’argomento ne mostra l’impotenza causale. Oltre a questo problema sono famosissimi gli argomenti di Ned Block contro questo tipo di funzionalismo, primo tra questi è l’argomento del “cervello-cinese[[28]](#footnote-28)”: Block immagina di provare a fare un esperimento, e di far simulare alla popolazione cinese[[29]](#footnote-29) un enorme cervello. In particolare ogni singolo abitante della Cina simulerà un neurone ed equipaggiato con appositi dispositivi di comunicazione, svolgerà il ruolo di sinapsi, comunicando le informazioni che deve comunicare ad un altro abitante della Cina. Supponendo che tutti gli abitanti della Cina riuscissero in tal compito, si avrebbe un macro-cervello riprodotto in tutto e per tutto, la cui descrizione dei ruoli funzionali sarà la stessa che si può dare ad un cervello umano. Secondo Block, anche se una cosa come questa fosse possibile, un cervello-cinese non svilupperebbe mai una coscienza capace di percepire qualia, concludendo così che il funzionalismo è falso. L’argomento di Block mostra che la costruzione di un cervello-cinese con le stesse caratteristiche funzionali di un cervello umano non implica la comparsa di una coscienza. Block arriva a questa conclusione facendo leva sul fatto che ci sembra altamente improbabile che una struttura non organica, al contrario di quella neuronale, possa svilupparla. Questo però non implica l’impossibilità di fare una descrizione funzionalista di organismi dotati di qualia, anche se questi ultimi rimarrebbero comunque fuori da una descrizione di questo genere. Posso infatti descrivere un essere umano nelle sue caratteristiche funzionali, ma questo tipo di descrizione potrebbe non differire in nulla rispetto a quella di un robot. Volevo proporre un esperimento mentale volto a spiegare meglio quest’ultimo punto. Mostrerò dopo l’esposizione dell’esperimento mentale come questo si differenzia da quello proposto da Block e che vantaggio teorico può offrirci. Immaginiamo di essere stati rinchiusi in una stanza dove l’unico modo per poter sapere cosa c’è all’esterno è attraverso lo schermo di un computer, che ci può descrivere attraverso parole scritte su un programma, gli oggetti che ci sono all’esterno. Supponiamo che l’unico tipo di descrizione che questo computer può darci è una descrizione funzionale. Supponiamo anche che questo computer sia collegato ad un sensore puntato verso l’esterno e che ha un’area di visione di dieci metri quadrati, tutto ciò che attraversa quest’area attiverà il computer che darà una descrizione funzionale di tutto ciò che “vede” il sensore. Supponiamo inoltre che l’area sia disseminata di pietre e che arrivi un essere umano, ed entrando in quell’area per una decina di secondi, data la presenza di pietre faccia attenzione a dove mette i piedi. Il computer, non appena il sensore “vedrà” l’essere umano, ci darà una descrizione funzionale del tipo: “c’è una tale cosa *x* (l’essere umano), che quando trova davanti al suo percorso un oggetto *y* (una pietra), si arresta, individua un’altra zona senza *y,* e una volta individuata, procede in quel percorso alternativo”[[30]](#footnote-30). Supponiamo inoltre che successivamente arrivi un robot[[31]](#footnote-31)con una forma umanoide, programmato in modo tale che quando incontra un ostacolo sul suo percorso, cerca un percorso alternativo ed una volta trovato, procede nel percorrerlo. Il computer darà una descrizione esattamente uguale a quella che aveva dato alla visione dell’essere umano, ovvero darà una descrizione del tipo: “c’è una tale cosa *x* (il robot), che quando trova davanti al suo percorso un oggetto *y* (una pietra), si arresta, individua un’altra zona senza *y*, e una volta individuata, procede in quel percorso alternativo”. Il computer non è in grado di discriminare la differenza tra l’essere umano ed il robot e noi non sapremmo mai se ad aver passeggiato al di fuori della stanza in cui siamo rinchiusi ci siano stati due uomini, due robot o un uomo e un robot. Eppure differenze tra un robot ed un essere umano ci sono e quella principale è che quest’ultimo durante la scelta del percorso alternativo ha provato delle sensazioni, cosa che un robot non può fare nel vero senso della parola, l’uomo ha provato qualia, ovvero proprietà fenomeniche, il robot no. Solo se il computer fosse riuscito a dirci che la “tale cosa *x*” possedeva proprietà fenomeniche allora saremmo stati in grado di dire che quella “tale cosa *x*” è un uomo[[32]](#footnote-32). L’argomento di Block era volto a mostrare la falsità del funzionalismo. Se il funzionalismo fosse vero, un organismo come il cervello-cinese deve ammettere la presenza di stati mentali come i qualia. Dato però che la presenza di stati mentali fenomenici in quell’organismo sembra altamente improbabile, Block conclude la falsità del fisicalismo. La forza dell’argomentazione di Block non sta nell’insufficienza teorica delle descrizioni funzionaliste in sé, ma nel far leva sull’assurdità del fatto che un organismo come un cervello-cinese possa avere stati mentali fenomenici, pur avendo in tutto e per tutto le stesse caratteristiche funzionali di un cervello umano. L’esperimento mentale di Block, è stato criticato, dato che il non riuscire a concepire che una mente-cinese possa possedere qualia potrebbe dipendere da una carenza della nostra capacità immaginativa, e che dunque in realtà sia possibile che un tale organismo possa avere esperienze fenomeniche. Famoso è infatti ciò che dice Tye: “Essendo molto più piccoli del sistema Cervello-Cinese, non riusciamo a vedere la foresta per gli alberi e balziamo a una conclusione non autorizzata sull’assenza dei qualia […], non più attendibile della conclusione a cui un minuscolo extraterrestre potrebbe giungere se gli capitasse di materializzarsi all’interno di un cervello umano e concludesse che gli impulsi elettrici che percorrono l’enorme numero di percorsi che lo circondano non supportano l’esperienza[[33]](#footnote-33)”. L’argomento che ho invece proposto, non ha il problema di andare incontro a questa obiezione, in linea di principio infatti, un cervello-cinese potrebbe anche avere stati mentali qualitativi, ma una descrizione funzionalista non potrebbe descriverli in alcun modo. L’esperimento mentale proposto è intuitivamente implicito in quello di Block, in quanto anche quest’ultimo fa emergere le lacune del funzionalismo ammettendo che sia possibile per due organismi essere funzionalmente identici, ma differire per quanto riguarda l’avere proprietà fenomeniche. La differenza consiste nel fatto che la buon riuscita dell’argomento proposto non dipende in alcun modo dalla possibilità che un cervello-cinese, come quello di Block, possa sviluppare una coscienza capace di possedere qualia, e questo mi sembra un buon vantaggio teorico.[[34]](#footnote-34) Come ha anche affermato Kim, una riduzione attraverso una descrizione funzionalista di questo tipo non è in grado di dare conto delle proprietà fenomeniche[[35]](#footnote-35), ed è esattamente questo che cerca di fare l’esperimento mentale. Questo problema verrà lasciato un attimo da parte per essere poi discusso più approfonditamente nel prossimo capitolo.

### Funzionalismo dell’occupante

Il funzionalismo dell’occupante è una forma di funzionalismo che consiste nell’identità dello stato mentale con il suo realizzatore fisico. In questo caso il realizzatore non è più ciò che permette ad una proprietà di ordine superiore di realizzarsi, ma lo stato mentale è letteralmente il realizzatore. Il funzionalismo dell’occupante non è altro che una teoria dell’identità, non di tipo, ma dell’occorrenza. Quello che dobbiamo vedere è se questo tipo di teoria possa resistere all’argomento di Kim, proprio come fanno le teorie dell’identità di tipo, ma che a differenza di queste ultime possa resistere anche all’argomento della realizzabilità multipla. Per rispondere alla prima domanda è utile fare una precisazione sulle teorie dell’identità in generale. Come fa notare Kim[[36]](#footnote-36), le teorie che descrivono il mentale come un qualcosa che accorre ogni qual volta avviene una stimolazione di determinate fibre-C non è una teoria dell’identità, poiché così descritto, il mentale sembra qualcosa di distinto dall’area neuronale stimolata e dunque irriducibile; Indipendentemente da come viene caratterizzata la dipendenza ontologica, se la teoria che si sostiene ammette l’irriducibilità tra mentale e fisico, essa non è compatibile con una teoria dell’identità. Le teorie dell’identità postulano un’identità stretta tra il mentale e il fisico allo stesso modo di come l’acqua è uguale all’H2O, l’acqua infatti non è qualcosa che sopravviene alla sua formulazione molecolare, ma è[[37]](#footnote-37) la sua formulazione molecolare. Il funzionalismo dell’occupante sembra resistere all’argomento di Kim, in quanto essendoci un’identità stretta tra realizzatore e stato mentale, quest’ultimo diventa letteralmente un qualcosa di fisico, non c’è più un dominio mentale irriducibile che rimaneva con i dualisti delle proprietà, ma ne rimane solo uno, quello materiale, il quale però contiene il mentale come sua parte. Se dunque rimane solo un dominio non c’è alcun tipo di sovradeterminazione e nessuna violazione del principio di chiusura causale; la causalità del mentale così facendo è preservata ed è completamente inseribile in una metafisica fisicalista. La tipica critica in cui ci si può imbattere a questo punto è che una teoria dell’identità non può essere utile per salvare la causalità del mentale, in quanto è il mentale in quanto tale che deve avere poteri causali, se lo si riduce al fisico, non si può parlare più veramente di qualcosa di “mentale”. Facendo così però si continua a perdere di vista il vero scopo dell’identità tra mentale e fisico. David Robb, un teorico dell’identità, esprime esattamente il punto con queste parole: ”Ma ancora una volta, questo sembra assumere che i poteri mentali nelle teorie dell’identità hanno una natura duale, […] non c’è tale divisione in un tropo: un potere mentale, la sua natura mentale e la sua natura fisica sono tutte una e la stessa. In questo senso, un tropo mentale è sia completamente mentale sia completamente fisico.”[[38]](#footnote-38) Dunque come riferito anche da Kim[[39]](#footnote-39), non dobbiamo farci confondere dalle parole, anche se parliamo del mentale, ci potremmo riferire senza problemi a qualcosa di fisico, è solo infatti una questione di rivedere i nostri preconcetti di senso comune che tendono a separare in modo ingiustificato il mentale dal fisico. Dunque la differenza tra mentale e fisico non è una differenza ontologica, si tratta di esattamente la stessa cosa detta in termini diversi, sono questi ultimi, ovvero i termini utilizzati, che fanno nascere l’ambiguità. Possiamo infatti chiamare un oggetto con nomi diversi, ma è sempre lo stesso oggetto quello a cui noi ci riferiamo.[[40]](#footnote-40) Per superare l’argomento della realizzabilità multipla con questo tipo di teoria bisogna accettare però una conseguenza particolare. Se in diverse specie di esseri viventi si dà un qualcosa come il dolore, e dato che il dolore si da solo tramite l’opportuna stimolazione di un realizzatore, allora essendoci realizzatori anche molto diversi tra una specie e l’altra[[41]](#footnote-41), dovremmo affermare che esistono anche “dolori” diversi, che per analogia chiamiamo dolore. Poiché un determinato tipo di dolore è dovuto al possesso di un determinato tipo di realizzatore fisico, se ci sono più realizzatori fisici diversi tra loro ci saranno anche “dolori” diversi tra loro. Osservando per esempio i comportamenti di un determinato animale molto diverso da noi, quello che per noi può sembrare dolore richiamandoci alla memoria una particolare sensazione fastidiosa, magari per l’animale sarà una sensazione completamente diversa dalla nostra, ma che svolge la stessa funzione che il dolore svolge per noi; dunque esisteranno tanti tipi di “dolore” tanti quanti saranno i realizzatori in grado di svolgere quella determinata funzione una volta stimolati. Questo discorso può essere anche ampliato all’interno di una stessa specie animale, due persone infatti, in quanto hanno due realizzatori fisici diversi, seppur magari con leggere differenze proveranno un “dolore” diverso; la somiglianza tra sensazioni di dolore tra individui diversi potrebbe essere direttamente proporzionale alla somiglianza tra realizzatori. Così facendo il funzionalismo dell’occupante sembra resistere sia all’argomento di Kim e sia alla realizzabilità multipla. Dunque possiamo salvare sia il mentale, come argomentato da Robb, purché lo si concepisca come completamente riducibile al fisico, sia il suo potere causale, evitando i grossi problemi che portano le teorie dualiste, che siano della sostanza o della proprietà. Nel prossimo capitolo sullo sfondo della teoria appena discussa ci occuperemo di uno dei più grandi problemi riguardanti la riduzione del mentale al fisico, ovvero i qualia, finché infatti gli aspetti fenomenici del mentale non saranno ridotti non si potrà avere una teoria fisicalista in grado di spiegare tutti gli aspetti del mentale.

