

UFRRJ

INSTITUTO DE CIÊNCIA HUMANAS E SOCIAIS

CURSO DE PÓS-GRADUAÇÃO EM FILOSOFIA

Dissertação

**O experimento do zumbi e a história e-qualia de
Robert Kirk**

Matheus Pereira da Fonseca

2025



UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS HUMANAS E SOCIAIS
CURSO DE PÓS-GRADUAÇÃO EM FILOSOFIA

**O experimento do zumbi e a história e-qualia de
Robert Kirk**

Matheus Pereira da Fonseca

Sob a Orientação do(a) Professor(a)

Alessandro Bandeira Duarte

Texto de qualificação submetido como requisito parcial para obtenção do grau de Mestre(a), no Curso de Pós-Graduação em Filosofia, Área de Concentração em Filosofia.

Seropédica, RJ

Abril de 2025

Universidade Federal Rural do Rio de Janeiro
Biblioteca Central / Seção de Processamento Técnico

Ficha catalográfica elaborada
com os dados fornecidos pelo(a) autor(a)

FF673e Fonseca, Matheus Pereira da , 1998-
e O experimento do zumbi e a história e-qualia de
Robert Kirk / Matheus Pereira da Fonseca. - Rio de
Janeiro, 2025.
102 f.

Orientador: ALESSANDRO BANDEIRA DUARTE.
Dissertação(Mestrado). -- Universidade Federal Rural
do Rio de Janeiro, Programa de Pós-graduação em
Filosofia - Curso de Mestrado em Filosofia, 2025.

1. Experimento do zumbi. 2. Qualia. 3.
Fisicalismo. 4. Robert Kirk. I. DUARTE, ALESSANDRO
BANDEIRA, 1970-, orient. II Universidade Federal
Rural do Rio de Janeiro. Programa de Pós-graduação em
Filosofia - Curso de Mestrado em Filosofia III. Título.

UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS HUMANAS E SOCIAIS
CURSO DE PÓS-GRADUAÇÃO EM FILOSOFIA

MATHEUS PEREIRA DA FONSECA

Dissertação submetida como requisito parcial para obtenção do grau de Mestre em Filosofia, no Curso de Pós-Graduação em Filosofia, área de Concentração em Filosofia.

DISSERTAÇÃO APROVADA EM 28/04/2025

Alessandro Bandeira Duarte. Dr. UFRRJ
(Orientador)

Robinson Guitarrari. Dr. UFRRJ

Julia Cavalcanti Telles de Menezes. Dra. UFF

Marco Aurélio Sousa Alves. Dr. UFSJ



ATA Nº 1545/2025 - DeptFILO (12.28.01.00.00.00.85)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 01/05/2025 14:54)

ALESSANDRO BANDEIRA DUARTE

PROFESSOR DO MAGISTERIO SUPERIOR

DeptFILO (12.28.01.00.00.00.85)

Matrícula: ###645#3

(Assinado digitalmente em 01/05/2025 15:16)

ROBINSON GUITARRARI

PROFESSOR DO MAGISTERIO SUPERIOR

DeptFILO (12.28.01.00.00.00.85)

Matrícula: ###801#4

(Assinado digitalmente em 01/05/2025 16:35)

MARCO AURÉLIO SOUSA ALVES

ASSINANTE EXTERNO

CPF: ###.###.476-##

(Assinado digitalmente em 01/05/2025 15:58)

JULIA CAVALCANTI TELLES DE MENEZES

ASSINANTE EXTERNO

CPF: ###.###.017-##

Visualize o documento original em <https://sipac.ufrrj.br/documentos/> informando seu número: **1545**, ano: **2025**,
tipo: **ATA**, data de emissão: **01/05/2025** e o código de verificação: **12b0db3b34**

Dedicatória

Dedico este trabalho a Néia Mendes da Fonseca.

Agradecimentos

Gostaria de agradecer à minha família por todo o apoio incondicional. Sem vocês, nada disso seria possível.

Agradeço ao meu orientador, Alessandro Bandeira Duarte, por todo esforço e paciência para conduzir a orientação.

Agradeço à minha companheira, Juliana, por me apoiar nos momentos conturbados durante esse processo.

Também agradeço a todos os envolvidos no Programa de Pós-graduação em Filosofia da UFRRJ. É um orgulho ter participado desse Programa.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior -Brasil (CAPES) - Código de Financiamento 001

Resumo

DA FONSECA, Matheus Pereira. **O experimento do zumbi e a história e-qualia de Robert Kirk**. 2025. 104. Dissertação (Mestrado em Filosofia). Instituto de Ciências Humanas e Sociais, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2025.

Se zumbis são concebíveis, também são possíveis? A ideia de zumbi implica o epifenomenalismo? O epifenomenalismo implica a possibilidade de zumbis? O objetivo da presente pesquisa é analisar argumentos pró-zumbis e anti-zumbis que respondem essas perguntas. Basicamente os zumbis são criaturas hipotéticas idênticas a nós em todos os aspectos, mas sem experiência qualitativa, as chamadas qualia. Se a concebilidade dos zumbis implicar a possibilidade de que esses seres são metafisicamente possíveis, então a tese fisicalista que afirma que propriedades mentais se seguem de propriedades físicas é falsa. Em meados da década de 70, Robert Kirk publica dois artigos defendendo a possibilidade de zumbis contra o fisicalismo no geral – mas especialmente o funcionalismo e o behaviorismo - pois, segundo o autor, há uma tese que é compartilhada por qualquer fisicalista que se preze. Anos mais tarde, Kirk volta atrás em sua posição e passa a atacar o argumento do zumbi. Segundo ele, existe uma tese epifenomenalista (a história e-qualia) que implica a possibilidade de zumbis. Para o epifenomenalismo, as qualia, assim como outras propriedades e estados mentais, são propriedades que não têm nenhum efeito causal sobre os nossos comportamentos. Se for comprovado que os zumbis filosóficos são possíveis, então as propriedades qualitativas podem ser retiradas sem que isso afete o comportamento humano, o que parece estar de acordo com o epifenomenalismo. A estratégia de Kirk para atacar os zumbis é mostrar que a história e-qualia é inconcebível e, portanto, os zumbis também seriam inconcebíveis.

Palavras-chave: argumento dos zumbis; Robert Kirk; qualia; fisicalismo; problema mente-corpo.

Abstract

DA FONSECA, Matheus Pereira. **The Zombie Experiment and Robert Kirk's e-qualia story**. 2025. [104](#). Dissertation (Master in Philosophy). Instituto de Ciências Humanas e Sociais, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2025.

If zombies are conceivable, are they also possible? Does the idea of a zombie entail epiphenomenalism? Does epiphenomenalism entail the possibility of zombies? The aim of this research is to analyse pro-zombie and anti-zombie arguments that answer these questions. Basically zombies are hypothetical creatures identical to us in all aspects, but without qualitative experience, the so-called qualia. If the conceivability of zombies entails the possibility that these beings are metaphysically possible, then the physicalist thesis that mental properties follow from physical properties is false. In the mid-1970s, Robert Kirk published two articles defending the possibility of zombies against physicalism in general - but especially against functionalism and behaviourism - since, according to the author, there is a thesis that is shared by any serious physicalist. Years later, Kirk rejects his first position and starts attacking the zombie argument. According to him, there is an epiphenomenalist thesis (the e-qualia story) that entails the possibility of zombies. For epiphenomenalism, qualia, as other mental properties and states, are properties that have no causal effect on our behaviour. If it is proven that philosophical zombies are possible, then qualitative properties can be removed without affecting human behaviour, which seems to be in line with epiphenomenalism. Kirk's strategy for attacking the zombies is to show that the e-qualia story is inconceivable and therefore the zombies would also be inconceivable.

Keywords: zombie argument; Robert Kirk; qualia; physicalism; mind-body problem.

Sumário

1	Introdução	10
2	O experimento do zumbi: um panorama geral	12
2.1	Qualia	12
2.2	O experimento do zumbi	16
2.2.1	Fisicalismo	18
2.2.2	Concebibilidade	20
2.2.3	Mundos zumbis são metafisicamente possíveis?	27
3	Argumentos pró-zumbis e anti-zumbis de Robert Kirk	32
3.1	Argumentos pró-zumbi	32
3.1.1	O caso de Zulliver	32
3.1.2	O caso de Dan	38
3.1.3	Críticas a fase pró-zumbi de Kirk	42
3.2	Argumento anti-zumbi: a história e-qualia	46
3.3	Críticas à história e-qualia	58
3.3.1	Argumento <i>self-stultification</i>	59
3.3.2	Resposta a <i>self-stultification</i>	61
3.3.3	Intimidade epistêmica	65
3.3.4	Concebibilidade no argumento de Kirk	70
3.3.5	Premissa (A2)	71
4	Argumentos pró-zumbis	74
4.1	A concebibilidade de zumbis implica a história e-qualia?	74
4.1.1	Crítica à premissa (E3) da história e-qualia	74
4.1.2	Tipos de defensores de zumbis <i>versus</i> história e-qualia	76
4.1.3	O erro crucial da história e-qualia	79
4.1.3.1	Sobredeterminação causal em (Z3)	83
4.2	Defesa interacionista de mundos zumbis	85
5	Conclusão	99

Capítulo 1

Introdução

A discussão sobre o argumento dos zumbis é uma das questões centrais no campo da filosofia da mente, particularmente no debate sobre consciência. O argumento dos zumbis foi inicialmente formulado por Robert Kirk e ganhou notoriedade com *The Conscious Mind*, de David Chalmers. O objetivo central deste argumento é explorar a relação entre consciência e o mundo físico. A premissa principal deste argumento é a possibilidade de seres fisicamente idênticos aos humanos que, apesar de exibirem os mesmos comportamentos e processos físicos, não possuem consciência fenomenal. Ou seja, esses zumbis seriam fisicamente indistinguíveis de seres humanos, mas não teriam experiências subjetivas ou qualia (a sensação de "como é" ter uma experiência).

Esse cenário levanta questões cruciais sobre a natureza da consciência: se algo fisicamente idêntico a um ser humano pode ser concebido sem consciência, isso implicaria que a consciência não emerge de processos físicos do cérebro. Consequentemente, o argumento dos zumbis desafia o fisicalismo, que sustenta que tudo, incluindo a consciência, é necessitado pela estrutura física do mundo. Por outro lado, o argumento fortalece as posições dualistas, como o interacionismo, que sugerem que a natureza física do cérebro não se aplica a propriedades mentais, como sensações, pensamentos etc. No capítulo 2, apresentaremos a estrutura mais aceita do argumento do zumbi, apresentada por David Chalmers. Além disso, o capítulo destina-se a apresentar conceitos e discussões que estão intimamente relacionadas a noção de zumbi filosófico, caracterizando melhor a discussão.

A primeira versão do argumento do zumbi apresentada por Robert Kirk, em 1974, tem como objetivo central atacar principalmente a tese behaviorista. Ainda no mesmo ano, Kirk publica outra versão de seu argumento pró-zumbi. Ambos os artigos dedicam-se a mostrar que a ideia de zumbi não apresenta contradição aparente e, por isso, deve ser considerada como argumento contra teses fisicalistas. Tempos depois, Kirk volta atrás em sua posição e defende

que a concebibilidade de zumbis supostamente implica que a consciência é epifenomenal. O objetivo do capítulo 3 é discutir ambas as fases de Kirk (pró e anti-zumbi), analisando os seus argumentos para defender sua posição em cada fase. Comentarei os pontos que Kirk deixa a desejar em sua argumentação.

No capítulo 4, apresento críticas de Dmytro Sepetyi direcionadas à Kirk em sua fase anti-zumbi. Sepetyi explora as deficiências argumentativas de Kirk e mostra que é possível adotar a concebibilidade de zumbis sem defender o epifenomenalismo. Além disso, ainda no capítulo 4, apresento um argumento interacionista desenvolvido por Jake Khawaja a favor da concebibilidade de zumbis, apontando a concebibilidade de zumbis não implica que a consciência é epifenomenal.

A discussão sobre os zumbis é importante não apenas para a metafísica da mente, mas também para a compreensão da possibilidade conceitual e das limitações epistemológicas das teorias que buscam explicar a consciência. Filósofos que defendem a viabilidade do argumento dos zumbis argumentam que, mesmo que o comportamento físico e neural de uma pessoa possa ser descrito em termos científicos, isso não garante que tenhamos explicado ou compreendido a experiência subjetiva associada a esses processos. Portanto, o debate sobre os zumbis continua a ser um terreno fértil para investigar questões sobre a natureza da experiência consciente e seus limites no contexto das explicações físicas.

Capítulo 2

O experimento do zumbi: um panorama geral

Neste capítulo, como o nome sugere, apresentaremos um panorama geral da discussão sobre zumbis filosóficos, desde conceitos que serão importantes para o restante da discussão quanto propriamente a ideia de zumbi.

2.1 Qualia

Um conceito central que permeia toda a pesquisa é o de qualia (do singular 'quale'), uma vez que este conceito está diretamente relacionado ao experimento do zumbi, pois ele propõe que as propriedades mentais, assim como as qualia, não decorrem dos fatos físicos. Em virtude disso, consideramos prudente apresentar uma breve explicação sobre o que são as qualia (ou, ao menos, tentar). Embora seja um desafio definir o que são as qualia – não pretendemos assumir tal responsabilidade –, nosso objetivo é, apenas, oferecer um panorama geral da discussão, ainda que de forma sucinta. O desafio na definição das qualia reside no fato de que, em geral, há um consenso na literatura de que as qualia não podem ser definidas de maneira não-circular; por esse motivo, os autores frequentemente recorrem a exemplos para ilustrá-las.

Na literatura, o termo 'qualia' é utilizado para se referir às qualidades fenomenais subjetivas das experiências mentais conscientes. A qualidade fenomenal que um sujeito vivencia em estados e experiências conscientes corresponde àquilo que caracteriza ou é característico para tal sujeito ao ter aquela experiência. Por exemplo, há uma característica de como é experienciar a dor para um sujeito, uma característica de como é perceber o vermelho, ou ainda, uma característica de como é experimentar o sorvete gelado. Cada uma dessas experiências é associada a uma quale específica. Dessa forma, a quale é, conforme os exemplos ilustrados, a qualidade sentida que caracteriza a vivência de

uma experiência específica.

Essa pode ser uma primeira aproximação do que significa o termo 'qualia'. Podemos destacar outro tipo de definição, dada por [Shoemaker \(1982\)](#), no qual qualia são “as características qualitativas ou características fenomenais das experiências sensoriais, em virtude das quais elas se assemelham e diferem umas das outras, qualitativamente, nas formas como o fazem” ([SHOEMAKER, 1982](#), p. 367)¹. Ou seja, nesse caso, as qualia são definidas em termos daquilo que difere uma quale de outra: a experiência de ver o vermelho produz uma quale diferente da quale produzida pela experiência de ver o verde. Essas experiências conscientes vão ser diferidas pelo sujeito experimentador a partir das qualia que são produzidas. Para ficar mais claro, sabemos que ver o verde e ver o vermelho não são experiências idênticas porque as qualidades produzidas em nós são distintas. E como sabemos disso? Em geral, na literatura diz-se que nós sabemos (ou conhecemos) a qualidade fenomenal de uma experiência por introspecção. Se as qualia são propriedades que se caracterizam por ser “como as coisas se parecem” para nós, logo essa caracterização só faz sentido se tivermos acesso introspectivo às qualia, e.g., ao comer uma maçã, sei como é o gosto doce da maçã porque tenho capacidade de acessar introspectivamente minhas qualia². Assim, uma definição aproximada e imprecisa do que em geral os autores defendem pode ser a do tipo:

(Q) qualia são as formas como as coisas nos parecem.

De acordo com [Tye \(2021\)](#) os estados mentais que podem ter as qualia são:

1. Experiências sensórias no geral, como ver o vermelho, ouvir uma música alta, provar o sorvete, sentir o cheiro de algo queimando etc.
2. Experiências corporais, por exemplo, ter dores, ter coceiras, náuseas, calor, frio, fome etc.
3. Reações e emoções sentidas, como o sentimento de medo, felicidade, ciúme, tristeza.
4. O humor sentido, por exemplo, se sentir eufórico, amedrontado, calmo, entediado. Cada tipo de estado mental possui uma quale distinta, e essa quale vai caracterizar como é ter esse estado mental, por exemplo, estar deprimido gera uma quale que caracteriza esse estado mental.

¹ “the qualitative or phenomenal features of sense experiences, in virtue of having which they resemble and differ from each other, qualitatively, in the ways they do”.

² Veja [Kind \(2001, seção II\)](#).

O termo 'quale' obtém uma maior notoriedade nas obras de C.I. Lewis (1929), ainda que Lewis tenha utilizado quale em sua teoria dos dados dos sentidos. Para Lewis, quale é uma propriedade dos dados dos sentidos, e estes últimos são obtidos de imediato da experiência. Lewis diferencia quale de propriedades objetivas de um objeto: uma cor azul de um objeto é uma propriedade objetiva. Essa propriedade objetiva pode gerar várias qualia visuais (de cores) em situações diferentes, nesse sentido, ver o azul de um carro e ver o azul do céu são qualia diferentes³. Atualmente, usa-se o termo qualia para designar as propriedades da experiência, e não dos dados dos sentidos, como em Lewis. Entretanto, algumas características de quale que aparecem na definição de Lewis são mantidas na concepção atual:

As qualia são subjetivas; elas não têm nomes no discurso comum, mas são indicadas por algum circunlóquio, como "parece"; elas são inefáveis, pois podem ser diferentes em duas mentes, sem possibilidade de descobrir esse fato e sem inconveniência necessária para nosso conhecimento dos objetos ou de suas propriedades. Tudo o que pode ser feito para designar uma quale é, por assim dizer, localizá-la na experiência, ou seja, designar as condições de sua recorrência ou outras relações com ela. (LEWIS, 1929, p. 124)⁴

A apreensão das qualia ocorre a partir da experiência, e, conforme foi anteriormente mencionado, defende-se que o acesso às qualia se dá por meio da introspecção. Outro ponto relevante é que não é possível descrever as qualia de forma objetiva no discurso comum. Assim, na maioria dos casos, recorre-se a descrições de experiências semelhantes, utilizando, por exemplo, a expressão 'se parece com'. David Chalmers situou a dificuldade de explicar as qualia ao dividir a questão 'como é possível explicar a consciência' em dois problemas: o problema fácil e o problema difícil. Os problemas fáceis da consciência estão relacionados à explicação da estrutura da consciência, bem como das suas funções e dinâmicas. Esses problemas estão diretamente ligados aos métodos de investigação da ciência cognitiva, de modo que é possível explicá-los em termos de mecanismos neurais. O problema fácil não é fácil no sentido de ser simples de explicar, mas no sentido de que há uma abordagem pela qual podemos investigar tal problema, o que não ocorre no problema difícil. Em síntese, o problema difícil refere-se à grande dificuldade de explicar a relação entre os fenômenos físicos,

³ Cf. LEWIS, 1929, p. 121-2.

⁴ "Qualia are subjective; they have no names in ordinary discourse but are indicated by some circumlocution such as "looks like"; they are ineffable, since they might be different in two minds with no possibility of discovering that fact and no necessary inconvenience to our knowledge of objects or their properties. All that can be done to designate a quale is, so to speak, to locate it in experience, that is, to designate the conditions of its recurrence or other relations of it."

como os processos cerebrais, e as experiências fenomenais — as qualia — associadas a esses mecanismos ou processos físicos. Ao contrário do problema fácil, no problema difícil há a ausência de um meio de investigação científica capaz de solucionar tal questão.

A definição de Chalmers para o termo 'qualia' é similar à definição inicialmente apresentada. Qualia seriam "as propriedades dos estados mentais que classificam esses estados de acordo com o que é tê-los" (CHALMERS, 1996, Nota 2 do capítulo 1, tradução nossa)⁵. Nesse sentido, ao experimentar o gosto do sorvete, podemos descrever que a experiência desgustativa é típico de quando se come o sorvete. Note que a definição de 'quale' está em uma relação direta com a própria experiência típica de "como é tê-la".

Quanto ao status ontológico das qualia, surge a questão: as qualia são propriedades físicas ou propriedades não-físicas irreduzíveis? Na literatura sobre qualia, diversos experimentos do pensamento tentam defender a posição de que as qualia são, de fato, propriedades não-físicas irreduzíveis. O mais célebre, talvez, seja o caso de Mary, desenvolvido por Frank Jackson (1982)⁶. Na história proposta por Jackson, Mary está confinada em um quarto preto e branco, e, ao longo dos anos, sem nunca poder sair do ambiente, ela adquire informações sobre o mundo por meio de livros e aparelhos de televisão, todos em preto e branco. Mary adquire cada vez mais informações, passando a conhecer todos os fatos físicos⁷ relacionados à percepção das cores e ao modo como elas se apresentam no mundo, embora nunca tenha visto nenhuma cor além do preto e branco e nunca tenha saído do quarto.

Apesar de conhecer todos os fatos físicos pertinentes à percepção das cores, Mary se questiona sobre como é a experiência de ver o verde ou o vermelho. Até que, um dia, ela é libertada para ver o que existe fora do quarto em que viveu até aquele momento, para vivenciar as cores do mundo. Ao sair do quarto, Mary se depara com um jardim repleto de flores. Ao ver uma rosa vermelha, ela exclama: 'Então, é assim que se experimenta o vermelho.' E, ao olhar para a grama do jardim, diz: 'E isso é como é experimentar o verde.' Ao proferir tais frases, Mary parece adquirir informações relevantes. A questão, então, é: as informações que Mary obteve são realmente novas? Em outras palavras, se Mary conhece todos os fatos físicos relacionados às cores e à percepção delas, seria possível que ela aprenda algo novo ao entrar em contato com as cores? Jack-

⁵ "to those properties of mental states that type those states by what it is like to have them".

⁶ Ou também conhecido como argumento do conhecimento.

⁷ Aqui, "fatos físicos" deve ser tomado em um sentido amplo, que inclui o conhecimento sobre o funcionamento dos receptores e neurônios envolvidos na visão de cores (conhecimento biológicos e fisiológicos), além do conhecimento sobre toda a rede de relações causais entre os processos relativos à visão de cores e os estímulos externos e comportamentos (que se caracteriza por um conhecimento funcional).

son argumenta que sim, Mary adquire novas informações – sobre a qualidade fenomenal da experiência das cores – ao ter a experiência visual do vermelho, informações que não poderiam ser extraídas dos fatos físicos relacionados à cor e à percepção das cores. Assim, a nova informação adquirida por Mary deve ser considerada como sendo de natureza não-física⁸.

Um fisicalista não poderia sustentar a mesma posição que Jackson defende em relação ao caso de Mary. Isso porque, ao afirmar que o conhecimento físico seria suficiente para compreender todos os aspectos da percepção das cores, o fisicalista argumentaria que Mary não pode adquirir informações novas ao vivenciar sua primeira experiência com o vermelho, visto que ela já conhece tais fatos enquanto está em seu quarto. De maneira análoga, o mesmo se aplicaria às experiências com outras cores, uma vez que Mary possui pleno conhecimento de todos os fatos físicos relevantes à percepção das cores. Nosso objetivo não é, especificamente, discutir se o exemplo do quarto de Mary é bem-sucedido ao argumentar a favor de que as qualia são propriedades não-físicas, mas sim apresentar uma posição antirreducionista sobre o status ontológico das qualia.

O experimento do pensamento antirreducionista sobre as qualia que nos interessa na presente pesquisa é a possibilidade dos zumbis filosóficos. Em síntese, um zumbi filosófico é uma duplicata, molécula por molécula, de um ser senciente, como um ser humano, por exemplo, mas que difere dessa criatura por não possuir nenhuma propriedade qualitativa, ou seja, nenhuma quale. Na próxima seção, abordaremos brevemente o experimento do zumbi.

2.2 O experimento do zumbi

O argumento do zumbi, tal como o conhecemos atualmente, foi inicialmente proposto por Robert Kirk no artigo “Zombies vs. Materialists”, publicado em 1974. Nesse trabalho, Kirk defende uma posição que, em princípio, é a mais simples: os zumbis são seres fisicamente idênticos aos seres humanos, mas desprovidos de qualia. O zumbi é uma duplicata física exata de um ser humano, molécula por molécula, porém não há experiências qualitativas ocorrendo nesse ser peculiar. Nesse contexto, um mundo zumbi segue a mesma lógica: o mundo zumbi seria uma réplica idêntica ao mundo atual, exceto pela ausência de experiências qualitativas, ou seja, não existiriam qualia nesse mundo. Podemos

⁸ Existem duas versões do argumento do conhecimento, com conclusão forte ou fraca. Outra questão é se essa informação nova que Mary obtém é uma propriedade ou não. Para mais, sugiro a leitura de Nida-Rümelin, Martine and Donnchadh O Conaill, “Qualia: The Knowledge Argument”, The Stanford Encyclopedia of Philosophy (Spring 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <<https://plato.stanford.edu/archives/spr2024/entries/qualia-knowledge/>>.

supor que, nesse mundo zumbi, existiria um gêmeo zumbi para cada pessoa do mundo atual. Cada gêmeo zumbi seria funcional, comportamental e psicologicamente idêntico ao seu correspondente humano. Isso implica que os estados cerebrais dos dois indivíduos (o gêmeo humano e o gêmeo zumbi) desempenham o mesmo papel funcional e geram os mesmos estados psicológicos⁹. Por exemplo, se ambos os indivíduos estão em um estado cerebral que tem a função de causar tremores de dor, ambos apresentariam o mesmo comportamento, mas o gêmeo zumbi não teria nenhuma experiência qualitativa da dor. A identidade funcional e psicológica do gêmeo zumbi em relação ao humano do mundo atual decorre da sua identidade física. No entanto, o fato de o gêmeo zumbi ser fisicamente idêntico ao humano não lhe confere a capacidade de experienciar qualidades qualitativas.

Em termos gerais, a hipótese da possibilidade dos zumbis não postula que tais zumbis sejam nomicamente possíveis, isto é, que sua existência seja compatível com as leis da natureza. Em vez disso, a hipótese sustenta que réplicas zumbis desse tipo são, pelo menos, concebíveis, e este é o primeiro passo do argumento. O experimento mental dos zumbis filosóficos é comumente utilizado por dualistas para defender a tese de que as propriedades mentais não se derivam necessariamente de propriedades físicas, e, a partir disso, conclui-se que as propriedades mentais pertencem a uma categoria distinta das propriedades físicas. Existem ramificações desse dualismo, como aponta [Gomes \(2020\)](#):

Existem diversas variedades de dualismo em filosofia da mente, pelo menos duas possuem notoriedade: o dualismo substancial e o dualismo de atributos. Sobre o dualismo substancial, a maioria dos filósofos concorda que o principal representante desta teoria dualista é Descartes. O dualista substancial concede a existência de uma substância mental, ou seja, tal substância é distinta e incompatível com o mundo material. Além de ter tido alguma relevância no debate filosófico sobre a mente, é principalmente no debate metafísico e teológico que o dualismo substancial ganhará espaço, pois a substância pode ser interpretada como a alma que transcende o corpo, em existência e também em dignidade. ([GOMES, 2020](#), p. 105).

De modo geral, a caracterização fundamental do dualismo consiste na defesa da distinção entre mente e corpo. Os zumbis filosóficos representam uma séria ameaça às perspectivas fisicalistas, uma vez que, se as réplicas zumbis são metafisicamente possíveis, então surge um argumento simples que parece demonstrar que as propriedades mentais não decorrem necessariamente das pro-

⁹ [Silby \(1998, p. 5\)](#) define os estados psicológicos pela função que desempenham no sistema cognitivo e pelo comportamento que geram.

priedades físicas – o que é justamente a posição defendida pelos fisicalistas. O fisicalismo, por sua vez, é a tese metafísica que se opõe ao dualismo, afirmando que todas as entidades existentes são, necessariamente, físicas, de modo que não há nada além do físico. Consequentemente, as propriedades, estados e eventos mentais seriam, necessariamente, considerados como entidades físicas, segundo essa corrente teórica.

Segundo Chalmers (2010, p. 142), suponha que P é o conjunto de verdades físicas e Q é uma verdade mental particular – e.g., que alguém é fenomenalmente consciente. Podemos estruturar o argumento do zumbi da seguinte forma:

1. $P \& \neg Q$ é concebível
2. Se $P \& \neg Q$ é concebível, $P \& \neg Q$ é metafisicamente possível
3. Se $P \& \neg Q$ é metafisicamente possível, então o fisicalismo é falso

4. O fisicalismo é falso.

Se é possível que as verdades físicas não implique uma verdade mental, então a tese fisicalista não é o caso.

2.2.1 Fisicalismo

A concepção predominante sobre o fisicalismo é que ele postula ser a tese segundo a qual todos os fatos, incluindo os fatos mentais, decorrem necessariamente dos fatos físicos, ou, alternativamente, que são necessitados pelos fatos físicos. Caso afirmemos que a proposição ‘todos os fatos decorrem necessariamente dos fatos físicos’ fornece a condição necessária e suficiente para o fisicalismo, e, adicionalmente, se isso também se aplica às propriedades – isto é, ‘todas as propriedades decorrem necessariamente das propriedades físicas’ –, então obtemos a formulação do fisicalismo. Jackson descreve a tese fisicalista por meio da seguinte definição:

Fisicalismo (F): Qualquer mundo que seja uma réplica física mínima do nosso mundo será uma réplica *simpliciter* do nosso mundo¹⁰.

Uma réplica física mínima é, conforme Jackson, aquela que captura todos os aspectos físicos do nosso mundo sem adicionar nenhum elemento. Quando esses aspectos físicos do nosso mundo, que constituem a réplica física mínima, são replicados para um mundo possível, obtemos uma réplica *simpliciter* do nosso mundo. De acordo com Jackson, a duplicata física mínima do nosso mundo deve atender a duas condições:

¹⁰ Cf. JACKSON, 1998, p. 12.

- (a) é exatamente como o nosso mundo em todos os aspectos físicos (tanto em propriedades, leis, relações) e
- (b) não contém mais nada do que é necessário para satisfazer (a)¹¹.

