

PROGRAMA DOUTORAL EM FILOSOFIA

A Natureza da Consciência

Uma Ontologia Monista-Informacional

Ricardo Fernando Fontes Jesus Serrado

2025



Ricardo Fernando Fontes Jesus Serrado

A Natureza da Consciência Uma Ontologia Monista-Informacional

Tese realizada no âmbito do Doutoramento em filosofia, orientada pela Professora
Doutora Sofia Gabriela Assis de Morais Miguens
e pelo Professor Doutor Adelino Dias Cardoso

Faculdade de Letras da Universidade do Porto

2025

Ao meu pai, Fernando, e à minha mãe, Eduarda.
Pelo privilégio de uma vida com sentido, com valores, com afeto,
pelo caminho percorrido na luz da vossa sombra
que projeta um eu tão vosso.

Tão nosso.

Tão nosso como este trabalho que agora apresento.

À Cláudia,
Pelo sorriso do Universo.

*“Podemos pensar tanto quanto quisermos
e continuaremos a não estar lá.”*

Nicholas Humphrey

*“So, however mysterious the mind-body problem may
be of us, we should always remember that it
is a solved problem for nature.
All we have to do is figure out that solution by naturalistic means.”*

E. C. Banks

Sumário

Declaração de honra	8
Agradecimentos	9
Resumo	11
Abstract	13
Introdução	15
1. Mente e consciência: teorias dominantes	28
1.1. O que é a consciência?	31
1.2. O dualismo de propriedades	37
1.2.1. A Dra. Mary e a lacuna explicativa	40
Os problemas da lacuna explicativa	41
1.2.2. Os princípios da teoria de David Chalmers	43
1.2.3. Outros dualismos de propriedades	47
1.2.4. Os problemas do dualismo de propriedades	51
1.2.5. Evidências empíricas contra a proposta de Chalmers	56
1.3. O fisicalismo cognitivista	58
1.3.1 O espaço neuronal de trabalho (GNW)	63
O córtex pré-frontal como centro executivo da consciência	65
Uma teoria promissora	67
1.3.2 A teoria da integração da informação (IIT)	69
A validade da IIT	71
1.3.3 O fisicalismo corporificado (António Damásio)	73
Biologia, mente e consciência	74
Emoções e sentimentos	77
O sentimento de si	79
A importância dos <i>qualia</i> para a percepção consciente: provas empíricas	87
A consciência damasiana	91
O corpo como unificador da consciência	96
Virtudes da teoria de Damásio	97
2. Limites da consciência na ação	101
2.1. Determinismo, libertarismo e livre-arbítrio	105
2.1.1. O determinismo	106

2.1.2. O libertarismo	117
2.1.3. O libertarismo simples (ou não-causal)	118
2.1.4. O libertarismo de causação agencial	122
2.1.5. O libertarismo causal	125
2.1.6. O libertarismo e a física quântica	131
2.1.7. A hipótese das possibilidades alternativas contestada	135
2.1.8. O compatibilismo	136
2.2. A impossibilidade ontológica do indeterminismo: entre contradições e confusões conceptuais	139
2.2.1. O indeterminismo ontológico: um conceito filosoficamente incoerente	140
2.2.2. A impossibilidade dos eventos <i>acausados</i> : o paradoxo do libertarismo	144
2.2.3. Probabilidade e aleatoriedade: conceitos epistemológicos, não ontológicos	154
2.2.4. O mito da física quântica como justificação do libertarismo	162
2.3. Revertendo o debate: como o determinismo torna o livre-arbítrio possível	168
2.3.1. A necessidade de um universo determinista	170
2.3.2. O livre-arbítrio tem que ser um evento causado	174
2.3.3. O sujeito como uma construção	179
2.4. O meio segundo de atraso da consciência	181
2.4.1. As experiências sobre o livre-arbítrio	184
2.4.2. Objeções aos resultados Libetianos	188
2.5. A consciência como ilusão	192
2.5.1. A consciência causada	195
2.5.2. O livre-arbítrio como uma percepção	197
2.5.3. O livre-arbítrio como um sentimento	200
2.5.4. A ilusão da ação consciente	203
2.5.5. O intérprete e a narrativa da consciência	211
2.5.6. A consciência como fenómeno físico emergente	218
 3. A ontologia da Informação: mente, consciência e Universo	 221
3.1. A substância unificada informacional (SUI)	224
3.1.1. A física como epistemologia da informação: superando a ilusão da dualidade	230
3.2. Os quatro princípios da substância unificada informacional (SUI)	237
3.2.1. A informação tem de ser conservada	237
3.2.2. A informação tem de ser processada	239
3.2.3. A informação tem de ser finita	241

3.2.4. A informação tem de ser causal	242
3.2.5. SUI e causalidade emergente: o determinismo como motor da complexidade	244
3.3. Os três aspetos da SUI	248
3.3.1. A informação ontológica	248
A informação ontológica estrutural	249
A informação ontológica computacional	252
Matéria, energia e informação: a ontologia informacional da realidade	254
3.3.2. A hipótese do desdobramento informacional	255
A causalidade (ou determinismo) como atualização de estados informacionais	265
As partículas virtuais e o desdobramento informacional	267
Como o desdobramento informacional pode explicar a dualidade partícula-onda	270
Como o desdobramento informacional pode explicar a não-localidade quântica	273
3.3.3. A informação representacional	274
3.4. As interações entre a informação computacional e informação representacional	277
3.4.1. O modelo de decodificação recursiva	284
A decodificação de informação do corpo	285
O córtex pré-frontal como centro de recursividade informacional	288
3.4.2. Subconsciente computacional e subconsciente representacional	291
3.5. Uma solução para o problema do conhecimento	296
3.5.1. Três princípios para o problema do conhecimento	298
A luz como interface do conhecimento	302
3.6. Uma solução para o problema mente-corpo	308
3.6.1. A recursividade e especificidade dos neurónios	310
A consciência de primeira e de segunda ordem	318
3.6.2. O fisicalismo do problema mente-corpo	321
3.6.3 O naturalismo/fisicalismo de Espinosa	325
3.7. Uma solução para o epifenomenismo da consciência	327
3.7.1. O papel dos <i>qualia</i> na decisão	340
3.7.2. Mentes e consciência não-humanas	343
3.8. Uma analogia: o SELF	348
3.8.1. A transparência da consciência do SELF	359
 Conclusão	 369
Referências bibliográficas	379

Declaração de honra / *Declaration of Honour*

Declaro que a presente tese é de minha autoria e não foi utilizada previamente noutro curso ou unidade curricular, desta ou de outra instituição. As referências a outros autores (textos, trabalhos, ideias) respeitam escrupulosamente as regras de atribuição de autoria e encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as normas de referência. Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

I hereby declare that this thesis/dissertation/report is of my authorship and has not been used previously in another course, degree, curricular unit or subject, at this or any other institution. References to other authors (statements, ideas, thoughts) scrupulously respect the rules of attribution and are duly indicated in the text and bibliographical references, in accordance with the rules of referencing. I am aware that the practice of plagiarism and self-plagiarism is an academic offence.

Declaro, ainda, que não utilizei ferramentas de inteligência artificial generativa (chatbots baseados em grandes modelos de linguagem) para realização de parte(s) da presente tese, encontrando-se todas as interações (*prompts* e respostas) transcritas em anexo.

I further declare that I have not used generative artificial intelligence tools (chatbots based on large language models) to carry out part(s) of this thesis, and that all interactions (prompts and responses) have been transcribed in the annex.

Barcarena, 31 de março de 2025

Ricardo Serrado

Agradecimentos

Uma tese de doutoramento em Filosofia é, no meu entender, um trabalho árduo mas extremamente gratificante e estimulante. Por um lado sinto um grande alívio ao escrever estas linhas, visto estar na fase final do meu trabalho, mas, por outro, sinto já falta das horas em que me refugiava neste mar de letras, de pensamento e de ideias, nas quais me perdia, não só para encontrar algumas respostas que me inquietam, como para, de certa forma, poder exprimir a minha interioridade, muitas vezes com uma banda sonora que, também ela, me guiava para um mundo que sentia ser tão meu, mas que agora partilho.

Gostaria de começar por agradecer aos meus orientadores. A professora Sofia Miguens é uma das filósofas mais consagradas na área da filosofia da mente, com um saber extremamente vasto, e uma das pessoas mais atentas e humanas que conheci no meu percurso académico em Filosofia, pelo que foi um privilégio ser seu orientando. Quanto ao professor Adelino Cardoso, já o conheço desde os meus tempos de doutorando em História e Teoria das Ideias, em 2012. Sempre me surpreenderam o seu rigor filosófico, o seu saber mas, acima de tudo, a sua postura extremamente humana e fraterna em relação aos alunos. Foi, igualmente, um privilégio ser seu aluno, seu orientando e, acho que o posso dizer, seu amigo.

Agradeço: aos meus colegas de turma, em particular ao João Silva, que me ajudou, mais do que uma vez, a ultrapassar algumas dificuldades que fui encontrando ao longo do curso; ao meu amigo Pedro Serra, que faz questão em acompanhar o meu percurso académico; ao meu amigo Tito, por ser um confidente, um ouvinte e uma das poucas pessoas com quem conseguia levar alguns dos temas desta tese para a rua.

Gostaria, também, de agradecer a alguns guias espirituais: Espinosa, António Damásio, Thomas Metzinger, Daniel Wegner, Gerard't Hooft, Nicholas Humphrey, David Bohm, entre outros.

Obrigado a todos aqueles que, com afeto, fizeram parte de uma rede familiar que, de uma forma ou de outra, marcou a minha vida, proporcionou-me segurança e estabilidade e contribuiu para a minha formação humana: a minha madrinha Luísa, a minha madrinha Lubélia, Zé Luís e Rosa, tio Zé, primo António Pedro, tia Patrocínio, tia

Assunção, prima Rosa, primo Zé Jorge, primo Henrique, prima Carla e prima Xana, primo Zézé, avô Zé, avó Maria...

Por fim, deixo o meu mais sentido agradecimento e admiração ao meu círculo familiar: ao meu pai, meu herói; à minha mãe, meu abrigo; ao meu padrinho, meu maior amigo; aos meus avós, que me correm no sangue e que estão sempre comigo (gostava que pudessem ver onde cheguei – é modesto mas atingido com muito trabalho honesto, como sempre me ensinaram pelo vosso exemplo).

Um obrigado muito particular ao Sorriso do Universo – por teres caído no meu jardim vestida de azul fada. Com olhos feitos de estrelas trouxeste-me uma luz que me guia desde menino e que me encontrou em adulto, uma luz que quero agarrar para a eternidade, sem nunca a largar. Obrigado por fazeres parte da minha vida!

Resumo

A presente tese propõe uma solução original para o problema da consciência, sustentada numa ontologia monista-informacional, segundo a qual tudo o que existe é informação. Rejeita-se qualquer forma de dualismo — seja de substância, seja de propriedades — defendendo-se que o corpo, a mente e a consciência são expressões de uma única estrutura ontológica: a *Substância Unificada Informacional* (SUI). A consciência emerge, no nosso modelo, não como um epifenómeno, mas como um nível particular de organização da informação representacional, resultante da capacidade de certos sistemas nervosos — como o cérebro humano — realizarem complexas operações recursivas entre computação e representação, conferindo-lhes, assim, flexibilidade decisional e comportamental.

No nosso modelo, a SUI é apresentada como a base última da realidade, regida por quatro princípios: conserva informação, processa informação, é finita e é causal (determinista). Esta substância desdobra-se em três níveis: 1) informação estrutural (leis fundamentais do universo); 2) informação computacional (processamento emergente e causal); 3) informação representacional (representações mentais, conscientes ou não). A tese propõe, neste contexto, a *hipótese do desdobramento informacional*, segundo a qual a realidade emerge da informação ontológica estrutural através de processos computacionais que, em sistemas complexos, podem dar origem à representação mental. Esse desdobramento informacional ocorre de forma recursiva, permitindo que, em seres dotados de sistema nervoso, a informação computacional (sinais do corpo e do ambiente) se transforme em informação representacional — como a percepção consciente e os *qualia* — e que esta, por sua vez, seja novamente decodificada em informação computacional. Deste modo, a recursividade do desdobramento permite, em certos seres vivos, que essas representações sentidas retroalimentem o sistema sob a forma de novas computações, enriquecendo o cérebro, promovendo a autorregulação, a antecipação de estados e a construção de um modelo interno do eu — algo decisivo para decisões comportamentais mais eficientes e flexíveis. Assim, sugerimos que a consciência funciona como um sistema de decodificação informacional que o cérebro utiliza para conhecer melhor o corpo e o ambiente.

A tese aborda ainda o problema do livre-arbítrio, rejeitando o libertarismo — cujas inconsistências, nomeadamente a hipótese de eventos *acausais*, entre outras, são aqui aprofundadas — e defendendo que a liberdade só pode emergir num Universo causal, através da determinação biológica de sentimentos como a vontade, o desejo e a volição.

A partir deste quadro teórico, a proposta aqui desenvolvida apresenta um modelo segundo o qual a consciência é um fenómeno complexo e emergente de níveis informacionais mais fundamentais. Essa emergência, possibilitada pelo desdobramento informacional recursivo, permite-nos propor soluções para três problemas filosóficos centrais, que serão aprofundados na parte final da tese: o problema do conhecimento, o problema mente-corpo e o problema do epifenomenismo da consciência.

Abstract

This thesis proposes an original solution to the problem of consciousness, grounded in a monist-informational ontology according to which everything that exists is information. It rejects any form of dualism — whether of substance or of properties — arguing that the body, mind, and consciousness are different expressions of a single ontological structure: the Unified Informational Substance (UIS). In our model, consciousness does not emerge as an epiphenomenon, but as a specific level of organization of representational information, resulting from the capacity of certain nervous systems — such as the human brain — to perform complex recursive operations between computation and representation, thus granting them decisional and behavioral flexibility.

In our framework, the SUI (Ontological Informational Substance) is presented as the ultimate basis of reality, governed by four principles: it conserves information, processes information, is finite, and is causal (deterministic). This substance unfolds across three levels: (1) structural information (the fundamental laws of the universe); (2) computational information (emergent and causal processing); and (3) representational information (mental representations, whether conscious or not). The thesis, in this context, proposes the hypothesis of informational unfolding, according to which reality emerges from structural ontological information through computational processes that, in complex systems, can give rise to mental representation. This informational unfolding occurs recursively, allowing, in organisms endowed with a nervous system, computational information (signals from the body and the environment) to be transformed into representational information — such as conscious perception and qualia — which, in turn, can be decoded back into computational information. In this way, the *recursivity* of the unfolding enables, in certain living beings, these felt representations to feed back into the system as new computations, enriching the brain, promoting self-regulation, anticipation of states, and the construction of an internal model of the self — something crucial for more efficient and flexible behavioral decisions. Thus, we suggest that consciousness functions as an informational decoding system that the brain uses to better understand the body and the environment.

The thesis also addresses the problem of free will, rejecting libertarianism — whose inconsistencies, notably the hypothesis of acausal events, are here explored in depth — and arguing that freedom can only emerge in a causal universe, through the biological determination of feelings such as will, desire, and volition.

From this theoretical framework, the proposal developed here presents a model in which consciousness is a complex phenomenon emerging from more fundamental informational levels. This emergence, made possible by recursive informational unfolding, allows us to propose solutions to three central philosophical problems, which will be explored in the final part the thesis: the problem of knowledge, the mind-body problem, and the problem of the epiphenomenal nature of consciousness.

Introdução

Acordo. Abro os olhos e não vislumbro nada, desde logo porque gosto de dormir sem qualquer tipo de luz, natural ou artificial. Sinto a almofada onde dormi e sou de imediato invadido por pensamentos e sensações, não sem antes, no entanto, demorar um pouco – umas frações de segundo - a saber onde estou ou mesmo *quem* sou. Estive ausente durante algum tempo mas voltei. Com efeito, será que estive ausente? Sempre me intrigou porque a minha esposa me diz, instantes antes de ir dormir, “até amanhã”. Pergunto-lhe muitas vezes, em tom jocoso: “Mas vais dormir ou vais-te embora?!”. Ela responde-me: “Vou dormir”. “Então”, digo eu, “se não te vais embora é porque vais passar a noite junto de mim. Porquê dizer até amanhã?”.

Este simples exemplo é suficiente para demonstrar que todos nós sabemos que, quando vamos dormir, o nosso corpo e o nosso cérebro mantêm-se no mesmo sítio mas o eu ausenta-se. Por isso dizemos “até amanhã”. Porque o eu vai, de facto, “embora”, não obstante o corpo e o cérebro permanecerem onde estão. Durante a noite, o meu corpo esteve sempre ali, assim como o meu cérebro — pelo menos assim o espero, partindo do princípio de que vivemos num Universo determinista, e não indeterminista, onde eventos irregulares e instáveis poderiam ocorrer a qualquer momento. Mas, ao mesmo tempo, não estive ali porque a minha consciência – o meu eu - retirou-se durante a noite, por algumas horas.

Levanto-me. Acendo a luz e vejo as minhas pernas e o meu corpo. Vou à casa de banho e olho-me ao espelho. Vejo o meu rosto. Sou eu – vejo que estou ali, mas, sobretudo, *sinto* que estou ali. Naquele instante *sou* o meu corpo mas, ao mesmo tempo, *sou mais* do que o meu corpo. O meu corpo parece ter sido animado no momento em que acordei. Abro a torneira do lavatório e lavo a cara. *Sinto* a sensação da água fria a tocar-me no rosto numa sensação de prazerosa frescura e leve desconforto. Estou a ter uma experiência integrada – um género de filme que passa mesmo por trás dos meus olhos – sobre *como* é ser eu, uma experiência que estive ausente durante todo o tempo que estive a dormir (sem sonhos). Estou ali mas, em nenhum momento, vejo ou sei como são fabricados os *meus* pensamentos, as *minhas* sensações e os *meus* sentimentos. Não faço a mínima ideia de como sinto prazer e leve

desconforto com o frio da água matinal. Muito menos sei como é que o meu eu surgiu “do nada”. Consigo ver o meu corpo mas onde estarei eu? Como diria David Hume, eu nunca me vejo (1739/2016, p. 299).

Se considerar, como já aflorei, que eu sou o meu corpo — incluindo, evidentemente, o meu cérebro — então eu estou ali. Mas o problema é mais complexo. Enquanto dormi, o meu corpo esteve na minha cama — e, claro, também o meu cérebro. Por isso, se o meu corpo sempre esteve ali mas o meu eu não, a minha primeira intuição é de que eu *tenho* um corpo e um cérebro, mas não *sou* nem o meu corpo nem o meu cérebro.

Quem sou eu? A resposta mais simples é “eu sou eu” - o eu que está mesmo por trás dos meus olhos, a ver e a experienciar o mundo, ao comando do meu corpo. Mas o meu eu parece esquivo. Como David Hume observou com argúcia, o meu eu está ali num momento, para desaparecer no outro, sem que o consiga alguma vez apanhar (1739/2016, p. 299). O mesmo sucede com os meus pensamentos e sensações. Ora estão ali, ora deixam de estar e são substituídos por outros. Ora estou a pensar como dormi bem ou mal para depois pensar sobre o que irá ocupar o meu dia. Ora estou bem-humorado porque dormi bem, ora sou invadido por alguma ansiedade porque vou ter uma reunião importante. A mente consciente é volátil, esquiva e aparentemente imaterial. Não a vejo nem a consigo agarrar. Ao contrário do corpo – que é composto por partes materiais –, a consciência não se vê. Pode, como vimos, desaparecer sem que demos por isso. Não é composta e, precisamente por isso, julgou-se durante séculos que a mente consciente era uma substância distinta do corpo (Descartes, 1641/1996) – o que ainda é defendido por algumas, mas residuais e pouco expressivas teorias filosóficas, como veremos mais à frente.

Segundo a perspectiva cartesiana, o corpo seria uma substância física ao passo que a mente seria uma substância não-física – uma visão partilhada pela doutrina cristã, cuja influência nas representações contemporâneas é, ainda, inegável (Gilson, 2002; Dononi, 1988). A própria psicologia cognitivista e algumas escolas filosóficas modernizaram este dualismo. Psicólogos cognitivistas (como Jerry Fodor (1990 e 1998)) postulam que a mente pode ser explicada em termos de funções computacionais, mas não necessariamente redutíveis a mecanismos neurofisiológicos, enquanto muitos filósofos (como David Chalmers (1995), Thomas Nagel (1974 e 1979) ou Frank Jackson

(1984 e 1986), entre outros) consideram que a mente possui propriedades não-físicas distintas da fisicalidade do cérebro.

Os filósofos fisicalistas, por outro lado, advogam que a mente é aquilo que o cérebro faz (Dennett, 1991) ou uma propriedade física emergente do cérebro (Searle, 2019; Damásio, 2010). Esta perspectiva levanta, no entanto, um problema: como é que uma substância física (como o cérebro) pode dar origem a pensamentos e experiências (aparentemente) não-físicas (uma mente e um eu). É este problema – conhecido como problema mente-corpo - com efeito, que está na origem do dualismo platónico, cartesiano, cristão, cognitivista e de propriedades, entre outros. Salvo algumas raras exceções, como Bento Espinosa (1677/1992) ou Thomas Hobbes (1658/1978; 1655/1999), a maior parte dos filósofos ocidentais (dos quais se destacam Platão, Aristóteles, Agostinho, Tomás de Aquino, Descartes, Leibniz, Locke, entre outros) defendeu, durante séculos, que o ser humano possuía uma alma imaterial e um corpo material. Estes filósofos, no entanto, nunca resolveram um problema filosoficamente embaraçoso: como é que uma substância imaterial pode interagir com uma substância material? Como é que uma alma imaterial pode observar um objeto material? É notório, aliás, o constrangimento de Descartes ao tentar responder à princesa Elisabete da Boémia quando esta lhe perguntou, muito assertivamente, como é que uma entidade de natureza imaterial, como a alma, poderia criar algum tipo de ação numa entidade de natureza material, como o corpo (1649/1991). É o mesmo problema que os dualistas de propriedades como David Chalmers (1995) ou Thomas Nagel (1974 e 1979) têm, atualmente, para explicar como é que uma consciência emergente não-física pode interagir com um cérebro físico. Estes autores nunca resolveram este problema porque, no nosso entender, não perceberam algo que nos parece ubíquo para uma teoria do problema mente-corpo, que é o seguinte: para que duas entidades possam interagir, as mesmas não podem possuir duas substâncias distintas, nem tão pouco duas propriedades diferentes, mas apenas uma – algo que Espinosa percebeu ao propor a unicidade da Substância.

Talvez só um Universo constituído por um único tipo de substância (e não por dois) e por propriedades da mesma natureza (e não por dois tipos de propriedades) possa permitir que a mente e o corpo se relacionem. Enquanto Espinosa defendeu, em *Ética* (1677/1992), a unicidade da substância, Hobbes, em *De Corpore* (1655/1999), já

argumentava que toda a realidade, incluindo os pensamentos e sentimentos, pode ser explicada exclusivamente em termos de matéria e movimento, rejeitando qualquer noção de substâncias imateriais.

Talvez nem exista nenhum problema mente-corpo ontológico. Como postulou, a nosso ver de forma assertiva, o filósofo Ernest Charles Banks:

“So, however mysterious the mind-body problem may be for us, we should always remember that it is a solved problem for nature. All we have to do is figure out that solution by naturalistic means” (in Rovelli, 2022, p. 149).

Talvez o único problema seja a nossa *percepção* de problema, não o problema em si. Talvez a ontologia da realidade seja informacional e pertençamos todos a uma rede interconectada – uma substância fundamental indivisível que nos liga uns aos todos e, simultaneamente, a todo o Universo.

O dualismo de substância e, como veremos, também o de propriedades, são respostas que anunciam e denunciam, como veremos, uma epistemologia mas não, necessariamente, uma ontologia. São propostas que se debruçam mais sobre aquilo que *conhecemos que existe* do que sobre aquilo que *causa o que existe*. O sistema monista parece-nos muito mais coerente porque desde logo explica, ontologicamente, como é que as entidades do mundo, que são todas partes (Espinosa dirá modos) da mesma Substância podem interagir entre si. Pode até explicar alguns dos maiores mistérios atuais da física quântica – a não-localidade, a dualidade onda-partícula e as partículas virtuais, como veremos no nosso trabalho.

Dividimos este trabalho em 3 partes. Na primeira parte (1), iremos analisar as mais proeminentes teses sobre a consciência; na segunda (2), iremos estudar as principais conclusões sobre o papel da consciência na decisão, focando-nos em particular sobre o livre-arbítrio e apresentando uma proposta inovadora sobre este tema; na terceira parte (3), apresentaremos o nosso modelo ontológico do Universo e o nosso modelo sobre a natureza da consciência, bem como uma solução para o problema do conhecimento, para o problema mente-corpo e ainda para o epifenomenismo da experiência consciente.

1.

Como referido, este trabalho é sobre a natureza da consciência: o que é a consciência, como surge a consciência e qual a função da consciência. Por conseguinte, na primeira parte desta tese procederemos à análise das principais teorias científicas e filosóficas da consciência, com especial destaque para três abordagens amplamente discutidas na literatura contemporânea: a teoria do duplo aspeto da informação (*double-aspect theory of information*), proposta por David Chalmers (1996); a teoria do espaço global de trabalho neuronal (*global neuronal workspace theory* ou GNW), desenvolvida por Bernard Baars (2005) e aprofundada por Stanislas Dehaene (2014); e a teoria da informação integrada (*integrated information theory* ou IIT), formulada por Giulio Tononi (2008) e posteriormente desenvolvida por Christof Koch (2016). Para além destas, iremos também estudar 4) a teoria da consciência de António Damásio, o *sentimento de si* (2010 e 2013), que nos oferece uma perspetiva extremamente válida e, em alguns aspetos até, mais completa do que as primeiras.

Esta análise será organizada em duas categorias epistemológicas. Na primeira, abordaremos as teorias que partem de um dualismo de propriedades, com destaque para a teoria do duplo aspeto da informação, de Chalmers (1995, 1996), integrando também contribuições de outros autores que exploram esta linha de pensamento. Na segunda, focar-nos-emos em perspetivas fisicalistas, analisando as teorias GNW, IIT e o sentimento de si, articulando-as com outras abordagens fisicalistas que consideremos relevantes para o nosso enquadramento teórico.

Iremos verificar que o dualismo de propriedades oferece-nos, por um lado, muitos problemas que nos parecem inconciliáveis com as leis da Natureza e por outro, um conjunto de ideias que nos parecem ingénuas e frágeis do ponto de vista filosófico. Encontramos no fisicalismo respostas mais consistentes, como veremos. Primeiramente, todas as teorias fisicalistas oferecem-nos uma perspetiva em que a consciência obedece às leis e aos princípios da Natureza, o que é, quanto a nós, uma condição necessária para qualquer teoria filosófica que almeje ser coerente com aquilo que sabemos sobre a natureza do Universo. Como foi já referido, do ponto de vista da Natureza o problema mente-corpo está resolvido (o problema apenas parece existir na mente humana), pelo que o caminho para resolvê-lo será, no nosso entender, seguir as leis da Natureza.

Depois, as três teorias fisicalistas que aprofundaremos com mais detalhe desenvolvem ideias que nos parecem centrais no estudo da consciência e que vão, aliás, ao encontro do modelo que iremos sugerir no capítulo 3. A GNW propõe um modelo em que o córtex pré-frontal funciona como um espaço decodificador e integrador de informação do cérebro. A IIT, por seu lado, propõe uma teoria segundo a qual o cérebro funciona como um sistema de processamento de informação, no qual a integração de informação atinge um nível de complexidade suficiente para gerar processos conscientes como fenómenos emergentes. Por fim, o sentimento de si defende que a consciência emerge de sinais interoceptivos associados intimamente à homeostasia do corpo que são transformados em sentimentos que, subsequentemente, geram o sentimento de si, ou seja, uma experiência consciente. Todas estas teorias sugerem uma ideia comum: é a capacidade dos cérebros de processar e decodificar informação que permite o surgimento da consciência.

Em suma, nesta primeira parte do livro iremos tentar responder a “o que é a consciência” e “como emerge a consciência” na ótica das 4 teorias referidas. Esta primeira parte do trabalho é, portanto, um capítulo descritivo e analítico sobre algumas das principais teorias da consciência, na qual tentamos identificar algumas pistas para uma melhor compreensão da fenomenologia.

2.

Para além do problema de “o que é a consciência” e “como emerge a consciência”, existe um outro: “qual a função da consciência?”. Pode a consciência ter algum papel no âmbito das ações, das decisões e dos comportamentos humanos? Pode a consciência explicar o livre-arbítrio? Estas são algumas questões que iremos levantar e tentar responder na segunda parte do trabalho.

O problema da consciência para o livre-arbítrio é central. Desde logo porque uma decisão livre implica, necessariamente, um autor consciente com capacidade para escolher qual a decisão que pretende realizar, e não apenas para agir – e muito menos para reagir. Um ser humano reage, age e toma decisões. Mas e, por exemplo, um réptil? Um réptil reage, mas será que age? Será que toma decisões? O que pretendemos enfatizar é que a própria definição de decisão obriga a que exista um sujeito – um eu - que delibere. Ao contrário de muitos animais que parecem não agir mas reagir, como os

répteis ou os peixes, os seres humanos decidem – pelo menos temos essa intuição e esse *sentimento*. E temo-lo porque, quando decidimos, temos a experiência *subjetiva* que decidimos livremente. Decidimos livremente porque temos consciência dessa liberdade. Quando decidimos alguma ação, *sabemos* e *sentimos* que essa ação é livre. Duvidamos, por isso, que o réptil delibere alguma coisa pela simples razão que julgamos que o réptil não *sabe* nem *sente* que está a fazer coisa alguma e, precisamente por isso, limita-se a reagir. A simplicidade do seu sistema nervoso, aliás, assim o sugere.

Neste sentido, o conceito de decisão implica que uma ação seja uma escolha que é realizada voluntariamente, com o propósito de se atingir um objetivo desejado. Quando o sujeito decide iniciar uma ação, ele está a *escolher* livremente um determinado comportamento que visa determinadas consequências e objetivos. A decisão é, portanto, um conceito que obriga, necessariamente, a refletir sobre o livre-arbítrio. Isto porque apenas somos livres quando decidimos o que *desejamos*. Ora, se o sujeito é livre para decidir – e se a decisão apenas pode ser livre, caso contrário seria obrigação ou imposição - onde reside essa liberdade? Qual a natureza da liberdade humana?

Diríamos que a capacidade de *escolhermos* conscientemente as nossas decisões é o garante da nossa liberdade de ação. Se escolho, conscientemente, realizar uma determinada ação, como levantar o meu braço direito, então isso parece significar que *eu* sou livre para realizar essa ação. Dificilmente poderíamos dizer que uma ação é livre se ela for determinada pelo cérebro, pois não temos controlo sobre o nosso cérebro. Também nos parece problemático dizer que a natureza da liberdade depende da nossa mente. Muitos animais possuem mente (desde logo todos os animais com sistema nervoso central) e não é consensual afirmar que a tartaruga que tenho em minha casa decida voluntariamente iniciar ou conduzir uma ação. Onde reside, então, a nossa liberdade? Talvez a resposta seja: na consciência.

A consciência é onde o eu se congrega; é o que permite ter conhecimento dos pensamentos, dos sentimentos e das emoções; é o que possibilita pensar sobre o passado e projetar o futuro. A consciência sou eu – o protagonista da minha história. Ora, se eu sou a minha consciência e se decido conscientemente, então sou livre para decidir. Dito de outra forma, se *eu* – conscientemente – escolho o meu comportamento, isso significa que *eu* sou livre para tomar decisões sem constrangimentos sempre que

tiver consciência do que estou a fazer (é o que, *grosso modo*, defendem muitos filósofos libertários). Se, em contraste, as nossas ações forem determinadas pelo subconsciente (pelo cérebro ou pela mente inconsciente), isto significaria que não teríamos conhecimento das mesmas, pelo que não seríamos livres para escolher as nossas decisões (é o que, *grosso modo*, defendem muitos filósofos deterministas).

Esta problemática será analisada através de três escolas que se debruçam sobre o livre-arbítrio: o determinismo (Espinosa, 1677/1992; Honderich, 1988; Earman, 2006; Pereboom, 2002, 2007; Hooft, 2017); o libertarismo (Kane; 2005; Balaguer, 2004; Ginet, 2014; O'Connor, 2000); e o compatibilismo (Hobart, 1934; Dennett, 2003). Iremos verificar as propostas destas escolas e demonstrar, entre outras coisas, que: 1) o libertarismo tem incongruências significativas, nomeadamente a proposta do ser humano apenas poder ser livre se for capaz de violar as leis da Natureza; 2) o conceito canónico de livre-arbítrio deve ser redefinido, desde logo porque está muito associado à perspectiva libertária que, como dissemos, propõe um livre-arbítrio que vai contra tudo o que sabemos sobre a Natureza; 3) a única forma de existir livre-arbítrio é se este for um evento *causado* num Universo determinista (Serrado, 2025).

Depois de analisarmos o problema filosófico do livre-arbítrio, iremos analisar uma série de estudos realizados sobre essa temática. Identificaremos, igualmente, vários problemas que alguns filósofos colocam a essas experiências e verificar se, de facto, essas críticas têm, ou não, validade. Veremos que, do ponto de vista científico, os variadíssimos estudos que têm vindo a ser realizados no âmbito das neurociências por Benjamin Libet (1985 e 1999), Dan Wegner (2002) e Patrick Haggard (2005 e 2008), entre muitos outros (Soon et al, 2008; Bode et al. 2011 e 2014; Jeannerod, 2006; Banks, 2006, Eagleman, 2012) têm vindo a demonstrar, de uma forma consistente, que a consciência não toma decisões. Por isso alguns autores, como Thomas Metzinger (2009), Daniel Dennett (1991) ou Nicholas Humphrey (2023), entre outros, consideram que a consciência é uma ilusão e, subsequentemente, o livre-arbítrio não pode existir.

Em síntese, na segunda parte do nosso trabalho iremos, por um lado, analisar o problema filosófico do livre-arbítrio — no qual propomos uma nova perspectiva determinista, ao mesmo tempo que identificamos as principais incongruências das posições libertaristas — e, por outro, apresentar um conjunto de experiências científicas realizadas nesse domínio.

3.

Na última parte do trabalho, apresentaremos o nosso modelo de consciência. Para esse efeito, desenvolveremos um modelo ontológico de base monista-informacional, no qual a consciência é concebida como uma propriedade informacional-física emergente.

No nosso modelo filosófico, a ontologia da informação é o ponto de partida para compreender a estrutura fundamental da realidade, na qual a informação não é apenas um epifenómeno da matéria, mas a própria substância da existência. Neste enquadramento, surge a *substância unificada informacional* (SUI), que propõe que toda a realidade se manifesta como informação em diferentes níveis de complexidade e organização. A física (a disciplina ou área do saber), nesse contexto, não é a realidade última, mas a epistemologia da informação, ou seja, é o conjunto de descrições formais e experimentais que traduzem a forma como a informação se manifesta no espaço-tempo e interage dentro das leis naturais.

A SUI assenta em quatro princípios fundamentais. Tem de ser conservada, processada, finita e causal. É o fundamento da emergência da complexidade no Universo. Desde os primeiros momentos após o *Big Bang*, a informação estruturalmente simples evoluiu através de interações informacionais progressivamente mais complexas, dando origem à matéria, às forças fundamentais e, eventualmente, a sistemas biológicos e conscientes.

Este processo pode ser compreendido através dos três aspetos fundamentais da SUI: a *informação ontológica estrutural*, a *informação ontológica computacional* e a *informação representacional*. A informação ontológica estrutural corresponde ao código-fonte do Universo – uma singularidade que contém todas as potencialidades. Já a informação ontológica computacional refere-se ao desdobramento dessa informação fundamental, resultando na organização dos fenómenos físicos e na emergência das leis da natureza. Por fim, a informação representacional surge quando a informação computacional adquire a capacidade de representar estados e significados, um processo característico de sistemas biológicos conscientes.

Esta perspetiva leva à hipótese do *desdobramento informacional*, segundo a qual a realidade emerge da transformação contínua de informação em diferentes níveis de

manifestação. No domínio da informação ontológica, encontramos três categorias fundamentais: informação, energia e matéria. A matéria e a energia são manifestações computáveis da informação, sendo que a causalidade pode ser entendida como a atualização de estados informacionais dentro da estrutura espaço-temporal. Neste contexto, as partículas virtuais (detetadas pelo efeito Casimir (1948)) desempenham um papel crucial, pois são expressões transitórias desse *desdobramento*, refletindo a forma como a informação se atualiza no tecido do espaço-tempo. Propomos, neste quadro, que tanto a dualidade onda-partícula como a não-localidade quântica não são paradoxos ontológicos, mas manifestações distintas de um mesmo processo de desdobramento informacional.

Por seu turno, a *informação representacional*, isto é a mente e a consciência, emerge quando certos sistemas biológicos adquirem a capacidade de decodificar informação de forma recursiva. O *modelo de decodificação recursiva* que iremos propor descreve este processo, sugerindo que a mente consciente não apenas processa informação, mas também gera camadas de representação cada vez mais refinadas e integradas. Dentro desta dinâmica, a *decodificação da informação do corpo* através de processos interoceptivos desempenha um papel central, pois é através dela que sistemas biológicos constroem uma representação interna do seu próprio estado e do ambiente, permitindo a emergência de fenómenos como a percepção, a intencionalidade e a consciência.

Neste contexto, iremos sugerir que *a rede de saliência e o córtex pré-frontal são centros de recursividade informacional*. A rede de saliência desempenha um papel ubíquo na decodificação da informação do corpo ao transformar sinais químicos em imagens mentais sentidas. São representações de primeira ordem, que podemos encontrar em todos os mamíferos. O córtex frontal desempenha um papel central na *re-representação* de estados mentais ao integrar e amplificar informações num espaço global de trabalho, como defende a GNW. São representações de segunda ordem que podemos encontrar nos seres humanos.

No que concerne ao problema do conhecimento, iremos propor que o conhecimento resulta de três princípios interligados: (1) a SUI como base da interação informacional, (2) a capacidade dos sistemas nervosos de transformar informação computacional em representacional, e (3) a luz como veículo de transmissão de

informação entre os objetos e o observador. Esse modelo permite explicar como o cérebro, apesar de estar isolado no crânio, acede e interpreta diretamente o mundo externo, resolvendo a questão kantiana do conhecimento dos númenos.

No que respeita ao problema mente-corpo, sugeriremos uma solução baseada no nosso sistema informacional, argumentando que a mente não é uma entidade separada do corpo, mas um aspeto emergente da informação. A mente consciente é informação computacional decodificada em informação representacional a representar-se. A interação entre informação computacional e representacional através de um sistema de recursividade informacional nos sistemas biológicos explica como os estados mentais conscientes surgem e se tornam funcionais. Assim, a mente consciente é vista como uma interface informacional que decodifica dados computacionais do ambiente e do corpo em *percepções conscientes*, permitindo decisões adaptativas. Ainda em torno do problema mente-corpo, iremos propor que os neurónios possuem uma codificação informacional específica que permite a emergência da consciência. Ou seja, ainda não sabemos de que modo neurónios podem causar a consciência, porque ainda não tivemos acesso à maquinaria informacional que permite esse desdobramento. Proporemos que essa capacidade resulta da dualidade funcional de alguns neurónios, que não apenas computam informação (bits computacionais) como também a representam (bits representacionais), possibilitando a transformação de dados brutos (computacionais) em experiências subjetivas (representacionais).

Iremos, ainda, apresentar uma proposta para o problema do epifenomenalismo da consciência. Defendido por autores como Huxley (1874) e Bombarda (1902), o epifenomenalismo da consciência sustenta que a consciência não tem impacto direto no comportamento, sendo apenas um subproduto do funcionamento cerebral. Estudos recentes nas neurociências (Libet, 1999; Haggard, 2008; Wegner, 2002) reforçam essa visão, ao sugerirem que as decisões ocorrem antes da consciência delas. No entanto, outros autores (Chisholm, 2003; O'Connor, 2000) argumentam que a consciência tem um papel causal, podendo modular processos cerebrais.

A nossa proposta de consciência rejeita tanto o epifenomenalismo quanto a visão dualista. Defenderemos que a consciência é uma função do cérebro, responsável por transformar informação de forma recursiva — isto é, decodificar computações em representações e, recursivamente, converter informação representacional em

informação computacional, enriquecendo assim o sistema com novas possibilidades de ação e de decisão. Isso permite ao sistema nervoso integrar e codificar informações do meio e do corpo, flexibilizando o comportamento e permitindo a adaptação. Assim, a consciência não causa diretamente decisões, mas informa o cérebro para decisões mais eficientes. Os *qualia*, longe de serem irrelevantes, enriquecem a percepção e influenciam a memória e a aprendizagem, melhorando a tomada de decisão e a interação social. Por isso, somos da opinião que a experiência consciente é a manifestação inevitável da funcionalidade da consciência. Ou seja, a subjetividade tem a função de integrar, codificar e conhecer informação, precisamente porque essa experiência — essa fenomenologia — é útil e, mais importante, necessária para que o organismo desenvolva comportamentos flexíveis.

Por fim, este terceiro capítulo encerra com a analogia do SELF, que é utilizada para ilustrar a interação entre os níveis de informação e a emergência da consciência. O SELF, um sistema artificial hipotético, poderia reconhecer padrões e responder a estímulos, mas sem subjetividade. Isto é semelhante ao que ocorre em organismos com mente mas sem consciência, como répteis e peixes. A capacidade de representar o meio permite um comportamento mais flexível, ainda que automatizado. Se, no entanto, o SELF adquirisse consciência, poderia re-representar as suas próprias representações, permitindo-lhe formar memórias, antecipar cenários e tomar decisões mais sofisticadas. Assim, a consciência acrescentaria flexibilidade e complexidade à tomada de decisão. A experiência consciente permite a integração de mais bits de informação do meio e do corpo, influenciando a decisão sem, no entanto, a causar diretamente. O cérebro utiliza a consciência para se informar sobre o ambiente e sobre si próprio, flexibilizando comportamentos e decisões.

A nossa abordagem, como veremos, rejeita tanto o dualismo cartesiano, como o dualismo de propriedades, como ainda o epifenomenalismo, sugerindo, no entanto, que a consciência desempenha um papel ativo na organização e simulação de informações complexas. O nosso modelo defende que tudo o que existe é físico, rejeitando a possibilidade de entidades não-físicas. Ou seja, o nosso sistema informacional é ostensivamente fisicalista — fisicalista emergentista, para sermos rigorosos —, na medida em que considera que a informação e a fisicalidade são duas faces de uma mesma realidade, sem qualquer dissociação ontológica entre elas. Ou, melhor dizendo, a

fisicalidade é um conceito que atribuímos a um Universo regido por leis informacionais que se manifestam, aos nossos olhos, como físicas. Neste sentido, sugerimos que o Universo é um sistema fisicalista, no qual não há qualquer espaço para substâncias ou propriedades não-físicas. A ontologia fisicalista do Universo proíbe, do ponto de vista da coerência e da lógica, que um sistema fisicalista possibilite a existência de substâncias ou propriedades não-físicas.

Em suma, a nossa proposta passa pela criação de um sistema monista (*a substância informacional unificada*) no qual a substância única – a informação – permite a interação entre todas as partes desse sistema. Neste sistema, a mente consciente pode facilmente interagir com o cérebro porque ambos são feitos de informação. Precisamente por isso, o cérebro criou a mente consciente para se informar melhor, tanto sobre o corpo que tem de proteger como sobre o meio ambiente que rodeia esse mesmo corpo. A mente consciente é uma porta de entrada de bits do exterior (do meio) e do interior (do corpo) que permite ao cérebro decidir mais eficazmente. Numa palavra, a mente consciente é uma forma que o cérebro encontrou para se informar melhor sobre o meio e sobre si.

1. Mente e consciência: teorias dominantes

"Não sou nada.

Nunca serei nada.

Não posso querer ser nada."

Álvaro de Campos

"Dentro de nós há uma coisa que não tem nome,
essa coisa é o que somos."

José Saramago

Mente e consciência são dois conceitos que, à primeira vista, podem parecer ser a mesma coisa. Muitas vezes, são utilizados como sinónimos, tanto na literatura científica quanto filosófica. Durante séculos, e ainda nos dias de hoje, encontramos autores que se referem à mente sem clarificar se estão a falar de processos mentais mais amplos (como a memória ou a atenção) ou da consciência enquanto experiência subjetiva.

David Chalmers, em *Facing Up to the Problem of Consciousness* (1995), explora a consciência como a dimensão experiencial do processamento de informação, distinguindo-a dos processos funcionais da mente. Giulio Tononi, em *Consciousness as Integrated Information* (2008), e Christof Koch, em *Neural correlates of consciousness* (2016), abordam a consciência enquanto integração de informação, focando-se na experiência subjetiva. No entanto, nesses trabalhos, a distinção clara entre consciência (como *qualia*) e mente (como conjunto de processos cognitivos) pode não ser sempre explícita. A perspetiva de Thomas Nagel (1979), que se aproxima de uma visão pampsiquista, também não estabelece uma distinção clara entre mente e consciência, na medida em que trata a mente essencialmente como experiência subjetiva — ou seja, como consciência — sem diferenciar entre propriedades mentais não conscientes e estados conscientes.

É certo que a nossa intuição pode, com alguma facilidade, confundir um estado mental com um estado consciente. Isto explica-se porque apenas *temos* consciência de estados mentais e porque *estamos* conscientes desses estados mentais. Por isso, podemos facilmente confundir a nossa mente com a nossa consciência. Por exemplo, apenas sabemos que pensamos, raciocinamos e sentimos porque temos *consciência* que pensamos, raciocinamos e sentimos. A mente pode ser considerada a base da consciência, na medida em que só temos consciência daquilo que é processado e representado mentalmente. No entanto, podemos estar a pensar em algo sem estarmos conscientes do que estamos a pensar. Basta pensar no seguinte exemplo. Podemos saber (e pensar) que a Terra gira à volta do Sol — temos essa ideia, pensamo-la — mas não precisamos de estar conscientes dessa ideia para a que mesma possa estar no interior da nossa mente. Isto é, não precisamos que esta ideia esteja sempre consciente para sabermos que ela corresponde a algo que sabemos. Também sabemos que $2+2 = 4$ mas não temos de estar sempre conscientes deste pensamento para sabermos o

resultado daquela operação. Isto significa que podemos ter representações mentais conscientes e inconscientes. Ou seja, enquanto as representações mentais podem ser consideradas *todo* o conteúdo mental, a consciência é o *conhecimento* de um determinado conteúdo mental de cada vez.

Um outro exemplo é o seguinte. Um lagarto possui um sistema nervoso, pelo que terá a capacidade de representar mentalmente o seu ambiente externo e interno. Neste sentido, um lagarto terá uma mente. Uma mente não muito sofisticada mas ainda assim, uma mente capaz de representar o ambiente em que se insere. Os seus órgãos sensoriais (como os olhos) mapeiam o meio ambiente e permitem, com certeza, cartografar mentalmente o mundo, o que é muito útil para a sobrevivência do lagarto. Mas será que o lagarto está consciente dessas representações? Será que possui alguma consciência *qualitativa* sobre as suas experiências? É bem provável que o lagarto possua uma mente – porque o seu sistema nervoso consegue representar o meio – mas não possua uma consciência, a capacidade de desenvolver, diríamos mesmo de *representar*, uma experiência subjetiva (Humphrey, 2023). Por conseguinte, importa dizer que os nossos conteúdos mentais não têm que ser sempre conscientes porque a mente e a consciência *não são* a mesma coisa. A menos que se lhe acrescente um eu, a mente permanece inconsciente (Damásio, 2010, 27).

Neste sentido, como podemos distinguir mente de consciência? No que concerne a este assunto, a explicação de António Damásio parece-nos muito assertiva ao distinguir dois processos que devem ser bem definidos. Na sua obra *O Livro da Consciência* (2010), Damásio define mente como a capacidade do cérebro em gerar imagens através de uma extensa e complexa cartografia neuronal. Todos os animais com sistemas nervosos terão uma mente. Já a consciência é o processo que permite *conhecer* os conteúdos desses mapas cartográficos, isto é, é o processo que permite a entrada de um protagonista no processo mental. Nas palavras do neurocientista: “A consciência não se resume a imagens mentais. Terá, no mínimo, a ver com a organização de conteúdos mentais centrada no organismo que produz e motiva esses conteúdos” (2010, p. 27). Por conseguinte, a consciência é a capacidade que um determinado sistema nervoso possui para introduzir um sujeito – um dono – às imagens mentais.

Ao longo deste trabalho, utilizaremos o termo mente para designar o conjunto das representações mentais, enquanto reservamos o termo consciência para referir as

representações centradas na identidade e na subjetividade — ou seja, no eu — do organismo que as produz.