# 2. Qualia in una prospettiva fisicalista

## 2.1. “Problemi facili” e il “problema difficile”

Nel primo capitolo ho argomentato a favore del funzionalismo dell’occupante, una posizione che si può legittimamente sostenere se si accetta sia l’argomento di Kim, sia l’efficacia causale del mentale. Questa posizione, oltre a resistere al problema della realizzabilità multipla, ci consente di considerare il mentale come un qualcosa di completamente fisico, e questo è indubbiamente un punto fondamentale per chi sostiene una metafisica fisicalista. È opinione abbastanza comune tra i filosofi della mente che qualsiasi tentativo di ridurre il mentale al fisico possa funzionare solo per ciò che David Chalmers ha chiamato i “problemi facili”[[42]](#footnote-42), ovvero tutti quei problemi che sono spiegabili e risolvibili attraverso i risultati delle neuroscienze[[43]](#footnote-43). Per quanto riguarda invece il “problema difficile”, ovvero quello della coscienza, sembra che nessun tipo di teoria possa efficacemente spiegare gli aspetti qualitativi, ovvero i cosiddetti *qualia*. I qualia rappresentano dunque un ultimo ostacolo da affrontare per qualunque teoria fisicalista. Se si riuscissero ad inserire le esperienze qualitative all’interno di teorie riduzioniste, allora potremmo a tutti gli effetti accettare il fisicalismo, ma come si vedrà nei prossimi paragrafi, il problema sembra irrisolvibile.

### 2.1.1. *The explanatory gap*

Il filosofo americano Joseph Levine fu il primo ad introdurre l’espressione “*explenatory gap”*, per riferirsi al divario esplicativo che sembra presentarsi ogni qual volta si provi a ridurre le esperienze qualitative a qualcosa di fisico. In un suo famoso articolo[[44]](#footnote-44), Levine cerca di far notare le differenze che ci sono tra l’identificare il calore con il movimento molecolare e il dolore con la stimolazione di fibre-C. Per quanto riguarda il primo, identificare il calore con il movimento di molecole ci dà una spiegazione completamente esaustiva di che cosa è il calore, per quanto riguarda invece il dolore, la questione è un po’ diversa. Posso infatti benissimo dare una spiegazione del ruolo causale e funzionale del dolore, ma non posso spiegare perché il dolore ha quel particolare carattere qualitativo che ha. Questa incapacità di rendere in termini fisicalisti le esperienze qualitative è stata utilizzata come principale argomento per difendere il fatto che i qualia sono qualcosa di diverso dal fisico. Se infatti sembra impossibile rendere in termini fisicalisti completamente il concetto di dolore, includendone anche gli aspetti qualitativi, sembrano plausibili scenari come il cervello-cinese di Ned Block. Il semplice fatto che una descrizione funzionalista di qualunque tipo non renda conto degli aspetti qualitativi, sembra implicare dunque almeno la possibilità logica dell’esistenza di organismi con una certa organizzazione funzionale, senza che essi presentino proprietà fenomeniche. Come però ho cercato di mostrare nel primo capitolo, nulla implica che effettivamente questi organismi non le possiedano, potrebbe essere semplicemente un nostro problema immaginativo il fatto che non riusciamo a concepire per esempio un cervello-cinese dotato di qualia. Il problema potrebbe riguardare solo l’insufficienza del linguaggio funzionalista nel descrivere stati qualitativi[[45]](#footnote-45), mentre poi nella realtà dei fatti qualsiasi organismo con caratteristiche funzionali potrebbe presentare qualia. Il problema del gap esplicativo è utilizzato come argomento epistemologico, cerca infatti di mostrare che nessun linguaggio fisicalista può descrivere i qualia. Questo argomento è stato utilizzato dai non riduzionisti, poiché l’impossibilità di utilizzare termini fisicalisti per descrivere i qualia può portare a pensare che questi ultimi non siano descrivibili proprio perché non sono fisici. Bisogna specificare però che questo argomento, non essendo un argomento metafisico, non porta direttamente alla conclusione a cui vorrebbero arrivare i non riduzionisti. Cercherò nei prossimi paragrafi di esporre un esperimento mentale che intende mostrare, proprio come il cervello-cinese di Ned Block, che i qualia non possono essere ridotti a qualcosa di fisico; questo esperimento mentale, a differenza dell’argomento di Levine è presentato come un argomento metafisico, ovvero cerca di stabilire che i qualia non sono qualcosa di fisico. Farò un’analisi di questo esperimento mentale, per poi mostrare tramite l’elaborazione di un argomento, che l’esperimento mentale così costruito non raggiunge il suo obiettivo.

## 2.2 Esperimento mentale degli zombie

Questo esperimento mentale è strettamente legato alla figura di David Chalmers, e l’idea centrale è che tutto ciò che è concepibile è anche metafisicamente possibile. L’esperimento mentale può essere riassunto nel modo seguente: si provi ad immaginare l’esistenza di individui esattamente uguali a noi, atomo per atomo, che si comportano esattamente come noi in ogni nostra azione quotidiana, ma che non presentano qualia. Questi individui, chiamati “zombie”, prendono la metro, mangiano, guardano la televisione e discutono di filosofia proprio come noi, dunque sono simili a noi sotto ogni aspetto, tranne per il fatto che non presentano proprietà fenomeniche[[46]](#footnote-46). Se questo scenario è concepibile, allora è anche possibile. Se è possibile, significa che la presenza di qualia non è necessaria data la presenza del mondo fisico, dunque i qualia non sono riducibili al fisico e il fisicalismo riduzionista è falso.

L’argomento può essere schematizzato nel modo seguente:

1. È possibile concepire un mondo di individui fisicamente uguali a noi, ma privi di esperienze fenomeniche (zombie).
2. Tutto ciò che è concepibile è anche possibile.
3. Un mondo-zombie è possibile [da (1) e (2)].
4. Se un mondo-zombie è possibile allora il fisicalismo è falso.
5. Il fisicalismo è falso [da (3) e (4) per modus ponens].

L’argomento è valido, ma non sembra molto chiaro se le premesse (2) e (3) sono vere. I prossimi paragrafi saranno utilizzati per esporre le principali ragioni di Chalmers per accettare (2) e (3), solo così facendo si può comprendere meglio la portata dell’argomento.

### 2.2.1 Diversi sensi di possibilità

Prima di iniziare a trattare la spiegazione delle premesse (2) e (3) è utile distinguere almeno tre[[47]](#footnote-47) sensi diversi di possibilità[[48]](#footnote-48). Il primo tipo di possibilità è la cosiddetta “possibilità logica”, dire che qualcosa è logicamente possibile significa dire che nel concepirlo non vi è alcun tipo di contraddizione. La proposizione “le mucche volano” non implica alcun tipo di contraddizione per quanto riguarda l’aspetto puramente logico, dunque è logicamente possibile; la proposizione “le mucche volano e non volano nello stesso momento” è una contraddizione logica, dunque è logicamente impossibile. Il secondo tipo di possibilità è la “possibilità metafisica”, dire che un qualcosa è “metafisicamente possibile” significa ammettere che quello stato di cose considerato avrebbe potuto darsi in accordo con le leggi metafisiche. A differenza della possibilità logica, la possibilità metafisica riguarda direttamente la realtà[[49]](#footnote-49), ad esempio sembra metafisicamente impossibile l’esistenza un oggetto senza alcun tipo di proprietà mentre dire che le mucche volano è metafisicamente possibile, in quanto non c’è nulla di contraddittorio nel considerare un ente con la proprietà di volare; è utile sottolineare che tutto ciò che è metafisicamente possibile è anche logicamente possibile. Il terzo tipo di possibilità che ci interessa è la “possibilità fisica[[50]](#footnote-50)”, dire che un qualcosa è fisicamente possibile significa dire che quella possibilità è in accordo con le leggi fisiche del mondo attuale. Tutto ciò che è fisicamente possibile è anche logicamente e metafisicamente possibile, ma non viceversa, una mucca che vola per esempio è fisicamente impossibile, ma metafisicamente e logicamente possibile. Dire che è possibile che sulla Terra ci sia una forza di gravità doppia di quella del mondo attuale, è fisicamente possibile, in quanto basta immaginare che la Terra abbia una massa doppia; in fisica infatti, la forza di attrazione gravitazionale di un oggetto è direttamente proporzionale alla sua massa.