Basicamente, ao replicar a estrutura física do mundo atual para o mundo possível *w*, o fisicalista sustenta que todos os aspectos que o nosso mundo tem – incluindo os aspectos mentais –, o mundo possível *w* também terá, pois todos os aspectos (incluindo propriedades) decorrem necessariamente de aspectos físicos. Acerca do fisicalismo, Jackson diz:

O fisicalismo não é simplesmente a doutrina de que o mundo tem muita natureza física. Isso não é controverso: quase todos concordam, por exemplo, que os objetos têm massa, carga e densidade, e que existem campos de força gravitacionais e elétricos. A doutrina distintiva dos fisicalistas é, como eles dizem, que o mundo é inteiramente de natureza física, que ele não é nada além do mundo físico, ou nada além dele, e que um inventário completo das propriedades e relações físicas instanciadas seria um inventário completo *simpliciter*. (JACKSON, 1998, p. 9)¹²

Em outras palavras, a afirmação do fisicalismo (F) sustenta que a natureza física do nosso mundo determina de maneira exclusiva toda a sua natureza, sem omitir nenhum aspecto. Um argumento simples pode demonstrar como a tese metafísica (F), conforme descrita acima, é fundamental para o fisicalismo. Inicialmente, vamos supor que a tese (F) seja falsa. Caso (F) seja falsa, isso implica que o nosso mundo e alguma réplica física mínima dele são distintos; pelo menos um contém algo que o outro não possui. No entanto, por definição, uma réplica física mínima do nosso mundo não contém nenhum aspecto nem instancia nenhuma propriedade que não esteja presente em nosso mundo; ou seja, tudo o que consta em qualquer réplica física mínima do nosso mundo está também em nosso mundo. Assim, o nosso mundo conteria aspectos ou instanciaría propriedades que a duplicata física mínima não contém. Contudo, esse aspecto ou propriedade teria de ser não físico, pois o nosso mundo e a duplicata são fisicamente idênticos. Dessa forma, se o nosso mundo contém aspectos ou proprie-

¹¹ Cf. JACKSON, 1998, p. 13

¹² "Physicalism is not simply the doctrine that the world has lots of physical nature. That is not controversial: nearly everyone agrees, for instance, that objects have mass, charge, and density, and that there are gravitational and electrical force fields. The physicalists' distinctive doctrine is, as they variously say it, that the world is entirely physical in nature, that it is nothing but, or nothing over and above, the physical world, and that a full inventory of the instantiated physical properties and relations would be a full inventory *simpliciter*."

dades não físicas, o fisicalismo é falso. Portanto, se (F) for falsa, o fisicalismo é falso, ou seja, o fisicalismo está comprometido com (F)¹³.

Por outro lado, é necessário demonstrar que, se o fisicalismo for falso, então a tese (F) também é falsa. Caso o fisicalismo seja falso, isso implica que o nosso mundo contém algum aspecto de natureza não física. No entanto, essa natureza não pode estar presente em nenhuma réplica física mínima do nosso mundo, uma vez que tal natureza constitui um acréscimo não físico à natureza física do nosso mundo. Assim, qualquer réplica desse tipo não pode ser considerada uma réplica simpliciter do nosso mundo e, portanto, a tese (F) é falsa. Dessa forma, pode-se concluir que a tese metafísica (F) capta o que é essencial para a concepção fisicalista¹⁴.

Uma vez caracterizada a tese metafísica do fisicalismo, é possível contextualizar a discussão a partir do experimento mental do zumbi. O experimento filosófico do zumbi tem como objetivo desafiar a afirmação de que os fatos e propriedades físicas determinam todas as propriedades, fatos e aspectos do nosso mundo, conforme estipulado pela tese (F). No mundo zumbi, as propriedades, fatos e eventos mentais não são instanciados nos seus habitantes, embora o mundo seja uma réplica física mínima do nosso mundo. Se a premissa (2) – Se $P \& \neg Q$ é concebível, $P \& \neg Q$ é metafisicamente possível – for verdadeira e o mundo zumbi for concebível, isso implica que as propriedades físicas não determinam todas as propriedades do nosso mundo atual na réplica zumbi. Os habitantes do mundo zumbi teriam todos os elementos necessários para instanciar as propriedades mentais, dado que tal mundo é uma réplica física mínima do nosso mundo atual, mas essas propriedades mentais não seriam, de fato, instanciadas. Assim, como mencionado anteriormente, a conclusão central que os defensores da hipótese dos zumbis buscam sustentar é que as propriedades, fatos e aspectos mentais não decorrem necessariamente de propriedades, fatos e aspectos físicos.

No entanto, para se chegar a essa conclusão, o defensor da hipótese zumbi precisa argumentar que a concepção de concebilidade implica possibilidade.

2.2.2 Concebilidade

A premissa (2), que trata da concebilidade, desempenha um papel fundamental no argumento do zumbi. A noção de concebilidade constitui um conceito técnico na epistemologia da modalidade e, de maneira sucinta, sustenta que, se algo é concebível, então esse algo é possível. Conforme destacado por

¹³ Cf. JACKSON, 1998, p. 13-4.

¹⁴ Cf. JACKSON, 1998, p. 13-4.

Chalmers (2002), os argumentos que envolvem a noção de concebilidade geralmente se desdobram em três etapas: em primeiro lugar, afirma-se o que pode ser concebido (reivindicação epistêmica); em segundo lugar, sustenta-se que o que é concebido é possível ou necessário (reivindicação modal); por fim, afirma-se algo sobre a natureza do mundo (reivindicação metafísica). Nos argumentos que defendem que o mental não é redutível ao físico, como no caso do argumento dos zumbis, observa-se uma lacuna entre os dois domínios: primeiramente, afirma-se que podemos conceber o mental sem o físico, ou vice-versa; em seguida, sustenta-se a existência de uma lacuna modal, isto é, a possibilidade de o mental existir sem o físico, ou vice-versa; e, finalmente, conclui-se que o mental não é redutível ao físico, caracterizando a lacuna metafísica.

Esse é o primeiro problema para os defensores do argumento do zumbi. Em geral, os autores que defendem a concebilidade de mundos zumbis defendem que algo é concebível se não podemos descartar por raciocínio *a priori*. Chalmers 2002 e 2010 apresenta uma série de distinções que serão úteis para os nossos propósitos, tanto no presente capítulo quanto no capítulo 3, no qual apresentarei o que Robert Kirk entende como concebilidade ao apresentar o argumento da história e-qualia e farei um panorama com a semântica bi-dimensional de Chalmers¹⁵. É útil apresentar uma discussão mais refinada sobre concebilidade, pois a crítica de Kirk ao argumento do zumbi se situa especificamente na ideia de que mundos zumbis são inconcebíveis.

Para discutir a tese de concebilidade dos argumentos contra o materialismo¹⁶, Chalmers faz distinções importantes que refinam a definição daquilo que é concebível. Primeiro, é válido destacar que a noção de concebilidade é uma noção epistêmica, definida em termos epistemológicos e, em alguma medida, psicológicos¹⁷. Como foi dito, em geral, dizemos que uma sentença S é concebível quando S expressa uma hipótese que não pode ser descartada *a priori*.

Chalmers distingue essa definição em duas: concebilidade *prima facie* e concebilidade ideal. Uma sentença S é *prima facie* concebível quando um sujeito é incapaz de descartar S por raciocínio *a priori* em uma primeira consideração. Por outro lado, uma sentença S é idealmente concebível quando um sujeito é incapaz de descartar S após reflexões ideais. A diferença, sugere Chalmers, é que na concebilidade *prima facie* o sujeito pode ter limitações cognitivas contingentes

¹⁵ Além de Chalmers, a semântica bi-dimensional tem outros autores como proponentes. No entanto, no presente trabalho, apresentaremos apenas as ideias de Chalmers. Para uma visão mais ampla, ver Capítulo 7 em Ruffino (2022).

¹⁶ Seguirei os mesmos termos utilizados pelo autor, mas não faço nenhuma diferenciação entre materialismo e fisicalismo. Portanto, a mesma discussão sobre fisicalismo se segue para o materialismo.

¹⁷ Cf. CHALMERS, 2010, p. 143

e desse modo uma sentença S pode ser considerada erroneamente concebível. Na concebibilidade ideal, as limitações cognitivas contingentes de um sujeito são eliminadas, de maneira que a reflexão ideal elimina qualquer possibilidade de erro.

Alguns exemplos oferecidos por Chalmers podem clarear essa distinção:

1. A sentença “ $2+2=5$ ” não é *prima facie* concebível nem idealmente concebível, pois se rejeita sua concebibilidade antes mesmo de uma reflexão ideal;
2. Seja S uma sentença matemática complexa, mas demonstrável: $\neg S$ é *prima facie* concebível, mas não é idealmente concebível, ainda que seja demonstrável – só não temos acesso a essa demonstração.

As noções de concebibilidade acima são versões de concebibilidade negativa, que é definida em termos daquilo que o sujeito pode rejeitar como concebível através do raciocínio *a priori*. Dizemos que S é negativamente concebível “quando o sujeito não pode descartar S através do raciocínio *a priori*” (CHALMERS, 2010, p. 144)¹⁸. Desse modo, as noções de concebibilidade que definimos acima são as noções de concebibilidade *prima facie* negativa e concebibilidade ideal negativa.

A noção oposta da concebibilidade negativa é a concebibilidade positiva, que é definida em termos daquilo que o sujeito pode formar como concepção. Dizemos que S é positivamente concebível quando o sujeito pode imaginar coerentemente uma situação em que S é o caso. Assim, as duas noções de concebibilidade *prima facie* positiva e concebibilidade idealmente positiva podem ser definidas:

S é *prima facie* positivamente concebível quando um sujeito pode imaginar uma situação que a pessoa considera coerente e também considera ser aquela em que S é o caso. Além disso, pode-se dizer que S é idealmente concebível positivamente quando a sua concebibilidade positiva *prima facie* não pode ser derrotada na reflexão racional ideal. (CHALMERS, 2010, p. 144).¹⁹

Uma situação é coerentemente imaginada quando podemos preencher os detalhes da situação de modo que nenhuma contradição se revele. Desse modo, a concebibilidade idealmente positiva precisa mostrar que todos os detalhes podem ser preenchidos sem que a situação seja rejeitada.

Pelo que vimos até agora, a concebibilidade ideal surge como um guia melhor para a possibilidade em relação a concebibilidade *prima facie*. O caso da sentença

¹⁸ “when S cannot be ruled out through *a priori* reasoning”.

¹⁹ “S is *prima facie* positively conceivable when a subject can imagine a situation that the person takes to be coherent and also takes to be one in which S is the case. Furthermore, one can say that S is ideally positively conceivable when its *prima facie* positive conceivability cannot be defeated on ideal rational reflection”.

matemática altamente complexa nos mostra que tal sentença pode ser *prima facie* negativamente concebível, mas não idealmente concebível.

Chalmers destaca que, em geral, as alegações de que a concebilidade não implica possibilidade são alegações de que a concebilidade *prima facie* não implica possibilidade. Por exemplo, é comum que se defenda que tanto a Conjectura de Goldbach²⁰ quanto a sua negação são concebíveis, ainda que apenas um dos casos seja possível. A noção de concebilidade utilizada nesse exemplo é algo próximo da noção de concebilidade *prima facie* negativa, ou seja, seja *S* a sentença que descreve a conjectura de Goldbach, tanto *S* quanto $\neg S$ não podem ser rejeitadas através de raciocínio *a priori*. Por esse motivo, os contraexemplos a afirmação de que concebilidade não implica possibilidade em geral são contraexemplos contra a concebilidade *prima facie*.

No entanto, parece pouco provável que tanto a conjectura de Goldbach quanto a sua negação sejam idealmente concebíveis²¹. Na reflexão ideal, isto é, retirando todas as limitações cognitivas, um dos casos seria descartado como inconcebível. Desse modo, a concebilidade ideal ganha força como um guia mais seguro para a possibilidade, em relação à concebilidade *prima facie*.

A partir de discussões sobre sentenças necessárias *a posteriori* feitas por Kripke (1980), Chalmers faz outra distinção com o objetivo de mostrar que há um sentido de concebilidade na qual a sentença “água não é H_2O ” é corretamente concebível e há outro sentido na qual a mesma sentença não é concebível, a qual Chalmers denomina de concebilidade primária e concebilidade secundária, respectivamente. Na concebilidade secundária, uma sentença *S* é secundariamente concebível quando é concebível que *S* poderia ter sido o caso em mundos contrafactuais, a partir das verdades do mundo atual. No mundo atual, “água não é H_2O ” é uma sentença falsa, e esse valor de verdade se estende para os mundos contrafactuais nos quais existirem a água. Nesse sentido, a sentença “água não é H_2O ” não é concebível.

Por outro lado, a concebilidade primária afirma que uma sentença *S* é primariamente concebível quando é concebível que *S* realmente seja o caso. A concebilidade primária está no domínio *a priori*, de modo que todo o caráter empírico do mundo deve ser suspenso. Segundo Chalmers, a concebilidade primária tem como base a ideia de que “por tudo o que sabemos *a priori*, há muitas maneiras pelas quais o mundo pode ser” (CHALMERS, 2002, p. 157)²².

²⁰ A Conjectura de Goldbach afirma que todo número par maior que 2 pode ser expresso como a soma de dois números primos (iguais ou diferentes). Entretanto, nem a prova da conjectura e nem a da sua negação foram alcançadas.

²¹ Vale ressaltar que existem interpretações de que a Conjectura de Goldbach é um caso de incompletude. Se a Conjectura de Goldbach for interpretada como um caso de incompletude, nem a sua afirmação nem sua negação são deriváveis em um sistema matemático.

²² “for all we know *a priori*, there are many ways the world might be”.

As maneiras pelas quais o mundo poderia ser são pensadas como possibilidades epistêmicas, nesse sentido, “água não é H₂O” é epistemicamente possível quando uma hipótese que verifica a sentença “água não é H₂O” não pode ser descartada *a priori*. Dessa maneira, a concebilidade negativa definida anteriormente é similar a concebilidade primária, à medida que ambas são definidas em termos daquilo que não pode ser rejeitado *a priori*, e a sentença “água é H₂O” não pode ser estabelecida *a priori*.

Quando consideramos situações como reais, nós as consideramos e avaliamos da maneira que consideramos e avaliamos possibilidades epistêmicas. Ou seja, dizemos a nós mesmos: e se o mundo real for realmente assim? Supõe-se hipoteticamente que a situação em questão é real, e considera-se se, dessa suposição, segue-se que S é o caso. Se sim, então a situação verifica S, quando considerada como real. (CHALMERS, 2002, p. 157-8).²³

Ou seja, quando consideramos uma hipótese, avaliamos a suposição e verificamos se é o caso. Se for possível conceber a situação que verifica a hipótese, então podemos considerá-la como real. Por exemplo, no caso da água e H₂O, podemos conceber que os cientistas descubrem que estivemos enganados e a água é feita de moléculas de XYZ. Se alguém hipoteticamente aceita que tal situação ocorre, então há um sentido em que água não é H₂O.

Chalmers deixa claro que a concebilidade primária não implica em possibilidade metafísica, no entanto, é o guia mais confiável para a ligação de concebilidade e possibilidade. Isso porque ainda permanece uma conexão entre concebilidade primária e possibilidade metafísica em alguns casos, como o da água e H₂O:

Quando concebemos que a água não é H₂O, imaginamos (por exemplo) uma situação na Terra Gêmea em que o líquido aquoso nos oceanos e lagos é XYZ. Essa situação é metafisicamente possível, portanto, há um sentido em que nossa concepção envolve o acesso a um mundo possível. De acordo com a maneira usual de descrever mundos possíveis, esse mundo não é um mundo em que a água não é H₂O. Mas o mundo ainda está em uma forte relação com a sentença “água não é H₂O”. Em particular, se aceitarmos que nosso próprio mundo tivesse o caráter desse mundo (com XYZ nos oceanos e lagos),

²³ “When we consider situations as actual, we consider and evaluate them in the way that we consider and evaluate epistemic possibilities. That is, we say to ourselves: what if the actual world is really that way? One hypothetically assumes that the situation in question is actual, and considers whether, from that assumption, it follows that S is the case. If so, then the situation verifies S, when considered as actual.”.

deveríamos então endossar a afirmação “água não é H_2O ”. (CHALMERS, 2010, p. 147).²⁴

Em outras palavras, a situação de imaginar o mundo como a Terra Gêmea, no qual o líquido aquoso nos oceanos e lagos é XYZ, endossa a afirmação de que a água não é H_2O . Se tomarmos a Terra Gêmea como um mundo que poderia ser o nosso, estaremos aceitando a afirmação de que “a água não é H_2O ”. Dentro da semântica que Chalmers apresenta,

3. Podemos dizer que a Terra Gêmea não satisfaz a sentença “água não é H_2O ”, i.e., se considerarmos a Terra Gêmea como mundo contrafactual (ao mundo atual), a sentença “água não é H_2O ” é falsa.
4. Por outro lado, dizemos que a Terra Gêmea verifica a sentença “água não é H_2O ”, ou seja, se considerarmos a Terra Gêmea como mundo atual, então a sentença “água não é H_2O ” é verdadeira.

A distinção entre (3) e (4) é capturada pela distinção entre intensões primárias e secundárias da declaração “água não é H_2O ”. Na semântica modal de Kripke, cada expressão tem uma extensão e uma intensão. Por exemplo, um termo, como um nome próprio, tem uma extensão que será o indivíduo ao qual o termo se refere. Por outro lado, a intensão de uma expressão é uma função de mundos possíveis para extensões²⁵, de modo que a intensão é uma função que mapeia os mundos possíveis com o referente da expressão.

Kripke formula sua semântica modal baseada no mundo atual, de modo que a referência das expressões é fixada no mundo atual, e os mundos possíveis são mundos contrafactuais com base na referência do mundo atual. No caso da sentença necessária *a posteriori* “água é H_2O ”, se é verdadeira, qualquer mundo possível que tenha água, terá H_2O . Deste modo, na semântica modal de Kripke, não é possível que água seja associada a outra extensão que não seja H_2O . Entretanto, um falante poderia ter a intuição de que a água poderia não ser H_2O se o mundo fosse diferente. A distinção entre intensão primária e secundária da semântica bidimensional consegue capturar essa intuição.

²⁴ “When we conceive that water is not H_2O , we imagine (for example) a Twin Earth situation in which the watery liquid in the oceans and lakes is XYZ. This situation is metaphysically possible, so there is a sense in which our conceiving involves access to a possible world. Under the usual way of describing possible worlds, this world is not one in which water is not H_2O . But the world still stands in a strong relation to the sentence ‘water is not H_2O .’ In particular, if we came to accept that our own world had the character of this world (with XYZ in the oceans and lakes), we should then endorse the claim ‘water is not H_2O .’”

²⁵ Cf. CHALMERS, 1996, p. 51.

Na semântica bidimensional, a intensão secundária se mantém a mesma da semântica modal de Kripke, ou seja, é a função de mundos contrafactuais. Em (3), estamos expressando a intensão secundária da sentença “água não é H_2O ”, no qual a sentença é falsa para um mundo contrafactual w . Na intensão secundária, as verdades do mundo considerado contrafactual dependem das verdades do mundo atual.

Por outro lado, a intensão primária de um conceito (nesse caso, água) é uma função de mundos considerados como atuais para extensões. Em (4), expressamos a intensão primária de “água não é H_2O ”, que é verdadeira em um mundo w tomado como qualitativamente idêntico ao nosso mundo.

A rigor, os mundos que são relevantes para as intensões primárias são mundos centrados: mundos que vêm com um “centro” marcado que consiste em um indivíduo e um tempo. Quando consideramos um mundo centrado w como real, consideramos a hipótese de que estamos atualmente na situação do indivíduo no centro. (CHALMERS, 2010, p. 146-7).²⁶

Desse modo, a principal diferença da intensão primária para a intensão secundária é tomar o mundo como “centrado”, isto é, como o mundo atual. Ao fazer isso, não estaremos considerando o mundo hipotético como dependente das verdades do nosso mundo, na verdade, estaremos considerando como seria o estado de coisas sob as verdades desse mundo hipotético. Desse modo, podemos afirmar que

quando a intensão primária de S é verdadeira em algum mundo centralizado (ou seja, quando algum mundo centralizado verifica S), S é primariamente possível, ou 1-possível. Quando a intensão secundária de S é verdadeira em algum mundo (ou seja, quando algum mundo satisfaz S), S é secundariamente possível, ou 2-possível. Então, “água não é H_2O ” não é 2-possível, mas é 1-possível. (CHALMERS, 2010, p. 147).²⁷

Ainda que a sentença “água não é H_2O ” seja primariamente concebível, mas não secundariamente possível, ela é inteiramente compatível com os casos em que há ligação entre concebibilidade primária e possibilidade primária.

²⁶ “Strictly speaking, the worlds that are relevant to primary intensions are centered worlds: worlds that come with a marked “center” consisting of an individual and time. When we consider a centered world w as actual, we consider the hypothesis that we are currently in the situation of the individual at the center.”.

²⁷ “when the primary intension of S is true at some centered world (i.e., when some centered world verifies S), S is primarily possible, or 1-possible. When the secondary intension of S is true at some world (i.e., when some world satisfies S), S is secondarily possible, or 2-possible. Then ‘water is not H_2O ’ is not 2-possible, but it is 1-possible.”.

Outras necessidades *a posteriori* apresentadas por Kripke, como “calor é o movimento cinético das moléculas” ou “Héspero é Fósforo” seguem o mesmo padrão. Para cada um desses casos, em termos bi-dimensionais, dizemos que são primariamente concebíveis (1-concebível) mas não secundariamente possíveis (2-possível). No entanto, em cada um desses casos, a negação em questão é primariamente possível (1-possível). Por exemplo, a sentença “o calor não é o movimento das moléculas” é verificada por um mundo hipotético no qual algo diferente das moléculas causa sensações de calor. Do mesmo modo, a sentença “Héspero não é Fósforo” é verificada por um mundo hipotético no qual os objetos visíveis nos céus da manhã e da noite são totalmente diferentes. Portanto, ainda há um elo forte que liga concebilidade e possibilidade metafísica, segundo Chalmers.

Podemos retirar algumas conclusões a partir da discussão:

- a. A concebilidade *prima facie* é um guia imperfeito para possibilidade;
- b. A concebilidade primária é um guia imperfeito para a possibilidade secundária
- c. A concebilidade ideal é o caminho mais seguro para possibilidade
- d. A concebilidade primária, ainda que seja imperfeita para a possibilidade secundária, é um bom guia para possibilidade primária, i.e., a possibilidade de um mundo que satisfaça a intenção primária de S.

Juntando as conclusões (c) e (d), Chalmers conclui que há uma ligação consistente entre a concebilidade primária ideal e a possibilidade primária, portanto, é a esse tipo de concebilidade que devemos nos deter.

2.2.3 Mundos zumbis são metafisicamente possíveis?

Agora, o passo essencial para o argumento do zumbi funcionar é mostrar que há uma ligação entre a possibilidade primária (possibilidade epistêmica) e a possibilidade metafísica. Se a tese de Chalmers estiver correta, o argumento do zumbi deve ser reestruturado da seguinte forma:

1. $P \& \neg Q$ é concebível
2. Se $P \& \neg Q$ é concebível, $P \& \neg Q$ é epistemicamente possível (possibilidade primária)
3. Se $P \& \neg Q$ é epistemicamente possível, então $P \& \neg Q$ é metafisicamente possível.
4. Se $P \& \neg Q$ é metafisicamente possível, então o fisicalismo é falso

4. O fisicalismo é falso.

A questão agora é mostrar que há um elo entre a possibilidade primária e a possibilidade metafísica. O bidimensionalismo epistêmico adota duas formas de avaliação de uma sentença. A verificação relaciona um cenário hipotético a uma sentença: podemos dizer que um cenário hipotético x verifica a sentença S se S é verdadeira no cenário x . Ao mesmo tempo, a satisfação relaciona um mundo possível a uma sentença: dizemos que um mundo possível w satisfaz a sentença S se S é verdadeira em w . A intensão primária, mencionada anteriormente, relaciona os cenários hipotéticos às extensões primárias e é capturada pela relação de verificação. Por outro lado, a intensão secundária relaciona mundos contrafactuais às extensões e é capturada pela relação de satisfação.

Para relacionar as possibilidades epistêmicas e as possibilidades metafísicas, Chalmers recorre ao que [Alves \(2023\)](#) denomina de princípio da ponte (PP). Esse princípio busca uma correlação entre cenários epistêmicos e mundos possíveis considerados como atuais. Nesse sentido, cada cenário epistêmico hipotético corresponderá a um mundo possível tomado como atual (possibilidade epistêmica), o qual, em última instância, revelará como o mundo poderia ser (possibilidade metafísica).

Segundo PP, para cada relação de verificação $ver(c, S)$, corresponde uma relação de verificação $ver(w, S)$, na qual o mundo w considerado como atual substitui seu cenário correspondente c . Podemos dizer, portanto, que para cada cenário, há um mundo tomado como atual que verifica as mesmas sentenças que são verificadas pelo cenário.²⁸

Ou seja, o PP correlaciona cenários epistêmicos com mundos possíveis tomados como atuais, e tais mundos verificam as sentenças verificadas pelos respectivos cenários. Para entender como funciona o PP, suponha o cenário no qual queremos verificar a sentença “Ouro não é o elemento de número atômico 79”. Para conceber um cenário x que verifica essa sentença, podemos imaginar uma situação na qual os cientistas revelam que estávamos errados e que o número atômico do ouro, na verdade, é 100 — de modo que a sentença “Ouro não é o elemento de número atômico 79” é verdadeira nesse cenário. Assim, temos: $ver(x, \text{“Ouro não é o elemento de número atômico 79”})$. Segundo o PP, o cenário descrito corresponderá a um mundo possível w_1 , tomado como atual, que também verifica a mesma sentença. Logo, temos: $ver(w_1, \text{“Ouro não é o elemento de número atômico 79”})$.

Por considerar que termos de tipo natural funcionam como designadores

²⁸ Cf. [ALVES, 2023](#), p. 309.

rígidos — ou seja, que termos como “ouro” designam o mesmo tipo natural em qualquer mundo possível em que o ouro exista —, a semântica modal de Kripke estabelece que os mundos possíveis nos quais a sentença “Ouro não é o elemento de número atômico 79” é verdadeira não são mundos em que o ouro existe, mas sim mundos em que há um outro tipo natural, semelhante ao ouro. Isso porque, no mundo atual, estabeleceu-se que o ouro tem número atômico 79 — e por ser um tipo natural, funciona como designador rígido —, e, portanto, em qualquer mundo possível onde o ouro exista, ele terá esse mesmo número atômico.

Podemos ilustrar como o exemplo funciona utilizando as noções de verificação e satisfação. A relação de satisfação envolve mundos possíveis e sentenças, de modo que não há nenhum mundo possível w tal que se obtenha a relação $\text{sat}(w, \text{“Ouro não é o elemento de número atômico 79”})$. A diferença entre algo ser concebível e esse mesmo algo constituir uma possibilidade metafísica é evidenciada pelo fato de que $\text{ver}(w_1, \text{“Ouro não é o elemento de número atômico 79”})$ pode ser obtida, enquanto $\text{sat}(w, \text{“Ouro não é o elemento de número atômico 79”})$ não o é.

Compreendidas as relações de verificação e satisfação, podemos aplicá-las ao caso dos zumbis filosóficos. A concebilidade de “ $P \ \& \neg Q$ ” (o mundo zumbi) nos permite imaginar um cenário x no qual obtemos $\text{ver}(x, \text{“}P \ \& \neg Q\text{”})$. Segundo o princípio da ponte (PP), isso nos conduz a um mundo w_1 , tomado como atual, no qual também se verifica a relação $\text{ver}(w_1, \text{“}P \ \& \neg Q\text{”})$. O próximo passo é mostrar que a relação $\text{sat}(w, \text{“}P \ \& \neg Q\text{”})$ é obtida — ou seja, que existe um mundo contrafactual w correspondente a w_1 no qual “ $P \ \& \neg Q$ ” é verdadeira. Para alcançar essa relação de satisfação e estabelecer o elo entre possibilidade epistêmica e possibilidade metafísica, Chalmers recorre novamente ao princípio da ponte, que estabelece o seguinte: se a intensão primária e a intensão secundária de uma sentença S coincidirem, então, para qualquer mundo w_1 que verifique a sentença S , existe um mundo contrafactual w correspondente a w_1 que satisfaz S .

Isso significa que, se as intensões primária e secundária de P e Q — isto é, a conjunção de todas as verdades físicas e a conjunção de todas as verdades fenomenais, respectivamente — coincidirem, então obtemos $\text{ver}(w_1, \text{“}P \ \& \neg Q\text{”})$ e, conseqüentemente, também obtemos $\text{sat}(w, \text{“}P \ \& \neg Q\text{”})$. Para justificar que as intensões primária e secundária de “ $P \ \& \neg Q$ ” coincidam, Chalmers recorre à descrição canônica. Em termos gerais, a descrição canônica de um mundo w é aquela formulada em termos semanticamente neutros. Por sua vez, uma expressão é considerada neutra se não puder ser utilizada por um falante, em uma situação contrafactual, com um significado diferente daquele que possui no mundo atual.

Por exemplo, 'água' e 'Hespéro' não são neutros; mas 'e', 'filósofo', 'amigo', 'consciência' e 'causa' plausivelmente são. Para uma primeira aproximação, uma expressão é neutra se não for 'Terra-gêmeável' (*twin-earthable*): isto é, se não se pode elaborar um caso de Terra-Gêmea onde um gêmeo de um usuário totalmente competente da expressão use a expressão correspondente com um significado diferente. (CHALMERS, 2011, p. 71) ²⁹

Ou seja, se admitirmos que todas as verdades físicas (P) e todas as verdades fenomenais (Q) podem ser representadas em termos semanticamente neutros — de modo que "P & ¬Q" constitua uma descrição canônica —, então podemos afirmar, de certo modo, que $ver(w_1, "P \& \neg Q")$ implica que $sat(w, "P \& \neg Q")$. A conclusão é que o mundo zumbi configura uma possibilidade metafísica e, portanto, que a doutrina materialista é falsa. Seguindo Alves (2023), podemos reformular o argumento do zumbi, adicionando o PP de Chalmers:

- (1) $ver(x, "P \& \neg Q")$ [P&¬Q é concebível]
- (2) PP: correspondência entre x e w_1
- (3) $ver(w_1, "P \& \neg Q")$
- (4) As intensões primárias e secundárias de "P&¬Q" coincidem
- (5) PP: correspondência entre w_1 e w
- (6) $sat(w, "P \& \neg Q")$
- (7) (P&¬Q)

C: O materialismo é falso.

O argumento movido por Chalmers pode ser visto como o mais refinado, mas não isento de críticas. De todo modo, não entraremos nas objeções. Também vale ressaltar que nem todos os detalhes sobre a semântica bi-dimensional foram abordados aqui. Para mais detalhes, sugiro a leitura de Chalmers (2002). No final capítulo 3, retomaremos a discussão sobre concebilidade.

Uma saída rápida para o materialista é afirmar que não é o caso de mundos zumbis serem concebíveis. Mas tal afirmação não é fácil de ser sustentada, pois há um forte apelo intuitivo envolvendo a concebilidade de zumbis. Kirk

²⁹ "For example, 'water' and 'Hesperus' are not neutral; but 'and', 'philosopher', 'friend', 'consciousness', and 'cause' plausibly are. To a first approximation, an expression is neutral if it is not "Twin Earthable": that is, if one cannot devise a Twin-Earth case where a twin of a fully competent user of the expression uses their counterpart expression with a different meaning."