1.1. O que é a consciência?

Como é que temos experiência subjetiva, raciocinamos e sentimos é algo que, de facto, tem intrigado pensadores de várias áreas e ocupado as mentes de inúmeros filósofos ao longo do tempo. Da filosofia clássica chegam-nos algumas das primeiras ideias sobre este problema. Sócrates defendia que a mente (ou alma) era o centro da racionalidade e virtude, sendo distinta do corpo e imortal (Brickhouse e Smith, 2010). Platão desenvolveu as ideias de Sócrates e apresentou uma visão dualista da mente e do corpo. Segundo Platão, a alma era imortal e pré-existente, adquirindo conhecimento no mundo das ideias antes de encarnar no corpo físico (Barney, 2012). Por seu turno, Aristóteles rejeitou o dualismo de Platão e apresentou uma visão funcionalista da mente, ainda assim dualista, em *De Anima* (Nussbaum e Rorty, 1995).

Entre a filosofia clássica e os contributos de René Descartes no século XVII, o problema da mente continuou a ser discutido, particularmente no pensamento medieval e no âmbito do cristianismo, que promoveu uma perspetiva claramente dualista (Gilson, 2002; Matthews, 2005; Kenny, 2010). Os filósofos cristãos, como Agostinho e Tomás de Aquino foram, aliás, fundamentais para manter o dualismo no pensamento ocidental. Agostinho, influenciado pelo platonismo, defendeu que a alma era imortal, divina e o elemento mais importante na relação do homem com Deus (Matthews, 2005). Já Tomás de Aquino integrou a filosofia aristotélica com a teologia cristã, propondo que a alma era a "forma" do corpo, mas mantendo uma visão dualista onde a alma tinha uma existência separada após a morte. Para Tomás de Aquino, a mente racional (ou alma racional) era o principal atributo que distinguia os humanos dos outros seres vivos, sendo essencial tanto para a experiência sensível quanto para o raciocínio abstrato (Kenny, 2010).

Alguns séculos depois, René Descartes trouxe uma transformação significativa ao problema mente-corpo. Descartes postulou que a mente (*res cogitans*) e o corpo (*res extensa*) eram substâncias separadas e distintas. A mente, segundo o filósofo nascido

em França, era imaterial, responsável pelo pensamento e pela experiência subjetiva, enquanto o corpo era material e sujeito às leis físicas (Descartes, 1641/1996; 1637/2008). Esta separação entre mente e corpo foi sintetizada na célebre máxima "*Cogito, ergo sum*"¹ (Descartes, 1637/2008, p. 52), que indicava a primazia da mente como certeza fundamental da existência. Descartes acreditava que a interação entre mente e corpo ocorria na glândula pineal (Descartes, 1649/2008, p. 102 e 103), embora esta explicação tenha sido alvo de críticas tanto no seu tempo quanto posteriormente.

Após Descartes, outros pensadores continuaram a explorar o dualismo, ampliando ou contestando algumas ideias mas sempre dentro de um escopo dualista. Nicolas Malebranchec desenvolveu o ocasionalismo, uma teoria que propunha que a interação entre mente e corpo não era direta. Em vez disso, Deus seria o intermediário que harmonizava as operações da mente com as do corpo (Robinson, 2011). Gottfried Leibniz elaborou o conceito de harmonia pré-estabelecida, sugerindo que mente e corpo eram como dois relógios perfeitamente sincronizados por Deus. Esta visão era, no entanto, claramente dualista, já que rejeitava qualquer interação direta entre mente e corpo, mas afirmava que ambos operavam em perfeita correspondência (Kulstad, 1997). John Locke, embora frequentemente associado ao empirismo, manteve algumas ideias que se relacionam com o dualismo, sugerindo que a alma era uma "tábua rasa" (*tabula rasa*) na qual as experiências sensoriais escrevem, mas deixando aberta a possibilidade de que Deus pudesse unir mente e matéria de formas além da compreensão humana (Chappell, 1994).

Durante séculos, o dualismo foi a perspectiva filosófica canonicamente aceita, o que se compreende tendo em conta as suas afinidades com a teologia cristã (Gilson, 2002). Neste período, dois homens confrontaram o paradigma dualista: Thomas Hobbes e Bento Espinosa. Hobbes defendeu uma abordagem materialista e mecanicista da mente. Ele argumentou que tudo, incluindo a mente, é composto por matéria, e que os processos mentais podem ser explicados em termos de movimento e mecânica natural. Para Hobbes, a consciência é o resultado de movimentos no cérebro causados por estímulos externos e o pensamento é uma sequência de associações causais entre ideias, comparável a um discurso mental (1658/1978; 1655/1999). Para Espinosa, a

¹ "Penso, logo existo" (Descartes, 1637/2008, p. 52).

realidade é composta por uma única Substância, que ele identificou como Deus ou Natureza (*Deus sive Natura*). A Substância espinosana estava subjacente a todo o Universo, pelo que a mente e corpo, nesta perspetiva, eram dois atributos desta Substância, expressando-se de formas diferentes. Esta visão monista elimina a necessidade de interação causal entre mente e corpo, tratando-os como paralelos (1677/1992).

Apesar das ideias monistas e naturalistas de Espinosa e Hobbes, o dualismo de substância foi a filosofia dominante até meados do século XIX – período em que o dualismo começou a ser criticado de uma forma mais ampla, sobretudo por autores ligados à medicina e às ciências naturais, como Karl Vogt, Thomas Henry Huxley, William James ou Miguel Bombarda. Segundo Vogt, todos os processos mentais são produtos da atividade física do cérebro, rejeitando qualquer dualismo entre mente e matéria. O pensamento não seria nada mais do que um subproduto do cérebro, assim como a bília é um produto da função hepática (Gregory, 1977). No mesmo sentido, Huxley (1874) defendia que todas as funções vitais, incluindo os processos mentais, podem ser explicadas em termos de causas físicas e biológicas. Ele rejeitava qualquer tipo de força vitalista ou explicação sobrenatural para a consciência, argumentando que o cérebro é o único responsável pelos pensamentos e comportamentos humanos. Esta posição estava alinhada com a ciência mecanicista do seu tempo, que procurava explicar todos os fenómenos naturais através de leis físicas (Serrado, 2021). Sensivelmente a mesma opinião tinha Miguel Bombarda, para quem a consciência, o pensamento e qualquer processo mental eram um produto do cérebro (1902; Serrado, 2024b). William James tinha uma perspetiva diferente, já que recusava explicações dualistas e materialistas. Em vez disso, via a mente como um processo dinâmico e adaptativo, ligado ao cérebro, mas com um papel próprio na experiência e na tomada de decisões (1890).

Embora a consciência seja algo intrínseco à experiência de qualquer ser humano (Searle, 2019), ao mesmo tempo parece difícil explicar, objetivamente, no que é que consiste. O próprio conceito de consciência pode ser difuso e ambíguo. Por isso, a consciência foi considerada pelos cientistas durante algum tempo demasiado esquiva e etérea – uma identidade sem características métricas passíveis de serem estudadas – pelo que o seu estudo foi evitado e, até, profundamente ostracizado durante

praticamente todo o século XX². Ou seja, não obstante alguns cientistas, como os acima referidos, terem dedicado atenção significativa à consciência no século XIX e inícios do século XX, a verdade é que durante uma boa parte do último século a consciência foi um tema erradicado do panorama científico. Devido a esta tendência, o behaviorismo (que dominou a psicologia durante a primeira parte do século XX) eliminou a mente consciente do seu campo de estudo, limitando-se a estudar o comportamento humano em função de processos de estímulo e resposta (Braunstein e Pewzner, 2003).

Atualmente, o panorama filosófico-científico é muito diferente. Não só a consciência é, sobretudo desde as décadas de 80 e 90, um objeto de estudo para inúmeros filósofos (Dennett, 1991; Churchland, 1985; Churchland, 1998; Metzinger, 2009; Searle, 1998; Chalmers, 1996; O'Connor, 2000) e neurocientistas (Baars, 1988 e 1997; Crick, 1998; Damásio, 2010; Tononi, 2004; Koch, 2016; Seth, 2021; Dehaene, 2014; Panksepp, 2005; Gazzaniga, 2011), como a sua definição parece ser suficientemente unânime.

Já vimos que a consciência pode ser definida, numa palavra, como o *eu*. Mas podemos defini-la, também, *como aquilo que permite a um determinado sujeito ter uma experiência subjetiva*. A consciência é, portanto, a experiência *qualitativa* da realidade – os filósofos chamam-lhe, também, fenomenologia ou *qualia* (Chalmers, 1995 e 1996). A consciência é, portanto, aquilo que nos permite sermos protagonistas dos conteúdos da nossa mente – ter uma identidade, um *eu* (ou um *self*³) e, subsequentemente, ter uma experiência fenomenológica (Damásio, 2010; Humphrey, 2023; Seth, 2021; Dennett, 1991).

É certo que alguns autores consideram que esse *eu* não existe – é uma ilusão (Dennett, 2015; Churchland, 1998, 1981 e 1985; Churchland, 1983 e 2015; Metzinger, 2009). Não uma ilusão no sentido de não existir, mas uma ilusão no sentido de *não* ser aquilo que julgamos que é. Os filósofos Patricia e Paul Churchland, bem como Daniel Dennett ou Thomas Metzinger, podem ser considerados filósofos eliminativos, mas nenhum defende que a consciência deve ser eliminada, por não existir. O que defendem,

2 Thomas Nagel (1974), David Chalmers (1995, 1996) e Frank Jackson (1984, 86), por exemplo, embora sejam autores que já refletiram e publicaram trabalhos pertinentes sobre a consciência, não creem que a mesma possa ser explicada do ponto de vista científico.

³ Podemos utilizar estes conceitos indiscriminadamente durante o trabalho.

com algumas diferenças entre si, é que *a consciência não é aquilo que parece*. Paul Churchland, por exemplo, um dos principais defensores do materialismo eliminativo, argumenta que o entendimento popular da consciência é limitado e deve ser substituído por uma concepção mais avançada baseada em neurociência. Ele não nega a existência da consciência, mas propõe que compreendê-la exige abandonar pressupostos intuitivos e adotar um novo paradigma científico (1984). Segundo Churchland:

“And when neuroscience has matured to the point where the poverty of our current conceptions is apparent to everyone, and the superiority of the new framework is established, we shall then be able to set about reconceiving our internal states and activities, within a truly adequate conceptual framework at last” (1984, p. 2).

Assim, Churchland não recusa que tenhamos estados internos, nem tão pouco sugere que a consciência não existe. O que ele postula é que o nosso entendimento sobre a consciência é superficial e apoiado num aparato pouco científico. Por isso, deve ser eliminado e substituído por outro mais avançado e rigoroso.

Thomas Metzinger, no seu livro *The Ego Tunnel* (2009), descreve a consciência como uma ilusão evolutiva criada pelo cérebro, o que ele denomina “túnel do ego” (*Ego Tunnel*). Para Metzinger, a consciência é uma construção representacional sofisticada que organiza informações de forma eficiente, mas não há um eu real no sentido ontológico. Metzinger diz-nos que:

“It must be emphasized that although our brains create the Ego Tunnel [consciousness], no one lives in this tunnel. We live with it and through it, but there is no little man running things inside our head- The Ego and the Tunnel [or our consciousness] are evolved representational phenomena, a result of dynamical self organization on many levels.” (2009, p. 8).

Por seu turno, Daniel Dennett defende na sua obra *Consciousness Explained* (1991) que a consciência é uma espécie de “ilusão útil”, rejeitando qualquer noção de qualia ou de um “teatro cartesiano” onde a mente reuniria percepções para um observador interno. Para Dennett, a consciência não é uma entidade unificada nem um fenómeno centralizado, mas sim um efeito emergente da atividade distribuída do cérebro, resultado de múltiplos processos computacionais inconscientes que operam em paralelo (Miguens, 2002).

Apesar de existir, atualmente, um consenso relativo sobre a consciência ser a subjetividade associada ao eu (Dennett, 1991; Churchland, 1985; Churchland, 1998; Metzinger, 2009; Searle, 1998; Chalmers, 1996; O'Connor, 2000; Crick, 1998; Damásio, 2010; Tononi, 2004; Koch, 2016; Seth, 2021; Dehaene, 2014; Panksepp, 2005; Gazzaniga, 2011), a *qualidade* da experiência que a torna individual ainda é um mistério. Como é que o cérebro, que é um órgão composto por átomos e moléculas, pode dar origem à experiência *qualitativa* de vermelho, de dor ou de doce?

O cérebro é um objeto. Pode ser visto, cortado, mexido e remexido. Mas a consciência não. Ninguém pode ver a minha consciência (ou qualquer outra consciência), muito menos mexer e remexer nas minhas experiências de vermelho, de dor e de doce. Se abrirmos o meu crânio, o meu cérebro está ali. Mas onde está a consciência? Se analisarmos com uma máquina de ressonância magnética funcional o meu cérebro a ver vermelho, a sentir dor ou a saborear doce, conseguimos identificar padrões neuronais que podem indicar que estou conscientemente a ver vermelho, a sentir dor ou a ter prazer enquanto degusto algum alimento. Mas onde está a experiência de vermelho, de dor e de prazer? Apenas eu – e mais ninguém – experiencio o vermelho, a dor e o prazer. Nenhuma pessoa que analise o meu cérebro a ter *essas* experiências fenomenológicas poderá ter *essas* experiências. Por conseguinte, o cérebro físico pode ser observado mas a experiência fenomenológica não, e, precisamente por isso, muitos defendem que essa experiência é não-física. O problema coloca-se, portanto, desta forma: como é que um objeto físico – o cérebro – pode causar uma experiência fenomenológica, aparentemente não-física?

Para o filósofo John Searle, não há nenhum mistério, porque à pergunta “Como pode o cérebro ser consciente?” Searle limita-se a responder: “Limita-se a sê-lo” (2019, p. 51), da mesma forma que o estômago limita-se a processar a digestão. John Searle, em *Da Realidade Física à Realidade Humana* (2019), adota uma posição naturalista, sugerindo que a consciência é simplesmente uma função biológica do cérebro, comparável à digestão no estômago. Para ele, o mistério da consciência resulta mais de uma confusão conceptual do que de uma limitação ontológica. Apesar de muitos filósofos e neurocientistas concordarem com esta ideia – de que não há nenhum mistério transcendente no problema mente-cérebro (Damásio, 2010; Gazzaniga, 2011; Churchland, 1985; Humphrey, 2023; Eagleman, 2012; Dennett, 2015; Metzinger, 2009;

Seth, 2021) – também é verdade que ainda não há respostas conclusivas para responder: 1) como é que existe consciência; 2) e como é que o cérebro causa a consciência.

Com efeito, a neurociência tem vindo a desvendar alguns problemas relacionados com o problema mente-corpo, nomeadamente ao nível da tomada de decisão, da aprendizagem, da memória, entre muitos outros, mas ainda se levantam questões sobre *como* e *porque* o sujeito *experiencia* qualidades subjetivas⁴. Chalmers refere-se a isto como o *problema difícil da consciência* (Chalmers, 1996). Deparamo-nos, então, ainda com três questões metafísicas para as quais ainda não existem respostas definitivas: 1) o que é a consciência? 2) como pode existir consciência? 3) e porque existe consciência? (Van Gulick, 2012).

1.2. O dualismo de propriedades

O dualismo de propriedades é a escola filosófica que defende que o Universo é constituído por uma única substância que possui duas categorias fundamentais de propriedades: propriedades físicas (como massa e carga) e propriedades mentais ou fenomenológicas (como consciência ou *qualia*) (Chalmers, 1995 e 1996; O'Connor, 2000a). Ou seja, o dualismo de propriedades postula que, apesar do Universo ser constituído por matéria, os estados fenomenológicos correspondem a um aspeto dual da matéria – das estruturas físicas (Chalmers, 1995, 1996; Nagel, 1979; O'Connor, 2000a). É uma tese com múltiplos seguidores, que obteve amplo reconhecimento no domínio filosófico (O'Connor, 2000; Chisholm, 2003), muito mais do que na comunidade científica, onde o dualismo de propriedades tende a ser rejeitado (Damásio, 2010; Humphrey, 2023; Edelman, 1995, Baars, 2005; Crick, 1998; Dehaene, 2014; Llinás e Ribary, 1993; Nani et al. 2019; Pinker, 1999; Bohm, 1980).

4 Sobre este assunto, concordamos com Searle, para quem não restam dúvidas “que a consciência é causada por processos neuronais” (2019, p. 50). Julgamos que a maior parte dos filósofos e cientistas defende esta ideia. A nosso ver, o problema centra-se em *como* é que a consciência é causada por processos neuronais.

Muito provavelmente devido ao avanço das neurociências, mas também ao surgimento de uma escola filosófica fisicalista que se apoiou muito nas conclusões científicas relacionadas com o problema mente-corpo (Rosenthal, 2005; Dennett, 1991; Metzinger, 2009; Searle, 2019; Churchland, 1998), o debate sobre se a consciência depende de uma substância imaterial parece ter-se tornado obsoleto nas últimas décadas, pelo que o dualismo de substância, embora defendido ainda por alguns autores (Foster, 1989; Swinburne, 1986), acabou por ser, pelo menos em grande parte, erradicado da discussão filosófico-científica sobre a natureza da consciência (Damásio, 2010; Crick, 1998; Chalmers, 1995; Dennett, 2005; Churchland, 1998; Rosenthal, 2005). No entanto, como vimos, embora aceitando uma perspetiva fisicalista do Universo, alguns teóricos questionam se a consciência poderá ter origem, não numa *substância* distinta da matéria como defendem os dualistas de substância, mas numa *propriedade* diferente da matéria.

No entender do filósofo David Chalmers, existem dois problemas centrais relacionados com a consciência: *the easy problems* e *the hard problem*. Quanto aos primeiros, os problemas têm a ver com compreender a atenção, a reação aos estímulos do ambiente, a aprendizagem, a diferença entre estar acordado e a dormir, entre muitas outras faculdades cognitivas (1995 e 1996). Para o filósofo, todas as características da mente que sejam funcionais (aprendizagem, ler, falar, observar, a atenção, a saliência, etc.) podem ser explicadas cientificamente, de uma forma materialista-reducionista, como de resto a neurociência tem vindo a fazer com sucesso. Como estas faculdades têm, sobretudo, a ver com funções cognitivas, poderíamos designar este tipo de “consciência” como consciência de acesso (Block, 1995). Por ser um tipo de consciência cognitiva é mais fácil de explicar, precisamente, porque está relacionada com a *computacionalidade* de determinadas faculdades cognitivas.

O outro tipo de consciência – que podemos designar por fenoménica (Block, 1995) – segundo Chalmers, é algo muito mais difícil de compreender, de estudar e de explicar, já que está intrinsecamente relacionado com a subjetividade, isto é, com a experiência interior e privada de cada um, o que extravasa a *computacionalidade* das faculdades cognitivas relacionadas com os *easy problems*. Nas palavras de Chalmers: “What makes the hard problem hard and almost unique is that it goes beyond problems about the performance of functions” (1995, 203). Numa palavra, a consciência, para

Chalmers, não possui qualquer função, pelo que não pode ser explicada sob um ponto de vista funcionalista ou reducionista. Segundo esta perspectiva, seria portanto possível um ser desenvolver *funções cognitivas* sem experiência fenomenológica porque a fenomenologia é, segundo Chalmers, algo que *acompanha* as funções cognitivas, não necessariamente algo que as *cause* ou que seja *causada* por elas. Estar consciente, para Chalmers, é experienciar a subjetividade, pelo que as funções cognitivas – *os easy problems* – ao não serem acompanhadas por subjetividade não têm nada a ver com a consciência, com o *hard problem*.

Se, como advoga Chalmers em *The Conscious Mind* (1996), a fenomenologia não está relacionada com as funções cognitivas, seria possível que um indivíduo realizasse todas as faculdades mentais associadas aos *easy problems* sem estar minimamente sensiente do que está a fazer, tal como um *zombie*. Um *zombie* seria, segundo esta perspectiva, um ser fisicamente idêntico ao ser humano em todos os aspetos, incluindo o comportamento, o funcionamento do corpo e as atividades cerebrais. Contudo, esse *zombie* não possuiria qualquer experiência subjetiva: toda a sua realidade interior estaria às escuras (Morch, 2023). Assim, o *zombie* poderia realizar todas as tarefas que um ser sensiente realiza – como falar, reagir a estímulos, raciocinar, perceber objetos, etc. –, mas sem qualquer fenomenologia associada.

A mesma opinião tem o filósofo Frank Jackson, para quem poderiam existir pessoas a realizar as mesmas funções que os seres humanos e que se comportassem exatamente como eles, sem que estivessem conscientes das suas experiências (Jackson, 1982). Ou seja, um *zombie* poderia desenvolver comportamentos humanos, como escrever, ler e falar, sem ter consciência fenoménica do que estava a escrever, a ler ou a falar. Para Chalmers, se podemos conceber um ser assim, é perfeitamente concebível que um *zombie* possa existir. E se um *zombie* pode existir, então o fisicalismo não pode explicar a fenomenologia, já que esse ser seria *fisicamente* igual a um ser humano. Dito por outras palavras, tanto para Chalmers como para Jackson, se podemos conceber um *zombie* – nomeadamente que um *zombie* pode possuir exatamente as mesmas características físicas e exibir os mesmos comportamentos de um sujeito consciente – então é porque o fisicalismo é falso (1996). Este cenário ilustra a separação proposta por Chalmers entre a consciência fenomenológica (*hard problem*) e as funções

cognitivas (*easy problems*), reforçando a tese de que as propriedades mentais subjetivas são distintas das propriedades físicas.

Por isso, estar consciente é “como ser alguma coisa” (“What it is like to be a bat”, na expressão de Thomas Nagel (1974)); é existir e agir *experienciando* a existência e a ação. Segundo estes filósofos, a neurociência poderá explicar a existência do sujeito e a ação do sujeito, mas não a *experiência* da existência nem a *experiência* da ação do sujeito. Porquê? De que forma podemos estudar objetivamente a subjetividade – a fenomenologia da experiência? De que modo podemos estudar *como é* ser um morcego? Parece-nos, de facto, impossível. Mesmo que um dia se saiba tudo sobre o cérebro e a consciência (de um ser humano ou até de um morcego), os neurocientistas terão, provavelmente, dificuldades em *experienciar* a *experiência* subjetiva de como é ser alguma coisa, a não ser através da observação, de modelos científicos e de uma linguagem descritiva, bem como de padrões de mapas neuronais que ficarão sempre aquém da interioridade experimentada pelo sujeito (Chalmers, 1995; Jackson, 1982, 1986)⁵.

Segundo os dualistas de propriedades existe, portanto, uma *lacuna explicativa* (Levine, 1983; Jackson, 1986) entre aquilo que é objetivo (físico) e subjetivo (fenomenológico). Esta lacuna explicativa tem sido um dos argumentos mais fortes contra o fisicalismo da consciência (Levine, 1983; Jackson, 1982; Nagel, 1974), visto questionar como é que um sistema físico como um cérebro (que é um objeto físico) pode dar origem a algo que tem experiência (que não se vê nem se localiza, pelo que não pode ser físico).

1.2.1. A Dra. Mary e a lacuna explicativa

A lacuna explicativa, descrita inicialmente pelo filósofo Joseph Levine (1983) e perfilhada por Frank Jackson (1986), refere-se à dificuldade em explicar como os processos físicos do cérebro dão origem à experiência subjetiva. Essa lacuna surge porque, mesmo que compreendamos detalhadamente os mecanismos cerebrais responsáveis por funções cognitivas – como a perceção, a memória ou o processamento

⁵ Ramachandran propõe um modelo interessante que refuta esta ideia, ao sugerir um mecanismo no qual um determinado sistema pudesse transmitir o que outro ser poderia estar a experienciar (Ramachandran e Hirstein, 1997).

sensorial –, continua a faltar uma explicação para o *porquê* e o *como* essas funções são acompanhadas por experiências fenomenológicas (*qualia*).

Para expor as fragilidades do fisicalismo e ilustrar a lacuna explicativa, o filósofo Frank Jackson elaborou uma experiência mental, com implicações semelhantes às do *zombie* filosófico proposto por David Chalmers. No artigo *Epiphenomenal Qualia*, publicado em 1982, Jackson imaginou um sujeito – de nome Mary – que sabia tudo o que havia para saber sobre a ciência da visão e das cores (poderia ser uma neurocientista especializada na visão), mas que vivera toda a sua vida num quarto a preto e branco com uma televisão a preto e branco na qual tinha acesso à realidade. Mary nunca vira uma cor em toda a sua vida, para além do preto e do branco. A questão que Jackson coloca é: será que Mary aprenderia algo de novo quando saísse desse quarto e experienciasse todas as outras cores de que tinha sido privada toda a sua vida? Se a resposta fosse positiva – que Mary tivesse aprendido algo de novo quando saísse da sua realidade a preto e branco e experienciasse cores pela primeira vez – então o fisicalismo teria que ser falso, já que o fisicalismo não conseguia explicar a natureza da experiência do sujeito (1982).

O filósofo John Foster, em *A Defense of Dualism* (1989), defendeu uma ideia em tudo semelhante quando questionou se um cego conseguiria compreender o que seria a experiência visual se estivesse na posse de todas as explicações do mundo físico e todo o conhecimento técnico sobre a visão (1989). Tanto Jackson como Foster utilizam estes exemplos para advogar que o fisicalismo não consegue captar toda a realidade dos *qualia*, mesmo que possa explicar, saber e decifrar todos os aspetos físicos que lhe podem estar inerentes (1986). O fisicalismo, segundo este ponto de vista, deixa sempre algo de fora – a experiência qualitativa de uma determinada perceção mental (Levine, 1983). Em poucas palavras, a lacuna explicativa não é contornada pelo fisicalismo pela simples razão que o fisicalismo não explica a fenomenologia - não explica a experiência subjetiva que Mary adquiriu quando saiu do seu quarto a preto e branco.

Os problemas da lacuna explicativa

No nosso entendimento, a lacuna explicativa – tal como as restantes ideias centrais do dualismo de propriedades – é um problema epistemológico, não ontológico.

No que concerne ao problema proposto por Nagel, em rigor, nunca poderemos saber como é ser um morcego, pois não possuímos os sistemas físicos que lhe permitem ter a sua própria experiência subjetiva. Da mesma forma, não sabemos como é ser outra pessoa porque não temos o seu cérebro, mas isso não implica que a subjetividade dessa pessoa tenha uma origem não-física. Como iremos defender, a ontologia da consciência (seja de um morcego ou de qualquer outro ser) é física e informacional.

Podemos nunca experienciar diretamente como é ser um morcego, mas isso não invalida que o cérebro seja a *causa* da experiência fenomenológica. A ontologia da consciência pode ser física (e iremos defender que só pode ser física), mesmo que um observador não possa aceder diretamente à subjetividade do outro. O problema é epistemológico – a dificuldade em *conhecer* os estados internos subjetivos de outro ser –, e não ontológico. Só poderíamos saber como é ser um morcego se tivéssemos acesso às suas estruturas físicas e processássemos a informação da mesma forma que ele. Veremos isto mais à frente em maior detalhe.

Relativamente à experiência da Dra. Mary, Jackson classificou-a, justamente, como *uma teoria do conhecimento* (1986), destacando que o verdadeiro problema reside na compreensão da experiência fenomenológica (como é ver cores), e não na explicação das causas da consciência. Deste modo, o dualismo de propriedades propõe um modelo para abordar o problema mente-corpo do ponto de vista do conhecimento, mas não oferece uma solução ontológica. Assim, a lacuna explicativa deve ser entendida como uma limitação epistemológica, e não como uma questão ontológica.

Como veremos, os argumentos a favor da lacuna explicativa podem ser falaciosos. A subjetividade de Mary e de um cego não depende de propriedades não-físicas, mas sim de condições *físicas* específicas. Mary, ao nunca ter tido contacto com cores, não terá desenvolvido as ligações sinápticas necessárias para representar cores. Da mesma forma, o cego não possui a estrutura neuronal que permite *criar* qualquer representação visual. A ausência de experiência fenomenológica é, assim, explicada pela falta de *sistemas físicos* adequados, e não pela existência de propriedades não-físicas. Em última análise, a lacuna explicativa reflete uma limitação do funcionamento cerebral nesses casos, e não uma falha ontológica.

Os dualistas de propriedades, contudo, tratam a lacuna explicativa como se fosse uma questão ontológica, argumentando que aquilo que *sabemos* deveria ser

equivalente àquilo que experienciamos. Esta abordagem ignora que a lacuna explicativa pode ser exclusivamente epistêmica. Ontologicamente, pode não haver lacuna alguma: o cérebro pode causar a consciência sem que sejamos capazes de compreender totalmente os meandros desse processo ou de aceder à subjetividade de outra pessoa ou animal. Por insistirem que a lacuna é ontológica, os dualistas de propriedades acabam por propor teorias extremas e pouco verosímeis à luz do conhecimento que temos sobre a Natureza. David Chalmers, por exemplo, defende que, se o fisicalismo não consegue compreender a consciência, então será necessário adotar uma teoria radical que considere a consciência um princípio fundamental do Universo, irreduzível ao cérebro ou a qualquer outra estrutura física. No entanto, este tipo de abordagem parece ignorar que o verdadeiro problema reside no âmbito do conhecimento. A lacuna explicativa tenta abordar epistemologicamente questões que são, na realidade, ontológicas. Isso leva os dualistas de propriedades a propor respostas ontológicas para questões que dizem respeito ao acesso cognitivo e à experiência subjetiva.

Neste sentido, defendemos que o que precisamos é de uma tese ontológica que compreenda a lacuna explicativa como uma limitação cognitiva, e não como uma verdade ontológica. A consciência pode ser totalmente explicada em termos físicos, mas as nossas capacidades epistemológicas podem não ser suficientes para captar a totalidade da experiência subjetiva. A lacuna explicativa, assim entendida, não é uma prova contra o fisicalismo, mas uma limitação inerente ao nosso aparato cognitivo.

1.2.2. *Os princípios da teoria de David Chalmers*

Como referido, ao não conseguir ultrapassar a lacuna explicativa, Chalmers desenvolve uma teoria radical na qual defende que a consciência, como não pode ser reduzida a sistemas físicos, é uma propriedade fundamental do Universo. Apesar de quase todos os fenómenos conhecidos poderem ser explicados inteiramente como entidades físicas passíveis de serem reduzidas a entidades mais simples, existem propriedades que não podem ser reduzidas porque são princípios fundamentais do Universo. A massa, a gravidade e o espaço-tempo, por exemplo, são princípios fundamentais do Universo e não podem ser reduzidas a nada mais elementar que elas próprias. São, para além do mais, princípios que estão em todo o lado no Universo, devido à sua *fundamentalidade*. Deste modo, dado que não se afigura a possibilidade

de a ciência conseguir explicar a experiência subjetiva em termos de função⁶, como referimos, Chalmers sugere uma teoria radical em que a consciência possa ser um princípio fundamental do Universo, como o espaço-tempo. Isto explicaria porque a lacuna explicativa existe. A consciência, como não pode ser reduzida ao físico, existe como fundamento do Universo. Por isso, não é necessário – nem tão pouco possível – reduzi-la a nada, nem ao cérebro. A lacuna é inevitável, não por falta de conhecimento, mas porque não consegue unir duas propriedades diferentes.

Um sistema psicofísico, no qual a consciência fosse uma propriedade fundamental, não interferiria com as leis da física pois, pelo contrário, a fenomenologia poderia relacionar-se com a física e, simultaneamente, seria explicado porque todas as ações humanas são acompanhadas por experiência (1995, p. 210). Os seres humanos teriam experiência do mundo porque a sensiência é uma propriedade fundamental do Universo, que acompanha tudo o que existe no Universo, tal como qualquer outro princípio fundamental. Neste sentido, não só os seres humanos teriam experiência subjetiva, como esta poderia estar embutida em todas as partículas do Universo — uma ideia partilhada, pelo menos em parte, pelo dualismo de propriedades com afinidades pampsiquistas de Thomas Nagel (1979) e pelo emergentismo de William Hasker (1999), cujas propostas iremos analisar mais adiante.

Para desenvolver a sua teoria, Chalmers desenvolve 3 princípios: 1) o da coerência estrutural, 2) o da invariância organizacional e 3) o do duplo aspeto da informação – a ideia central da sua teoria. O primeiro princípio afirma que existe uma relação entre propriedades físicas e a sensiência, pelo que seria possível encontrar algum nível de consciência em qualquer estrutura produtora de informação⁷. Este princípio sugere que para cada estado consciente há um estado físico correspondente no sistema subjacente, de forma que as propriedades fenomenológicas da consciência – como os *qualia* – refletem diretamente a estrutura e organização da informação processada pelo sistema. Assim, a consciência não é vista como algo separado da

6 Iremos verificar que a fenomenologia possui uma função – a de representar estados homeostáticos e qualificar estímulos exteriores como positivos e negativos.

7 Esta teoria é, pelo menos em parte, partilhada por algumas teorias fisicalistas da consciência (Tononi e Koch, 2018).

realidade física, mas como uma propriedade que emerge de maneira coerente a partir da organização dos processos informacionais.

De acordo com Chalmers, a estrutura da experiência consciente está intrinsecamente alinhada com a estrutura da informação que o sistema processa. Por exemplo, no caso do cérebro humano, as experiências fenomenológicas, como a percepção de uma cor ou a sensação de dor, correspondem a padrões específicos de atividade neuronal. Isso significa que os estados conscientes não são independentes dos processos físicos, mas dependem diretamente da maneira como a informação é organizada e processada. Este princípio também implica que a consciência não é exclusiva de sistemas biológicos. Qualquer sistema capaz de processar informação de maneira estruturada poderia, em teoria, estar associado a algum grau de consciência. Por exemplo, um computador ou sistema de inteligência artificial suficientemente avançado, se projetado com uma arquitetura informacional compatível, poderia apresentar uma forma de consciência. A universalidade desse princípio expande a aplicação da teoria para além de organismos vivos, sugerindo que a consciência pode ser encontrada em qualquer estrutura informacional complexa (1995 e 1996). Além disso, Chalmers vai além da aplicação biológica e tecnológica e sugere uma implicação cosmológica: como todas as partículas fundamentais processam algum nível de informação, mesmo que rudimentar, a consciência poderia ser vista como uma propriedade fundamental do Universo. Nesta perspectiva, a consciência não é apenas um subproduto de sistemas organizados, mas algo inerente à própria estrutura da realidade física. Este posicionamento aproxima-se de uma visão pansiquista, na qual a consciência, em graus variados, está presente em todo o Universo.

O segundo princípio da teoria de David Chalmers, conhecido como *o princípio da invariância organizacional*, complementa e expande o primeiro princípio, o da coerência estrutural. Enquanto o primeiro princípio estabelece que existe uma correspondência sistemática entre os estados físicos de um sistema e a experiência consciente associada, o segundo princípio afirma que dois sistemas funcionalmente equivalentes, com o mesmo nível de organização informacional, terão experiências conscientes qualitativamente idênticas, independentemente do substrato físico que os compõe.

De acordo com o princípio da coerência estrutural, a consciência está diretamente ligada à estrutura da informação processada por um sistema. Isso significa

que as propriedades fenomenológicas da consciência, como os *qualia*, refletem os padrões de organização da informação subjacente. Já o princípio da invariância organizacional dá um passo em frente ao postular que essa relação não depende do material específico que implementa os processos informacionais. O que realmente importa para a geração da consciência é o *isomorfismo computacional* – isto é, a equivalência funcional entre sistemas. Por exemplo, se o cérebro humano fosse replicado de forma exata num substrato diferente, como um sistema computacional ou uma máquina artificial, e se a organização funcional e os processos informacionais fossem mantidos, esse sistema seria capaz de gerar as mesmas experiências conscientes que o cérebro humano original. Este princípio desafia a visão de que a consciência está exclusivamente ligada a substratos biológicos, como neurónios e sinapses, sugerindo que qualquer sistema suficientemente refinado no processamento de informação pode ser consciente.

A relação entre o primeiro e o segundo princípio é evidente: enquanto o primeiro princípio fundamenta a ligação entre as propriedades físicas e a consciência, o segundo princípio enfatiza que essa ligação é universal e independente do tipo de material físico. Ambos sustentam a ideia de que a consciência emerge da organização da informação e essa organização é o fator determinante para a experiência fenomenológica. As implicações do segundo princípio são profundas. Ele sugere que a consciência não é exclusiva de sistemas biológicos e pode, em teoria, existir em sistemas artificiais, desde que possuam o mesmo grau de refinamento computacional e a mesma organização funcional. Isso significa que sistemas de inteligência artificial suficientemente avançados poderiam gerar experiências conscientes semelhantes às humanas. Além disso, o princípio da invariância organizacional reforça a ideia de que a consciência é um fenómeno universal, podendo estar presente em qualquer sistema funcionalmente equivalente.

O terceiro princípio – o duplo aspeto da informação – é o mais importante na teoria de Chalmers. Segundo Chalmers, e como já foi afluado, toda a informação possui dois aspetos: o físico e o fenomenológico. O filósofo explica esta dualidade porque, por um lado, como a consciência existe como *building blocks* do Universo, então estaria em qualquer partícula do Universo; por outro, porque todas as estruturas físicas, ao possuírem algum nível fundamental de consciência e ao processarem informação,

poderiam produzir algum grau de informação fenomenológica. Segundo esta ideia, a informação física, porque estaria imbuída de consciência (devido à sua fundamentalidade), desdobrar-se-ia em informação fenomenológica, pelo que onde existem sistemas físicos, existem estados fenoménicos. Em poucas palavras, Chalmers sugere um isomorfismo entre determinados espaços informacionais e determinados espaços físicos. A subjetividade seria, segundo este ponto de vista, um estado fenomenológico da estrutura física (1995, p. 217).

O duplo aspeto da informação procura explicações da consciência na física e não na biologia, desde logo porque a física pode explicar o mundo em função da noção de informação – um campo onde Chalmers crê poder encontrar o ingrediente especial para explicar a consciência (1995, p. 210). O ingrediente especial que Chalmers sugere é a fenomenologia (isto é, informação em estado subjetivo), que uma estrutura física (isto é, informação em estado físico) pode criar. Ou seja, segundo Chalmers, toda a estrutura física possui algum tipo de organização informacional que produz algum tipo de experiência fenomenológica, pelo que a informação possui dois aspetos: o físico e o fenomenológico. Este último podemos encontrar, em diferentes graus, em todas as estruturas físicas (1996, p. 216) – o mesmo seria dizer, em todo o Universo (1995, p. 210). Chalmers sugere, em poucas palavras, uma ontologia psicofísica, na qual a física e a fenomenologia são dois lados da mesma moeda.

Em suma, como todas as teorias reducionistas falharam para explicar a subjetividade da consciência e a lacuna explicativa, precisamos de uma teoria não-reducionista que, segundo Chalmers, não reduza a consciência a nenhuma estrutura física. Pelo contrário, a consciência é um pilar do Universo, pelo que existe em cada partícula que processa informação. O processamento de informação psicofísico permite não só explicar como é que a fenomenologia pode surgir num mundo físico, como ainda de que modo uma estrutura pode ter um duplo aspeto: um físico e outro fenomenológico.

1.2.3. *Outros dualismos de propriedades*

O filósofo Thomas Nagel considera, tal como Chalmers, que a consciência é uma propriedade fundamental do Universo, numa abordagem semelhante ao pansiquismo. Nagel diz-nos de uma forma clara no seu artigo *Panpsychism* que “os

constituintes básicos do Universo possuem propriedades mentais” (1979, 181). Ou seja, a consciência está espalhada por todo o Universo, já que, na essência, toda a matéria possui “microestruturas” não-físicas capazes de criar consciência quando organizadas de uma forma específica. Essas propriedades mentais são intrínsecas à matéria, não emergindo dela, mas estando presentes de forma fundamental em cada partícula. Elas desempenham o papel de "sementes" da consciência, permitindo que, sob certas condições organizacionais, surjam experiências subjetivas nos organismos.

Nagel argumenta que tudo o que existe é feito de matéria, incluindo criaturas conscientes. Contudo, se as propriedades mentais de um organismo não podem ser explicadas exclusivamente pelas suas propriedades físicas, então essas propriedades mentais devem derivar de constituintes fundamentais que possuem essas características. Nas palavras de Nagel,

“Se as propriedades mentais de um organismo não estão implícitas em nenhuma propriedade física, mas devem derivar de propriedades constituintes do organismo, então esses constituintes devem ter propriedades não-físicas das quais o aparecimento das propriedades mentais resulta quando a combinação é do tipo certo⁸” (p. 182).

Assim, segundo Nagel, a matéria que compõe o Universo possui propriedades não-físicas fundamentais que, quando combinadas de maneira adequada, permitem o surgimento da consciência em organismos feitos de matéria. Essas propriedades mentais não são observáveis diretamente, mas são postuladas como necessárias para explicar o aparecimento da consciência. Elas funcionam como elementos básicos que, quando organizados de forma complexa, dão origem à experiência subjetiva. Desta forma, Nagel rejeita o fisicalismo tradicional, pois considera que ele não pode explicar a consciência. Para ele, as propriedades mentais não emergem das propriedades físicas da matéria – elas já estão presentes, de forma latente, na própria matéria, em níveis microscópicos.

Embora a visão de Nagel possua semelhanças com a de Chalmers – ambos defendem que a lacuna explicativa não pode ser resolvida pelo fisicalismo –, eles divergem nas suas propostas ontológicas. Chalmers sustenta que a consciência é um

8 Tradução do autor.

fenômeno psicofísico, resultado da dualidade entre os aspectos físico e fenomenológico da informação. Para ele, a informação é o elo entre o físico e o mental, com ambos os aspectos sendo fundamentais e coexistentes. Nagel, por outro lado, argumenta que a consciência não é um fenômeno dual, mas uma propriedade fundamental das partículas materiais. Não se trata de um aspecto fenomenológico que emerge da organização física, como Chalmers sugere, mas de algo que está na própria matéria desde o início. Enquanto Chalmers propõe que o duplo aspecto da informação reflete uma ontologia fundamental, Nagel vê esta dualidade como uma questão meramente epistêmica, ou seja, uma forma de compreender a relação entre as propriedades físicas e mentais, mas não como uma divisão ontológica real. Para Nagel, não há uma separação intrínseca entre os aspectos físico e mental; o mental é um componente intrínseco e inseparável da matéria, presente em todos os constituintes fundamentais do Universo. Embora esta perspectiva tenda para um monismo pampsiquista, Nagel acaba por propor, em grande medida, um dualismo de propriedades, ao sustentar que a consciência e a experiência subjetiva possuem características inconciliáveis com o fisicalismo.

O filósofo Philip Goff tem uma ideia semelhante à de Nagel, mas com uma abordagem canonicamente pampsiquista. Em *Galileo's Error* (2019), Goff defende que a consciência é uma propriedade fundamental e universal do Universo. Em oposição ao fisicalismo, que não explica satisfatoriamente a experiência subjetiva, e ao dualismo substancial, que introduz complexidade desnecessária, Goff propõe que as partículas fundamentais possuem formas rudimentares de consciência, que se combinam para formar experiências mais complexas, como as humanas (2019). Embora Goff rejeite o dualismo substancial, a sua teoria mantém traços de dualismo de propriedades, na medida em que distingue as propriedades físicas das propriedades fenomenológicas da matéria. Assim, apesar de adotar um monismo pampsiquista, a sua posição implica que a consciência não pode ser reduzida às propriedades físicas convencionais, aproximando-se, nesse aspecto, do dualismo de propriedades.

O filósofo William Hasker defende ostensivamente em *The Emergent Self* (1999) que a consciência *emerge* de propriedades físicas. Segundo o dualismo emergentista, a consciência não é, ao contrário do que defende o dualismo de substância, uma entidade ou propriedade adicional que exista independentemente do cérebro. Em vez disso, é

algo que emerge do cérebro, sendo inerente à sua organização funcional e relacional (p. 189).

Segundo Hasker, tal como a natureza é constituída por fenómenos emergentes – como a afinidade química entre moléculas que formam a água, os gases ou outras substâncias – também a consciência é uma propriedade emergente de algo mais simples, como o cérebro. Há uma base fisicalista na teoria de Hasker, já que uma determinada organização molecular possibilita a emergência da consciência. Da mesma forma que um íman cria campos magnéticos através de uma determinada organização atómica (Hasker, 1999, p. 190), também o cérebro produz um campo da consciência através de uma determinada organização molecular – esta muito mais complexa do que a primeira. Um campo magnético é uma entidade real mas é distinta do objeto físico que lhe está na origem, o que pode ilustrar os dois tipos de realidade que podem existir - da mesma forma que o campo magnético é distinto do íman, também a consciência é distinta do cérebro. Deste modo, a consciência, tal como os campos magnéticos, possui uma causalidade própria, “seguramente no cérebro (telepatia) e provavelmente também nos outros aspetos do mundo material (*telequinesia*)” (1999, 191). Por outras palavras, a consciência, tal como os campos magnéticos, ocupa um lugar no espaço que extravasa a massa cerebral que podemos encontrar dentro dos crânios dos seres vivos (1999, p. 190).

No livro *Persons and Causes* (2000a), Timothy O'Connor, na mesma linha de Hasker, argumenta que a consciência emerge e depende do cérebro físico, mas, enquanto fenómeno emergente, adquire propriedades distintas das propriedades estritamente físicas. Em particular, a consciência manifesta propriedades fenomenológicas que não podem ser reduzidas à fisicalidade do cérebro. Segundo O'Connor, estas propriedades fenomenológicas conferem à consciência poder causal, permitindo-lhe influenciar os processos cerebrais e, por extensão, possibilitar o livre-arbítrio humano (O'Connor, 2000a).

Em suma, as versões do dualismo de propriedades — quer nas suas formulações mais canónicas, quer nas mais próximas de uma perspetiva pampsiquista — sustentam, essencialmente, que o fisicalismo não consegue responder de forma satisfatória ao *hard problem* da consciência. Assim, propõem modelos em que a mente, ou emerge do físico

(Chalmers, 1995, 1996; Hasker, 1999; O'Connor, 2000a), ou está ontologicamente imbuída no próprio tecido da realidade física (Nagel, 1979; Goff, 2019).

1.2.4. *Os problemas do dualismo de propriedades*

A teoria de David Chalmers não deixa de ser, em alguns aspetos, fundamentada em princípios que, atualmente, se afiguram promissores, como a ideia da informação estar na origem do Universo. No entanto, identificamos muitos problemas, que passamos a expor.

1) Tendemos a concordar com duas ideias de Chalmers: A) o Universo é constituído e construído por informação (ver cap. 3); B) a consciência é uma manifestação da informação (ver cap. 3). No entanto, discordamos ostensivamente da ideia de que a informação fenomenológica possui propriedades não-físicas. Consideramos que é um argumento desnecessário, visto o aspeto dual da informação não refutar a fisicalidade da informação (Serrado, 2024 e 2024c). Ou seja, não concordamos, nem sequer consideramos que é necessário, classificar o aspeto dual da informação de não-físico vs. físico. Porque a informação fenomenológica é não-física, quando pode ser considerada física? Para quê dualizar a fenomenologia da física? Aliás, se Chalmers defende que a consciência é um *building block* do Universo físico, presume-se que faça parte das leis da física e não as ultrapasse de nenhum modo. No mesmo sentido, se o filósofo compara a *fundamentalidade* da consciência com a *fundamentalidade* física do espaço-tempo, porquê atribuir propriedades não-físicas à consciência? É o espaço-tempo não-físico? Se o é, como é que impõe leis físicas num sistema regido por leis físicas? A não ser que Chalmers considere que o espaço-tempo é um *building block* do Universo não-físico – não nos parece ser o caso – então não vemos razões para atribuir à consciência qualquer propriedade não-física. A ideia de informação fenomenológica como propriedade distinta da informação computacional parece-nos, portanto, conter problemas.

Sabe-se hoje que, de facto, o Universo é constituído não apenas por matéria no sentido clássico do termo, mas por campos de energia, sem que estes fenómenos sejam desprovidos de fisicalidade (Rovelli, 2018). O som, por exemplo, parece não ter qualquer matéria, bem como a luz, a radiação ou mesmo o vazio. Tudo isto é, no entanto, físico. A eletricidade, que durante séculos constituiu um mistério, é uma substância

completamente física e constituída por partículas materiais – os fótons. As telecomunicações – o tipo de informação que Chalmers utiliza no seu trabalho para explicar a informação – são, igualmente, criadas por sistemas físicos que emitem sinais carregados de informação física. Dualizar a estrutura física – digamos, um telemóvel, um rádio ou um PC – dos sinais que emitem, dizendo que um é físico e o outro é fenomenológico (como também Hasker sugere), é estar a criar um problema que ontologicamente não existe. Como questiona Daniel Dennett, será necessário dualizar as imagens, a cor e o som que estão armazenados num DVD do DVD em si (Dennett, 2015)? Da mesma forma que a radiação e o eletromagnetismo são fenómenos tão físicos como as estruturas que os criam e emitem, também a consciência pode ser tão física como o cérebro que a causa (Rovelli, 2018).