### 2.2.2 Argomento della concepibilità

L’argomento degli zombie non è altro che un’esemplificazione di quello che Chalmers ha chiamato “argomento della concepibilità”. L’argomento della concepibilità può essere formalizzato come segue[[51]](#footnote-51):

1. p˄ ̴ q è concepibile.
2. Se p˄ ̴ q è concepibile, allora p˄ ̴ q è metafisicamente possibile.
3. Se p˄ ̴ q è metafisicamente possibile, allora il fisicalismo è falso.
4. Il fisicalismo è falso.

Dove “p” rappresenta tutte le verità fisiche del mondo e ” ̴ q” l’assenza di proprietà fenomeniche. La prima premessa afferma che la congiunzione di tutte le verità fisiche sul mondo (p) insieme all’assenza di proprietà fenomeniche è concepibile, la seconda premessa è un condizionale che afferma che la concepibilità di questa congiunzione ne implica la possibilità metafisica e la terza è un altro condizionale che porta alla conclusione che il fisicalismo è falso, in quanto non sarebbe metafisicamente necessario il fatto che le proprietà mentali sopravvengano alle proprietà fisiche. Da notare è il fatto che sia l’argomento degli zombie, sia questo appena presentato, si basano sullo stesso presupposto, ovvero il fatto che tutto ciò che è concepibile è anche metafisicamente possibile. Prima di portare delle obiezioni contro l’esperimento mentale, è utile definire come Chalmers sta usando il concetto di “concepibilità” e come questa possa implicare la possibilità metafisica.

### 2.2.3 Concepibilità

Chalmers distingue due sensi di “concepibilità[[52]](#footnote-52)”, la concepibilità “primaria” (o “1-concepibile”) e la concepibilità “secondaria” (o “2-concepibile”). Per spiegarne la differenza in modo adeguato è utile tenere presente la seguente proposizione: “l’acqua è H2O”. Questo tipo di proposizione è quello che Kripke definisce “analitico a posteriori”; a posteriori in quanto l’unico modo per sapere che l’acqua ha una certa composizione molecolare è attraverso l’osservazione con determinati strumenti scientifici, e analitica in quanto l’acqua è per Kripke un designatore rigido, ovvero in tutti i mondi possibili in cui vi è acqua essa non può che avere la formulazione molecolare H2O; l’acqua dunque è necessariamente H2O. Prendiamo per esempio il famoso esperimento mentale di Hilary Putnam di “Terra Gemella[[53]](#footnote-53)”. Supponiamo dunque l’esistenza di un pianeta esattamente uguale alla Terra, con abitanti, fiumi, mari e laghi e dove però la sostanza liquida che compone questi ultimi, pur avendo le stesse identiche caratteristiche visibili dell’acqua, ha una composizione molecolare diversa, per esempio XYZ. Avendo presente questo scenario, non è 2-concepibile che l’acqua non sia H2O, in quanto solo la sostanza che ha quella composizione molecolare può essere chiamata “acqua”; al limite si può dire che in Terra Gemella sembra che l’acqua non sia H2O, ma quella sostanza che sembra acqua, non è in realtà acqua. Affermare dunque che è concepibile che l’acqua non sia H2O in questo caso, sarebbe come affermare che l’H2O non è H2O, il che è un’evidente contraddizione. La nozione di concepibilità primaria, ci permette di affermare che è concepibile che l’acqua non sia H2O, in quanto il significato della parola “acqua” dipende dal mondo che consideriamo attuale. Se noi ci trovassimo su Terra Gemella, la sostanza che bagna i mari e i fiumi la chiameremmo comunque acqua, anche se ha XYZ come sua composizione molecolare. In quel mondo inoltre, la proposizione “l’acqua è XYZ” verrebbe considerata analitica a posteriori, dunque in ogni mondo possibile in cui vi è una sostanza con la composizione molecolare XYZ ci sarebbe acqua (prendendo sempre come punto di riferimento il mondo considerato attuale, in questo caso Terra Gemella, in cui la sostanza chiamata acqua è XYZ); in questo senso si può dire che è concepibile che l’acqua non sia H2O, siccome sarebbe come dire che XYZ non è H2O, il che non implica nessuna contraddizione. Dunque sembra si possa dire che l’acqua non è H2O è 1-concepibile, ma non 2-concepibile. Se si pensa bene, da questo punto di vista, sembra non ci sia alcun problema ad ammettere l’esistenza di un mondo metafisicamente possibile in cui l’acqua che riempie i laghi e i mari abbia XYZ come composizione molecolare, dunque si potrebbe dire che è “primariamente possibile” (o 1-possibile) che l’acqua sia XYZ. Ammettere questo significa confermare l’esistenza di un “collegamento” tra la concepibilità primaria (o possibilità logica) con la possibilità metafisica, ammettendo così che la prima implica la seconda. Tutte queste precisazioni vengono ritenute da Chalmers essenziali per spiegare la legittimità nell’ammettere che la semplice possibilità logica possa implicare la possibilità metafisica, in quanto sembra che la semplice concepibilità possa dipendere dal mondo che prendiamo di riferimento come attuale.

### 2.2.4 Esperimento mentale contro la concepibilità di un mondo-zombie

Date queste premesse[[54]](#footnote-54), proviamo ora ad osservare in maniera più dettagliata cosa significa dire che un mondo-zombie è concepibile. Per essere almeno logicamente possibile, un mondo-zombie non deve contenere alcun tipo di contraddizione, e così come spiegato nelle premesse dell’argomento della concepibilità, deve essere esattamente uguale al nostro mondo dal punto di vista fisico. Se riusciamo ad immaginarci questo mondo senza la presenza di qualia, sembra che dovremmo inferirne anche la possibilità metafisica. A tal proposito, volevo proporre un esperimento mentale che può essere considerato come una versione particolare di quello zombie. Lo scopo dell’esperimento mentale è mostrare che, almeno intuitivamente, la concepibilità di uno scenario-zombie sembra concepibile solo ad un costo molto alto. Prima di esporlo è utile precisare esattamente come dovrebbe essere strutturato un mondo-zombie per confutare il fatto che i qualia sopravvengano metafisicamente al fisico. Per farlo, espliciterò nuovamente le premesse dell’esperimento mentale di Chalmers. Questo mondo-zombie deve essere esattamente simile al nostro, in tutti i suoi aspetti fisici e micro-fisici. Gli abitanti di questo mondo, che d’ora in avanti chiamerò “controparti-zombie”, devono essere capaci di compiere esattamente tutte le stesse cose che normalmente una persona dotata di qualia farebbe, con la differenza che queste controparti-zombie non hanno qualia. Assumeremo anche, in accordo con Chalmers, che la possibilità logica (o concepibilità) implichi la possibilità metafisica. Se un tale scenario dunque è concepibile, allora esiste un mondo metafisicamente possibile in cui esistono controparti-zombie esattamente come noi, ma senza qualia. La possibilità metafisica di uno scenario del genere mostrerebbe che l’esistenza di qualia non è necessaria in un mondo fisicamente uguale al nostro e che dunque essi non sono fisici, perché altrimenti sopravverrebbero direttamente sul fisico. Precisati i termini per la buona riuscita dell’esperimento mentale degli zombie, passiamo al proporre il nuovo esperimento mentale. Proviamo ad immaginarci un mondo possibile in cui sono presenti tutti i fatti fisici e micro-fisici del nostro mondo, ma in cui non sono presenti persone, almeno per il momento. Tutto ciò sembra poter essere fatto senza alcun tipo di contraddizione, e possiamo dire che fino ad ora, tranne per l’assenza di persone, questo mondo è uguale al mondo-zombie; dunque un mondo fisicamente uguale al mondo attuale, senza la presenza di persone è metafisicamente possibile. Assumiamo anche che effettivamente i qualia siano un qualcosa di distinto dalla costituzione fisica del mondo, come sembra suggerirci Chalmers. Supponiamo ora di popolare finalmente questo mondo, che è fisicamente uguale al mondo attuale, e supponiamo che gli individui che lo popolano possiedano qualia, proprio come noi (teniamo presente però che il mondo che stiamo osservando deve avere almeno qualcosa di diverso dal mondo attuale, che sia anche il fatto che la controparte di me in quel mondo non faccia filosofia, altrimenti per la legge di Leibniz[[55]](#footnote-55) staremmo parlando del mondo attuale); avremmo in questo modo un mondo che è come il nostro. Osserviamo ora le nostre controparti di questo mondo durante la loro vita, mentre lavorano, mangiano, studiano, e supponiamo che la mia controparte vada a comprare dei vestiti in un negozio. Immaginiamo che entrando in questo negozio, la mia controparte veda delle magliette appese e che sia indecisa tra due magliette dello stesso modello e misura, una di colore blu ed una di colore rosso, e che infine scelga quella di colore blu. Ora proviamo ad immaginare che proprio in quel momento, poco prima di scegliere, tutti i qualia in quel mondo spariscano e dunque la mia controparte non avesse più le esperienze qualitative del rosso e del blu. Dato che sembra che scegliamo una preferenza su un colore in base all’effetto che ci da, sembra difficile immaginare che la mia controparte possa ancora scegliere quella particolare maglietta. Se un mondo-zombie fosse concepibile allora la mia controparte avrebbe continuato senza alcun tipo di problema a prendere la maglia blu, ma sembrano sorgere delle difficoltà in più a concepire un tale scenario. Intuitivamente ci verrebbe da dire che la mia controparte non possa più scegliere in base alla preferenza di colore, ma sembra che nulla implichi logicamente che non possa comunque afferrare la maglietta blu, pagare e uscire dal negozio, proprio come farebbe un robot, o almeno così sembra. Bisogna ricordare però che la scelta della maglietta blu, è dovuta ad una preferenza dell’effetto che ci da il blu rispetto a quello che ci da il rosso. Dunque l’azione sembra in un certo senso essere legata ad aspetti qualitativi della coscienza, a patto che questo movimento non fosse già in un certo senso predeterminato, proprio come è predeterminato un robot programmato a muoversi in un certo modo. Dopotutto l’esperimento mentale degli zombie ci chiede di immaginarci proprio una situazione simile, con la differenza che in quel mondo i qualia non sono mai stati presenti. Ma intuitivamente, verrebbe da dire che ciò sarebbe possibile solo nel caso in cui le controparti-zombie fossero “programmate[[56]](#footnote-56)” a muoversi in quel determinato modo per tutta la durata della loro presenza in quel mondo. Se questa intuizione si dimostrerà vera, tutte le scelte fatte dalla mia controparte-zombie non sarebbero vere e proprie scelte, ma sarebbero solo parte di un programma già scritto. Ma se non ci fosse questa “programmazione” le controparti-zombie non potrebbero replicare totalmente i comportamenti delle controparti del mondo attuale, siccome molti di questi, almeno intuitivamente, dipendono dalla presenza di qualia. L’esperimento mentale degli zombie funziona solo se si presume che tutte le scelte fatte nel mondo attuale non dipendono dall’esistenza di qualia. Se anche solo la presenza di un quale può determinare gli esiti di un certo comportamento, allora l’esperimento non funziona, perché senza qualia quel determinato comportamento non sarebbe accaduto. Se è vera infatti l’intuizione che c’è dietro l’esperimento mentale, i comportamenti che riguardano scelte che dipendono dal particolare effetto che ci danno alcune cose come i colori, allora senza i qualia non potrebbero esserci certi nostri comportamenti. Se invece nessun quale è responsabile delle azioni umane, e se dunque il mondo-zombie è concepibile, allora siamo totalmente determinati dalla nostra costituzione fisica, una conseguenza che non tutti accetterebbero[[57]](#footnote-57). Intuitivamente, come cerca di mostrare l’esperimento mentale, sembra che alcune nostre azioni dipendano dall’esistenza di qualia e che questi ultimi, siano in qualche modo responsabili di alcune nostre azioni, come la scelta dovuta ad una preferenza tra due oggetti di colore diverso. Se non avessi mai avuto le esperienze qualitative del rosso o del blu, sembra che non avrei mai potuto preferire il blu. Compito del prossimo paragrafo sarà proprio cercare di dimostrare questo, ovvero cercare di argomentare a sostegno dell’intuizione che sembra esserci dietro la buona riuscita dell’esperimento mentale, vale a dire, dare una spiegazione del fatto che alcuni comportamenti necessitino l’esistenza di qualia. Se questa affermazione è vera, solo un mondo di controparti-zombie predeterminate dal fisico sarebbe logicamente non contraddittorio, siccome un mondo-zombie non predeterminato non potrebbe garantire che tutte le sue controparti-zombie possano replicare tutte le azioni di persone dotate di qualia. Ma l’esperimento mentale degli zombie non ci dice nulla di questa predeterminazione, ci chiede semplicemente di immaginare un mondo in cui senza qualia le controparti-zombie fanno esattamente le stesse cose delle controparti del mondo attuale, ma se la tesi del prossimo paragrafo è vera, ciò sarà impossibile senza accettare la predeterminazione fisica. Come sosterrò nella parte finale dell’elaborato, se la tesi secondo cui alcuni comportamenti necessitano qualia è vera e se accettiamo anche l’argomento di Kim, non si può far altro che ammettere che i qualia devono essere qualcosa di fisico, seppur probabilmente non potranno mai essere descritti in un linguaggio fisicalista.