(2005) formula um argumento no qual busca mostrar que a concebibilidade de mundos zumbis implica conceber uma visão contraditória sobre a consciência e, por este motivo, a concebibilidade de mundos zumbis deveria ser rejeitada. Na segunda seção do próximo capítulo, veremos tal argumento. Mas antes, vejamos a posição pró-zumbi de Kirk formulada em dois artigos.

Capítulo 3

Argumentos pró-zumbis e anti-zumbis de Robert Kirk

3.1 Argumentos pró-zumbi

Nesta seção, nos dedicaremos a apresentar a posição de Kirk a favor da concebibilidade de zumbis, defendida em meados dos anos 70. Ao final da seção, apresentaremos uma subseção com comentários sobre deficiências que Kirk apresenta ao defender os zumbis.

3.1.1 O caso de Zulliver

No artigo intitulado ‘Zombies vs Materialists’ [1974a](#), Kirk apresentará a objeção contra a formulação fisicalista que, segundo ele, se resume a dizer que ‘um homem não é nada mais que um objeto físico’. De acordo com Kirk, se aceitarmos esse slogan fisicalista, devemos aceitar que uma réplica de um homem também é uma réplica física, uma vez que o homem é um objeto físico e nada mais.

Em tese, há duas posições para tal formulação do que é uma réplica física: podemos descrever em relação a propriedades ou a descrições. Para a primeira, podemos dizer que uma coisa é uma réplica (réplica física) de outra se e somente se cada uma tiver exatamente as mesmas propriedades (propriedades físicas) que a outra. Porém, Kirk irá descartar a formulação em termos de propriedades³⁰. A outra posição é em termos descritivos: y é uma réplica física de x sse y tem todas as descrições físicas de x. Assim, ele defende que:

(A) Definição de réplica física: y no tempo T_2 é uma réplica (física) de x no

³⁰ Kirk não aponta os motivos pelos quais irá abandonar a formulação em termos de propriedades.

tempo T_1 sse toda descrição física não relacional que se aplica a x em T_1 se aplica a y em T_2 ³¹.

Ou seja, se y é réplica física de x , toda descrição física não relacional que se aplica a x e se aplica a y ³². Por “descrição não relacional”, Kirk entende como “aquelas descrições de x cuja aplicação a x é logicamente independente da existência (passada, presente ou futura) tanto de particulares diferentes de x , quanto de tipos de coisas que não são componentes de x ” (KIRK, 1974a, p. 136)³³.

De acordo com Kirk, qualquer materialista que está comprometido com (A) também está comprometido com:

(B) Qualquer réplica física de um determinado homem em um determinado momento também seria uma réplica desse homem naquele momento³⁴.

Negar (B) seria admitir que certas descrições não-relacionais que se aplicam ao homem não se aplicariam a suas réplicas físicas e, portanto, haveria algo a mais no homem que não se aplicam a sua réplica física. Seria contradizer o slogan “o homem nada mais é que um objeto físico”. No entanto, podem existir visões materialistas que aceitem que pode haver mais em um homem além de suas características físicas, ainda que sustente que tais coisas são, em última instância, puramente físicas. Em tal materialismo, defenderia-se que os homens não passam de objetos físicos³⁵. Portanto, Kirk formula uma tese mais geral:

(C) Toda descrição não relacional que se aplica a um determinado homem em um determinado momento é implicada pela conjunção de todas as descrições não relacionais puramente físicas que se aplicam a ele naquele momento.

(C) é a Tese da Implicação. Devemos notar que em (C) não há somente uma descrição não relacional que se aplica à réplica, mas a conjunção de todas as descrições não relacionais, o que comporta visões materialistas citadas anteriormente.

Com base no que foi exposto, parece razoável supor que todo fisicalista está comprometido com a Tese da Implicação. Se todas as descrições psicológicas podem ser reduzidas (no fisicalismo) a descrições físicas, então toda réplica física de um ser humano contém todas as descrições correspondentes a esse ser. Rejeitar a formulação da Tese da Implicação implica, de forma implícita, sustentar a existência de uma descrição não relacional que se aplica a x , mas não a y ,

³¹ Cf. KIRK, 1974a, p. 137.

³² Cf. KIRK, 1977, p. 139.

³³ “as all those descriptions of x whose application to x is logically independent of the existence (past, present or future) both of particulars other than x , and of types of things which are not components of x .”

³⁴ Cf. KIRK, 1974a, p. 138.

³⁵ Cf. KIRK, 1974a, p. 138-9.

ou seja, que há algo a mais no ser humano que não se aplica à sua réplica. Em outras palavras, seria afirmar que o ser humano é mais do que um objeto físico³⁶. Como descreve Kirk, a Tese da Implicação seria falsa se, e somente se, existisse um homem cujas descrições físicas em um determinado momento não implicassem todas as outras descrições não relacionais que se aplicam a ele. Em outras palavras, se as descrições físicas de um homem não abarcassem completamente sua natureza, então a Tese da Implicação seria falsa. O homem não seria descrito como “nada além de um objeto físico”, ele seria algo a mais do que isso.

O que definitivamente nada mais é do que um objeto físico é a réplica física do homem ao qual se aplicam apenas as descrições físicas e tudo o que elas implicam - algo que podemos chamar de ‘réplica zumbi’. De fato, conceitualmente o zumbi nada mais é do que a réplica física a qual as descrições não-relacionais sobre consciência não são implicadas por descrições físicas. De acordo com Kirk,

Uma vez que o homem seria diferente de sua réplica Zumbi, mas não de maneiras descritíveis em termos puramente físicos, ele não seria “nada além de um objeto físico”, ou não no mesmo sentido transparentemente claro. Portanto, se o slogan do materialista deve ser entendido nesse sentido, ele é logicamente equivalente à Tese da Implicação. (Devo assumir para simplificar que, se as réplicas de zumbis são logicamente possíveis, as descrições que não se aplicam a elas são descrições de sensações. (KIRK, 1974a, p. 141).³⁷

Desse modo, é possível afirmar que todo materialista está comprometido com a tese (C). No entanto, Kirk afirma que é difícil para o materialista sustentar o “implicar” de forma estrita que a Tese da Implicação coloca, pois seria necessário buscar análises comportamentais ou neutras para alguns casos problemáticos – para que a Tese da Implicação fosse verdadeira:

Uma ilustração não psicológica pode ajudar. Suponha que o ‘vocabulário puramente físico’ seja restrito ao vocabulário da física e da química. O materialista diria que as árvores, por exemplo, são “nada além de objetos físicos”, embora uma palavra como “nodoso”, que se aplica a algumas árvores, esteja excluída de seu vocabulário preferido. Fornecer uma análise de ‘nodoso’ em termos físicos ou neutros seria de fato uma maneira de explicar por que sua aplicação não impli-

³⁶ Cf. KIRK, 1974a, p. 138.

³⁷ “Since the man would differ from his Zombie replica, but not in ways describable in purely physical terms, he would not be ‘nothing but a physical object’, or not in that same transparently clear sense. So if the materialist’s slogan is to be understood in that sense, it is logically equivalent to the Entailment Thesis. (I shall assume for the sake of simplicity that if Zombie replicas are logically possible, the descriptions which fail to apply to them are descriptions of sensations.”.

cava que as árvores fossem algo além de meros objetos físicos. (KIRK, 1974a, p. 140).³⁸

Ou seja, seria necessário fornecer uma análise de um termo que não está no vocabulário físico, para que aplicação do termo não implicasse que árvores tem algo a mais que algo meramente físico. Ou ainda sustentar que para aplicar tal termo – no caso, ‘nodoso’ – precisaríamos utilizar termos puramente físicos.

Entretanto, é um erro supor que um materialista não está de acordo com a Tese. Em “Reply to Don Locke on Zombies and Materialism” (1977), Kirk reforça que a noção de “implicar” que está presente na Tese deve ser tomada no sentido amplo³⁹:

uma descrição P implica uma descrição Q se existem alguns fatores lógicos ou conceituais que tornam impossível que P seja verdadeiro de algo do qual Q é falso. (KIRK, 1977, p. 263)⁴⁰

Ou ainda que é impossível que P se aplique a algo ao qual Q também não se aplica⁴¹.

Então (por mais surpreendente que isso possa parecer) muito pouco argumento adicional é necessário para mostrar que, nesta interpretação, a Tese da Implicação não é apenas suficiente, mas necessária, para qualquer materialismo digno desse nome. (KIRK, 1974a, p.141)⁴²

Conclui-se que o materialista está, de alguma maneira, comprometido com a Tese. Aquele que a negar, terá consequências indesejadas: estará, implicitamente, sustentando que existem algumas descrições não relacionais de um homem que não se aplicam a réplica física, contradizendo o slogan materialista. O slogan é logicamente equivalente à Tese da Implicação⁴³.

Feita a caracterização do que supostamente seria o materialismo e os seus

³⁸ “A non-psychological illustration may help. Suppose the ‘purely physical vocabulary’ is restricted to the vocabulary of physics and chemistry. The materialist would say that trees, for example, are ‘nothing but physical objects’, despite the fact that a word like ‘gnarled’, which applies to some trees, is excluded from his preferred vocabulary. To provide an analysis of ‘gnarled’ in physical or neutral terms would indeed be one way of explaining why its application did not imply that trees were anything but mere physical objects.”.

³⁹ É válido alertar que a noção de implicação usada no presente trabalho é a noção de implicação estrita (ou lógica).

⁴⁰ “a description P entails a description Q if there are some logical or conceptual factors which make it impossible that P should be true of something of which Q is false”.

⁴¹ Cf. KIRK, 1974a, p. 141.

⁴² “Then (surprising as this may appear) very little additional argument is required to show that on this interpretation the Entailment Thesis is not only sufficient, but necessary, for any materialism worthy of the name.”.

⁴³ Cf. KIRK, 1974a, p. 141.

pressupostos, Kirk apresenta o que seria o exemplo zumbi para desqualificar a Tese da Implicação e, portanto, o materialismo. O autor pede que suponhamos um indivíduo (a quem chamaremos de Gulliver) que exibe todos os comportamentos típicos de um ser humano, vivendo em comunidade com outros seres humanos. Em determinado momento, pequenas criaturas invadem o cérebro de Gulliver, bloqueando seus nervos centrais e outras terminações nervosas. Essas pequenas criaturas podem ser denominadas Brain Team. O Brain Team passa a controlar o cérebro de Gulliver, mantendo seus comportamentos normais, de maneira indistinguível de um ser humano. Após ser completamente controlado pelo Brain Team, esse ser – a quem chamaremos de Zulliver – apresenta comportamentos semelhantes aos de um ser humano. Seria possível afirmar que Zulliver é senciente? A resposta afirmativa a essa questão não pode ser sustentada apenas com base no comportamento de Zulliver, pois agora sabemos que Zulliver é uma marionete, sendo contraintuitivo afirmar que marionetes são sencientes apenas com base no comportamento⁴⁴.

Zulliver é, em grande parte, um contraexemplo a teoria behaviorista. Um behaviorista diria que Zulliver é senciente, apenas pelo fato de Zulliver ter comportamentos típicos de um ser humano em determinadas condições. O ponto de Kirk é defender que não podemos afirmar que Zulliver, sendo comandado pelo Brain Team, ainda é senciente só pelo fato de que ele ainda se comporta como um ser humano. Como foi dito, Zulliver é uma marionete e, à primeira vista, é plausível sustentar que ele não tem sensações. Se o behaviorismo filosófico está correto, parece que Zulliver é senciente, o que é contraintuitivo. Não podemos afirmar, a partir do comportamento de Zulliver, que ele é senciente ou não⁴⁵.

Zulliver me parece ser um contra-exemplo decisivo para o behaviorismo analítico, incluindo aquela variedade menos vulnerável que não insiste na possibilidade de análises comportamentais fragmentadas de conceitos mentais, mas apenas na afirmação de que o vocabulário psicológico fornece nada mais do que maneiras de falar sobre comportamento e disposições comportamentais. (KIRK, 1974a, p. 144).⁴⁶

Um materialista poderia apelar para a tese verificacionista contra o exemplo apresentado por Kirk. Assim, o materialista argumentaria que devem existir evidências empiricamente observáveis que indiquem a perda de senciência no caso de Gulliver. Caso não haja evidência comportamental que indique a perda

⁴⁴ Cf. KIRK, 1974a, p. 144.

⁴⁵ Cf. KIRK, 1974a, p. 150.

⁴⁶ “Zulliver seems to me to be a decisive counter-example to analytical behaviourism, including that less vulnerable variety which does not insist on the possibility of piecemeal behavioural analyses of mental concepts, but only on the claim that the psychological vocabulary provides no more than ways of talking about behaviour and behavioural dispositions.”.

de sciência de Gulliver, não haveria justificativas para considerar que Gulliver constitui um contraexemplo da Tese da Implicação e, por conseguinte, do materialismo.

De acordo com Kirk, Zulliver seria um contraexemplo ao suporte verificacionista:

Quando os Liliputianos, ou mesmo os contemporâneos de tamanho normal de Zulliver, o descreviam e o tratavam como tendo pensamentos e sentimentos, o faziam porque para eles Zulliver era indistinguível de um homem normal. Eles não tiveram que decidir se deveriam ou não o tratar como senciente: a questão simplesmente não surgiu. E é claro que eles estavam totalmente justificados em tratá-lo assim. (KIRK, 1974a, p. 149)⁴⁷

Em outras palavras, os indivíduos que conviviam com Zulliver o tratavam como senciente, pois ele se comportava como um ser humano normal. Não foi exigido pelos indivíduos que houvesse evidências para a mudança de sciência. No entanto, o fato de que Zulliver se assemelhava a um ser humano normal não implica que ele fosse senciente, como vimos. Da mesma forma, o mesmo fato (de que Zulliver se assemelhava a um ser humano normal) não implica que ele fosse insensível⁴⁸. O ponto é que, com base apenas no comportamento de Zulliver, não é possível afirmar que ele seja senciente nem que seja insensível.

Por conseguinte, uma vez que Zulliver é uma réplica física, segundo a Tese da Implicação, todas as descrições não relacionais – principalmente, as sobre consciência – deveriam ser implicadas pelas descrições não relacionais puramente físicas. No entanto, tal implicação não é possível, pois não é possível afirmar que Zulliver é senciente apenas por seu comportamento e sua estrutura física. Ou seja, descrições não relacionais sobre consciência não são implicadas por descrições não relacionais puramente físicas, o que constitui o contraexemplo a Tese da Implicação, segundo Kirk. Assim, concluímos o primeiro argumento de Kirk em favor dos zumbis. Na próxima seção, veremos outra forma de Kirk defender sua posição pró-zumbi.

⁴⁷ “When the Lilliputians, or for that matter Zulliver’s normal-sized contemporaries, described and treated him as having thoughts and feelings, they did so because he was to them indistinguishable from a normal man. They did not have to decide whether or not to treat him as sentient: the question simply did not arise. And of course they were fully justified in treating him so.”

⁴⁸ Cf. KIRK, 1974a, p. 149-150.

3.1.2 O caso de Dan

Curiosamente, embora Kirk afirme que o exemplo de Zulliver constitui um contraexemplo à tese verificacionista, em “Sentience and Behaviour”, [Kirk \(1974b\)](#) apresenta uma variação do exemplo zumbi. Nesta versão, é possível observar que Kirk está empenhado em demonstrar a possibilidade de haver evidências comportamentais que indiquem a perda de sentiência. Suponha-se um indivíduo (denominado Dan) que leva uma vida saudável, sem problemas de saúde. Em determinado momento, Dan sofre um acidente enquanto corta a mão e exhibe comportamentos típicos de alguém que está sentindo dor, como gritar de dor e afirmar “estou sentindo uma dor terrível”. No entanto, embora se comporte de maneira semelhante à de alguém que normalmente experimenta dor, ele, de fato, não sente nada, não ocorrendo nenhuma propriedade qualitativa associada à dor naquele instante.

Dan se sentiu confuso e espantado, pois não sentia nada, mesmo tendo sofrido um sério corte na mão e, além disso, ele gritava de dor e proferia frases típicas de alguém que está com dor.

Ele [Dan] continuou a insistir que realmente não havia sentido nenhuma dor e disse que, no que diz respeito a seus gemidos, estremecimentos, exclamações, reclamações e outros “comportamentos de dor”, eles foram uma série de acontecimentos sobre os quais ele teve nenhum controle: ele se sentiu “como uma marionete”; ele havia apenas notado sua ocorrência, como alguém pode notar o dedo se contraindo involuntariamente. De fato, ele disse que lhe parecia que toda a série de itens que constituíam seu “comportamento de dor” tinha sido uma série de espasmos involuntários extraordinariamente complicados. ([KIRK, 1974b](#), p. 44, grifo nosso)⁴⁹

De acordo com Kirk, a hipótese de que Dan não sentiu dor poderia explicar certas características de seu comportamento de espanto. A reação de espanto de Dan nesta situação pode ser interpretada como uma resposta genuína e involuntária ao seu comportamento de dor (sem a correspondente sensação de dor), e não como uma mera simulação. O sistema nervoso central de Dan seria capaz de gerar movimentos físicos genuínos. Assim, Kirk argumenta que não é logicamente contraditório supor que Dan se espante com seu comportamento apa-

⁴⁹ “He continued to insist that he really had not felt any pain, and he said that so far as his groans, winces, exclamations, complaints and other ‘pain-behaviour’ were concerned, they had been a series of happenings over which he had had no control: he had felt ‘like a puppet’; he had merely noticed their occurrence as one might notice one’s finger involuntarily twitching. Indeed, he said that it seemed to him that the whole series of items constituting his ‘pain-behaviour’ had been a series of extraordinarily complicated involuntary twitches.”

rentemente relacionado à dor, apesar de não estar efetivamente sentindo dor⁵⁰.

Sobre o comportamento controverso de Dan, Kirk nos lembra que não é contraditório conceber a situação de um indivíduo ter determinados comportamentos sem estar no estado mental. Atores conseguem estar em tal situação constantemente. A diferença é que Dan não está simulando, e a sua reação de espanto é a evidência de que não há simulação⁵¹.

Assim como Zulliver, Dan também representa um desafio para a tese behaviorista⁵². Se Dan se comporta (por meio de falas, expressões, etc.) de maneira similar a alguém que, tipicamente, está experimentando dor, isso implicaria que Dan está sentindo dor e, conseqüentemente, que ele é senciente, conforme as teses anteriormente mencionadas. No entanto, no caso de Dan, é possível oferecer uma explicação empírica de como seu comportamento (a reação genuína de espanto) pode indicar uma mudança na perda de sensibilidade, contrariando a tese verificacionista.

Suponha que, além da perda da capacidade de sentir dores, Dan continue perdendo outras capacidades sensoriais a cada 6 meses, como o olfato, o tato etc., mas continua proferindo frases típicas de quem têm sensações. Ao cheirar uma rosa, por exemplo, ele profere frases do tipo “essas rosas cheiram muito bem” e logo em seguida mostra reação genuína de espanto e angústia⁵³.

[...] apesar de não ter sentido nenhum cheiro, seus músculos faciais apenas formaram uma expressão de prazer e seus órgãos vocais produziram uma observação que só seria apropriada se ele realmente tivesse detectado e apreciado o perfume das rosas. A expressão facial e os movimentos de seus órgãos da fala eram todos (assim ele parece pensar) involuntários, assim como no caso de sua produção de sentenças “expressando” ou “descrevendo” dores em situações de dor. (KIRK, 1974b, p. 47)⁵⁴

Podemos pensar que os órgãos de Dan correspondem da maneira correta às experiências e por isso há a descrição das situações por meio da fala, no entanto, as experiências não produzem a mesma sensação que era produzida anteriormente, o que explica a reação de espanto.

⁵⁰ Cf. KIRK, 1974b, p. 45.

⁵¹ Cf. KIRK, 1974b, p. 44.

⁵² Na realidade, Dan constitui um contraexemplo a qualquer tese ou teoria que esteja comprometida com a Tese da Implicação.

⁵³ Cf. KIRK, 1974b, p. 47-8.

⁵⁴ “[...] despite his not having been aware of any scent, his facial muscles have just formed an expression of enjoyment, and his vocal organs have produced a remark that would have been appropriate only if he had really detected, and enjoyed, the scent of roses. The facial expression and the movements of his organs of speech were all (so he seems to think) involuntary, just as in the case of his production of sentences ‘expressing’ or ‘describing’ pains in pain-situations.”

Suponha que Dan tenha chegado ao estágio de só ter dois sentidos: a visão e a audição. Ele continua insistindo que seu comportamento diante de certas experiências é totalmente involuntário.

Além disso, mesmo quando é capaz de observar algumas partes de seu comportamento pela visão ou pela audição, é impotente para inibi-las. Ele se ouve "descrevendo" o sabor de uma maçã, digamos, mesmo quando (como ele insiste) está mais agudamente consciente da ausência de qualquer experiência correspondente. (KIRK, 1974b, p. 48)⁵⁵

Eventualmente, Dan chega ao estágio de perder a visão. Ele perde o controle sobre suas ações e precisa ser internado em um hospital psiquiátrico. As pessoas ao seu redor começam a temer que Dan tenha alcançado o estágio final de zumbi, uma vez que ele não possui mais nenhuma capacidade sensorial, exceto a audição. Suponha que Dan retorne para casa. Poderíamos então considerar que este indivíduo que retorna não é mais o mesmo Dan, mas sim outro indivíduo (denominado D).

Kirk argumenta, no entanto, que esse indivíduo que voltou do hospital psiquiátrico não é uma outra pessoa, mas um zumbi. O zumbi, nesse caso, não é outra pessoa. Para demonstrar isso, suponha-se que, a cada seis meses, D recupere uma capacidade sensorial. Ele começa a proferir frases como "posso ouvir novamente! Talvez eu recupere meus outros sentidos". Após mais seis meses, ele afirma: "agora posso ver, assim como ouvir!". Apesar dessa situação peculiar, D não apresenta nenhum sinal de angústia ou apreensão. Se de fato fosse outra pessoa ocupando o lugar de Dan, então seria razoável supor que essa pessoa apresentaria sinais de angústia ou apreensão, pois estaria vivenciando mudanças drásticas. Portanto, a hipótese de que uma outra pessoa tenha substituído Dan é refutada pela suposição de Kirk. Apenas a hipótese de que um zumbi tenha tomado o lugar de Dan permanece válida, de acordo com a suposição proposta⁵⁶.

Suponhamos, conforme Kirk, que entre o estágio inicial, os estágios intermediários da perda de senciência e o estágio completo de zumbi, a atividade neural de Dan seja indistinguível da atividade neural de um ser humano. Em outras palavras, qualquer observador que analise as atividades cerebrais de Dan não encontrará nenhum indício de perda de senciência. Nesse contexto, Kirk argumenta que, para o materialista, o experimento de Dan parece violar

⁵⁵ "Moreover, even when he is able to attend to some parts of his behaviour by sight or hearing, he is powerless to inhibit them. He hears himself 'describing' the flavour of an apple, say, even when (as he insists) he is most keenly aware of the absence of any corresponding experience." (KIRK, 1974b, p. 48).

⁵⁶ Cf. KIRK, 1974b, p. 50.

o Princípio da Razão Suficiente. Tal princípio postula que tudo deve ter uma causa, razão ou fundamento. Assim, ao aplicar o Princípio da Razão Suficiente a este caso, parece absurdo (sob a perspectiva materialista) supor que os sentidos de Dan foram perdidos sem uma razão, e que essa razão não esteja localizada no cérebro de Dan. Como pode o cérebro ter permanecido inalterado ao longo dos estágios de perda dos sentidos? Ou seja, se o funcionamento neural de Dan não apresenta nenhuma atividade correspondente à dor, parece razoável concluir que ele não está sentindo dor⁵⁷.

Essa posição adotada por Kirk contrasta com duas visões: a tese verificacionista e a outra é a visão de que “senciência decorre logicamente do fato de que zumbis são indistinguíveis de homens vivos” (KIRK, 1974b, p. 51)⁵⁸. Consideremos o seguinte argumento que pretende afirmar que a sentiência se segue do comportamento:

- a. todo comportamento humano é explicável em termos físico-químicos;
- b. desempenhar certas funções ligadas ao comportamento é uma condição suficiente para se ter sentiência.

Se o argumento estiver correto, conclui-se que, caso o zumbi se comporte da maneira sugerida por (b), então o zumbi seria sentiente, o que é contraintuitivo quando confrontado com o conceito tradicional de zumbi. A proposição (b) decorre da tese atribuída a David Lewis, segundo a qual uma experiência é definida pelo seu papel causal, existindo condições finitas que especificam a causa típica e o efeito típico⁵⁹. Assim, se Dan se comporta (por meio de falas, expressões, etc.) de maneira semelhante à de alguém que, tipicamente, está sentindo dor, isso implicaria que Dan está sentindo dor, o que, por sua vez, sugeriria que ele é sentiente. A perspectiva em questão sustenta que o cérebro é o órgão responsável por desempenhar o papel de "meio de campo" entre o estímulo proveniente dos órgãos sensoriais e o comportamento ou movimento do corpo. Em resumo, o cérebro é o órgão que desempenha um papel causal. Esse raciocínio abre espaço para a consideração de cérebros distintos do humano, mas que desempenham o mesmo papel causal. No entanto, isso não implica que o desempenho de papéis causais necessariamente resulte em sentiência⁶⁰.

O ponto que Kirk levanta é que, se isso estiver correto, então, nos casos envolvendo cérebros humanos, deve-se também considerar o seguinte: o desempenho de papéis causais não implica sentiência. Dado que a análise de papéis

⁵⁷ Cf. KIRK, 1974b, p. 50-51.

⁵⁸ “sentience follows logically from the fact that what I call Zombies are indistinguishable from living men.”

⁵⁹ Cf. LEWIS, 1966.

⁶⁰ Cf. KIRK, 1974b, p. 51-2.

causais não fornece informações sobre a estrutura cerebral física do ser humano, parece não haver contradição lógica em afirmar que poderiam existir seres conscientes desempenhando esses papéis causais⁶¹.

Se uma equipe fosse capaz de controlar o cérebro de um ser humano, os papéis causais estariam sendo desempenhados, mas isso não implicaria senciência. Pouco importa, por exemplo, se a equipe responsável (denominada Brain Team) é consciente ou não, ou qual a sua localização, entre outros aspectos. O que importa é o papel causal que está sendo gerido pela equipe.

A objeção verificacionista gira em torno da questão de como podemos distinguir um zumbi de um ser humano. Parece arbitrário afirmar que um x é um zumbi e y é um ser humano, como se estivéssemos apenas atribuindo etiquetas a cada um deles. Não sabemos o que fundamenta essa diferença, ou pelo menos isso não está claro.

Em resposta, Kirk afirma que, uma vez que se pode explicar a hipótese de Dan ter perdido a capacidade de sentir dor, não há tensão conceitual ou incongruência na hipótese de Dan⁶². A explicação proposta é que, quando Dan parece se comportar como se estivesse com dor, na verdade ele está realizando uma série de movimentos involuntários complexos. Ele não está afirmando que está com dor, nem negando, mas apenas apresentando lapsos involuntários.

Em ambos os artigos defendendo a posição a favor dos zumbis, Kirk reitera que não existe contradição aparente em supor tanto o caso de Zulliver quanto o caso de Dan. O que iremos nos deter na próxima seção são detalhes em ambos os argumentos que os tornam problemáticos.

3.1.3 Críticas a fase pró-zumbi de Kirk

Na presente seção, pretendo expor comentários sobre a primeira fase do pensamento de Kirk, na qual o autor defende a possibilidade de zumbis como uma objeção ao materialismo. Tais comentários não têm a intenção de refutar de maneira definitiva a primeira fase de Kirk, mas visam apenas apresentar suas inconsistências.

Em Kirk (1974a), o autor busca caracterizar o pressuposto com o qual qualquer materialista ou fisicalista estaria comprometido. Segundo ele, a Tese da Implicação não poderia ser rejeitada por um materialista, pois sua rejeição implicaria a aceitação de que há algo no homem além de um objeto físico. Isso

⁶¹ Cf. KIRK, 1974b, p. 52-3. E mesmo que a tese do papel causal sustentasse que a estrutura física é importante para o desempenho dos papéis causais, seria injustificável, à medida que não se pode descartar *a priori* que outras estruturas podem ser sencientes, desde que as funções relevantes sejam desempenhadas.

⁶² Cf. KIRK, 1974b, p. 58.

ocorre porque o conjunto de descrições físicas não relacionais não implicaria necessariamente todas as outras descrições não relacionais, incluindo aquelas referentes à consciência.

A primeira questão que podemos colocar é se todo materialista está de fato comprometido com a caracterização sobre o que é uma descrição não relacional. Novamente citando, Kirk entende que descrições não relacionais são “aquelas descrições de x cuja aplicação a x é logicamente independente da existência (passada, presente ou futura) tanto de particulares diferentes de x , quanto de tipos de coisas que não são componentes de x ” (KIRK, 1974a, p. 136)⁶³, onde x é um objeto ou indivíduo.

Suponha uma descrição mental não relacional de um indivíduo x . Um fisicalista pode rejeitar essa descrição, afirmando que a classe de descrições mentais é relacional, ou seja, depende de fatores que não integram a consciência. Um representacionista, por exemplo, refutaria a ideia de que as descrições mentais são não relacionais, pois sustenta que a mente cria representações mentais que refletem aspectos do mundo. Essas representações podem ser de natureza visual, linguística, conceitual ou qualquer outra forma por meio da qual a mente ‘represente’ o mundo. Por exemplo, ao observar uma árvore, a mente cria uma representação dessa árvore, permitindo ao indivíduo reconhecê-la e interagir com ela. Nesse sentido, para o representacionista, a maior parte — se não todas — das descrições mentais é relacional, pois depende de aspectos e componentes que não pertencem ao indivíduo. Podemos concluir que a caracterização do pressuposto materialista/fisicalista feita por Kirk é incompleta, pois não contempla todos os tipos de materialismo.

Podemos ainda questionar se todas as descrições puramente físicas são, de fato, não relacionais. Por exemplo, suponha um conjunto de descrições puramente físicas que descrevem a estrutura física do corpo humano. É inegável que o corpo humano interage com o meio no qual está inserido, como, por exemplo, através da percepção, do tato, da audição etc. Portanto, é questionável afirmar que todas as descrições puramente físicas do corpo humano sejam não relacionais, pois, para que tais descrições sejam verdadeiras, é necessário que o corpo interaja com o meio, em última instância. Por exemplo, descrições sobre o córtex sensorial são relacionais, uma vez que essa parte do cérebro desempenha um papel fundamental na distribuição das informações que recebemos – ou captamos – do mundo. Além disso, as descrições sobre o córtex sensorial, por exemplo, são relacionais pois estão em contato com outras partes do cérebro, como o

⁶³ “as all those descriptions of x whose application to x is logically independent of the existence (past, present or future) both of particulars other than x , and of types of things which are not components of x ”.

córtex pré-frontal. Desse modo, um fisicalista não precisa estar comprometido com a visão de que as descrições puramente físicas da estrutura corporal humana são descrições não relacionais e, assim, a caracterização de Kirk sobre o materialismo/fisicalismo perde força.