Acrescente-se ainda que a fenomenologia pode ser vista como uma dimensão funcional específica da organização física, resultante da complexidade de sistemas biológicos. Em vez de a considerarmos algo "não-físico", podemos entendê-la como uma manifestação emergente das leis físicas, mas sem transcender o domínio da física. Ao adotar este enquadramento, evitamos criar um dualismo desnecessário entre propriedades fenomenológicas e propriedades físicas. Reconhecemos, em vez disso, que a fenomenologia pode ser uma extensão intrínseca das propriedades físicas de sistemas altamente organizados, como o cérebro humano. Tal perspetiva elimina a necessidade de categorizações ontológicas que não encontram suporte nas evidências científicas atuais.

2) O dualismo de Chalmers é uma teoria exclusivamente fundamentada na física, o que nos parece insuficiente para explicar a consciência. Iremos defender que, embora a consciência possa emergir das leis da física e da informação física, apenas processos biológicos informacionais – nomeadamente sistemas nervosos e a dupla função dos neurónios - permitem o surgimento da consciência. Ou seja, defendemos que apenas informação na sua manifestação biológica (e representacional; ver cap.3) poderá criar consciência. Neste sentido, não vemos necessidade de explicar a consciência fora da biologia, designadamente, da neurobiologia (Damásio, 2010; Edelman, 1995).

3) Também não concordamos com a ideia de Chalmers de que a fenomenologia não tem função. Chalmers ignora, ou prefere ignorar, as descobertas no âmbito da consciência realizadas na neurociência, que vão muito para além das

computacionalidades cognitivas. Por um lado, vários estudos têm demonstrado que certos sistemas neuronais contribuem decisivamente para a experiência subjetiva e podem, de facto, *causar* consciência fenomenológica (Damásio, 2010; Dehaene, 2014; Tononi, 2004, 2008; Koch, 2016; Nani et. al, 2019; Ledoux & Brown, 2017); por outro, diferentes estudos têm demonstrado que a consciência fenomenológica possui uma função tão central que dela dependem a própria experiência percetiva (Humphrey, 2023; Ramachandran e Hirstein, 1997) e a capacidade de planear e antecipar decisões (Rosenthal, 2005 e 2008; Damásio, 1993/2011).

Por exemplo, as experiências fenomenológicas, como o prazer ou a dor, desempenham papéis cruciais na regulação do comportamento, servindo como mecanismos que orientam a aprendizagem, a adaptação e a tomada de decisão. Sem a qualificação emocional e fenomenológica de uma experiência, os organismos teriam dificuldade em priorizar ações ou evitar perigos (Damásio, 1993/2011; Humphrey, 2023). Além disso, várias teorias sugerem que a fenomenologia é essencial para criar um modelo interno do mundo que integre perceções sensoriais e estados internos, permitindo previsões mais eficazes e respostas adaptativas (Cleeremans e Tallon-Baudry, 2022). Ao argumentar que a fenomenologia não tem função, Chalmers parece ignorar evidências que ligam diretamente a experiência subjetiva à sobrevivência e ao sucesso adaptativo dos organismos (Rosenthal, 2005 e 2008). Esta relação é particularmente evidente em estudos que exploram como emoções e sensações conscientes moldam escolhas e comportamentos, indicando que a fenomenologia não é apenas um subproduto, mas sim uma característica funcional indispensável (Damásio, 2003, 2010 e 2011).

4) Existe, de facto, uma lacuna explicativa – de difícil mas não impossível resolução (Ramachandran e Hirstein, 1997) – que é sentir ou conhecer a subjetividade de uma outra criatura. Mas isso não significa que, na realidade, essa mesma subjetividade não possa ser estudada ao ponto de existir uma compreensão total sobre que sistemas podem *causar* todos os estados fenomenológicos. Acredito que, uma vez conhecidos todos os substratos neuronais da consciência, esta ficará compreendida na sua totalidade. Ficaremos a ter a certeza daquilo que hoje se considera o mais plausível: *a consciência é causada exclusivamente pelo cérebro e pelo corpo*. Se um dia estivermos na posse do conhecimento de quais são os sistemas neuronais que sustentam a

consciência – como acredito que será possível em breve – não me parece viável continuar a defender que a consciência é um processo não-físico. Afirmar que nunca poderemos conhecer o interior da subjetividade (como Chalmers e Nagel sugerem) não é o mesmo que afirmar que a subjetividade não pode ser explicada (Skytt, 2014). Claro que ficará a faltar a resposta para a lacuna explicativa (Levine, 1983), mas esse poderá ser um problema intrínseco ao entendimento humano, como defende o filósofo Colin McGinn⁹ (2011), semelhante a outros problemas, tais como o que existiu antes do nosso Universo ter sido criado. O nosso aparato cognitivo e os limites da própria física e biologia poderão impor condicionalismos ao nosso conhecimento, mas isto não invalida que o conhecimento total do cérebro leve à conclusão de que o cérebro é a origem da consciência.

5) David Chalmers desenvolveu uma teoria que tende a ser um beco sem saída e que, de certa forma, moderniza o problema mente-corpo proposto por Descartes (2008) de uma maneira desnecessária. Da mesma forma que Descartes não concebia que a consciência pudesse ser matéria, também os dualistas de propriedades parecem ter dificuldades em aceitar que a fenomenologia pode ser entendida como uma manifestação (e não uma propriedade distinta da) física (Dehaene, 2001; Skytt, 2014). Segundo Chalmers (1996), Nagel (1974 e 1979), O'Connor (2000a) e Hasker (1999), entre outros dualistas, o fisicalismo funciona e pode ser aplicado a quase tudo no Universo – menos à consciência. Ou seja, o dualismo de propriedades defende que o fisicalismo é verdadeiro para as rochas, para as árvores, para os corpos, mas não para a consciência (Stoljar, 2010), o que não nos parece plausível.

6) A teoria do *zombie* como forma de atacar o fisicalismo parece-nos ser extremamente limitada e até ingénua. É, aliás, meramente especulativa, sem possuir qualquer suporte científico ou até lógico. Podemos, de facto, conceber a ideia abstrata de que um *zombie* pode manifestar todas as funções cognitivas humanas, comportando-se como um ser humano completamente funcional, sem com isso manifestar qualquer experiência fenomenológica. Poder-se-ia, aliás, pensar que o sonambulismo é um bom exemplo – provavelmente o mais próximo – de um sujeito que realiza ações

⁹ McGinn considera que a consciência possui uma essência epistemologicamente transcendente (2011, p. 79).

completamente funcionais sem que as mesmas sejam acompanhadas por experiência fenomenológica. Um sonâmbulo pode agir com elevada complexidade e fazer as mesmas coisas que faria se estivesse consciente, tal como um *zombie*.

No entanto, a questão é: pode um sujeito sonâmbulo decidir tão bem como o mesmo sujeito consciente? Seria esse sujeito capaz de decidir e agir da mesma forma que funcionaria com a consciência? Seria uma sociedade de sonâmbulos igual a uma sociedade de seres sensientes? Não nos parece. O sonambulismo envolve, na maior parte dos casos, circuitos cerebrais automáticos e processos reflexivos, mas não demonstra capacidades como aprendizagem complexa, tomada de decisão deliberada ou introspeção, o que condiciona a tomada de decisão e os comportamentos em geral. Desde logo porque sabemos que muitas das coisas que os seres humanos aprendem a fazer são desenvolvidas porque eles estão conscientes durante essa aprendizagem (Baars, 1988 e 2005; Damásio, 2010; Dehaene, 2014; Edelman, 1995; Saur et al. 2008; Adolph et al, 2003; Corbetta e Shulman, 2002; Earl, 2014). A consciência parece ser necessária para aprendermos as faculdades e capacidades que nos tornam humanos, nomeadamente andar, falar, socializar, construir e fabricar, entre muitas outras (Dehaene & Naccache, 2001). Por isso, não nos parece que uma pessoa que nascesse *zombie* pudesse alguma vez manifestar comportamentos que são adquiridos através da sensiência, como os acima enumerados. Caso nunca tivesse sido consciente, a capacidade de aprendizagem de um *zombie* seria muito reduzida e o seu comportamento não poderia ser em nada igual ao de um ser sensiente. Como refere Nicholas Humphrey na sua obra *A Magia da Alma*: “Tem de haver coisas que um *zombie* faria de uma forma diferente precisamente porque está tudo escuro no seu interior” (2012, 3).

Tal como o dualismo de substância, que parece ir muito mais facilmente ao encontro da nossa intuição dualista que muitos consideram inata (Bloom, 2005), também o dualismo de propriedades parece denunciar uma dificuldade extrema para ver a realidade de uma forma não-dualista. Ou seja, da mesma maneira que os dualismos de substância parecem ter dominado o panorama filosófico durante séculos por irem, mais facilmente, ao encontro das intuições gerais e da comunidade intelectual (dominadas e condicionadas pela tradição e pela influência aristotélica e cristã), também o dualismo de propriedades parece insistir na defesa de uma ideia que pode ter tanto

de intuitiva como de enganadora. O monismo espinosano (1677/1922), embora muito menos intuitivo, desde logo porque não vê o ser humano como uma criatura detentora de duas substâncias – o que à primeira vista é difícil de aceitar, devido à impressão clara de que temos um corpo material (porque se vê) e uma mente imaterial (porque não se vê) – pode-nos oferecer, a nosso ver, respostas ontológicas mais válidas. Como veremos ao longo deste trabalho, o monismo (espinosano e não só) oferece uma coerência ao problema corpo-mente que nenhum dualismo parece oferecer, devido não só à incapacidade crónica do dualismo de propriedades em explicar como e porque um sistema físico necessita de criar propriedades não-físicas mas também à dificuldade, tanto do dualismo de substância como do de propriedades, em explicar como é que propriedades (e substâncias) físicas se podem relacionar com propriedades (ou substâncias) não físicas. Bem sei que Chalmers entende ter resolvido este problema no seu princípio da coerência organizacional, em que defende que um sistema psicofísico não impediria que o físico se relacionasse com o não-físico. Mas uma coisa é afirmar e outra, muito diferente, é sustentar. Em rigor, Chalmers não sustenta essa afirmação porque, ou a consciência é física é isso permite a interação com o mundo físico, ou a consciência é não-física e isso torna problemática a relação de uma propriedade não-física com uma propriedade física. Se se trata de uma questão de semântica – e não nos parece sê-lo – então é desnecessário caracterizar a consciência como não-física já que, para a consciência existir num mundo físico, ela terá que ser física.

1.2.5. *Evidências empíricas contra a proposta de Chalmers*

No livro *O Erro de Descartes* (1993/2011), António Damásio verificou que indivíduos com lesões no córtex ventromedial ficavam diminuídos na sua capacidade de sentir emoções, o que afetava a sua capacidade de decisão. Estas pessoas mantinham todas as suas funções cognitivas inalteradas – o seu raciocínio, a memória, a linguagem, o seu pensamento lógico e abstrato – mas as suas decisões deixaram de ser eficientes, saudáveis e benéficas para si e para quem as rodeava. Damásio demonstrou que era a sua menor capacidade para sentir emoções – e não uma menor capacidade intelectual, já que esta se mantinha inalterada – que afetava a sua capacidade de decidir convenientemente. Em grande medida, a experiência fenomenológica destas pessoas (a

sua experiência da realidade) fora afetada, o que prejudicou a sua eficiência cognitiva de forma extraordinária.

Um bom exemplo do que queremos afirmar é o famoso e bem documentado caso de Phineas Gage, analisado em *O Erro de Descartes* (1993/2011). Gage sofreu um acidente em 1848, quando trabalhava nos caminhos-de-ferro e uma barra de ferro trespassou o seu córtex ventromedial. Para espanto de todos, depois do acidente, Gage manteve as suas funções cognitivas e intelectuais intactas. Conseguia falar normalmente, lembrar-se de episódios passados, raciocinar com lucidez e manter um discurso completamente coerente. As suas memórias de curta e de longa duração estavam intactas e o seu discurso era fluente e coerente como sempre tinha sido. Parecia ser a prova de que realmente a consciência não dependia do cérebro, mas sim de uma alma. Contudo, o seu comportamento e personalidade mudaram drasticamente porque a sua experiência fenomenológica mudou. Como muitos disseram na altura, “Gage já não era Gage”. Um homem outrora honesto, responsável, ambicioso, com sentido moral e trabalhador, bem visto por todos os que o rodeavam, tinha-se tornado mentiroso, agressivo, preguiçoso e irresponsável. Era agora impulsivo e incapaz de manter um emprego. Damásio demonstrou, anos mais tarde, que o acidente de Gage no córtex ventromedial tinha afetado a sua capacidade de sentir emoções (a sua consciência fenomenológica) e que essa dificuldade amputava-o de uma conduta ordeira, empática e honesta.

Gage mudara de personalidade mas – diria mais – Gage mudara a sua experiência interna sobre quem era e sobre como é o mundo. A sua percepção de si e do mundo fora afetada. Sem capacidade de sentir emoções com a mesma intensidade (o mesmo é dizer, com a sua experiência fenomenológica alterada), a sua experiência emocional tornara-se apática, cinzenta, inerte, *flat*, o que afetou drasticamente a forma como ele via o mundo. Este exemplo demonstra como a experiência fenomenológica não é apenas uma característica acessória da consciência, mas um elemento central para a integração cognitiva, social e moral do indivíduo. Ao perder tecido neuronal onde são processados sentimentos de emoção, a experiência fenomenológica de Gage fora profundamente modificada e alterada. A sua experiência subjetiva mudara e, com isso, mudara igualmente a percepção que tinha de si, dos outros e do próprio meio que o rodeava. Este exemplo, entre outros entretanto documentados (Montoya, 2021;

Damásio, 2003), parece demonstrar que a consciência possui, de facto, uma função: a de permitir uma percepção sentida do seu corpo capaz de conferir vantagens nas decisões.

Em suma, enquanto Chalmers argumenta que a fenomenologia é um fenómeno sem função, os casos documentados por Damásio e outros autores (Humphrey, 2023; Seth, 2021) indicam o oposto: que a consciência fenomenológica não só integra informações sensoriais e emocionais, mas também orienta o comportamento de maneira adaptativa e socialmente relevante. A mudança drástica na fenomenologia observada no caso de Phineas Gage demonstra que as emoções e os sentimentos não são meros subprodutos da consciência, mas constituem parte integrante da experiência subjetiva e, consequentemente, da interação com o mundo e do processo de tomada de decisão.

1.3. O fisicalismo cognitivista

O fisicalismo é a escola filosófico-científica que defende que se vivemos num universo regido por leis físicas, então: 1) todos os eventos que existem nesse Universo podem ser reduzidos às leis da física; 2) todos os eventos podem ser explicados em função das leis da física; 3) não é possível existirem eventos, propriedades ou causas não-físicas num sistema físico, regido por leis físicas. Por outras palavras, o fisicalismo defende que tudo o que existe num sistema físico existe e pode existir porque, nesse sistema, tudo é, por necessidade, físico (Stoljar, 2010). É importante referir, desde já, que não consideramos o fisicalismo à luz do conceito canónico de materialismo, em que todas as coisas têm de ser reduzidas à materialidade dos objetos tangíveis. O fisicalismo defende, sobretudo, que o Universo é constituído não apenas por matéria (no sentido clássico de objetos sólidos), mas também por campos de forças físicas, como a gravidade, o eletromagnetismo e outros fenómenos que, embora intangíveis, são igualmente físicos (Rovelli, 2018; Smolin, 2019).

No que concerne à consciência, o fisicalismo sustenta que a consciência é um processo ou evento físico que emerge exclusivamente das leis da física e pode, em princípio, ser explicado apenas pelas leis da física e da Natureza. O filósofo J.J.C. Smart

contribui para esta ideia com a teoria da identidade mente-cérebro (1959). Segundo Smart, estados mentais, como a dor, não são entidades separadas, mas idênticas a estados físicos do cérebro. Ele defendia que os fenômenos mentais podem ser totalmente explicados pelas ciências físicas, rejeitando qualquer dualismo de propriedades ou substâncias. Esta perspectiva rejeita qualquer necessidade de recorrer a entidades ou propriedades não-físicas para compreender a consciência.

Outro exemplo clássico desta abordagem é a teoria de percepção introspectiva, defendida por David Armstrong (1968). Segundo Armstrong, a consciência emerge de processos cerebrais físicos que monitorizam os estados internos do organismo, reforçando o entendimento de que todos os fenômenos mentais podem ser explicados pelas leis da física. Em *A materialist theory of the mind* (1968), Armstrong reforça esta perspectiva com a teoria da identidade mente-cérebro, defendendo, tal como Smart (1959), que os estados mentais são idênticos a estados físicos no cérebro.

No seu influente livro *Consciousness Explained* (Dennett, 1991), Daniel Dennett apresenta uma posição fortemente fisicalista e funcionalista, rejeitando qualquer explicação dualista ou essencialista da consciência. A sua tese central é que a consciência pode ser explicada inteiramente em termos de processos físicos e computacionais no cérebro, sem recurso a qualidades fenomenológicas irreduzíveis (os chamados *qualia*, que ele considera ilusórios).

Mais recentemente, o filósofo David Rosenthal defendeu a consciência como um fenómeno relacional e representacional, resultante da interação entre estados mentais de primeira ordem (como percepções ou emoções) e pensamentos de ordem superior que os representam. Tanto os estados de primeira ordem, quanto os pensamentos de ordem superior, têm bases *físicas* nos processos cerebrais. Assim, mesmo que a consciência seja descrita como uma "funcionalidade emergente" no pensamento de Rosenthal, esta emergência ocorre dentro do domínio físico (2005 e 2008).

O fisicalismo é, portanto, uma perspectiva atualmente perfilhada por alguns filósofos que estudam a consciência (Dennett, 1991, 1993, 2015, 2016; Metzinger, 2009; Churchland, 1981, 1985, 1998; Churchland, 1983, 2013; Rosenthal, 2005; Searle,

1998¹⁰), mas tem encontrado sobretudo na ciência um maior número de seguidores (Dehaene, 2014; Tononi, 2004, 2008; Koch et al., 2016; Damásio, 2010; Gazzaniga, 2011; Eagleman, 2012; Humphrey, 2012; Seth, 2021; Crick, 1998; LeDoux & Brown, 2017; LeDoux, 2020; Edelman, 1995). Embora as perspectivas fisicalistas possam divergir entre si, há uma ideia central que me parece ubíqua em todas: para *todos* os fisicalistas, os fenómenos mentais e conscientes, ou *são* físicos, ou *são causados* pelas leis da Natureza, *nunca deixando, portanto, de ser físicos*.

As teorias fisicalistas que aqui iremos estudar com maior profundidade (Dehaene, 2014; Tononi, 2004; Koch, 2016; Tononi e Koch, 2015; Damásio, 2010) possuem, em grande medida, uma matriz de identidade clara – ou seja, pretendem demonstrar que uma determinada organização neuronal, ou *é* a causa de determinado estado mental consciente ou *é* esse estado mental consciente. Isto é, de um modo geral, os autores fisicalistas postulam que a consciência ou *é* um determinado estado cerebral ou *é causada* por um determinado estado cerebral. Existem, aparentemente, diferenças entre estas duas afirmações. Uma defende que a consciência *é* o cérebro (Churchland, 1983), a outra advoga que a consciência *é causada* pelo cérebro (Damásio, 2010). No entanto, não consideramos que existam diferenças ontológicas pertinentes, visto ambas as perspectivas defenderem que a consciência *é* um processo físico, independentemente da consciência *ser* o cérebro ou *ser causada* pelo cérebro.

Tanto em autores que são considerados fisicalistas eliminativos (Paul Churchland, 1981; Patricia Churchland, 1983) ou radicais (Dennett, 2015; Metzinger, 2009) como em autores que são fisicalistas biológicos (Damásio, 2010; Humphrey, 2023; Seth, 2021; Dehaene, 2014), fisicalistas informacionais (Tononi, 2016; Koch, 2016) ou

10 John Searle é visto como um dualista por muitos filósofos. Discordamos desta ideia. Não só Searle é ostensivamente crítico para com os dualistas de propriedades, acusando-os de quererem perpetuar o dualismo cartesiano (1998, 16-17), como ainda julgamos que esta atribuição é fruto de uma interpretação errada das ideias de Searle. Aliás, Searle não só defende que “O dualismo, sob qualquer forma, é hoje em dia geralmente considerado fora de questão, porque julga-se inconsistente com a conceção científica do mundo” (1998, p. 17), como ainda considera o dualismo de propriedades uma “doutrina idiota” (1998, 29). Embora o filósofo defenda que a consciência emerge do cérebro, isto não significa, no seu entender, que aquilo que emerge do cérebro seja uma substância não-física, por um lado, nem significa que aquilo que emerge do cérebro (em forma de imagens) tenha que possuir propriedades ontologicamente distintas do cérebro. Neste sentido, parece-nos que Searle é um filósofo de matriz fisicalista. Ele próprio classifica a sua teoria como um “naturalismo biológico” (1995, 15). Searle prefere designar a sua teoria como naturalista, mas parece-nos uma questão de semântica, visto as perspectivas naturalistas serem ostensivamente fisicalistas (Stoljar, 2010), embora não necessariamente materialistas.

ainda naturalistas com um fisicalismo subjacente (Searle, 1998, 2019), a ideia fisicalista da consciência está bem presente. Vejamos. Paul Churchland (fisicalismo eliminativo) defende que “estados mentais são meramente processos e estados sofisticados de um complexo sistema físico: o cérebro” (Churchland, 1998, 7). Damásio (fisicalista biológico) defende que “A essência do eu [da consciência] é vista como um focar da mente sobre o organismo material que ela habita (2010, 227 e 228). Por seu turno, John Searle (naturalista e, no nosso entender, fisicalista) diz-nos que “estados conscientes são causados por processos do cérebro” (2019, 50). Este último, que muitos consideram (Dennett, 1991; Feser, 2004; Chalmers, 1996), a nosso ver erradamente, um dualista de propriedades, possui, ostensivamente, um pensamento fisicalista. Searle escreveu, numa frase que poderia ser atribuída a muitos materialistas canónicos: “os fenómenos mentais são causados por processos neurofisiológicos cerebrais e são, eles próprios, características do cérebro” (1998, p. 15). Ou seja, para Searle, a mente está para o cérebro como a digestão está para o sistema digestivo (Searle, 1998, p. 15). É a digestão um processo não-físico? Será necessário dualizar a digestão do sistema digestivo? Não nos parece. Da mesma forma, sob o ponto de vista de Searle, também não é necessário dualizar a consciência do cérebro.

Por isso, o facto de Searle afirmar que a consciência emerge do cérebro, ou que o cérebro causa a consciência (1988), não é, a nosso ver, suficiente para afirmar que ele é um dualista de propriedades, como alguns defendem. Esta interpretação errada deve-se, em parte, à ênfase de Searle na irreducibilidade da experiência subjetiva (*qualia*), que ele considera uma característica física emergente, e não uma propriedade não-física ou separada do cérebro. Desde logo, Searle advoga ostensivamente que aquilo que emerge do cérebro sob a forma de consciência é uma propriedade naturalista (e subjacentemente física) do cérebro. Por outras palavras, Searle possui uma perspetiva da consciência que ele considera naturalista mas que é, ontologicamente, biologista e fisicalista, já que, na sua filosofia da consciência, não há nada de não-físico ou que não seja explicado de acordo com as leis da física/natureza (1998). Embora rejeite o reducionismo estrito, Searle enquadra-se num fisicalismo emergentista, onde os estados mentais, embora irreducíveis a propriedades mais básicas, permanecem fenómenos inteiramente físicos. Esta abordagem é compatível com um fisicalismo

amplo que admite níveis superiores de organização (como os defendidos por Rosenthal (2005 e 2008) sem comprometer a base física dos fenómenos.

Como vemos, todos estes autores fisicalistas defendem, essencialmente, a mesma coisa: a consciência *é o que o cérebro faz*, *é causada* pelo cérebro, *existe* no cérebro e *emerge* do cérebro pelo que *se o cérebro é físico, a consciência é física*. Se a consciência *é* o cérebro ou *é causada* pelo cérebro não importa para a certeza que estes autores têm da sua fisicalidade. Não há espaço para qualquer tipo de dualismo (de substância ou de propriedades) nestas afirmações porque em nenhum momento se defende que a consciência *é* ou *possui* alguma característica ou propriedade não-físicas. Neste sentido, as teorias fisicalistas não defendem, necessariamente, que um determinado estado cerebral *seja* literalmente um determinado estado mental. Para Damásio (2010), tal como para Searle (2019) e outros autores fisicalistas (Dehaene, 2014; Tononi, 2008; Koch, 2016; Humphrey, 2023; Seth, 2021), a consciência não é, necessariamente, o cérebro — é, antes, algo que o cérebro faz. Ou seja, estes pensadores não dualizam mente e cérebro; pelo contrário, sublinham que *a consciência é um fenómeno físico gerado por um órgão físico*. Em rigor, o sistema digestivo também *não é* a digestão. A digestão é, isso sim, aquilo que *o sistema digestivo faz*. Qual a necessidade de dualizar o cérebro da consciência se não dualizamos o sistema digestivo da digestão? Sugerimos que, tal como a função do sistema digestivo é processar e digerir os alimentos, a função do sistema nervoso é, entre inúmeras coisas, criar a mente e a consciência. Searle, aliás, compara a fisicalidade da consciência não só com a fisicalidade da digestão mas também com a fisicalidade da mitose, da meiose ou da secreção de enzimas (1998, 15), o que muitos materialistas eliminativos igualmente defendem (Churchland, 1985; Churchland, 1983). Neste contexto, o facto de a consciência *ser* o cérebro ou *ser causada* pelo cérebro é irrelevante — trata-se de uma questão semântica — porque quer fisicalistas restritos (materialistas eliminativos como Churchland ou funcionalistas como Dennett) quer fisicalistas emergentistas (Damásio e Searle) concordam que a consciência é um processo físico.

Nas últimas décadas, surgiram alguns modelos da consciência que, de alguma forma, se tornaram de referência no âmbito científico-filosófico. O espaço neuronal de trabalho (*global neuronal workspace* ou GNW), de Stanislas Dehaene (2021), e a teoria

da integração de informação (*integration information theory* ou IIT), de Giulio Tononi (2004), são, provavelmente, as teorias de consciência mais conhecidas e consagradas. No entanto, reconhecendo os méritos das primeiras, optamos por analisar uma terceira teoria que recebeu, também, amplo reconhecimento: o *Sentimento de Si*, de António Damásio. Em comum, todas estas teorias têm o objetivo de identificar os correlatos neurais da consciência e convergem, por isso, na ideia de a consciência pode ser compreendida, exclusivamente, através do estudo do cérebro¹¹.

1.3.1. O espaço neuronal de trabalho (*Global Neuronal Workspace*)

Na década de 80 do século XX, o psicólogo cognitivista Bernard Baars desenvolveu a teoria *espaço de trabalho global* (*global neuronal workspace* - GNW). Esta teoria cognitivista, baseada em modelos computacionais, defende que o processamento neuronal torna-se consciente através do GNW. Isto é, esta teoria propõe que o cérebro realiza um extraordinário número de processamentos de dados inconscientes em simultâneo, sendo que apenas um de cada vez – aquele que consegue a atenção da GNW – se torna consciente (Baars, 1988 e 2005). Como refere Baars no seu artigo *In the Theatre of Consciousness: Global Workspace Theory*: “In the working theatre, focal consciousness acts as a “bright spot” on the stage, directed there by the selective “spotlight» of attention” (1997, 3). A consciência corresponde, neste sentido, à “luz” da mente que permite tornar conhecível e experienciada uma determinada realidade – seja algo extrínseco ao indivíduo, como uma percepção sensorial, seja algo intrínseco ao sujeito, como um pensamento ou uma memória.

Na década de 90, o neurobiólogo Francis Crick conseguiu demonstrar que as oscilações de 40Hz das ondas gama correspondiam à entrada do estímulo da atenção da consciência. Ou seja, Crick verificou que o aumento de atividade neuronal, medido pelas ondas gama acima dos 40Hz, correspondia à tomada de consciência de um determinado estímulo (1998). Em grande medida, o GNW proposto por Baars (1997) corresponderia, do ponto de vista da atividade neural, aos 40Hz de ondas gama propostos por Crick. Estes resultados foram promissores, contudo, a proposta de Baars era

11 Uma das primeiras evidências de que a consciência podia estar dependente do cérebro veio das experiências de estimulação elétrica de Wilder Penfield e Edwin Boldrey na década de 30 do século XX (Rosa, 2018).

fundamentalmente teórica, ao passo que as ideias de Crick, embora válidas e igualmente corroboradas por Llinás e Ribary (1993), apenas demonstravam como é que um estímulo poderia recrutar a atenção do cérebro, o que é muito diferente da consciência fenomenológica e da subjetividade.

No início do século XXI, os neurocientistas Stanislas Dehaene, Lionel Naccache e Jean Pierre Changeux, influenciados pelo trabalho de Baars e Crick, desenvolveram a teoria *global neuronal workspace* - *espaço neuronal de trabalho* (GNW) (Dehaene e Naccache, 2001; Dehaene e Changeaux, 2003). O avanço tecnológico nas neurociências, nomeadamente as técnicas de ressonância magnética funcional, permitiram realizar experiências no cérebro humano que até então eram impossíveis, nomeadamente identificar os sistemas neuronais aptos a suportar empiricamente o GNW proposto por Baars.

Segundo a GNW, o cérebro humano possui, no córtex pré-frontal, no córtex parietal posterior e no córtex cingulado — estruturas que, em conjunto, integram o chamado córtex frontoparietal — um conjunto particularmente denso de neurónios, com capacidade para estabelecer ligações de longa distância e integrar sinais provenientes de diversas regiões cerebrais. Ora, segundo esta teoria, é este denso grupo de neurónios frontoparietais que corresponde ao GNW — ou os correlatos neuronais da consciência. O GNW, com a sua capacidade de recolher informação de todo o cérebro, corresponde, segundo esta teoria, a uma área representacional (Dehaene e Changeaux, 2003). Por outras palavras, a área frontoparietal permite o surgimento de imagens conscientes — é como se estas áreas *traduzissem* as computações subconscientes (sinais “duros”) para imagens mentais conscientes¹². Esta ideia é importante porque uma das propostas centrais do nosso trabalho é a de que a consciência emerge através da transformação de informação computacional em informação representacional através de um processo de *recursividade informacional*. No capítulo 3, iremos verificar que esta ideia de Dehaene valida a nossa proposta; que informação “dura” (computacional) é decodificada para representações (consciência).

¹² Embora não corroborem totalmente a GNW, esta ideia parece-nos pertinente no âmbito da nossa teoria, que apresentaremos mais à frente.

O córtex pré-frontal como centro executivo da consciência

Na obra *Consciousness and the Brain* (2014), Dehaene refere que uma boa parte do cérebro é especializada em processar um determinado tipo de informação. Ou seja, determinados neurónios são especializados na cor, outros na forma, outros na profundidade, outros na coordenação motora, outros na motricidade fina, etc., mas estes sinais apenas se *transformam* em representações integradas no GNW, isto é, no córtex frontoparietal que tem a função de transformar informação “dura” em informação representacional. Deste modo, em vez de serem especializados em *processar* informação, os neurónios do GNW são caracterizados pela sua habilidade para enviarem e receberem projeções para inúmeras áreas do cérebro de modo a poderem *representar informação*. Estes neurónios, neste sentido, não obedecem a uma lógica local, mas pelo contrário, têm a função de trocar informação de uma forma global, como se a sua função fosse transformar processos computacionais de todo o cérebro em representações¹³ (Dehaene e Naccache, 2001; Dehaene & Changeaux, 2001).

O neurocientista Stanislas Dehaene verificou também que as áreas frontoparietais do GNW são, igualmente, responsáveis pela atenção, o que promove abertura para que a informação do meio seja disponibilizada e possa entrar no sistema conscientemente, algo crucial para os processos de aprendizagem. Isto não invalida que a informação que não entre pela GNW seja, igualmente, processada pelo cérebro, mas de uma forma subconsciente. Deste modo, o GNW tem a flexibilidade necessária para codificar um estímulo que, depois de ter captado a atenção das áreas frontoparietais, é processado por outros sistemas neuronais subconscientes, ao mesmo tempo que a GNW é dirigida para outros estímulos do meio. A consciência, nomeadamente os sistemas frontoparietais, funcionaria, portanto, como porta de entrada sensorial e seria crucial para uma série de capacidades humanas, como memória de trabalho, memória de longo prazo, tomada de decisão, entre outras (Dehaene & Naccache, 2001).

A GNW teria, portanto, a função de tornar conhecível a informação do meio num processo de *top-down* – informação processada por sistemas de corticais de primeira

¹³ Importa referir que, para o GNW desempenhar a sua função, é necessário o contributo do tálamo, que, neste modelo, pode ser entendido como um género de motor do córtex, já que ativa e desativa o córtex através de neurotransmissores como dopamina e noradrenalina, entre outros (Dehaene e Naccache, 2001; Dehaene & Changeaux, 2001).

ordem (GNW), e depois “armazenada” em sistemas de segunda ordem (sistemas subcorticais e outros que não o GNW). No mesmo sentido, a GNW permite igualmente processamento de informação numa lógica *bottom-up*. Isto verifica-se, por exemplo, quando algum tipo de informação subconsciente processada por sistemas de segunda ordem tem força suficiente para captar a atenção da GNW¹⁴. Deste modo, podemos considerar que o GNW possui não só a capacidade de *representar* informação do meio como também a de captar a atenção de determinados estímulos do cérebro e do corpo, *representando-os* em forma de sentimentos, de pensamentos ou de imagens mentais (Dehaene *et al*, 2001, 2004, 2014). No capítulo 3 iremos sugerir que a consciência pode influenciar o cérebro ao transformar informação representacional em informação computacional através de um sistema de recursividade informacional, o que é uma ideia muito próxima da sugerida pela GNW.

Desde modo, o GNW permite dois tipos de processo: um em que estímulos do meio se tornam conscientes, e outro em que a atividade consciente regula o processamento interno, podendo incluir a representação do eu: 1) no primeiro processo, um determinado estímulo do meio capta a atenção dos córtices frontoparietais (o GNW), que representam esse estímulo de uma forma consciente, passando-o depois para sistemas de segunda ordem, que os armazenam no subconsciente (*top-down*); 2) no segundo processo, um conjunto de neurónios de segunda ordem chama a atenção de sistemas frontoparietais (o GNW), o que permite que um determinado estímulo do cérebro (um pensamento, uma memória) ou do corpo (um sentimento de dor) se torne consciente (*bottom-up*). A atividade neuronal de segunda ordem (do subconsciente) pode ou não ser, portanto, suficientemente forte para captar a atenção dos sistemas de primeira ordem (nomeadamente do GNW). Se for suficientemente amplificada pode entrar na GNW, mas se não o for mantém-se subliminal (Dehaene & Changeaux, 2003; Dehaene & Naccache, 2001). Este processo está de acordo com a nossa proposta, analisada com detalhe no capítulo 3: a informação computacional pode gerar informação representacional (*bottom-up*), mas também a informação representacional pode gerar informação computacional (*top-down*). Este é

14 Curiosamente, Dehaene *et al* (2001, 2004, 2014) demonstraram que o GNW corresponde ao espaço onde estão mais ativas as ondas gama, que Crick demonstrou estarem associadas a processos de atenção consciente, como já foi referido (Crick, 1998).

o tipo de consciência a que Dehaene, no seu livro *How We Learn* (2021), chama *global availability*: “having the information in mind” (p. 44).

Esta ideia converge com as ideias de Daniel Dennett sobre a consciência, que vale a pena trazer à discussão. Dennett, no seu livro *Consciousness Explained* (1991), considera que o cérebro produz constantemente torrentes de informação (a que Dennett chama *modelo de esboços múltiplos*) que competem entre si por fama – pela capacidade de captar a atenção da consciência. Dito de outra forma, o cérebro processa informação que compete entre si, e apenas a vencedora dessa “competição” pode entrar na consciência. Esta é a informação que podemos reportar e que faz parte da nossa experiência consciente. Tal como no modelo de GNW, também Dennett considera que o nosso cérebro é um órgão computacional que processa informação através de sistemas inconscientes que produzem esboços, dos quais apenas alguns, e um de cada vez, consegue emergir na consciência. No entanto, enquanto Dennett sugere que se esqueça um teatro cartesiano para a consciência (1991), Dehaene parece sugerir, não um centro, mas uma rede *central* apoiada nos córtices frontoparietais na qual a consciência emerge (2021).

Importa referir ainda que, apesar de, inicialmente, Dehaene ter sugerido a GNW apenas para a consciência relacionada com processos de atenção, mais recentemente (2014) o neurocientista referiu-se, igualmente, à autoconsciência – a capacidade de refletirmos sobre os nossos pensamentos e sobre a nossa identidade, o que está mais próximo da consciência fenomenológica. Dehaene defende que, embora a consciência “atencional” seja distinta da autoconsciência, ambos os processos residem nos córtices frontoparietais, como parecem indicar as lesões nesta área do cérebro que, quando muito extensas, afetam a capacidade do sujeito tanto de desenvolver meta cognição como de reter a atenção num determinado estímulo.

Uma teoria promissora

O GNW é visto por alguns segmentos da comunidade científica como a mais promissora das teorias e outros cientistas têm-na adotado nas suas perspetivas sobre a consciência (LeDoux e Brown, 2017; Ledoux, 2020). No artigo *A higher-order theory of emotional consciousness* (2017), inspirado no trabalho de David Rosenthal (2005), designado por *Higher Order Theory* (HOT) e que sugere que a consciência é um sistema

de re-representação (2005 e 2008), Joseph LeDoux converge em grande medida com as ideias da GNW sobre a importância do córtex frontoparietal para o surgimento da consciência. No seguimento das ideias de Dehaene (2014) e Rosenthal (2008), LeDoux defende, num modelo que ele designou por *General Networks of Cognition* (GNC), que a consciência fenomenológica é uma re-representação cognitiva de informação produzida por sistemas de primeira ordem, que apenas são sistemas de representação (não de re-representação). Ou seja, ele distingue sistemas de primeira ordem (sistemas de representação localizados nos córtices parietais) – dos sistemas de ordem superior (sistemas de re-representação localizados no córtex pré-frontal), cuja função é re-representar os sistemas de primeira ordem (2017).

Este modelo converge com as teorias de Dehaene e de Rosenthal, que também atribuem um papel central ao córtex frontoparietal como uma estrutura crítica para a consciência, identificando-o como o local onde ocorre a integração de informações globais necessárias ao surgimento da consciência. Assim, enquanto Dehaene se concentra na coordenação das informações em todo o cérebro, LeDoux especifica que os córtices pré-frontais desempenham um papel essencial na re-representação introspectiva, permitindo a autoconsciência. O córtex-pré-frontal surge, tanto no GNW como no HOT e no GNC, como uma zona de superconvergência, que de certa forma coordena todo o processamento do resto do sistema nervoso, recebendo e enviando informação representacional, re-representando-a. Os três modelos sugerem, portanto, que a consciência requer uma transformação qualitativa da informação inicial, embora Dehaene (2014) descreva o GNW como uma rede integrativa e LeDoux (2017) e Rosenthal (2008) sublinhem a re-representação em sistemas de ordem superior.

Em suma, vimos neste capítulo um conjunto de teorias que poderíamos designar por espaço de trabalho: GNW-GNW / HOT-GNC. Todas elas defendem, essencialmente, que o córtex-frontoparietal (Baars e Dehaene) ou o córtex-pré-frontal (LeDoux e Rosenthal) possui um conjunto de áreas cruciais para o surgimento da experiência consciente, nomeadamente neurónios que possuem a faculdade de transformar *computações* neuronais em *representações* neuronais (Dehaene, 2021) ou em representações de representações (LeDoux, 2020).

1.3.2. A teoria de integração de Informação (*Integration Information Theory*)

Em 2004, no artigo *Measuring Information Integration*, o neurocientista Giulio Tononi propôs um modelo de consciência designado por *teoria de integração de informação* (*integration information theory* – IIT) (Tononi, 2004, 2008), que, a par da GNW, se tornou uma referência científica no âmbito da temática. A IIT postula, em poucas palavras, que um determinado sistema consegue ter experiências conscientes se tiver capacidade de produzir e integrar informação. Segundo este modelo, quanto maior for a capacidade do sistema de produzir e integrar informação (o que os autores designam de Phi de integração de informação), maior a capacidade fenomenológica do sistema (Tononi *et al.*, 2016). Neste sentido, uma menor capacidade de integrar informação (isto é, um Phi mais baixo) corresponderia a uma consciência mais reduzida, ao passo que uma maior capacidade de integrar informação (e, portanto, um Phi mais elevado) corresponderia a uma consciência mais amplificada (Tononi, 2004, 2008).

A relevância desta ideia é que ela redefine a consciência de maneira quantitativa e funcional, destacando que não há um limite rígido que diferencie sistemas conscientes de não conscientes, mas sim um contínuo onde o grau de consciência depende do nível de integração de informação. Contudo, esta abordagem enfrenta desafios, uma vez que pressupõe que a consciência pode ser medida diretamente através de uma métrica (Phi) — algo que, embora elegante, permanece controverso na comunidade científica. Este modelo permite que a IIT seja aplicada não apenas a sistemas biológicos, mas também à investigação sobre sistemas artificiais, como inteligência artificial e redes neuronais artificiais, levantando questões filosóficas profundas sobre a possibilidade de consciência em sistemas não biológicos.

O Phi de consciência está associado ao que temos vindo a designar por correlatos neuronais da consciência (NCC). Os NCC correspondem, nas palavras de Tononi e de Koch: “the minimal neural mechanisms that are jointly sufficient for any one conscious percept, thought, memory, under constant background” (2018, p.2). Ou seja, os NCC são uma estrutura física — os pré-requisitos neuronais — que permite a experiência consciente e que, na teoria GNW, por exemplo, corresponde a circuitos localizados nos córtices frontoparietais. Essa abordagem destaca a necessidade de correlacionar a estrutura física dos sistemas neuronais com os processos subjetivos de experiência

consciente, embora não resolva completamente a lacuna explicativa entre mecanismos físicos e *qualia*.

A IIT assenta em cinco axiomas fundamentais. O primeiro é o da existência intrínseca, segundo o qual a consciência existe e é a única coisa da qual podemos ter certeza. O segundo é o da composição, que afirma que o sistema consciente deve ser estruturado em mecanismos de causa-efeito que atuem de forma integrada. O terceiro é o da informação, implicando que a consciência depende da capacidade do sistema para gerar informação diferenciada, ou seja, estados únicos e não redundantes. O quarto axioma é o da integração, que exige que o sistema seja capaz de unificar os seus conteúdos, de modo a que a experiência não seja fragmentada. Por fim, o quinto axioma é o da exclusão, segundo o qual a experiência consciente é definida pelo seu conteúdo e pela sua distribuição espaço-temporal, não permitindo que múltiplos sistemas concorrentes partilhem simultaneamente o mesmo conteúdo (Tononi, 2004).

Segundo a IIT, o cérebro produz uma assinatura neuronal específica para cada experiência: uma para ver vermelho, outra para ouvir uma música, outra para saborear um gelado, etc. Se, por exemplo, um determinado grupo de neurónios é ativado quando uma pessoa vê uma cor, a estimulação artificial desse mesmo grupo de neurónios deveria recriar a mesma experiência consciente. Para cada grupo de neurónios ativado existiria uma experiência distinta, o que implica que a fisicalidade neuronal e a experiência fenomenológica estão profundamente correlacionadas (Koch, 2016). No entanto, é importante destacar que esta relação ainda é uma hipótese e não uma conclusão universalmente aceite, especialmente no que diz respeito à definição exata de Phi e à sua aplicabilidade a experiências conscientes complexas.

Embora Tononi e Koch reconheçam a importância do corpo e das emoções para a consciência humana concreta, defendem que a consciência fenomenológica básica pode, em teoria, emergir unicamente da integração de informação, independentemente da existência de um corpo biológico. A IIT, nesse sentido, propõe uma abordagem universalista: qualquer sistema fisicamente integrado e dotado de capacidade para gerar informação diferenciada e unificada poderá, potencialmente, ter consciência, independentemente da sua constituição biológica.

Em contraste, autores como António Damásio (2010), Nicholas Humphrey (2023) e Anil Seth (2021) sustentam que a mente é inseparável do corpo e do self, defendendo

que o corpo é uma fonte essencial de informação interoceptiva para a construção da experiência consciente. Para estes autores, os sentimentos e emoções que emergem do corpo não são simples moduladores da consciência, mas constituem o seu núcleo gerador — sendo a consciência, desde as suas formas mais primordiais, profundamente enraizada na biologia do organismo.

Esta divergência de perspectivas prolonga-se também na identificação dos correlatos neuronais da consciência. Enquanto a GNW — e autores como Dehaene (2021) ou LeDoux e Brown (2017) — apontam para o papel central do córtex frontoparietal, nomeadamente do córtex pré-frontal, Tononi e Koch (2016) refutam essa ideia, argumentando que essas regiões estão mais envolvidas em processos como atenção, memória e tomada de decisão. Para os proponentes da IIT, a consciência fenomenológica reside no córtex posteromedial, sendo este o centro nevrálgico da experiência subjetiva.

O caso de pacientes com lesões nessa região, que frequentemente entram em estados vegetativos apesar de manterem um tronco cerebral funcional, é apresentado como evidência a favor desta hipótese. Contudo, para autores como Damásio (2010), Seth (2021) e Panksepp (2005), essa explicação é insuficiente. Estes sublinham o papel decisivo de estruturas subcorticais — como o tronco cerebral, o tálamo, a amígdala, a ínsula, o córtex cingulado e o córtex orbitofrontal — na mediação das emoções e sentimentos que sustentam a própria consciência.

A validade da IIT

A IIT tem-se afirmado como uma das abordagens mais ambiciosas e influentes no debate contemporâneo sobre a consciência (Nani et al., 2019). Ao propor que a experiência consciente corresponde a um padrão específico de integração de informação dentro de um sistema físico, a IIT oferece um modelo teoricamente robusto e formalizável, centrado na métrica Φ (phi), que quantifica o grau de integração de um sistema. Essa formalização permite, em teoria, aplicar o modelo a qualquer sistema físico, biológico ou artificial, o que a torna particularmente atrativa num contexto de crescente interesse por consciência não-humana e inteligência artificial.

Contudo, essa mesma abstração tem gerado críticas significativas. Diversos autores sustentam que o conceito de informação, tal como operado pela IIT, é

funcionalmente descritivo mas ontologicamente opaco. Isto é, a teoria permite descrever como a informação é integrada num sistema, mas não explica porque ou como certos padrões de integração dão origem à experiência subjetiva — os chamados *qualia*. Autores como Damásio (2010), Seth et al. (2012, 2021) e Nani (2019) argumentam que a informação, por si só, pode dar conta de processos como atenção, categorização ou decisão, mas permanece incapaz de atravessar o hiato explicativo entre os mecanismos físicos e a qualidade intrínseca da experiência consciente. Em suma, a IIT corre o risco de descrever as condições formais da consciência sem abordar a sua natureza fenomenológica.

Apesar destas objeções, acreditamos que a teoria pode ser enriquecida se integrada com perspectivas oriundas das neurociências afetivas e da interocepção. Propomos que sentimentos e emoções — frequentemente ignorados ou secundarizados na formulação da IIT — constituem elementos centrais para compreender a emergência da consciência. Argumentaremos, no capítulo 3, que é precisamente a capacidade do sistema nervoso representar internamente os estados fisiológicos do corpo que permite a formação de sentimentos conscientes. Nesse sentido, os sentimentos seriam representações informacionais de estados somáticos, dotadas de valor adaptativo e significado subjetivo.

Ao incorporar este eixo afetivo-interoceptivo no formalismo da IIT, é possível conceber uma ponte entre a estrutura informacional da experiência e o seu conteúdo qualitativo. Não se trata de resolver de imediato o problema difícil da consciência, mas de propor que os *qualia* não são entidades isoladas ou inexplicáveis, mas sim o resultado de sistemas que atribuem significado interno à informação, sobretudo quando essa informação diz respeito ao próprio estado do organismo. Esta abordagem, ao articular integração de informação com autorreferência somática, poderá abrir novas vias para compreender a natureza da consciência encarnada, sem abandonar a formalização rigorosa que a IIT oferece.