### 2.2.4.1 Precisazione su qualia e colori

In letteratura e nel corso degli anni il termine “qualia” è stato utilizzato in modi diversi[[58]](#footnote-58). In questa sottosezione cercherò di specificare nel modo migliore possibile a che cosa mi riferisco quando utilizzo il termine “qualia” in questo elaborato. Per avere un’idea di come viene utilizzato il termine qualia è utile tenere a mente che cosa si intende quando si parla di “spettro invertito”. L’ipotesi dello spettro invertito è un caso particolare di “qualia invertiti”, e in particolare si riferisce ai qualia che riguardano il colore. Uno scenario di spettro invertito è uno scenario in cui i colori che si osservano sono, per esempio, invertiti con i loro complementari, dunque il sole apparirà di colore viola, l’erba di colore rosso e il cielo di colore arancione[[59]](#footnote-59). In questo senso due persone pur indicando lo stesso oggetto potrebbero avere due esperienze qualitative diverse e dunque vedranno due “colori” diversi. Supponiamo ci siano due persone, che per distinguerle chiameremo “Marco” e “Noemi”, Marco ha un’esperienza qualitativa del colore “normale[[60]](#footnote-60)” e Noemi ha lo spettro dei colori invertito rispetto a quello di Marco. Sia Marco che Noemi quando parlano dell’erba riferiscono che essa è di colore verde, ma se Marco potesse anche solo per un istante avere la stessa esperienza in prima persona che sta vivendo Noemi dirà che l’erba che vede Noemi è in realtà rossa. Viceversa se Noemi potesse anche solo per un momento guardare un pomodoro maturo come lo guarderebbe Marco, direbbe che il pomodoro è verde. Questo mostra che quando si parla di colori, non si può fare altro che identificare il colore con l’esperienza qualitativa che si prova, non si può infatti nemmeno concepire il “rosso” se non attraverso l’esperienza qualitativa che vi si associa. Questo è esattamente il modo in cui sto trattando i qualia dei colori[[61]](#footnote-61) in questo elaborato, e questo può far capire la difficoltà che può esserci nel concepire una scelta tra due colori diversi, come l’esperimento mentale intende fare, se i qualia non sono presenti in quel mondo.

## 2.3 Analisi della possibilità che i qualia possano essere causa di comportamenti

Prima di iniziare la trattazione di questa sezione è utile tenere a mente i risultati del paragrafo precedente. L’esperimento mentale ha mostrato che i qualia potrebbero essere responsabili di alcune nostre azioni, in particolare quelle che riguardano alcune scelte. Accettare questo fatto, significa accettare che in un mondo-zombie le controparti non possono replicare tutte le nostre azioni e che dunque un tale mondo non è concepibile. Si può replicare che un mondo di questo genere è invece concepibile, in quanto basta immaginare che le controparti-zombie siano in un certo senso predeterminate dalla loro costituzione fisica. Ammettere questo ha però come conseguenza, quella di ammettere che la coscienza fenomenica non ha alcun tipo di funzione nel determinare anche solo minimamente qualche nostro comportamento. Se infatti un mondo-zombie in cui i comportamenti delle controparti sono fisicamente predeterminate è concepibile, allora deve essere anche metafisicamente possibile. Se è metafisicamente possibile, si deve ammettere che anche le nostre azioni nel mondo attuale potrebbero essere completamente predeterminate dal fisico. Ammettere invece che la coscienza abbia qualche influenza causale nel nostro mondo, dopo che basterebbe il mondo fisico per spiegare ogni azione umana, significherebbe ammettere un caso di sovradeterminazione causale. Se però accettiamo, come in questo elaborato stiamo facendo, l’argomento di Kim, ciò è impossibile e l’unica causa responsabile delle nostre azioni sarebbe la nostra costituzione fisica, quindi o la coscienza è solamente epifenomenica o è fisica. Ma se i qualia si dimostrano causalmente efficaci e dunque la coscienza si dimostra, almeno in un certo senso, responsabile di alcune azioni umane, allora i primi devono essere fisici. Per rendere più evidente la portata dell’argomento è utile schematizzarlo come segue:

1. O i qualia sono responsabili di azioni umane o i qualia non sono responsabili di azioni umane.
2. Se i qualia sono responsabili di qualche azione umana, allora un mondo-zombie non fisicamente predeterminato, non è concepibile.
3. Se i qualia non sono responsabili di qualche azione umana, allora un mondo zombie fisicamente predeterminato, è concepibile.
4. Solo una premessa tra (2) e (3) può essere vera [dal principio del terzo escluso].
5. Solo un mondo-zombie fisicamente predeterminato è concepibile [da (2) e (3)].

Indipendentemente dal fatto, di quale tra (2) o (3) sia vera, possiamo completare l’argomento in questo modo:

1. Se un mondo-zombie fisicamente predeterminato è concepibile, allora i comportamenti delle controparti sono completamente determinati dalla loro costituzione fisica. [da (5)].
2. Se un mondo-zombie fisicamente predeterminato è concepibile, allora è metafisicamente possibile. [accettando che la concepibilità implichi la possibilità metafisica]
3. Se un mondo-zombie fisicamente predeterminato è metafisicamente possibile, allora la fisica del nostro mondo è sufficiente a spiegare tutti i nostri comportamenti [da (5)].
4. Nel mondo attuale sono presenti esperienze qualitative.
5. O le esperienze qualitative sono inerti o sono qualcosa di fisico [da (8) e (9)].
6. Se i qualia determinano causalmente qualche azione umana (se dunque (2) è vera), allora sono fisici. [da (2) e (10)].

La buona riuscita dell’argomento dipende dunque dalla presunta capacità dei qualia di determinare causalmente le azioni umane. Se i qualia si dimostrano inefficaci dal punto di vista causale, come dovrebbe ammettere chi sostiene l’esperimento mentale degli zombie, l’esperienza qualitativa è solo epifenomenica e solo in questo caso, potrebbero essere qualcosa di non riducibile al fisico[[62]](#footnote-62).