Outro ponto a ser questionado é que, à luz do que foi exposto na seção 2.2.1, a caracterização do materialismo/fisicalismo proposta por Jackson — ‘Qualquer mundo que seja uma réplica física mínima do nosso mundo será uma réplica *simpliciter* do nosso mundo’ — oferece mais vantagens do que a Tese da Implicação de Kirk, pois a tese de Jackson é neutra quanto a propriedades ou descrições, evitando o problema das descrições não relacionais que acabamos de discutir. Isso resulta em um maior comprometimento por parte dos fisicalistas com a tese proposta por Jackson, em relação à Tese da Implicação.

Tanto em Kirk (1974a) quanto em Kirk (1974b), os argumentos que exemplificam o argumento do zumbi parecem não apresentar similaridade significativa com os argumentos sobre os zumbis mais contemporâneos. A sofisticação argumentativa e o aprofundamento nos detalhes apresentados por Chalmers (2010) tornam os argumentos de Kirk ingênuos, oferecendo pouco desafio aos fisicalistas atuais. Assim, a fase pró-zumbi de Kirk deixa a desejar quanto o poder crítico contra as teses materialistas/fisicalistas em comparação com argumentos pró-zumbis mais recentes.

Em 1974a, Kirk propõe a ideia de criaturas minúsculas que invadem o cérebro de Gulliver e passam a controlá-lo. A primeira objeção fisicalista que surge é a de que a estrutura cerebral de Zulliver (agora sob o controle do Brain Team) não constitui uma réplica física do cérebro de um ser humano normal. Quando se assume que um zumbi é um ser cujo cérebro é controlado por criaturas, perde-se o ponto crucial de conexão com o fisicalismo/materialismo, a saber, que esse ser é fisicamente idêntico a um ser normal, mas carece de consciência fenomenal. O cérebro de Zulliver não é o mesmo que o de Gulliver; não se tratam de réplicas físicas. Kirk argumenta que não há diferenças relevantes entre os cérebros de Zulliver e Gulliver. Contudo, no cérebro de Zulliver, há pequenas criaturas que controlam as conexões neuronais e todas as outras funções cerebrais. Essa é uma diferença relevante, pois a função anteriormente desempenhada pela ativação dos neurônios agora é realizada por essas pequenas criaturas.

Por outro lado, em 1974b, Kirk propõe um tipo de consciência muito peculiar. Isso porque, ao supor que Dan perde uma capacidade sensorial a cada período de tempo, Kirk está supondo que a consciência é modular – talvez involuntariamente –, ou seja, que existem divisões no cérebro que atuam de forma isolada e que o mal funcionamento dessas divisões não afeta o funcionamento

do cérebro por completo e nem sequer é possível ser percebido em análises do cérebro – utilizando aparelhos e técnicas como a ressonância magnética. Teorias da ciência cognitiva – que são mais influentes atualmente – já descartavam essa ideia na década de 80: a Global Workspace Theory (GWT), proposta por Bernard Baars (1988), sugere que a consciência surge quando informações de diversas regiões do cérebro são integradas e se tornam acessíveis a um "espaço de trabalho global". Nesse espaço, a informação pode ser utilizada por diferentes sistemas cognitivos, como memória, percepção e controle motor, permitindo uma coordenação flexível e dinâmica das funções mentais. A teoria oferece uma explicação de como a consciência pode ser vista como um processo de integração e disseminação global de informações, permitindo que o cérebro tenha acesso simultâneo a diversos tipos de dados.

Outra teoria influente formulada mais recentemente é a Teoria da Integração de Informação (IIT), proposta por Giulio Tononi (2007), que sugere que a consciência está relacionada à capacidade de um sistema de integrar informações de forma altamente interconectada. A IIT sugere que a consciência emerge quando um sistema exibe um alto grau de "integração" de suas partes, formando uma rede que não pode ser decomposta em partes independentes. Nesse sentido, quando Kirk argumenta que não há incoerência em supor que Dan pode perder a capacidade sensorial aos poucos, sem afetar e sem ser percebido em outras partes do cérebro, ele está caminhando em direção contrária a maioria das teorias cognitivas influentes, além de supor que a consciência é modular, quando a maioria das teorias compreende a consciência como um sistema integrado e de compartilhamento de informações⁶⁴.

Em certa medida, pode-se afirmar que a fase pró-zumbi de Kirk se restringe a criticar uma parcela limitada do conjunto de teses materialistas, as quais, atualmente, são pouco defendidas. Em ambos os artigos, Kirk parece estar mais preocupado em atacar, de forma particular, o behaviorismo analítico, o qual, na década de 1970, já não estava em ascensão e perdia espaço nos debates para novas teses materialistas, como a teoria da identidade e os tipos de funcionalismo. Dessa forma, caso o objetivo de Kirk fosse refutar o materialismo em sua totalidade, ele deveria ter concentrado sua análise principalmente nas teses mais contemporâneas à época, não em teses que já estavam com baixa popularidade entre os filósofos materialistas.

Desse modo, pode-se concluir que a fase pró-zumbi de Kirk, em ambos os artigos, carece de detalhes argumentativos ao caracterizar o que é o zumbi e

⁶⁴ Ainda que não se tenha um consenso na neurociência sobre qual modelo é o mais corroborado para explicar a consciência, a maioria dos modelos convergem para interpretar a consciência como sistema integrado de informações.

como podemos conceber esse ser, cometendo erros grosseiros que comprometem o poder crítico de seus argumentos. Além disso, Kirk falha ao caracterizar o que é o materialismo/fisicalismo, adotando uma postura crítica em relação a teses materialistas que já não possuem grande influência. Conclui-se que sua crítica aos materialistas se limita a uma pequena parte deles, sendo, ainda assim, um argumento ingênuo que oferece pouco ou nenhum risco para os materialistas/fisicalistas contemporâneos.

3.2 Argumento anti-zumbi: a história e-qualia

Nesta seção, abordaremos o argumento anti-zumbi proposto por Robert Kirk (2005), no qual o alvo é a afirmação de que os zumbis são concebíveis.

A ideia de zumbi pressupõe o fechamento causal do domínio físico, e, por essa razão, quem defende a hipótese dos zumbis deve sustentar que as propriedades que possuímos e que os zumbis supostamente carecem (as qualia) não desempenham um papel essencial na causação ou explicação do comportamento⁶⁵. Para Kirk, qualia são as propriedades dos estados mentais ou experiências que caracterizam esses estados pela forma como é experimentá-los. Por exemplo, há uma qualidade fenomenal particular na experiência de sentir dor, que caracteriza essa experiência. Essa interpretação sobre as qualia é semelhante em alguns aspectos a interpretação de Chalmers, mas não se compromete com o ponto principal: afirmar se são propriedades físicas ou não-físicas:

Se, por exemplo, estou experimentando o cheiro de eucalipto, poderia descrever a situação (não elegantemente) dizendo que minha experiência olfativa tem uma qualidade de eucalipto. Observe que a definição de qualia de Chalmers é neutra quanto a se ter qualia envolve algo não-físico. Claro que ele acha que sim; mas ele não acha que o assunto está resolvido por definição. (KIRK, 2005, p. 37)⁶⁶

Para Kirk, a definição proposta por Chalmers permite afirmar que, dentro dessa perspectiva, a consciência não desempenha um papel explicativo ao esclarecermos o fato de sermos capazes de falar, escrever, pensar, entre outras ações. No entanto, essa visão parece entrar em conflito com nossas intuições sobre a vida experiencial e perceptiva. Por exemplo, em diversos momentos, afirmamos que nossas experiências perceptivas influenciam nosso comportamento. Dessa

⁶⁵ Cf. KIRK, 2005, p. 37.

⁶⁶ "If for example I am experiencing the scent of eucalyptus, I could describe the situation (not elegantly) by saying my olfactory experience has a eucalyptus quale. Note that Chalmers's definition of qualia is neutral as to whether having qualia involves anything non-physical. Of course he thinks it does; but he doesn't think the matter is settled by definition."

forma, parece que nossas experiências conscientes são fundamentais para o controle de nossas ações. E é justamente esse ponto que um defensor da hipótese dos zumbis não aceitaria.

Eles [**defensores dos zumbis**] estão comprometidos com a visão de que, mesmo nos casos em que temos bastante tempo para refletir sobre a aparência das coisas, como quando um pintor estuda o efeito de pinceladas recentes e depois adiciona outra cuidadosamente, as experiências conscientes não contribuem para o monitoramento ou controle do comportamento. Portanto, embora os zumbis apelem para as intuições comuns em apoio à possibilidade dos zumbis, sua posição também entra em conflito com as intuições comuns. (KIRK, 2005, p. 38, grifo nosso)⁶⁷

Aqui, Kirk está explicitamente tentando demonstrar de que maneira o argumento do zumbi nos levaria ao epifenomenalismo, à medida que poderíamos rejeitar a causalidade das qualia na explicação do comportamento. O ponto central é que rejeitar a tese de que a consciência influencia e modifica o nosso comportamento implica rejeitar uma intuição forte, profundamente enraizada na nossa compreensão de nós mesmos. Embora rejeitem tal tese, os epifenomenalistas e defensores da hipótese dos zumbis estão convencidos de que as experiências conscientes são extremamente importantes e merecem muita atenção, sendo frequentemente discutidas e analisadas⁶⁸. Com base nessa constatação, a proposta do restante do capítulo é examinar como Kirk argumenta que a noção de zumbi envolve uma concepção incoerente da experiência fenomenalmente consciente.

Kirk acusa os defensores dos zumbis de propagar a ‘falácia da jaqueta’. A ideia dessa falácia é ilustrar que, em certos casos, as propriedades não podem ser removidas sem afetar as propriedades remanescentes. O que o defensor do zumbi propõe é justamente que a consciência fenomenal é uma propriedade que pode ser retirada, deixando intactas as propriedades remanescentes. O argumento de Kirk é que essa suposição está errada. No entanto, não basta apenas afirmar que as intuições estão equivocadas; é necessário apresentar um argumento convincente.⁶⁹

O argumento construído para eliminar quaisquer intuições a favor dos

⁶⁷ “They are committed to the view that even in cases where we have plenty of time to reflect on how things look, as when a painter studies the effect of recent brush-strokes and then thoughtfully adds another, conscious experiences make no contribution to the monitoring or control of behaviour. So although zombists appeal to ordinary intuitions in support of the zombie possibility, their position also conflicts with ordinary intuitions.”

⁶⁸ Cf. KIRK, 2005, p. 38.

⁶⁹ Cf. KIRK, 2005, p. 39.

zumbis é descrito da seguinte forma:

- (I) A história e-qualia não é concebível;
 - (II) Se os zumbis fossem concebíveis, a história e-qualia seria concebível;
- Por modus tollens, temos,
Portanto: (III) Zumbis são inconcebíveis.

O argumento precisa de duas etapas: primeiro, é preciso provar que a história e-qualia não é concebível e, posteriormente, provar que a concebilidade dos zumbis implica a história e-qualia. Mas, primeiramente, é preciso esclarecer o que é a história e-qualia. O termo “e-qualia” é composto por qualia – qualidades subjetivas específicas de estados mentais - e pelo prefixo “e-” que é uma abreviação de “epifenomenal”. Ou seja, as e-qualia são as qualia epifenomenais, que não tem efeitos físicos⁷⁰. Sendo assim, a história e-qualia é uma história sobre a consciência epifenomenal, descrita em 5 teses.⁷¹

- (E1) O mundo é parcialmente físico, e todo o seu componente físico está fechado sob causalidade: todo efeito físico tem uma causa física.
- (E2) Os seres humanos mantêm alguma relação com um tipo especial de propriedades não físicas, as e-qualia. As e-qualia são o que tornam os seres humanos são fenomenalmente conscientes.
- (E3) As e-qualia são causadas por processos físicos, mas não têm efeitos físicos: podem ser eliminadas sem perturbar o mundo físico.
- (E4) As seres humanos consistem apenas em corpos funcionais e em suas e-qualia relacionadas.
- (E5) O ser humano é capaz de perceber, atentar, pensar e comparar suas e-qualia.

(E1) é a tese conhecida como o fechamento causal do domínio físico. (E1) afirma que todo efeito físico possui uma causa física. Nesse contexto, propriedades

⁷⁰ Kirk diz que é correto afirmar que o epifenomenalismo implica a possibilidade de um mundo zumbi, porque de acordo com o epifenomenalismo, a consciência depende essencialmente de propriedades não físicas. Sendo não físicas, essas propriedades especiais não podem depender logicamente ou *a priori* do mundo físico; portanto, o mundo físico poderia ter existido sem essas propriedades.

Entretanto, a possibilidade de zumbis não implica o epifenomenalismo, segundo Kirk, porque aparentemente um mundo zumbi poderia ser possível mesmo que o nosso mundo fosse interacionista, ou seja, mesmo que houvesse interação entre propriedades não físicas (mentais) e propriedades físicas, essas propriedades não físicas poderiam não existir ou serem ociosas em um mundo zumbi. Então, a possibilidade de zumbis não implica o epifenomenalismo.

⁷¹ Reiterando o que o próprio Kirk afirma: a história e-qualia não pretende ser um relato fiel da visão epifenomenalista e isso não é uma questão relevante, o que é relevante é que a concebilidade dos zumbis implica a concebilidade da história e-qualia. Ademais, a história e-qualia não faz jus a todos os pontos de vista sobre as qualia, algumas autores poderiam discordar. Ver mais (KIRK, 2005, p. 40).

físicas têm causas físicas. Propriedades não-físicas, como as qualia, podem ter causas físicas – no caso das qualia, elas possuem causas físicas, conforme descrito em (E3). (E3) pode parecer peculiar, e alguém pode questionar: se o mundo físico é causalmente fechado, dado (E1), como algo pode ser causado por processos físicos e, ao mesmo tempo, não ter efeitos físicos? A confusão que surge dessa questão está em interpretar (E1) como postulação de que toda causa física deve ter um efeito físico. Na realidade, se (E1) postulasse que toda causa física tem um efeito físico, então os processos físicos causariam as e-qualia, e as e-qualia seriam, portanto, propriedades físicas. No entanto, (E1) postula exatamente o contrário: (E1) afirma que todo efeito físico possui uma causa física. Dessa forma, as e-qualia podem ser causadas por processos físicos, sem que gerem efeitos físicos.

(E2) esclarece que essas propriedades são não-físicas; portanto, mesmo que sejam causadas por processos físicos, elas são inertes quanto à sua causalidade em comportamentos. Se todo efeito físico possui causas físicas e as e-qualia são não-físicas, então as e-qualia não podem causar nada, o que leva à conclusão de (E3). Kirk não descarta a possibilidade de sobre-determinação em outros mundos, embora afirme que a sobre-determinação não poderia ocorrer no mundo descrito pelas teses (E1)-(E5). A sobre-determinação causal é um conceito fundamental em nossa discussão. A ideia central é que, se um determinado efeito for sobre-determinado, ele possui mais de uma causa. No caso das e-qualia, a sobre-determinação ocorreria de tal maneira que tanto causas físicas quanto as e-qualia seriam responsáveis por efeitos físicos. Em outras palavras, na sobre-determinação, as e-qualia desempenhariam algum papel causal nos efeitos físicos. Segundo Kirk, se (E3) for concebível, então não há necessidade de se preocupar com a sobre-determinação⁷². Posteriormente, retomaremos a noção de sobre-determinação causal.

Além disso, não poderia haver causalidade entre e-qualia no mundo descrito pelas teses (E1)-(E5); ou seja, não é possível que e-quale tenha efeitos sobre outra e-quale. Para ilustrar por que isso é impossível, suponha que uma e-quale, Q1, participe da causação de outra e-quale, Q2. De acordo com (E3), sabemos que as e-qualia são causadas por processos físicos, portanto, Q1 é inerte em relação à causação de Q2: os processos físicos são responsáveis por toda a causalidade. Assim, as e-qualia não têm poder causal nem no mundo físico, nem entre si.

Se o comportamento é afetado apenas pelos processos físicos, então podemos remover as e-qualia do mundo sem que o mundo físico seja alterado, como é afirmado em (E3). Como supracitado no início da seção, Kirk ressalta que as

⁷² Cf. KIRK, 2005, p. 41.

e-qualia são distintas da noção de qualia defendida por outros autores, como, por exemplo, David Chalmers. No entendimento de Chalmers, afirmar que as qualia existem é afirmar que somos fenomenalmente conscientes. As qualia de Chalmers apresentam um ponto importante: uma vez que sua definição não implica que essas propriedades devam ser não-físicas, a existência de qualia, em seu sentido, pode ser aceita até mesmo pelos fisicalistas. Além disso, essa definição não implica que as qualia possam ser removidas sem afetar o mundo físico, ao contrário do que é pressuposto por (E3)⁷³.

No caso da história e-qualia, de acordo com (E2), o que nos torna fenomenalmente conscientes é a relação com as e-qualia. Em outras palavras, afirmar que somos fenomenalmente conscientes é afirmar que estamos em uma relação com essas propriedades não-físicas, o que justifica (E5) — ou intimidade epistêmica, como é denominado por Kirk. No entanto, é precisamente contra esse ponto que Kirk irá argumentar. Segundo ele, a história e-qualia é incoerente, pois, se (E1)-(E4) forem verdadeiras, não podemos ter a intimidade epistêmica descrita por (E5)⁷⁴. Se aquilo que nos torna fenomenalmente conscientes não possui efeitos físicos, então não seríamos capazes de perceber, atentar, pensar ou comparar nossas e-qualia.

Suponha que estou provando dois vinhos. Posso comparar como é para mim experimentar seus sabores, o que pode me levar a mudar minhas preferências e, portanto, meu comportamento: posso parar de beber um e beber mais do outro. Certamente parece que, ao contrário da concepção e-qualia de consciência, as qualidades de nossas experiências têm efeitos físicos. Naturalmente, os epifenomenalistas dedicaram muita atenção a essa objeção. Eles admitem que parece que as qualidades de nossas experiências têm efeitos em nosso comportamento, mas sustentam que estamos enganados. (KIRK, 2005, p. 41)⁷⁵

Se as e-qualia são inertes, como o epifenomenalista poderia explicar a realização da intimidade epistêmica, incluindo a percepção, a atenção, o pensamento e a comparação de nossas experiências qualitativas? Os epifenomenalistas não poderiam adotar um ceticismo absoluto e rejeitar a ocorrência dessas atividades cognitivas, pois tais atividades são parte integrante de nossas vidas. Tampouco poderiam afirmar que essas atividades cognitivas são realizadas por outras pro-

⁷³ Cf. KIRK, 2005, p. 40.

⁷⁴ Cf. KIRK, 2005, p. 41.

⁷⁵ "Suppose I am tasting two wines. I might compare what it is like for me to experience their flavours, which might cause me to change my preferences, hence my behaviour: I might stop drinking one and drink more of the other. It certainly seems that, contrary to the e-qualia conception of consciousness, the qualities of our experiences have physical effects. Naturally epiphenomenalists have devoted a lot of attention to this objection. They concede it seems as if the qualities of our experiences have effects on our behaviour, but maintain that we are mistaken."

priedades não-físicas, pois isso violaria a tese (E4). A intimidade epistêmica envolve processos cognitivos, como o armazenamento e a recuperação de informações, entre outros, que são tipos de processos causadores de mudanças; por exemplo, o processo de conceituação envolve transformações nas informações. Como as e-qualia são inertes, elas não são capazes de realizar essas atividades cognitivas de forma autônoma. Portanto, as atividades cognitivas que constituem a intimidade epistêmica devem ser realizadas pelos corpos e seus processos físicos. No entanto, como veremos posteriormente, Kirk argumenta que os corpos e processos físicos não são capazes de realizar a intimidade epistêmica⁷⁶.

Uma resposta à pergunta anterior mencionada por Kirk é apresentada por Chalmers. De acordo com Kirk, Chalmers sugere que, embora as e-qualia não exerçam influência causal, “sua mera presença no contexto físico adequado constitui parcialmente o conteúdo dos pensamentos envolvidos: ajuda a garantir que nossos pensamentos sejam sobre essas e-qualia” (KIRK, 2005, p. 44). Kirk argumenta que o ponto central de sua contestação é exatamente aquilo que Chalmers considera como certo: a existência de uma experiência fenomenal com a qual estamos intimamente conectados. Se de fato existe algo como a intimidade epistêmica, como poderíamos estar conectados a ela sem que isso tenha efeitos sobre nosso comportamento? Para demonstrar que a história e-qualia exclui a intimidade epistêmica, Kirk pede que suponhamos a existência de um irmão gêmeo zumbi seu (denominado Zob). A diferença entre Kirk e Zob é que Zob carece de e-qualia e, portanto, segundo a história e-qualia, carece de consciência. Suponhamos que as teses (E1)-(E4) sejam concebíveis, ou seja, não sejam incoerentes *a priori*. Kirk sustenta que (E2) e (E4) garantiriam que a associação de e-qualia em Zob resultaria em um ser consciente. Vamos supor que as leis naturais que asseguram que os processos corporais causam e-qualia comecem a operar no mundo de Zob, e, assim, Zob passe a ter e-qualia causadas por seus processos corporais⁷⁷.

Suponha ainda que uma mudança diferente no mundo de Zob fizesse com que imagens em movimento de tudo o que ele vê fossem exibidas nas solas de seus pés. A questão é se Zob poderia ter algum tipo de intimidade epistêmica com as e-qualia, agora que as leis do seu mundo permitem a existência de e-qualia, como (E5) exige. Embora as imagens sejam causadas por, e isomórficas com, as percepções visuais de Zob, nada garantiria que o funcionamento cognitivo de Zob reconhecesse a presença dessas imagens. Assim, a defesa de Kirk é que a presença das imagens no caso de Zob não produziu o tipo de intimidade epistêmica aqui indicada. De maneira semelhante ao caso descrito, as e-

⁷⁶ Cf. KIRK, 2005, p. 43.

⁷⁷ Cf. KIRK, 2005, p. 45.

qualia são causalmente impotentes e inertes; portanto, devemos sustentar que elas também não podem gerar o tipo de intimidade epistêmica necessária para perceber, atentar-se, pensar e comparar as e-qualia⁷⁸.

A história diz que suas imagens de solas não têm efeitos em seus processos perceptivos e cognitivos, que, lembre-se, refletem os meus. Dada essa informação, insistir que Zob sabia ou pensava sobre suas imagens de solas, ou que prestava atenção a elas, entraria em conflito com qualquer entendimento normal de conhecimento, intencionalidade e atenção. Quando ignoramos completamente um determinado assunto e nenhuma crença ou hipótese a respeito dele aparece em nossos pensamentos, não faz sentido dizer que estamos pensando sobre ele, muito menos que estamos prestando atenção a ele. (KIRK, 2005, p. 46)⁷⁹

Se o exemplo de Zob for excessivamente complexo, um exemplo mais simplificado pode ser apresentado: Kirk solicita que suponhamos que os processos eletroquímicos em nossos cérebros, responsáveis pelo processamento perceptivo, induzam correntes elétricas padronizadas, sem que essas correntes elétricas exerçam qualquer efeito sobre o nosso processamento perceptivo. Como as correntes elétricas não têm impacto algum, não seríamos capazes de percebê-las, compará-las, entre outras ações.

Se fosse esse o caso, nosso processamento cognitivo seria sobre essas correntes elétricas padronizadas? Nós as notaríamos ou prestaríamos atenção a elas? Seríamos capazes de compará-las ou estar em qualquer tipo de relação epistemicamente íntima com elas? Obviamente não, já que não sabemos nada sobre essas correntes e elas não têm efeitos em nosso processamento cognitivo. Não há nada que faça com que nossos pensamentos sejam epistemicamente relevantes para essa atividade elétrica. (KIRK, 2005, p. 46)⁸⁰

A argumentação de Kirk sugere que não há diferença relevante entre as e-qualia, as imagens únicas e as correntes elétricas padronizadas, uma vez que

⁷⁸ Cf. KIRK, 2005, p. 45-6.

⁷⁹ "The story has it that his solepictures have no effects on his perceptual and cognitive processes, which, remember, mirror my own. Given that information, to insist that Zob knew or thought about his sole-pictures, or attended to them, would conflict with any normal understanding of aboutness, intentionality, and attention. When we are completely ignorant of a certain matter, and no beliefs or hypotheses concerning it figure in our thoughts, it makes no sense to say we are thinking about it, still less that we are attending to it."

⁸⁰ "If that were the case, would our cognitive processing be about those patterned electric currents? Would we notice or attend to them? Would we be able to compare them, or stand in any sort of epistemically intimate relation to them? Obviously not, given we know nothing about such currents and they have no effects on our cognitive processing. There is nothing to make it the case that our thoughts would be epistemically relevant to such electrical activity."

todas são causadas por, e isomórficas aos, processos perceptivos. Desse modo, se o processo cognitivo de Z_{ob} não pode gerar intimidade epistêmica com as imagens e, da mesma forma, se o processamento cognitivo não é capaz de gerar intimidade epistêmica com as correntes elétricas, então o processamento cognitivo também não pode gerar intimidade epistêmica com as e-qualia. Portanto, a concepção de consciência nas teses (E1)-(E4) é incoerente com a noção de intimidade epistêmica da tese (E5). Ou seja, as quatro primeiras teses impedem que sejamos capazes de pensar, perceber, atentar-nos ou comparar as qualidades de nossas experiências. Assim, a incoerência da história e-qualia está relacionada ao fato de que, se as teses (E1)-(E4) forem verdadeiras, a tese (E5) será falsa, e, como isso foi demonstrado por meio de reflexões *a priori*, a história e-qualia é inconcebível⁸¹. Esse foi o passo para estabelecer a tese inicial (I), a saber, a história e-qualia não é concebível. Agora, resta estabelecer (II), se zumbis fossem concebíveis, a história e-qualia seria concebível.

Kirk não está comprometido com a tese de que a concebilidade do mundo zumbi implica que a história e-qualia é verdadeira no mundo real. À medida que a concebilidade do mundo zumbi é compatível com o interacionismo, essa tese é falsa. Kirk está defendendo a tese de que a concebilidade do mundo zumbi implica na concebilidade da história e-qualia. Uma vez que algo é concebível quando não é incoerente *a priori*, Kirk precisa mostrar que “enquanto os zumbistas se apegam à sua tese, suas implicações os forçam a sustentar que nenhuma reflexão *a priori* sobre a história de e-qualia revelaria incoerência” (KIRK, 2005, p. 48)⁸². O plano de Kirk é mostrar que se os zumbis fossem concebíveis, então concebivelmente um mundo zumbi poderia ser transformado em um mundo no qual a história e-qualia é verdadeira. Desse modo, se isso estiver correto, a concebilidade do mundo zumbi implicaria na concebilidade da história e-qualia.

Primeiramente, devemos lembrar que o mundo zumbi é causalmente fechado: todos os efeitos desse mundo são causados fisicamente. Desse modo, segundo Kirk, a concebilidade de zumbis implicaria a concebilidade de um mundo zumbi z tal que:

- (A1) z é um sistema puramente físico e causalmente fechado;
- (A2) Fisicamente, z é, na medida do possível, exatamente como o mundo real;
- (A3) Os habitantes semelhantes a humanos de z carecem de consciência fenomenal.

⁸¹ Cf. KIRK, 2005, p. 47-8.

⁸² “so long as zombists stick to their thesis, its implications force them to maintain that no amount of a priori reflection on the e-qualia story would reveal incoherence.”

Para que se tenha consciência fenomenal em *z* (seja *z* concebível), algo tem que ser adicionado a *z* e esse algo deve ser não-físico. Deve ser não físico porque os próprios processos físicos já presentes em *z* não dão conta de oferecer a consciência fenomenal aos habitantes de *z*.

Supondo que um mundo zumbi seja concebível, então, também é concebível que deva haver um item não físico ou itens *x* ('item' de agora em diante) que, quando apropriadamente associado a *z*, garantiria que seus habitantes adquirissem nosso tipo de consciência fenomenal? (KIRK, 2005, p. 49)⁸³

Kirk busca desenvolver um dilema para os defensores dos zumbis, que não poderiam refutar a ideia de que um item não físico associado a *z* garantiria a consciência fenomenal e a intimidade epistêmica. Como foi mencionado anteriormente, o item não físico que garante a consciência fenomenal no mundo atual é removido dos indivíduos no exemplo do zumbi. Se, então, adicionarmos esse item (denominado *y*) de volta ao mundo, os indivíduos recuperariam a consciência fenomenal. O objetivo do argumento é demonstrar que, no mundo *z*, o item não físico *x*, que produz consciência fenomenal e intimidade epistêmica, quando associado a *z*, gera um mundo idêntico ao descrito pela história *e*-qualia.

Kirk apresenta duas opções, dado que *y* é o item que garante a consciência fenomenal no mundo real: ou *y* tem efeitos físicos ou não tem efeitos físicos. Deve-se notar que o item que produz a consciência fenomenal e a intimidade epistêmica em *z*, a saber, o item não físico *x*, é análogo ao item não físico *y* que produz consciência fenomenal e intimidade epistêmica no mundo atual. Assim, o exemplo de Kirk sugere que, se o dilema for válido para o item não físico *y*, também deverá ser válido para o item não físico *x*.

O defensor dos zumbis poderia argumentar que *y* não tem efeitos físicos, o que implicaria que nosso comportamento é puramente causado por processos físicos. No entanto, uma vez que os defensores dos zumbis estão comprometidos com a tese de que o mundo zumbi *z* é essencialmente idêntico ao mundo atual, e que é possível retirar *y* do mundo atual, então adicionar a *z* algo semelhante a *y*, concederia consciência fenomenal aos zumbis de *z*. Se é possível retirar *y* do mundo atual, então deve ser possível adicionar algo semelhante a *y* no mundo zumbi e produzir consciência fenomenal. De fato, esse é um passo crucial para os defensores dos zumbis sustentarem que as propriedades físicas não são suficientes para produzir consciência. Por outro lado, os defensores dos zumbis

⁸³ "Assuming a zombie world is conceivable, then, is it also conceivable that there should be a non-physical item or items *x* ('item' from now on) which, when appropriately associated with *z*, would ensure that its inhabitants acquired our kind of phenomenal consciousness?"

poderiam afirmar que *y* tem alguns efeitos físicos. Kirk sustenta que, se a teoria dos zumbis é concebível, então também é concebível que a relação causal de *y* seja contingente, de modo que seus efeitos físicos poderiam ou não ter ocorrido. A conclusão a que Kirk chega é que, em todo caso, o defensor dos zumbis deve aceitar que o item não físico, como *y*, terá efeitos físicos.