1.3.3. Fisicalismo corporificado – António Damásio

*"We are not thinking machines,
We are feeling machines that think."*

António Damásio

Tanto a GNW como a IIT são teorias essencialmente cognitivistas. Não só os modelos assentam numa perspetiva cortical como é nos processos cognitivos que as teorias procuram explicar a experiência consciente. Por isso designamos essas teorias como fisicalismo cognitivista. O neurocientista António Damásio desenvolveu uma teoria que, embora fisicalista, distancia-se do *corticentrismo* e do *cognitivismo* dos modelos analisados. Como a teoria de Damásio assenta, sobretudo, nas emoções que emergem do corpo em forma de sentimentos, designamos o seu modelo de *fisicalismo corporificado*. Isto não significa que o modelo damasiano rejeite sistemas corticais e cognitivos. Apenas significa que, por um lado, as estruturas subcorticais e, por outro, as emoções e os sentimentos possuem, ao contrário das outras teorias analisadas, um papel ubíquo no surgimento da experiência consciente.

António Damásio, tal como Dehaene e Tononi/Koch, reduz o fenómeno da consciência ao funcionamento neuronal, numa lógica de identidade. Se para Tononi e Koch uma determinada ativação neuronal corresponde à experiência fenomenológica, para Damásio um determinado mapa neuronal *causa* uma determinada imagem mental consciente. Afirmações como "a mente emerge da atividade do tecido cerebral" (2010, p. 91) ou "a atividade mental é provocada pelos fenómenos cerebrais que a antecedem" (2010, p. 34) evidenciam a sua visão de que os processos mentais estão intrinsecamente ligados à dinâmica do cérebro.

É um facto que, ao afirmar que a mente é *provocada* pelo cérebro, Damásio está a dizer-nos que a mente *emerge* do cérebro. Poderíamos interpretar esta ideia à luz de um certo dualismo (que Damásio rejeitaria veementemente) se considerarmos, como os autores dualistas de propriedades, que a consciência que emerge do cérebro é constituída por propriedades não-físicas (Chalmers, 1996; Jackson, 1982, 1986; Nagel, 1979; Hasker, 1999; O'Connor, 2000). O mesmo sucede com a interpretação que muitos (como Dennett (1991); Chalmers (1996)) fazem do pensamento do filósofo John Searle,

já referido e analisado. No que respeita a Damásio, apesar de o neurocientista não se referir ostensivamente a este problema, é suficientemente claro que Damásio considera que a mente consciente é um processo físico que emerge de um cérebro físico. Quando Damásio refere que os sentimentos possuem uma dupla face – “which are «physical» processes and yet yield subjective «mental experiences» (2023, p. 282) - não está a desenvolver um pensamento dualista como muitos podem, erroneamente, julgar. Está apenas a defender que a experiência mental é um processo *físico* causado pela *fisicalidade* de um cérebro. A distinção entre o físico e mental que Damásio faz – e que Searle também aplica - e que emprega no seu vocabulário é pura e simplesmente epistémica. A consciência damasiana (bem como a de Searle) é física.

Biologia, mente e consciência

António Damásio é, como Dehaene e Tononi/Koch, um reducionista – no sentido de reduzir a consciência a processos mais simples. No entanto, é importante clarificar que o reducionismo de Damásio deve ser entendido num sentido fisicalista integrativo e emergentista, pois ele considera não apenas o cérebro, mas também o corpo e os sistemas homeostáticos como partes fundamentais para a experiência consciente.

O reducionismo de Damásio no entendimento da consciência é evidente quando ele procura reduzir determinados estados fenoménicos – como tristeza, alegria, medo, entre outros sentimentos de emoção – a determinados mapas neuronais. Num estudo de referência realizado em 2000 e intitulado *Subcortical and Cortical Brain Activity during the Feeling of Self-Generated Emotions* (Damásio et al., 2000), Damásio procurou encontrar o substrato neuronal de um conjunto de sentimentos que deveriam, segundo o autor, deixar uma assinatura neurofisiológica no cérebro e no corpo. A hipótese de Damásio era que um determinado padrão neurofisiológico (X, Y, Z) deveria corresponder a um determinado estado mental – tristeza (X), alegria (Y), medo (Z)... Ou seja, segundo Damásio, o estado mental tristeza (X) deveria corresponder ao estado neurofisiológico X, ao passo que o estado mental alegria (Y) deveria corresponder ao estado neurofisiológico Y e o estado mental medo (Z) deveria corresponder ao estado neurofisiológico Z. Esta abordagem converge naquilo que Tononi/Koch designaram por substratos neuronais da consciência e que, como já aludimos, está associado a uma lógica de identidade entre o substrato neuronal e a imagem mental consciente, como

os filósofos David Armstrong (1968) e J.C.C. Smart (1959) postularam entre as décadas de 50 e 60 do século XX.

No referido estudo, a equipa de Damásio submeteu 41 voluntários a uma experiência que permitiu analisar a atividade neuronal através de um PET (*positron emission tomography*) SCAN – uma tecnologia que permite analisar com elevado detalhe a atividade neuronal. O estudo foi ao encontro da hipótese ao demonstrar que determinados estados físicos (neurais e somáticos) correspondiam a paisagens mentais do estado do organismo e que as diferenças na paisagem neuronal e somática constituem a razão crítica de porque cada emoção é sentida de forma diferente (2000, pp. 1051). Os diferentes sentimentos de emoção não só ativavam sistemas neurais com padrões distintos, como deixavam uma paisagem corporal diferente. Cada emoção possuía uma assinatura neuronal e somática distinta, sendo que estas assinaturas eram muito semelhantes entre os voluntários (Damásio *et al.*, 2000). Estes resultados, entretanto, foram replicados por outros autores ao longo do tempo (Barret *et al.*, 2004; Barret & Simmons, 2015; Bechara & Naqvi, 2004).

Esta experiência demonstrou, em suma, três coisas: 1) que determinados sistemas neurais e corporais (sobretudo a ínsula, o córtex cingulado, o hipotálamo, a amígdala, o tronco cerebral, o córtex sensorial...) causavam determinados estados mentais, nomeadamente sentimentos de emoções; 2) que uma determinada paisagem corporal correspondia a uma determinada emoção, cada uma distinta das outras; 3) que um determinado padrão neuronal, a par de uma determinada paisagem corporal, correspondia a uma determinada experiência consciente. Nas palavras de Damásio: “Those neural patterns are different for each emotions, as shown by comparing the foci of activation and deactivation noted in subregions of the insula, SII and cingulate” (2000, p. 1051).

Importa notar que o reducionismo de Damásio não se limita ao cérebro mas à própria biologia nas suas formas mais simples. Por isso, em rigor, Damásio não reduz apenas a consciência à biologia do cérebro, mas redu-la à biologia do corpo. Por conseguinte, o autor considera que a consciência não resulta exclusivamente de sistemas neurais, mas também de sistemas não-neurais ligados à homeostasia do corpo, que evoluíram de sistemas simples (unicelulares) para sistemas complexos (multicelulares). Por isso, o neurocientista, na sua obra *O Livro da Consciência* (2010),

sugere compreender a evolução das mentes conscientes em função de uma evolução de sistemas simples para os complexos, desde os organismos unicelulares aos organismos multicelulares, cujos comportamentos são, por vezes, surpreendentemente semelhantes. Segundo Damásio, muitos seres vivos unicelulares exibem comportamentos “eficientes e adaptativos que, para todos os efeitos, se assemelham aos que se observam nas criaturas conscientes” (2010, p. 52), nomeadamente uma “vontade” de viver o tempo que geneticamente lhes é possível, num esforço contínuo e por vezes admirável (2010, p. 53).

Em grande medida, Damásio sugere que os comportamentos primitivos de organismos unicelulares são replicados por organismos complexos, não obstante segundo uma forma mais sofisticada, que conduz a comportamentos mais flexíveis e elaborados, apenas possíveis em organismos dotados de mente e de consciência. As emoções e os sentimentos surgem, assim, quando certos organismos adquirem um sistema nervoso que lhes permite, não apenas reagir ao meio em função de sistemas de reação, mas representar esse mesmo meio – diríamos transformar sinais fisiológicos homeostáticos em programas de ação, como as emoções. Neste sentido, a consciência damasiana pode ser compreendida como um fenómeno emergente que manifesta (torna conhecidos) necessidades, apetites e impulsos presentes desde os primeiros organismos vivos.

Como iremos defender mais à frente, é possível que a consciência, como processo decodificador de informação, permita o conhecimento das “vontades ocultas” de organismos não sensientes, como a variedade de células que compõem os corpos. Talvez a informação “pura”, mas tão ubíqua em todos os organismos vivos, possa ser representada em seres vivos dotados de mentes conscientes. Por outras palavras, talvez a consciência represente a capacidade de auscultar e decifrar a linguagem natural inscrita nos processos vitais das unidades biológicas mais elementares. Quando esses organismos unicelulares passam a integrar um sistema multicelular — como no caso dos seres humanos e de outros animais complexos —, essa linguagem biológica torna-se uma mensagem unificada e integrada, passível de ser sentida e conhecida na forma da experiência consciente (Damásio, 2010).

Emoções e sentimentos

Antes de surgirem a mente e a consciência no curso da evolução, os primeiros organismos possuíam dispositivos de reação – aproximação e fuga – ao meio ambiente, como os que podemos encontrar nos organismos sem sistema nervoso. Mais tarde, os primeiros organismos, dotados de sistemas nervosos simples, adquiriram uma vantagem ao poderem reagir ao meio ambiente com respostas mais flexíveis. Estes não deixaram de possuir os sistemas anteriores, mas tinham agora, uma vez que possuíam um sistema nervoso, uma nova forma de se relacionarem com o meio – através de mecanismos de processamento que possibilitam o surgimento de representações rudimentares tanto do meio quanto do próprio estado interno (Damásio, 2010). Estas representações, ainda que simples, diferem das respostas puramente automáticas de organismos sem sistema nervoso, que apenas reagem diretamente aos estímulos, sem uma interpretação ou processamento adicional.

António Damásio advoga que, em organismos com sistema nervoso, esses sistemas primitivos evoluíram para sistemas de ação mais complexos e flexíveis, a que chamamos emoções. As emoções seriam, neste sentido, dispositivos naturais de regulação homeostática que vieram aprimorar os circuitos básicos de sobrevivência dos organismos sem sistema nervoso – ou com sistemas nervosos mais simples. As emoções são, portanto, programas de ação surgidos para aumentar a eficiência comportamental de organismos que inicialmente reagem ao meio de uma forma mais inflexível. As emoções, ao entrarem em cena, permitiram que os organismos que as tivessem pudessem ampliar e flexibilizar as suas ações. Dito de outra forma, as emoções proporcionaram uma flexibilização das respostas homeostáticas, permitindo a adaptação a situações mais variadas e complexas. Enquanto os organismos unicelulares dependem de mecanismos básicos de reação (aproximação e fuga), as emoções nos organismos multicelulares mais complexos ofereceram uma capacidade de responder com maior sofisticação e eficácia às exigências do ambiente (Damásio, 2010).

Nicholas Humphrey desenvolveu, na sua obra *Poeira da Alma* (2012), uma ideia em tudo semelhante, ao defender que antes da existência da consciência, os organismos reagiam ao meio ambiente em função das respostas – de aceitação e de rejeição – do seu corpo. A sensação, que para Humphrey é a essência da consciência, seria a capacidade de os organismos transformarem esses sinais primitivos em sensações (no

modelo que apresentamos, esta capacidade é possível porque um organismo com sistema nervoso pode representar informação outrora oculta, isto é, não representável). Ou seja, para Humphrey, os primeiros sinais de consciência surgiram quando os primeiros animais conseguiram ter as primeiras sensações – sensações essas que apenas eram possíveis devido à existência de um sistema nervoso (2012).

Também Anil Seth, no seu livro *A Criação do Eu* (2021), defende que a consciência emerge de um primeiro sentimento do corpo vivo. Os sentimentos de emoção são centrais para o conteúdo da consciência surgir como expressão de um corpo que, em esforço alostático permanente, se expressa pela interocepção. Por outras palavras, a interocepção emite sinais sobre a forma como o corpo é regulado, o que é crítico não só para a sobrevivência do organismo como para o emergir da consciência. Neste sentido, a interocepção é crucial para a consciência porque permite não só criar uma unidade fenomenológica através dos sentimentos do corpo, mas também criar uma representação dessa unidade, que Seth designa por inferência interoceptiva (2021).

Embora Damásio perfilhe as ideias de Humphrey e de Seth, o modelo do primeiro obriga a uma distinção entre emoções e sentimentos, o que nos parece muito pertinente. As emoções são, para António Damásio, programas de ação automáticos que se passam no corpo, ao passo que os sentimentos de emoção são o substrato mental das emoções. Muito influenciado pela teoria das emoções jamesiana (James, 1894), Damásio considera que as emoções correspondem à libertação neuronal e química que altera a paisagem do corpo (nas vísceras, no coração, nos músculos, etc.). Esta paisagem emite sinais para o cérebro que, por sua vez, leva à codificação da paisagem corporal em mapas neuronais que permitem o surgimento dos sentimentos em relação àquilo que se passa no corpo – e que é possível ser decifrado através da interocepção (Damásio, 2023). Ou seja, se a emoção X é espoletada por um estímulo, o corpo assume uma determinada paisagem. Esta paisagem é, posteriormente, “interpretada” pelo cérebro em forma de sentimentos de emoção. Desta forma, cada paisagem corporal corresponde a um sentimento ou a um conjunto de sentimentos. Estudos subsequentes (Barret, 2004; Barret e Simmons, 2015; Bechara e Naqvi, 2004) têm vindo a demonstrar, de forma consistente, os resultados de Damásio.

Por isso, Damásio defende que *as emoções são processos que se passam no corpo*, enquanto os *sentimentos são processos que se passam na mente*. Dito de outra

forma, as emoções são respostas neurofisiológicas a um determinado estímulo, enquanto os sentimentos de emoção são percepções (mentais) dessas respostas neurofisiológicas (Damásio, 2017). Neste sentido, as emoções são programas que flexibilizam sinais homeostáticos básicos (como os que podemos observar em organismos unicelulares), ao passo que os sentimentos de emoção são a experiência consciente dessas emoções. Os sentimentos de emoção são, portanto, a consciência – o conhecimento ou a percepção – das emoções que é conseguido através da interocepção (do sentimento do que se passa no interior do nosso corpo) (Damásio, 2002).

Em suma, as emoções são programas de ação que ocorrem no corpo, enquanto os sentimentos são as percepções mentais desses programas. Segundo esta perspectiva, um animal sem experiência subjetiva pode ter emoções mas não terá sentimentos de emoções.

O sentimento de si

Até agora temos referido a consciência como a subjetividade que permite ao indivíduo ter experiências qualitativas. É aquilo que designamos por experiência fenomenológica. Mas, como o filósofo Ned Block verificou em *On a Confusion About a Function of Consciousness* (1995), para além da consciência fenomenológica, também existe a consciência de acesso. Uma coisa é a consciência *qualitativa* da experiência, enquanto outra é o *acesso* que temos à experiência. Isto, em certa medida, parece verdadeiro. Parece-nos que uma coisa é percecionarmos um objeto e uma outra muito diferente é termos a experiência qualitativa desse objeto. Ou seja, quando estamos conscientes temos a capacidade de conhecer o objeto percecionando-o (ver uma maçã vermelha) e, ao mesmo tempo, temos a experiência fenomenológica dessa maçã vermelha (como é experienciar esse objeto). Em certa medida, estes dois tipos de consciência estão relacionados com aquilo que David Chalmers refere quando divide os problemas da consciência em problemas fáceis e o problema difícil. Os problemas fáceis da consciência estariam relacionados com a percepção (com a consciência de acesso), explicável pelas *computacionalidades* neuronais que permitem a uma pessoa percecionar um objeto. Já o problema difícil estaria relacionado com a questão – para Chalmers, impossível de responder do ponto de vista científico – de porque a percepção vem acompanhada de fenomenologia (1995 e 1996).

No entanto, talvez Block e Chalmers estejam errados e tenham confiado excessivamente na intuição. De facto, parece-nos que, ao mesmo tempo que acedemos a um objeto também temos, simultaneamente, a experiência *qualitativa* dessa experiência. Parece que são dois processos a acontecer em simultâneo, relativamente autónomos. Foi esta interpretação do problema que fez com que Chalmers sugerisse a figura bizarra de um *zombie* (1996) – alguém que poderia ter consciência de acesso mas não teria consciência fenomenológica. Segundo esta perspetiva, seria possível estar na posse de um tipo de consciência e não do outro. Mas, e se a perceção consciente (aquilo que Block designa por consciência de acesso e Chalmers identifica como problema fácil da consciência) apenas existir se acompanhada com a faculdade de *sentir* a perceção? Ou seja, e se a perceção apenas for possível porque vem acompanhada por um sentimento – de *qualia* - que possibilita a fenomenologia? E se, sem fenomenologia, não existir consciência de acesso? Dito de outra forma: e se a consciência de acesso apenas é possível porque existe a consciência fenomenológica? Isto significaria que sem sentimentos e, subsequentemente, sem fenomenologia, não existiria perceção (pelo menos perceção consciente) (Damásio, 2010; Humphrey, 2023). Significaria também que, uma vez demonstrado que os sentimentos *causam* a perceção consciente, o problema dos *qualia* ficaria (pelo menos em parte) resolvido. Os *qualia* teriam a função de possibilitar a perceção consciente. É precisamente isto que António Damásio sugere.

Damásio demonstrou no seu livro *O Sentimento de Si* (2013) que o cérebro faz representações percetuais (em forma de perceções) e representações sentidas (em forma de sentimentos). As primeiras não são conscientes, as segundas são; as primeiras estão relacionadas com a exteroceção – a capacidade de o organismo mapear os objetos do meio na forma de imagens mentais – ao passo que as segundas estão relacionadas com a interoceção, isto é, com a capacidade de o organismo mapear o corpo e causar sentimentos que sinalizam o estado do mesmo. Quando um organismo interage com um objeto, existem modificações viscerais que fazem emergir uma emoção que, através da interoceção, é sentida conscientemente. Isto significa que o sujeito percebe conscientemente um estímulo através das modificações fisiológicas que causam um *sentimento que permite esse conhecimento*. Sem sentimento não há perceção consciente porque sem sentimento não há nada que *qualifique* o objeto a ser conhecido.

Numa palavra, é o sentimento que torna o objeto conhecido (Damásio, 2023) e a interocepção é, nas palavras de Damásio, “the enabler of consciousness” (2023, 282).

Nicholas Humphrey possui uma ideia muito semelhante à de Damásio. Para o psicólogo, o cérebro cria duas representações diferentes mas complementares: a percepção e a sensação. Enquanto a primeira é uma representação conceptual, dependente de conceitos e modelos do mundo (poderíamos dizer cognitiva), e segunda é uma representação sensitiva – a capacidade de um organismo ser afetado pelos estímulos do meio (associada aos sentimentos) (Damásio, 2023). Um bom exemplo da percepção e da sensação pode ser visto da seguinte maneira. Se cortarmos a cabeça de um sapo e, de seguida, aplicarmos um químico nocivo para a pele do sapo, o sapo decapitado vai reagir para o retirar – e vai continuar a fazê-lo depois do seu cérebro ser destruído (Humphrey, 2023, p. 107). Poder-se-á dizer que, nestas circunstâncias, o sapo decapitado terá tido um género de percepção de dor, desde logo porque teve um comportamento de dor, mas não pode ter tido a sensação de dor – não tinha como tê-la porque ficara sem cérebro. Em poucas palavras, ter a percepção de dor não significa ter a sensação de dor. Ao ter percepção da dor mas não ter sentimento de dor, o sapo não tem acesso ao conhecimento da dor. A percepção é exclusivamente comportamental, já que não existe do ponto de vista consciente.

O neurocientista Vilayanur Subramanian Ramachandran advogou, em 1997, no seu artigo *Three Laws of Qualia* uma hipótese em tudo semelhante às de Damásio (2023) e Humphrey (2023). Para ele, qualquer percepção que não viesse acompanhada de *qualia* não seria consciente. Segundo esta perspetiva, são os *qualia* – a experiência consciente - que permitem o significado à percepção. Uma pessoa apenas dá significado ao conjunto de letras (ex. “mesa”) porque possui os *qualia* que lhe dão significado. Sem os *qualia*, essa palavra seria apenas uma sequência de letras sem significado (ex. m - e - s - a). No mesmo sentido, uma pessoa, ao ser espetada com uma agulha na mão, reage e tira a mão porque tem a percepção da dor, mas a fenomenologia da dor (o sentimento da dor) é posterior à percepção e à reação. A experiência da dor sucede à reação, não a inicia nem tão pouco a acompanha. Neste sentido, qualquer animal – humano ou não-humano (como o sapo do exemplo acima referido) – pode *reagir* após uma percepção que *não* vem acompanhada de experiência fenomenológica. Pode reagir, mas não pode conhecer. Por conseguinte, para se ter uma percepção consciente é necessário *qualia*. A

consciência subjetiva de um estímulo existe porque existe uma qualidade fenomenológica que acompanha esse estímulo que a torna conhecida (Damásio chama-lhe sentimento, Humphrey chama-lhe sensação e Ramachandram chama-lhe *qualia*) e permite conhecer esse estímulo (Ramachandran e Hirstein, 1997). Com efeito, segundo este ponto de vista poderíamos dizer que o sujeito apenas possui consciência da sua existência se essa percepção de si vier acompanhada pelo *qualia* de si – aquilo que Damásio designa por *sentimento de si* (Damásio, 2013).

Isto contradiz a teoria de Chalmers, para quem: 1) é possível existir percepção sem fenomenologia (como num *zombie*); 2) a fenomenologia não possui função. Como vimos, segundo Damásio (2023), Humphrey (2023) e Ramachandran (Ramachandran e Hirstein, 1997), a 1) percepção depende dos sentimentos para existir (pelo menos conscientemente), pelo que, subsequentemente, a consciência fenomenológica teria a 2) função de dar a conhecer a percepção. O filósofo Immanuel Kant propôs algo muito parecido há alguns séculos atrás, quando sugeriu que o sujeito constrói o conhecimento através das formas *à priori* do entendimento (que podemos associar à razão) e da sensibilidade (que podemos associar à experiência sensorial). Sem sensibilidade seria impossível existir fenomenologia porque seria através das *sensações* que o sujeito seria afetado pelos objetos do mundo (Kant, 1781/ 2008). Algumas hipóteses recentes, como as referidas, parecem dar razão a Kant. Iremos aprofundar este tema mais à frente e verificar que, de facto, os *qualia* parecem ser necessários à percepção consciente.

Damásio sugere que o sentimento de si constitui uma percepção fenomenológica do eu que só é possível graças à mediação dos sentimentos. Em outras palavras, assim como não é possível haver percepção consciente de um estímulo sem a ocorrência de sentimentos — isto é, sem *qualia* —, também não é possível que emerga um sentido de eu sem a experiência sentida do corpo. O eu surge precisamente porque os sentimentos permitem sentir o corpo como próprio, isto é, permitem que o organismo se represente a si mesmo como sujeito da experiência. Sem sentimentos, um organismo poderia, em princípio, processar estímulos e reagir a eles — tal como ocorre com animais com processamento sensório-motor automatizado, como o sapo —, mas essas percepções, por não serem sentidas, não seriam conscientes. Assim, não promoveriam comportamentos orientados pela autorreferência nem possibilitariam a complexidade adaptativa que observamos em seres humanos. Os sentimentos não são, portanto,

meros acompanhamentos afetivos da percepção: são a base da consciência fenomenológica e da construção do self.

Segundo este ponto de vista, o sujeito obtém conhecimento de si e do mundo através de um aparato de molde kantiano, no qual a sensibilidade e a razão constroem a experiência consciente. A sensibilidade (os *qualia*/sentimentos/sensações) permite que o sujeito identifique o objeto a ser conhecido no meio – permite que o sujeito seja afetado por ele – ao passo que a razão (o acesso) permite ao sujeito elaborar ideias sobre o objeto (Kant, 1781/ 2008). Na teoria de Humphrey (2012 e 2023), a sensação, ao contrário da percepção, é centrada no corpo, como se nascesse da interpretação que fazemos de como o nosso corpo foi afetado, o que foi defendido por Espinosa no século XVII de uma forma pioneira e visionária (Espinosa, 1677/1992) e corroborado em grande medida pela neurociência atual (Damásio, 2003). Neste sentido, sem a *qualia*/sensação/sentimento – isto é, sem o corpo ser afetado pelo estímulo – a percepção não é consciente já que a experiência fenomenológica não teve lugar. Em poucas palavras, o que Damásio, Humphrey, Ramachandran, Kant e Espinosa defendem é que sem experiência fenomenológica (sem sentimentos, isto é, sem o corpo ser afetado) não temos consciência da percepção. Se isto for verdade, o *zombie* chalmarsiano seria impossível. Não só porque o *zombie*, ao não ter fenomenologia não poderia perceber o meio, como, subsequentemente, o comportamento do *zombie* teria de ser, necessariamente, muito diferente do de um ser com experiência fenomenológica, que tem acesso consciente às percepções do seu corpo. Chalmers vê os *qualia* como um artifício da Natureza – um género de elemento decorativo, um apêndice, sem função comportamental e sem utilidade cognitiva – e é esta sua percepção que permite a ideia do *zombie*. No entanto, parece que os *qualia* possuem uma função central no comportamento humano ao possibilitarem a percepção consciente do meio e do eu que é, aliás, o que permite a flexibilidade comportamental que caracteriza os seres humanos. Sem *qualia* não há percepção consciente, o que é o mesmo que dizer que sem *qualia* os seres humanos não se comportariam, de todo, como se comportam.

Na sua obra *À Procura de Espinosa* (2003), António Damásio refere que existem vários tipos de sentimentos: sentimentos de fundo ancorados no corpo¹⁵, sentimentos homeostáticos (fome, dor, sede, etc.) e sentimentos de emoção (alegria, tristeza, medo, etc.). Embora pertençam a “classes” de sentimentos diferentes, todos os sentimentos têm origem nos sentimentos de fundo. Os sentimentos de fundo são centrais para a consciência e correspondem a um processo de interconexão entre o cérebro (sobretudo o tronco cerebral) e o corpo do qual resulta uma primeira sensação de estar vivo (Damásio, 2003¹⁶).

Os sentimentos são, também, marcadores biológicos, isto é, qualificadores do estado da vida – da homeostasia. Em criaturas com sistema nervoso, sempre que a homeostasia é desequilibrada, o cérebro identifica esses desequilíbrios através de sentimentos homeostáticos, como fome, dor, sede, frio, falta de ar, etc. (Damásio, 2010, 63-66). Quando o organismo funciona bem, dentro dos limites homeostáticos, o sujeito sente-se bem, experienciando *qualidades* a que podemos chamar prazer e recompensa, ao passo que quando o organismo está a funcionar muito perto dos limiares homeostáticos, o sujeito sente *qualidades* como dor e castigo (2010, 77).

Por sua vez, os sentimentos de emoção como alegria, tristeza, medo, nojo, etc., nascem dos sentimentos de fundo e dos sentimentos homeostáticos, com o objetivo de mobilizar o corpo para melhores oportunidades de sobrevivência. Quando um sujeito sente alegria ao acabar de comer um alimento saboroso ou quando acaba de ter relações sexuais, esse sentimento, ao *qualificar* a situação como prazerosa, estimula o sujeito a repetir esse comportamento. A alegria funciona como um motivador para que se repitam determinadas ações que favorecem a sobrevivência e o bem-estar. O mesmo princípio pode ser aplicado ao medo. Quando um sujeito sente medo porque foi confrontado com uma dor ou um perigo, esse sentimento, ao *qualificar* a situação como dolorosa, está a reforçar o sujeito a *não* repetir esse comportamento. A experiência *qualitativa* do medo funciona como motivador para que não se repitam determinadas ações perigosas para a sobrevivência. De notar que tanto o medo como a alegria são sentimentos de emoções que, como já foi referido, surgem como tal através de

15 A que poderíamos chamar sensiência, à imagem da teoria de Humphrey (2023), também ela apoiada na capacidade do corpo de servir de teatro de sentimentos interpretados pela consciência.

16 Seth possui uma ideia semelhante (Seth, 2012).

alterações de paisagens do corpo (Damásio *et al.* 2000; Barret *et al.* 2004; Barret & Simmons, 2015; Bechara & Naqvi, 2004).

Os sentimentos de emoção contribuem para aquilo que Damásio designa como *hipótese dos marcadores somáticos* (1993/2011). Esta teoria, que tem vindo a ser confirmada através de vários estudos (Barrett, 2004; Barrett e Simmons, 2015; Bechara e Naqvi, 2004) postula que estados emocionais qualificam e catalogam experiências emocionais que são frequentemente recrutados na tomada de decisão (supostamente racional). Segundo este ponto de vista, o sujeito utiliza as suas emoções (e o seu corpo) para decidir, o que significa que a cognição depende dos sentimentos de emoção para ser orientada. É através dos marcadores somáticos que o sujeito *qualifica* qualquer situação como boa ou má, utilizando esse mesmo mecanismo para decidir sempre que uma situação semelhante surgir (Damásio, 1993/2011).

Os sentimentos (ou *qualia*) são, portanto, ubíquos para a consciência porque *qualificam* objetos, situações e experiências. Sem sentimentos não é possível qualificar alguma coisa pelo que, se a consciência existe como o sentimento *qualitativo* da realidade, então sem sentimentos não pode existir consciência. Se não temos forma de qualificar um objeto, de que forma podemos ter conhecimento desse estímulo? Sem sentimentos não há consciência porque não se pode estar consciente daquilo que não pode ser qualificado. Segundo esta perspetiva, o “penso, logo existo” (Descartes, 1637/2008, p. 52) é um erro assente numa perspetiva cartesiana que considera a razão, não só uma faculdade superior como ainda a sede da consciência. Para Damásio, seria mais correto afirmar “sinto, logo existo”, já que, segundo o modelo de consciência damasiano, a consciência nasce quando o cérebro consegue representar um eu que *sente* (e logo sabe) que existe. São os sentimentos que permitem ao organismo a sensação – um primeiro tipo de conhecimento – de estar vivo. A consciência corresponde, segundo esta ótica, ao processo de conhecer através do sentimento, pelo que um organismo é consciente quando *sente* – mais do que *sabe* – que está vivo. Por outras palavras, são os sentimentos que animam a consciência (Damásio, 2023). De certo modo, as emoções animam o corpo ao passo que os sentimentos animam a consciência. O pensamento – ou a cognição – não é suficiente para este processo. Esta ideia é central no pensamento de Damásio, porque este considera que a consciência é, sobretudo, um sentimento de existência originado num corpo vivo. Aliás, o corpo vivo

só sabe que está vivo porque *sente* que está vivo. E este conhecimento é possível devido a um tipo de célula especial: os neurónios.

A especificidade dos neurónios e a sua capacidade de transmitir informação, bem como a relação intrínseca que estabelecem com o corpo terá possibilitado transformar *sinais* fisiológicos em *imagens mentais* sentidas. Os pensamentos e os sentimentos podem ser, neste sentido, uma manifestação consciente (representacional) das “vontades ocultas” do corpo (Damásio, 2010) – algo que iremos sugerir no capítulo 3. Dito de outra forma, os neurónios e a sua capacidade de recolher e transmitir informação podem ter a capacidade de decifrar sinais fisiológicos e transformá-los em representações conscientes. Damásio questiona, aliás, se a vontade feroz que temos de viver não será uma manifestação dessas forças internas implícitas das células que compõem o nosso corpo, de viverem o tempo que lhes é geneticamente determinado (Damásio, 2010, p. 60). Ao permitirem a transformação de sinais do corpo em representações mentais conscientes, os neurónios possibilitam que as vontades do corpo se tornem conhecíveis, não só em forma de sentimentos de emoção, mas também de pensamentos.

No que concerne aos *qualia* causarem a perceção consciente, António Damásio sugere algo que dificilmente é refutável. Podemos, de facto, reduzir ou anular os *qualia* se reduzirmos ou anularmos a capacidade de ter sentimentos. Ou seja, ao desligarmos a reatividade emocional de um organismo através de tranquilizantes como *valium* ou antidepressivos como *prozac*, ou ainda um bloqueador beta como *propranolol*, estamos a “desligar” os *qualia*, já que sem sentimentos não há *qualia*. Dito de outra forma, estas drogas, se utilizadas de uma certa forma, podem reduzir ou anular os sentimentos emocionais, o que, inevitavelmente, conduz a uma perda dos *qualia* – uma perda da subjetividade (Damásio, 2010, 315) – o mesmo é dizer, um enfraquecimento da representação do eu. Estas drogas, se administradas com uma sobredosagem podem, aliás, manter o sujeito acordado mas, inevitavelmente, adormecido do ponto de vista da experiência qualitativa, sem ser capaz de perceber fenomenologicamente a realidade. Isto permite-nos observar que, reduzindo os sentimentos e, subsequentemente, os *qualia*, alteramos a experiência fenomenológica, nomeadamente a perceção consciente do ambiente e da própria individualidade.

Ao ter a sensação – ou o sentimento – de dor ou de prazer, nenhum organismo pode duvidar da sua existência. Como refere Milan Kundera no seu romance *A Insustentável Leveza do Ser*: “A frase «penso, logo existo» (Descartes, 1637/2008, p. 52) foi dita por um intelectual que subvaloriza a dor de dentes” (in Humphrey, 2023, p. 150). Ora, Humphrey, no seu livro *Senciência*, diz-nos que: “Quando se masturba, nem sequer um banobo seria capaz de por em causa a sua essência única e imutável” (2022, 150). Segundo este ponto de vista, os *qualia* existem porque o cérebro consegue representar sensorialmente o seu mundo interno, o que permite criar uma experiência qualitativa do estímulo e do *eu*. Isto, segundo Humphrey, poderia explicar o *hard-problem* da consciência, ou seja, explicaria porque o conhecimento vem acompanhado de experiência subjetiva. Sem *qualia* não existe experiência fenomenológica, sim, mas segundo este ponto de vista de Damásio/Humphrey não existe experiência de todo, pelo que o mundo interior e exterior do sujeito nunca é conhecido. Em poucas palavras, sem sentimento/sensação o sujeito está às escuras.

Dito isto, importa reter cinco ideias centrais na teoria de Damásio: 1) as emoções são programas de ação que dependem de um corpo; 2) os sentimentos de emoção são o substrato mental dessas ações; 3) os sentimentos de fundo são a sensação mais ubíqua que sinaliza um organismo que está vivo; 4) são os sentimentos que permitem a consciência de acesso, isto é, sem sentimentos não existe consciência de acesso; 5) são os sentimentos que permitem o surgimento de um eu – a representação consciente de um eu. Por conseguinte, a experiência de estar vivo depende: 1) de um corpo onde se manifestem emoções, 2) de um sistema nervoso capaz de transformar emoções em sentimentos e, subsequentemente, 3) da experiência de sentimentos de fundo que indiquem ao organismo que está vivo; 4) de sentimentos que façam emergir um eu e, deste modo, permitam a percepção consciente do eu e do meio.

A importância dos qualia para a percepção consciente: provas empíricas

Aprofundemos um pouco mais a relação entre os *qualia* e a percepção. Damásio, ao longo da sua vasta obra (2003, 2010, 1993/2011, 2013, 2017 e 2023), enumera uma série de pacientes que mantiveram as suas capacidades cognitivas intactas mas que perderam, em grande medida, a capacidade de sentir, com dramáticas repercussões para a sua vida e para o seu eu.

Podemos dar o exemplo (mais famoso) de Phineas Gage, já aqui referido, cuja redução na intensidade dos sentimentos de emoção o levou a mudar completamente a sua personalidade. Gage mantinha as suas funções intactas ao nível da atenção, comunicação, linguagem e memória, o que lhe permitia ter um discurso perfeitamente normal e competente. No entanto, incapaz de sentir sentimentos de emoção acima de uma determinada intensidade, Gage não só não conseguia tomar decisões racionais como desenvolveu comportamentos completamente desadequados. Apesar da cognição de Gage estar normal, a redução afetiva substancial condicionou a forma como ele experienciava a realidade (Damásio, 1993/2011). A sua experiência fenomenológica ficou severamente afetada, o que afetou a sua percepção, do meio e de si. Claro que Gage conseguia perceber objetos e o mundo, desde logo porque não ficou sem sentimentos de emoção. No entanto, a sua percepção ficou amplamente comprometida, como de resto pode ser comprovado pelas decisões completamente desadequadas e desajustadas que tomava. A sua fenomenologia, ao ter sido afetada, condicionou a sua percepção e, subsequentemente, a sua tomada de decisão.

Uma das condições em que podemos assistir à redução de sentimentos de emoção é o *síndrome de Cotard*. Esta doença psiquiátrica, caracterizada por uma redução substancial na afetividade, leva a que o paciente acredite que esteja morto ou podre. Estas pessoas, tal como na condição que afligiu Gage, podem pensar, comunicar e memorizar com facilidade. No entanto, a redução drástica na sua afetividade prejudica-lhes a experiência fenomenológica como se os sentimentos não conseguissem animar a consciência a partir de um certo nível. Ao terem uma reduzida experiência fenomenológica – isto é, ao não apresentarem sensações e sentimentos suficientemente intensos – estas pessoas acham que não existem ou que estão mortas. A ausência quase total de fenomenologia condiciona a *percepção* ao ponto desta se tornar completamente distorcida. Com reduzida fenomenologia, estas pessoas não possuem o impulso interior que lhes permite ter a certeza que estão vivos. Por isso, a percepção do eu é afetada. A redução drástica na intensidade dos sentimentos de emoção condiciona e limita o sentimento de existência (Ananthaswamy, 2015).

Nicholas Humphrey apresenta em *Senciência* (2023) vários casos em que os sujeitos perderam a visão sensorial consciente, mas não necessariamente a capacidade de percepção visual inconsciente, o que é revelador do papel central dos *qualia* na

percepção, como veremos. Um dos exemplos referidos é o de Helen, uma pequena macaca cujo córtex visual foi destruído, sem que isso a impedisse de reagir aos estímulos visuais. Embora Helen não tivesse experiência visual consciente, conseguia desviar-se de objetos colocados à sua frente, sugerindo que mantinha uma forma de percepção visual inconsciente. Através da análise do seu comportamento, Humphrey propôs que Helen sofria de "visão cega" (*blindsight*): sem *qualia* visuais, não possuía uma percepção fenomenológica dos objetos, mas parecia capaz de detectá-los a nível cognitivo. Segundo Humphrey, ao não ter experiência fenomenológica, era como se Helen não "soubesse" que via — apesar de, funcionalmente, ver. Faltava-lhe uma experiência consciente e integrada do campo visual, o que a tornava cega em termos subjetivos, ainda que conseguisse orientar-se no espaço com base em informação visual inconsciente.

A teoria de Humphrey pôde ser testada em seres humanos nas mesmas condições que Helen. O paciente D.B., na tentativa de curar enxaquecas impossíveis de tratar, foi submetido a uma cirurgia que lhe removeu o córtex visual do lado direito do cérebro. Ora, isto levou a que o paciente perdesse de imediato a visão do campo visual esquerdo. Após a operação, quando se apresentava uma luz ao seu campo visual esquerdo, D.B. negava que era capaz de a ver ou de a identificar. Contudo, pediam-lhe para ele indicar onde é que ele imaginava que estava a luz. Para espanto de todos, D.B. acertou sempre onde estava a luz, não obstante insistir que não sabia para onde estava a apontar. Mais espantoso ainda é que novos testes demonstravam que ele era capaz de identificar a forma e a cor de objetos que lhe eram apresentados ao campo cego. Os resultados sugeriam que D.B., tal como Helen, tinha percepção inconsciente de objetos porque tinha perdido a experiência fenomenológica (2023, p. 59).

Um outro exemplo é o de H.D., que ficou cega aos 3 anos quando contraiu varíola. Aos 27 anos, um cirurgião ótico conseguiu restaurar-lhe a ótica através de enxertos de córnea. Em teoria, H.D. poderia voltar a ver, já que tinha recuperado a computacionalidade dos olhos. No entanto, não obteve grandes melhoras porque, muito provavelmente, ocorreram alterações degenerativas no seu córtex visual por não ter sido exercitado durante tantos anos (2023, 65). Tal como D.B. e Helen, H.D., apesar de cega, conseguia desviar-se dos objetos, mesmo que o fizesse de forma totalmente inconsciente. A falta de um córtex visual a funcionar em condições privava H.D. de experiência sensorial e, subsequentemente, de experiência fenomenológica. Como

defendera Kant há séculos atrás, a razão não é suficiente para representar o conhecimento. É necessária a sensibilidade (as sensações) para que os fenómenos surjam ao sujeito num estado percetível, e depois possam ser pensados e idealizados pela razão (Kant, 1781/2008).

Em suma, tanto Helen como D.B. e H.D., ao terem ficado com os córtices visuais destruídos ou lesados, desenvolveram o que hoje em dia é conhecido por visão cega. Estas pessoas conseguem ter perceção visual mas não têm perceção sensorial, o que significa que não têm experiência fenomenológica daquilo que veem. Não obstante, o mais espantoso é que, apesar de conseguirem percecionar inconscientemente os objetos, estes indivíduos continuam a considerar que não conseguem ver. Isto acontece porque, ao perderem a experiência fenomenológica, a sua perceção tornou-se inconsciente. Ao perderem a visão sensorial – ou fenomenológica – deixaram de ter a capacidade de experienciar o que veem (Humphrey, 2023). Estas experiências remetem-nos para outras muito semelhantes que analisaremos mais à frente mas de que devemos dar nota breve. Michael Gazzaniga verificou que pacientes com os hemisférios separados têm visão cega quando lhes são apresentados estímulos ao hemisfério direito. Tal como os sujeitos acima indicados, os pacientes de Gazzaniga não conseguem ver conscientemente o que lhes é apresentado ao campo visual esquerdo (hemisfério direito) mas escrevem ou desenhavam os objetos que (não) veem. Estes pacientes percecionam inconscientemente os objetos mas, ao não terem perceção sensorial dos mesmos, acabam por não ter experiências fenomenológicas (2011).

Estas experiências demonstram dois aspetos fundamentais: (1) que os *qualia* desempenham a função de possibilitar o conhecimento consciente, pelo que um sujeito desprovido de determinado *qualia* sensorial não poderá ter uma perceção consciente dessa modalidade; (2) que, embora seja possível haver perceção inconsciente na ausência de *qualia*, tal não implica que o comportamento resultante seja tão eficiente ou flexível quanto aquele que se baseia em perceções conscientes. Com efeito, em todos os casos analisados, os sujeitos revelavam comportamentos significativamente mais eficazes e adaptativos quando a perceção era acompanhada por experiência fenomenológica, em comparação com situações em que a perceção ocorria de forma meramente inconsciente.

A consciência damasiana

Damásio, ao contrário das teorias da consciência cognitivistas, considera que a consciência é um processo gradual, pelo que envolve três níveis (2003, 2010, 2011 e 2013). O que para Chalmers (1995, 1996), Dehaene (2014) e Tononi/Koch (2015) é a consciência, para Damásio é o último estágio de consciência – a consciência alargada e autobiográfica, que mais facilmente podemos comparar com a GNW ou a IIT.

A consciência alargada que Damásio sugere está alicerçada em dois “tipos” de consciência mais estrutural e mais ubíqua - o proto-eu e o eu-nuclear -, pelo que não pode funcionar sem os sentimentos que emergem de um corpo vivo (2013). Segundo Damásio, as teorias cognitivas têm dificuldade em explicar: 1) como é que lesões em sistemas subcorticais (fundamentais no mapeamento do corpo e na manutenção da homeostasia e associados ao proto-eu e ao eu-nuclear damasiano) levam inevitavelmente à suspensão da consciência alargada e fenomenológica; e 2) como é que o “simples” processamento de informação neuronal permite a integração de uma experiência unificada; 3) como é que a percepção consciente é condicionada – e pode desaparecer – através da manipulação dos sentimentos.

Damásio oferece uma resposta para estes problemas: 1) os sistemas subcorticais são dispositivos que estão dedicados à gestão da vida em forma de emoções e de sentimentos e são estes impulsos que vêm do corpo que permitem, primeiro, o surgimento de uma mente e, depois, de uma consciência; 2) a consciência depende de *um* corpo (mapeado no tronco cerebral, no córtex somatosensorial e na insula) para possuir uma dimensão unificada e integrada¹⁷, e isso pode ser demonstrado pela suspensão da consciência quando o corpo é “retirado” do cérebro através de lesões em determinados sistemas (como o tronco cerebral) que mapeiam o corpo; 3) a percepção consciente do objeto – e do eu – apenas é possível se a fenomenologia permitir qualificar o que se está a experienciar (2010, 2013). Poderíamos colocar este problema na seguinte proposição: 1) a consciência fenomenológica permite o conhecimento; 2) a fenomenologia depende dos sentimentos (da sensação, dos *qualia*); 3) logo, a percepção consciente apenas é possível se acompanhada de sentimentos.

17 Humphrey defende o mesmo quando diz que “À medida que a vida foi decorrendo e o nosso corpo – o nosso único corpo – começou a interagir com o mundo exterior, as partes separadas uniram-se, orquestradas, por assim dizer, pela partitura isolada que nos compôs a vida” (2023, 19).

No mesmo sentido, as teorias cognitivas acima analisadas não explicam como é que uma mente se torna conhecedora, isto é, são omissas sobre como é que uma mente adquire o sentimento de identidade – como é que um eu como objeto (mente) se transforma num eu como conhecedor (consciência). Ora, Damásio postula que a consciência (o eu enquanto conhecedor) emerge da relação do eu com o objeto através das modificações que o objeto exerce no corpo e que, uma vez que causam emoções (isto é, modificações fisiológicas) no corpo, manifestam-se como sentimentos. A lógica inverte-se na teoria de Damásio, já que não é na mente que reside o eu, mas sim no corpo e nos sentimentos que dele emergem e que conduzem à consciência (Damásio, 2003; 2010)¹⁸.

Para que os sinais do corpo e, subsequentemente, os sentimentos animem a consciência é primordial o proto-eu – um primeiro nível de consciência. O proto-eu é a consciência essencial centrada no organismo e dependente dos sinais que vêm do corpo em forma de sentimentos de fundo – uma forma de sensiência, na teoria de Humphrey (2023). Neste sentido, o proto-eu “Consiste num aglomerado de imagens que descrevem aspetos relativamente estáveis do corpo e criam sentimentos espontâneos do corpo vivo (sentimentos primordiais)” (2010, 228). É um eu centrado no corpo – aquilo que Rosa designou com mestria por “sentimento de corporeidade” (2018, 77). Está associado aos sentimentos de fundo que, para Damásio, correspondem aos sentimentos de um corpo vivo, sem a necessidade de ser mediado por um estímulo externo. No proto-eu existe um sentimento primordial produzido pelo tronco cerebral, em particular, a Formação reticular, o núcleo do trato solitário, núcleos parabranciais, periaqueduto, hipotálamo e a insula posterior. O sentimento primordial do proto-eu é o antecessor de todos os sentimentos de emoção e, portanto, “a base de todos os sentimentos provocados pela interação entre os objetos e o organismo” (2010, 233).

O proto-eu é constituído por mapas neuronais que criam imagens do corpo “mas também imagens sentidas do corpo” (2010, 239), isto é, do estado homeostático do organismo que surge na mente em forma de sentimentos. Aliás, sempre que se verificam alterações na consciência verificam-se igualmente alterações nas emoções e nos

18 Com exceção da GNC de LeDoux, que postula que a mente torna-se consciente através de processos de re-representação (nos quais as emoções possuem um papel determinante), a GNW e a IIT são omissas a este respeito (2020).

sentimentos (Damásio, 2013, 34), como temos vindo a referir. No modelo de Humphrey, semelhante ao de Damásio em alguns aspetos, o facto da sensação estar apoiada no corpo permite como “expressar uma opinião corporal em relação àquilo que o estímulo significa para nós” (2023, 94). Desta forma, não há consciência sem sentimentos porque sem sentimentos – *sem os qualificadores do estado homeostático do corpo* – não existem imagens sentidas do corpo, pelo que não há sentimento de existência e, subsequentemente, não há conhecimento sem o corpo (Damásio, 2010; Seth, 2021), algo que Espinosa já tinha sugerido de forma visionária no século XVII (1677/1992).

O segundo nível de consciência damasiana é o eu-nuclear. O eu-nuclear continua a precisar do tronco cerebral (continua a ser fundamental) mas inclui agora outros sistemas, tais como o tálamo, a ínsula anterior, o córtex cingulado anterior, o córtex somatossensorial e a ínsula posterior. O eu-nuclear surge quando o proto-eu (e mapas corporais associados) é modificado pelos sentimentos através de uma interação com algum estímulo do meio. Nas palavras de Damásio: “o eu nuclear é criado através da ligação do proto-eu modificado ao objecto que provocou essa modificação, um objecto que foi agora destacado pelo sentimento...” (2010, 254). Ou seja, o organismo interage com o mundo através do corpo e são as modificações no corpo (emoções transformadas em sentimentos) feitas pela interação com o meio ambiente, que fazem emergir o eu-nuclear.