### 2.3.1 Spiegazione controfattuale

Questo paragrafo sarà incentrato sul dare contenuto all’intuizione emersa nell’esperimento mentale del paragrafo 2.2.4, ovvero il fatto che i qualia sembrano essere responsabili di alcune nostre scelte nel nostro quotidiano. Queste scelte si caratterizzano per il fatto che sembra che dipendano da una nostra preferenza tra alcune esperienze qualitative. Fare una scelta basata su una preferenza tra esperienze qualitative diverse consiste nell’avere ben chiare almeno due esperienze qualitative e sceglierne una, in base a quale delle due viene considerata più piacevole o appetibile. Esempi sono quando, come presentato nell’esperimento mentale, scegliamo tra due magliette di colore diverso o per esempio quando decidiamo cosa mangiare. Supponiamo di entrare in una gelateria[[63]](#footnote-63), che contiene dodici gusti di gelato diversi e immaginiamo che i nostri gusti preferiti siano menta, cioccolato e fragola. Se questi gusti sono presenti in gelateria, inevitabilmente tenderemo a prendere quei gusti. Ma anche solo per affermare che menta, cioccolato e fragola sono i nostri gusti preferiti dobbiamo avere ben presente l’esperienza qualitativa che in passato ci ha dato l’aver mangiato il gelato alla menta, al cioccolato o alla fragola. Immaginiamo, inoltre, che fortunatamente la gelateria ha tutti e tre i nostri gusti di gelato preferiti, ma supponiamo che tre palline di gelato costino troppo e possiamo permetterci solo due gusti. In questo caso, ognuno di noi, proverà a rievocare alla memoria i tre gusti per valutare quale esperienza qualitativa tra quella della menta, del cioccolato e della fragola preferiamo ripetere, sapendo bene che solo due di queste tre potrà effettivamente ripetersi. La scelta, in breve, non è altro che una scelta tra esperienze qualitative, sulla base del ricordo di quelle passate, per replicarle nuovamente; come in questo caso i qualia di alcuni particolari gusti di gelato. Messa così, sembra impossibile affermare che degli zombie senza esperienze qualitative possano scegliere, per esempio, gusti di gelato diversi, contando sul fatto che le scelte si basano proprio sull’effetto particolare che un determinato gusto ha su di noi, ovvero la sua esperienza qualitativa. Il caso particolare della scelta del gusto di gelato può essere spiegato con un condizionale controfattuale del seguente tipo: “se non avessi avuto[[64]](#footnote-64) l’esperienza qualitativa del gelato alla menta, non avrei scelto il gusto menta”. L’esempio può essere esteso a qualsiasi altro tipo di scelta che ha a che fare con qualia, per generalizzare possiamo formulare il seguente condizionale controfattuale: “se non avessi esperienze qualitative, allora non potrei fare alcuni tipi[[65]](#footnote-65) di scelte”, se questo controfattuale è vero allora implica che se non avessi qualia, molti miei comportamenti non potrebbero avvenire, in quanto molti di questi derivano da scelte. Dunque se questo controfattuale è vero, mette in crisi la possibilità di concepire un mondo-zombie del tipo di Chalmers.

### 2.3.2 Soluzione basata su un controfattuale

Basare un argomento su una premessa costruita con un condizionale controfattuale ha il problema che tutti i controfattuali normalmente hanno, ovvero la difficoltà di stabilire il loro valore di verità. Un controfattuale è un condizionale in cui l’antecedente nel mondo attuale è falso, ma assumendo che sia vero cerchiamo di stabile un certo conseguente che potrebbe seguire. Per valutare quando un conseguente potrebbe seguire a partire da un antecedente falso, è utile tenere presente i lavori di David Lewis[[66]](#footnote-66). Sebbene alcune parti della sua teoria dei controfattuali hanno suscitato qualche controversia[[67]](#footnote-67), possiamo utilizzarla per stabilire se il nostro controfattuale è vero. Lewis presenta un controfattuale come un tipo particolare di condizionale che è legato ad un condizionale stretto. I condizionali stretti non sono altro che condizionali materiali della forma “se p, allora q”, a cui viene fatto precedere un operatore modale, e questi ultimi si comportano come quantificatori che operano ristrettamente su mondi possibili[[68]](#footnote-68). In particolare, l’operatore di necessità opera come un quantificatore universale su tutti quei mondi che soddisfano quella determinata ristrettezza, mentre l’operatore di possibilità opera come un quantificatore esistenziale su qualche mondo che soddisfa quella determinata ristrettezza. Esempi di ristrettezza sono i diversi sensi di possibilità descritti nel paragrafo 2.2.1. Senza andare troppo nei dettagli, a cui rimando la lettura dell’opera, Lewis afferma che un controfattuale è vero nel mondo attuale se e solo se, o non ci sono mondi in cui è presente l’antecedente, oppure quando mondi, in cui sono veri sia l’antecedente e sia il conseguente, sono “più vicini” di mondi in cui l’antecedente è vero e il conseguente è falso (ovvero in quei mondi in cui il condizionale materiale con quelle proposizioni è falso). Per mondi più vicini, si intendono mondi più simili, e questi sono quelli in cui ci sono meno differenze con il mondo attuale, rispetto ad altri mondi[[69]](#footnote-69). Per rendere più chiare le condizioni di verità di un controfattuale è utile fornire un esempio. Consideriamo il seguente condizionale controfattuale: “se Putin non fosse presidente della Federazione Russa, allora non ci sarebbe la guerra in Ucraina”. Il mondo attuale, è un mondo in cui Putin è presidente della Federazione Russa e in cui c’è la guerra in Ucraina. Se non esistessero mondi in cui Putin non è presidente della Federazione Russa allora il controfattuale sarebbe vero a prescindere, se infatti un condizionale materiale ha l’antecedente falso, sicuramente il condizionale è vero. Se invece ci sono mondi in cui Putin non è presidente della Federazione russa, allora bisogna considerare, tra quei mondi, se sono più vicini quelli in cui non c’è la guerra o quelli in cui c’è la guerra. Se i mondi più vicini a quello attuale, in cui Putin non è presidente della Federazione Russa, e in cui non c’è la guerra, sono più vicini di quelli in cui, Putin non è presidente e in cui c’è la guerra, allora il controfattuale è vero. Per stabilire quali mondi sono quelli più vicini ci basta capire in quali, il cambiamento apportato, provocherebbe maggiori differenze dal mondo attuale. Torniamo al nostro controfattuale, “se non avessi avuto l’esperienza qualitativa del gusto menta, allora non avrei potuto scegliere il gelato alla menta”. Esso risulterà vero se e solo se nei mondi in cui non ho avuto l’esperienza qualitativa della menta e in cui non potrei scegliere il gusto menta sono più vicini di quelli in cui pur non avendo mai avuto l’esperienza qualitativa del gusto menta scelgo il gelato alla menta. Andiamo a considerare allora quei mondi in cui non ho mai avuto l’esperienza qualitativa della menta. Questi mondi, includono sia quelli in cui magari non ho mai assaggiato la menta, ma dove ho potuto avere le esperienze qualitative di altri gusti, sia quelli in cui non esistono qualia. Sicuramente, i mondi in cui l’unica esperienza qualitativa che non ho mai avuto è la menta sono mondi più vicini del mondo attuale rispetto a quelli in cui non vi sono qualia. In questi mondi più vicini, il controfattuale ha più possibilità di risultare falso, siccome ci sono molti scenari in cui pur non possedendo il quale della menta, prendo lo stesso il gelato alla menta. Per esempio, in questo mondo, degli amici che hanno già assaggiato la menta potrebbero consigliarmi di provarla, o magari il colore verde del gelato mi attirerebbe poiché è il mio colore preferito, o ancora la gelateria potrebbe essere famosa in tutto il paese per il gusto alla menta e allora potrei convincermi a provarlo. Si noti che in tutti questi casi c’è sempre bisogno della presenza di qualia per fare in modo che la mia controparte nei diversi mondi assaggi la menta. Nel primo caso i miei amici mi consigliano di assaggiare la menta, siccome per loro, l’esperienza qualitativa della menta è stata piacevole e vorrebbero farla provare anche a me. Nel secondo caso è la preferenza del quale verde rispetto ad altri qualia a determinare la mia scelta, e nella terza è la possibilità della piacevole esperienza qualitativa che potrei provare a condizionarmi nel decidere di assaggiare il gusto, persuaso dal fatto che la gelateria è famosa per quel gusto in tutto il paese. Quindi finché nei mondi compaiono almeno altri qualia, il condizionale può essere falso. Proviamo ora a valutare nei mondi in cui non sono presenti qualia, in quali è possibile che la mia controparte assaggi il gelato alla menta. I casi più evidenti sono quelli in cui tutte le controparti si muovono in modo programmato, proprio come fossero dei robot. In linea di principio, se noi programmassimo alla perfezione otto miliardi di robot, in modo tale da replicare in tutto e per tutto i nostri comportamenti, potremmo senza troppi sforzi, immaginare dei robot che vanno a prendere il gelato. In questo senso, l’unico modo per immaginare un mondo-zombie è uno scenario in cui tutti i loro movimenti sono programmati, proprio come quelli di un mondo di robot. Tenendo a mente questi punti, il condizionale controfattuale “se non ci fossero i qualia, allora molti nostri comportamenti non potrebbero darsi” risulta vero, in quanto i mondi in cui pur non essendoci qualia, i comportamenti si danno, sono più lontani dei mondi in cui pur non essendoci qualia, i comportamenti non si danno. E la differenza consiste nel fatto che per permettere a degli zombie di replicare tutte le mie azioni, io devo fare in modo che la fisica di quel mondo sia sufficiente a predeterminare tutte le loro azioni e che queste siano uguali a quelle del mondo attuale. Ma se come abbiamo mostrato, alcuni comportamenti dipendono dalla presenza di qualia, significa che, o la fisica del nostro mondo non basta per spiegare tutti questi comportamenti, per cui dovremmo cambiarla per predeterminare tutte le azioni umane[[70]](#footnote-70), o che i qualia fanno parte della fisica del nostro mondo. In modo assolutamente bizzarro e contraddittorio, l’unico modo per concepire un mondo-zombie con esattamente le nostre verità fisiche è concepirlo con la presenza di qualia, dunque non sarebbe un mondo-zombie. Per riassumere l’argomento:

1. Se il controfattuale è vero, allora i qualia sono responsabili di alcune azioni umane
2. Il controfattuale è vero, quindi i qualia sono responsabili di alcune azioni umane [modus ponens da (1)]
3. I qualia sono distinti dal fisico [assunzione]
4. La fisica del mondo attuale non è in grado di rendere conto di tutte le azioni umane [da (2) e (3)]
5. Concepire un mondo-zombie con le stesse verità della fisica del mondo attuale è un mondo in cui alcune azioni non potrebbero esserci [da (4)]
6. Per far si che le azioni del mondo-zombie possano replicare tutte le azioni del mondo attuale bisogna cambiarne la fisica, in modo tale che le controparti siano “programmate” come dei robot, oppure bisogna includere i qualia nella fisica del nostro mondo
7. L’unico modo di concepire un mondo fisicamente uguale al nostro è includere i qualia nel mondo fisico. [da (6)]

Questo argomento dovrebbe mostrare che dunque l’esperimento mentale degli zombie, così come è costruito, non può funzionare.