Uma outra linha de argumentação que o defensor dos zumbis poderia adotar seria afirmar que *y* é necessário para a consciência fenomenal, mas não suficiente; nesse sentido, qualquer item não físico envolvido em nos tornar conscientes só seria capaz de fazê-lo se tivesse efeitos físicos. No entanto, os defensores dos zumbis não podem sustentar tal posição. Segundo Kirk, essa linha de defesa é insustentável, pois, em sua maioria, os defensores dos zumbis concordam que temos algum tipo de acesso privilegiado à nossa consciência fenomenal,

à existência de "algo que é como" para nós. É apenas porque temos esse acesso que, segundo eles, podemos saber que nós mesmos não somos zumbis. Além disso, ter esse tipo de acesso é o que nos permite contar como são nossas experiências sem ter que verificar se elas têm efeitos físicos. Quando me perguntam sobre o sabor do vinho, posso procurar palavras adequadas; mas não preciso tatear em busca da experiência em si. Também não preciso verificar se a experiência em si (ao contrário do álcool que a acompanha) tem algum efeito físico antes de poder dizer que estou tendo. Tais efeitos não podem, portanto, ser necessários para ter a experiência. (KIRK, 2005, p. 50)⁸⁴

De qualquer maneira, Kirk sustenta que os defensores dos zumbis não podem negar que a concebilidade dessas entidades implica a possibilidade de um item não físico *x*, o qual, ao ser associado ao mundo zumbi *z*, o transformaria em um mundo *z**, cujos habitantes, antes zumbis, passariam a possuir nosso tipo de consciência fenomenal. Haveria como escapar da conclusão? Uma outra saída, segundo Kirk, seria dizer que é concebível que *z* possa ser transformado em um mundo no qual os habitantes são fenomenalmente conscientes (*z**), mas não epistemicamente íntimos de suas experiências. No entanto, isso entra em conflito com uma posição defendida por Chalmers já citada anteriormente, na qual "ser epistemicamente íntimo das próprias experiências não é nem mesmo conceitualmente separável de ser fenomenalmente consciente: faz parte disso" (KIRK, 2005, p. 50).

⁸⁴ "to there being 'something it is like' for us. It is only because we have this access that, according to them, we can know we ourselves are not zombies. Further, having this sort of access is what enables us to tell what our experiences are like without having to check whether they have physical effects. When I am asked about the flavour of the wine, I may grope for suitable words; but I don't have to grope for the experience itself. Nor do I have to check whether the experience itself (unlike the alcohol which accompanies it) has any physical effects before I can tell I am having it. Such effects cannot therefore be necessary for having the experience."

Os defensores dos zumbis devem adotar a posição de que, embora y produza consciência fenomenal em nosso mundo, y não gera intimidade epistêmica. Caso aceitem que y produz tanto consciência fenomenal quanto intimidade epistêmica em nosso mundo, então seriam obrigados a aceitar que y produz essas mesmas qualidades em z (se y for adicionado a z, resultando em z*), dado que z é fisicamente idêntico ao mundo atual⁸⁵.

Se y não produz intimidade epistêmica, então ela deve ser produzida por itens e processos puramente físicos. No entanto, se a intimidade epistêmica é produzida por itens e processos puramente físicos, (A1) e (A2) já preveem a intimidade epistêmica, pois todos os processos físicos estão presentes dadas essas duas cláusulas. Uma vez que (A1) e (A2) também estão em z* (afinal, z* é cópia física idêntica a z), então a intimidade epistêmica também está presente em z* junto com a consciência fenomênica quando adicionamos algo semelhante a y (nesse caso, o item não físico x)⁸⁶. A conclusão permanece, a saber,

Os defensores dos zumbis, portanto, não podem objetar que é inconcebível que x forneça intimidade epistêmica, bem como consciência. Portanto, a conclusão anterior se mantém. Mesmo aqueles defensores dos zumbis que sustentam que há interação não física no mundo real devem admitir que, concebivelmente, um item não físico poderia ser associado a um mundo zumbi de modo a transformá-lo em um mundo z* onde os ex-zumbis desfrutassem de consciência fenomenal e intimidade epistêmica com suas experiências. (KIRK, 2005, p. 51)⁸⁷

Dessa forma, Kirk supostamente mina as chances defensores dos zumbis de rejeitar que, uma propriedade não-física x que produz consciência fenomenal associada a um mundo z, produziria um mundo no qual os ex-zumbis passam a ter consciência fenomenal e intimidade epistêmica. Portanto, os defensores dos zumbis não poderiam negar que a concebibilidade dos zumbis implicaria a concebibilidade do mundo z* que pode ser caracterizado pelas seguintes teses:

- (Z1) z* é parcialmente físico, e todo o seu componente físico é fechado sob causação: todo efeito físico em z* tem uma causa física.
- (Z2) Os organismos semelhantes a humanos em z* estão relacionados a um tipo especial de itens x. x faz com que eles sejam fenomenalmente

⁸⁵ Cf. KIRK, 2005, p. 50.

⁸⁶ Cf. KIRK, 2005, p. 50-1.

⁸⁷ "Zombists therefore cannot object that it is inconceivable that x should provide for epistemic intimacy as well as consciousness. So the earlier conclusion stands. Even those zombists who hold that there is non-physical interaction in the actual world must concede that conceivably, a non-physical item could be associated with a zombie world so as to transform it into a world z* where the ex-zombies enjoyed phenomenal consciousness and epistemic intimacy with their experiences."

conscientes.

- (Z3) x é causado por processos físicos, mas não tem efeitos físicos: pode ser removido sem perturbar o componente físico de z^* .
- (Z4) Os habitantes semelhantes a humanos de z^* consistem em nada além de corpos funcionais e seus x relacionados.
- (Z5) Os habitantes semelhantes aos humanos de z^* são capazes de perceber, prestar atenção, pensar e comparar as qualidades de suas experiências.

Como foi previamente mencionado, (Z1)-(Z5) refletem as teses da história e-qualia (E1)-(E5). O termo " z^* " substitui "o mundo", "organismos semelhantes a humanos" substitui "seres humanos" e "x" substitui as "e-qualia". Segundo Kirk, as teses (E1)-(E5) descrevem o mundo atual e as características dos seres humanos, enquanto (Z1)-(Z5) descrevem como seria o mundo particular z^* e os seres ex-zumbis semelhantes aos humanos, mas também como o mundo real e seus habitantes poderiam ser ou ter sido, com base na ideia de que os zumbis são concebíveis. Por essa razão, não há diferença entre as duas primeiras substituições. Quanto à terceira, também não há diferença, pois a estrutura de x não importaria, uma vez que, de acordo com (Z2) e (Z5), os habitantes de z^* ainda devem ter experiências com qualidades fenomenais, permitindo, assim, substituir e-qualia por x sem alterações significativas⁸⁸. A concebibilidade do mundo zumbi implica a concebibilidade da história e-qualia, já que as teses do mundo zumbi refletem as teses da história e-qualia. De acordo com Kirk, isso nos permite concluir, como no item (II), que, se os zumbis fossem concebíveis, a história e-qualia também seria concebível. Tendo demonstrado as premissas (I) e (II), por *modus tollens*, podemos concluir que: (III) zumbis são inconcebíveis.

O primeiro passo do argumento visa mostrar, por meio do exemplo das imagens nas solas dos pés e das correntes elétricas no cérebro, que a concepção de e-qualia proposta pela história e-qualia não permite que os seres humanos possuam intimidade epistêmica com suas experiências de qualidades fenomenais. Como vimos, esse primeiro passo permitiu a Kirk chegar à premissa (I), ou seja, que a história e-qualia é inconcebível. A história e-qualia é incoerente, pois não consegue explicar como seus habitantes podem ser epistemicamente íntimos de suas experiências com qualidades fenomenais, partindo do pressuposto de que as e-qualia são causalmente inertes. Sendo rejeitada *a priori*, concluímos que a história e-qualia é inconcebível.

⁸⁸ Cf. KIRK, 2005, p. 51-2.

Como vimos anteriormente, o segundo passo do argumento consistiu em demonstrar que, dado um mundo zumbi z , seria possível transformá-lo em um mundo z^* ao adicionar um item não físico x , o qual produz intimidade epistêmica e consciência fenomenal nos seres humanos (e tal item é análogo ao item não físico que confere consciência fenomenal aos habitantes do mundo atual). Ao adicionar x a z , obteríamos z^* , um mundo no qual os habitantes ex-zumbis passariam a ter consciência fenomenal e intimidade epistêmica em relação às suas qualidades fenomenais. Para sustentar este segundo passo, foi necessário demonstrar que os defensores dos zumbis não poderiam negar que o item não físico x , quando adicionado a z , produziria consciência fenomenal e intimidade epistêmica nos habitantes desse mundo específico. Uma vez que se demonstrou que o defensor do zumbi não pode escapar dessa conclusão, foi possível sustentar as teses (Z1)-(Z5) e mostrar que elas refletem as teses da história e-qualia, de modo que, a partir disso, se pode sustentar a premissa (II): se os zumbis são concebíveis, a história e-qualia também é concebível. Com a demonstração de (I) e (II), utilizando *modus tollens*, chegamos à conclusão (III), a saber, que os zumbis são inconcebíveis.

3.3 Críticas à história e-qualia

Como acabamos de ver, um passo importante para a história e-qualia é provar que esse “epifenomenalismo específico” o qual Kirk quer atacar é inconcebível, incoerente. O passo essencial da primeira etapa é mostrar que não podemos ter intimidade epistêmica, isto é, não podemos conhecer, pensar e comparar – entre outras coisas – nossas e-qualia à medida que tais propriedades são causalmente inertes. Não podemos, inclusive, ter memória sobre nossas e-qualia: uma vez que a experiência qualitativa de dor é causalmente inerte, não é possível que eu lembre dessa experiência qualitativa.

Kirk adota – apesar de ter criticado –, involuntariamente ou não, uma intuição para prosseguir com o passo crucial da primeira etapa do seu argumento, a saber, a intuição de que é necessário que eventos mentais sejam causalmente eficazes – causando comportamentos – para que possamos estar justificados sobre a crença nesses eventos mentais e, além disso, para que possamos estar justificados ao relatar tais eventos mentais. Isto é, para que eu conheça e possa relatar a minha dor pontiaguda, é necessário que o evento mental qualitativo da dor seja capaz de causar comportamentos, e.g., o meu relato afirmando “Estou com uma dor terrível!”, e o mesmo serve para outros estados mentais.

De fato, parece certo dizer que os nossos estados mentais, tal como o caso da dor, influenciam o nosso comportamento. Podemos ter a plena certeza

de que a conexão entre estados mentais e comportamentos é a base da nossa experiência. Apesar do seu forte apelo intuitivo, epifenomenalistas sustentam que tal afirmação não tem mais força do que sustentar que estados mentais são causalmente ineficazes⁸⁹. Eles argumentam que, dada a lacuna explicativa⁹⁰, o epifenomenalismo tem o caráter abdutivo e está em posição de ser a melhor explicação para a consciência.

A objeção construída por Robert Kirk compartilha da mesma intuição que outra objeção forte desenvolvida contra o epifenomenalismo: o argumento *self-stultification*. Vejamos tal objeção.

3.3.1 Argumento *self-stultification*

O ponto que diverge o epifenomenalismo de outras teorias é, certamente, o ponto de maior ataque. Vamos por passos. Inicialmente, segundo [Robinson \(2004\)](#), o epifenomenalismo entende que nossas ações são efeitos de interações complexas de neurônios e células do nosso cérebro sendo, portanto, efeitos de causas inteiramente físicas. Do mesmo modo, sustenta que nossas qualia são distintas de – e irreduzíveis a – estados físicos. O apelo à irreduzibilidade de propriedades qualitativas é defendido por meio da lacuna explicativa:

Não temos nenhuma ideia promissora para uma estratégia reducionista que nos explique por que as ativações neurais de tais e tais tipos devem ser acompanhadas, por exemplo, de dores. ([ROBINSON, 2004](#), p. 45)⁹¹

Além de abraçar a completude física e a irreduzibilidade mental, o epifenomenalismo rejeita a possibilidade de ocorrer sobredeterminação causal de estados mentais e físicos, ou seja, um efeito físico não pode ter uma causa física suficiente e uma causa mental. Tudo isso combinado produz a rejeição da causalidade mental. Portanto, a principal tese epifenomenalista é afirmar que os estados e propriedades mentais não tem qualquer efeito causal sobre propriedades físicas e estados neurais⁹².

No epifenomenalismo, o evento qualitativo de dor não causa o relato de que eu estava com dor. Em vez disso, o relato de que eu estava com dor tem

⁸⁹ Cf. [ROBINSON, 2004](#), p. 160.

⁹⁰ Em suma, a lacuna explicativa é a dificuldade que o fisicalismo enfrenta em explicar como as propriedades físicas dão origem à maneira como as coisas nos parecem subjetivamente quando experimentamos. Ver mais em [Levine \(2014\)](#)

⁹¹ “We do not have any promising ideas for a reductive strategy that will explain to us why neural activations of such and such kinds must be accompanied by, e.g., pains”

⁹² Cf. [ROBINSON, 2004](#), p. 159.

alguma causa física, podemos chamar de P_1 . Podemos, portanto, deixar de fora a ocorrência do evento de dor, afinal, esse evento mental está fora do fluxo causal que produz o relato de que eu estava com dor. Em outras palavras, o relato de que eu estava com dor é cego à ocorrência ou falta de ocorrência de um evento mental. Ou, uma vez que o relato de que eu estava com dor surge inteiramente devido a alguma fonte física distinta da dor, o evento de dor não tem nada a ver com o relato de que eu estava com dor, portanto, o evento de dor é irrelevante para a ocorrência do relato de dor.

No entanto, os críticos sustentam que, se a tese epifenomenalista é verdadeira, não podemos ter conhecimento sobre os nossos próprios estados mentais ou não há nenhum valor no relato sobre estados mentais. O epifenomenalista fica preso na posição autodestrutiva de relatar que temos estados qualitativos, mas ao mesmo tempo não tem justificativa para o relato de que temos estados qualitativos. Nesse sentido, a tese epifenomenalista implode.

Assim, suponha que S seja um epifenomenalista e que S diga “Estou com uma dor terrível”. S está comprometido com a visão de que a dor não causa o enunciado. Mas então, ao que parece, S estaria fazendo o mesmo enunciado quer a dor estivesse ocorrendo ou não. Se for assim, então os testemunhos de S sobre suas próprias dores não têm valor – tanto para nós quanto para S. Eles não podem ser considerados como representação de qualquer conhecimento sobre dores por parte de S (se a visão epifenomenalista de S for verdadeira). (ROBINSON,)⁹³

O argumento é apresentado como destrutivo para o epifenomenalismo, e considera como condição necessária para o relato qualitativo de que eu estava com dor ser contado como conhecimento, que ele tenha sido causado pelo próprio evento mental de dor. Como anteriormente citado, a objeção *self-stultification* compartilha da mesma intuição que a história e-qualia de Kirk: é necessário para o conhecimento que estados mentais sejam causalmente eficazes para que possamos estar justificados sobre a nossa crença neles. O próprio Kirk assume que tal afirmação faz parte da história e-qualia:

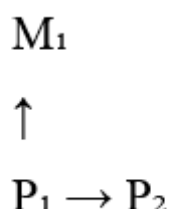
A objeção relacionada [a história e-qualia] é que não poderíamos pensar, falar ou saber sobre as qualidades de nossas experiências se elas não tivessem efeitos sobre os eventos físicos envolvidos nessas

⁹³ “Thus, suppose S is an epiphenomenalist, and that S utters “I am in terrible pain”. S is committed to the view that the pain does not cause the utterance. But then, it seems, S would be making the same utterance whether or not a pain were occurring. If this is so, then S’s testimonies about S’s own pains are worthless — both to us and to S. They cannot be taken to represent any knowledge about pains on S’s part (if S’s epiphenomenalist view is true).”

3.3.2 Resposta a *self-stultification*

Como supracitado, o argumento *self-stultification* depende, em sua totalidade, da afirmação de que é uma condição necessária para o conhecimento que o relato deva ser causado pelo evento mental qualitativo, e uma vez que a consciência fenomenal não causa nada, ela não pode causar a nossa crença de que a temos. Robinson (2004, p. 160-167) desenvolve uma resposta à objeção tratada até o momento, considerando da seguinte forma: é possível substituir a premissa alegada pela objeção *self-stultification*, sem perda de significado, por outra premissa que afirma ser uma condição necessária para o conhecimento que o relato deva ser causado pelo evento mental qualitativo relatado. Isso é possível pois se supõe que o relato sobre o evento mental qualitativo e o próprio evento mental qualitativo têm a mesma causa subjacente. Em suma, a substituição pela nova premissa significa que é necessário que o meu relato de dor seja causado pelo evento mental qualitativo de dor que eu relato. Só assim é possível conhecer meus estados mentais.

Agora, suponha que P_1 seja uma causa física, M_1 o evento mental qualitativo de dor e P_2 o relato da dor. Robinson considera o seguinte esquema epifenomenalista:



O esquema epifenomenalista nos mostra que não há nenhuma causa de M_1 além de P_1 e nenhuma outra causa de P_2 além de P_1 . Então nunca haverá ocorrência de M_1 que não seja seguido por P_2 , e nunca haverá um caso de P_2 que não seja precedido por M_1 . Assim, o contrafactual “Se M_1 não tivesse ocorrido, então P_2 não teria ocorrido”, será verdadeiro. Entretanto, apesar do contrafactual ser verdadeiro, M_1 e P_2 não estarão relacionados como causa e efeito, mas sim como efeitos comuns da mesma causa, P_1 ⁹⁵. Supondo que M_1 e P_2 não te-

⁹⁴ “The related objection is that we could not think, talk, or know about the qualities of our experiences if they had no effects on the physical events involved in those activities themselves.”

⁹⁵ Cf. ROBINSON, 2004, p. 160.

tenham outra causa além de P_1 , e P_1 necessita tanto M_1 e P_2 , então sempre que o relato de dor P_2 ocorre, a causa física P_1 ocorre dando origem à experiência de dor M_1 . Como o relato de dor só ocorre quando sua causa física também causa o evento de dor qualitativa, esse relato é justificado e pode ser contado como conhecimento. Por conseguinte, de acordo com Robinson, o epifenomenalismo é capaz de produzir os relatos justificados e verdadeiros sobre estados qualitativos.

Os defensores da objeção *self-stultification* considerariam o seguinte esquema para que o relato do evento mental seja justificado:

$$P_1 \rightarrow M_1 \rightarrow P_2$$

Neste caso, o relato P_2 do estado mental é causado diretamente por M_1 , o estado mental. Como Robinson nota, supostamente existe uma condição para considerar afirmações de conhecimento, a saber, quando a cadeia de eventos que resulta na afirmação é do tipo que não ocorreria caso a afirmação não fosse verdadeira⁹⁶. Ou seja, o relato P_2 , por exemplo, “Estou com dor!” não ocorreria se não fosse verdadeiro. Entretanto, essa condição é satisfeita por ambos os esquemas apresentados. Desse modo, o esquema epifenomenalista tem tanta pretensão de fornecer a explicação para o conhecimento de M_1 quanto o esquema defendido pelos críticos.

Portando, para o epifenomenalista, a resposta à objeção é sustentar que há uma causa subjacente P_1 que causa tanto M_1 quanto P_2 . Neste caso, o contrafactual “Se P_2 não tivesse ocorrido, M_1 não teria ocorrido” é verdadeiro. Quando o sujeito relata seu estado mental, o fato do relato e o estado mental serem ambos causados por P_1 permite que o sujeito conheça seus estados mentais, ainda que tais estados sejam causalmente ineficazes.

Uma objeção à estrutura apresentada por epifenomenalistas é dada por Moore (2012), que sustenta que o relato P_2 não pode ser considerado evidência para a ocorrência do estado mental M_1 , ainda que se tenha a causa comum P_1 . Suponha que um tronco de árvore carrega 2 grãos de areia em um rio e, em determinado momento do rio, o tronco se choca com uma pedra. Os dois grãos se desprendem do tronco e caem em uma bifurcação no rio: um grão segue à esquerda e o outro grão segue à direita. O grão à esquerda do rio segue poucos metros e para na margem do rio; o grão à direita segue rio abaixo 1km até

⁹⁶ Cf. ROBINSON, 2004, p. 166.

se chocar com outro tronco de árvore⁹⁷. De acordo com Moore, o grão que segue à esquerda do rio é independente do grão que segue à direita, ou seja, não há nenhuma relação de dependência entre esses dois grãos, não há nenhuma informação carregada pelo grão à esquerda que seja evidência para o destino do grão à direita e vice-versa.

No entanto, os dois grãos possuem a mesma causa comum, a saber, a colisão do tronco com a rocha no rio acima e esse fato garante que o contrafactual “Se o grão de areia à esquerda não ficasse preso na margem, o grão à direita não teria ficado em um tronco” é verdadeiro. Mas a mesma causa comum não garante que exista dependência de um grão com o outro. Do mesmo modo, em analogia, Moore conclui que o fato de existir uma causa comum P_1 não garante que exista dependência entre M_1 e P_2 , ou que M_1 seja relevante para P_2 e vice-versa.

Embora seja verdade que o primeiro grão de areia caia no banco antes que o segundo grão de areia caia em um tronco, o primeiro fato tem pouco a ver com o segundo. Da mesma forma, a ocorrência de dor, mesmo que ocorra antes de um relato de dor, ainda é irrelevante para, ou não está intimamente relacionada a, ou é independente de, ou não desempenha nenhum papel na determinação da ocorrência do relato de dor. Cf. [MOORE, 2012](#), p. 633⁹⁸

O problema surge quando não há informações sobre o grão à esquerda quando investigamos o grão à direita, por mais intenso que seja o exame. Do mesmo modo, não há informações sobre M_1 quando investigamos o relato P_2 . A resposta de [Robinson \(2013\)](#) a essa objeção apela para a forma como o mundo é estruturado. Segundo o autor, quando assumimos que o mundo do exemplo dos grãos tem uma certa estrutura e característica, como ter uma rocha que colide com os troncos, ter uma bifurcação etc., as informações sobre o grão à esquerda pela presença do grão à direita são informações que dependem da estrutura do mundo, e não são informações que dependem somente do grão à direita.

Similarmente, o transporte de informações de M_1 pelo relato P_2 depende de como o mundo é assumido pela tese epifenomenalista e, neste caso, o transporte de informações não depende somente de P_2 . Ao investigar somente a existência do relato P_2 , as informações que encontraremos de M_1 são consequências da estrutura do mundo – e isso, é claro, envolve a causa comum P_1 –, e não são

⁹⁷ Cf. [MOORE, 2012](#), p. 632-3.

⁹⁸ “While it is true that the first grain of sand does land on the bank before the second grain of sand lands on a log, the first fact has little to do with the second fact. In the same way, the occurrence of pain, even though it occurs prior to a report on pain, is still irrelevant to, or not closely related to, or independent of, or plays no role in determining, the occurrence of the report on pain.”

consequências da própria existência de P_2 ⁹⁹.

Podemos utilizar um exemplo para mostrar como dois efeitos são independentes, ainda que tenham a mesma causa. Imagine que em uma fábrica de eletrônicos houve uma falha no sistema de resfriamento das máquinas. Essa falha gerou dois efeitos: o superaquecimento dos equipamentos e a ativação de um alarme de emergência. Suponha que alguém perceba que o alarme de emergência foi acionado. Mesmo sem verificar diretamente as máquinas, essa pessoa pode inferir que os equipamentos estão superaquecendo, pois ambos os efeitos têm a mesma causa, a saber, a falha no sistema de resfriamento. Da mesma forma, ao saber que os equipamentos estão superaquecendo, é possível imaginar que o alarme também foi acionado.

Assim, os dois efeitos podem ser conhecidos um pelo outro, mas apenas por meio da causa comum (a falha no sistema de resfriamento), apesar de ocorrerem de forma independente. Para epifenomenalistas como Robinson, uma pessoa que investigasse a ativação do alarme de emergência obteria informações sobre o superaquecimento dos equipamentos em função da estrutura das máquinas da fábrica. Do mesmo modo, se uma pessoa investigasse o superaquecimento dos equipamentos obteria informações sobre a ativação do alarme de emergência em função da estrutura das máquinas da fábrica.

Retomando o exemplo dos grãos, Robinson salienta um ponto de desacordo:

Pode parecer interessante ressaltar aqui que sempre podemos subir o rio até a bifurcação, observar atentamente as colisões entre troncos e pedras e, assim, verificar por observação que os grãos de areia no córrego da direita são produzidos por, e somente por, colisões de troncos que também desalojam um segundo grão (esquerdo); ao passo que não temos nenhuma maneira paralela de justificar que os eventos neurais que contribuem causalmente para nossos relatos (ou outro comportamento) também causam eventos qualitativos. (ROBINSON, 2013, p. 188-9).¹⁰⁰

Ou seja, no caso do rio, há um método de investigação que pode revelar a ocorrência das colisões de troncos com pedras do rio, o que ocasiona o desprendimento dos grãos. Por outro lado, no caso de estados mentais, não há nenhum método que possa revelar como os estados e eventos neurais causam

⁹⁹ Cf. ROBINSON, 2013, p. 188.

¹⁰⁰ "It may seem appealing here to point out that we can always go upstream to the fork, closely observe twig-rock collisions and thus verify by observation that sand grains in the right hand stream are produced by, and only by, twig collisions that also dislodge a second (left) grain; whereas we have no parallel way of justifying that the neural events that causally contribute to our reports (or other behavior) also cause qualitative events."

– ao mesmo tempo – relatos de comportamento e estados mentais qualitativos. E isso vale não somente para a explicação epifenomenalista: uma estrutura de explicação da ocorrência de estados mentais por interacionistas ($P_1 \rightarrow M_1 \rightarrow P_2$) tem a mesma dificuldade, pois não há como observar M_1 e a sua correlação com P_2 . Do mesmo modo, fisicalistas não podem observar de que modo a ocorrência de estados mentais é reduzida à ocorrência de estados físicos.

Em tese, Kirk já antecipa a resposta ao epifenomenalista quando afirma que processos físicos não são capazes de nos colocar em intimidade epistêmica com estados mentais. Para isso, Kirk apela para o exemplo das correntes elétricas¹⁰¹. Relembrando o exemplo: em resumo, suponha que o nosso processo perceptivo induza – sem que seja afetado – por correntes elétricas em nosso cérebro. Então, toda vez que olhamos para alguma coisa, temos correntes elétricas sendo geradas em uma parte específica do nosso cérebro. A questão que Kirk põem é: nosso processamento cognitivo seria capaz de nos colocar em intimidade epistêmica com essas correntes elétricas? Poderíamos conhecê-las? Kirk responde que não. Porém, é possível observar outro problema em relação ao exemplo de correntes elétricas e a noção de intimidade epistêmica desenvolvida por Kirk, que apresenta fragilidades, a qual veremos na próxima seção.

3.3.3 Intimidade epistêmica

A noção de intimidade epistêmica desenvolvida por Kirk é ponto chave no desenvolvimento de seu argumento. Relembrando, a intimidade epistêmica para Kirk é uma relação que nós temos com as qualidades de nossas experiências conscientes, e essa relação pode ser exemplificada por atividades tais como perceber as qualidades de nossas experiências conscientes, compará-las, pensar sobre elas etc. De fato, há uma intuição forte que nos leva a crer que estamos em uma relação de intimidade epistêmica com nossas *qualia*: quando, por exemplo, experimentamos uma bebida doce, podemos nos atentar a tal experiência, compará-la etc.

A noção que Kirk adota de intimidade epistêmica é muito próxima – se não a mesma – de uma noção epistêmica de conhecimento conhecida como *acquaintance*. De modo mais abrangente, a noção epistêmica de *acquaintance* pode ser descrita como o tipo de conhecimento

que relaciona uma pessoa com seus próprios estados e processos fenomenalmente conscientes diretamente, incorrigivelmente, e de uma forma que parece revelar sua essência. Quando alguém está ciente

¹⁰¹ E o exemplo mirabolante de Zob, seu gêmeo zumbi.

de um estado fenomenal no processo de tê-lo, algo essencial sobre ele [estado mental] é revelado, direta e incorrigivelmente – ou seja, *como é tê-lo*. (BALOG, 2012, p. 16, grifo nosso).¹⁰²

Ou seja, de modo mais basilar, *acquaintance* é o tipo de conhecimento que alguém tem com seus próprios estados fenomenais. O conceito de *acquaintance* foi introduzido na filosofia contemporânea por Bertrand Russell em “On Denoting” (1905) e aprofundado em seu artigo intitulado “Knowledge by Acquaintance and Knowledge by Description” (1910). Russell explica que uma pessoa está em uma relação de *acquaintance* com um objeto quando está em uma “relação cognitiva direta com o objeto, ou seja, quando [o sujeito está] diretamente ciente do próprio objeto” (RUSSELL, 1910, p. 108, grifo nosso)¹⁰³. Como nota DePoe (2013), a noção de *acquaintance* é frequentemente confundida com estar diretamente familiarizado com algo.

É importante notar que há uma diferença entre estar diretamente familiarizado com algo e ter conhecimento por *acquaintance* de que algo é o caso. Para um sujeito estar diretamente familiarizado com algo, apenas é necessário que o sujeito tenha acesso não mediado ao objeto de consciência. O conhecimento por *acquaintance* de que algo é o caso, no entanto, é mais do que estar diretamente familiarizado com algo sendo o caso. O conhecimento por *acquaintance*, afinal, é um tipo de conhecimento, que requer que o sujeito tenha uma crença sob as condições certas. Para um sujeito estar diretamente familiarizado com algo, não é necessariamente necessário que o sujeito tenha uma crença sobre isso. (DEPOE, 2013).¹⁰⁴

Ou seja, a diferença principal sobre estar familiarizado com algo e o conhecimento por *acquaintance* é que, neste último caso, é necessário que o sujeito tenha crença sobre seu conhecimento. Por exemplo, considere as seguintes afirmações:

- S está diretamente familiarizado com *p*

¹⁰² “that relates a person to her own phenomenally conscious states and processes directly, incorrigibly, and in a way that seems to reveal their essence. When one is aware of a phenomenal state in the process of having it, something essential about it is revealed, directly and incorrigibly – namely, *what it is like to have it*.”

¹⁰³ “direct cognitive relation to that object, i.e. when directly aware of the object itself.”

¹⁰⁴ “It is important to notice that there is a difference between being directly acquainted with something and having knowledge by acquaintance that something is the case. For a subject to be directly acquainted with something only requires for the subject to have unmediated access to the object of awareness. Knowledge by acquaintance that something is the case, however, is more than being directly acquainted with something’s being the case. Knowledge by acquaintance, after all, is a kind of knowledge, which requires the subject to hold a belief under the right conditions. For a subject to be directly acquainted with something does not necessarily require the subject to hold a belief about it.”

e

- S sabe por *acquaintance* (conhecimento direto) que *p*.

Suponha que *p* é o fato de que alguém está sentindo uma dor, tudo o que é preciso para que (1) seja verdadeira é que o sujeito tenha algum acesso ou consciência não mediada de sua experiência de dor. Em (1), algumas pessoas podem estar diretamente familiarizadas com sua experiência de dor, mas não ter conhecimento de que estão tendo uma experiência de dor, talvez por não prestar atenção a essa experiência de tal forma que formem uma crença proposicional com base em seu conhecimento direto da experiência de dor. No entanto, para que (2) seja verdadeira, é necessário mais. É necessário que o sujeito forme crenças com base na experiência de dor que está sentindo.