Em Damásio, o corpo é, neste contexto, um filtro da realidade, porque é nele e através dele que o sujeito não só apreende a realidade como adquire o seu eu. É a interocepção – a capacidade de um sujeito sentir as modificações fisiológicas que ocorrem no seu corpo na relação com o objeto - que permitem criar uma ideia sobre a interação do eu com o meio. Dito de outra forma, o corpo do sujeito, ao ser afetado pelo estímulo vai iniciar um processo de conhecimento *sentido* do estímulo, fazendo com que o sujeito centre a sua atenção não apenas no organismo (proto-eu) mas na relação do organismo com o estímulo (através dos sentimentos), emergindo um eu-nuclear. Deste modo, as modificações que o estímulo causou no corpo do sujeito levam a que este conheça o seu corpo através dos sentimentos que emergem das modificações fisiológicas conduzindo, subsequentemente, a um sentimento de propriedade do corpo (Rosa, 2018). O que está a fornecer algo mais ao processo mental simples do proto-eu é, segundo Damásio, uma série de imagens sentidas – imagens do organismo vivo

(proto-eu) e imagens do organismo vivo a serem alteradas pelas reações emocionais provocadas pela interação com o objeto. Neste sentido, o sentimento de conhecer contribui para o surgimento do eu-nuclear. Esta é a consciência que mais temos presente através da relação que estabelecemos com o meio (2010, 252-260).

Espinosa defendeu uma ideia semelhante ao postular que era pela forma como o corpo era afetado que a mente surgia (Espinosa, 1677/1992, pp. 277-283). Nas suas palavras: “Ora como a alma, pelas ideias do Corpo, tem necessariamente consciência de si mesma, tem consciência do seu esforço (p. 277). Para o filósofo de origem portuguesa, as ideias que surgiam na mente eram ancoradas na forma como o sujeito era afetado pelas modificações da interação entre o corpo e o meio, o que é algo muito próximo da interoção (Damásio, 2003). O eu-nuclear surge, neste sentido, quando um protagonista é acrescentado ao processo do proto-eu através das modificações sentidas no corpo, resultantes do contacto com o meio.

O eu-nuclear, no entanto, não tem capacidade para pensar no passado e projetar o futuro, pelo que é um eu refém do aqui e agora. O eu-nuclear é semelhante à *consciência primária* do neurocientista Gerald Edelman – uma consciência em que um indivíduo tem experiência perceptual e fenomenológica do momento, mas sem capacidade para a categorizar ou relembrar do ponto de vista simbólico ou semântico. A consciência primária de Edelman é um primeiro estágio de consciência que pode ser encontrada em muitos organismos com sistema nervoso mas à qual não foi acrescentada uma consciência elaborada, que corresponde à consciência autobiográfica de Damásio (Edelman, 1995).

Ora, se o eu-nuclear apenas consegue estar no aqui e agora, os seres humanos, como animais capazes de reter informação do passado e de projetar o seu futuro, terão de possuir aquilo que Damásio designa por consciência autobiográfica – o terceiro nível de consciência que Damásio sugere. Enquanto o proto-eu é a experiência do corpo vivo (conhecido pelos sentimentos) e o eu-nuclear corresponde às transformações do corpo vivo na relação com o meio (possível através das modificações realizadas pelos sentimentos de emoção), o eu autobiográfico é a memória das experiências passadas e a antecipação dessas experiências no futuro.

O eu autobiográfico leva uma vida dupla, segundo Damásio. Pode surgir como imagem recordada ou antecipada (em forma de memória ou de antecipação do futuro)

ou pode funcionar nos bastidores, longe da consciência, à espera de poder aceder à luz (quando o sujeito vive no aqui e no agora) (2010, 263-264). A *memória autobiográfica* possui muitas semelhanças com a *consciência elaborada* de Gerald Edelman – uma consciência superior, ancorada na linguagem e na capacidade do córtex desenvolver formas de comunicação sofisticadas, categorizações e compreensão simbólica, o que permite ao seu detentor libertar-se, segundo Edelman, “da tirania de um quadro temporal imediato ou dos acontecimentos concomitantes que ocorrem em tempo real”, possibilitando deste modo que o sujeito possa viajar para o passado e projetar-se no futuro (1995, p. 192).

Ao contrário dos outros níveis de consciência, o eu-autobiográfico depende de forma crítica da memória, pois é no alojamento de informação e de recordações que o sujeito ergue o seu sentido de identidade, embora não necessariamente de personalidade e temperamento, que estão bem mais relacionados com os outros dois níveis de consciência. Do ponto de vista neuronal, Damásio sugere que os núcleos associativos do tálamo, o hipocampo, o claustró e o córtex posteromedial (a conjugação do córtex cingulado posterior, o córtex retrosplenial e o precuneus) são fundamentais para o tipo de consciência alargada devido à função que têm “na coordenação do conteúdo que define o eu autobiográfico” (2010, pp. 270-286).

Em suma, Damásio sugere três níveis de consciência: 1) uma consciência vital baseada nos sentimentos de fundo, relacionados com a homeostasia – o proto-eu; 2) uma consciência nuclear, fundamentada no conhecimento do meio através das modificações corporais, que permitem o conhecimento sentido no eu e do meio – o eu nuclear; 3) uma consciência identitária, na qual estão depositadas memórias, histórias e desejos - o eu-autobiográfico. Os sentimentos alimentam todos estes níveis de consciência e permitem o conhecimento, tanto do meio como do eu.

O corpo como unificador da consciência

Uma das características da consciência é ser um processo integrado e unificado. Neste sentido, um dos problemas em torno da consciência tem que ver com como é que ela surge e existe de forma unificada – como é que o sujeito possui uma experiência integrada.

As teorias cognitivistas, por exemplo, tentam explicar como surge a consciência mas são omissas em relação a como é que essa experiência é unificada. Enquanto a GNW defende que a consciência surge no córtex frontoparietal, sendo omissa quanto à forma como esse sistema cria uma experiência integrada, a IIT postula que a consciência emerge do processamento de informação no córtex posteromedial sem, no entanto, revelar como é que esse processamento de informação consegue criar uma experiência unificada.

António Damásio, por seu turno, oferece-nos uma resposta plausível para este problema em *O Livro da Consciência* (2010). Segundo o neurocientista, são os sentimentos de *um* corpo unificado que permitem a estabilidade necessária à consciência para que esta permita uma experiência integrada. Como temos vindo a referir, sem sentimentos a consciência seria algo vazio, desprendido, desintegrado, sem um *ownership*. Os sentimentos de um corpo – e que emergem de *um* corpo - oferecem a estrutura integrada que suporta o eu. Por isso o tronco cerebral é crítico para a consciência – porque oferece ao cérebro uma *cartografia unificada* do estado do corpo. Dito de outra forma, o tronco cerebral, o tálamo e a insula, ao descodificarem sinais do corpo em informação neuronal, permitem que o corpo se apresente como *uno* de modo a que abarque nessa *unidade* a consciência (Damásio, 2013).

Como o próprio Damásio refere, nenhuma ideia é mais importante do que: “o corpo é o alicerce da mente consciente” (2010, 39). A mente é feita de imagens e a consciência é o *sentimento* de que essas imagens me pertencem. Mas para que esse sentimento surja – isto é, para que eu sinta que sou dono da minha mente – é necessária uma referência relativamente estável. Essa referência é o corpo e a forma como ele integra a consciência é através dos sentimentos homeostáticos. Esta estabilidade oferecida pelo corpo à mente permite a esta produzir um *duplo neuronal do corpo* – um género de eu “virtual” incorporado.

Damásio postula, na esteira de Espinosa (1677/1992, p. 277¹⁹), que todas as interações que o sujeito tem com o meio são feitas pelo corpo, pelas sensações do corpo e pelos sentimentos do corpo, que não só catalogam e qualificam a experiência, como a

¹⁹ “Ora como a alma, pelas ideias das afecções do Corpo, tem necessariamente consciência de si mesma, tem consciência do seu esforço” (p. 277).

unificam e estabilizam. Neste sentido, todos os estímulos com que o sujeito lida e recorda espoliam os ajustamentos motores necessários da experiência original, pelo que quando o sujeito recorda ou se depara com um estímulo, o corpo está sempre presente em forma de imagens mentais sentidas (2013, 184). Posto de outra forma: quando um sujeito se depara ou recorda um estímulo, o que surge na mente são imagens das modificações do corpo quando esse estímulo afetou o organismo. Por isso o corpo serve de intermediário entre o sujeito e o meio – um género de filtro segundo o qual o sujeito acede à realidade e ao conhecimento do mundo, como Espinosa sugeriu (1677/1992).

Em suma, a consciência é um processo integrado e unificado, e um dos desafios teóricos é explicar essa unificação. Modelos como a GNW e a IIT propõem locais cerebrais para o surgimento da consciência, mas não explicam como a experiência se mantém integrada. António Damásio sugere que a chave para essa integração é o corpo e os sentimentos que dele emergem, os quais fornecem estabilidade e uma referência contínua à mente consciente. O tronco cerebral, ao processar sinais do corpo, cria uma cartografia unificada que ancora a consciência.

Virtudes da teoria de Damásio

António Damásio parece oferecer-nos uma robusta teoria da consciência que, apesar de contraintuitiva, fornece algumas respostas para alguns problemas filosóficos. Desde logo, a teoria de Damásio assenta num conjunto de situações bem documentadas que dão força à sua hipótese. Vejamos algumas dessas situações.

Pacientes com anosognosia²⁰, por exemplo, ficam incapazes de compreender o grave estado em que se encontram porque, segundo Damásio, o seu cérebro perde a capacidade de produzir representações de mapas do corpo. Ao não terem acesso aos mapas do seu corpo, eles perdem o acesso aos sentimentos de emoção, o que diminui drasticamente a sua capacidade de *qualificarem* eficazmente a situação em que encontram. Dito de outra forma, estas pessoas perdem, em parte, o sentimento de

²⁰ Anosognosia é a condição neurológica em que o indivíduo não tem consciência de um défice ou incapacidade que possui, como a paralisia de um membro ou a cegueira. É comum em lesões do hemisfério direito, especialmente no lobo parietal. A pessoa nega ou ignora a sua limitação, mesmo perante evidências claras (Damásio, 2003 e 2010).

corporeidade e, subsequentemente, têm uma fenomenologia alterada que afeta a sua percepção consciente do que se está a passar. Por conseguinte, ao perder sentimentos de fundo, o sujeito perde igualmente acesso aos marcadores somáticos, os quais são cruciais para o sujeito ser informado que possui um corpo e qual o estado desse corpo, como ainda perde acesso ao catálogo emocional que permite tomadas de decisão funcionais, flexíveis e adaptativas (Damásio, 2010, 2011; Rosa, 2018). Por conseguinte, com uma fenomenologia alterada, estas pessoas têm uma percepção do corpo distorcida, o que as incapacita de avaliar com rigor a sua situação clínica extremamente dramática.

Na síndrome de Cotard, a que já aqui aludimos (em que se observa atrofia do córtex insular e do córtex frontoparietal, ambos responsáveis por sentimentos do corpo), as pessoas desenvolvem um género de depressão profunda em que deixam praticamente de sentir emoções, ao ponto de considerarem que estão mortas, podres ou que não existem. Não têm qualquer problema na linguagem nem no pensamento lógico, mas estão completamente comprometidos no sentimento de emoções, nomeadamente dos sinais que vêm do corpo (Ananthaswamy, 2015). A incapacidade de sentirem emoções, nomeadamente sentimentos do corpo, afeta o sentimento do *self* – a fenomenologia fica severamente afetada -, o que os leva a acreditar que estão mortos. Com uma fenomenologia alterada, a percepção que têm de si é extremamente distorcida. Faz sentido. Sem capacidade de *sentirem* que existem, duvidam que, de facto, existem. A síndrome de Cotard parece sugerir que, de facto, o sentimento do corpo, a par da capacidade de sentir emoções como defende Damásio, parece ser crucial para um *self* unificado e sentido. Uma pessoa com o síndrome de Cotard pode ser o exemplo típico de alguém que pode *saber* que existe mas não *sente* que existe. Trata-se, portanto, de um exemplo paradigmático que demonstra como a afetação da fenomenologia compromete inevitavelmente a percepção subjetiva da própria existência.

A doença de Alzheimer, um problema dramático em que se assiste a uma quebra gradual dos níveis de consciência, pode ser utilizada para explicar quais os sistemas neuronais que melhor contribuem para a experiência subjetiva. Por exemplo, existem vários sistemas que não são afetados pela doença, como o cerebelo ou os córtices sensoriais, pelo que estes parecem não ser críticos para a subjetividade. No entanto, o córtice insular, os córtices postero-mediais e o núcleo parabraquial (do tronco cerebral), por exemplo, são severamente afetados, o que parece sugerir que estes sistemas (entre

outros) são necessários para a experiência consciente, como sugere Damásio. Curiosamente, tanto a insula como o núcleo parabraquial são sistemas de interocepção, isto é, são áreas neuronais dedicadas aos sentimentos e emoções, mas também ao sentimento do corpo e à homeostasia (Damásio, 2010, p. 290; Seth *et al.*, 2012), o que parece dar validade à teoria damasiana.

Para além destas doenças em que a fenomenologia e a percepção de si e da realidade são afetadas, existem outras doenças em que podemos observar a necessidade de um corpo para o fenómeno da consciência. Nas doenças que provocam ausências de consciência, por exemplo, e que levam ao coma, é possível verificar que o cérebro processa informação. A mente está ativa e muitos destes pacientes exibem alguns comportamentos simples. Contudo, a sua consciência não existe. Os pacientes quando muitas vezes recuperam a consciência confirmam que durante o período em que estavam em coma não sentiram rigorosamente nada. Era como se estivessem mortos. Mas, no entanto, as suas mentes estavam, muitas vezes, ativas (Damásio, 2013, 121-136).

Isto parece sugerir que para uma pessoa estar consciente é necessário que a integração e o processamento de informação transmitam informação do corpo e não apenas sinais do cérebro desligado do resto do corpo. É preciso algo mais do que a capacidade de um organismo produzir e integrar informação cerebral, como sugere a IIT (Tononi, 2004; Koch *et al.*, 2016). Segundo António Damásio, é imperativo que a informação do corpo – realizada, em particular, através do tronco cerebral - seja decodificada em sentimentos de fundo. Dito de outra forma, retire-se o sistema que decodifica informação proveniente do corpo e a consciência desvanece-se (Damásio, 2003, 2010, 2011, 2013, 2017; Seth, 2021).

Por isso, precisamente devido a essa ligação entre o corpo e cérebro realizada pelo tronco cerebral, qualquer sujeito com uma lesão neste sistema, em particular no núcleo parabraquial ou na formação reticular, entra em coma. Porquê? Segundo Damásio, porque o cérebro deixou de receber informações sobre o seu elemento agregador e estabilizador (o corpo), pelo que deixa de conseguir gerar mapas do corpo, cuja representação em forma de sentimentos é necessária, tanto para o sentimento de si como para a criação das imagens mentais unificadas. Sem corpo, a mente consciente

desagrega-se, simplesmente porque sem informação do corpo não existem sentimentos de fundo que permitam o seu suporte homeostático²¹.

Em suma, sem os sinais do corpo oferecidos pela cartografia unificada do tronco cerebral, o cérebro torna-se incapaz de integrar os sinais necessários ao surgimento coeso da mente consciente através dos sentimentos. Em consequência, a consciência desliga-se — a luz apaga-se (Damásio, 2010, p. 304).

21 Recentemente, alguns autores defenderam que os NCC parecem residir num sistema muito próximo do proposto por Damásio, mas também pela GNW e pela IIT (Nani, 2019; Boly *et al.* 2017). Segundo alguns estudos recentes, a consciência parece depender do tronco cerebral, do tálamo (proto-eu damasiano), do córtex cingulado (eu-nuclear damasiano), do córtex frontoparietal (GNW) e do córtex posteriomedial (eu autobiográfico damasiano e IIT). Em grande medida, este estudo recente corrobora na íntegra o modelo damasiano. No entanto, não rejeita os modelos cognitivistas que possuem, igualmente, validade e podem ser integrados no modelo de Damásio. Este modelo (Nani, 2019) defende que o córtex frontoparietal (que Dehaene e LeDoux atribuem à consciência fenomenal) é responsável pela atenção, ao passo que o córtex posteriomedial (que Tononi e Koch associam à consciência fenomenal) está associado à experiência fenomenológica.

2.Limites da consciência na ação

“...os homens julgam-se livres apenas porque são conscientes das suas ações e ignorantes das causas pelas quais são determinados...”

Espinosa

Em relação ao problema o que é a consciência, vimos que, apesar de existirem vários modelos que a tentam explicar e compreender de forma distinta, existe um consenso relativamente alargado que considera que a consciência é o eu (ou *self*) – é a *experiência subjetiva e integrada* do sujeito (Damásio, 2010; Gazzaniga, 2011; Eagleman, 2012; Dennett, 1991; Wegner, 2002; Chalmers, 1995; Nagel, 1974; Tononi *et al.*, 2016; Koch *et al.* 2016; Dehaene, 2014; Baars, 2005). No que respeita à questão sobre *como* surge a consciência, vimos também que, apesar de estarmos longe de uma unanimidade, existe pelo menos um consenso num grupo amplo de autores fisicalistas que advogam que a consciência é um processo físico emergente *causado* pelo cérebro em sistemas neuronais, como os córtices frontoparietais (Dehaene, 2014), os córtices posteromediais (Tononi *et al.* 2016; Koch *et al.*, 2016) ou o tronco cerebral (Damásio, 2010).

O problema que propomos agora aprofundar refere-se à função da consciência. Em particular, pretendemos compreender se a consciência possui algum poder causal — isto é, se pode efetivamente influenciar decisões e iniciar ações. A questão central consiste, portanto, em saber se a consciência é apenas um epifenómeno de processos cerebrais automáticos ou se, pelo contrário, desempenha um papel ativo na orientação do comportamento. Esta questão é central para percebermos qual a liberdade que possuímos para iniciar e conduzir as nossas escolhas.

Num período em que o debate sobre se a alma controla o corpo tornou-se obsoleto e de certo modo esvaziado, o problema tem vindo, mais recentemente, a gravitar em torno de saber se a consciência – o eu – controla as ações e o comportamento (Libet, 1999; Wegner, 2002; Waggard, 2005 e 2008; Gazzaniga, 2011; Eagleman, 2012; Caruso, 2012; Kane, 2005; Jeannerod, 2006; Pockett *et al.* 2006; Chisholm, 2003; O'Connor, 2000; Morsella *at al.*, 2022). Se a consciência (ou o eu) causa ações, isso poderia significar que o ser humano possui livre-arbítrio, já que seríamos nós – o nosso eu – a *causar* a decisão. Por conseguinte, o problema central que discutiremos de seguida prende-se em saber se a consciência *causa* comportamento ou, melhor dizendo, se o eu comanda o corpo a agir.

Com o objetivo de aprofundar essa compreensão é necessário analisarmos, brevemente, o que entendemos por decisão e por ação.

O conceito de *decisão* é diferente do de *ação*, não obstante serem complementares. A decisão implica, obrigatoriamente, a necessidade de uma escolha livre. Como refere o psicólogo Daniel Kahneman em *Thinking Fast and Slow* (2011), a decisão é o processo cognitivo pelo qual um indivíduo escolhe uma opção entre várias alternativas disponíveis. Esse processo pode envolver avaliação de informação, análise de consequências e seleção da melhor ação para atingir um objetivo específico. Decidir envolve, portanto, uma escolha entre várias alternativas (Gallagher, 2005 e 2006). O sujeito *decide*, livremente, fazer X e não Y, pelo que a *decisão* está assente na possibilidade de se decidir intencional e conscientemente entre, pelo menos, duas opções (fazer ou não fazer). Caso contrário – se apenas existisse uma opção – o sujeito não teria que decidir, teria apenas de *agir*. Não existiria liberdade, já que o sujeito era *obrigado* a agir apenas de acordo com uma opção. A decisão pressupõe, portanto, a faculdade de *optar* voluntariamente por uma escolha em detrimento de outra (Mele, 2003).

A decisão é, por isso, uma faculdade que está dependente da consciência (do eu) - isto é, da capacidade do sujeito em *conhecer* as opções que tem ao seu dispor e de *escolher* livremente uma delas. Se o sujeito não tiver consciência do que decide, nunca poderá ocorrer verdadeiramente uma decisão já que, ao não saber o que decide, ele está limitado a agir de uma forma totalmente cega. Por conseguinte, a decisão só é livre se for causada pelo eu que *escolhe* a decisão e não por processos que fogem ao controlo do eu. Como refere o filósofo Gregg Caruso (2012) na sua obra *Free Will and Consciousness*, julgamo-nos livres quando 1) somos nós que escolhemos uma solução de um conjunto de possibilidades que livremente não escolhemos; 2) a natureza e origem das nossas decisões é o eu consciente e não nada que não controlemos (2012). Neste sentido, quem decide somos nós – o eu - não o nosso cérebro. Se o nosso cérebro decidisse sem que tivéssemos *conhecimento* dessas decisões, então, dificilmente poderíamos dizer que seríamos livres para decidir, já que as nossas decisões e subsequentes escolhas ser-nos-iam impostas pelo nosso cérebro.

A ação, ao contrário da decisão, está sobretudo relacionada com a execução motora, que pode ou não decorrer de um processo introspetivo. Enquanto a decisão envolve predominantemente processos internos e deliberativos (Kahneman, 2011), a ação refere-se à dimensão externa e motora do comportamento (Mele, 2003; Gallagher,

2005, 2006). Ao contrário da decisão, que, como referido, tem de ser uma escolha consciente, a ação pode não ser consciente, já que é unânime que muitas ações são automáticas (ou instintivas), como fugir de um carro em movimento ou retirar a mão de um forno a esquentar. Estes comportamentos não são escolhas livres, são, isso sim, ações automáticas. No entanto, se excluirmos as ações automáticas, a ação está relacionada com o resultado consciente de uma decisão. O sujeito *agiu* porque *decidiu* agir. Neste sentido, no que concerne a ações conscientes (isto é, não automáticas), o sujeito primeiro decide (conscientemente) e, depois, age em função dessa escolha (Gallagher, 2005 e 2006).

Se a decisão consciente é aquilo que inicia a ação – se é ela que a causa – então isso implicaria que a consciência – o eu – é livre para agir sem qualquer intervenção de processos subconscientes. Isto é, se é o eu que dá origem à decisão, e se só somos livres quando é o eu que inicia a ação, então isso significa que é o eu – a consciência – quem decide, e não o nosso subconsciente, que nos escapa. Mas isto coloca-nos um problema. Como pode o eu existir sem processos subconscientes? Como pode a consciência existir *acausadamente*? Se, por um lado, defendermos que o eu existe sem causa, deparamo-nos com uma dificuldade ainda maior: como explicar que algo, como o eu, possa existir sem ser causado por nada? Se, por outro lado, defendermos que alguma coisa causa o eu, então aquilo que o eu pode causar foi, na verdade, causado por aquilo que causou o eu. Neste sentido, o eu apenas pode causar algo que foi previamente determinado pela causa que lhe deu origem.

Parece-nos, portanto, que o problema do livre-arbítrio não pode escapar a esta questão: ou as ações são causadas pelo eu, sem que o eu resulte de qualquer causa (como defendem, em termos gerais, os libertaristas); ou são causadas pelo eu, sendo que o eu é, por sua vez, causado por algo anterior, nomeadamente pelo cérebro, através de processos subconscientes que não estão sob o conhecimento nem controlo direto do eu (Serrado, 2025).

2.1. Determinismo, libertarismo e livre-arbítrio

O tema do livre-arbítrio é um dos problemas mais antigos da filosofia, tendo vindo a inquietar as mentes de inúmeros autores ao longo dos séculos, até aos dias de hoje. É um assunto que tem evoluído pouco ao longo do tempo sendo, também, uma temática teorizada, a nosso ver, em torno de conceitos mal definidos, o que por vezes dificulta a discussão filosófica e a evolução do pensamento.

De uma forma geral, podemos designar o problema do livre-arbítrio, pelo menos sob um ponto de vista ontológico, como saber se o sujeito (o eu) é determinado ou não nas suas decisões e ações. Sob uma perspetiva canónica (que nos parece ter lacunas), se o sujeito for determinado por alguma coisa, dificilmente podemos dizer que é livre (determinismo) (Caruso, 2012; Honderich, 1988; Double, 1991); mas se *não for* determinado por nada, isto é, se as ações do sujeito não forem causadas por nenhum evento anterior nem pelas leis da Natureza, então podemos afirmar que ele é livre (libertarismo) (Ginet, 1962 e 2014; O'Connor, 2000; Kane, 2005). Assim, o livre-arbítrio tem sido tradicionalmente entendido, do ponto de vista canónico, como a ideia de que o sujeito só pode ser livre se, e apenas se, nada causar a sua decisão e o comportamento que dela decorre, permitindo-lhe assim agir de outro modo do que aquele que as leis da Natureza determinariam num dado momento.

Assim, do ponto de vista canónico, tanto para o determinismo como para o libertarismo, o livre-arbítrio é *incompatível* com um universo determinista (Pereboom, 2007). Este é um ponto essencial sobre o qual estas duas escolas, apesar de antagónicas nas suas conclusões, convergem. Ou seja, quer os filósofos deterministas (como Honderich (1988), Caruso (2012), Double (1991), entre outros), quer os libertários (como Ginet (2014), Chisholm (2003), Kane (2005), entre outros), são incompatibilistas, na medida em que defendem que o livre-arbítrio não pode coexistir com o determinismo. Assim, os primeiros concluem que o livre-arbítrio *não* existe, dado que vivemos num Universo em que tudo está determinado por leis causais; os segundos, pelo contrário, sustentam que, se o livre-arbítrio existe, então o determinismo é falso e que, consequentemente, alguns eventos — nomeadamente os actos livres — não podem ser determinados por causas antecedentes, implicando, em última análise, a existência de eventos *acausais*.

Uma outra teoria, designada de compatibilismo²² (ou determinismo moderado), advoga que o livre-arbítrio pode existir num Universo determinístico desde que o sujeito possa agir de forma racional, de acordo com a sua vontade e desejos, sem coações. Ou seja, enquanto os incompatibilistas defendem que o livre-arbítrio é incompatível com um universo determinista, os compatibilistas sugerem que, se o ser humano tem controlo sobre as suas ações — se pode inibir comportamentos nocivos e tomar decisões ponderadas e racionais sem ser coagido —, então pode ter livre-arbítrio, mesmo num universo determinista (Hoefer, 2023). Em poucas palavras, enquanto o determinismo defende que o livre-arbítrio não existe devido à natureza causal do Universo, e o libertarismo postula que o livre-arbítrio existe e implica a falsidade do determinismo, o compatibilismo sustenta que o sujeito é livre mesmo num Universo totalmente determinístico, ainda que a metafísica da liberdade que cada escola defende seja substancialmente distinta — o que, a nosso ver, dificulta uma discussão clara entre o compatibilismo e as outras duas correntes.²³

Mais adiante, iremos sustentar uma proposta original: a de que o livre-arbítrio não apenas é compatível com um Universo determinista, como só pode existir num Universo regido por leis determinísticas — uma posição que, a nosso ver, se distingue de forma substancial das teses compatibilistas tradicionais.

2.1.1. *O determinismo*

O determinismo é o conceito segundo o qual todos os acontecimentos são resultado necessário de causas anteriores, em conformidade com as leis da Natureza

22 No entanto, os filósofos compatibilistas são ontologicamente deterministas (Dennett, 2006; Hobart, 1934; Searle, 2019). Apenas tentam compatibilizar o determinismo com o livre-arbítrio. Como refere o filósofo John Searle: as ações “são determinadas pelo nosso livre-arbítrio, mas estão ainda assim determinadas” (Searle, 2019, p. 285). Daniel Dennett, na mesma senda de Searle, defende que o determinismo é compatível com o livre-arbítrio porque é o determinismo, aliás, que nos confere a capacidade de escolher e de ter responsabilidade moral. Ou seja, Dennett postula que qualquer ser humano com um cérebro saudável tem responsabilidade moral porque é precisamente a biologia neuronal que confere a responsabilidade moral (2003). Por conseguinte, poderíamos afirmar que o ser humano é livre porque é determinado pelo seu livre-arbítrio.

²³ Parece-nos que o libertarismo defende um livre-arbítrio *acausal*, fundamentado em algum tipo de indeterminismo que permite ao ser humano, em grande medida, contornar as leis da Natureza. O compatibilismo rejeita veementemente este tipo de livre-arbítrio e defende, em grande medida, precisamente o mesmo tipo de livre-arbítrio dos filósofos deterministas — a ideia de que o ser humano está subordinado às leis da Natureza e que apenas pode ser livre se, e se só se, desenvolver comportamentos racionais e responsáveis, mesmo que completamente determinados.

(Caruso, 2012). Mas o determinismo é mais do que um conceito e muitos autores desenvolvem teses fundamentadas usando o determinismo como teoria. Espinosa, por exemplo, utiliza o determinismo para desenvolver a sua ontologia monista, mas desenvolve, igualmente, uma teoria filosófica determinista, em que tudo é determinado pelas leis da Natureza (1677/1992). De forma análoga, o físico Gerard 't Hooft recorre ao determinismo para desenvolver uma teoria determinista da física quântica, baseada na ideia de um *celular automaton* subjacente à realidade física (2016). Por conseguinte, iremos utilizar ao longo deste trabalho o conceito de determinismo para identificar uma teoria filosófica que se pronuncia sobre a ontologia e o funcionamento do Universo (Earman, 1986).

No princípio do século XIX, o matemático Pierre-Simon Laplace desenvolveu uma ideia, mais tarde designada por Demónio Laplaciano, que defendia que, em teoria, seria possível determinar o futuro se um intelecto tivesse conhecimento de todas as leis da Natureza. Nas suas palavras:

"Podemos supor que, para um intelecto que conhecesse todas as forças de que a natureza está cheia e a posição dos seres que a compõem, nada seria incerto para ele e o futuro, tal como o passado, estaria presente aos seus olhos.²⁴" (in Wegner, 2002, [s.p.]).

Esta ideia de Laplace tem sido frequentemente evocada para ilustrar os princípios do determinismo. Contudo, importa sublinhar que o determinismo não é uma teoria de natureza epistemológica. É precisamente por isso que Laplace admite que, para se poder prever o futuro, seria necessário deter o *conhecimento* total da configuração física do Universo — isto é, todas as leis da Natureza, as posições e os movimentos de todos os corpos existentes. No mesmo sentido, a proposta de Laplace é meramente teórica: ainda que, em princípio, um intelecto onisciente pudesse prever todos os eventos futuros, tal capacidade está fora do alcance de qualquer ser real. As teorias do caos demonstram isso mesmo: a mínima perturbação num sistema pode originar resultados profundamente divergentes (Sautoy, 2018), tornando a previsão exata de estados futuros intrinsecamente impossível. Esta impossibilidade de previsão, porém, não invalida o determinismo: este não é — nem pretende ser — uma teoria

²⁴ Tradução do autor.

premonitória, como sugerem erradamente alguns autores (Kane, 2005, p. 8; Balaguer, 2010, p. 68; Kaku, 2021, p. 69), pois não se refere ao nosso *conhecimento* dos eventos, mas à existência de relações causais necessárias que regem o funcionamento do Universo (Earman, 1986).

Apesar de existirem várias definições de determinismo (Earman, 1986), podemos começar por uma formulação muito simples, formulada pelo vencedor do Prémio Nobel da Física de 1999, Gerard't Hooft, que utilizaremos ao longo deste trabalho: o determinismo é a teoria segundo a qual “todos os processos físicos são completamente controlados por leis” (2016, p. 5). Podemos, também, dizer que o determinismo é a teoria que defende “que todo o acontecimento ou ação, incluindo a ação humana, é o resultado inevitável de acontecimentos e ações anteriores e das leis da natureza” (Caruso, 2012, [s.p.]); ou, de forma equivalente, que “every event is necessitated by antecedent events and conditions with the laws of nature” (Hoefer, 2023, [s.p.]). O determinismo é, portanto, em poucas palavras, uma teoria que defende que *todos os acontecimentos e eventos são determinados por leis e por causas antecedentes* (Honderich, 1988; Hooft, 2017). Aliás, poderíamos remontar esta definição a Espinosa, para quem o determinismo consiste na ideia de que tudo o que existe ou acontece é causado pelas leis da Natureza (Espinosa, 1677/1992, p. 131).

O determinismo não é, portanto, como afirma equivocadamente o físico Michio Kaku, “a filosofia de que todos os eventos podem ser previstos com exatidão” (Kaku, 2021, p. 69). Previsão e medição são ferramentas científicas com propósitos epistemológicos — não ontológicos — frequentemente utilizadas como base metodológica da ciência, nomeadamente da física quântica (Hooft, 2016; Romeijn, 2022). Em contraste, o determinismo deve ser compreendido como uma filosofia ontológica, que se pronuncia sobre o modo como o Universo e as suas leis causam eventos — independentemente da nossa capacidade de os conhecer ou prever (Hooft, 2017; Bombarda, 1902; Espinosa, 1677/1992). De modo conciso, o determinismo pode ser visto como uma teoria que defende que causas determinam causas, não que podemos *determinar* essas causas. Assim, a ideia de previsibilidade — frequentemente invocada, também, por autores críticos do determinismo (Kane, 2005, p. 8; Balaguer, 2010, p. 68) — não representa a essência desta posição filosófica. Com efeito, dificilmente um autor determinista (Double, 1991; Pereboom, 2007; Bombarda, 1902;

Espinosa, 1677/1992) advoga qualquer ambição premonitória. O determinismo limita-se a afirmar que, num sistema que se rege por leis naturais (ou físicas)²⁵, tudo o que acontece — as causas e o futuros dos eventos — é determinado por essas leis naturais, e apenas por elas.

Podemos postular o determinismo segundo a seguinte proposição:

- 1) O Universo e tudo o que nele existe, existe porque existem leis;
- 2) As leis do Universo causam eventos;
- 3) logo, tudo o que existe no Universo existe porque existem leis que causam eventos.

O determinismo defende, por conseguinte, que quando um evento (X) é *causado* por outro evento (Y), o primeiro evento (X) não poderia ter acontecido de forma diferente pela simples razão que o evento que o causou (Y) atuou de forma determinística, de acordo com as leis da Natureza (Pereboom, 2007). Por conseguinte, os filósofos deterministas acreditam que os nossos pensamentos, ideias e sentimentos — assim como a vontade, os desejos e os atos de volição — são causados por eventos precedentes, nomeadamente pelo funcionamento dos nossos neurónios, espoletado por estímulos externos e internos (Honderich, 1988).

Ora, se os nossos pensamentos, ideias e sentimentos — bem como a vontade, os desejos e a volição — são causados pelo funcionamento dos nossos neurónios, então, quando pensamos numa mesa de madeira em vez de numa cadeira de ferro, ou quando sentimos desejo de comer uma salada em vez de um bolo de chocolate, esses pensamentos e esses desejos não poderiam ser outros. O que os causou foi uma determinada organização neuronal e um determinado contexto — e não outros. Se a organização neuronal, bem como o contexto, fossem diferentes, então os pensamentos e desejos também seriam diferentes, o que demonstra que tudo o que acontece tem de acontecer como aconteceu, pelo simples facto de que o que causou o evento foram aquelas causas e não outras. Se X (uma determinada organização neuronal e um

²⁵ Podemos dizer que as leis da Natureza são as constantes e regularidades objetivas que existem na realidade, enquanto a física são as descrições dessas constantes e regularidades. Por conseguinte, a física é a interpretação das leis da Natureza. Não sendo conceitos sinónimos são ideias que se referem à mesma realidade.

determinado contexto) causa Y (o desejo de comer salada), então Y acontecerá sempre que X acontecer. Dito de outra forma, o determinismo defende que, se X ocorrer, Y ocorrerá de forma inevitável; se X não ocorrer, então, ou Y não ocorre, ou ocorre devido a outra causa (Honderich, 1988).

Note-se, contudo, que esta causalidade entre X e Y não é epistémica. Não se trata, como defendia David Hume (1739/2016), de uma mera percepção subjetiva de ligação causal. Trata-se de uma causalidade real, fundada em leis naturais que a ciência tem vindo a identificar e descrever com crescente precisão (Newton, 1687/1934; Einstein, 1905; Feynman et al, 1964; Rovelli, 2018; Smolin, 2019). A causalidade que o determinismo defende assenta em leis físicas e constantes naturais, as quais causam efetivamente eventos, que, por sua vez, originam outros eventos, num encadeamento contínuo. A causalidade determinística não é, também, uma causalidade assente numa semântica ou numa compreensão simbólica, apoiada numa matemática que tenta prever probabilidades, como a física quântica (Ismael, 2021). A causalidade entre X e Y é ontológica, com ou sem possibilidades epistemológicas ou teleológicas. Por esta razão, Y pode ser igualmente causado por Z, mediante outras circunstâncias (interiores e exteriores). Dito de outra forma, X causa Y no contexto A mas Z pode causar Y no contexto B. No mesmo sentido, a organização neuronal X pode causar Y mas a organização neuronal S pode, igualmente, causar Y. O que importa para o determinismo é que os eventos são *causados* por eventos naturais anteriores.

Posso estar, por exemplo, a ver televisão e, de repente, a televisão desliga-se devido a uma interrupção de energia (evento Y). Este corte de energia pode ser causado por um relâmpago (evento X) que atingiu a fonte de energia, tendo causado uma descarga elétrica. Mas o corte de energia (evento Y) também pode ser causado por um curto-circuito no sistema de alimentação (evento Z). Tanto o relâmpago (X) como o curto-circuito (Z) podem causar Y. Neste sentido, podem existir múltiplas causas possíveis para o mesmo efeito Y, e essas causas até podem ocorrer em simultâneo ou em momentos distintos. O que o determinismo não admite é que, uma vez verificada uma causa suficiente — por exemplo, X (o relâmpago que atinge a fonte de energia) ou Z (o curto-circuito no circuito de alimentação) —, o efeito Y (corte de energia) pudesse não ter ocorrido. Ou seja, dado X ou dado Z, Y era inevitável. Simplesmente, porque todos os eventos são *causados* por eventos anteriores (Honderich, 1988). Em poucas

palavras, num sistema físico como o Universo, as causas são necessárias, essencialmente, para que existam eventos.

No mesmo sentido, os autores deterministas sustentam que o alegado indeterminismo da física quântica não corresponde, de facto, a um verdadeiro indeterminismo (Hooft, 2016; Caruso, 2012; Serrado, 2024). O que se verifica, de facto, na física quântica são probabilidades, o que não deve ser confundido nem com determinismo nem com indeterminismo. Isto porque, se os eventos quânticos são *causados* por leis da Natureza (Rovelli, 2022; Smolin, 2019; Hooft, 2016), mesmo que não consigamos prever o seu comportamento, então, o *determinismo continua a ser verdadeiro*, e o indeterminismo é uma ilusão de percepção. O indeterminismo é, aliás, um conceito que, frequentemente, não dialoga bem com o determinismo. Isto porque o primeiro funda-se, na maior parte das vezes, em problemas epistemológicos, como probabilidades dos acontecimentos e aleatoriedade dos eventos, ao passo que o segundo fundamenta-se em problemas ontológicos, como as causas subjacentes aos eventos, independentemente da nossa capacidade de os prever. Este é um dos problemas que, a nosso ver, deve ser ultrapassado. O determinismo e o indeterminismo fundamentam-se em pressupostos distintos e, precisamente por isso, tratam de coisas diferentes. Por isso, a tentativa de utilizar o indeterminismo para combater o determinismo é estéril e infrutífera.

Os filósofos libertários, como veremos, fundam o seu pensamento — de forma mais direta ou indireta — na ideia de que a decisão do agente tem de ser totalmente livre, isto é, não causada por quaisquer eventos antecedentes nem pelas leis da Natureza (Kane, 2005; Ginet, 2014; Balaguer, 2004; O'Connor, 2000). Para que o livre-arbítrio seja possível, a decisão tem de ser *acausal*. Por isso, os filósofos libertários procuram, nos diferentes tipos de indeterminismo — como o indeterminismo simples, o indeterminismo causal ou o indeterminismo de causação agencial —, evidências de eventos *acausais* que permitam refutar o determinismo e, assim, sustentar a possibilidade do livre-arbítrio. Mas esta procura encerra um equívoco: ontologicamente, o indeterminismo não desmente o determinismo — limita-se a pronunciar-se sobre a *probabilidade* de um evento ocorrer de determinada forma, não sobre o que causou esse evento. A procura de uma ontologia indeterminista é, como veremos com maior detalhe, um dos maiores problemas das teorias libertárias.

Partiremos de uma premissa simples, mas que consideramos estrutural, tanto para a ontologia do Universo como para qualquer teoria coerente do livre-arbítrio: tudo o que existe na Natureza, para poder existir, tem de ser causado pelas leis da Natureza. Defendemos, portanto, ostensivamente uma causalidade ontológica que nos parece, não só necessária como inevitável. Podemos colocar esta premissa numa proposição, também ela, muito simples:

- 1) O Universo é regido por leis físicas;
- 2) Essas leis determinam causalmente todos os acontecimentos no Universo;
- 3) Tudo o que existe no Universo é, ou resulta de, um acontecimento físico;
- 4) Logo, tudo o que existe no Universo teve de ser causado.

A causalidade ontológica que propomos é o princípio segundo o qual tudo o que existe tem uma causa — não por convenção teórica, mas porque a própria estrutura do Universo assim o impõe. De acordo com tudo o que a ciência tem vindo a demonstrar, vivemos num Universo regido por leis físicas que determinam a sucessão dos eventos, e todos os sistemas observáveis — desde partículas elementares até galáxias — emergem da aplicação dessas leis. Esta afirmação não assenta apenas num raciocínio lógico-dedutivo, mas em evidência empírica robusta: em nenhuma observação científica foi alguma vez identificado um fenómeno físico sem causa, e todos os modelos explicativos de física, cosmologia ou biologia assentam em cadeias causais — desde a formação das partículas e dos átomos no pós-Big Bang à emergência dos sistemas vivos por evolução (Rovelli, 2018 e 2022; Hawking e Mlodinov, 2011; Krauss, 2017; Hooft, 2016; Smolin, 2019; Sapolsky, 2024). Como demonstrámos na proposição anterior, se todas as entidades e acontecimentos que existem são, em última instância, eventos físicos regidos por leis causais, então nada pode existir sem ter sido causado. Ainda que Hume (1739-2016) tenha tentado reduzir a causalidade a uma mera expectativa psicológica baseada na repetição de padrões, essa abordagem ignora o facto de que a própria ciência moderna pressupõe a eficácia causal das leis para explicar e prever com sucesso o comportamento dos sistemas físicos. A causalidade não é, pois, um hábito mental, mas uma estrutura fundacional do real, validada empiricamente e indispensável à consistência ontológica do Universo.

Comecemos, pois, por um simples exercício de lógica filosófica: existirá algo no Universo que não tenha sido causado? O que poderia existir sem uma causa? Os átomos existem sem causas? E as moléculas, os genes, as células, os neurónios? A água, as plantas, os organismos? Existem cérebros sem causas? E pessoas? Planetas, estrelas, galáxias? Se algo disto existe sem ter sido causado, então foi acausado. Mas se foi acausado, não teve origem, nem relação causal com qualquer outro elemento do Universo. E se assim fosse, como poderia esse algo emergir, persistir ou interagir num Universo causalmente estruturado? Num sistema onde tudo o que existe depende de encadeamentos causais, a simples ideia de algo existir sem causa é logicamente incoerente e ontologicamente impossível.

E quanto aos estados mentais — o desejo, a vontade, a volição? Poderão eles existir sem causa? Os filósofos libertários — quaisquer que sejam as suas variantes teóricas (Kane, 2007; Ginet, 2008; O'Connor, 2000) — sustentariam que é o agente (o eu) que causa as suas decisões, a sua vontade e a sua volição, sem que essas decisões tenham, por sua vez, sido causadas por qualquer evento anterior.

No entanto — refutaremos nós — alguma coisa causou, inevitavelmente, não apenas a existência do agente, mas também os seus desejos, a sua vontade, a sua volição e, em última análise, aquilo a que chamamos livre-arbítrio. Nenhum desses elementos se forma espontaneamente, ou emerge *ex nihilo*, como que por magia; todos exigem causas (ou leis) anteriores que os tragam à existência (Serrado, 2025). Acreditar que um estado mental possa surgir sem qualquer precedente causal é violar não apenas o princípio da causalidade, mas toda a coerência que sustenta o edifício científico e ontológico da realidade. Por isso, o determinismo sustenta que, se tudo o que existe foi causado, então também os nossos pensamentos e desejos, a nossa vontade, volição, decisões e comportamentos têm de ser causados por algo que os antecede — em conformidade com as leis da Natureza.

Por conseguinte, se admitirmos que tudo o que existe tem uma causa — mas os nossos pensamentos, crenças, decisões e comportamentos, a nossa volição e o nosso livre-arbítrio não têm qualquer causa, como sustentam diversos filósofos libertários — então enfrentamos um problema ontológico grave. É difícil — diria mesmo impossível — identificar, ou sequer conceber logicamente, eventos *acausais*, isto é, eventos que existam *ex nihilo*, esquivando-se às leis da Natureza que compõem — e são — o

Universo, tal como propõem autores como Ginet (2008), McCann (1998), Chisholm (2003), Kane (2005), Balaguer (2014) e O'Connor (1995). Esta dificuldade é visível, por exemplo, em certas tentativas de alguns libertários contemporâneos em compatibilizar causalidade e indeterminação. No artigo *Can an Indeterministic Cause Leave a Choice up to the Agent* (2014), Carl Ginet, por exemplo, diz-nos – num paradoxo ingênuo – que a Natureza pode causar eventos *acausados* (pp. 24-26). Nas suas palavras:

“But why could not entirely uncaused events be part of “the natural order (even of a thoroughly physical natural order)? Nature has held many surprises. I see no a priori reason for thinking it impossible that one surprise that nature might have in store for us is that some events are not caused, that empirical investigation will persuade us that the best account of the laws of nature dictates for some kinds of events that their antecedents do not causally produce them but only limit what the ensuing event will be to one or another of a certain proper subset of the logical or metaphysical possibilities” (p. 25).

Ora, ou a Natureza causa eventos (e podem existir eventos causados sem que os possamos compreender) ou não causa eventos; e se não causa eventos nunca poderia causar eventos *acausados*. Dizer que a Natureza causa eventos *acausados* é defender que a Natureza causa eventos, pelo que não podem existir eventos *acausados* se estes foram causados pela Natureza. Sugerimos, por isso, que não podem existir eventos *acausais* num universo em que tudo o que existe exige uma causa para poder existir. Afirmar que o agente causa a decisão sem ser ela própria causada é apenas uma forma de deslocar a *acausalidade* — não de resolvê-la.

Tal como Ginet, as várias teorias libertárias têm dificuldades em sustentar logicamente a existência de eventos *acausais*. Veremos isto com mais detalhe mais adiante, mas, para já, parece-nos importante realçar a necessidade da existência de eventos causais, em contraste com as teses libertárias que defendem a sua inexistência. Se, como defendem os filósofos libertários causais (Kane, 2005; Balaguer, 2014), as nossas decisões têm uma causa aleatória e, subsequentemente, indeterminística, então, isso não invalida que não tenham sido *determinadas* – isto é, causadas – por meios que não conseguimos antecipar. Decisões aleatórias não conferem liberdade mas, independentemente disso, têm que ser causadas por alguma coisa. Se, por outro lado, como defendem os filósofos libertários não-causais (Ginet, 1962, 1997 e 2008; McCann,

1998), as nossas decisões são causadas pela nossa vontade e pela nossa volição mas nada causa a vontade e a volição, então, como explicar que possa existir vontade e volição? O que causou – o que trouxe à existência – a vontade e a volição? Se, como defendem os filósofos libertários de causação agencial (O'Connor, 2000; Chisholm, 2003), as nossas decisões são causadas por um eu imaterial — uma consciência não-física —, mas nada causa esse eu decidir, então coloca-se um problema fundamental: como explicar que esse eu tenha tomado uma decisão sem que nada o tenha causado a fazê-lo? Mais ainda, como justificar que uma entidade imaterial possa intervir causalmente no mundo físico, violando o princípio da causalidade natural, sem que ela própria esteja submetida a qualquer cadeia causal? Estas dificuldades — centrais nas teorias libertárias de base dualista — serão analisadas com maior detalhe mais adiante.