### 2.3.3 Risultati teorici raggiunti

Da come si è potuto osservare, l’unico mondo-zombie che inizialmente sembra essere fisicamente uguale al nostro è un mondo in cui tutti i comportamenti delle controparti zombie sono fisicamente predeterminati. Se accettiamo ciò, la fisica del nostro mondo sarebbe in grado di spiegare tutti i nostri comportamenti. Se però si accetta l’argomento basato sui controfattuali, si deve ammettere che i qualia sono responsabili di alcune azioni umane nel mondo attuale. Se non accettiamo la sovradeterminazione[[71]](#footnote-71), allora la fisica del nostro mondo non è in grado di spiegare tutte le azioni umane, quindi un mondo-zombie fisicamente uguale al nostro non è concepibile. Dire che l’unico mondo-zombie concepibile è un mondo-zombie predeterminato dal fisico, implica che la fisica di questo mondo deve essere diversa da quella del mondo attuale. Quindi l’unico mondo fisicamente uguale al nostro, in grado di rendere conto di tutte le azioni umane, lo si può concepire solo se si includono i qualia tra le entità fisiche. Come però ho già sottolineato, questo non significa che siamo in grado di descrivere i qualia con un linguaggio fisicalista, ma significa che potremmo inserirli in un’ ontologia fisica. Come a tutti può sembrare, i qualia sono entità effimere ed evanescenti, non stupisce che si faccia fatica a concepirli come entità fisiche di tutto rispetto, soprattutto se si pensa che normalmente la fisica abbia a che fare con entità più concrete. Questo però non deve indurre a pensare che considerarle come fisiche vada a snaturarle della loro particolarità, significa semplicemente guardarle a partire da un altro punto di vista. Come è stato già detto per quanto riguarda le proprietà mentali non fenomeniche, ridurre la coscienza completamente a qualcosa di fisico, non significa privarla della propria importanza, ma al contrario significa trattarla con l’importanza che merita. Così facendo infatti si può benissimo affermare, senza cadere nei problemi dei dualismi, che ha un’efficacia causale proprio come potrebbe averla una qualsiasi altra entità fisica. Risultando così un’entità fisica, non si violerebbe in alcun modo il principio di chiusura causale del mondo fisico, resistendo così anche all’argomento di Kim, che le uniche relazioni causali che ammette sono quelle del dominio fisico.

## 2.4 Considerazioni finali

In questo elaborato, ho cercato di presentare le diverse difficoltà che si affrontano nel cercare di rendere conto della causalità del mentale. Nel primo capitolo, ho mostrato che accettando l’argomento di Kim, le uniche relazioni causali che possono esserci sono quelle che avvengono nel dominio fisico, dunque o si concepisce la mente come un qualcosa di fisico o come semplicemente epifenomenica. Concepirla come epifenomenica significa privarla completamente di efficacia causale, quindi ho preferito concentrarmi sui modi in cui è possibile ridurre la mente a qualcosa di fisico. Ho esposto che il problema della realizzabilità multipla non permette di adottare una teoria dell’identità di tipo, dunque mi sono concentrato sulle alternative date dal funzionalismo, esponendo le sue due principali varianti, il funzionalismo del ruolo e quello dell’occupante. Il principale problema che sembra mostrare il funzionalismo è quello di non essere in grado di descrivere i qualia. Ho esposto a tal proposito l’esperimento mentale del cervello-cinese di Block, che facendo leva sull’improbabilità che una gigantesca mente-cinese dotata di coscienza possa svilupparsi da una certa organizzazione funzionale, ha cercato di mostrare che il funzionalismo è falso. Come Tye ha argomentato però, il problema potrebbe dipendere solo da una nostra carenza immaginativa. Ho allora esposto un esperimento mentale volto a mostrare che il linguaggio funzionalista non è in grado di rendere conto delle proprietà fenomeniche, ma che ciò non implica che quell’entità descritta non le possieda. Ecco dunque che seppur il cervello-cinese presentasse effettivamente proprietà fenomeniche, una descrizione funzionalista non sarebbe in grado di dircene nulla, ed è per lo stesso motivo che intuitivamente l’esperimento mentale di Block è tanto convincente. Ho poi presentato il funzionalismo dell’occupante come un’ottima posizione che permette sia di resistere all’argomento di Kim, riducendo tutte le proprietà mentali ai loro realizzatori, sia di riservare comunque efficacia al mentale. Come ho cercato di mostrare infatti, riportando anche le parole di Robb, è un errore linguistico distinguere il mentale dal fisico, è solo una questione di rivedere i nostri preconcetti. Comunemente infatti tendiamo ad utilizzare i termini “mentale” e “fisico” per riferirci a cose diverse, una teoria dell’identità come quella del funzionalismo dell’occupante pretende invece che i due termini si riferiscano in realtà alla stessa cosa, e così facendo non avremmo alcuna difficoltà nel riservare anche al mentale, la propria causalità. In quest’ultimo capitolo invece, ho cercato spiegare il problema del gap esplicativo che sembra esserci nel trattare gli aspetti qualitativi del mentale, i cosiddetti “qualia”. Ho esposto l’esperimento mentale degli zombie filosofici, che aveva come obiettivo quello di mostrare a priori che i qualia sono qualcosa di diverso dal fisico poiché non sarebbe metafisicamente necessario che essi sopravvengano al fisico. L’esperimento mentale si basa sul fatto che sia possibile concepire un mondo fisicamente uguale al mondo attuale, ma che non presenta qualia, e che gli abitanti-zombie di questo mondo possano replicare esattamente tutte le azioni che una persona dotata di qualia farebbe normalmente. Ho assunto, come Chalmers fa, che la concepibilità implichi anche la possibilità metafisica. Detto questo ho elaborato un altro esperimento mentale volto a mostrare che almeno intuitivamente, i qualia sembrano essere necessari per il darsi di alcuni comportamenti umani e che se l’esperimento mentale degli zombie prevede l’assenza di qualia allora la fisica di quel mondo dovrebbe essere in grado di predeterminare tutti i comportamenti degli zombie. Così facendo però si ammette che è concepibile un mondo zombi fisicamente predeterminato, e dunque si ammette il fatto che è anche metafisicamente possibile. Se è metafisicamente possibile significa che allora la fisica del nostro mondo dovrebbe essere in grado di spiegare tutti i nostri comportamenti anche senza la presenza di qualia, ma come mostrato dall’esperimento mentale proposto, i qualia sembrano essere necessari per alcuni comportamenti. Ho cercato allora di analizzare questa intuizione attraverso un’analisi controfattuale incentrata sulla somiglianza tra mondi possibili, mostrando che un controfattuale del tipo “se non ci fossero qualia, allora alcuni comportamenti non potrebbero darsi” risulta vero. Infine ho proposto un argomento che dovrebbe mostrare, che se effettivamente i qualia sono necessari per alcuni tipi di comportamenti, allora la fisica del mondo-zombie è diversa dalla fisica del mondo attuale. La fisica del mondo zombie infatti deve permettere il realizzarsi di azioni che nel mondo attuale avverrebbero solo per la presenza di qualia e se non accettiamo la sovradeterminazione causale, allora nel nostro mondo la nostra fisica non può permetterlo, tranne nel caso in cui i qualia non siano a tutti gli effetti considerati come qualcosa di fisico. La coscienza viene spesso considerata come l’ultima difesa, da parte dei non riduzionisti, per affermare un’indipendenza del mentale dal fisico, nonché l’ultimo scoglio con cui il fisicalismo deve confrontarsi. Questa indipendenza può essere vista o come un modo per far sopravvivere “a tutti i costi” forme di dualismo, o un modo per sottolineare la peculiarità della mente. Il problema rimane quello mostrato da Kim nel suo argomento, rendere irriducibile il mentale significa privarne l’efficacia causale. Alcuni filosofi, come Armstrong[[72]](#footnote-72), hanno specificato che “esistere” significa avere poteri causali e che dunque ammettere che la mente o anche solo i qualia non abbiano potere causale, significherebbe privarli di consistenza ontologica. Come infatti ho cercato di mostrare, il fatto che si concepisca la mente come un qualcosa di fisico non deve essere visto come un’eliminazione del mentale. Non vedo il motivo per cui non si possa concepire il mentale come facente parte a tutti gli effetti del mondo fisico. Quasi nessuno nega l’esistenza dei qualia[[73]](#footnote-73), e farli entrare dentro una metafisica fisicalista non può far altro che sottolineare la loro efficacia causale. Così facendo, con questo elaborato, ho cercato di mostrare che per difendere l’efficacia causale del mentale in generale[[74]](#footnote-74), si dimostra molto più utile far entrare a tutti gli effetti l’intero dominio del mentale in un’ontologia fisicalista che invece cercare di far sopravvivere a tutti i costi forme di dualismo.

# Bibliografia

## Primaria:

* Block N., *The China Brain*, da *Troubles With Functionalism*, “Minnesota Studies in the Philosophy of Science”, 9, University of Minnesota Press, Minneapolis 1978, p. 279.
* Chalmers D., *Can Consciousness Be Reductively Explained?*, cap. 3, in *The Conscious Mind. In Search of a Fundamental Therory*, Oxford University Press, New York 1996.

—, *The Two-Dimensional Argument against Materialism*, cap. 6, in *The Character of Consciousness*, Oxford University Press, New York 2010.

* Kim J., *Physicalism or Something Near Enough*, Princeton University Press, Princeton 2005.
* Levine J., *Materialism and Qualia: The Explanatory Gap,* “Pacific Philosophical Quarterly” 64, 1983.
* Putnam H., “*La natura degli stati mentali”*, De Palma A., Pareti G. (a cura di), in *Mente e corpo. Dai dilemmi della filosofia alle ipotesi della neuroscienza*., Bollati Boringhieri editore, Torino 2004, pp.63-78.