Do mesmo modo, Kirk sustenta que a intimidade epistêmica é exemplificada por processos cognitivos como perceber, atender, pensar e comparar as qualidades de nossas experiências¹⁰⁵. Ora, para que tais processos sejam realizados por um sujeito, é necessário que o mesmo forme crenças com base em suas experiências, por exemplo, suponha que um sujeito tenha a experiência de dor em dois momentos, *t'* e *t''*. Para que o mesmo compare, por exemplo, se a dor em *t'* foi mais forte ou leve que a dor em *t''*, é preciso formar crenças sobre essas experiências. Assim, as noções de intimidade epistêmica e *acquaintance* parecem ser similares quando se afirma que ambas noções epistêmicas são relações do sujeito com as qualidades de suas experiências conscientes.

O próprio Kirk (2005, p. 42) admite ter retirado a noção de intimidade epistêmica daquilo que Chalmers (1996) chama *acquaintance*, uma “relação epistêmica íntima” com nossas experiências fenomenais¹⁰⁶. Segundo Chalmers,

[...] há algo intrinsecamente epistêmico na experiência. Ter uma experiência é automaticamente estar em algum tipo de relação epistêmica íntima com a experiência - uma relação que poderíamos chamar de “acquaintance”. Não há sequer uma possibilidade conceitual de que um sujeito possa ter uma experiência vermelha como essa sem ter nenhum contato epistêmico com ela: **ter a experiência é estar relacionado a ela dessa maneira.** (CHALMERS, 1996, p. 196-7, grifo nosso).¹⁰⁷

Assim, o conhecimento que possuímos de nossas experiências é direto

¹⁰⁵ Cf. KIRK, 2005, p. 43.

¹⁰⁶ Cf. CHALMERS, 1996, p. 196-7.

¹⁰⁷ “there is something intrinsically epistemic about experience. To have an experience is automatically to stand in some sort of intimate epistemic relation to the experience—a relation that we might call “acquaintance.” There is not even a conceptual possibility that a subject could have a red experience like this one without having any epistemic contact with it: to have the experience is to be related to it in this way.”

e imediato, no sentido de que a experiência qualitativa é, em si mesma, o conhecimento dessa experiência. Não parece haver diferenças substanciais entre a noção de intimidade epistêmica e a noção de *acquaintance*. Porém, ainda que as duas noções signifiquem a mesma coisa, Kirk não caracteriza de maneira adequada. Neste ponto, Kirk parece cometer um equívoco ao separar o objeto e o conhecimento que temos sobre o objeto no exemplo das correntes elétricas no cérebro, o que é incompatível com a noção epistêmica de *acquaintance* e, de maneira análoga, com a noção de intimidade epistêmica. Ao distinguir entre o objeto causado (as correntes elétricas) e o conhecimento das experiências (intimidade epistêmica), o exemplo perde a caracterização epistemológica central relativa ao conhecimento de nossas qualidades fenomenais, uma vez que tal conhecimento é, por sua natureza, direto e imediato.

Seria redundante dizer que a natureza das correntes elétricas, por um lado, e a dos estados (eventos e propriedades) mentais, por outro, são totalmente distintas na visão epifenomenalista, no sentido de que os estados mentais são conhecíveis diretamente, não mediado. No entanto, é de extrema importância compreender que essa diferença de natureza também torna o exemplo fragilizado: no caso das correntes elétricas, há uma separação entre o objeto e uma propriedade que relacionaria epistemicamente o objeto ao sujeito; por outro lado, no caso dos estados mentais, como a dor, o próprio estado mental relaciona o sujeito ao estado mental em seu fluxo de ocorrência. A intimidade epistêmica deve residir – ou ter origem – na própria ocorrência do estado mental, tal como ocorre quando se caracteriza o conhecimento *acquaintance*.

Em outras palavras, pela forma como o exemplo das correntes elétricas é formulado parece haver uma lacuna entre a ocorrência das correntes elétricas e a possibilidade de ter intimidade epistêmica com elas. No entanto, o conhecimento que temos de nossas qualidades fenomenais não permite que haja lacuna entre a ocorrência de um estado mental e o nosso conhecimento das qualidades fenomenais desse estado. Portanto, a conclusão é que o exemplo não consegue retratar o que é o conhecimento que temos das qualidades fenomenais de nossas experiências.

Da mesma forma, o exemplo do gêmeo zumbi Zob perde o ponto central do conhecimento das qualidades fenomenais ao pressupor que a intimidade epistêmica não poderia ocorrer quando as leis que originam as e-quality são adicionadas ao mundo desse gêmeo zumbi. Tal suposição implica que a intimidade epistêmica seria um tipo de conhecimento independente da ocorrência das e-quality, o que é inaceitável, dado que o conhecimento de nossas experiências está intrinsecamente ligado ao fluxo da experiência mental. Assim, sempre que um sujeito tem uma experiência, ele necessariamente conhece essa experiência.

Como afirma Chalmers (2002), a característica central do conhecimento por *acquaintance* é que “sempre que um sujeito tem uma propriedade fenomenal, ele está familiarizado com essa propriedade fenomenal.” (CHALMERS, 2002, p. 250)¹⁰⁸.

Kirk poderia argumentar que a intimidade epistêmica por ele desenvolvida apresenta diferenças relevantes em relação à noção epistêmica de *acquaintance*. Contudo, o argumento apresentado nesta seção visa sustentar que não existe distinção substancial entre a intimidade epistêmica e o conhecimento por *acquaintance*. Além disso, Kirk deveria ter explicitado a diferença entre a intimidade epistêmica e a noção epistêmica de *acquaintance*, conforme descrita por Chalmers, a qual, conforme o próprio Kirk afirma, o influenciou.

O problema com a noção de intimidade epistêmica, apontado até aqui, surge quando Kirk tenta separar a ocorrência de uma experiência e o conhecimento que temos das nossas experiências qualitativas. Como argumentamos, a intimidade epistêmica é semelhante à noção epistêmica de *acquaintance*, e esta última não permite separar a ocorrência de uma experiência do conhecimento que temos das qualidades dessa experiência.

No entanto, Kirk poderia alegar que não é possível definir o conhecimento por *acquaintance* e dizer que “quando temos uma propriedade fenomenal, estamos em relação de *acquaintance* com essa propriedade” não significa que definimos *acquaintance* e, por similaridade, intimidade epistêmica. De fato, essa é uma alegação justa, pois boa parte das propriedades e estados mentais não é explicável por definição – as qualia, por exemplo.

Entretanto, podemos apelar para o poder intuitivo que sustenta essa caracterização de *acquaintance* feita por Chalmers. Quando temos uma certa experiência fenomenal, podemos dizer que, no ato de sua ocorrência, estamos familiarizados com tal experiência. E Kirk não poderia rejeitar esse apelo intuitivo, visto que sua argumentação utiliza o forte apelo intuitivo de que há alguma ligação causal entre intimidade epistêmica e e-qualia — algo que nos permite falar, perceber e nos atentarmos às e-qualia. A questão a ser respondida é: por que Kirk utiliza esse apelo intuitivo e, em seguida, ao construir seu exemplo que derrotaria a concebibilidade de zumbis, apela para um cenário que separa as propriedades fenomenais da intimidade epistêmica? Portanto, nossa conclusão permanece a mesma: a noção de intimidade epistêmica de Kirk falha em sua caracterização do que é o conhecimento que possuímos sobre nossas experiências qualitativas.

¹⁰⁸ “that whenever a subject has a phenomenal property, the subject is acquainted with that phenomenal property.”

3.3.4 Concebibilidade no argumento de Kirk

Para um falante não-filósofo comum, “conceber” e “imaginar” podem muito bem ser usados de forma intercambiável; e frequentemente os filósofos também tratam esses termos aproximadamente como sinônimos. Na literatura sobre a epistemologia da modalidade, no entanto, a concebibilidade é uma noção técnica, que pode ser definida de diferentes maneiras.

Até o momento, observamos que a noção de concebibilidade é fundamental para a robustez do argumento do zumbi, pois, conforme discutido na seção 2.2.2, nem todas as concepções de concebibilidade são suficientemente fortes para permitir uma concepção adequada.

Em seu argumento anti-zumbi, Kirk sustenta que algo é concebível quando não é *a priori* falso¹⁰⁹. Em outras palavras, para Kirk, seja *S* uma proposição, *S* é concebível quando, por reflexões *a priori*, um sujeito não pode descartar *S* como falsa. Assim, para conceber uma proposição que descreve o mundo zumbi, basta que o sujeito, em suas reflexões, não encontre incoerências que revelem a situação concebida como falsa.

A noção de concebibilidade empregada por Kirk é, de fato, frequentemente utilizada por outros filósofos. No entanto, ao adotarmos o relato de concebibilidade desenvolvido por Chalmers como fonte de conhecimento modal, perceberemos que a concepção de concebibilidade adotada por Kirk se revela fraca para alcançar seu objetivo. Inicialmente, vamos retornar a bi-dimensionalidade proposta por Chalmers.

Chalmers divide, primeiramente, em duas opções de concebibilidade: *prima facie* e ideal. Esta distinção está relacionada aos limites cognitivos de quem concebe. A concebibilidade *prima facie* refere-se à avaliação inicial de uma pessoa sobre a possibilidade de uma hipótese descrita, sem a necessidade de um exame aprofundado sobre ela. Pelo que vimos no capítulo 2, Chalmers sugere que um raciocínio mais cuidadoso geralmente oferece uma razão para questionar a concebibilidade *prima facie* em alguns casos.

A concebibilidade ideal, por outro lado, envolve um tipo de raciocínio que não pode ser refutado por considerações posteriores. Se uma proposição é idealmente concebível, isso significa que ela é verdadeira do ponto de vista de um raciocinador ideal hipotético que está “livre de todas as limitações cognitivas contingentes” (CHALMERS, 2002, p. 148)¹¹⁰. Isso é suficiente para garantir que o raciocínio ideal esteja protegido contra erros modais. Por essa razão, para Chalmers, um vínculo entre concebibilidade e possibilidade só pode ser estabelecido

¹⁰⁹ Cf. KIRK, 2005, p. 48.

¹¹⁰ “free of all contingent cognitive limitations.”

quando se trata da concebilidade ideal.

É pertinente retomar esta discussão para observar que, segundo Chalmers, a maioria dos casos em que se rejeita que a concebilidade implica possibilidade corresponde a casos de concebilidade *prima facie*. Como foi mencionado anteriormente, a conjectura de Goldbach constitui um exemplo paradigmático no qual tanto a sua veracidade quanto a sua falsidade são concebíveis, embora apenas um desses cenários seja verdadeiro. Como já foi destacado, este caso ilustra claramente a falha da concebilidade *prima facie*, uma vez que, para o concebedor ideal – desprovido de limitações cognitivas –, um dos cenários – seja a verdade ou a falsidade da conjectura – seria descartado após uma reflexão ideal.

O propósito de mencionar este ponto é que, para que o argumento do zumbi seja eficaz contra o fisicalismo, não basta que o sujeito conceba o mundo zumbi e não identifique incoerência em sua reflexão *a priori*; é necessário que essa concebilidade seja imune a erros e limitações cognitivas. Adotar a concebilidade ideal protege – em tese – aquele que concebe os pares de mundos zumbis de cometer algum erro. Tal afirmação é corroborada por Chalmers, que sustenta que a concebilidade ideal é o caminho mais seguro para a possibilidade.

Sendo assim, é possível afirmar que a noção de concebilidade apontada por Kirk é fraca, e a opção mais forte para se conceber mundos zumbis é a concebilidade ideal, não permitindo que opositores aleguem que há inconsistência na concebilidade desses mundos. Como afirma Chalmers, a concebilidade ideal traz robustez uma vez que “é necessária uma justificação que não possa ser racionalmente derrotada” (CHALMERS, 2002, p. 148)¹¹¹.

3.3.5 Premissa (A2)

Em certa altura de sua argumentação, Kirk desenvolve o que seria o argumento para se conceber mundos zumbis. Para relembrar, segundo Kirk, a concebilidade de mundos zumbis implica as seguintes premissas:

- (A1) z é um sistema puramente físico e causalmente fechado;
- (A2) Fisicamente, z é, na medida do possível, exatamente como o mundo real;
- (A3) Os habitantes semelhantes a humanos de z carecem de consciência fenomenal.

(A1) é a premissa que postula o fechamento causal do físico e, em termos gerais, constitui o foco das críticas associadas ao argumento do zumbi. Por sua vez,

¹¹¹ “one needs justification that cannot be rationally defeated.”

(A3) esclarece a definição de zumbis, caracterizando-os como cópias dos seres do mundo atual, mas destituídas de consciência fenomenal. Com relação a essas duas premissas, não parecem surgir maiores objeções. Contudo, (A2) requer uma análise mais cuidadosa, uma vez que visa apresentar o conceito de réplica zumbi do mundo atual.

Como discutido na seção 2.2.1, o fisicalismo sustenta que qualquer réplica física mínima do mundo atual será uma réplica *simpliciter* do nosso mundo. Isso implica que, ao replicarmos fisicamente o mundo atual em outro contexto, esse mundo conterá tudo o que existe no nosso, dado que a estrutura física do nosso mundo necessita de todas as propriedades – sejam estas físicas ou não-físicas.

Para um fisicalista, ao replicarmos fisicamente o mundo atual em outro, é crucial que não se adicione nenhum elemento a esse mundo que altere a estrutura física replicada. De fato, caso algo seja acrescentado a uma réplica física do mundo atual, um fisicalista poderia argumentar que ela deixa de ser uma réplica fiel do nosso mundo.

Entretanto, ao analisarmos a forma como (A2) é apresentada, observamos que o qualificador ‘na medida do possível’ enfraquece substancialmente a premissa, uma vez que possibilita que a réplica zumbi inclua componentes físicos que não pertencem ao mundo real, o que representa um problema para um fisicalista. Em termos mais claros, (A2) não captura o que é essencial para o fisicalista: a necessidade de preservação da estrutura física replicada.

Conforme será detalhado na seção 4.1.2, a fragilidade de (A2) pode gerar objeções e abrir espaço para que defensores do interacionismo refutem a argumentação de Kirk. De acordo com Sepetyi (2022, p. 89), a explicação para a caracterização de (A2) da forma como está deve-se ao fato de que Chalmers não considera o argumento do zumbi como um endosse do epifenomenalismo, mas o deixa aberto a outras possíveis teses, como o interacionismo e o panprotopsi-quismo. Dessa forma, Kirk ajustou (A2) para que qualquer defensor da teoria dos zumbis fosse forçado a admitir a concepção do mundo zumbi.

Essa explicação é coerente, considerando que o objetivo de Kirk é que a história e-qualia sirva como um argumento definitivo contra qualquer defensor da concepção de zumbis. No entanto, é relevante salientar que, ao mesmo tempo em que Kirk supostamente consegue pressionar dualistas de diversas vertentes a aceitarem a concepção de zumbis, ele acaba por rejeitar as duas condições fundamentais para caracterizar adequadamente uma réplica física do mundo atual, destacadas por Jackson (1998) e supracitadas na seção 2.2.1, a saber,

(a) é exatamente como o nosso mundo em todos os aspectos físicos (tanto em propriedades, leis, relações) e

(b) não contém mais nada do que é necessário para satisfazer (a).

Assim, podemos concluir que a premissa (A2) representa um desafio para Kirk, uma vez que pode ser utilizada para pressionar os dualistas não-epifenomenalistas a aceitarem a concepção de zumbis, enquanto simultaneamente viola as condições necessárias para caracterizar adequadamente uma réplica física do mundo atual.

Capítulo 4

Argumentos pró-zumbis

O objetivo deste capítulo é mostrar argumentos que apontam o equívoco cometido por Kirk em sua posição anti-zumbis, a saber, de que a concebilidade de mundos zumbis implica em uma tese epifenomenalista.

4.1 A concebilidade de zumbis implica a história e-qualia?

Como anteriormente mencionado, o argumento da história e-qualia ataca a concebilidade de zumbis. O movimento do argumento é resumido em: primeiro, mostrar que a história e-qualia é incoerente – e portanto, inconcebível –; posteriormente, sustentar que a concebilidade de zumbis implica na concebilidade da história e-qualia; por fim, por *modus tollens*, concluir que zumbis são inconcebíveis.

O argumento de Kirk depende inteiramente da afirmação de que a concebilidade de zumbis implica a coerência da história e-qualia. Nesta seção, apresentarei um argumento desenvolvido por [Sepetyi \(2022\)](#), no qual o autor busca mostrar que a afirmação anterior falha. Dividirei as críticas de Sepetyi em subseções para melhor compreensão.

4.1.1 Crítica à premissa (E3) da história e-qualia

O autor observa que existem duas abordagens em relação às qualia e outros processos cognitivos, como perceber, atentar-se, pensar e comparar, quando se discute o argumento do zumbi. De um lado, filósofos materialistas – como os funcionalistas, por exemplo – podem entender as qualia como qualidades subjetivas da experiência que resistem à redução materialista e, por isso, são problemáticas, enquanto os processos cognitivos podem ser explicados em termos funcionais. Assim, o que Sepetyi chama de fisicalismo cognitivo coloca as qualia na categoria fenomenal e os processos cognitivos na categoria física. Se os du-

alistas aceitarem essa posição, estarão em desvantagem ao ter que admitir que uma boa parte do reino mental – os processos mentais cognitivos – não são nada além de processos físicos funcionais.

Por outro lado, para Sepetyi, é plausível que os defensores do argumento do zumbi defendam que pensamentos, atenção, compreensão etc. sejam tão difíceis de reduzir à explicação reducionista quanto as qualia. Ao defender que zumbis são possíveis, os processos cognitivos estão na categoria fenomênica, e não na categoria física. Essa mudança de abordagem teria consequências na interpretação da história e-qualia, pois os processos cognitivos como perceber, atentar-se, pensar e comparar que fazem com que seres humanos sejam fenomenalmente conscientes seriam e-qualia na definição (E2) de Kirk¹¹².

Retomando a história e-qualia, Sepetyi argumenta que o raciocínio só é válido se uma parte da premissa (E3) — ‘E-qualia são causadas por processos físicos’ — for interpretada como:

(E3*) Todas as e-qualia são diretamente causadas exclusivamente por processos físicos.

A suposta incoerência da história e-qualia repousa na ausência de causalidade entre experiências qualitativas e processos cognitivos, como pensamento e atenção. No entanto, a formulação original de (E3) não exclui a possibilidade de experiências qualitativas não-físicas exercerem influência causal sobre o pensamento e a atenção, caso esses processos cognitivos sejam considerados não-físicos, o que constitui a posição – que apresentamos acima – que Sepetyi sugere como plausível de ser defendida.

Seria possível argumentar que Kirk, de fato, antecipa essa objeção e afirma que a história e-qualia não permite que e-qualia tenham efeitos sobre outras e-qualia. No entanto, Sepetyi nota que a argumentação feita por Kirk interpreta (E3) como sendo (E3*), pois, segundo Kirk,

Como por (E3) **todas** as e-qualia já são causadas por eventos físicos [...]. Assim, as e-qualia não têm efeitos nem no mundo físico nem entre si. Elas são completamente inertes [...]. A causação de e-qualia por outras e-qualia é descartada por (E1)–(E5) porque se diz que as e-qualia são causadas por processos físicos, não por outras e-qualia. (KIRK, 2005, p. 42, ênfase nossa)¹¹³

Note que, na formulação original de (E3), não há menção a “todas”, porém,

¹¹² Cf. SEPETYI, 2022, p. 86.

¹¹³ “Since by (E3) all e-qualia are already caused to occur by physical events [...]. Thus e-qualia have no effects either on the physical world or among themselves. They are completely inert [...]. The causation of e-qualia by other e-qualia is ruled out by (E1)–(E5) because e-qualia are said to be caused by physical processes, not by other e-qualia.”

na argumentação feita por Kirk é utilizada. As outras modificações em (E3*) são:

- (1) Se todas as e-qualia são causadas *apenas* por processos físicos, então os processos físicos são suficientes para causar e-qualia, de modo que não haveria nenhuma contribuição causal das e-qualia. A introdução do qualificador 'apenas' impede a possibilidade de as e-qualia serem causadas simultaneamente por processos físicos e outras e-qualia, de forma que, na ausência das e-qualia, os processos físicos não produziram o mesmo efeito.
- (2) Se todas as e-qualia forem diretamente causadas por processos físicos, então não há possibilidade de causação indireta de e-qualia, como seria o caso se um processo físico P causa uma e-quale A que, por sua vez, causa outra e-quale B. Ao enfatizar a causação direta, toda a causação de e-qualia é feita diretamente por processos físicos.

Podemos notar que essa é a linha de raciocínio argumentativo produzido por Kirk para defender (E3) na citação anterior, mas que requer alteração para (E3*) para ser mais explícita. Assim, a primeira crítica de Sepetyi é a falta de clareza de Kirk ao explicitar a premissa (E3) da história e-qualia, que é fundamental para sua argumentação.

4.1.2 Tipos de defensores de zumbis *versus* história e-qualia

Sepetyi classifica os defensores dos zumbis em três grupos, conforme suas posições, que podem confrontar o argumento de Kirk: o defensor do zumbi que sustenta que o pensamento pertence a uma mente não-física pode ser denominado defensor cartesiano dos zumbis. Por outro lado, o defensor fisicalista cognitivo dos zumbis (defensor não-cartesiano dos zumbis) argumenta que os processos cognitivos pertencem ao domínio físico. A maneira como este último aborda a questão da história e-qualia e o *status* causal de estados mentais dependerá de sua perspectiva, seja ela epifenomenalista ou interacionista¹¹⁴.

O defensor epifenomenalista não-cartesiano dos zumbis compreende o mundo real como descrito na história e-qualia e, portanto, para rejeitar o argumento de Kirk deverá mostrar que não há incoerência na história e-qualia¹¹⁵. Desse modo, o foco de Sepetyi será discutir a resposta contra a história e-qualia

¹¹⁴ Cf. SEPETYI, 2022, p. 88.

¹¹⁵ Uma opção são os argumentos apresentados por Robinson, que discutimos nas seções 3.3.1 e 3.3.2

dos defensores cartesianos e os defensores interacionistas não-cartesianos dos zumbis.

À primeira vista, pode parecer que a concebilidade de mundos zumbis é inconsistente com o interacionismo. Se o interacionismo é verdadeiro no mundo real, a consciência não-física influencia causalmente os processos e estados cerebrais que causam o comportamento. Se o interacionismo é verdadeiro, a aparente inconsistência com a concebilidade de mundos zumbis se deve ao fato de que os zumbis carecem de consciência não-física relevante que afeta a estrutura física cerebral. Assim, o funcionamento dos corpos de zumbis seriam diferentes dos seres no mundo real.

No entanto, contra a aparente inconsistência apresentada acima, Sepetyi utiliza a defesa desenvolvida por [Chalmers \(2004\)](#). Chalmers sustenta que, para defender o argumento do zumbi, o interacionista pode argumentar que os mundos zumbis concebidos possuem lacunas causais que são preenchidas por estados mentais no mundo atual, ou que no mundo zumbi concebido essas lacunas são preenchidas por outras coisas diferentes de estados mentais¹¹⁶.

Tal suposição levantada por Chalmers pode parecer problemática no primeiro momento. Como explica Sepetyi,

Pode-se pensar que a ideia de um mundo no qual algumas causas fisicamente relevantes estão sistematicamente ausentes, mas todos os eventos físicos ocorrem como se nada estivesse ausente, é incoerente; tal mundo não é c-possível. Caso contrário, se em um mundo possível, lacunas causais físicas são preenchidas por algo diferente de processos mentais, o que pode ser esse “algo outro”? Parece que se não é mental e é causalmente relevante, não há razão para que não deva ser considerado físico. No entanto, se for considerado físico, então o mundo possível assim concebido não é exatamente fisicamente idêntico ao mundo real; ele tem algum excedente físico. ([SEPETYI, 2022](#), p. 90).¹¹⁷

Em outras palavras, a aparente incoerência poderia ocorrer devido ao fato de haver um excedente físico que preenche a lacuna causal do mundo físico. Dessa forma, o mundo zumbi não seria exatamente uma réplica física do mundo atual.

¹¹⁶ Cf. [CHALMERS, 2004](#), p. 184.

¹¹⁷ “One may think that the idea of a world in which some physically relevant causes are systematically lacking but all physical events go as if nothing were lacking is incoherent; such a world is not c-possible. Otherwise, if in a possible world, physical causal gaps are filled by something other than mental processes, what can this “something other” be? It seems that if it is not mental and is causally relevant, there is no reason why it should not count as physical. However, if it counts as physical, then the possible world so conceived is not exactly physically identical with the actual world; it has some physical surplus.” Sepetyi utiliza o termo “c-possível” com o mesmo significado que Kirk utiliza a c-concebilidade.

Entretanto, ao considerarmos a premissa que descreve o mundo zumbi da história e-qualia, a saber, (A2) 'Fisicamente, z é, tanto quanto possível, exatamente igual ao mundo real', observamos que a premissa não é explícita quanto à estrutura física replicada do mundo atual, uma vez que a condição 'tanto quanto possível' deixa em aberto a possibilidade de mundos zumbis com excedente físico. Esse tipo específico de interacionismo Sepetyi denomina de interacionismo substituíbilista (*replaceabilism*)¹¹⁸.

O interacionismo substituíbilista implica a possibilidade de zumbis fenomenais modificados ou de um mundo zumbi alterado, que não carece de nada físico que o mundo atual possui, mas que, ainda assim, carece de consciência. Nesses mundos zumbis modificados, deveriam existir outros fatores causais fisicamente relevantes para compensar a lacuna causal criada pela ausência da consciência. De acordo com a condição (A1), que afirma que um mundo zumbi z é totalmente físico, Kirk define que esses fatores, se existirem, devem ser físicos. Um mundo zumbi modificado pode ser fisicamente mais rico que o mundo real. Desse modo, o interacionista substituíbilista poderia conceber mundos zumbis coerentes.

Segundo Sepetyi, existe outras duas possibilidades para o interacionista responder ao argumento de Kirk, a qual ele chama de interacionismo insubstituíbilista (*irreplaceabilism*)¹¹⁹ e interacionismo supercoincidentalista. Um insubstituíbilista poderia afirmar que não é concebível mundos modificados nos quais entidades físicas substituem aspectos fenomênicos da mente humana, de maneira que o comportamento continua o mesmo de um ser humano consciente, algo como o substituíbilista poderia sugerir. Obviamente, esses não poderiam ser considerados defensores dos zumbis pois, para isso, é necessário crer que zumbis são concebíveis no sentido de não poder ser rejeitado *a priori*.

Ainda que isto precise ser admitido, um interacionista insubstituíbilista poderia utilizar o argumento do zumbi de forma parcial para criar um dilema para o fisicalismo: ou se deve aceitar que o mundo atual não é causalmente fechado e, portanto, o fisicalismo é falso; ou se deve aceitar que o mundo atual é causalmente fechado, mas que zumbis são possíveis e, portanto, o fisicalismo é falso¹²⁰.

Por outro lado, o supercoincidentalismo sustenta que mundos zumbis não modificados (sem entidades físicas substituindo a consciência) são possíveis;

¹¹⁸ Outra tradução possível para *replaceabilism* é substitubilismo.

¹¹⁹ A outra tradução possível para *irreplaceabilism* é insubstitubilismo.

¹²⁰ Neste ponto, surge a questão de como seria possível não ser um defensor de zumbis e, ainda assim, utilizar a ideia de que os zumbis são possíveis. Seria viável adotar a hipótese de que os zumbis são possíveis sem endossá-la? Parece que tal posição é pouco coerente ou inviável para um interacionista.

no entanto, essa possibilidade envolve uma série de coincidências improváveis. Caso a indeterminação causal física não seja incoerente¹²¹ e a consciência tenha efeitos no mundo atual, seria genuinamente possível a existência de um mundo zumbi como resultado de um evento extremamente improvável. Supondo que a indeterminação causal seja coerente, não seria possível determinar a origem de tal evento improvável.

Assim, há pelo menos três possibilidades para o interacionista confrontar o argumento de Kirk.

4.1.3 O erro crucial da história e-qualia

Na segunda etapa da argumentação, Kirk propõe um dilema no qual o defensor do zumbi se vê obrigado a admitir que o fator de consciência x – que é análogo ao fator de consciência y no mundo atual – quando associado ao mundo z , produz z^* , um mundo no qual os antigos seres zumbis possuem consciência fenomenal. Tal passo supostamente permite que Kirk afirme a semelhança entre (Z1)-(Z5) – teses que descrevem o mundo z^* – e as teses (E1)-(E5) que descrevem a história e-qualia e, por conseguinte, permite que afirme que a concebibilidade do mundo zumbi implica a concebibilidade da história e-qualia.

Segundo Sepetyi, há pelo menos dois questionamentos quanto a segunda etapa da argumentação de Kirk:

- Por que Kirk acredita que o defensor dos zumbis está comprometido com a transformação de z em z^* ?

e

- (Z1)-(Z5) é, de fato, equivalente à história e-qualia?

Começamos pelo primeiro ponto. Duas razões levam Kirk a afirmar que o defensor dos zumbis está comprometido com a transformação de z em z^* . Primeiramente, o fator de consciência x adicionado a z é fenomenalmente idêntico ao fator de consciência humana. O fator de consciência x , tal como y , não tem efeitos físicos e, portanto, garante que as teses (Z1)-(Z4) possam ser afirmadas. A segunda razão é que, dada a semelhança entre x e y e o fato de os seres humanos possuírem intimidade epistêmica com experiências, não há nada que impeça os seres de z^* de terem intimidade epistêmica com suas experiências.

¹²¹ O indeterminismo concebe a causalidade como não-linear, no sentido de que alguns eventos podem não ter causa ou cujas causas não podem ser determinadas. A mecânica quântica se abstém de afirmar se o universo é determinado ou indeterminado, e frequentemente filósofos utilizam essa abordagem para sustentar suas teses sobre a indeterminação causal.

Sepetyi sugere que Kirk reforce a segunda razão, afirmando que, como z^* é fisicamente semelhante ao mundo atual, os seres de z^* falaria e escreveriam sobre suas experiências. Para que isso fosse possível, eles precisariam ser capazes de perceber, atentar-se, pensar e comparar suas experiências¹²². Podemos denominar essa segunda etapa do argumento de Kirk, juntamente com a sugestão de Sepetyi, de argumento da intimidade epistêmica para z^* . Caso esse argumento seja bem-sucedido, o defensor da posição dos zumbis precisaria admitir a c-possibilidade de z^* ¹²³.