Por isso, de acordo com a necessidade de uma existência causal, segundo os filósofos deterministas, apenas podemos escolher entre duas afirmações: ou todos os eventos são causados (isto é, determinísticos), ou todos os eventos são *acausados*, (isto é, indeterminísticos). Não há fuga possível a esta dicotomia (Double, 1991). Caso sejam indeterminísticos, então: (1) ou são causados por leis da Natureza — sendo apenas a nossa capacidade de previsão limitada (indeterminismo epistemológico) — e, nesse caso, o determinismo permanece verdadeiro (Rovelli, 2022); ou (2) são efetivamente *acausais* (indeterminismo ontológico), o que é filosoficamente incoerente e metafisicamente insustentável. Com efeito, se um evento não for causado, terá de ser, obrigatoriamente, *acausado*. Se é *acausado*, como pode existir? Inversamente, se um evento não é *acausal*, então é causado — o que implica a validade do princípio da causalidade e, com ele, do determinismo. Negar esta relação implicaria aceitar que um acontecimento possa existir sem causa e, simultaneamente, sem ser *acausal*, o que constituiria uma contradição lógica e, portanto, um absurdo filosófico.

Imaginemos uma porta que se abre: ou foi causada por algo, ou abriu-se sem qualquer causa. Esta segunda hipótese, que implicaria a existência de um evento *acausal*, parece logicamente insustentável. Ora, se este princípio se aplica a objetos simples num Universo físico, por que razão não se aplicaria também às decisões humanas? Mesmo que se argumente que as decisões envolvem experiência fenomenológica, isso não as isenta das leis físicas. A tomada de decisão ocorre num

sistema físico — o Universo — e, como tal, tem de obedecer às mesmas leis que regem todos os outros fenómenos. Pensar o contrário seria admitir que os eventos *acausais* existiam por magia mas, nesse caso, a própria magia seria a causa desses eventos, pelo que deixariam de ser *acausais* (Slattery, 2014).

Em suma, o nosso raciocínio pode ser colocado, também, desta forma:

- 1) A existência implica causalidade;
- 2) se algo existe, foi trazido à existência;
- 3) isso vale para decisões, vontades, crenças ou até entidades como agente, consciência ou alma;
- 4) logo, toda a teoria que afirme que algo pode existir sem causa subverte a própria consistência ontológica da realidade.

Assim, se os nossos pensamentos, desejos e vontades têm de ser causados para poderem existir, então, segundo a conceção canónica de livre-arbítrio já analisada, não pode existir livre-arbítrio — pois não é o sujeito que os determina, mas sim um evento anterior (Double, 1991). Neste sentido, se o determinismo for verdadeiro, então cada ação humana, cada pensamento consciente (e inconsciente), cada decisão e cada desejo são necessariamente determinados por acontecimentos que ocorreram antes (Bargh, 2017; Pereboom, 2007; Caruso, 2012). Se isto for verdade, como responsabilizar os comportamentos? Se somos determinados a agir e se não podemos ter comportamentos diferentes daqueles que são causados por eventos antecedentes, então como responsabilizar alguém por uma ação criminosa? Como exigir a um criminoso que agisse de forma diferente, se foram precisamente as causas anteriores à sua ação que determinaram o seu comportamento? Como sustentar que ele poderia ter feito algo diferente do que estava determinado? Estas questões têm ocupado as mentes de muitos filósofos, juristas e neurocientistas (Eagleman, 2012; Sapolsky, 2024).

Nos capítulos seguintes, analisaremos com maior detalhe as principais teorias libertárias e o modo como estas recorrem ao indeterminismo ontológico (*acausal*) para justificar o livre-arbítrio. Argumentaremos, depois, que tais abordagens assentam em pressupostos ontologicamente incoerentes e em confusões conceptuais entre epistemologia, causalidade e estrutura da realidade.

2.1.2. O libertarismo

O libertarismo é uma teoria filosófica que defende a existência do livre-arbítrio com base na ideia de que o ser humano é livre para agir sem que as suas decisões sejam determinadas por eventos anteriores ou pelas leis da Natureza. Neste sentido, um agente é livre se, e apenas se, num dado contexto e perante os mesmos antecedentes causais — incluindo o seu próprio estado neurobiológico num determinado momento —, ele puder agir de forma diferente. Esta capacidade de agir de outro modo, apesar de todas as condições antecedentes e as leis presentes se manterem constantes, constitui, para os libertários, o critério essencial da liberdade (Kane, 2005 e 2007).

Para sustentar a ideia de que o sujeito pode decidir livremente, sem constrangimentos internos ou externos, os filósofos libertários recorrem ao conceito de indeterminismo ontológico — e não apenas epistemológico —, isto é, à tese de que certos eventos não são determinados pelas leis da Natureza e, por isso, são *acausais* (Morch, 2023). Entre esses eventos incluem-se, consoante a variante libertária, a volição, as vontades e os desejos (libertarismo simples), os fenómenos quânticos (libertarismo causal) ou a própria consciência (libertarismo de causação agencial).

Como veremos, alguns filósofos, como Ginet (1968) e McCann (1998), no âmbito do libertarismo simples (ou não-causal), defendem que existem eventos *acausais* e que decisões e atos de volição são exemplos de eventos humanos que não são antecidos por causas nem determinados pelas leis naturais. Outros, como Kane (2005) e Balaguer (2004), representantes do libertarismo causal, sustentam que existem eventos ontologicamente *acausais*, dos quais as decisões humanas emergiriam como manifestações de uma *acausalidade* fundamental, alegadamente evidenciada na física quântica. Neste sentido, o indeterminismo epistemológico, frequentemente mobilizado pela física quântica para descrever a imprevisibilidade de certos comportamentos, é interpretado por alguns filósofos libertários como evidência de um indeterminismo ontológico — ainda que essa leitura não seja sustentada por grande parte da comunidade científica (Rovelli, 2022; Smolin, 2019). Por fim, veremos que filósofos como O'Connor (2000) e Chisholm (2003), associados ao libertarismo de causação agencial, defendem que é o agente quem determina as suas ações, mas que nada determina o agente. Em última análise, este é concebido como uma entidade *acausal* — uma fonte de causalidade que não é ela própria causada por nada. Esta conceção,

contudo, implica admitir a existência de algo que causa sem ter sido causado, o que, como temos vindo a argumentar, é ontologicamente insustentável. Podemos resumir o libertarismo segundo a seguinte proposição:

- 1) O livre-arbítrio é incompatível com o determinismo;
- 2) temos livre-arbítrio;
- 3) logo, o determinismo é falso.

Apesar das diferenças entre estas correntes, todas partilham a mesma premissa de base: para que exista livre-arbítrio genuíno, as decisões humanas não podem ser determinadas por eventos antecedentes nem pelas leis da Natureza. Em todas elas, o agente é concebido como uma fonte originária de causalidade, capaz de iniciar ações de forma autónoma — ainda que isso implique aceitar a existência de eventos *acausais* ou propriedades não-físicas. Nos parágrafos seguintes, analisaremos com maior detalhe cada uma destas vertentes libertárias: o libertarismo simples ou não-causal, o libertarismo de causação agencial e o libertarismo causal.

2.1.3. *O libertarismo simples (ou não-causal)*

Em torno do libertarismo, o libertarismo simples é a posição filosófica que nos parece aproximar-se mais do conceito essencial de libertarismo: a ideia de que o sujeito pode iniciar as suas ações sem que estejam ou sejam determinadas por algum evento antecedente (Clarke, 2021). Segundo esta perspetiva, um evento mental, como uma razão ou uma intenção, pode iniciar uma ação e causar uma decisão sem que nada cause esse evento mental. Ou seja, a origem do comportamento humano residiria numa espécie de espontaneidade radical, onde o agente é visto como uma fonte primária de causalidade, não dependente de condições antecedentes nem de estruturas neurobiológicas. Os defensores desta posição, sustentam que certas decisões humanas surgem *ex nihilo*, sem qualquer base determinística, sendo o agente o único responsável pela emergência desses estados mentais. A liberdade, nesta visão, é maximizada: o sujeito é completamente autónomo na origem das suas ações.

Segundo o filósofo Carl Ginet (1962), qualquer ação humana é, ou começa, com um simples evento mental que não foi causado por nenhum evento anterior. Ginet defende que: 1) é conceptualmente impossível uma pessoa saber qual a decisão que vai

tomar antes de, efetivamente, a tomar; ora, 2) se as decisões fossem causadas por eventos antecedentes, então seria possível ao sujeito antecipar a sua decisão antes de a realizar. Como ambas as premissas são consideradas verdadeiras, conclui-se que a decisão não é causada por qualquer evento anterior (Ginet, 1962, 2014). No artigo *Freedom, Responsibility, and Agency*, o filósofo afirma: “Every action, according to me, either is or begins with a causally mental action, that is, a mental event that does not consist of one mental event causing others” (1997, p. 89).

Neste sentido, uma decisão corresponde, nas suas palavras, a “a flash or a bang” (1962, p. 54): algo que irrompe *acausalmente* na mente e que tem o poder de iniciar uma cadeia de ações livres, totalmente independentes das leis da natureza e de qualquer determinação anterior. Um evento mental – como uma razão ou uma vontade – é, para Ginet, uma experiência fenomenológica não causada, mas que causa. Isto é, uma experiência que, não sendo efeito de nada, tem o poder de gerar consequências. Este tipo de evento mental possui, segundo o autor, uma *actish quality* (1997, p. 89), ou seja, uma qualidade intrínseca de agência que distingue estes estados mentais de outras ocorrências psicológicas passivas. Esta qualidade de agir seria suficiente, por si só, para conferir ao evento mental a capacidade de iniciar uma ação voluntária.

É esta liberdade que confere a possibilidade daquilo que Ginet designa por *princípio das possibilidades alternativas* – a existência de múltiplos futuros que o sujeito pode livremente seguir. Este princípio só se sustenta se as ações do sujeito não forem, nas suas palavras, “causalmente determinadas por estados e eventos antecedentes” (1997, p. 86). Esta formulação permite ao agente aquilo que é uma das ideias centrais do libertarismo: a capacidade de fazer algo diferente daquilo que os eventos antecedentes – ou as leis da natureza – poderiam determinar. Por conseguinte, para Ginet, é o sujeito quem causa as suas ações, e não o encadeamento causal do Universo ou qualquer estado anterior ao momento da decisão (Ginet, 1997). Esta ideia de causalidade originária seria, mais tarde, objeto de crítica por parte do filósofo Harry Frankfurt (1969)²⁶, que propôs uma concepção compatibilista da responsabilidade moral, sustentando que um agente pode ser responsável mesmo que não possua alternativas de ação, desde que a sua decisão decorra de uma vontade própria, não constrangida.

²⁶ Trazido à discussão mais à frente.

Por outro lado, Ginet sustenta que a liberdade humana assenta na possibilidade de a Natureza permitir eventos acausados. De forma bastante ousada, defende que é plausível – e conceptualmente legítimo – que a Natureza possa causar eventos que ela própria não determina, como se certas ações humanas surgissem como “surpresas” da realidade. No seu artigo *Can an Indeterministic Cause Leave a Choice up to the Agent?*, afirma: “I see no a priori reason for thinking it impossible that one surprise that nature might have in store for us is that some events are not caused” (2014, p. 25). A decisão, a vontade ou a intenção de agir seriam, assim, exemplos de eventos que, embora ocorram no seio da Natureza, não são causados por ela.

Deste modo, a vontade de realizar uma ação não é algo que acontece ao sujeito – como se fosse desencadeada por forças exteriores ou internas –, mas, pelo contrário, é algo que o sujeito faz acontecer. A vontade, para Ginet, não é causada por nada, mas tem capacidade de causar. O simples ato mental de desejar levantar o braço ou mudar de canal é, nesta ótica, um acontecimento fenomenológico com estatuto ontológico originário: ele causa diretamente o movimento, sendo simultaneamente o início e o motor da ação. A experiência fenomenológica de querer agir é, por si só, o início do agir. Assim, se o sujeito S carrega no botão do comando, não há nenhuma outra causa para esse gesto senão o desejo de S de mudar de canal – um desejo entendido como evento *acausal*, não determinado por qualquer estado anterior do cérebro, do corpo ou do meio (Ginet, 1962).

Este modelo, ainda que conceptualmente coerente dentro da tradição libertária, exige uma ontologia na qual certos eventos mentais não apenas escapam ao domínio das leis naturais, como são dotados de um estatuto causal fundamental. Trata-se de um Universo no qual a liberdade é real porque não derivada: emerge diretamente do sujeito como fonte não derivada de causalidade – algo que, como veremos, levanta profundas questões filosóficas.

O filósofo Hugh McCann, na sua obra *The Works of Agency* (1998), defende uma conceção de liberdade que, à semelhança de Ginet, rompe com a causalidade antecedente como condição da ação. Para McCann, o sujeito é verdadeiramente livre quando a ação é um evento espontâneo, não derivado, e é, simultaneamente, intrinsecamente intencional — ou seja, quando existe volição. A volição, neste modelo, não é uma causa externa ou precedente da ação: ela é a própria ação. Assim, a liberdade

não decorre de um processo causal, mas da estrutura fenomenológica da ação enquanto manifestação direta da intenção do agente.

Segundo McCann, a ação não é desencadeada por fatores anteriores, sejam eles crenças, desejos ou estados mentais antecedentes, mas é um fenómeno primário, que se basta a si mesmo. Esta autossuficiência da ação livre significa que não há uma sequência de causas que leve à ação: a ação emerge, em vez disso, da intencionalidade volitiva que a constitui. A volição é um fenómeno teleológico, e não causal: ela dirige-se a um fim, carrega em si um objetivo, mas não depende de antecedentes para ser iniciada. É, por isso, um acontecimento intencional que possui em si mesmo a justificação do seu início e do seu curso. Como afirma o autor:

“Even where the intention out of which an action is performed precedes it [...] what makes it the case that the action is performed from that intention is not a causal relation but rather the intrinsic intentionality of the volitional activity on which the action is based” (1998, p. 128).

Esta perspectiva tem implicações profundas. Em primeiro lugar, rompe com qualquer concepção mecanicista da mente e do comportamento: não é uma cadeia de eventos que conduz à ação, mas uma estrutura intencional interna, imediata e autorrealizadora. Em segundo lugar, implica que a liberdade é irreduzível a componentes psicológicos ou *disposicionais*. Ou seja, a volição não é explicada por crenças ou pensamentos prévios — embora possa coexistir com eles —, pois a sua natureza fenomenológica e diretiva precede qualquer articulação discursiva ou racional da ação. Assim, neste modelo, a liberdade não é um epifenómeno, mas a própria origem da agência. A volição contém já um plano de ação e, ao ser exercida, realiza diretamente esse plano, sem mediações causais. A ação livre é, assim, um fenómeno fundamental: não precisa de ser justificada por causas antecedentes, porque é, ela própria, a expressão imediata da autonomia do agente. A atividade volitiva, ao contrário das concepções deterministas ou compatibilistas, não é um elo numa cadeia causal, mas o ponto de partida da ação humana.

Com isto, McCann rejeita não só o determinismo, mas também todas as teorias que reduzem a ação a uma sequência encadeada de estados mentais — sejam crenças, desejos ou intenções. Para ele, a liberdade reside no próprio ato de agir enquanto

expressão volitiva autónoma. Agir livremente é, portanto, iniciar uma sequência no mundo a partir de uma fonte originária: o próprio sujeito, entendido como entidade fenomenologicamente ativa, cuja liberdade se manifesta sem dependência de causas exteriores ou antecedentes.

Em suma, para os filósofos do libertarismo simples, nada causa o sujeito a decidir; pelo contrário, é o sujeito quem causa as suas próprias ações, seja pela decisão ou pela volição - eventos mentais que, como vimos, são *acausais*.

2.1.4 O libertarismo de causação agencial²⁷

O libertarismo de causação agencial pode ser considerado uma versão mais radical do que o libertarismo causal (que analisaremos mais à frente) e com um compromisso metafísico mais forte do que o libertarismo simples.

Durante séculos, a ideia de que a alma conferia livre-arbítrio foi comum (Descartes, 1637/2008). Mais recentemente, num período em que, no âmbito filosófico-científico, a defesa da existência de uma alma que controla o corpo se tornou menos comum, vários filósofos libertários procuraram encontrar justificação para a liberdade humana através de uma propriedade distinta no ser humano, nomeadamente uma consciência fenomenológica não-física. Deste modo, não seria na alma que poderíamos encontrar a explicação para a liberdade humana, mas numa especificidade própria do ser humano – a sua consciência imaterial (O'Connor, 2000; Chisholm, 2003).

Tal como outras versões do libertarismo, esta perspetiva defende que o sujeito pode, perante determinadas condições e eventos antecedentes, exercer livremente o seu comportamento — isto é, agir de forma diferente mesmo quando todas as variáveis causais e físicas se mantêm constantes, mesmo que isso viole as leis da Natureza. Isto é possível, segundo esta perspetiva, porque o agente é detentor de (ou é) uma consciência imaterial que, apesar de nascer da física, emerge como entidade superior e especial capaz de se autodeterminar de um modo completamente independente das leis da Natureza.

O filósofo Timothy O'Connor, em *Persons and Causes* (2000), advoga que o ser humano possui certas propriedades especiais, nomeadamente uma consciência

²⁷ *Agent-causal libertarianism*.

emergente e imaterial, distinta da matéria, que permite volição e, subsequentemente, causar eventos. Segundo O'Connor, a consciência é causada por processos físicos, no entanto, emerge como uma propriedade fundamental, distinta da física. Ou seja, na mesma linha filosófica do dualismo de propriedades, que, entre outros, David Chalmers (1995 e 1996) defende, O'Connor advoga que o mental e o físico são duas propriedades fundamentais do Universo. Como duas propriedades fundamentais, o físico e o mental, embora distintos, podem relacionar-se. O mental está imbuído no físico mas emerge do físico, adquirindo autonomia, da mesma forma que o físico pode ser autónomo da fenomenologia. Por isso, a consciência emerge da física mas adquire uma importância especial, já que se assume como um processo fundamental de outro nível, nomeadamente fenomenológico. Por outras palavras, a consciência é uma propriedade fundamental que surge através de uma estrutura física (o cérebro), mas como é emergente e não-física, existe acima dessa estrutura, como uma propriedade fenomenológica imaterial e independente (O'Connor, 2000a; Chalmers, 2006).

O ser humano é, segundo esta versão do libertarismo, uma entidade fundamentalmente diferente em relação aos outros animais. Possui uma natureza fenomenológica distinta. Da mesma forma que certos objetos possuem propriedades que causam eventos (a água causa certas coisas acontecerem, por exemplo), o ser humano, ao contrário dos objetos e de outros animais, possui propriedades fenomenológicas que conferem poderes causais, designadamente, uma fenomenologia mais desenvolvida que permite ao ser humano realizar decisões completamente livres. Isto é, segundo O'Connor, as pessoas são entidades biológicas mas "também possuem estados mentais emergentes" (O'Connor, 2000, p. 110), pelo que a mente consciente possui propriedades fenomenológicas que, apesar de nascerem da física, adquirem uma independência da física, que garante o livre-arbítrio.

Por isso, a consciência é um processo emergente que, apesar de estar dependente da física, influencia e condiciona o mundo físico, nomeadamente o cérebro. A consciência pode, portanto, controlar o cérebro porque é uma propriedade emergente independente. Assim, esta capacidade não-física da consciência confere liberdade ao sujeito, já que as propriedades emergentes da consciência – do eu - podem intervir nas propriedades físicas. Ou seja, segundo esta perspetiva, o ser humano possui (ou é) uma consciência distinta da matéria do cérebro, a qual, embora originada nos

processos físicos, adquire a capacidade de exercer causalidade própria, não redutível à causalidade física. Dessa forma, o agente pode iniciar uma nova cadeia causal, agindo sobre a matéria — nomeadamente sobre o próprio cérebro — sem, contudo, violar as leis da Natureza (O'Connor, 2000a). O ser humano pode, assim, intervir na cadeia de eventos porque, pertencendo a essa cadeia mas com qualidades especiais (nomeadamente uma consciência não-física), pode autodeterminar o seu comportamento ao intervir diretamente no seu sistema nervoso.

O filósofo Roderick Chisholm, em linha com a proposta de Timothy O'Connor, defende em *Human Freedom and the Self* (2003) que a consciência, ao emergir do cérebro, adquire a capacidade de atuar sobre o próprio cérebro — estabelecendo, assim, um *loop* causal. O sujeito — entendido como experiência fenomenológica — pode, nesta perspetiva, intervir nos seus próprios processos neuronais e, dessa forma, causar ações. Chisholm admite que é possível afirmar que o ser humano move a mão, e que esse movimento foi ativado por determinados músculos e circuitos do sistema nervoso. Contudo, sustenta que alguns dos eventos neuronais envolvidos “foram causados pelo agente e não por outros eventos” (2003, p. 31).

Neste modelo, o sujeito possui, portanto, a capacidade de intervir no seu próprio cérebro precisamente porque ele não é o cérebro: é uma entidade fenomenológica emergente que pode, de forma autónoma, controlar — ou ao menos influenciar — a atividade neuronal. Assim, as ações humanas não são causadas por desejos ou crenças, como sustentam os libertários não-causais (Ginet, 2008; McCann, 1998), pois o agente pode agir contra os próprios desejos e crenças, mesmo quando estes são originados no cérebro. Um desejo forte de comer um bolo de chocolate, causado por processos neuronais, pode ser livremente rejeitado pelo sujeito consciente, que opta por não ceder ao impulso. Ou seja, a decisão não é tomada pelo cérebro, mas por uma instância superior que o transcende: a consciência. O agente é, assim, o verdadeiro iniciador da ação, pois é ele quem desencadeia eventos neuronais que culminam em comportamentos livres (Chisholm, 1966).

Na teoria de Chisholm, mesmo que existam automatismos no comportamento, a liberdade é preservada se a origem do processo estiver numa decisão consciente inicial. A ação livre não depende de antecedentes causais, mas da intencionalidade fenomenológica que a inicia. Desse modo, o agente pode romper com a causalidade

natural, exercer controlo sobre o sistema nervoso e iniciar ações que se tornam automáticas apenas após terem sido originalmente voluntárias. Estas ações permanecem livres precisamente porque derivam de um ato volitivo fenomenológico e, por isso, não têm causa anterior — são, segundo Chisholm, *acausadas*. Nesta perspetiva, a liberdade não depende de eventos antecedentes, mas também não assenta num indeterminismo ontológico que introduz aleatoriedade ou sorte — como acontece nas outras variantes do libertarismo. Pelo contrário, a liberdade advém da capacidade da consciência de iniciar ações por si mesma, de forma intencional e autónoma. A consciência fenomenológica, enquanto propriedade emergente e não-física, é concebida como fonte originária de causalidade. A liberdade do agente resulta, assim, da sua aptidão para causar diretamente eventos cerebrais sem estar condicionada por antecedentes físicos ou leis naturais.

Nesta visão, o fator adicional que confere liberdade ao ser humano é a imaterialidade da consciência. Embora emergente de uma base física (o cérebro), a mente consciente possui propriedades que a transcendem e que lhe permitem intervir na realidade física — nomeadamente no funcionamento do sistema nervoso. O agente não é determinado pelas circunstâncias, porque, perante a mesma situação X, pode sempre escolher entre diferentes possibilidades, sem estar subordinado aos acontecimentos prévios. Dada esta soberania fenomenológica, a consciência tem o poder de decidir livremente, inclusive contrariando o seu próprio cérebro. Neste quadro teórico, a consciência pode, em última instância, violar as leis da física, precisamente porque é uma propriedade emergente, mas *acausal*, cuja existência não se submete às estruturas deterministas da Natureza (Chisholm, 2003).

Em suma, este tipo de libertarismo — o libertarismo de causação agencial, que poderíamos também designar como libertarismo dualista — defende, essencialmente, que o sujeito possui a capacidade de violar as leis da física e de, inclusive, controlar o seu próprio sistema nervoso, permitindo-lhe, assim, ter poderes causais únicos.

2.1.5. *O libertarismo causal*

O libertarismo causal, por seu turno, distingue-se dos outros tipos de libertarismo ao tentar encontrar uma explicação para o livre-arbítrio nas leis da física quântica, contornando a dificuldade, presente nos outros libertarismos em explicar

como é que existem eventos mentais *acausados* (libertarismo simples), ou como é que uma consciência não-física pode interagir com um sistema nervoso físico (libertarismo de causação agencial).

Segundo o libertarismo causal (Kane, 2005; Balaguer, 2004), o Universo funciona de modo indeterminista, uma vez que a física quântica tem vindo a demonstrar que os fenómenos fundamentais da realidade não obedecem a uma cadeia causal estrita. É com base nesta leitura da física quântica que o libertarismo causal tenta fundamentar a possibilidade do livre-arbítrio. Segundo esta perspetiva, certos eventos — como o colapso da função de onda — são interpretados como ocorrências ontologicamente indeterminadas, isto é, como eventos cuja manifestação não resulta de qualquer causa anterior. Assim, a física quântica não apenas demonstraria limites à previsibilidade (como seria o caso num indeterminismo meramente epistemológico), mas revelaria uma estrutura do Universo onde determinados eventos seriam genuinamente *acausais*. Este indeterminismo quântico, por seu turno, parte de três princípios (muito discutidos e discutíveis, diga-se desde já): 1) que o indeterminismo quântico é ontológico; 2) que o indeterminismo quântico é *acausado*; 3) que o indeterminismo quântico pode ser verificado ao nível dos sistemas neuronais (Kane, 2005; Balaguer, 2004).

Em *A Contemporary Introduction to Free Will* (2005), Kane defende, por exemplo, que a amplificação de “saltos quânticos no nível neuronal” (p. 134) poderia explicar o indeterminismo no disparo dos neurónios, permitindo futuros alternativos não determinados pelas condições iniciais. De forma semelhante, Balaguer, em *A Coherent, Naturalistic, and Plausible Formulation of Libertarian Free Will* (2004), sustenta que as decisões humanas poderiam resultar de um “enorme conglomerado de eventos quânticos” (p. 403) existentes no cérebro. Mesmo sem controlo direto sobre cada evento quântico, o sujeito poderia, segundo Balaguer, considerar como suas as decisões emergentes dessa agregação aleatória, desde que se identificasse com elas.

Tanto Kane como Balaguer sustentam, portanto, duas ideias fundamentais: por um lado, que o indeterminismo quântico comprova que o Universo é indeterminístico; por outro, que esse indeterminismo pode estender-se ao comportamento humano através de saltos quânticos nos neurónios. Com base nestes princípios, o libertarismo causal postula que, sobretudo nas decisões mais complexas (*torn decisions*), o indeterminismo quântico permitiria que as escolhas do agente não fossem

determinadas. Nestes casos — em que o sujeito está dividido entre duas ou mais alternativas igualmente justificáveis —, a presença de processos neuronais sensíveis à aleatoriedade quântica poderia permitir que a decisão final não resultasse diretamente de causas antecedentes, mas emergisse de um estado de indeterminação genuína, o que permitiria possibilidades alternativas. Assim, embora o agente se envolva racional e emocionalmente na deliberação, o desfecho da escolha conteria um elemento de indeterminismo que, segundo esta perspectiva, salvaguardaria a liberdade. Esta posição, no entanto, contraria não apenas as mais consagradas abordagens da neurociência (Damásio, 2010; Rosenthal, 2005; Gazzaniga, 2011; LeDoux, 2020; Swaab, 2014), mas também a interpretação de vários físicos que, mesmo aceitando o indeterminismo epistemológico, não o consideram um fenómeno *acausal* no sentido requerido (Rovelli, 2018; Smolin, 2019; Hooft, 2016; Hawking e Mlodinov, 2011).

Vejamos o processo sugerido pelo libertarismo causal com mais pormenor. Segundo o filósofo Robert Kane, quando uma pessoa delibera sobre alguma coisa importante como, por exemplo, se deve ou não mudar de casa ou de emprego, uma série de pensamentos, memórias e imagens surgem na mente do agente. Surgem cenários mentais, sentimentos, memórias e futuros antecipados, num processo em que o sujeito está a ponderar as suas decisões. Este processo pode demorar dias, semanas e meses, durante os quais a pessoa pensa em inúmeros cenários possíveis. Segundo Kane, este é um processo completamente indeterminístico – ontologicamente *acausado*. E é indeterminístico segundo o princípio da incerteza de Heisenberg²⁸ que Kane transporta para o cérebro humano, num processo de aleatoriedade e indeterminismo que, segundo o filósofo, nem sequer o Demónio Laplaciano poderia prever. Ou seja, o processo de deliberação do sujeito é indeterminístico porque é completamente imprevisível, pelo

²⁸ O princípio de incerteza de Heisenberg, formulado em 1927, afirma que não é possível conhecer simultaneamente, com precisão absoluta, a posição e o momento (quantidade de movimento) de uma partícula subatômica. Quanto mais precisamente se mede uma dessas grandezas, maior é a incerteza na medição da outra. Este princípio não resulta de limitações tecnológicas, mas sim da própria natureza quântica da realidade. Ele marca uma rutura com o determinismo clássico e introduz um limite fundamental ao conhecimento, já que demonstra que o comportamento de uma partícula não pode ser deterministicamente medido. Não é que a partícula "não tenha" posição ou velocidade, mas sim que essas propriedades não existem de forma definida antes da medição (Rovelli, 2022).

que mesmo alguém que estivesse na posse de todo o conhecimento e factos do Universo não poderia adivinhar o que essa pessoa iria deliberar.

Esta aleatoriedade permite, segundo Kane, que o sujeito possa ter múltiplos futuros alternativos já que, segundo o libertarismo causal, essa possibilidade é conferida pela aleatoriedade quântica, que o filósofo associa ao indeterminismo²⁹. Isto, no entanto, como é evidente, não garante qualquer tipo de liberdade, já que os pensamentos, memórias e sentimentos surgem de forma aleatória, o que Kane não rejeita. O que confere liberdade, no entender do filósofo, é a capacidade do sujeito em *escolher*, de forma consciente, com qual dos pensamentos, memórias, sentimentos, desejos e vontades quer agir. Ou seja, segundo Kane, não obstante o processo estocástico e indeterminístico pelo qual a mente quântica do sujeito gera aleatoriamente pensamentos, memórias e sentimentos, ele tem liberdade de escolher, num esforço consciente, com qual desses pensamentos, memórias e sentimentos ele quer decidir. Essa liberdade de escolher o pensamento X em vez do pensamento A, B ou C é, segundo Kane, o que confere o livre-arbítrio (Kane, 2005).

Neste sentido, o libertarismo causal de Kane confere a capacidade de, em determinadas circunstâncias e segundo as leis da física, o sujeito fazer o contrário do que estava determinado – tal e qual uma partícula quântica que, em estados de sobreposição, possui um estado indeterminado. Mas, podemos questionar, se os pensamentos são indeterminados e fruto do acaso, onde está a liberdade do sujeito? Para Kane, como referido, a liberdade reside na *escolha consciente* do sujeito a todos esses pensamentos indeterminísticos, que lhe confere controlo no processo de decisão. Isto é, o agente não tem a liberdade sobre os pensamentos, sentimentos e memórias que lhe surgem de forma indeterminística, porque simplesmente não os controla. Mas, segundo Kane, o agente pode controlar a reação a esses pensamentos, e assumir o controlo da decisão, tomando-a livremente sem que nada a causasse a não ser a decisão consciente do sujeito. A decisão do agente seria, portanto, uma decisão racional e responsável, completamente livre, e não arbitrária, já que ele assumiu o controlo do processo indeterminístico que foi utilizado no processo de deliberação. Isto confere um

²⁹ Veremos mais à frente que a aleatoriedade não é uma condição para um Universo ontologicamente indeterminista.

futuro aberto e não determinado, desde logo porque as escolhas indeterminadas que surgem na mente, por não serem determinadas, alargam as possibilidades de futuro. Kane reconhece que isto não confere liberdade mas permite a existência de futuros alternativos. A liberdade consegue-se, isso sim, quando o sujeito assume controlo, não do surgimento do processo indeterminístico dos seus pensamentos, que não controla, mas do posicionamento consciente a esses pensamentos em forma de um controlo livre sobre a decisão que quer tomar (Kane, 2005)³⁰.

Robert Kane dá o exemplo de uma mulher de negócios que, estando atrasada para uma reunião importante, depara-se com uma pessoa a ser assaltada numa viela. A mulher terá de escolher de acordo com o seu carácter, mas essa escolha é considerada livre e autodeterminada, pois é a vontade consciente do sujeito que determina a decisão. Segundo Kane, a mulher tem de fazer um esforço consciente para contrariar a tentação de seguir em frente, e tentar ajudar a pessoa que está a ser agredida. A tentação natural, diz Kane, seria prosseguir o seu caminho, e é precisamente por ter resistido a essa inclinação que a decisão de ajudar é uma ação livre e responsável — já que poderia ter feito o contrário mas escolheu agir segundo o que acreditava ser o mais correto. Estas decisões são o que Kane designa por *self-forming actions*, ações formadoras do eu, que conferem liberdade ao agente. No caso referido, a decisão moral de ajudar o outro, em detrimento de um interesse individual imediato, constitui um ato de autoformação, pois permite à mulher moldar a pessoa que quer ser, em função dos seus valores morais (Kane, 2007).

No caso da mulher de negócios, a decisão de ajudar foi inicialmente indeterminada, pois nasceu da aleatoriedade quântica, mas tornou-se *auto-determinada pelo sujeito* quando este escolhe, num esforço consciente, ajudar a pessoa que estava a ser assaltada. Desta forma, como refere Kane no seu livro *A Contemporary Introduction to Free Will*, “Ações livres podem por isso ser *acausadas* sem termos de

30 O filósofo Alfred Mele analisa o processo de tomada de decisão humana e sugere que a liberdade de ação está relacionada com a capacidade do agente de deliberar e agir com base nas suas crenças, desejos e intenções. Mele explora teorias libertárias que associam o indeterminismo ao livre-arbítrio, mas não defende que as escolhas ou os pensamentos surjam de forma indeterminada. Para Mele, a liberdade está associada ao controle racional do agente sobre as suas deliberações e decisões, em vez de depender diretamente de processos aleatórios ou indeterminísticos. Nesse sentido, a liberdade de um sujeito fundamenta-se na sua capacidade de avaliar diferentes opções e escolher de acordo com razões que refletem as suas intenções e objetivos (Mele, 2000).

recorrer simplesmente à sorte ou ao acaso [inerentes ao indeterminismo quântico]. As ações ocorreram por uma razão e um propósito³¹” (Kane, 2007, p. 54).

O filósofo Mark Balaguer, embora dentro da mesma escola libertária de Kane, tem uma posição um pouco diferente sobre o modo como o sujeito é livre. Ele defende, no seu artigo *A Coherent, Naturalistic, and Plausible Formulation of Libertarian Free Will* que as escolhas livres, nomeadamente as escolhas deliberadas, não são “determinadas por nada³²” (2004, p. 387). Para Balaguer, as decisões não são determinadas por razões ou crenças nem “são determinadas pelo estado físico dos nossos cérebros antes do momento da escolha; em vez disso, apenas fazemo-las³³” (2004, p. 387). Tal como Kane, Mark Balaguer defende que o cérebro é um sistema estocástico que funciona de forma indeterminista, de acordo com princípios quânticos que ele extrapola para as redes neuronais. Contudo, Balaguer tem uma perspetiva mais radical porque, embora Kane advogue que é o esforço consciente do sujeito que confere liberdade, Balaguer defende que a liberdade não é causada por nada mais do que 1) a existência de eventos *acausais* no processo de decisão e 2) o facto de essas decisões serem ações genuinamente humanas, ou seja, produzidas pelo próprio agente, ainda que sem qualquer causa anterior determinante.

De acordo com a perspetiva de Balaguer, imagine-se o José, que está indeciso entre comer arroz doce ou bolo de chocolate. Pensa e tenta deliberar sobre o que comer mas não consegue tomar nenhuma decisão até que opta, sem qualquer razão, por comer bolo de chocolate. O José decidiu apenas porque teve de decidir, e não por algum desejo maior de comer um em relação ao outro. Assim, ao contrário do que é defendido por outros libertários, como Kane ou Ginet (embora segundo perspetivas e processos distintos), não são necessariamente as razões (Ginet) ou o esforço consciente (Kane) que causam as ações livres. A ação de comer o bolo de chocolate, para Balaguer, não teve causa. Foi completamente indeterminada – o mesmo é dizer *acausada*. E foi indeterminada não apenas porque emerge do indeterminismo estocástico dos processos neuronais — cuja base última reside no indeterminismo quântico —, mas também porque o sujeito não escolhe conscientemente essa decisão no sentido

³¹ Tradução do autor.

³² Tradução do autor.

³³ Tradução do autor.

tradicional, e ainda assim se identifica plenamente com ela. Ou seja, embora o agente não tenha deliberado de forma determinista nem controlado causalmente o processo, ele considera que a decisão que emergiu foi sua. Nesse sentido, o agente “escolheu” a ação não por a ter causado, mas por a ter assumido como expressão autêntica de si, ainda que a sua origem seja completamente indeterminada. A liberdade, nesta perspectiva, não requer controle causal anterior, mas apenas concordância pós-factual com o resultado de um processo acausal interno.

Ao contrário de Kane, a liberdade, para Balaguer, não reside no poder de deliberação entre diferentes opções, mas sim no facto de o sujeito expressar um indeterminismo que emerge no seu interior. Ora, se esse indeterminismo nasce no próprio sujeito sem ter sido causado por fatores externos ou antecedentes, então isso confere-lhe liberdade de ação, desde que o sujeito aja de forma consciente e responsável. Neste sentido, as decisões livres são causadas por eventos mentais conscientes — isto é, pelo próprio sujeito — mas, sendo essas decisões indeterminadas no momento em que ocorrem, é o sujeito, enquanto entidade consciente que se identifica com elas, que as realiza. Em poucas palavras, as ações são livres quando não são causadas por quaisquer eventos anteriores e, simultaneamente, são materializadas pelo próprio sujeito (Balaguer, 2014, pp. 75–76).

Os filósofos libertários admitem que esta liberdade não é a liberdade que gostariam de ter, já que possui uma dimensão ampla de aleatoriedade, mas ainda assim é uma liberdade suficiente para garantir que o sujeito é livre para tomar as suas decisões, já que é ele que as escolhe através de um processo de controlo do indeterminismo quântico, seja através da vontade, de crenças e de razões (Kane, 2005), seja através de uma afirmação do sujeito de realizar uma ação sem qualquer razão (Balaguer, 2014).

2.1.6. *Libertarismo e física quântica*

A perspectiva metafísica do indeterminismo quântico aplicada ao livre-arbítrio libertário que acabamos de analisar nasceu no âmbito da interpretação de Copenhaga da física quântica e da sua (suposta) natureza probabilística. Ancorada nas contribuições de Max Born (Ismael, 2021), Werner Heisenberg (1958) e na interpretação probabilística de Niels Bohr (Bohr, 1958), a interpretação de Copenhaga postulou, de forma pioneira,

uma ideia central: a natureza da física quântica é probabilística – o mesmo é dizer indeterminada.

É considerado de forma unânime que o estudo da física quântica começa no início do século XX com Max Planck, secundado por Einstein que vence o seu Prémio Nobel nessa área devido ao seu estudo da natureza da luz (Rovelli, 2018; Smolin, 2019). Einstein foi fundamental na compreensão da luz como uma onda e como uma partícula (no contexto do efeito fotoelétrico), enquanto Louis de Broglie estendeu a dualidade onda-partícula para as partículas em geral (Rovelli, 2018 e 2022; Smolin, 2019). Ora, de acordo com a dualidade onda-partícula, um fóton – ou outra partícula – existe numa sobreposição em que existe tanto como onda tanto como partícula³⁴. Esta dualidade desaparece, ou quando a onda-partícula interage com algum sistema de medição ou quando um sistema quântico interage com um sistema clássico. Em ambos os casos, a função de onda colapsa e esta assume-se agora como uma partícula bem definida no espaço (Hawking e Mlodinov, 2011; Smolin, 2019).

Naquela que é, provavelmente, a experiência mais emblemática da física quântica — a experiência da dupla fenda —, fótons (ou outras partículas) são lançados em direção a uma barreira com duas fendas, antes de atingirem um anteparo. Quando os fótons são emitidos um a um, e nenhuma medição é feita para saber por onde passam, observa-se no anteparo um padrão de interferência típico das ondas, o que indica que cada fóton se comportou como se tivesse passado pelas duas fendas ao mesmo tempo – estava num estado de sobreposição de onda e partícula. No entanto, se se colocarem sensores para detetar por qual das fendas o fóton passou, o padrão muda radicalmente: o padrão de interferência desaparece e os fótons passam a comportar-se como partículas clássicas, formando duas faixas bem definidas no anteparo — uma por cada fenda. Isto significa que a simples medição colapsa a função de onda, fazendo com que o fóton assuma uma trajetória definida como partícula. Antes da medição, a evolução do sistema é descrita de forma determinista pela equação de Schrödinger, que governa a propagação da função de onda. Mas no momento da medição, o sistema já não pode ser descrito por uma única função de onda: ocorre um

34 De notar que alguns autores defendem que as partículas não se comportam como onda e como partícula ao mesmo tempo mas que, em algumas circunstâncias, as partículas podem comportar-se, ou como ondas ou como partículas (Gribbin, 2007; Berenguer, 2016).

colapso, e a realidade observável passa a ser expressa em termos de probabilidades, refletindo a limitação epistemológica imposta pelo ato de medir (Krauss, 2017; Rovelli, 2022).

Werner Heisenberg postulou que existe um limite para sabermos a posição de partículas quânticas, ficando este limite conhecido por *Princípio da Incerteza* (1958), já aqui referido. No mundo do dia-a-dia, sabemos medir com alguma exatidão tanto a posição como a velocidade de um objeto. Na física quântica, esta medição torna-se mais difícil porque as partículas não se comportam desta forma – “algo que é óbvio, em retrospectiva, se pensarmos um pouco na dualidade onda-partícula” (Gribbin, 2007). A posição de uma partícula é um comportamento típico de uma partícula, enquanto o movimento é um comportamento típico de uma onda. Desta forma, se uma propriedade quântica tiver aspetos tanto de partícula como de onda, não é surpreendente que seja difícil de ser exatamente localizada. Heisenberg verificou, no entanto, que a quantidade de incerteza da posição de uma partícula quântica (ou seja, a incerteza da sua localização) está intimamente relacionada com a incerteza do seu movimento (a incerteza sobre para onde se está a deslocar). Ou seja, quanto mais assertivas forem as medições para a posição de uma partícula, menos corretas serão as verificações para o movimento dessa partícula, e vice-versa (Hawking e Mlodinov, 2011). O físico Werner Heisenberg demonstrou, portanto, que apenas conseguia determinar a probabilidade do movimento das partículas, e não de uma forma exata, o que levou muitos cientistas a defenderem que o Universo possui uma natureza probabilística, isto é, indeterminística.

Por seu turno, o físico Max Born defendeu que não é possível prever o comportamento das partículas, mesmo quando as condições necessárias foram especificadas. Segundo Born, nunca podemos ter X à espera que aconteça Y mas apenas ter X e esperar que aconteça Y ou Z (Ismael, 2021). Neste sentido, segundo Heisenberg (1958) e Born, a natureza do Universo é probabilística porque não conseguimos determinar o comportamento das partículas-ondas – apenas prever hipóteses.

É importante notar que o indeterminismo a que estes estudos se referem — bem como outros mais recentes (Rovelli, 2022; Smolin, 2019) — tem natureza essencialmente epistemológica, isto é, diz respeito aos limites do nosso conhecimento sobre o comportamento das partículas, e não a uma ausência real de causalidade

subjacente (Hooft, 2016; Tegmark, 2014; Hossenfelder e Palmer, 2020). Apesar de a interpretação de Copenhaga ter deixado entrever a possibilidade de uma realidade ontologicamente indeterminada, essa leitura nunca foi formalmente desenvolvida pelos seus autores (Heisenberg, 1958; Bohr, 1958). Ainda assim, a tendência para interpretar esse indeterminismo como ontológico favoreceu não só uma visão anti-realista da estrutura do real (Bohr, 1958), como também serviu de fundamento especulativo para várias teses libertárias sobre o livre-arbítrio (Kane, 2005; Balaguer, 2004).

Por conseguinte, embora o determinismo tenha adquirido bastante adesão recentemente, não só no interior da filosofia (Honderich, 1998; Caruso, 2012; Earman, 2006; Pereboom, 2007; Dennett, 2003; Double, 1991) e das neurociências (Damásio, 2010; Gazzaniga, 2011; Eagleman, 2012; Humphrey, 2012; Seth, 2021; Crick, 1998; LeDoux & Brown, 2017; LeDoux, 2020; Edelman, 1995) mas também no âmbito da física teórica (Hooft, 2016 e 2017; Hawking e Mlodinov, 2011; Hossenfelder e Palmer, 2020; Smolin, 2019), alguns resultados observados na mecânica quântica, interpretados de acordo com a escola de Copenhaga (Ismael, 2021), conduziram não só a alguma resistência às teorias deterministas, como ao surgimento de uma escola antirrealista liderada pela interpretação de Copenhaga (Bohr, 1958; Heisenberg, 1958) que acabou, de facto, por atrair autores com tendências místicas (Briegel e Muller, 2015; Masi, 2023; Kane, 2005; Balaguer, 2014) o que, no entender de alguns (Smolin, 2019), minou a procura de uma física mais realista.

Curiosamente, apesar da proposta de Copenhaga ser a mais popular, existe uma interpretação alternativa — a Bohmiana³⁵ —, considerada por alguns como a interpretação mais coerente e realista (Smolin, 2019), que explica a física quântica de uma forma totalmente determinística. Segundo os físicos David Bohm (1952) e De Broglie (1956), e a sua teoria onda-piloto, não existe qualquer indeterminação na física quântica. Para eles, a partícula comporta-se de uma forma indeterminada porque o seu movimento e posição estão a ser determinados por uma onda que continua a existir num nível mais fundamental e a coordenar deterministicamente o comportamento da partícula. Dito de outra forma, a teoria de David Bohm propõe que as partículas não se comportam de maneira probabilística, como na interpretação de Copenhaga, mas

35 Existe também a interpretação de muitos mundos (Tegmark, 2014).

seguem trajetórias bem definidas, guiadas por uma "onda-piloto" invisível. As previsões teóricas de Bohm e de De Broglie foram posteriormente confirmadas experimentalmente (Smolin, 2019).

Importa ainda referir, ou reforçar, três ideias centrais que raramente são mencionadas no âmbito da física quântica: 1) a física quântica é, à luz do modelo de Copenhaga, uma ciência de medição e é por isso que funciona em torno de probabilidades; 2) as probabilidades da física quântica não invalidam que o Universo seja determinístico, já que as primeiras são medições matemáticas que não possuem a capacidade de interpretar a realidade como ela é, mas apenas em torno de modelos (Hooft, 2016); 3) a interpretação de David Bohm oferece, como vimos, uma explicação totalmente determinística da mecânica quântica (Bohm, 1952). Talvez não seja tão apelativa para cientistas e filósofos místicos, mas, ainda assim, é muito mais realista e coerente na compreensão da natureza da realidade (Smolin, 2019). Iremos analisar os problemas da física quântica com mais detalhe mais à frente.

2.1.7.A hipótese das possibilidades alternativas contestada

Um dos principais temas que gravitam em torno do livre-arbítrio é o da responsabilidade moral. Apesar de não ser o foco central do presente trabalho, parece-nos importante destacar a relação entre o livre-arbítrio e a responsabilidade, uma vez que muitos dos debates em torno da liberdade de ação se desenvolvem precisamente a partir desta ligação.

Um dos argumentos mais recorrentes contra o determinismo é a ideia de que, num universo determinista, o ser humano não poderia ser considerado responsável pelas suas ações, uma vez que estas seriam o resultado inevitável de cadeias causais antecedentes. Tal conceção é, para muitos autores, não apenas filosoficamente problemática, mas socialmente perigosa, na medida em que poderia minar os fundamentos morais e jurídicos da noção de responsabilidade individual. Assim, os libertários sustentam que a única forma de preservar a responsabilidade moral é através da rejeição do determinismo: apenas se as ações humanas forem causadas exclusivamente pelo próprio agente — e não por fatores antecedentes — poderá o sujeito ser considerado verdadeiramente responsável. A liberdade, nesta perspetiva,

reside precisamente na possibilidade real de agir de outra forma, sendo esta condição necessária para que a responsabilidade moral seja atribuída.

No entanto, Harry Frankfurt (1969) veio questionar diretamente essa premissa, ao propor que a responsabilidade moral não exige, necessariamente, a possibilidade de agir de forma diferente. No seu influente artigo *Alternate Possibilities and Moral Responsibility* (1969), Frankfurt apresentou uma série de contraexemplos — os chamados *Frankfurt cases* — em que um agente é, intuitivamente, responsável pelas suas ações, ainda que não tenha tido alternativas reais de escolha.