## Secondaria:

* Armstrong D. M., A Theory of Universals: Universals and Scientific Realism, Volume II , Cambridge University Press, Cambridge 1978, pp. 45-46.
* Block N., *Inverted Earth* “Philosophical Perspectives” Vol. 4, Action Theory and Philosophy of Mind 1990, pp. 53-79.
* Byrne A., "Inverted Qualia", The Stanford Encyclopedia of Philosophy (Fall 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2020/entries/qualia-inverted/>.
* Dennet D. “*Quainare i qualia*” in De Palma A., Pareti G. (a cura di), *Mente e Corpo. Dai dilemmi della filosofia alle ipotesi della neuroscienza*., Bollati Boringhieri editore, Torino 2004, pp. 189-233.
* Kment B., "Varieties of Modality", The Stanford Encyclopedia of Philosophy (Spring 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2021/entries/modality-varieties/>.
* Levin, Janet, "Functionalism", The Stanford Encyclopedia of Philosophy (Winter 2021 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/win2021/entries/functionalism/.
* Lewis D., *Counterfactuals*, Harvard University Press 1973.
* Mallozzi A., Anand V., Michael W., "The Epistemology of Modality", The Stanford Encyclopedia of Philosophy (Winter 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/win2022/entries/modality-epistemology/>.
* McLaughlin B.P. *Varieties of Sopravvenience*, cap. 8, in *Sopravvenience* di Kim J.,London, Routledge 2002.
* Menzies P., Beebee H., "Counterfactual Theories of Causation", The Stanford Encyclopedia of Philosophy (Winter 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2020/entries/causation-counterfactual/>.
* Paolini Paoletti M. *E.J. Lowe* “APhEx PORTALE ITALIANO DI FILOSOFIA ANALITICA” N° 7, 2013, Tripodi V. (ed).
* Putnam H., *Meaning and Reference*, “Journal of Philosophy” 70, 1973, pp. 669-771.
* Robb D. *The Identity Theory as a Solution to the Exclusion Problem*, in *Mental Causation and Ontology*, Gibb S., Lowe E., J., Ingthorsson R. (a cura di), Oxford University Press, Oxford 2013 Cap. 9 p. 222.
* Tye M,. "Qualia", The Stanford Encyclopedia of Philosophy (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2021/entries/qualia/>.

—, *Absent Qualia and the Mind-Body Problem*, “The Philosophical Review”, Vol. 115, 2006, p 142.