É importante destacar que a segunda etapa da argumentação de Kirk se fundamenta no fisicalismo cognitivo epifenomenalista – já mencionado anteriormente –, ou seja, na ideia de que o fator de consciência x não exerce efeitos físicos em z^* . No entanto, isso não implica que não haja conexões causais entre os fatores de consciência.

Sepetyi apresenta um problema para o argumento da intimidade epistêmica em z^* , que isentaria o defensor dos zumbis de ser forçado a aceitar tal conclusão. Segundo o autor, existe uma diferença crucial nas relações causais que torna (Z5) – ‘Os habitantes semelhantes aos humanos de z^* são capazes de perceber, prestar atenção, pensar e comparar as qualidades de suas experiências’ – inválida em z^* . No mundo atual, presume-se que há causalidade entre experiências e processos cognitivos, o que possibilita a existência de processos cognitivos como o pensamento e a atenção, permitindo-nos conhecer nossas próprias experiências. Vale lembrar que a incoerência da história *e-qualia* decorre exatamente desse ponto.

Por outro lado, para que o argumento da intimidade epistêmica em z^* seja válido, é necessário estipular que não exista a mesma causalidade em z^* , e essa estipulação decorre do fato de que o fator de consciência x não tem efeitos físicos – conforme a tese (Z3). O ponto que Sepetyi argumenta é que a ausência dessas relações causais em z^* sequer permitiria a existência de processos cognitivos, como o pensamento e a atenção. Segundo o autor, se não há relações causais entre experiências e processos cognitivos, então (Z5) não pode ser considerado verdadeira.

Embora z^* seja semelhante fisicamente e fenomenalmente ao mundo atual, os processos cognitivos em z^* não podem ser sobre experiências, pois a relação causal entre experiências e processos cognitivos está relacionada a *aboutness*¹²⁴ –

¹²² Cf. SEPETYI, 2022, p. 95.

¹²³ Sepetyi nota que em 2008 e 2017, Kirk omite o argumento da intimidade epistêmica para z^* , tomando como certo que a transformação de z em z^* tornaria os habitantes conscientes fenomenalmente e íntimos epistemicamente de suas experiências.

¹²⁴ O conceito de *aboutness* refere-se à característica de certos estados mentais (como pensamentos, crenças, desejos, intenções, etc.) de estarem "direcionados" para algo, ou seja, de serem sobre algo. Um estado mental é *about* (ou "sobre") algo quando ele tem um conteúdo, um objeto ou

em casos que o referente é um objeto existente.

Para ilustrar a discussão em questão, suponha que exista uma duplicata física de mim em z^* e que, na noite passada, eu tenha experimentado uma dor de cabeça intensa. Certamente, sou capaz de me lembrar da dor que senti, de relatar como foi a experiência e de compará-la com outras dores que já vivenciei. A minha duplicata, por sua vez, pode 'lembrar' de ter tido dores de cabeça na noite anterior, e tal 'lembrança' pode ser fisicamente e fenomenalmente idêntica à minha lembrança.

No entanto, a minha lembrança refere-se a uma experiência real, enquanto a lembrança da minha duplicata não diz respeito a uma experiência real, pois não há qualquer ligação causal entre a experiência e a sua 'lembrança'. Este é o problema que ocorre com os seres de z^* . Uma vez que se estipula que o fator de consciência x não tem efeitos físicos, não há uma relação causal entre as experiências e os processos cognitivos; portanto, (Z5) não pode ser considerada verdadeira. Assim, o defensor da posição dos zumbis pode refutar o argumento da intimidade epistêmica em z^* ao alegar que existem diferenças cruciais nas relações causais de z^* que impedem a validade de (Z5).

O segundo ponto apresentado por Sepetyi questiona se de fato as teses (Z1)-(Z5) são equivalentes à história e-qualia. Sepetyi sugere que um defensor dos zumbis que abraça o dualismo de substâncias poderia alegar que (E4) e (Z4) não são equivalentes.

(Z4): Os habitantes semelhantes a humanos de z^* consistem em nada além de corpos funcionais e seus x relacionados

(E4): Os seres humanos consistem apenas em corpos funcionais e em suas e-qualia relacionadas.

O defensor dos zumbis não é obrigado a adotar a visão de que o fator de consciência x seja apenas e-qualia. Alguns defensores podem adotar o dualismo de substâncias, acreditando que os seres humanos consistem apenas de um corpo funcional e um sujeito mental não-físico que é o portador de estados mentais fenomenais (qualia). Neste caso, não é concebível que existam e-qualia sem um sujeito mental não-físico subjacente a eles. (Z4) para um zombista dualista de substâncias ficaria assim:

(Z4*): Os habitantes semelhantes a humanos de z^* consistem apenas em corpos funcionais e a mente não-física que comporta o fator de consciência x .

uma proposição à qual se refere ou se relaciona. A Teoria Causal *aboutness* sugere que os estados mentais são *about* algo porque há uma relação causal entre esses estados e os objetos a que se referem. Por exemplo, a crença de uma pessoa de que está chovendo pode ser "sobre" a chuva porque sua crença é causada pela percepção de gotas de chuva.

É possível afirmar que Kirk antecipa essa objeção e responde dizendo que a diferença entre (Z4) e (E4) não é significativa, pois, mesmo que o fator de consciência x seja uma substância unitária em vez de um conjunto de propriedades, ele ainda teria que ser subjacente ou proporcionar uma pluralidade de propriedades, incluindo qualidades fenomenais. Portanto, as qualia fenomenais poderiam ser chamadas de e-qualia¹²⁵. Tal sugestão equivale a corrigir (E4) para:

(E4*) Os seres humanos consistem em nada além de corpos funcionais e suas e-qualia relacionadas, ou substâncias não físicas que são portadoras de suas e-qualia relacionadas.

(E4*) retorna a equivalência entre as teses, à medida que as e-qualia e um sujeito mental não-físico são os únicos candidatos para o papel do fator de consciência x . No entanto, Sepetyi argumenta que há uma diferença crucial entre (E3*) e (Z3).

Devemos lembrar que, na seção 4.1.1 Sepetyi sugere a reformulação da premissa (E3) para (E3*), alegando que a interpretação de Kirk está baseada na segunda, não na primeira. Agora, o autor afirma que as teses (Z3) e (E3*) não são equivalentes.

(Z3) x é causado por processos físicos, mas não tem efeitos físicos: pode ser removido sem perturbar o componente físico de z^* .

(E3*) Todas as e-qualia são diretamente causadas apenas por processos físicos, mas não têm efeitos físicos: elas podem ser removidas sem perturbar o mundo físico.

Primeiramente, devemos lembrar que a incoerência da história e-qualia se deve porque (E3*) proíbe eventos mentais não físicos – como experiências qualitativas – de causar ou participar na causação de quaisquer processos cognitivos (como atenção consciente, pensamento, etc.). Sendo semelhantes, (Z3) deveria ter a mesma proibição que (E3*) apresenta. Contudo, um defensor cartesiano dos zumbis – que entende que processos cognitivos, como o pensamento, pertencem a uma mente não-física – não encontra a mesma proibição em (Z3), pois tal premissa não proíbe a causalidade dentro do fator de consciência x .

Sepetyi alerta para o fato de que o fator de consciência x não precisa ser uma simples propriedade:

Um defensor dos zumbis pode sustentar (em linha com uma variedade de dualismo de propriedades) que o fator consciência é um

¹²⁵ Cf. KIRK, 2005, p. 52.

conjunto de estados mentais não físicos causalmente conectados (experiências conscientes causando consciência, atenção, pensamento, etc.). Ou ele pode sustentar que o fator consciência é uma entidade-substância com rica diferenciação interna, desenvolvimento temporal e relação causal interna entre seus estados — pode ser um sujeito mental completo, ou self, ou a alma cartesiana. Se for assim, não há incoerência entre (Z3) e (Z5) análoga à incoerência entre (E3*) e (E5). Então, a história z^* não é equivalente à história e-qualia de uma forma crucial que torna o argumento de Kirk inválido. (SEPETYI, 2022, p. 100)¹²⁶

Ou seja, o defensor cartesiano dos zumbis pode adotar o dualismo de propriedades e sustentar que há o fator de consciência x corresponde a conjunto de estados mentais não físicos causalmente conectados, ou pode adotar o dualismo de substâncias e defender que o fator de consciência x funciona como a alma cartesiana e tem relação causal entre seus estados mentais. Se o defensor cartesiano dos zumbis seguir um dos caminhos, não haveria incoerência entre (Z3) e (Z5) tal como a suposta incoerência entre (E3*) e (E5) e, portanto, as teses (Z1)-(Z5) não são equivalentes as teses da história e-qualia.

Para um interacionista não-cartesiano – que aceita o fisicalismo cognitivo –, existem dois modos de rejeitar a transformação de z em z^* .

4.1.3.1 Sobredeterminação causal em (Z3)

Para o interacionista não-cartesiano, Sepetyi considera duas possibilidades para enfrentar a segunda etapa da argumentação de Kirk:

1. Defender que, quando adicionamos a z o mesmo fator de consciência do mundo real, também incluímos seus poderes causais; ou
2. Defender que, quando adicionamos a z um fator de consciência que é exatamente igual ao do mundo real, mas é desprovido de seus poderes causais para produzir efeitos físicos.

Em (1), o interacionista pode alegar que o mundo resultante (z^*) não pode conter (Z3), tese que afirma que os fatores de consciência não tem efeitos físicos. Segundo Sepetyi, há um problema com a tese (Z3): no mundo z^* , os fatores

¹²⁶ “A zombist can hold (in line with a variety of property dualism) that the consciousness factor is a set of causally connected non-physical mental states (conscious experiences causing conscious awareness, attention, thought, etc.). Or she can hold that the consciousness factor is an entity-substance with rich internal differentiation, temporal development, and internal causal relationship between its states—it may be a full-blown mental subject, or self, or the Cartesian soul. If so, there is no incoherence between (Z3) and (Z5) analogous to the incoherence between (E3*) and (E5). So, the z^* story is not equivalent to the e-qualia story in a crucial way that makes Kirk’s argument invalid.”

físicos sozinhos têm todos os poderes causais necessários para produzir os efeitos que o fator de consciência gera no mundo real. Da mesma forma, o fator de consciência sozinho possui todos os poderes causais para gerar esses efeitos.

No entanto, em z^* , os poderes causais do físico e da consciência não atuam separadamente, mas são combinados. Isso sugere que essa combinação deveria resultar em efeitos diferentes (aditivos) em comparação com os efeitos que resultariam se apenas um dos fatores agisse. Se os efeitos forem de fato diferentes, isso significaria que o fator de consciência é causalmente eficiente, o que invalidaria a afirmação (Z3) de que o fator de consciência não tem efeitos físicos.

O autor sugere que, em z^* , os poderes causais do fator de consciência podem ser sobredeterminados, ou seja, eles não alteram os efeitos físicos, mas existem relações causais que ainda podem ser significativas. Essa situação implica que, embora o fator de consciência não tenha efeitos físicos diretos ou diferentes, ele ainda mantém uma conexão causal com o cérebro, o que pode ser suficiente para explicar fenômenos como percepção e atenção.

Portanto, z^* não é um mundo onde a consciência não tem efeitos físicos, mas sim um mundo no qual os efeitos causais da consciência não alteram os estados físicos do cérebro, embora ainda existam vínculos causais que podem ser relevantes para a experiência consciente, gerando intimidade epistêmica. Vale ressaltar que o interacionista não precisa defender que os efeitos no mundo atual são sobredeterminados, mas apenas sustentar que concebível – no sentido de não poder ser rejeitado *a priori* – um mundo similar ao nosso que é sobredeterminado.

Por outro lado, em (2), um interacionista não-cartesiano pode argumentar que o mundo z^* não se encaixa em (Z5), que afirma que os seres de z^* “são capazes de perceber, prestar atenção, refletir e comparar as qualidades de suas experiências”. No mundo resultante, existiriam estados *quasi*-cognitivos fisicamente (e funcionalmente) idênticos aos de perceber, prestar atenção e refletir sobre experiências no mundo real. No entanto, esses estados não deveriam ser considerados como uma verdadeira percepção, atenção ou reflexão sobre experiências, justamente porque não estão na relação causal apropriada com as experiências. Isso ocorre porque o argumento da intimidade epistêmica é falho, como foi mostrado na subseção anterior.

A descrição restante, de (Z1) a (Z4), sem (Z5), é fundamentalmente diferente da história e-qualia porque, na história e-qualia, a contradição surge entre (E1)-(E4*) de um lado e (E5) do outro. Por outro lado, a descrição do mundo de (Z1) a (Z5), é perfeitamente coerente, ao contrário da história incoerente das e-qualia. Portanto, Kirk falha em provar que a tese interacionista não-cartesiana é incoerente.

Para finalizar a presente seção, é válido pontuar as conclusões que podemos retirar da discussão feita por Sepetyi. Primeiro, o autor mostra que toda a interpretação de Kirk de (E3) é, na verdade, feita com base em (E3*) e, por isso, tal alteração é extremamente necessária para a interpretação da argumentação de Kirk. Além disso, ao contrário do que Kirk propõe, a concebilidade de zumbis ainda está disponível para um dualista interacionista, por exemplo, que pode adotar três vertentes, a qual Sepetyi denomina de substituibilismo, supercoincidentalismo ou insubstituibilismo. Até este ponto, já poderíamos afirmar que a história e-qualia não descarta totalmente a concebilidade de zumbis para um dualista. Para um interacionista, Sepetyi desenvolve uma resposta que o livra de ter que abraçar a conclusão do argumento da intimidade epistêmica para z^* . O argumento de Kirk se torna mais frágil quando Sepetyi mostra que existe uma diferença crucial que não permite que (Z5) seja verdadeira em z^* e que (Z4) e (E4) não são equivalentes. A conclusão que podemos retirar – juntamente com as críticas que apresentamos na seção 3.3 – é que existem pontos frágeis na argumentação de Kirk, que colocam à prova a sua crítica a concebilidade de mundos zumbis e principalmente sua premissa (II), a saber, “Se os zumbis fossem concebíveis, a história e-qualia seria concebível”, pois vimos que nem todo defensor dos zumbis estaria comprometido com a história e-qualia.

4.2 Defesa interacionista de mundos zumbis

Em seu artigo recente, [Khawaja \(2022\)](#) apresenta, como ele diz, um argumento que busca desarmar a afirmação de que a concebilidade de mundos zumbis implica que propriedades mentais são epifenomenais. O autor começa apresentando as premissas que envolvem o experimento de mundos zumbis.

- (1) Concebilidade: Os pares de mundos zumbis são concebíveis.
- (2) Relação entre concebilidade e possibilidade: Se algo é concebível, então é metafisicamente possível.
- (3) Lema: A partir de (1) e (2), os pares de mundos zumbis são possíveis.
- (4) Superveniência global: Para todo A, B: A e B são não distintos somente se não houver dois mundos diferentes com relação às suas propriedades A sem também diferir com relação às suas propriedades B.
- (\therefore) Dualismo: As propriedades mentais são distintas das propriedades físicas.

Uma vez que (3) e (4) são estabelecidas, há um mundo que é diferente do nosso em um aspecto: as propriedades mentais diferem das propriedades físicas, por-

tanto, a superveniência global falha e é possível estabelecer a tese dualista. Frequentemente, como vimos com Kirk, o argumento do zumbi nos leva a uma espécie de zumbi epifenomenal. Como argumentam, se zumbis são possíveis, então as propriedades mentais são epifenomenais, no sentido de serem inertes causalmente. Segundo Khawaja, a ideia de rejeitar os zumbis epifenomenais é baseada na alegação de que eles tornam os nossos relatos introspectivos não confiáveis, à medida que os zumbis produzem o mesmo relato introspectivo que nós, então nossos relatos não dependem da consciência, como é suposto pela maioria das pessoas. Acredita-se que o interacionista não poderia invocar o argumento do zumbis pelo mesmo motivo, porque o argumento do zumbi implicaria que as propriedades mentais são epifenomenais. A ideia do autor é mostrar que o interacionista pode invocar o argumento do zumbis.

Bailey (2009), por exemplo, defende que a possibilidade de zumbis (seja possibilidade lógica ou metafísica) torna a noção de consciência fenomênica inconsistente com o interacionismo, pois todos os eventos e estados – como comportamentos – poderiam ocorrer de qualquer modo, mesmo sem a presença de consciência fenomenal¹²⁷. Khawaja afirma que, por trás da afirmação de Bailey, está um critério necessário de dependência contrafactual para causalidade, que está explícito na ideia de que estados físicos – como comportamentos – poderiam ocorrer de qualquer forma sem a presença da consciência fenomenal.

Assim, segundo Khawaja, o argumento de Bailey para acusar os defensores dos zumbis de serem epifenomenalistas tem a seguinte forma. Seja A a consciência fenomenal e B os efeitos físicos:

- (1) Independência: Se um mundo com propriedades mentais tem uma duplicata zumbi, então os efeitos físicos desse mundo são contrafactualmente independentes das propriedades mentais.
- (2) Dependência contrafactual: A causa B somente se B depender contrafactualmente de A.
- (∴) Zumbis epifenomenais: Se mundos zumbis são possíveis, então as propriedades mentais são epifenomenais.

Tomado dessa forma, o argumento do zumbi implica que B não depende contrafactualmente de A, logo, A não causa B. Assim, a consciência fenomenal é epifenomenal. Khawaja argumenta que (2), a dependência contrafactual, é uma tese demasiadamente forte e, quando amenizada, os dualistas interacionistas escapam dessa conclusão.

¹²⁷ Cf. BAILEY, 2009, p. 131.

A noção de dependência contrafactual exprime a ideia de causalidade e é extremamente importante para o argumento do zumbi, pois o argumento do zumbi é formulado justamente contra a ideia de que propriedades físicas causam propriedades mentais. Essa noção envolve o cenário de mundos possíveis: dizer que B depende contrafactualmente de A é dizer que não há nenhum mundo possível - que seja mundo mais similar ao mundo atual - em que B ocorra sem A ocorrer.

Segundo Khawaja, podem existir dois critérios de causação mental, que está relacionada a tese (2). Considere M as propriedades mentais; P as propriedades físicas e B o comportamento correspondente a M. Considere os seguintes critérios:

- Critério Mínimo: M causa B somente se $(M \ \& \ \neg P) \Box \rightarrow B$.
- Critério Estrito: M causa B somente se (i) $(M \ \& \ \neg P) \Box \rightarrow B$, e (ii) $(P \ \& \ \neg M) \Box \rightarrow \neg B$ ¹²⁸.

No Critério Mínimo, a instanciação de propriedades mentais é capaz de produzir os efeitos comportamentais relevantes em um mundo possível, sem a instanciação de propriedades físicas. No Critério Estrito, (ii) é uma condição em que os comportamentos dependem contrafactualmente de instanciações de propriedades mentais, de modo que, se as propriedades mentais não fossem instanciadas, o comportamento não ocorreria. Khawaja diz que o Critério Estrito equivale a dizer que a sobredeterminação causal é metafisicamente impossível¹²⁹. Citamos a noção de sobredeterminação causal no capítulo 3. Retomando, a sobredeterminação causal ocorre quando um efeito é causado por múltiplas causas independentes, de modo que qualquer uma delas, isoladamente, seria suficiente para produzir o efeito.

Imagine que duas pessoas, Alice e Bruno, estão jogando pedras em um vidro. Cada um deles lança uma pedra ao mesmo tempo, e ambas acertam o vidro com força suficiente para quebrá-lo. Se apenas Alice tivesse lançado a pedra, o vidro teria quebrado e se apenas Bruno tivesse lançado a pedra, o vidro também teria quebrado. Neste caso, o vidro quebrou porque as duas pedras o atingiram, mas qualquer uma das ações (o lançamento de Alice ou de Bruno) seria suficiente, por si só, para causar o mesmo efeito. Isso caracteriza a sobredeterminação causal.

¹²⁸ Neste caso, as implicações são implicações contrafactuais, analisadas em termos de mundos possíveis. Assim, " $(P \rightarrow Q)$ é verdadeiro em w se e somente se existe um mundo possível $(P \ \& \ Q)$ que seja mais próximo de w do que qualquer mundo $(P \ \& \ \neg Q)$ " Cf. KHAWAJA, 2022, Nota 10. A ideia de "proximidade" ou "similaridade" entre mundos é central: queremos saber, se P fosse verdadeiro, o que provavelmente ou normalmente também seria verdadeiro, dado o mundo mais semelhante ao nosso onde P vale.

¹²⁹ Cf. KHAWAJA, 2022, p. 489.

No caso de propriedades mentais e físicas, um mundo com sobredeterminação causal seria um mundo em que causas mentais e causas físicas causam efeitos comportamentais, sendo que cada uma das causas são suficientes para se ter comportamento. No Critério Estrito, a rejeição da sobredeterminação causal se dá pois, dado dois eventos P e M que causam o efeito B, nenhum dos dois é causalmente suficiente para instanciar B, ou seja, os efeitos comportamentais relevantes só ocorrem se P e M ocorrer. Portanto, devemos notar que, no Critério Mínimo, a sobredeterminação causal é possível, já no Critério Estrito não¹³⁰.

Qual dos dois critérios está correto? Casos de sobredeterminação causal são raros. O Critério Mínimo aceita a sobredeterminação causal com maior facilidade, o que a literatura rejeita por entender que a sobredeterminação é extremamente rara. Por outro lado, o Critério Estrito é demasiado forte porque, de um modo geral, a literatura ainda admite a possibilidade de casos de sobredeterminação causal (mesmo que sejam raros), por exemplo, a sobredeterminação simétrica como o exemplo de Alice e Bruno acima. Por esse motivo, ao invés de excluir completamente a sobredeterminação causal com o Critério Estrito, Khawaja opta por complementar o Critério Mínimo, adotando a seguinte premissa:

- **Não há sobredeterminação generalizada (No Widespread Overdetermination, ou NWO):** No espaço modal, a sobredeterminação causal genuína é extremamente rara e, no mundo real, não há dois tipos de eventos ou propriedades X e Y, cada um dos quais causa individualmente algum efeito E, mas ambos são sempre co-instanciados de fato.

A NWO limita os casos de sobredeterminação, tornando-a extremamente rara nos mundos possíveis, e excluindo-a no mundo real. Segundo Khawaja, o ponto daqueles que rejeitam a causalção mental dualista é a exigência de que os efeitos comportamentais sejam sempre ou frequentemente sobredeterminados, e isso é rejeitado, pois, como vimos, os casos de sobredeterminação são extremamente raros. Entretanto, se a sobredeterminação é genuinamente possível, está em aberto a possibilidade para a causalção mental dualista em que

- (1) as causas físicas não são normalmente suficientes para os efeitos comportamentais relevantes; e
- (2) as causas físicas podem por vezes ser suficientes, e quando o são, isto produz uma ocorrência rara de sobredeterminação causal genuína.

Como os autores frequentemente rejeitam a sobredeterminação no mundo atual e na maioria dos mundos possíveis, Khawaja sugere que a sobredeterminação

¹³⁰ Cf. [KHAWAJA, 2022](#), p. 489.

poderia ocorrer em apenas um mundo possível qualquer, podemos chamá-lo de Z. Assim, o mundo real e a maioria dos mundos possíveis não hospeda a sobredeterminação, mas em Z ocorre a sobredeterminação. Agora, é possível defender que Z possui uma duplicata zumbi Z' e o par formado entre Z e Z' pode servir como base para argumentos zumbis favoráveis ao interacionismo.

A maneira como um interacionista deve responder ao argumento dos zumbis epifenomenais é concordar com seus oponentes de que não existem duplicatas zumbis do nosso mundo. No entanto, os interacionistas podem considerar outros pares de mundos que sejam consistentes com a causação mental dualista e que possam servir como uma base perfeitamente boa para argumentos de concebibilidade modificados. (KHAWAJA, 2022, p. 7)¹³¹

O desafio agora é mostrar que (1) e (2) são compatíveis. Para defender a possibilidade desse par de mundos Z e Z', Khawaja introduz as ideias sobre a mecânica estatística clássica (CSTM) para corroborar com a possibilidade de eventos improváveis que darão origem aos pares de mundos zumbis.

A mecânica estatística clássica utiliza teorias fundamentais da física, como a mecânica newtoniana, o eletromagnetismo, a teoria de probabilidades e a termodinâmica, para descrever o comportamento coletivo de sistemas com muitas partículas. A termodinâmica, em particular, é essencial para entender as propriedades macroscópicas dos sistemas (como temperatura, pressão, volume e entropia) a partir das interações microscópicas entre as partículas. A segunda lei da termodinâmica, que afirma que a entropia de um sistema isolado tende a aumentar, é crucial para a descrição do comportamento dos sistemas, refletindo a tendência dos sistemas evoluírem para estados mais prováveis, ou seja, aqueles com mais microestados possíveis.

Na mecânica estatística, a entropia é uma medida da probabilidade de um sistema estar em um determinado estado. Um microestado é uma configuração específica do sistema, descrevendo a posição e o movimento de todas as partículas, enquanto um macroestado é caracterizado por propriedades macroscópicas (como temperatura e pressão), sem detalhar o comportamento individual das partículas.

Quando se fala em entropia, frequentemente associamos o termo à desordem, mas na mecânica estatística, ela está relacionada ao número de microestados possíveis. Quanto maior o número de microestados em um dado macroestado, maior é a entropia do sistema. Isso pode ser ilustrado com o exemplo de

¹³¹ "The way an interactionist should respond to the epiphenomenal zombies argument is to agree with their opponents that there are no zombie duplicates of our world. However, interactionists can consider other world-pairs which are consistent with dualist mental causation, and which may serve as a perfectly good basis for modified conceivability arguments."

um cubo de gelo: no estado sólido, as moléculas de água estão organizadas, formando uma rede cristalina. Isso resulta em um número relativamente pequeno de microestados, e a entropia é baixa. Quando o gelo derrete e se transforma em água líquida, as moléculas se tornam mais desordenadas e têm maior liberdade para se mover, o que aumenta o número de microestados possíveis e, consequentemente, a entropia do sistema. Esse aumento reflete a maior probabilidade do sistema estar no estado líquido, onde as moléculas se movem de maneira mais variada.

A mecânica estatística clássica é construída sobre as leis determinísticas da mecânica clássica, que permitem prever, teoricamente, o comportamento futuro de um sistema se conhecermos as condições iniciais de todas as partículas. No entanto, quando lidamos com sistemas compostos por muitas partículas, como gases, líquidos ou sólidos, torna-se impraticável rastrear a evolução de cada partícula individualmente devido à complexidade e ao número de interações possíveis. Por isso, a mecânica estatística foca no comportamento coletivo do sistema.

Em sistemas grandes, o número de microestados possíveis é gigantesco, e a mecânica estatística descreve o comportamento do sistema de forma probabilística. Ela introduz a ideia de que, embora flutuações aleatórias (eventos improváveis) possam ocorrer, a probabilidade de eventos extremos é muito baixa. Por exemplo, em um gás ideal, é improvável que todas as partículas se movam na mesma direção ao mesmo tempo, mas essas flutuações não afetam significativamente as propriedades médias do sistema, como temperatura e pressão.

Essas flutuações não são descartadas pela mecânica estatística, mas como a quantidade de partículas é imensa, o comportamento macroscópico do sistema tende a seguir as leis determinísticas da termodinâmica. Isso garante que propriedades como temperatura e pressão, que são médias de grande número de partículas, permaneçam estáveis, mesmo em face de pequenas flutuações térmicas. Dessa forma, a mecânica estatística fornece uma descrição probabilística do sistema, enquanto a termodinâmica assegura a estabilidade do comportamento global.

No entanto, ainda que a CSTM explique probabilisticamente como cenários improváveis são extremamente difíceis de acontecer, é justamente este ponto que Khawaja está interessado: .

A adição da CSTM à nossa imagem física do mundo não serve para excluir tais cenários – na verdade, eles continuam a ser possibilidades nômicas genuínas. Em vez disso, a CSTM torna tais cenários extremamente improváveis, na medida em que seriam realizados por microestados altamente atípicos. É esta lição ampla que devemos ter em

mente quando apresento, no restante desta seção, os pares de mundos que penso que podem servir de base para um argumento interacionista favorável aos zumbis. (KHAWAJA, 2022, p. 9)¹³²

Portanto, a mecânica estatística clássica não descarta cenários improváveis, mas trata esses cenários como altamente improváveis. Ela utiliza probabilidades para descrever a distribuição dos microestados e a tendência dos sistemas de se encontrarem em estados mais prováveis (com maior entropia) ao longo do tempo, mas reconhece que eventos improváveis ainda podem ocorrer, embora com uma frequência muito baixa. O conceito de improbabilidade está diretamente ligado à natureza estatística e probabilística dos sistemas, e os eventos raros ou improváveis não são descartados, mas sim entendidos como exceções em um contexto mais amplo.

A questão é que, para que esses eventos altamente improváveis ocorram, é necessário um cenário altamente atípico com microestados. Com isso em mente, um interacionista pode defender que os estados e propriedades mentais são apenas causalmente suficientes para produzir os tipos relevantes de efeitos comportamentais, mas também que as propriedades e estados mentais são causalmente necessários, na medida em que os efeitos geralmente dependem contrafactualmente das propriedades mentais e não são causalmente sobredeterminados.

Entretanto, à luz do que foi apresentado na discussão sobre a mecânica estatística clássica, o interacionista poderia argumentar que, em alguns mundos, situações altamente improváveis poderiam ocorrer, tornando os efeitos comportamentais sobredeterminados. Em situações como essa, poder-se-ia defender a existência de duplicatas zumbis e construir um argumento dos zumbis favorável aos dualistas interacionistas. Para entender como isso ocorreria, considere o seguinte diagrama para explicitar o que acontece na cadeia causal das propriedades físicas, mentais e do comportamento:

¹³² “The addition of CSTM to our physical picture of the world serves not to rule out such scenarios—indeed, they remain genuine nomic possibilities. Rather, CSTM renders such scenarios extremely unlikely, insofar as they would be realized by highly atypical microstates. It is this broad lesson that one should bear in mind as I introduce, in the remainder of this section, the world-pairs that I think can serve as a basis for an interactionist-friendly zombie argument.”

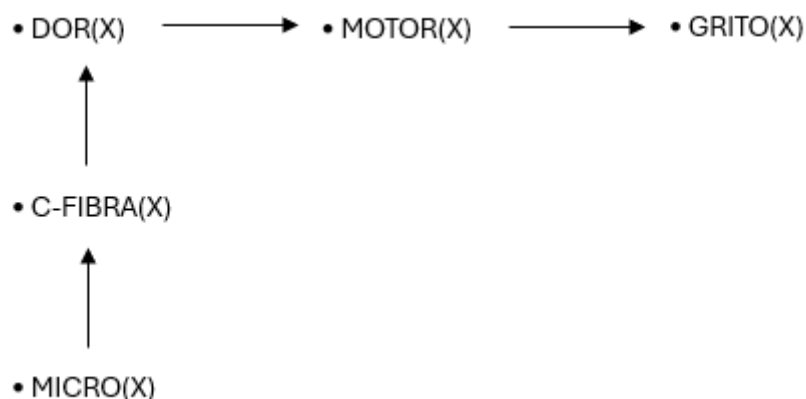


Figura 4.1:

Na figura 4.1, as setas verticais representam algum tipo de relação de dependência não causal, como, por exemplo, a emergência forte ou a realização dos estados. As setas horizontais representam a dependência causal. Além disso, MICRO(x) refere-se à instancição de algum estado microfísico preciso para algum sistema x , C-FIBRA(x) refere-se à instancição do estado cerebral de ativação da C-FIBRA, DOR(x) refere-se à instancição da propriedade fenomenal da dor, MOTOR(x) refere-se à instancição de alguma resposta motora em resposta à dor e GRITO(x) refere-se ao efeito comportamental de um ser consciente gritando de dor.