Estes casos envolvem cenários em que existe um mecanismo externo oculto — como, por exemplo, um dispositivo que monitoriza as intenções do agente e interviria caso este decidisse não executar a ação desejada. No entanto, esse mecanismo nunca chega a ser ativado porque o agente, por sua própria vontade, escolhe realizar a ação. Apesar de não poder ter agido de outra forma, a sua ação resulta da sua própria deliberação, pelo que, segundo Frankfurt, não há qualquer razão para negar a sua responsabilidade moral.

Com estes exemplos, Frankfurt desafia o Princípio das Possibilidades Alternativas (*Principle of Alternative Possibilities* - PAP), segundo o qual um sujeito só pode ser responsabilizado se pudesse ter feito outra coisa. O autor rejeita este princípio, argumentando que a responsabilidade moral depende, não da liberdade para agir de forma diferente, mas da origem da ação no próprio agente. Para Frankfurt, é suficiente que o sujeito aja voluntariamente, sem coação externa, para que possa ser considerado responsável.

Esta posição abre caminho a uma forma de compatibilismo que dispensa o PAP, defendendo que o determinismo não exclui a responsabilidade moral, desde que as ações do agente sejam expressão da sua vontade e não resultado de manipulação externa ou coerção. Com isso, Frankfurt marca um ponto de viragem fundamental no debate contemporâneo sobre livre-arbítrio, deslocando o foco da discussão da liberdade de alternativas para a autenticidade da origem da ação.

2.1.8. O compatibilismo

Além do determinismo e do libertarismo, existe outra filosofia que trata do livre-arbítrio: o compatibilismo. Filósofos compatibilistas (Fisher, 2007; Dennett, 2003;

Hobart, 1934; Locke, 1689/2008; Schopenhauer, 1839/2005) aceitam que vivemos num mundo determinista, mas consideram que, mesmo assim, o determinismo ontológico é compatível com a existência do livre-arbítrio. Como muitos filósofos argumentam que o livre-arbítrio é necessário para que exista responsabilidade moral, o compatibilismo pode ser visto como a filosofia que busca reconciliar esses dois aspetos (McKenna e Coates, 2024).

Em *Free Will as Involving Determination and Inconceivable Without It* (1934), o filósofo R. E. Hobart argumenta que não apenas o livre-arbítrio é compatível com o determinismo, mas que ele só é possível num mundo determinista. Ou seja, o livre-arbítrio exige o determinismo. Segundo Hobart, o indeterminismo é problemático porque, se as ações do sujeito forem determinadas pelos desejos, crenças ou carácter do sujeito, então, o determinismo é verdadeiro. Para ele, ou as ações são determinadas pelo agente ou não são. Se são determinadas pelo agente, como defendem os filósofos libertários, então o indeterminismo é falso e o determinismo é verdadeiro. Para Hobart, dizer que as ações vêm do eu significa que elas são determinadas pelo próprio eu (1934, p. 3). Se não forem determinadas, então as ações não têm causa – ou seja, não são causadas por nada nem por ninguém – o que tornaria impossível a existência, tanto de ações como de livre-arbítrio.

Para Hobart, existe uma relação causal implícita nas teorias libertárias que, paradoxalmente, acabam por confirmar o determinismo. Isto porque, como os próprios libertários admitem, uma ação precisa de ser determinada por algo. Ora, se uma ação tem de ser determinada por algo, então o determinismo é verdadeiro. Assim, se há uma relação causal entre o agente (nomeadamente, o seu carácter) e as suas ações — como alega o libertarismo simples — essa relação causal tem de ser, necessariamente, determinista. Portanto, ou o carácter do agente determina as suas ações, ou não o determina.

Schopenhauer (1839/2005) apresenta uma ideia semelhante. O filósofo defende que ou o carácter do agente determina certas ações, ou não determina nada; mas, se não as determina, essas ações tornam-se aleatórias, desordenadas e incoerentes. Neste sentido, ou o carácter do agente determina as suas ações — boas ou más — e ele é, por isso, livre e responsável por elas; ou então o carácter não determina nada, e a vontade perde consistência: o agente já não poderia ser responsabilizado por ações que se

tornariam arbitrárias. Esta mesma intuição é retomada, mais tarde, por Hobart (1934), que afirma que, se o carácter do agente não for determinado — ou seja, se não for determinado a ser bom ou mau —, então “o que se tem não é um eu moralmente admirável ou condenável, nem um eu moralmente caracterizável” (p. 3). De certa forma, se o agente é capaz de realizar X — isto é, se tem o poder de executar a ação X — é porque existem leis que o determinam a fazê-lo. E se há leis que tornam X possível, é porque certas condições naturais tornaram essa ação realizável. Isso só é concebível num universo determinista (Hobart, 1934).

Na mesma linha, no livro *Freedom Evolves* (2003), Daniel Dennett argumenta que o determinismo é compatível com o livre-arbítrio, porque é justamente o determinismo que nos dá a capacidade de evitar perigos, tomar decisões e exercer autocontrole. Para Dennett, o livre-arbítrio depende de um cérebro saudável, governado por leis físicas e biológicas, que de facto conferem liberdade ao agente, evitando que ele seja refém de compulsões ou da falta de autocontrole — problemas que resultam de síndromes mentais, desequilíbrios químicos ou lesões cerebrais. Por outras palavras, Dennett postula que qualquer ser humano com um cérebro saudável tem responsabilidade moral, porque é precisamente a neurobiologia funcional que confere essa responsabilidade. Assim, tanto a responsabilidade moral como a volição — elementos tradicionalmente associados ao livre-arbítrio — são, nesta perspetiva, causados pelo funcionamento do cérebro (Dennett, 2003).

Por outro lado, Dennett considera que a capacidade de alguns seres de *evitar* o futuro é uma demonstração clara de livre-arbítrio, especialmente nos humanos, que têm a habilidade de antecipar o futuro melhor do que qualquer outra espécie — incluindo a própria morte, que pode, em certos casos, ser evitada. A evolução produziu organismos com sistemas nervosos capazes de representar o ambiente e antecipar o futuro. Diferentes sistemas nervosos, ao longo da evolução, permitiram que várias espécies se adaptassem melhor ao ambiente em que estavam inseridas, evitando perigos futuros. Dennett chama essas espécies de *evitadores* (*avoiders*). Nesse sentido, segundo Dennett:

"Um modelo mecânico do determinismo demonstra que, no vasto espaço de configurações possíveis da matéria, existem algumas que persistem melhor do que outras, porque foram projetadas para evitar danos"³⁶ (2003, p. 62).

De entre todos os *evitadores*, os humanos são os mais eficientes. Nas palavras de Dennett:

"Somos evitadores virtuosos, prevenidores, interferidores, precavidos. Conseguimo-nos colocar na feliz situação de termos tempo livre para sentar e analisar sistematicamente o futuro, perguntando-nos o que fazer a seguir"³⁷ (2003, p. 54).

Portanto, o determinismo, segundo Dennett, não implica inevitabilidade, pois milhões de seres vivos evitam perigos constantemente, o que mostra que a previsibilidade não significa falta de liberdade.

Em suma, o compatibilismo sustenta que o livre-arbítrio é compatível com o determinismo, na medida em que a liberdade consiste na capacidade de agir de acordo com o carácter, as motivações e a deliberação racional do agente. Para autores como Hobart, Schopenhauer e Dennett, essa relação causal entre o agente e as suas ações não elimina a liberdade, mas torna-a possível, ao garantir consistência moral, responsabilidade e adaptação ao futuro.

2.2. A impossibilidade ontológica do indeterminismo: entre contradições e confusões conceptuais

Como vimos, o libertarismo utiliza o indeterminismo ontológico para argumentar a favor de um livre-arbítrio *acausal*. Tanto o libertarismo simples como o libertarismo causal defendem (embora sobre perspectivas muito distintas) que é a natureza indeterminística do Universo que permite a existência de livre arbítrio. Ambas assumem, aliás, que o indeterminismo é uma posição ontológica sobre o Universo. No entanto,

36 Tradução do autor.

37 Tradução do autor.

parece-nos claro que o indeterminismo não é nem pode ser – em função das suas premissas – um conceito ontológico.

Como foi anteriormente referido, o libertarismo fundamenta-se, sobretudo, em três princípios centrais: 1) na ideia de que o ser humano pode violar as leis da Natureza — ou seja, que é possível que, num Universo regido por leis, existam eventos mentais que escapam a essas mesmas leis, o que constitui uma contradição lógica (Ginet, 2014; McCann, 1998; Balaguer, 2004; Kane, 2005); 2) na possibilidade de existirem eventos *acausais* — segundo diversas formulações libertárias — como a vontade (Ginet, 2008), a volição (McCann, 1998) ou uma consciência não-física (O'Connor, 2000; Chisholm, 2003), o que implicaria aceitar a existência de acontecimentos sem qualquer causa, o que é logicamente incoerente; 3) que o Universo, por ser indeterminístico possui uma natureza probabilística e aleatória (Kane, 2005; Balaguer, 2004), o que não só é uma conclusão equivocada, como é uma ideia claramente epistemológica (Hooft, 2016; Romeijn, 2022; Hossenfelder, 2022; Rovelli, 2022), que nada nos diz sobre as *causas* dessas probabilidades e dessa aleatoriedade.

Para além destes problemas, que analisaremos com cuidado, parece-me igualmente que existem problemas semânticos e conceptuais, que iremos também identificar.

2.2.1. Indeterminismo ontológico: um conceito filosoficamente incoerente

Como referido, as filosofias libertárias procuram fundamentar a defesa da liberdade humana através do recurso a alguma forma de indeterminismo. No entanto, como será demonstrado, o indeterminismo é não só um conceito ontologicamente incoerente, como tem sido erroneamente utilizado pelos teóricos libertários.

Podemos dizer que o indeterminismo divide-se em duas dimensões fundamentais: 1) a epistemológica, frequentemente utilizada na física quântica (Rovelli, 2022; Smolin, 2019); 2) e a ontológica, implícita na interpretação de Copenhaga (Bohr, 1958; Heisenberg, 1958) e recrutada de forma explícita por alguns filósofos libertários (Ginet, 2014; Kane, 2005; Balaguer, 2014). Na primeira, o indeterminismo visa descrever fenómenos causados por condições antecedentes, mas cujos resultados não podem ser previstos com exatidão — devido a limitações tecnológicas ou cognitivas. Trata-se, portanto, de um indeterminismo aparente, amplamente aceite por muitos físicos, como

Carlo Rovelli (2022), Lee Smolin (2019) e Sabine Hossenfelder (2022), entre outros. Já o indeterminismo ontológico procura sustentar que existem fenômenos fundamentalmente *acausais* e que, precisamente por não possuírem qualquer causa antecedente, são indeterminísticos na sua essência. Este tipo de indeterminismo é frequentemente invocado por filósofos libertários, como Robert Kane (2005) ou Mark Balaguer (2004).

Apenas um destes conceitos nos parece filosoficamente sustentável: o indeterminismo epistemológico. Desde logo porque não postula uma realidade *acausal*, tese que consideramos logicamente insustentável. O indeterminismo ontológico, ao defender que a Natureza pode *causar* eventos *acausais* (Ginet, 2014; p. 25; Balaguer, 2014, p. 381), incorre numa contradição categorial: se a Natureza *acausa* um evento, então o evento é causado — e não *acausado*. Assim, o próprio conceito de indeterminismo ontológico colapsa na sua formulação. O conceito de indeterminismo possui, portanto, dois problemas fundamentais. Por um lado, o indeterminismo ontológico é filosoficamente insustentável — é um absurdo filosófico —, pelo que não tem capacidade para refutar o determinismo; por outro, o indeterminismo epistemológico refere-se apenas à nossa incapacidade de determinar certos eventos, pelo que não dialoga bem com o determinismo que é um conceito que se pronuncia sobre as causas fundamentais, não sobre a nossa capacidade de as prevermos.

Relembremos que o determinismo afirma que todos os eventos têm uma causa suficiente (Earman, 1986; Hooft, 2016; Caruso, 2012). Se o indeterminismo se limitar a defender que nem sempre conseguimos conhecer tais causas, mas que elas existem, então está a utilizar a terminologia de forma incorreta — pois está, na prática, a enunciar uma posição determinística. O único modo de o indeterminismo contrariar efetivamente o determinismo é assumir, de forma clara, que o universo é ontologicamente indeterminístico — ou seja, que contém eventos *acausais*. Contudo, como argumentamos, essa hipótese compromete-se com uma ontologia que consideramos incoerente e filosoficamente indefensável.

Para além do problema central do indeterminismo ontológico (a partir daqui designado apenas por indeterminismo³⁸), surge um outro, que emana de um erro de

³⁸ Visto o indeterminismo epistemológico não conseguir dialogar com o determinismo.

categorização – a ideia de que o determinismo é falso porque a física quântica demonstrou que existem eventos cujo comportamento não pode ser calculado com exatidão, o que significa que o Universo é indeterminístico. Como já aqui foi citado, o físico Michio Kaku refere-se ao determinismo como a filosofia que pretende prever eventos com exatidão (2021, p. 69). Os filósofos Robert Kane e Mark Balaguer, no mesmo sentido de Kaku, são claros a este respeito quando gravitam a sua discussão em torno da possibilidade de as partículas elementares poderem, ou não, ser (epistemologicamente) determinadas, como se isso pudesse validar, ou não, o determinismo ou indeterminismo ontológico. Ou seja, ao considerar que o comportamento das partículas quânticas não é previsível, Kane extrapola essa imprevisibilidade para a conclusão de que o Universo não pode possuir uma natureza determinística. Nas suas palavras:

“You are probably familiar with the claim that modern quantum physics has introduced indeterminism or chance into the physical world. Much of the behavior of elementary particles, it is said, from quantum jumps in atoms to radioactive decay, is not precisely predictable and can be explained only by statistical, not deterministic, laws. [...] the uncertainty and indeterminacy of this world of quantum physics, according to the standard view of it, is not due to our limitations as knowers, but to the unusual nature of elementary particles themselves [...] their future behavior is not precisely predictable or determined.” (2005, p. 8).

Na mesma linha de Kane, o filósofo Mark Balaguer sugere que o Universo pode ter uma natureza indeterminística, já que a interpretação de Copenhaga tem vindo a demonstrar que o Universo é probabilístico. Balaguer postula que:

“The standard (and probably best) interpretations of quantum mechanics take the world to be indeterministic at the most fundamental level. It may be that the world *is causally indeterministic*³⁹ all over the place, and so it may be that neural events and, hence, decisions are indeterministic.” (2014, p. 381).

Tanto Kane como Balaguer utilizam, portanto, a aleatoriedade quântica para dar um salto extrapolativo entre a imprevisibilidade epistemológica — isto é, a impossibilidade de conhecer ou prever certos eventos — e a defesa de um Universo

³⁹ Sublinhado do autor.

indeterminista, ou seja, *acausal*. Em ambos os casos, parece verificar-se uma confusão entre indeterminismo epistemológico (utilizado pela física quântica) e indeterminismo ontológico (extrapolado pelos filósofos libertários), ao sugerirem a possibilidade de um Universo indeterminístico e, necessariamente, *acausal* apenas porque, como demonstra a física quântica, não conseguimos prever o comportamento das partículas fundamentais.

Assim, estas propostas parecem confundir dois planos distintos: por um lado, um Universo determinista, no qual as leis naturais causam eventos — o que não é refutado pela física quântica (Rovelli, 2022; Smolin, 2019; 't Hooft, 2016); por outro, um Universo *determinável*, isto é, um Universo cujos eventos poderiam, em princípio, ser previstos (Kane, 2005; Balaguer, 2014). Ora, o determinismo defende, de forma ostensiva, que o Universo é regido por leis causais estritas, mas afirma, com igual clareza, que o seu desenrolar não pode ser conhecido ou previsto na totalidade (Sapolsky, 2024; Bombarda, 1902; Dennett, 2003; Pereboom, 2007; Caruso, 2012). Como refere Gerard 't Hooft no seu livro *The Cellular Automaton Interpretation of Quantum Mechanics*: “Deterministic cannot imply that the outcome of the evolution process can be foreseen” (2016, p. 68).

Importa por isso enfatizar, desde já, uma ideia central. *O determinismo não é a tese de que podemos determinar o futuro do Universo; ou a tese de que podemos determinar o comportamento humano; nem tão pouco a tese de que podemos determinar coisa alguma.* Qualquer teoria que defenda que a física quântica contradiz o determinismo ao provar que o Universo é aleatório (Kane, 2005; Balaguer, 2004 e 2014; Kaku, 2021) é falsa. É falsa porque o determinismo não se funda em ideias ou conceitos epistemológicos — como aleatoriedade ou probabilidade. O determinismo não defende que é possível medir, determinar ou prever o que quer que seja, pelo que as probabilidades e a aleatoriedade que podem existir no Universo são características sobre as quais o determinismo não se pronuncia (Hooft, 2016; Caruso, 2012; Earman, 2006; Bombarda, 1902). O Universo pode ser aleatório e probabilístico e ainda assim determinístico, já que a aleatoriedade e a probabilidade são dois conceitos epistemológicos utilizados para *conhecer* e *medir* o Universo (Hooft, 2016; Romeijn, 2022; Smolin, 2019), e o determinismo não se propõe nem a conhecer nem a medir o Universo. O determinismo apenas defende, como temos vindo a referir, que *todos os*

eventos que existem são causados por eventos anteriores e pelas leis da natureza. Esses eventos, podem ser probabilísticos e as leis da natureza podem ser aleatórias e ainda assim o determinismo ser verdadeiro. A nosso ver, não existe indeterminismo mas apenas a *percepção* de indeterminismo (Tegmark, 2014; Hooft, 2016 e 2017).

Assim, o indeterminismo proposto por filósofos libertários como Kane ou Balaguer — seja ele ostensivamente *acausal*, ou causal como quando Balaguer afirma que “the world is causally indeterministic all over the place...” (2014, p. 381) — revela-se profundamente contraditório. Ou o indeterminismo é, de facto, *acausal* e estes filósofos têm de explicar como existem eventos na Natureza, que se regem segundo as leis da Natureza sem que tenham sido causados pelas leis da Natureza; ou é um indeterminismo causal, como defende Balaguer, e neste caso ele terá de clarificar qual é, exatamente, a diferença entre um evento indeterministicamente causado e um evento deterministicamente causado. É que, do ponto de vista dos filósofos deterministas, qualquer evento que é causado é, por definição, determinístico.

Em síntese, consideramos que, por um lado, o determinismo deve ser entendido como a tese segundo a qual todos os eventos possuem uma causa — não como a ideia de que esses eventos podem ser previsíveis; e, por outro lado, que o indeterminismo ontológico é filosoficamente incoerente por duas razões fundamentais: 1) porque postula uma realidade *acausal*, o que nos parece logicamente insustentável; 2) e porque é frequentemente extrapolado por filósofos libertários para defender a existência de uma causalidade indeterminística — uma formulação paradoxal, já que, ao admitir causalidade, acaba por pressupor uma estrutura determinista.

2.2.2.A impossibilidade dos eventos *acausais*: o paradoxo do libertarismo

Como referimos anteriormente, o determinismo deve ser entendido como uma lei estrutural da Natureza: uma tese segundo a qual todos os eventos possuem uma causa e ocorrem em conformidade com leis naturais. Trata-se, assim, de uma propriedade constitutiva da própria estrutura ontológica do Universo (Hooft, 2016; Pereboom, 2007; Espinosa, 1677/1992; Earman, 2006). Isso significa que, se o Universo possui leis, essas leis não só o regem como são inerentes à sua existência (Hooft, 2016 e 2017). Poderíamos, aliás, conceber o Universo como o conjunto de leis que regulam e

causam efeitos no próprio tecido da realidade física, sendo essas leis a condição necessária para que qualquer evento possa existir ou ocorrer.

É, portanto, possível que o Universo nos pareça indeterminado do ponto de vista epistemológico — afinal, não o conseguimos medir nem observar integralmente na sua estrutura —, mas isso não implica que o seja ontologicamente. O que rejeitamos é que, em algum momento, alguma parte do Universo possa comportar-se à margem das suas próprias leis, como defendem certas propostas libertárias (Ginet, 2014, p. 25; Balaguer, 2014, p. 381; Kane, 2005, p. 8). Se o Universo é regido por leis, então todas as suas partes — incluindo as partículas que compõem o ser humano — têm de obedecer a essas mesmas leis (Hooft, 2016; Rovelli, 2018). Por isso, o indeterminismo só pode ser epistémico. Ou o Universo é um sistema constituído por leis que causam eventos, ou não o é — e, nesse caso, não poderia sequer existir. Num sistema físico regido por leis naturais e constituído por leis naturais, é logicamente necessário que essas leis determinem os eventos, já que são simultaneamente causa e estrutura da realidade. Mesmo que o indeterminismo pareça verdadeiro à nossa percepção, a natureza física do Universo obriga à existência de causalidade (Hooft, 2017; Hawking e Mlodinov, 2011; Krauss, 2017; Smolin, 2019; Rovelli, 2018).

Aplicando isto à física quântica, a situação é semelhante. Afirmar que fótons ou elétrons são indeterminados significaria que nada causou o seu comportamento — o que é insustentável num sistema físico. Podem parecer *acausados* à nossa percepção, mas se emergem do Universo, emergem de um "nada" que é ele próprio estruturado fisicamente. Esse "nada" é, na verdade, uma expressão de um todo (Bohm, 1980), regido por leis. O que muitos autores defendem não é que os *quanta* sejam ontologicamente indeterminados, mas apenas que não conseguimos prever o seu comportamento com precisão (Hooft, 2016 e 2017; Hawking e Mlodinov, 2011; Gribbin, 2007; Krauss, 2017; Earman, 2006; Rovelli, 2022). Assim, mesmo que partículas como os fótons se comportem ora como onda, ora como partícula (Rovelli, 2022), isso tem uma causa — mesmo que não saibamos ainda qual. Podemos especular que esta dualidade resulta de uma estrutura mais primordial da realidade, anterior à manifestação clássica da partícula, mas, o facto de não conseguirmos prever ou localizar o comportamento dessas partículas não implica que não haja uma posição ou um movimento determinado (Bohm, 1952 e 1980; Hooft, 2016; Gribbin, 2007).

A física quântica é, como refere Gerard't Hooft, um modelo estatístico e instrumental — um *template* de previsões, não a ontologia da realidade em si: “...a mathematical tool, not a mystic departure from ordinary, classical, logic” (2016, p. 81). As suas leis são operatórias, e não fundacionais. Assim, é expetável que certos fenómenos pareçam aleatórios, mas essa aleatoriedade é, como em muitos sistemas caóticos (clima, neurobiologia), apenas aparente. Os comportamentos quânticos, mesmo imprevisíveis, são causados e, aliás, muito mais bem determinados do que muitos sistemas macroscópicos, já que é muito fácil prever a localização de um fotão ou de um eletrão, do que um dia de chuva (Rovelli, 2022). Deste modo, não vemos como se possa aplicar à física quântica o conceito de indeterminismo ontológico, que pressupõe a existência de eventos sem causa anterior (Kane, 2005; Balaguer, 2004). O único indeterminismo compatível com esta ciência é o das medições — uma limitação epistémica, não ontológica (Hooft, 2016).

É neste contexto que se torna especialmente problemática a tese libertária segundo a qual o Universo pode conter eventos verdadeiramente *acausais*. O indeterminismo, segundo os teóricos libertários (Ginet, 2014; Kane, 2005; Balaguer, 2014) abriria a possibilidade de escolhas não causadas por estados anteriores nem regidas pelas leis da Natureza. Mas o que significa, exatamente, afirmar que o Universo pode conter eventos que escapam à sua própria estrutura? No nosso entender, isto seria o mesmo que dizer que há leis ou ocorrências que não pertencem ao Universo — o que constitui uma contradição em termos.

Balaguer sugere, como referido, que o Universo pode ser “causally indeterministic all over the place” (2014, p. 381). Mas esta formulação é conceptualmente instável, pois mistura categorias contraditórias: causalidade e indeterminação. Se algo é causado, é porque há uma relação causal; se é indeterminado, não há causa suficiente. A expressão “causally indeterministic” é, portanto, internamente contraditória. Ao admitir causalidade, pressupõe-se já uma estrutura determinista. No mesmo sentido, Ginet afirma, como também já foi anteriormente analisado, que a Natureza pode causar eventos *acausais* (2014, p. 25). Mas, afirmar que a Natureza causa eventos *acausais* é um paradoxo lógico. Ou um evento é causado — e está subordinado às leis naturais —, ou é *acausal* — e não foi causado por nada, nem

sequer pela Natureza. Não é possível sustentar logicamente que algo seja simultaneamente causado e *acausal*.

Consideramos que, mesmo que a Natureza causasse eventos inéditos, desconhecidos ou imprevisíveis, esses eventos estariam, necessariamente, subordinados à sua estrutura e às suas leis. O conceito de violação das leis da Natureza parece-nos, portanto, contraditório neste contexto, já que qualquer violação aparente seria, na verdade, expressão das possibilidades abertas pelas próprias leis. A Natureza não pode agir contra si mesma. Neste sentido, a proposta libertária é insustentável por duas razões: 1) todos os eventos que existem num Universo regido por leis têm de ser causados por essas leis; 2) se os eventos existem, então foram causados, mesmo que não sejam previsíveis. Assim, 3) postular que algo acontece fora da ordem causal do Universo equivale a afirmar que o Universo possui elementos que escapam às condições que o permitem existir.

Por conseguinte, não só é impossível afirmar que o ser humano pode violar as leis da Natureza, como é igualmente insustentável postular que a própria Natureza o possa fazer. Se o Universo não pode violar as suas próprias leis, então nenhum dos seus componentes — incluindo a consciência — o pode fazer. Afirmer que o comportamento humano é indeterminístico equivale, neste quadro, a enunciar um evento ontologicamente impossível, da mesma forma que é impossível que o Universo possa violar a sua própria estrutura. Em termos formais podemos verificar a inconsistência dos eventos *acausais* proposto pelos teóricos libertários desta forma:

- 1) Se o indeterminismo é a tese de que certos eventos não precisam de causas anteriores nem de leis naturais;
- 2) e se o Universo contém tais eventos;
- 3) então o Universo possui eventos que não obedecem às suas leis;
- 4) logo, o Universo viola a sua própria estrutura, o que é logicamente incoerente.

Analisemos, agora, o problema dos eventos *acausais* no âmbito de cada uma das filosofias libertárias.

Como anteriormente analisado, o indeterminismo simples defende, que uma razão, uma decisão ou a volição são eventos mentais *acausados*. No entanto, quando Ginet (1968, 2014) diz que uma razão surge sem causas, o que significa isso? Que uma

ideia surgiu na mente transpondo as leis da física e da química que governam os cérebros? Quando McCann (1998) refere que a volição confere liberdade ao ser humano, sugerindo que nada causou essa intencionalidade, está ele a sugerir que essa intenção vive à margem das leis que governam o Universo e os cérebros? Pode uma razão, uma crença ou a volição surgir na mente de forma *acausada*?

Se aceitarmos que uma razão ou uma volição pode emergir sem qualquer causa, estaríamos a admitir que eventos mentais complexos, como uma decisão consciente ou uma intenção deliberada, poderiam existir sem qualquer antecedente físico ou biológico, o que implica uma rutura ontológica radical com a estrutura causal que organiza o Universo. Esta hipótese levanta, aliás, um paradoxo ontológico incontornável: se a consciência, enquanto fenómeno físico ou dependente do físico, pode gerar eventos mentais *acausados*, então a própria causalidade que estrutura o Universo e, subsequentemente o Universo, seria violada dentro do cérebro, o que gera um paradoxo evidente.

No mesmo sentido, seria necessário postular que a mente humana possui uma capacidade única de escapar às leis da natureza, o que exigiria, não apenas uma explicação para a emergência dessa capacidade, mas sobretudo uma justificação plausível para a sua existência num sistema físico que é regido por leis físicas. Esta exigência, a nosso ver, não é apenas logicamente inconsistente, mas também cientificamente insustentável. Por conseguinte, admitir a existência de eventos mentais *acausados*, como fazem Ginet (1962; 2014) e McCann (1998), não resolve o problema da liberdade; apenas transfere a incoerência para o domínio da mente, sugerindo que um fenómeno físico — o cérebro — pode gerar espontaneamente eventos que escapam às condições físicas que o tornam possível. Podemos verificar os paradoxos do libertarismo simples desta forma:

1. O Universo é um sistema físico regido e constituído por leis;
2. essas leis são a causa dos eventos que ocorrem nesse sistema;
3. logo, todos os eventos do Universo são causados por essas leis;
4. por conseguinte, não é possível a existência de eventos *acausais* num sistema cuja existência e funcionamento dependem dessas mesmas leis.

Essencialmente no mesmo sentido, o indeterminismo de causação agencial (O'Connor, 2000; Chisholm, 2003) postula, como vimos, que a liberdade existe porque o ser humano possui uma consciência não-física que permite não só atuar sobre o cérebro como ainda violar as leis físicas segundo as quais os cérebros podem, e devem, funcionar. A consciência, sendo uma propriedade fundamental emergente, desprende-se das leis físicas, pelo que pode ter uma autonomia que confere ao sujeito a possibilidade de decidir livremente, designadamente interferir no seu sistema nervoso, conforme argumentado.

Ora, estas ideias colocam-nos os seguintes problemas: 1) se a consciência emerge num sistema físico, regido por leis físicas, como pode ela ser uma propriedade diferente das leis da física? 2) se a consciência imaterial é causada pelo cérebro, como é que pode interagir com a estrutura física que permite a sua existência; 3) se a consciência imaterial é causada pelo cérebro, então não vemos como é que podemos desprender a fisicalidade do cérebro da fisicalidade da consciência, isto é, não vemos como é que a consciência pode ser uma propriedade não-física; 4) mesmo que a consciência seja um processo emergente do cérebro, não vemos qualquer necessidade de dualizá-la, atribuindo-lhe características não-físicas, já que a consciência pode existir como propriedade fenomenológica física. A fisicalidade não é antagónica à fenomenologia. Pelo contrário. Suporta-a e, ao mesmo tempo, é a fenomenologia. Por conseguinte:

- 1) O Universo é um sistema físico regido e constituído por leis;
- 2) não é possível existir nada no Universo que não seja físico;
- 3) se a consciência pertence ao Universo;
- 4) então a consciência humana tem de se reger por essas leis e tem de ser, aliás, constituída por leis físicas.

O libertarismo de causação agencial parece-nos bizarro porque, mais do que todas as outras teorias, desafia os alicerces das leis da Natureza. Como refere Daniel Dennett em *Freedom Evolves*: “Agent causation is a frankly mysterious doctrine, positing something unapparelled by anything we discover in the causal process...” (2003, p. 100).

Por isso, postulamos que, se a consciência existe num Universo físico e regido por leis físicas, e se é causada pelo cérebro, então não pode interagir com a estrutura

que lhe dá origem senão *por intermédio dessas mesmas leis*. Se os libertários pretendem que a consciência tenha poder causal no mundo físico têm de, necessariamente, atribuir fisicalidade à consciência. Só assim a consciência pode “comunicar” com o cérebro e, subsequentemente, só assim a consciência pode assumir poder causal. Somos da opinião que a consciência pode participar nas decisões e influenciar o cérebro (e veremos isso mais à frente, no capítulo 3), mas isso só é possível se obedecer às leis do Universo, pelo que tudo o que a consciência possa realizar tem de ser causado pelas leis físicas. Deste modo, as decisões conscientes do agente, se obedecerem às leis da física, são causadas pela consciência, que por sua vez, é causada pelo cérebro. Não é possível que o cérebro, enquanto sistema físico, crie um sistema não-físico que interaja com ele. Em primeiro lugar, porque qualquer sistema que exista num Universo físico só pode gerar sistemas físicos; em segundo lugar, porque para que dois sistemas possam interagir é crucial que partilhem não apenas a mesma substância, mas também as mesmas propriedades. Segundo estes princípios, como veremos com mais detalhe no capítulo 3, a consciência, apesar de poder ser um processo emergente, não só é causada pelo cérebro, como está dependente desse cérebro para existir e funcionar, desde logo porque a sua emergência depende de leis que *causam* e suportam essa mesma emergência (Damásio, 2010; Gazzaniga, 2011; Eagleman, 2012; Humphrey, 2012; Seth, 2021; Crick, 1998; LeDoux & Brown, 2017; LeDoux, 2020; Edelman, 1995; Ramachandran e Hirstein, 1997).

Por sua vez, o libertarismo causal defende que o ser humano é livre porque pode escolher livremente as suas decisões após um processo indeterminístico que cria as múltiplas opções futuras, como analisado (Kane, 2005; Balaguer, 2014). Segundo este tipo de libertarismo, o indeterminismo quântico rege os sistemas neuronais, o que permite múltiplas possibilidades futuras, mas, conforme argumentado por esta perspetiva, esse indeterminismo culmina num controlo por parte do agente, que *escolhe* – e *causa* livremente – a decisão. No entanto, esta proposta tem vários problemas (nomeadamente relacionados com a adaptação da física quântica ao funcionamento neuronal, que aprofundaremos mais à frente), desde logo a ideia do controlo que o agente faz sobre os processos indeterminísticos do seu cérebro ser (ainda que implicitamente) *acausadamente causado* pelo sujeito.

Deste modo, o libertarismo causal possui dois tipos de *acausalidade*: uma *acausalidade* ontológica, sustentada no indeterminismo ontológico; e uma *acausalidade* mental, ora implícita quando Kane refere que o sujeito decide conscientemente sem referir o que causa essa decisão, ora explícita quando Balaguer sugere ostensivamente que nada causa a decisão do agente, que apenas decide.

Isto coloca-nos, desde logo um problema evidente. Como podem as decisões do sujeito ser *acausadas*? Isto significaria que, não só o Universo era capaz de causar eventos *acausados*, que nos parece um absurdo filosófico, como o cérebro era capaz de gerar eventos mentais *acausados* o que é, igualmente, uma impossibilidade ontológica. Quando Kane (2005) postula que o sujeito toma uma decisão consciente, ele não só não explica como é que essa decisão é causada por uma ontologia *acausal*, como é omissa em relação àquilo que leva à causa da decisão consciente. No mesmo sentido, quando Balaguer refere que o agente toma uma decisão mas nada causa essa decisão, ele, tal como Kane, não só não explica como é que uma ontologia *acausal* pode causar o que quer que seja, como não é claro como é que o sujeito toma uma decisão completamente *acausada*. Um outro problema nesta teoria é a ideia de que as crenças, as ideias e os pensamentos, entre muitos outros eventos mentais, são *acausadas* porque surgiram das probabilidades inerentes à física quântica, anteriormente analisado. Ora, para além de terem de explicar como é que o indeterminismo quântico — que é, em rigor, um fenómeno probabilístico e não necessariamente indeterminístico — poderia causar eventos *acausados*, estes filósofos têm igualmente de justificar de que modo as regras da física quântica se poderiam aplicar ao funcionamento macroscópico dos neurónios, algo que é veementemente rejeitado pelas neurociências (Damásio, 2010; Gazzaniga, 2011).

Tomando como exemplo uma situação frequentemente utilizada pelos defensores do libertarismo causal, imaginemos um sujeito que, confrontado com a necessidade de escolher uma profissão, passa vários dias em introspeção. Durante esse processo — caracterizado, segundo esta teoria, pela emergência indeterminística de ideias, memórias e intenções — o agente teria a capacidade de exercer controlo sobre as decisões que toma, precisamente, por conseguir orientar-se dentro desse fluxo interno de possibilidades. A escolha consciente e responsável entre várias opções possíveis confere, segundo esta perspetiva, a liberdade necessária para que o sujeito

consiga superar, simultaneamente, o *acaso* do indeterminismo e a *causalidade* determinista.

Mas isto parece-nos um equívoco ingênuo, não só porque as múltiplas possíveis opções tiveram de ser causadas pelas leis da Natureza (mesmo que probabilisticamente) como pelo facto de, quando o sujeito escolhe, decide e delibera – quando decide conscientemente o que fazer perante as múltiplas opções que lhe emergem na mente – alguma coisa tem de causar essa escolha, essa decisão e essa deliberação. Ou o sujeito escolhe conscientemente através de um processo *acausado* — e os indeterministas causais têm de explicar, não só como é que alguma coisa pode existir sem ser causada, mas também como podem existir decisões criadas *ex nihilo* —; ou a decisão é, igualmente, causada pelas leis da física, nomeadamente pela organização neuronal e pelo contexto em que está inserida.

Se decisões e escolhas forem ontologicamente *acausadas*, então perdem o seu conteúdo racional e intencional — sobretudo na perspetiva defendida por Balaguer (2014), segundo a qual não existe qualquer decisão consciente por parte do sujeito. As decisões surgem de forma totalmente indeterminada, sem ligação causal com estados mentais precedentes, como desejos, crenças ou deliberações. Neste modelo, a liberdade resulta exclusivamente da ausência de determinação, e não de um processo consciente de escolha, o que esvazia a decisão de qualquer estrutura racional ou volitiva. Nessa lógica, nada impediria que uma bactéria — ou até um objeto inanimado — pudesse tomar uma "decisão livre", já que a ausência de causalidade permitiria que qualquer entidade gerasse escolhas sem que estas resultassem de processos internos organizados e regulares. Por outras palavras, o indeterminismo quântico poderia ter (a)causado uma série de ideias numa bactéria, permitindo à bactéria tomar uma decisão. Assim, se decisões e escolhas são *acausadas* e aleatórias, elas poderiam simplesmente acontecer a qualquer ser vivo ou sistema físico, esvaziando completamente a noção de agência.

A questão que colocamos é, portanto, a seguinte. E se as decisões e as escolhas só existirem porque são causadas pelo sistema nervoso, não num processo estocástico e *acausado* mas, pelo contrário, através de leis e de *causas* que *causam* esses eventos mentais? Defender, como Kane (2005), que a mulher de negócios que ia para uma

reunião optou por ajudar um homem em apuros, em vez de ir para uma reunião que era mais importante para a senhora, é apenas dizer que ela demonstrou empatia em relação ao senhor assaltado, e que essa empatia foi mais forte do que a sua vontade egoísta de ir para uma reunião importante. Não há nada de *acausal* nem de indeterminístico neste processo. Aliás, nem é um comportamento difícil de prever, já que as pessoas que conhecessem essa senhora saberiam (e conseguiriam prever deterministicamente) que, naquelas circunstâncias, ela iria decidir daquela forma. A senhora de negócios não precisou de ideias indeterminísticas nem de um controlo *acausal* para decidir o que quer que seja. Bastou efetuar um comportamento empático *causado* pelo seu cérebro e completamente previsível.

O exemplo de Kane sobre a senhora que está atrasada para uma reunião de negócios pode muito bem, aliás, ter exatamente a mesma conclusão à luz do determinismo. Segundo o determinismo, a mulher de negócios poderia ter duas opções alternativas (ir para a reunião ou ajudar a pessoa assaltada) e optar pela mesma razão – ajudar a pessoa agredida, prejudicando assim a sua vida profissional. Segundo o determinismo, ela até poderia ter dez alternativas e escolher qualquer uma delas. Poderiam, inclusive, fazer-se probabilidades em relação ao que a senhora decidiria, mas isso não invalida, em absoluto, que o seu comportamento esteja determinado por 1) aquilo que *causa* a vontade de agir (emoções, ética, moral, o carácter...); 2) e pela vontade de agir. Se a senhora tomou a decisão de ajudar a pessoa que estava a ser assaltada, isso significa que esse segundo evento (aquela ação), dadas as exatas circunstâncias (estrutura cerebral e contextos) do primeiro evento (a escolha), não poderia ter acontecido de outra forma. Ou a senhora de negócios é um elemento da Natureza, que se rege segundo as leis da Natureza, o que permite tomar decisões que, não obstante serem causadas (e sobretudo por serem causadas) podem causar eventos; ou essa senhora tem poderes sobrenaturais para contornar as leis da Natureza e, deste modo, realizar ações que *não são causadas por nada* e, ainda assim, causam eventos.

Por conseguinte, em poucas palavras, aquilo que o determinismo defende em relação a esta situação é que, qualquer que seja aquilo que a senhora de negócios escolha, 1) será inevitavelmente causado por eventos antecedentes aos quais ela – o sujeito – não tem acesso, nomeadamente a sua organização neuronal, que foi desenvolvida por fatores que ela também não controla, designadamente genéticos,

educacionais e ambientais (Sapolsky, 2024); e 2) a decisão (fosse qual fosse), uma vez ocorrida, não podia ter existido de forma diferente pela simples razão de que, para que a opção se desse, era necessária uma *causa* que a *causasse*.

Poderíamos dizer o seguinte em relação ao libertarismo causal:

1. Se o Universo é regido por leis que causam eventos;
2. então todos os eventos são causados;
3. logo, eventos *acausais* são impossíveis;
4. e, por isso, o indeterminismo ontológico (que postula eventos *acausais*) é impossível.

Segundo Richard Double (1991), os filósofos libertários confundem liberdade com *ilusão de liberdade*, ou seja, não distinguem o desejo de realizar uma ação livre com as *causas* que determinaram essa ação, desde logo porque ignoram essas causas, algo que Espinosa já tinha sugerido no século XVII quando disse: “Os homens julgam-se livres porque são conscientes das suas ações e ignorantes das causas por que são determinadas” (Espinosa, 1677/1992, p. 272).

Em suma, os eventos *acausais* são impossíveis num Universo regido por leis e causas físicas. Por um lado, a existência de eventos obriga à necessidade de causalidade, já que só podem existir coisas se forem trazidas à existência por *causas*. Por outro, o indeterminismo, ou causa eventos através das leis da física, mesmo que continuem a ser aleatórios e probabilísticos, e morre como conceito, já que passa a ser determinismo; ou então, continua a defender que o Universo possui eventos *acausais*, que podem violar as leis da física, tendo de explicar, entre outras coisas, como é que o Universo pode violar as próprias leis que criou e permitem a sua existência, e como é que o ser humano pode violar essas mesmas leis.

2.2.3 Probabilidade e aleatoriedade: conceitos epistemológicos, não ontológicos

Para além da existência de eventos *acausais* (cuja existência nos parece uma impossibilidade), verificamos também que o libertarismo causal defende que alguns eventos são probabilísticos. Aliás, por serem probabilísticos são aleatórios, o que no entender desta filosofia, refuta o determinismo.

Como foi anteriormente referido, no entendimento dos filósofos defensores do libertarismo causal (Kane, 2000, 2005; Balaguer, 2004, 2014), se os eventos são aleatórios e probabilísticos, então é porque o determinismo é falso. Esta visão indeterminística do Universo permitiria, como vimos, abrir mais possibilidades ao livre-arbítrio já que, segundo esta visão ontológica, o futuro estaria em aberto, ao contrário do que se sucederia no determinismo (Brieger e Muller, 2015). Nas palavras de Mark Balaguer:

“...a física quântica possui regras probabilísticas e algumas das mais importantes e mais aceites interpretações da mecânica quântica adotam essas leis para capturar indeterminações irreduzíveis que existem no nível mais fundamental da natureza⁴⁰” (2010, p. 68).

No mesmo sentido, Robert Kane diz-nos que a física quântica introduziu:

“indeterminismo ou o acaso no mundo físico” e que “Muitos dos comportamentos das partículas elementares [...] não é completamente previsível e apenas pode ser explicado por leis estatísticas, não determinísticas⁴¹” (2004, p. 8).

São essas leis probabilistas que, segundo Balaguer e Kane, podem governar os sistemas neuronais e, assim, possibilitar o livre-arbítrio. No entanto, parece-nos, desde logo, que nestas afirmações existe, claramente, uma interpretação ontológica de um indeterminismo que é, sobretudo, epistemológico (Rovelli, 2022; Smolin, 2019; Hooft, 2016). Como temos vindo a referir, é um erro comum, mesmo entre filósofos e cientistas, confundirem determinismo com previsibilidade (Kaku, 2021; Kane, 2005; Balaguer, 2004; Ginet, 2008; O’Connor, 2000, 2000a). Nada nas conclusões indeterministas da física quântica, mesmo no âmbito da interpretação e Copenhaga (Bohr, 1958; Rovelli, 2022; Smolin, 2019) fundamenta ostensivamente uma ontologia mas, sim, uma epistemologia (Hooft, 2016; Hossenfelder, 2022).

A crença numa ontologia indeterminística está, portanto, em grande parte sustentada numa interpretação incorreta da física quântica, por um lado, e na suposição equivocada de que a aleatoriedade é sinónima de *acausalidade*, por outro. A física quântica, em rigor, não é indeterminística. *As medições, que resultam em*

⁴⁰ Tradução do autor.

⁴¹ Tradução do autor.

probabilidades, é que o são. De facto, um elemento importante (que raramente é mencionado) da interpretação de Copenhaga é que apenas podemos perguntar o *resultado* da experiência (Hooft, 2016; Smolin, 2019), pelo que esta interpretação é um *modelo* de uma experiência probabilística – que, aliás, tem na probabilidade a sua essência.

O modelo de Copenhaga é, aliás, um algoritmo cujo processo é, intrinsecamente, probabilístico, pelo que a probabilidade é inerente à interpretação de Copenhaga, não à física quântica em si (Fuchs & Peres, 2000). Na verdade, segundo Hooft, não existem sobreposições quânticas – apenas modelos matemáticos que nos ajudam a idealizar (erradamente) o Universo. A função de onda na interpretação de Copenhaga não é, aliás, uma entidade real: “trata-se de um instrumento de cálculo” (Rovelli, 2018, p. 69). Neste sentido, as conclusões da interpretação de Copenhaga têm muito mais a ver com o nosso parco conhecimento da realidade do que com o que acontece na realidade. Por isso é que a sobreposição do gato de Schrödinger⁴² é aceitável como modelo de medição, mas incapaz de “descrever um estado que ocorre no mundo real⁴³” (Hooft, 2016, p. 36). Ou seja, conceitos como sobreposição, colapso de onda e medição não são mais do que instrumentos que nos permitem compreender uma realidade muito mais complexa.

O que é uma probabilidade? Segundo o físico Carlo Rovelli no seu livro *O Abismo Vertiginoso* (2022): “Falamos de probabilidades quando não temos todos os dados do problema e não sabemos com certeza o que acontecerá” (p. 36). A probabilidade é um conceito que nasce da incerteza – uma característica refém do entendimento humano

⁴² O chamado “gato de Schrödinger” é uma experiência mental proposta pelo físico austríaco Erwin Schrödinger, em 1935, para ilustrar os paradoxos da mecânica quântica, em particular a noção de superposição e o problema da medição. A experiência consiste em imaginar um gato colocado dentro de uma caixa fechada, juntamente com um mecanismo que contém um átomo radioativo, um detetor de radiação, um martelo e um frasco com veneno. Se o átomo decair (algo que ocorre de forma probabilística), o detetor regista a emissão, o martelo parte o frasco e o veneno é libertado, matando o gato. Se o átomo não decair, nada acontece e o gato permanece vivo. Segundo a interpretação dominante da física quântica (a interpretação de Copenhaga), antes de qualquer observação, o sistema encontra-se numa superposição de estados: o átomo está simultaneamente decaído e não decaído, e, por consequência, o gato está simultaneamente vivo e morto. Só no momento em que o observador abre a caixa — isto é, quando se realiza uma medição — é que a função de onda colapsa para um dos estados possíveis, determinando se o gato está vivo ou morto (Rovelli, 2022).

⁴³ Tradução do autor.

– pelo que apenas poderá ser aplicada em modelos epistemológicos do mundo, não a modelos ontológicos. O *Dicionário Priberam de Língua Portuguesa* define probabilidade como:

“Quociente entre o número de casos favoráveis à ocorrência de um acontecimento ou fenómeno e o número total de casos possíveis; número que expressa o grau de aleatoriedade da ocorrência de um acontecimento ou fenómeno, medido numa escala de zero a um, da impossibilidade à certeza”⁴⁴.

No mesmo sentido, o filósofo Alan Hájek define probabilidade da seguinte forma:

“It assigns probabilities in the absence of any evidence, or in the presence of symmetrically balanced evidence. The guiding idea is that in such circumstances, probability is shared equally among all the possible outcomes, so that the classical probability of an event is simply the fraction of the total number of possibilities in which the event occurs” (2023).

Em poucas palavras, a probabilidade é uma ferramenta de medida; é um instrumento de previsão utilizado, designadamente, pela física quântica e pela matemática. Tanto a física quântica, como a matemática que é nela aplicada, são um modelo do mundo, não o mundo em si. Como refere Gerard’t Hooft em *The Cellular Automaton Interpretation of Quantum Mechanics* (2016), o Universo possui uma ontologia desconhecida, sobre a qual utilizamos a matemática visto ser uma ferramenta conveniente, não obstante as suas limitações na descrição da realidade (Hooft, 2016).