1. Intendo per essenza ciò che contraddistingue una sostanza da un’altra, ovvero le sue condizioni necessarie e sufficienti. [↑](#footnote-ref-1)
2. *Res cogitantes* è il plurale di *Res cogitans*. [↑](#footnote-ref-2)
3. Sebbene alcuni filosofi hanno messo in discussione la verità delle premesse dell’argomento di Kim, ad oggi non ci sono argomenti incontrovertibili che ne dimostrino la falsità. Assumerò comunque per valido l’argomento di Kim, poiché ritengo che sia in grado di darci un quadro chiaro del dibattito in filosofia della mente riguardo la causalità del mentale. [↑](#footnote-ref-3)
4. Con “effetto particolare” intendo la particolare esperienza qualitativa che si prova quando si osservano colori o quando si assaggia qualcosa. [↑](#footnote-ref-4)
5. Si pensi per esempio al Fedone di Platone, dove in alcuni passi sembra che il corpo venga visto come un ostacolo al raggiungimento della conoscenza e della verità. [↑](#footnote-ref-5)
6. Le chiamo “premesse generali” per distinguerle dalle premesse vere e proprie dell’argomento. Potrebbero essere definite anche “presupposti dell’argomento” in quanto chi le accetta deve accettare anche l’argomento. [↑](#footnote-ref-6)
7. McLaughlin B.P. *Varieties of Sopravvenience*, cap. 8, in *Sopravvenience* di Kim J.,London, Routledge 2002. [↑](#footnote-ref-7)
8. Dico “quasi nessuno” perché seppure ormai il dualismo delle sostanze sembra una posizione metafisica ormai obsoleta, alcuni filosofi della mente hanno proposto versioni di dualismo delle sostanze diverse dal classico dualismo cartesiano, come per esempio Jonathan Lowe. Paolini Paoletti M. *E.J. Lowe* “APhEx PORTALE ITALIANO DI FILOSOFIA ANALITICA” N° 7, 2013, Tripodi V. (ed). [↑](#footnote-ref-8)
9. Tutti questi problemi sono discussi in Kim J., *Physicalism or Something Near Enough*, Princeton University Press, Princeton 2005. [↑](#footnote-ref-9)
10. Kim J. *Physicalism or Something Near Enough*, op. cit. pp. 126-130. [↑](#footnote-ref-10)
11. Smart J. J. C., *Sensations and Brain Processes*, “Philosophical Review”, LXVIII, 1959. [↑](#footnote-ref-11)
12. Per proprietà fisiche intendo ovviamente quelle cerebrali. [↑](#footnote-ref-12)
13. Dico quasi perché come si vedrà poco dopo, l’epifenomenismo verrà escluso perché si ammette la causalità del mentale. [↑](#footnote-ref-13)
14. Kim J., *Physicalism or Something Near Enough,* op. cit. p. 50, (traduzione mia). [↑](#footnote-ref-14)
15. *Ibi*. pp. 50-52. [↑](#footnote-ref-15)
16. Queste versioni dell’argomento si trovano *Ibi.* pp 39-45. [↑](#footnote-ref-16)
17. *Ibi.* p. 17. (Traduzione mia.) [↑](#footnote-ref-17)
18. *Ibi.* p. 15. (Traduzione mia) [↑](#footnote-ref-18)
19. Primo tra tutti e il problema della realizzabilità multipla, che verrà discusso più avanti. [↑](#footnote-ref-19)
20. *Ibi.* p. 159, (traduzione mia). [↑](#footnote-ref-20)
21. Le informazioni che riporto sul saggio sono della traduzione in De Palma A., Pareti G. (a cura di), in *Mente e corpo. Dai dilemmi della filosofia alle ipotesi della neuroscienza*., Bollati Boringhieri editore, Torino 2004. [↑](#footnote-ref-21)
22. Uso il singolare “fibra-C” perché rende meglio il senso delle teorie dell’identità davanti all’ipotesi della realizzabilità multipla, ma ovviamente potrebbero essere perfettamente anche un intero gruppo di fibre-C responsabile di una determinata attività mentale. [↑](#footnote-ref-22)
23. A scopo dell’esempio utilizzerò “C” come variabile, dunque quando mi riferirò a fibre-D, fibre-E e fibre-F mi riferirò a loro in quanto esemplificazioni di fibre-C. [↑](#footnote-ref-23)
24. Per “dolore” qui intendo dolore fisico. [↑](#footnote-ref-24)
25. Ho dedicato tempo a questi dettagli perché se il lettore non comprende esattamente come funzionano le teorie dell’identità di tipo farà fatica a comprendere l’argomento della realizzabilità multipla. [↑](#footnote-ref-25)
26. Mi riferisco ovviamente ad umani senza particolari problemi di salute, dato che per determinati motivi medici e di malfunzionamento del sistema nervoso non è escluso che alcune persone possano sperimentare assenze di dolore anche se esposte a traumi fisici. [↑](#footnote-ref-26)
27. Questa formulazione è una formulazione leggermente modificata trovata in Levin, Janet, "Functionalism", The Stanford Encyclopedia of Philosophy (Winter 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2021/entries/functionalism/>, consultato nella versione del 2018. [↑](#footnote-ref-27)
28. Block N., *The China Brain*, da *Troubles With Functionalism*, “Minnesota Studies in the Philosophy of Science”, 9, University of Minnesota Press, Minneapolis, 1978, p. 279. [↑](#footnote-ref-28)
29. Block utilizza la popolazione cinese per l’esperimento mentale dato che il totale della popolazione cinese attuale si avvicina al numero di neuroni medio in un cervello. [↑](#footnote-ref-29)
30. L’input è data dalla visione di una pietra, ciò provoca un cambiamento di stato (l’uomo si ferma a pensare ad un percorso alternativo) e un output (l’inizio di un altro percorso); forse esistono esempi migliori, ma in ogni caso la forza di questo esperimento mentale (come di ogni altro esperimento mentale) è dato dalla sua forza intuitiva e spero che la sua sia tale da mostrare almeno intuitivamente la sua plausibilità. Sono inoltre convinto che l’esperimento mentale possa funzionare con altri tipi di descrizione funzionale, e che dunque attraverso una qualsiasi descrizione funzionale non riusciremmo a distinguere un essere umano ed un robot anche in altri contesti. [↑](#footnote-ref-30)
31. Attualmente esistono robot in grado di fare ciò, alcuni vengono utilizzati in Giappone all’interno di ristoranti per portare le pietanze. [↑](#footnote-ref-31)
32. Anche un qualsiasi altro animale sarebbe rientrato nella descrizione, ma senza sapere se l’animale ha proprietà fenomeniche, poteva benissimo anche essere un robot a forma di animale. [↑](#footnote-ref-32)
33. Tye M., *Absent Qualia and the Mind-Body Problem*, “The Philosophical Review”, Vol. 115, 2006, p 142, (traduzione mia). [↑](#footnote-ref-33)
34. La differenza principale sta nel fatto che l’esperimento mentale di Block risulta efficace se e solo se è vero che una mente-cinese di quel tipo non sviluppa una coscienza, dunque la buona riuscita dell’esperimento dipende da un fatto del mondo. La buona riuscita dell’esperimento mentale proposto invece non dipende in alcun modo dall’esistenza di un fatto del mondo, ma dal semplice funzionamento di un computer in grado di dare descrizioni funzionali. Sicuramente è meno impegnativo pensare ad un computer capace di fare ciò, invece che stabilire se una mente-cinese di enormi dimensioni possa sviluppare una coscienza. [↑](#footnote-ref-34)
35. Kim J.,*Physicalism or Something Near Enough,* op. cit p. 170. [↑](#footnote-ref-35)
36. *Ibi.* p. 131. [↑](#footnote-ref-36)
37. Qui la lettera “è” viene utilizzata nel senso di identità stretta. [↑](#footnote-ref-37)
38. Robb D. *The Identity Theory as a Solution to the Exclusion Problem*, in *Mental Causation and Ontology*, Gibb S., Lowe E., J., Ingthorsson R. (a cura di), Oxford University Press, Oxford 2013 Cap. 9 p. 222. (traduzione mia). [↑](#footnote-ref-38)
39. Kim J. *Physicalism or Something Near Enough*, op. cit., p. 160. [↑](#footnote-ref-39)
40. Detto in altri termini, è come quando due persone parlanti lingue diverse si riferiscono allo stesso oggetto, indipendentemente dal fatto che un inglese ed un francese chiamano l’acqua *eau* o *water,* si riferiranno sempre alla stessa cosa; il problema tra mentale e fisico potrebbe essere inteso esattamente allo stesso modo, abbiamo infatti due termini, “mentale” e “fisico” che pur avendo intensioni diverse hanno la stessa estensione. [↑](#footnote-ref-40)
41. Basta pensare anche solo alle diversità strutturali che ci sono tra il cervello dei polpi e quello degli umani. [↑](#footnote-ref-41)
42. De Palma A., Pareti G. (a cura di), *Mente e Corpo. Dai dilemmi della filosofia alle ipotesi della neuroscienza*., op. cit. pag. 236. [↑](#footnote-ref-42)
43. Esempi di “problemi facili” sono il controllo deliberato del comportamento, la capacità di accedere ai propri stati mentali interni e la capacità di distinguere stimoli esterni. Per ulteriori informazioni, vedi *Ibidem.* [↑](#footnote-ref-43)
44. Levine J., *Materialism and Qualia: The Explanatory Gap,* “Pacific Philosophical Quarterly” 64, 1983. [↑](#footnote-ref-44)
45. Come ho cercato di dimostrare nell’esperimento mentale proposto nel primo capitolo. [↑](#footnote-ref-45)
46. Una descrizione simile si può trovare in Chalmers D., *Can Consciousness Be Reductively Explained?*, cap. 3, in *The Conscious Mind. In Search of a Fundamental Therory*, Oxford University Press, New York 1996. [↑](#footnote-ref-46)
47. Altri tipi di possibilità sono quella “deontica” ed “epistemica”, ma entrambe non verranno trattate siccome non rientrano negli scopi dell’elaborato. [↑](#footnote-ref-47)
48. Kment B., "Varieties of Modality", The Stanford Encyclopedia of Philosophy (Spring 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2021/entries/modality-varieties/> e Mallozzi A., Anand V., Michael W., "The Epistemology of Modality", The Stanford Encyclopedia of Philosophy (Winter 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/win2022/entries/modality-epistemology/>. [↑](#footnote-ref-48)
49. Alcuni filosofi tendono a far coincidere la possibilità logica e la possibilità metafisica, penso però sia utile tenerle distinte per comprendere al meglio l’argomento degli zombie. [↑](#footnote-ref-49)
50. La possibilità fisica viene spesso anche chiamata “nomologica”. [↑](#footnote-ref-50)
51. Chalmers D., *The Two-Dimensional Argument against Materialism*, cap. 6, in *The Character of Consciousness*, Oxford University Press, New York 2010. [↑](#footnote-ref-51)
52. In realtà prima di fare questa distinzione Chalmers distingue anche tra “concepibilità negativa” e “concepibilità positiva”, che a loro volta possono essere intese in due sensi diversi, ovvero nella “concepibilità prima facie” (che dipende dal limite cognitivo del soggetto) e dalla “concepibilità ideale” (che non dipende dai limiti cognitivi dell’individuo). Queste distinzioni, seppure importanti, non sono utili alla trattazione. Nel caso il lettore fosse interessato ved. *ibidem*. [↑](#footnote-ref-52)
53. Putnam H., *Meaning and Reference*, “Journal of Philosophy” 70, 1973, pp. 669-771. [↑](#footnote-ref-53)
54. L’argomento della concepibilità di Chalmers viene ulteriormente sviluppato arrivando ad una riformulazione dell’argomento in cui come conclusione vi è una disgiunzione tra il fatto che il materialismo sia falso e che il monismo russelliano sia vero. Non saranno trattati gli esiti dell’argomento in questa sede siccome solo chi ne accetta le premesse ne può accettare la conclusione e compito di questa sezione sarà proprio tentare di mostrare la falsità di almeno una delle premesse. [↑](#footnote-ref-54)
55. La legge di Leibniz ammette che se due oggetti possiedono esattamente le stesse proprietà, inclusa la collocazione spazio-temporale allora sono lo stesso oggetto. In logica formale: . [↑](#footnote-ref-55)
56. Per “programmate” intendo che i movimenti sono predeterminati, nello stesso senso in cui si può dire predeterminato un robot, su cui viene installato un programma operativo, che prestabilisce già tutte le azioni che quel robot farà. Un tipo particolare di esempio nella storia della filosofia che può essere utile a comprendere in che senso intendo “programmare” è dato dall’organizzazione interna di una monade nell’apparato metafisico di Leibniz. Tutto ciò che ogni monade farà, è già determinato in partenza, in quanto è in un certo senso “iscritto” nella sua essenza. In questo caso però dobbiamo immaginare che il fisico sia in grado di predeterminare tutte le azioni degli zombie. [↑](#footnote-ref-56)
57. Ci sarebbe anche l’opzione della sovradeterminazione causale, ma se accettiamo l’argomento di Kim, non possiamo prenderla in considerazione. [↑](#footnote-ref-57)
58. Per una panoramica dei diversi modi in cui il termine “qualia” è stato utilizzato, si veda la prima parte di Tye M,. "Qualia", The Stanford Encyclopedia of Philosophy (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2021/entries/qualia/>. [↑](#footnote-ref-58)
59. Si veda Block N., *Inverted Earth* “Philosophical Perspectives” Vol. 4, Action Theory and Philosophy of Mind (1990), pp. 53-79. Per avere una panoramica completa sulla questione dei qualia invertiti si veda Byrne A., "Inverted Qualia", The Stanford Encyclopedia of Philosophy (Fall 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2020/entries/qualia-inverted/>. [↑](#footnote-ref-59)
60. Con “normale” intendo non inverito. [↑](#footnote-ref-60)
61. Cosa analoga può essere detta con i qualia del gusto, che saranno trattati successivamente in questo elaborato. [↑](#footnote-ref-61)
62. È utile ricordare, come già è stato fatto, che se l’argomento funziona e dunque i qualia sono qualcosa di fisico, questo non implica che si possano effettivamente descrivere all’interno un linguaggio fisicalista. Come il problema del gap esplicativo ha mostrato, sembra che qualsiasi linguaggio fisicalista non sia in grado di rendere conto dell’esperienza qualitativa, ma, questo non implica che quest’ultima non possa essere qualcosa di fisico. [↑](#footnote-ref-62)
63. Questo esempio non è altro che una spiegazione più accurata dell’esperimento mentale proposto, utilizzando una diversa esperienza qualitativa, al posto del colore, il gusto. [↑](#footnote-ref-63)
64. Ovviamente non serve che io debba aver avuto in passato una certa esperienza qualitativa, potrebbe essere benissimo la prima volta che vedo il colore blu e preferirlo rispetto ad altri colori perché mi ha colpito di più. La cosa importante però è che ci sia l’esperienza qualitativa. [↑](#footnote-ref-64)
65. I “tipi di scelta” sono quelli che hanno a che fare con qualia. [↑](#footnote-ref-65)
66. In particolare, tutto ciò a cui mi riferirò, è basato su Lewis D., *Counterfactuals*, Harvard University Press 1973. [↑](#footnote-ref-66)
67. Le controversie a cui mi riferisco sono l’asimmetria temporale, la transitività e la prelazione. Per una visione di questi problemi si veda: Menzies P., Beebee H., "Counterfactual Theories of Causation", The Stanford Encyclopedia of Philosophy (Winter 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2020/entries/causation-counterfactual/>. [↑](#footnote-ref-67)
68. Lascio libero il lettore di considerare i mondi possibili, sia come semplici strutture teoriche, sia come entità a sé spaziotemporalmente isolate. L’accettazione o meno del realismo modale è indipendente dalla funzionalità della teoria. [↑](#footnote-ref-68)
69. Per esempio, un mondo in cui la Germania ha vinto la seconda guerra mondiale è un mondo più vicino rispetto ad un mondo in cui nel 1945 degli extraterrestri hanno colonizzato il pianeta Terra. In un certo senso si può dire che la vittoria della Germania nazista sarebbe stata più possibile di una colonizzazione da parte di extraterrestri, siccome il primo mondo (quello in cui la Germania vince) è più vicino al mondo attuale che il secondo (quello della colonizzazione extraterrestre). [↑](#footnote-ref-69)
70. È da ricordare che fino ad ora abbiamo assunto che i qualia fossero effettivamente un qualcosa di diverso dal fisico. Dunque se effettivamente i qualia sono responsabili di almeno alcune azioni umane e non ci sono casi di sovradeterminazione causale, la fisica del mondo attuale non è in grado di rendere conto di tutte le azioni umane. Quindi per concepire un mondo-zombie, dobbiamo concepire una fisica diversa da quella del mondo attuale, ovvero un mondo fisico che determini causalmente tutte le azioni proprio come fa il programma di un robot. Ma questo significa che allora non si può concepire un mondo-zombie con le stesse verità fisiche del mondo attuale. L’unico modo di concepire un mondo con la stessa fisica del mondo attuale che permette di spiegare tutte le azioni umane è un mondo in cui tra le entità fisiche ci siano anche i qualia. [↑](#footnote-ref-70)
71. In questo caso nel mondo attuale almeno alcuni comportamenti deriverebbero da due cause distinte e sufficienti ciascuna, la presenza di qualia e la predeterminazione da parte del fisico. [↑](#footnote-ref-71)
72. Armstrong D. M., A Theory of Universals: Universals and Scientific Realism, Volume II , Cambridge University Press, Cambridge 1978, pp. 45-46. [↑](#footnote-ref-72)
73. Famoso è il caso di Dennet D. “*Quainare i qualia*” in De Palma A., Pareti G. (a cura di), *Mente e Corpo. Dai dilemmi della filosofia alle ipotesi della neuroscienza*., op. cit., pp. 189-233. [↑](#footnote-ref-73)
74. Inclusi dunque anche gli aspetti fenomenici. [↑](#footnote-ref-74)