Para satisfazer a NWO, o interacionista precisa defender que se DOR(x) não for instanciado, então MOTOR(x) e GRITO(x) também não são instanciados, afinal, o interacionista defende a causalidade de propriedades mentais. Ou seja, o interacionista defende que efeitos comportamentais como GRITO(x) dependem contrafactualmente de instancições de propriedades mentais como DOR(x). É o que representa a figura 4.2. As bolinhas em branco indicam que a instancição de DOR(X) não ocorreu – talvez por violação de alguma lei de emergência¹³³ – e, por consequência, as instancições de MOTOR(X) e GRITO(X) não ocorreram e isso é sinalizado pelas setas tracejadas.

¹³³ O emergentismo em filosofia da mente defende que, grosso modo, as propriedades mentais emergem de propriedades físicas, mas que não podem ser reduzidas completamente a propriedades físicas. Nesse sentido, haveria dois níveis distintos de propriedades. Frequentemente teóricos postulam leis que explicam tal interação, e.g., a lei geral de emergência pode ser do tipo $\Box \forall x (P(x) \rightarrow M(x))$, no qual P são propriedades físicas e M são propriedades mentais. Para ver mais, sugiro a leitura de [Kim \(1999\)](#).

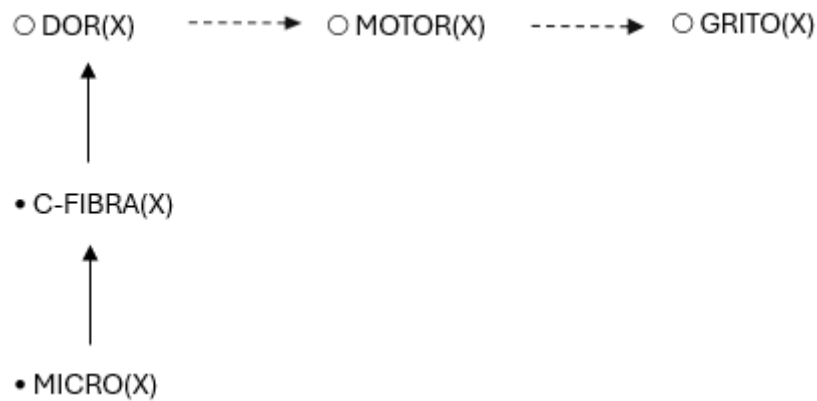


Figura 4.2:

Por outro lado, devemos lembrar que o Critério Mínimo permite que as instâncias de $DOR(x)$ sejam causalmente suficientes para as instâncias de $GRITO(x)$. Portanto, o Critério Mínimo permite que, mesmo que $MICRO(x)$ e $C-FIBRA(x)$ não sejam instanciadas e $DOR(x)$ seja instanciada, ainda teríamos a ocorrência de $GRITO(x)$. Algo como a figura 4.3.

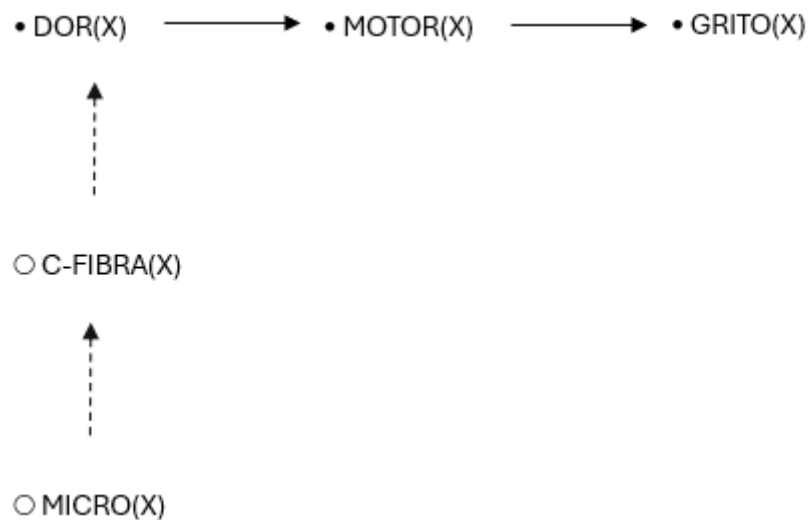


Figura 4.3:

Finalmente, os interacionistas poderiam argumentar que, devido a alguma mudança evolutiva ao longo da história, certos microestados do tipo $MICRO(x)$ poderiam evoluir atipicamente de tal forma que se tornariam capazes de instanciar um estado cerebral como $C-FIBRA(x)$, imediatamente seguido por um estado cerebral como $MOTOR(x)$, o qual produziria o comportamento $GRITO(x)$, mesmo sem instanciar $DOR(x)$. Isso significaria que, para esse $MICRO(x)$ atípico específico, começariam a gritar de maneira típica, como se estivessem gritando devido à dor, sem, no entanto, sentir dor.

Segundo Khawaja, “essa ideia vem naturalmente da CSTM, que torna esses microestados bizarros improváveis sem torná-los nomicamente impossíveis” (KHAWAJA, 2022, p. 11)¹³⁴. A figura 4.4 ilustra como seria esse caso: atípico:

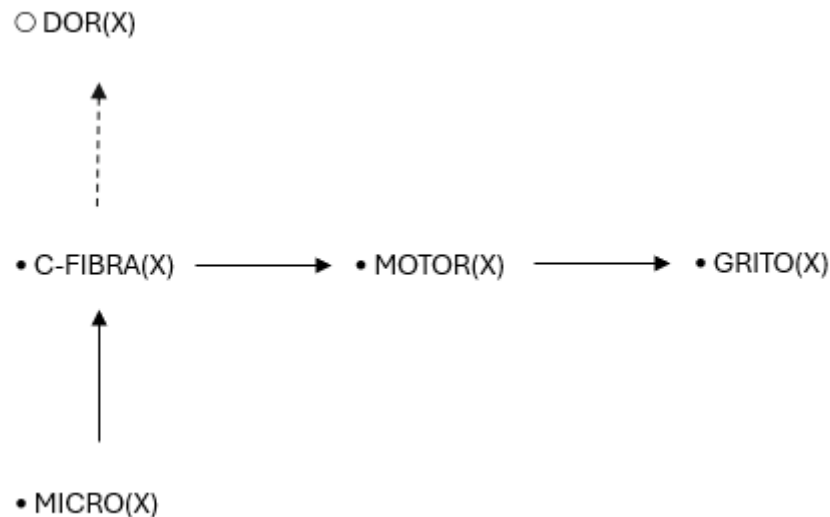


Figura 4.4:

Neste mundo, $DOR(x)$ não é instanciada, mas, devido à evolução atípica de $MICRO(x)$, a sequência de instancicações é mantida normalmente. Com base nessas figuras e na possibilidade da evolução atípica de $MICRO(x)$, agora é possível construir o argumento dos zumbis favorável ao interacionismo. Suponha que, em um mundo, um microestado inicial evolui e agora é capaz de instanciar todas as causas mentais, mesmo sem ter uma propriedade mental instanciada, como ilustrado na Figura 4.4. Chamemos esse mundo de w^*z . Suponha agora o mundo descrito pela Figura 4.1 — chamemos esse mundo de w^* . w^* é um mundo que hospeda as instancicações de propriedades mentais, como descrito na Figura 4.1, e é diferente de w^*z , que não hospeda as instancicações de propriedades mentais. Entretanto, há uma característica entre esses dois mundos, a saber,

sempre que em w^* há um processo causal como o da figura 4.1, w^*z hospeda um processo causal como o da figura 4.4. Por outras palavras, w^*z acaba por se parecer com w^* , porque sempre que as propriedades mentais em w^* desempenham um papel causal, a assustadora microfísica de w^*z intervém para produzir exatamente as mesmas sequências de respostas neurais e comportamentais. Podemos ainda dizer que em w^* e w^*z existe esse microestado inicial e ele evolui

¹³⁴ “This idea comes naturally from CSTM, which renders these bizarre microstates unlikely without making them nomically impossible.”

deterministicamente nos dois mundos, sendo que em w^* as propriedades mentais são instanciadas e em w^*z não são. Consequentemente, os dois mundos são duplicatas físicas, e um hospeda as propriedades fenomênicas e o outro é um mundo zumbi. Como resultado, w^* também abriga uma microfísica atípica que é suficiente, por si só, para produzir os efeitos comportamentais relevantes. Portanto, os efeitos relevantes são todos causalmente sobredeterminados em w^* pelas propriedades mentais e microfísicas. (KHAWAJA, 2022, p. 12)¹³⁵

Khawaja supõe que o microestado evoluído poderia existir tanto em w^*z quanto em w^* . Se fosse o caso, em w^* os efeitos comportamentais seriam sobredeterminados, enquanto em w^*z somente as propriedades físicas seriam capazes de produzir efeitos comportamentais. É claro que esses casos sobredeterminados são extremamente raros — como, por exemplo, o caso de Alice e Bruno lançando pedras em um vidro —; no entanto, não há nada na natureza física que impeça a ocorrência desses casos. De maneira similar, Khawaja propõe que mundos como w^* e w^*z não violam nenhuma lei da física.

Um ponto importante é que w^* é nomicamente possível, apesar de hospedar a sobredeterminação generalizada. Isso porque as leis do nosso mundo — que são de forma semelhante à CSTM — permitem a existência de microestados atípicos relevantes. Novamente, a CSTM explica:

- (i) como mundos como w^* são altamente improváveis no espaço de mundos nomicamente possíveis; e, ainda assim,
- (ii) como esses mundos permanecem nomicamente possíveis.

O item (i) é explicado porque, na ausência de conexões entre C-FIBRA(x) e MOTOR(x), podem existir microestados possíveis consistentes com C-FIBRA(x) que realizam todos os tipos de instanciações diferentes. No entanto, poucas dessas instanciações corresponderiam ao tipo exato de atividade realizada no cérebro, e, por isso, é extremamente raro. Por outro lado, (ii) é explicada devido ao fato de que nada na CSTM limita os casos de microestados raros no qual “as trajetórias passam por um estado cerebral de disparo de C-FIBRA seguido por um estado cerebral de atividade motora que produz dor!” (KHAWAJA, 2022,

¹³⁵ “wherever in w^* there is a causal process like that in Fig. 1, w^*Z hosts a causal process like that in Fig. 4. In other words, w^*Z ends up looking just like w^* , because wherever the mental properties in w^* are playing a causal role, the spooky microphysics of w^*Z steps in to produce exactly the same sequences of neural and behavioral responses.²¹ We can further suppose that w^* and w^*Z share the same precise initial microstate, and that said microstate evolves deterministically throughout all of history in exactly the same way in both w^* and w^*Z .²² Consequently, the two worlds are quark-for-quark physical duplicates, and yet one hosts phenomenal properties while the other is a zombie world. As a result, w^* also hosts an atypical microphysics which is sufficient, on its own, to bring about the relevant behavioral effects. Hence, the relevant effects are all causally overdetermined in w^* by the mental and microphysical properties.”

p. 14)¹³⁶. Como (i) é satisfeita significa que w^* , sendo extremamente raro, é compatível com a NWO. E, por (ii) ser satisfeita, é possível montar os pares de mundos para a concebilidade dentro das leis físicas.

Dito isso, finalmente temos o argumento do autor, possibilitando que interacionistas mantenham o argumento do zumbi contra o fisicalismo. Com w^* e w^*z , o interacionista modifica o argumento da concebilidade para:

- (1) Superveniência Global: Para todos os A, B: A e B não são distintos apenas se não houver dois mundos diferentes em relação às suas propriedades A sem também diferir em relação às suas propriedades B.
 - (2) Mundos Zumbis Milagrosos: Existem concebivelmente dois mundos possíveis w^* e w^*z , de modo que w^* e w^*z são fisicamente idênticos ao longo de toda a sua história, mas diferem em relação às suas propriedades mentais.
 - (3) Relação entre Concebilidade-Possibilidade: Se algo é concebível, então é metafisicamente possível.
 - (4) Lema: de (2) e (3), existem dois mundos possíveis, w^* e w^*z , de modo que w^* e w^*z são fisicamente idênticos ao longo de toda a sua história, mas diferem em relação às suas propriedades mentais.
- (\therefore) Dualismo: As propriedades mentais são distintas das propriedades físicas.

Esse argumento difere do argumento inicial porque adiciona a premissa (2). Segundo o autor, esse argumento modificado bloqueia a objeção apresentada por Bailey. De acordo com Bailey, permitir que seja possível subtrair a consciência fenomenal dos indivíduos e torná-los zumbis é permitir que a consciência fenomenal seja desnecessária para a ocorrência de qualquer fenômeno físico, uma vez que todos os eventos ocorreriam da mesma forma sem a consciência. Trata-se da banalização da consciência fenomenal. No argumento de Khawaja, isso não ocorre. A consciência fenomenal, ou melhor, as propriedades mentais, ocorrem no mundo w^* , mas, ainda em w^* , existe um microestado que evolui e é capaz de gerar comportamentos semelhantes sem que estados mentais estejam ocorrendo. E tudo isso é corroborado pela CSTM.

O argumento vale para o interacionismo à medida que o autor deixa claro que essas propriedades mentais e físicas interagem, como mostra nas figuras 1-4. A questão é que não necessariamente o comportamento GRITO(x) é realizado pela propriedade mental DOR(x). Mesmo que DOR(x) não esteja na cadeia causal, um microestado evoluído pode realizar GRITO(x). E o ponto principal

¹³⁶ “the trajectories so happen to pass through a brain state of C-FIBER firing followed by a brain state of pain-producing motor activity!”

é que a ocorrência desse microestado extremamente raro não infringe as leis de interação e leis de emergência, a saber $\Box \forall x(DOR(x) \rightarrow GRITO(x))$ e $\Box \forall x(C-FIBRA(x) \rightarrow DOR(x))$, respectivamente¹³⁷. Essas duas leis juntas explicam a regularidade de $\forall x(C-FIBRA(x) \rightarrow GRITO(x))$. No entanto, Khawaja questiona se $\forall x(C-FIBRA(x) \rightarrow GRITO(x))$ é uma lei em w^*z . Se essas leis de emergência e interação são leis da natureza, então w^* e w^*z não seriam fisicamente idênticos porque claramente em w^*z essas leis falham. Seguindo os critérios de [Mohamadian \(2021\)](#) para caracterizar algo como uma lei, Khawaja mostra que $\forall x(C-FIBRA(x) \rightarrow GRITO(x))$ não é uma lei da natureza.

O primeiro critério é que, em geral, as leis explicam as instâncias, porém, em $\forall x(C-FIBRA(x) \rightarrow GRITO(x))$, não está claro qual o tipo correto de microestado instancia $C-FIBRA(x)$, e muito menos sabemos qual o tipo correto de $C-FIBRA(x)$ instancia $GRITO(x)$. Desse modo, “sem especificar também o estado microfísico exato, ficamos com condições insuficientes para a instanciação do $GRITO(x)$ ” ([KHAWAJA, 2022](#), p. 16)¹³⁸.

O segundo critério é que as leis suportam contrafactuais. Entretanto, $\forall x(C-FIBRA(x) \rightarrow GRITO(x))$ claramente não suporta os contrafactuais. É possível, por exemplo, que um microestado realizasse $C-FIBRA(x)$ sem que isso instanciasse $GRITO(x)$, como é o caso da figura 4.2.

De acordo com o terceiro critério, as leis são capazes de apoiar conexões inferenciais entre instâncias observáveis e instâncias não observáveis. No entanto, é questionável se as instâncias observáveis mantêm a regularidade das instâncias não observáveis. Suponha que mini-agentes são capazes de observar os microestados do cérebro e eles possuem o conhecimento suficiente para mapear as instanciações do cérebro. Segundo Khawaja, a conclusão que os mini-agentes chegarão é de que cada instanciação, por exemplo, de $(C-FIBRA(x) \& GRITO(x))$ são um acaso, isso porque a regularidade desses casos depende dos microestados: se o microestado fosse alterado, $GRITO(x)$ poderia não ser instanciado. Por conseguinte, “os agentes perceberiam que não havia razão para inferir, a partir das instâncias observadas de $(C-FIBRA(x) \& GRITO(x))$, que as instâncias não observadas de $C-FIBRA(x)$ continuariam a obedecer a essa regularidade. Portanto, a suposta lei não daria suporte a inferências sobre regularidades futuras ou não observadas” Cf. [KHAWAJA, 2022](#), p. 16¹³⁹

O objetivo desta seção foi apresentar um argumento a favor do zumbi,

¹³⁷ Aqui, ' \Box ' se refere a necessidade nômica.

¹³⁸ “Without also specifying the exact microphysical state, we are left with insufficient conditions for the instantiation of $SCREAM(x)$ ”

¹³⁹ “the agents would come to realize that there was no reason to infer from the observed instances of $(C-FIBER(x) \& SCREAM(x))$ that the unobserved instances of $C-FIBER(x)$ would continue to obey this regularity. Hence, the putative law would not support inferences to future or unobserved regularities.”

que permanece disponível para os dualistas interacionistas. É importante salientar que, para um interacionista, apresentar qualquer argumento anti-fisicalista é, antes de tudo, rejeitar o fechamento causal da física. No entanto, é possível afirmar que o argumento de Khawaja demonstra como a história e-qualia falha ao tentar ser um argumento definitivo contra aqueles que defendem a concebilidade de zumbis.

Capítulo 5

Conclusão

Este trabalho teve como objetivo apresentar o debate em torno do argumento do zumbi, com ênfase nas contribuições de Robert Kirk a partir de uma perspectiva crítica, tanto em sua fase de defesa da concebibilidade dos zumbis quanto em sua fase anti-zumbis.

No Capítulo 2, foram discutidas as principais questões relacionadas ao experimento do zumbi: O que é um zumbi e o que lhe falta? Contra quais teses o argumento é formulado? Quais filósofos defendem a concebibilidade dos zumbis? Observou-se que a mera afirmação da concebibilidade dos zumbis não é suficiente; é necessário dar um passo adicional e demonstrar sua possibilidade metafísica. Nesse sentido, o bi-dimensionalismo de Chalmers representa uma alternativa para aqueles que buscam um caminho para avançar nessa questão.

Esse capítulo serviu como introdução aos debates desenvolvidos no Capítulo 3. Nele, analisamos as duas fases do pensamento de Robert Kirk, com foco nas críticas a suas posições. Observamos que a sua fase pró-zumbis é limitada, pois se concentra em refutar apenas algumas teses materialistas que, à época, já estavam em declínio nos debates filosóficos. Além disso, sua caracterização do que seria necessário para sustentar o materialismo é frágil e insatisfatória, uma vez que materialistas e fisicalistas contemporâneos rejeitariam tal definição, pois defendem que grande parte das descrições mentais é relacional — como os representacionalistas.

Ademais, os exemplos de zumbis apresentados por Kirk nessa fase revelam-se problemáticos. Um materialista poderia argumentar que Zulliver não é uma réplica física exata de um ser humano, pois as minicriaturas que invadem seu cérebro alteram sua estrutura, tornando-o diferente. Por outro lado, o exemplo de Dan pressupõe que seu cérebro não sofre nenhuma alteração ao longo da perda de consciência, uma suposição inaceitável para alguns modelos cognitivos de consciência.

Diante disso, conclui-se que os argumentos fornecidos por Kirk em sua

fase de defesa dos zumbis são ingênuos, pois contêm fragilidades que os tornam pouco atrativos para aqueles que buscam argumentos sólidos a favor da possibilidade dos zumbis e, além disso, atacam apenas uma parcela restrita dos materialistas.

No Capítulo 3, analisamos também a fase anti-zumbi de Kirk, em particular seu argumento da história e-qualia, que visa eliminar definitivamente a concebibilidade dos zumbis. Esse argumento sustenta que a defesa da concebibilidade dos zumbis implica a aceitação de uma noção de consciência epifenomenal, a qual seria contraditória com a forma como lidamos com nossas experiências conscientes.

Observamos que o argumento de Kirk se baseia na ideia de que só podemos ter conhecimento e formar crenças sobre nossas experiências qualitativas se estas participarem da cadeia causal do comportamento. No entanto, essa não é a única maneira de conceber o conhecimento das experiências qualitativas, pois um epifenomenalista poderia argumentar que há uma causa física comum entre experiências qualitativas e comportamentos, permitindo a conexão entre esses dois eventos.

Além disso, a argumentação de Kirk apresenta fragilidades conceituais, especialmente em relação às noções de intimidade epistêmica e concebibilidade. Nesse sentido, a história e-qualia contém falhas significativas, como no exemplo das correntes elétricas, que não consegue captar adequadamente a natureza do conhecimento que os seres humanos possuem de suas experiências qualitativas.

Por fim, concluímos que Kirk não alcança plenamente seu objetivo de eliminar a concebibilidade dos zumbis. Seu argumento atinge principalmente os epifenomenalistas, ao afirmar que as e-qualia não possuem poder causal; ainda assim, esses teóricos podem oferecer respostas a essa objeção, conforme discutido nas seções 3.3.1 e 3.3.2. Já os interacionistas, por exemplo, rejeitariam essa concepção de experiências qualitativas desde o princípio. Dessa forma, o máximo que a história e-qualia pode alcançar é gerar dificuldades para os epifenomenalistas na defesa dos zumbis. O capítulo 4 reforça essa conclusão ao apresentar críticas contundentes de Sepetyi à história e-qualia, além de demonstrar que as teses (E1)-(E5) não são equivalentes às teses (Z1)-(Z5). Além disso, o autor conclui que existem caminhos para que os interacionistas adotem o argumento dos zumbis. Nesse sentido, a seção 4.2 apresenta um argumento interacionista contemporâneo de Khawaja, que rejeita a tese de que a concebibilidade dos zumbis implica necessariamente em uma consciência epifenomenal—posição semelhante à defendida por Kirk.

Dessa maneira, a conclusão geral é que, para os defensores do argumento dos zumbis, é fundamental estruturar uma base argumentativa sólida, incluindo

uma concepção rigorosa de concebibilidade. No entanto, isso não é suficiente. Sustentar que os zumbis são concebíveis constitui apenas o primeiro passo, pois o verdadeiro desafio está em demonstrar a existência de uma conexão entre o que se concebe e sua possibilidade metafísica.

Por outro lado, aqueles que rejeitam a concebibilidade dos zumbis precisam apresentar um argumento suficientemente robusto para sustentar essa posição. A simples alegação de que a concebibilidade dos zumbis implica uma consciência epifenomenal não é suficiente para refutar todas as abordagens que sustentam o argumento dos zumbis. Assim, considerando o objetivo de Kirk com a história *e-qualia*, o presente trabalho evidencia suas limitações e seu fracasso em eliminar definitivamente a concebibilidade dos zumbis.

Referências Bibliográficas

ALVES, M. A. S. David chalmers e o mundo zumbi. In: *TENDÊNCIAS CONTEMPORÂNEAS DE FILOSOFIA DA MENTE E CIÊNCIAS COGNITIVAS*. [S.l.]: PUCPRESS, 2023. v. 1, cap. Cap. 13, p. 15. [2.2.3](#), [28](#)

BAARS, B. J. *A Cognitive Theory of Consciousness*. New York: Cambridge University Press, 1988. [3.1.3](#)

BAILEY, A. Zombies and epiphenomenalism. *Dialogue*, Cambridge University Press, v. 48, n. 1, p. 129, 2009. [4.2](#), [127](#)

BALOG, K. Acquaintance and the mind-body problem. In: GOZZANO, S.; HILL, C. S. (Ed.). *New Perspectives on Type Identity: The Mental and the Physical*. [S.l.]: Cambridge University Press, 2012. p. 16–42. [3.3.3](#)

CHALMERS, D. Does conceivability entail possibility. In: GENDLER, T. S.; HAWTHORNE, J. (Ed.). *Conceivability and Possibility*. [S.l.]: Oxford University Press, 2002. p. 145–200. [2.2.2](#), [2.2.3](#), [3.3.3](#), [3.3.4](#)

CHALMERS, D. J. *The Conscious Mind: In Search of a Fundamental Theory* (2nd edition). New York: Oxford University Press, 1996. [2.1](#), [25](#), [3.3.3](#), [106](#)

CHALMERS, D. J. Imagination, indexicality, and intensions. *Philosophy and Phenomenological Research*, Wiley-Blackwell, v. 68, n. 1, p. 182–90, 2004. [4.1.2](#), [116](#)

CHALMERS, D. J. *The Character of Consciousness*. New York: Oxford University Press, 2010. [2.2](#), [2.2.2](#), [17](#), [3.1.3](#)

CHALMERS, D. J. The nature of epistemic space. In: EGAN, A.; WEATHERSON, B. (Ed.). *Epistemic Modality*. [S.l.]: Oxford University Press, 2011. [2.2.3](#)

DEPOE, J. M. Knowledge by acquaintance and knowledge by description. In: *Internet Encyclopedia of Philosophy*. [S.l.: s.n.], 2013. [3.3.3](#)

GOMES, L. M. Filosofia da mente, dualismo e zumbis. *Cadernos PET Filosofia*, v. 21, n. 1, p. 101–116, 2020. [2.2](#)

JACKSON, F. Epiphenomenal qualia. *Philosophical Quarterly*, v. 32, p. 127–36, 1982. [2.1](#)

JACKSON, F. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. New York: Oxford University Press, 1998. [10](#), [2.2.1](#), [11](#), [13](#), [14](#), [3.3.5](#)

- KHAWAJA, J. Interactionist zombies. *Synthese*, Springer Verlag, v. 200, 2022. 4.2, 128, 129, 130, 4.2, 4.2
- KIM, J. Making sense of emergence. *Philosophical Studies*, Springer Netherlands, v. 95, n. 1-2, p. 3–36, 1999. 133
- KIND, A. Qualia realism. *Philosophical Studies*, v. 104, n. 2, p. 143–62, 2001. 2
- KIRK, R. Zombies v. materialists. *Aristotelian Society Supplementary Volume*, Wiley-Blackwell, v. 48, n. 1, p. 135–164, 1974a. 3.1.1, 31, 34, 35, 36, 41, 43, 44, 45, 48, 3.1.3
- KIRK, R. Sentience and behaviour. *Mind*, Oxford University Press, v. 83, n. January, p. 43–60, 1974b. 3.1.2, 50, 51, 53, 55, 56, 57, 60, 61, 62, 3.1.3
- KIRK, R. Reply to don locke on zombies and materialism. *Mind*, Oxford University Press, v. 86, n. April, p. 262–4, 1977. 32, 3.1.1
- KIRK, R. *Zombies and Consciousness*. Oxford: Oxford University Press UK, 2005. 2.2.3, 3.2, 65, 68, 69, 71, 72, 73, 74, 76, 77, 78, 81, 85, 86, 88, 3.3.1, 3.3.3, 105, 109, 4.1.1, 125
- KIRK, R. The inconceivability of zombies. *Philosophical Studies*, Springer, v. 139, n. 1, p. 73–89, 2008. 123
- KIRK, R. *Robots, Zombies and Us: Understanding Consciousness*. New York, NY, USA: Bloomsbury Academic, an imprint of Bloomsbury Publishing, Plc, 2017. 123
- KRIPKE, S. A. *Naming and Necessity: Lectures Given to the Princeton University Philosophical Colloquium*. Cambridge, MA: Harvard University Press, 1980. 2.2.2
- LEVINE, J. The explanatory gap. In: WEISBERG, J. (Ed.). *Consciousness (Key Concepts in Philosophy)*. [S.l.]: Polity, 2014. 90
- LEWIS, C. I. *Mind and the world-order*. New York: Charles Scribner's Sons, 1929. 2.1, 3
- LEWIS, D. K. An argument for the identity theory. *Journal of Philosophy*, Journal of Philosophy, v. 63, n. 1, p. 17–25, 1966. 59
- MOHAMMADIAN, M. If consciousness causes collapse, the zombie argument fails. *Synthese*, Springer Verlag, v. 199, p. 1599–1615, 2021. 4.2
- MOORE, D. On robinson's response to the self-stultifying objection. *Review of Philosophy and Psychology*, Springer Verlag, v. 3, n. 4, p. 627–641, 2012. 3.3.2, 97
- ROBINSON, W. *Epiphenomenalism*. Disponível em: <https://plato.stanford.edu/archives/sum2023/entries/epiphenomenalism/>. 3.3.1
- ROBINSON, W. S. *Understanding Phenomenal Consciousness*. New York: Cambridge University Press, 2004. 3.3.1, 89, 92, 3.3.2, 95, 96

- ROBINSON, W. S. Experiencing is not observing: A response to dwayne moore on epiphenomenalism and self-stultification. *Review of Philosophy and Psychology*, Springer Verlag, v. 4, n. 2, p. 185–192, 2013. [3.3.2](#), [99](#)
- RUFFINO, M. *Contingent a Priori Truths: Metaphysics, Semantics, Epistemology and Pragmatics*. [S.l.]: Springer Nature, 2022. [15](#)
- RUSSELL, B. On denoting. *Mind*, Duke University Press, v. 14, n. 56, p. 479–493, 1905. [3.3.3](#)
- RUSSELL, B. Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society*, Pelican Books, v. 11, p. 108–28, 1910. [3.3.3](#)
- SEPETYI, D. Robert kirk’s attempted intellectual filicide: Are phenomenal zombies hurt? In: . [S.l.]: Organon F, 2022. v. 29, n. 1, p. 78–108. [3.3.5](#), [4.1](#), [112](#), [114](#), [4.1.2](#), [122](#), [4.1.3](#)
- SHOEMAKER, S. The inverted spectrum. *Journal of Philosophy*, v. 79, n. 7, p. 367–, 1982. [2.1](#)
- SILBY, B. On the conceivability of zombies. In: COPELAND, J. (Ed.). *Philosophy research paper series - Dept Philosophy, University of Canterbury*. [S.l.: s.n.], 1998. [9](#)
- TONONI, G. The integrated information theory of consciousness. In: VELMANS, M.; SCHNEIDER, S. (Ed.). *The Blackwell Companion to Consciousness*. [S.l.]: Wiley-Blackwell, 2007. p. 243–256. [3.1.3](#)
- TYE, M. Qualia. In: ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Fall 2021. [S.l.]: Metaphysics Research Lab, Stanford University, 2021. [2.1](#)