É importante notar que probabilidades não são propriedades intrínsecas da natureza, mas ferramentas epistemológicas que descrevem o nosso conhecimento limitado do sistema estudado. Como referem os físicos Chris Fuchs e Asher Peres no seu influente artigo *Quantum theory needs no interpretation*:

“Quantum theory does not describe physical reality. What it does is provide an algorithm for computing probabilities for the macroscopic events (“detector clicks”) that are the consequences of our experimental interventions. This strict definition of the scope of quantum theory is the only interpretation ever needed, whether by experimenters or theorists” (Fuchs & Peres, 2000).

44 “Probabilidade” in Dicionário Priberam da Língua Portuguesa [em linha], 2008-2024, <https://dicionario.priberam.org/probabilidade>

A interpretação probabilística da física quântica também é muitas vezes mal compreendida. Alguns físicos consideram, aliás, que a física quântica é uma ciência extremamente eficiente, mas incompleta. Como refere o físico Lee Smolin em *Einstein's Unfinished Revolution*: “Indeed, the fact that quantum mechanics only makes sense when read as a theory of a part of the universe is, by itself, a sufficient reason for regarding quantum mechanics as incomplete” (2019, [s.p.]), algo que Einstein já tinha identificado nos anos 30 (Smolin, 2019).

A probabilidade é, numa palavra, um cálculo. É, por isso, um conceito epistemológico e, ao mesmo tempo, uma ferramenta científica que visa o *conhecimento* (Romeijn, 2022). Podemos, aliás, dizer que o que é probabilístico são as medidas, não a natureza da física quântica que continua a ser regida por leis determinísticas (Hooft, 2016; Krauss, 2017; Hossenfelder, 2022). Como refere Gerard't Hooft: “every particle behaves in a completely deterministic way, even though its behavior cannot be determined completely by any known observation technique” (2017, p. 13). Numa palavra: são os modelos – que Hooft designa por *template* – que são indeterminados, não a realidade em si. A aparente aleatoriedade quântica é epistemológica (Hooft, 2016).

E o que é aleatoriedade? O *Dicionário Priberam de Língua Portuguesa* refere que a aleatoriedade é uma probabilidade⁴⁵. Segundo o filósofo Antony Eagle:

“Randomness, as we ordinarily think of it, exists when some outcomes occur haphazardly, unpredictably, or by chance. These latter three notions are all distinct, but all have some kind of close connection to probability” (2021).

Ou seja, se a probabilidade é um conceito de medida, a aleatoriedade é um conceito que *qualifica* acontecimentos incertos, sujeitos ao acaso e que não cumprem uma ordem. Os eventos aleatórios, como são incertos e desconhecidos, não se podem determinar, pelo que se diz que são probabilísticos – não necessariamente indeterminísticos no sentido de serem *acausados*. Por conseguinte, se a probabilidade é, sobretudo, um conceito de medida, a aleatoriedade é um conceito explicativo do mundo. Tanto a probabilidade como a aleatoriedade são, portanto, conceitos

45 “Aleatoriedade” in Dicionário Priberam da Língua Portuguesa [em linha], 2008-2024, <https://dicionario.priberam.org/aleatoriedade>

epistemológicos. Não são conceitos ontológicos, pelo que não provam, de todo, que o Universo é indeterminístico, isto é, que existem eventos sem causas e que não obedecem às leis da Natureza. Assim, se probabilidade e aleatoriedade são epistemológicas, então o indeterminismo ontológico colapsa. Os filósofos libertários utilizam conceitos epistemológicos para fundamentar um indeterminismo ontológico, o que nos parece profundamente incoerente.

Por exemplo, quando lançamos um dado com 6 faces, podemos dizer que temos a probabilidade de 1 em 6 de nos sair, por exemplo, um 4. Dizemos que o resultado é aleatório simplesmente porque não conseguimos *determinar* o seu desfecho. É probabilístico porque sabemos apenas que em 6 vezes que atiramos o dado, temos a probabilidade de *uma* delas ser um quatro. Para podermos medir as probabilidades, podemos elaborar modelos matemáticos, que nos permitem ajudar a apurar o comportamento do dado, nomeadamente estudar todas as variáveis que influenciam o comportamento do dado, embora não tenhamos acesso a todas essas variáveis (não só do dado como do estado do Universo num determinado momento), o que não nos permite chegar a probabilidades muito seguras. Contudo, se do ponto de vista epistemológico a probabilidade de sair um determinado número é de 1 em 6 do ponto de vista ontológico, a probabilidade de sair um resultado específico é de 1 (ou seja, é de 1 em 1). As condições exatas que antecedem o lançamento dos dados vão causar, deterministicamente, um determinado resultado. O nosso desconhecimento, aliado à nossa incapacidade de prevermos o comportamento do dado, conduz à necessidade de criar probabilidades mas, do ponto de vista ontológico, determinadas causas e leis influenciam um determinado resultado. Se, após o lançamento do dado, o resultado for um 4, dizemos que a probabilidade acertou ou falhou, pelo que esse resultado é probabilístico. Mas apenas é probabilístico para o sujeito que está a efetuar uma medição já que, do ponto de vista ontológico, o resultado é determinístico. As probabilidades são epistemológicas, ao passo que o determinismo é ontológico, pelo que *são as probabilidades que são indeterminadas* (não indeterminísticas), não a realidade em si que continua a ser determinística (Slattery, 2014).

Aleatoriedade é, portanto, um conceito que confere significado a uma *percepção* de aleatoriedade. É a percepção enviesada, originada numa ação de medição, que conduz à ideia de aleatoriedade. O ser humano, não tendo capacidade de prever

comportamentos quânticos (e clássicos), precipita-se a considera-los aleatórios, quando na realidade essa aleatoriedade é meramente uma interpretação fruto de uma ignorância epistemológica.

Para o físico Lee Smolin, as probabilidades na mecânica quântica *não* são o reflexo de uma realidade profundamente probabilística ou indeterminística, mas sim uma maneira de expressar a limitação de nosso conhecimento sobre os sistemas quânticos. Ele acredita que a realidade subjacente tem de ser determinística ou ter uma estrutura mais profunda do que a que a física quântica tradicional sugere, e que as probabilidades surgem da impossibilidade da teoria atual mais convencional (a interpretação de Copenhaga) em descrever completamente os sistemas quânticos na sua totalidade. Smolin (2019) diz, aliás, que David Bohm (1952) já resolvera, pelo menos parcialmente, o problema do aparente indeterminismo da física quântica através de um modelo determinístico, testado, realista e coerente. Como refere Lee Smolin na sua obra *Einstein's Unfinished Revolution: The Search for What Lies Beyond the Quantum*:

“In 1952, David Bohm solved the biggest of all problems in quantum mechanics, which is to provide an explanation of quantum mechanics. Unfortunately, it is widely underappreciated. It achieves something that was often (before and even after 1952) claimed impossible: to explain the rules of quantum mechanics through a coherent picture of microscopic reality” (2019, s.p.)

As medições na física quântica envolvem problemas metodológicos, desde logo a extrema sensibilidade destas partículas que têm de ser medidas em ambientes extraordinariamente controlados para evitar que o comportamento quântico possa ser destruído pelas interações do ambiente. Estas, aliás, podem explicar tanto as dificuldades para medir o comportamento quântico como algumas conclusões bizarras no âmbito das medições quânticas. De facto, até pequenas moléculas de ar ou de luz podem perturbar as medições quânticas, o que amplifica a dificuldade na medição. É a *percepção* de indeterminismo que conduz a ideias indeterministas.

Em contraste, não temos percepção de determinismo porque, como Laplace sugeriu, as nossas mentes não têm capacidade de conhecer todas as forças, leis e sistemas que envolvem o Universo. O determinismo existe — e tem de existir — porque,

num sistema regido por leis, que é simultaneamente constituído por essas mesmas leis, é inevitável que todos os eventos que nele ocorrem sejam causados por essas leis. Por conseguinte, onde existem eventos, esses eventos, para existirem, tiveram de ser causados. Por outras palavras, ou vivemos num sistema físico e as leis da Natureza causam eventos, ou não vivemos num sistema físico, e as leis da Natureza não podem causar eventos devido à não-fisicalidade do sistema — o que significa que essas leis não poderiam causar coisa alguma, nomeadamente leis físicas, pelo que a existência de um Universo como o conhecemos seria uma impossibilidade. Admitindo que vivemos num mundo físico, o determinismo é a única teoria que pode dar sentido a este sistema, pelo que o determinismo é muito mais que uma percepção ou um conceito — é uma lei⁴⁶, uma lei tão ubíqua e tão basilar para o funcionamento do Universo como o evolucionismo para o funcionamento dos organismos (Honderich, 1988). O filósofo John Earman, aliás, designa o determinismo de doutrina ontológica no sentido de ser, mais do que uma teoria, um princípio sob o qual o Universo evolui (2004).

Inversamente, a aleatoriedade é uma percepção que nasce da nossa incompreensão do sistema em que estamos inseridos. A aleatoriedade não é nenhuma lei porque é um conceito epistemológico (é uma percepção ou interpretação), não um princípio ontológico. É uma teoria fundamentada numa semântica que pretende descrever ou conceptualizar a realidade mas não é a realidade em si. Como postula o filósofo Ted Honderich em *Mind and the Brain*:

“Such connections, called nomic or lawlike connections to distinguish them from logically necessary connections rooted in words and other symbols, are pervasive in the natural world, and fundamental to it.” (1988, p. 1).

A realidade faz-se de causas necessárias e é essa mesma necessidade que possibilita a realidade. Isto porque se a realidade existe é porque foi causada. No mesmo sentido, se o ser humano existe é porque foi causado, da mesma forma que se o livre-arbítrio existe é porque foi causado (Serrado, 2025). Pela necessidade de existirem, essas conexões necessárias são, segundo Honderich, “conexões não no meio das

⁴⁶ Ou pelo menos, a abstração de uma lei.

palavras [não são conexões lógicas ou semânticas] mas conexões no mundo real. Elas são o cimento do Universo.”⁴⁷ (1988, p. 1)⁴⁸.

Em suma, não só a aleatoriedade como a probabilidade são conceitos epistemológicos que nada nos dizem sobre a ontologia do Universo – apenas nos informam em relação à nossa percepção da realidade – tal como o funcionamento quântico é uma representação da realidade, não a realidade em si.

2.2.4. *O mito da física quântica como justificação do libertarismo*

No debate filosófico sobre o determinismo e o libertarismo, é frequente encontrar-se uma confusão entre aleatoriedade, probabilidades e as reais implicações da física quântica, como analisámos. Esta confusão estende-se também à compreensão do funcionamento neurofisiológico, dando origem a especulações libertárias desprovidas de fundamentação filosófica e empírica. Assim, o indeterminismo quântico levou alguns filósofos a defender, como vimos, que a própria natureza do Universo poderia ser indeterminística, mas também a especular que o funcionamento neurofisiológico, supostamente ancorado nessa incerteza quântica, poderia também seguir um regime indeterminista (Kane, 2005; Balaguer, 2004).

Outros autores (Briegel & Muller, 2015), na mesma linha, têm vindo a propor que o indeterminismo quântico pode estar presente não só no funcionamento da memória e na formação de ideias, como também na constituição de outros processos fisiológico, como as enzimas. No entanto, essa correlação é infundada. Não há qualquer evidência de que os fenómenos quânticos possam ser diretamente aplicáveis, nem ao funcionamento neuronal nem tão-pouco ao mundo macroscópico humano (Damásio, 2010; Gazzaniga, 2011; Eagleman, 2012; Humphrey, 2012, 2023; Seth, 2021; Banks,

47 Tradução do autor.

48 É importante referir, ainda, que o facto de uma partícula (como um fóton) ter um comportamento aleatório não significa que objetos com uma composição atómica – isto é, que são compostos por partículas, como tudo o que conhecemos – possam comportar-se de igual modo. É perfeitamente possível e natural que um eletrão se comporte de uma forma isoladamente, e se comporte de outra forma se estiver fundido com outras partículas. O hidrogénio e o oxigénio, isoladamente, têm um comportamento físico e químico muito diferente de quando estes átomos estão fundidos e criam a água. A própria água possui comportamentos distintos, como todos aprendemos em tenra idade – os estados líquido, gasoso e sólido – pelo que o comportamento imprevisível de uma partícula isolada não nos diz absolutamente nada sobre como são os comportamentos dessas mesmas partículas quando estão fundidas com outras, como podemos observar nos objetos, em células ou proteínas.

2006; Sapolsky, 2024). Pelo contrário, os indícios sugerem que os princípios da física quântica são relevantes apenas para partículas elementares isoladas, e não para sistemas complexos como o cérebro (Smolin, 2019; Rovelli, 2018, 2022; Hooft, 2016).

Parece-nos que esses autores confundem, mais uma vez, aleatoriedade, e até processos caóticos, com indeterminação. Como muitos autores apontam (Sapolsky, 2024; Damásio, 2010; Gazzaniga, 2011; Eagleman, 2012), a complexidade do funcionamento fisiológico e neuronal é regida por leis determinísticas, ainda que imprevisíveis devido à nossa incapacidade de mensurar todas as variáveis relevantes. Como destaca Daniel Dennett:

o “caos aparente em sistemas determinísticos não é o mesmo que indeterminismo, mas uma manifestação da imprevisibilidade prática («practical unpredictability») em sistemas newtonianos complexos”⁴⁹ (2003, p. 105).

A física quântica é, como temos vindo a referir, uma ciência de medição em ambientes altamente sensíveis e controlados. Ao contrário da física clássica, cujos fenómenos podem muitas vezes ser verificados diretamente através da observação natural, a física quântica exige ambientes altamente controlados e experiências tecnicamente sofisticadas, como as realizadas em aceleradores de partículas. Esta dependência de tecnologia avançada e de contextos experimentais artificiais introduz um viés significativo na forma como os resultados são obtidos e interpretados (Hooft, 2016; Smolin, 2019).

Já a neurociência, em contraste, tem objetivos muito mais amplos do que estudar a previsão do comportamento neuronal. O facto de a física quântica ser uma ciência de medição (Hooft, 2016) – de previsão – não só influencia os resultados da própria física quântica, como acaba por incompatibilizar os resultados da física quântica com os resultados da neurociência. Esta tendência de aplicar a física quântica a processos neurofisiológicos humanos conduz a que, como foi referido, se tente aplicar o indeterminismo em processos macroscópicos (Kane, 2005; Briegel e Muller, 2015; Balaguer, 2010), o que não é válido no âmbito do experiencialismo científico.

49 Tradução do autor.

Um exemplo paradigmático do tipo de argumentação que contestamos encontra-se no artigo de Hans Briegel e Thomas Müller, *A Chance for Attributable Agency* (2015). Nele, os autores procuram sustentar um indeterminismo ontológico que possa ser transposto para o domínio da neurofisiologia. Para tal, começam por defender que fótons individuais — por exemplo, em feixes de laser — exibem um comportamento indeterminístico. Nas suas palavras: “Consider a laser beam: its individual photons are produced by stimulated emission, which is an indeterministic quantum mechanical process” (p. 264). Contudo, esta interpretação parece-nos profundamente equivocada. O comportamento que os autores classificam como indeterminístico é, na verdade, desencadeado por sistemas macroscopicamente determinísticos, como o acionamento de um laser por um operador humano. O facto de um fóton individual poder ter trajetórias probabilísticas não anula a origem causalmente determinada do processo no seu conjunto. O mundo pode ser, e é, totalmente aleatório e, ainda assim, completamente determinístico.

No mesmo sentido, Briegel e Müller (2015) afirmam que o indeterminismo quântico pode ser observado em diferentes níveis da biologia, desde a constituição de enzimas até a fenómenos cognitivos complexos, como os pensamentos e a memória. Segundo os autores, os blocos moleculares responsáveis pela síntese proteica mover-se-iam aleatoriamente no interior das células, permitindo, ainda assim, uma produção eficiente. De igual modo, os canais iónicos das membranas celulares comportar-se-iam de forma estocástica (e, segundo estes autores, subsequentemente indeterminista), possibilitando a polarização e despolarização necessárias à transmissão de sinais eléctricos (2015, p. 267). Para além destes processos neurofisiológicos, defendem que esse mesmo indeterminismo — com origem na aleatoriedade quântica — manifesta-se também no funcionamento cognitivo, sustentando que a consolidação da memória, bem como o surgimento de ideias e desejos, dependeriam de flutuações *acausais* que o sistema nervoso conseguiria, de algum modo, integrar e explorar funcionalmente (2015, p. 270). A imprevisibilidade observável nesses fenómenos é usada como argumento para atribuir-lhes uma natureza ontologicamente indeterminista. Nas palavras dos próprios autores:

“Ora, o nosso mundo é indeterminístico e, ainda assim, contém mecanismos fiáveis. As células biológicas são exemplos perfeitos da utilização bem-sucedida do indeterminismo ao serviço de um todo funcional. Os blocos construtores responsáveis pela síntese de proteínas movem-se aleatoriamente no interior da célula, mas a síntese proteica é um processo altamente eficiente. Os canais iónicos nas membranas celulares comportam-se de forma estocástica, mas a polarização e despolarização das células é um meio eficaz de transmissão de sinais. Estes são exemplos — entre muitos outros — da utilização bem-sucedida (ainda que não sem exceções) do indeterminismo, isto é, da capacidade de contrariar uma microdinâmica errática através de uma organização funcional em níveis superiores.” (2015, p. 267).

Esta conclusão, no entanto, carece de fundamento. Não só o funcionamento fisiológico, que estes autores consideram aleatório e estocástico é totalmente determinístico, como o funcionamento neuronal é sustentado por interações bioquímicas altamente organizadas e regidas por leis físicas, químicas e genéticas deterministas — como a propagação dos potenciais de ação, a libertação de neurotransmissores ou a ativação de circuitos sinápticos. Não existe, até à data, qualquer evidência empírica ou teórica robusta que comprove uma influência causalmente relevante de processos quânticos na dinâmica cerebral. Assim, extrapolar comportamentos quânticos para justificar o livre-arbítrio humano, como fazem Briegel e Müller (2015), representa uma analogia especulativa e infundada, que ignora a complexidade e a consistência determinista dos sistemas biológicos (Sapolsky, 2024). Como refere o psicólogo William P. Banks: “The attempts to use quantum theory to save libertarian freedom are myopic” (2006, s.p.).

Do mesmo modo, note-se a tendência dos autores em fundamentar o seu sistema indeterminístico de acordo com a aleatoriedade dos processos neuronais e fisiológicos. Quando afirmam que “eventos verdadeiramente aleatórios podem ser gerados através da exploração de processos físicos fundamentais, como o decaimento radioativo de um núcleo ou a passagem de fotões individuais por um *beam splitter*⁵⁰” (Briegel & Müller, 2015, p. 269), os autores incorrem, ostensivamente, em duas imprecisões: 1) nada nos processos aleatórios acima descritos é, efetivamente, indeterminístico, uma vez que são comportamentos governados e causados por

⁵⁰ Tradução do autor.

processos físicos; 2) ao sugerirem a existência de um mundo “verdadeiramente aleatório”, estão a realizar uma extrapolação indevida de uma ferramenta epistemológica, como o conceito de aleatoriedade — intrinsecamente ligada à previsão dos eventos — para a descrição ontológica de um sistema.

Os processos fisiológicos não operam segundo leis indeterminísticas, nem os neurónios se regem por qualquer forma de aleatoriedade ontológica (Sapolsky, 2024; Damásio, 2010; Gazzaniga, 2011; Eagleman, 2012; Humphrey, 2012, 2023; Seth, 2021; Banks, 2006). Pelo contrário, tais processos estão fortemente condicionados por mecanismos genéticos e biológicos bem definidos (Sapolsky, 2024; Damásio, 2010; Gazzaniga, 2011; Eagleman, 2012; Humphrey, 2012; Seth, 2021; Crick, 1998; LeDoux & Brown, 2017; LeDoux, 2020; Edelman, 1995). A nossa incapacidade de prever, com precisão, o comportamento celular ou as projeções neuronais — tal como a dificuldade em prever flutuações atmosféricas — não implica qualquer indeterminismo ontológico. Trata-se, antes, de limites epistemológicos e não de falhas ontológicas nas cadeias causais. Nesse sentido, a tendência de aplicar interpretações indeterministas da física quântica a sistemas biológicos complexos, como o cérebro humano, revela-se conceptualmente problemática e cientificamente infundada.

Podemos imaginar um ser hipotético, incapaz de observar diretamente o nosso mundo clássico, que quisesse compreendê-lo. Esse ser saberia da existência do nosso mundo, mas não teria acesso direto aos fenómenos físicos tal como nós os percebemos. Para estudar esta realidade inacessível, ele teria de criar tecnologias avançadas que interferissem no mundo clássico e de formular modelos matemáticos que traduzissem os dados obtidos. No entanto, as suas observações seriam inevitavelmente enviesadas pelas manipulações necessárias para tornar a realidade acessível. No mesmo sentido, os seus modelos matemáticos, embora úteis para descrever padrões e criar previsões, não seriam o mundo em si, mas uma aproximação limitada. De forma análoga, os seres humanos estão para a física quântica como este ser estaria para o nosso mundo clássico. Incapazes de aceder diretamente aos pilares da realidade quântica devido às limitações do nosso aparato cognitivo e sensorial, dependemos de instrumentos altamente sofisticados e de equações matemáticas para compreender aquilo que não conseguimos observar diretamente. Essa dependência introduz enviesamentos que limitam a compreensão plena da física quântica, tal como as limitações hipotéticas daquele ser

restringiriam a sua compreensão do nosso mundo clássico. É fundamental, portanto, notar que a física quântica não refuta o determinismo no seu núcleo. Pelo contrário, muitos físicos veem-na como uma extensão determinista, embora probabilística, das leis físicas (Rovelli, 2018; Smolin, 2019; Earman, 2006).

Finalmente, é importante reconhecer a física quântica como uma ferramenta epistemológica incrivelmente eficaz, mas não como uma descrição final da realidade. Assim como os avanços tecnológicos da história permitiram avanços na compreensão de fenômenos antes considerados místicos, é provável que futuras inovações nos permitam compreender melhor os fundamentos da mecânica quântica. No entanto, essas inovações devem ser vistas como complementares, e não substitutivas, ao entendimento determinista da natureza.

Em síntese, a ideia de que a física quântica apoia teorias indeterministas no funcionamento neuronal ou fisiológico é profundamente falaciosa. Não só os resultados da física quântica são influenciados por limitações epistemológicas e tecnológicas, como também a transposição dos seus princípios para sistemas biológicos carece de total fundamento empírico. Como ocorre com outros fenômenos naturais historicamente mal compreendidos, é necessário abordar a física quântica com ceticismo crítico, evitando interpretações mirabolantes e especulações sem base sólida. Resta-nos questionar se muitos dos cientistas e filósofos não olham para a física quântica com a mesma imaginação e maravilhamento com que muitos dos nossos antepassados olhavam para outros fenômenos aparentemente mágicos (o Sol, a Lua, as estrelas, depois a eletricidade e os genes, entre muitos outros mistérios que o deixaram de o ser), criando ideias e teorias mirabolantes, provenientes de uma grande ignorância. Com o tempo, esses mistérios foram totalmente explicados e encaixados de acordo com um entendimento muito mais simples do que as especulações sugeriam, alinhado num edifício científico construído solidamente ao longo de séculos.

2.3. Revertendo o debate: como o determinismo torna o livre-arbítrio possível

Publicámos recentemente um artigo (Serrado, 2025) com uma proposta que gostaríamos de discutir no âmbito deste trabalho – a ideia de que a única forma de existir livre-arbítrio (ou *volição*)⁵¹ possa ser através de um Universo determinista.

Sabemos que os filósofos libertários rejeitam que o ser humano possa ser livre se estiver sujeito às leis da física e condicionado por causas que, de algum modo, limitem a sua liberdade (Kane, 2005; Ginet, 2008; O'Connor, 2000; Balaguer, 2014; McCann, 1998). Se o libertarismo simples e o de *causação agencial* parecem aceitar que as leis da física se aplicam a tudo, menos aos eventos mentais (libertarismo simples) ou à consciência (libertarismo de *causação agencial*), já o libertarismo causal (Kane, 2005) propõe uma versão híbrida, em que as leis do Universo geram eventos *acausais* que, depois, podem ser controlados pelo agente através de um esforço consciente – esforço esse cuja origem (a)causal nunca é verdadeiramente explicada.

Sabemos, igualmente, que os compatibilistas defendem que, apesar de o determinismo ser verdadeiro, o ser humano possui liberdade desde que, por um lado, não seja coagido (Frankfurt, 1969) e, por outro, possua um sistema nervoso que lhe permita controlar e inibir o seu comportamento, de modo a evitar eventos perigosos e nocivos (Dennett, 2003). Ora, gostaríamos de inverter a perspetiva e questionar o seguinte: e se o livre-arbítrio apenas puder existir num Universo determinista? E se o determinismo for a única teoria que permite, de facto, a liberdade individual? É isso que propomos.

Vimos que a proposição base do libertarismo pode ser formulada, por exemplo, desta forma:

1. O Universo é determinista;
2. temos livre-arbítrio;
3. logo, o determinismo tem de ser falso.

Ou assim:

⁵¹Embora sejam conceitos distintos, o livre-arbítrio e a *volição* mantêm uma relação intrínseca, uma vez que apenas podemos exercer o livre-arbítrio — isto é, a capacidade de escolher consciente e intencionalmente as nossas ações — se possuímos *volição*, entendida como a vontade consciente de decidir e agir. Por essa razão, é possível utilizar ambos os termos de forma intercambiável, sempre que o contexto o justificar.

1. A física quântica demonstrou que o Universo é indeterminista;
2. temos livre-arbítrio;
3. o determinismo é falso, pelo que o indeterminismo é o que possibilita o livre-arbítrio.

Mas julgamos que esta conclusão é precipitada. É precipitada, desde logo, porque os libertários partem da premissa de que temos livre-arbítrio, o que é, à partida, discutível – pelo menos no sentido de podermos agir sem que as nossas ações sejam causadas por eventos anteriores ou pelas leis da Natureza. E é, sobretudo, precipitada porque concluem que a única forma de existir livre-arbítrio é num Universo indeterminista – quando, muito provavelmente, o contrário é que se verifica.

De facto, os filósofos libertários apressam-se ao concluir que o determinismo inviabiliza o livre-arbítrio, porque consideram que: 1) o agente tem de ser a causa da liberdade, sem que nada (nem as leis da Natureza) a cause (Ginet, 2014; McCann, 1998); 2) o agente deve ser capaz de controlar o indeterminismo, mesmo que isso viole as leis da física (Kane, 2005; Balaguer, 2004); 3) ou o agente pode intervir nas próprias leis da Natureza através de uma consciência não-física (O'Connor, 2000; Chisholm, 2003).

Tendo isto em conta, passemos agora a uma análise mais pormenorizada das principais propostas libertárias sobre a relação entre o livre-arbítrio e o indeterminismo. Robert Kane, por exemplo, postula que o libertarismo exige a seguinte condição: “The agent should be able to act and act otherwise (choose different possible futures), given the same past circumstances and laws of nature” (2005, p. 38). Mark Balaguer defende o mesmo ao afirmar que o livre-arbítrio requer decisões humanas indeterminadas e acausais: “if our torn decisions are wholly undetermined [...] so that, at the moment of choice, which option is chosen is uncaused” (2010, p. 123).

Hugh McCann, no mesmo sentido, sustenta que o indeterminismo implica que: “...decision and action are, at least usually, exempt from nomic causation, and that if these were not so we would not be free or responsible” (1998, p. 8). Carl Ginet, de forma semelhante, refere que: “...a choice that is up to the agent at the time it occurs cannot be caused, even indeterministically” (2014, p. 24).

Em poucas palavras, os libertários sustentam uma ideia central: *o ser humano só pode ser livre se violar as leis da Natureza*, isto é, se o evento designado por livre-arbítrio (a volição, uma decisão ou uma ação) não for causado por nenhum evento anterior.

Ora, estas ideias apresentam três problemas centrais: 1) partem do princípio de que o livre-arbítrio só pode existir num Universo indeterminista, o que é insustentável, como veremos; 2) assentam numa noção de livre-arbítrio como sendo *acausal*, o que constitui um absurdo filosófico, pois não podem existir eventos absolutamente *acausados*; 3) baseiam-se numa conceção de agente (ou *self*), frequentemente dualista ou não-fisicalista, que interage com o mundo físico – o que é intrinsecamente contraditório.

2.3.1. *A necessidade de um Universo determinista*

Já aprofundámos a ideia de que, num Universo regido por leis, nada pode existir que não seja causado. Se algo existe, é porque foi trazido à existência por uma causa — isto é, por alguma lei do Universo (Hooft, 2016, 2017). Esta tese não é apenas uma especulação metafísica, mas uma consequência necessária da observação empírica de todos os fenómenos naturais, em todas as áreas da ciência. A causalidade ontológica necessária que aqui propomos não é apenas um princípio filosófico abstrato, mas a própria condição de inteligibilidade da realidade física e da existência de qualquer fenómeno organizado.

Na física (Rovelli, 2018; Hooft, 2016; Krauss, 2017; Hawking e Mlodinov, 2011), por exemplo, não se conhece um único evento que não tenha origem em condições antecedentes — seja a formação de partículas subatómicas, a agregação de matéria em estrelas ou a propagação da radiação eletromagnética. Mesmo na mecânica quântica (Rovelli, 2022; Smolin, 2019), frequentemente invocada como exceção, o comportamento aparentemente aleatório das partículas não invalida a estrutura causal do sistema como um todo. As leis que regulam o colapso da função de onda, por mais probabilísticas que nos pareçam, só operam em ambientes tecnicamente preparados e altamente controlados, o que reforça a ideia de que não são fenómenos *acausais*, mas causalmente condicionados por estruturas que ainda não compreendemos completamente.

Na biologia (Sapolsky, 2024; Ledoux, 2020; Curado, 2017; Damásio, 2010), a causalidade é ainda mais evidente. A formação de células, tecidos, órgãos e organismos obedece a regras químicas e genéticas precisas. O ADN, os processos de transcrição e tradução, a expressão génica e a própria embriogénese são exemplos de cadeias causais

altamente organizadas, regidas por leis físico-químicas deterministas. Mesmo o fenómeno da mutação genética, muitas vezes invocado como exemplo de aleatoriedade, não escapa à causalidade ontológica: as mutações ocorrem por radiações, erros de replicação, exposição a agentes químicos — causas todas elas rastreáveis e, em última instância, sujeitas às leis naturais.

Na neurociência (Seth, 2021; Dehaene, 2014; Swaab, 2014; Damásio, 2011; Gazzaniga, 2000 e 2011; Eagleman, 2012; Edelman, 1995), o princípio da causalidade manifesta-se na organização funcional do sistema nervoso. Cada decisão consciente, cada memória evocada, cada emoção sentida é acompanhada por um conjunto de alterações neurofisiológicas que decorrem de processos bioquímicos regulares. A libertação de neurotransmissores, a ativação de circuitos sinápticos, a propagação de potenciais de ação — todos estes fenómenos obedecem a leis determinísticas. A imprevisibilidade de certos comportamentos humanos não se deve à ausência de causas, mas sim à complexidade do sistema e à nossa limitação em medir com exatidão todas as variáveis envolvidas. Como refere Dennett (2003), a imprevisibilidade prática em sistemas deterministas complexos não equivale, de forma alguma, ao indeterminismo.

Portanto, em todas as ciências empíricas — da física à biologia, da química à neurociência — não há um único exemplo credível de um fenómeno *acausal*. A causalidade não é um postulado externo à realidade, mas uma propriedade intrínseca da própria estrutura ontológica do Universo. Se o Universo é um sistema físico regido por leis, então tudo o que nele ocorre tem de ser causado por essas leis, independentemente da nossa capacidade de prever ou medir os efeitos. A existência de qualquer evento, para ser pensável e para poder ser descrita cientificamente, exige uma causa — e é esse encadeamento causal que constitui o próprio tecido da realidade.

Neste sentido, como é que o livre-arbítrio pode existir no nosso Universo sem que tenha sido causado? E como é que um sistema ontologicamente indeterminista⁵²

⁵² Filósofos libertários (Kane, 2005; Balaguer, 2004; Ginet, 2014) referem frequentemente, como vimos, que o nosso Universo é aleatório. No entanto, consideramos que o conceito de aleatoriedade (ou acaso) não se aplica à ontologia do Universo, uma vez que é um conceito estritamente epistemológico, isto é, relacionado com os limites da nossa perceção e do nosso conhecimento causal, e não com a inexistência de causalidade na realidade. Por isso, no respeito à ontologia do Universo, utilizamos o termo

pode garantir um livre-arbítrio que, segundo os filósofos libertários, não só é um evento *acausado* como é um evento que, por ser *acausado* e existir num Universo indeterminista, não pode garantir coisa alguma devido à irregularidade dos eventos e subsequente instabilidade do próprio sistema? Ou seja, de que forma podemos ter livre-arbítrio num Universo indeterminista que gera eventos *acausados*?

No final do século XIX, o médico, psiquiatra e filósofo Miguel Bombarda sugeriu uma ideia muito interessante, que expunha as fragilidades do indeterminismo. Do ponto de vista do comportamento humano, o determinismo seria, no entender de Bombarda, crucial para o equilíbrio e existência do Universo, já que esse determinismo não era mais do que a funcionalidade dinâmica das leis que faziam o Universo funcionar de forma estável, regular e sistémica. Sem essas leis – isto é, sem a determinação de um conjunto de princípios fundamentais invariáveis –, o Universo não poderia possuir equilíbrio e, subsequentemente, vida. Miguel Bombarda questionou como seria se o Universo deixasse de estar regido por leis precisas:

“O mesmo corpo seria n’este momento transparente para ser opaco passados instantes; estas e aquellas vibrações manifestar-se-hiam agora como electricidade para logo se offerecerem sob a fórma de calor ou magnetismo...” (Bombarda, 1902, p. 13).

Seria a desordem absoluta, e a vida — possivelmente o próprio Universo — seria de impossível existência. O mesmo se passaria no cérebro. Imaginemos um ser humano sujeito a “leis” indeterministas (“leis indeterministas” é, em si mesmo, um paradoxo, na medida em que apenas um sistema determinista pode gerar eventos regulares, como leis, mas utilizamos a expressão para podermos imaginar um sistema governado pela aleatoriedade). Segundo Bombarda:

“O mesmo individuo indifferentemente praticaria o mal e o bem, conduzir-se-hia agora d’um modo para logo se conduzir de outro inteiramente diverso... Tudo isto seria feito independentemente da organização mental, independentemente da educação e quaisquer outros motivos de determinação.” (Bombarda, 1902, p. 20)

indeterminismo — com todas as ressalvas que ele implica — e não aleatoriedade ou acaso, por considerarmos que estes últimos expressam apenas a ignorância do observador, e não a estrutura do real.

Concordamos inteiramente com o médico-filósofo. Sem leis — isto é, num Universo ontologicamente indeterminista — não seria possível garantir estabilidade causal nem a regularidade dos eventos necessária à emergência e manutenção de fenómenos estáveis (Tegmark, 2014). Isto aplica-se, em particular, ao comportamento humano, que, se deixasse de estar ancorado às regularidades causais, tornar-se-ia um veículo de imprevisibilidade e instabilidade — precisamente o contrário do que exige o exercício do livre-arbítrio.

O filósofo Daniel Dennett, passados mais de 100 anos, apresenta-nos uma ideia semelhante. Dennett questiona como é que o acaso ou um evento indeterminado pode conferir liberdade ao ser humano, já que este, se quer ser livre, quer que as suas escolhas sejam determinadas por ele — pelas razões do agente — e não por eventos indeterminados sobre os quais não tem qualquer controlo. Se o indeterminismo fosse verdadeiro, o agente seria apenas um elemento no Universo sujeito às “leis” da indeterminação e da sua subsequente instabilidade, sem qualquer capacidade de intervir no seu destino, já que este, ao ser indeterminado, seria igualmente irregular e instável (Dennett, 2003). Dito de outra forma: se o sujeito realizou X porque o Universo é indeterminístico, isso apenas significa que o sujeito realizou X porque a causa foi Y (um evento indeterminístico — *acausal*). Se isto for verdade, qual o controlo do ser humano num Universo indeterminístico — *acausal*? Qual a liberdade do ser humano num sistema desordenado e instável, sem leis causais? (Pereboom, 2007).

De facto, sem um sistema ontologicamente determinista, não só o Universo seria um sistema instável, desordenado, sem leis que causassem coisa alguma — e os eventos suceder-se-iam sem ordem nem qualquer relação causal — como o sujeito que existiria nesse sistema seria, ele próprio, um veículo de irregularidade, instabilidade e imprevisibilidade. Querer-se-ia agora uma coisa para querer o seu oposto passados uns segundos; desejar-se-ia ir para a direita para depois decidir ir para a esquerda e, logo de seguida, querer voltar para trás; acreditar-se-ia num momento que se queria ajudar uma senhora a ser assaltada, para depois acreditar que seria melhor deixá-la ser assaltada e seguir para uma reunião mais importante, para logo de seguida apetecer roubar um carro e ir para a praia.

Num sistema ontologicamente indeterminista, governado por leis *acausais* (mesmo admitindo, por hipótese, a possibilidade da sua existência — o que nos parece

conceitualmente impossível), os eventos ocorreriam *acausalmente*, de forma irregular, instável e sem continuidade estrutural. O indeterminismo afetaria todos os eventos do Universo, incluindo os processos cerebrais, que seriam invadidos por ocorrências *acausais*, tornando-se, por isso, irregulares e desestruturados (Hooft, 2017; Bombarda, 1902; Dennett, 2003; Pereboom, 2007). Em contraste, um Universo ontologicamente determinista proporciona estabilidade e estrutura suficientes para permitir mudança e evolução sustentadas por leis dinâmicas que, por serem fixas mas flexíveis, possibilitam regularidades sobre as quais podem emergir fenômenos complexos.

A vida biológica, nomeadamente o sistema nervoso, é fruto de um Universo a laborar há milhares de milhões de anos, segundo leis que possibilitaram uma regularidade e estabilidade suficientes para que, pontualmente (pelo menos uma vez), a vida pudesse surgir e florescer (Dennett, 2003). Se, num Universo determinista e regido por leis estáveis, a vida não abunda — pelo menos no cosmos que nos é conhecido —, podemos imaginar o quão improvável seria a emergência da vida num Universo indeterminista, caracterizado por instabilidade causal, ausência de regularidade e desestruturação sistémica.

2.3.2. *O livre-arbítrio tem de ser um evento causado*

Parece-nos, portanto, insustentável considerar que possa existir liberdade num Universo indeterminístico pela simples razão que um Universo indeterminístico, pela sua ontologia *acausal*, não pode causar nada. Ora:

- 1) se o Universo é ontologicamente indeterminista;
- 2) se alguma coisa para existir tem que ser causada;
- 3) se o Universo indeterminista não pode causar nenhum evento, não só pela sua natureza *acausal*, como pela instabilidade inerente a um Universo aleatório;
- 4) então, não só um Universo ontologicamente indeterminístico não pode causar alguma coisa, como o livre-arbítrio não pode existir se não for causado por alguma coisa.

O livre-arbítrio (ou a volição) só pode existir se for causado e apenas um sistema ontologicamente determinístico — regular, estável e estruturado — pode causar eventos

estáveis mas flexíveis. Dito de outra forma, gostaríamos de sugerir que não é possível o livre-arbítrio – que é um evento que existe no Universo – existir se não for *causado*, isto é, se não for trazido à existência por uma causa, o mesmo é dizer pelas leis do Universo. Como um Universo indeterminístico não só é instável e irregular, como, precisamente por isso, não pode gerar eventos estáveis nem sustentar leis, então tal Universo não pode causar nada de estável ou permanente — e, por conseguinte, o livre-arbítrio não pode existir nesse Universo.

Se o livre-arbítrio não pode existir caso não seja causado – não pode existir em nenhum Universo sem leis (eventos causados segundo determinadas leis constantes) – então apenas num Universo que cause eventos de acordo com leis o livre-arbítrio pode existir. Só num Universo criador de eventos constantes – como *quarks*, neutrões, átomos, células, animais, cérebros e consciência – através de leis permanentes é que o livre-arbítrio pode existir. Podemos dizer o seguinte:

- 1) o Universo causa eventos através de leis regulares;
- 2) se o Universo causa eventos através de leis regulares é porque é determinístico;
- 3) o livre-arbítrio existe;
- 4) se o Universo causa eventos através de leis constantes é porque é determinístico, pelo que se o livre-arbítrio existe teve de ser causado;
- 5) logo, *o determinismo causou o livre-arbítrio*.

É esta a nossa proposta: que apenas num Universo que cause eventos segundo leis regulares, sistémicas ou determinísticas podem existir eventos mentais, como desejo, vontade e volição – portanto, livre-arbítrio. O neurocientista Anil Seth refere que não podemos ter liberdade em relação às leis da física e do Universo (Seth, 2021). Concordamos, mas sugerimos que apenas as leis da física nos podem dar liberdade. Esta nossa ideia entronca numa afirmação do neurocientista Gerald Edelman que nos parece basilar: “Free will – whatever you think about it, we’re determined to have it” (in Seth, p. 201). Ou seja, são as leis da física que permitem a existência do livre-arbítrio.

Sugerimos, também, que o cérebro é um evento causado pelas leis da Natureza que, por sua vez, causa o sentimento de liberdade de ação — tal como causa muitos outros sentimentos, como a alegria, a fome ou a sede. O que designamos por “livre-

arbítrio” poderá ser compreendido, neste enquadramento, como um *sentimento* gerado pelo cérebro, indissociável de outros estados subjetivos, como o desejo, a vontade ou a volição (Wegner, 2002). Neste sentido, defendemos que o livre-arbítrio é real enquanto efeito causal de um cérebro capaz de gerar esse sentimento. Claro está que esse mesmo cérebro foi causado por processos biológicos, nomeadamente por um genoma, que por sua vez resulta do cruzamento de dois genomas anteriores — numa cadeia contínua de causalidade que remonta ao início do Universo (Sapolsky, 2024). Neste sentido, o nosso futuro poderá ter sido inscrito no Big Bang, mas isso não invalida o livre-arbítrio: 1) porque o livre-arbítrio é um evento causado — e, por isso, existe; 2) e porque o sujeito possui livre-arbítrio na medida em que possui um sistema nervoso capaz de o gerar.

Daniel Dennett apresenta uma ideia semelhante à nossa ao postular que o livre-arbítrio é uma propriedade *criada* pela evolução que, com o tempo, aperfeiçoou-se, designadamente nos seres humanos. Para o filósofo, somos livres a partir do momento em que o nosso cérebro nos permite sermos responsáveis pelos nossos atos e realizar as nossas decisões com sentido moral. Por isso, o sujeito quer que as suas decisões sejam *determinadas* pelas suas razões e intenções (que são deterministicamente causadas) e não que sejam indeterminadas, porque isso significaria que o ser humano seria dominado pela aleatoriedade e pela imprevisibilidade, o que não só é o contrário do livre-arbítrio como é o oposto do que é desejável por qualquer ser humano. Por outro lado, Dennett considera que o determinismo não é uma lei de inevitabilidade, já que o ser humano pode evitar uma série de acontecimentos — capacidade que, aliás, lhe confere livre-arbítrio. Ou seja, a existência daquilo a que Dennett chama *avoiders* é uma demonstração de que o livre-arbítrio existe mesmo num universo determinístico⁵³ (2003). Na mesma linha de Dennett, o físico e biólogo Paul Davies diz-nos, no seu livro *The Demon in the Machine*: “To get free will we don’t really want indeterminism: we want our will to *determine* our actions” (Davies, 2019, p. 202). Por isso, o ser humano é constituído por uma identidade estável ao longo da vida — pelo menos nas identidades saudáveis.

53 Este argumento parece-nos fraco, já que a existência de *avoiders* não invalida que esses *avoiders* tenham liberdade para não evitar o que está determinado. Ou seja, é perfeitamente plausível que um Universo determinista tenha determinado a existência de *avoiders*, até por uma questão de sobrevivência dos organismos no âmbito da evolução.

Todos nós podemos imaginar o quão dramática pode ser a vida de uma pessoa com múltipla personalidade, com perturbações de personalidade que manifestam uma crónica instabilidade do eu evidente. É por isso que não queremos que as *nossas* decisões e ações não sejam outras para além daquelas que queremos determinar (como alega o indeterminismo), mas que sejam, isso sim, aquelas que possamos determinar em função da nossa identidade e da nossa personalidade, mesmo que essas decisões, essa identidade e essa personalidade sejam determinadas pelas leis da Natureza que atuam na biologia do nosso cérebro – algo de que, na maior parte das vezes, nem estamos conscientes. Somos livres, sim, mas é porque fomos determinados a ser livres.

Por conseguinte, propomos que o sujeito é livre desde que possua um cérebro funcionalmente saudável, capaz de gerar comportamentos em conformidade com um eu estável e estruturado — mesmo que essa regulação neuronal anteceda e cause a própria consciência. O livre-arbítrio surge, portanto, de um cérebro dotado de redes neuronais que possibilitem (diríamos, que *causam*) a existência do sentimento de volição, isto é, de liberdade de ação (Wegner, 2002). Um sentimento extremamente importante para manter a identidade – o eu – do sujeito, mas, ainda assim, um sentimento causado. Ou seja, o sujeito é livre porque pode, voluntariamente, expressar a sua volição, mesmo que não controle essa volição. Dito de outra forma, o ser humano é livre mas não controla a sua liberdade, simplesmente porque a sua liberdade foi-lhe imposta – foi causada pelas leis da física, isto é, pelo seu sistema nervoso.

O sujeito pode, portanto, fazer o que desejar. No entanto, o sujeito não pode deixar de desejar o que deseja porque *não tem* liberdade de deixar de desejar o que deseja. Isto é, o desejo sendo causada pelo cérebro, não permite que o sujeito deixe de desejar o que desejou. No mesmo sentido, a volição, sendo causada pelo cérebro, não permite que o sujeito deixe de sentir que tem volição – que tem livre-arbítrio. Mesmo que, como sugere o indeterminismo de causação agencial, o sujeito atue sobre essa vontade ou essa volição; mesmo que o sujeito possa contrariar os seus desejos e decidir não agir de acordo com os mesmos, essa decisão (essa suposta liberdade que contraria um desejo inicial) é igualmente *causada* pelo cérebro na forma de um outro desejo – o desejo de contrariar o desejo inicial, ou o desejo maior de controlar um outro desejo menos forte. Contudo, uma vez que o desejo de contrariar o desejo inicial se impõe, o sujeito deixa de ser livre para contrariar o desejo de contrariar.

Por essa razão, o neurocientista Anil Seth considera que o livre-arbítrio é, tal como a consciência, uma percepção alucinatória controlada. O livre-arbítrio existe, assim como a consciência, mas são causados pelo cérebro e são, em rigor, percepções (Seth, 2021). Da mesma forma que o vermelho é uma percepção de algo que não existe como a nossa percepção sugere, também o livre-arbítrio existe como uma percepção de algo que não é o que pensamos ser. Ou seja, o vermelho é uma percepção, da mesma forma que o livre-arbítrio é uma percepção. No entanto, existe uma realidade ontológica associada ao vermelho, da mesma forma que existe uma realidade ontológica associada ao livre-arbítrio. Tanto o vermelho como o livre-arbítrio – como realidade ontológica e como percepção – são causados, pelo que existem no mundo real. Apenas não são o que pensamos que são. Podemos, aliás, dizer que o vermelho e o livre-arbítrio, como são percepções, são conceitos epistemológicos para uma realidade ontológica mais complexa. Dito de outra forma, o vermelho e o livre-arbítrio existem no mundo real, mas existem como percepções (a primeira sensorial e a segunda emocional), já que a realidade ontológica inerente a essas percepções é distinta da percepção em si. Por isso Anil Seth considera que são alucinações controladas. São experiências fenomenológicas (Kant, 1781/2008).

Claro que os filósofos libertários objetariam estas ideias alegando que o sujeito apenas pode ser livre se nada causar a sua decisão ou volição, isto é, se o sujeito agir livremente sem qualquer condicionalismos e livre de quaisquer leis da Natureza. Esta é a perspetiva canónica de livre-arbítrio, mas deve ser repensada, pois trata-se de uma visão não só completamente incompatível com a Natureza causal do Universo, como também profundamente dissonante daquilo que sabemos sobre a realidade do Universo (Rovelli, 2018; Smolin, 2019; Hooft, 2016). É, portanto, um equívoco dizer que o ser humano só pode ser livre se violar as leis da Natureza.

Sugerimos que o livre-arbítrio pode existir desde que, tal como todas as coisas que existem num Universo como o nosso, tenha sido causado. E não há nada que invalide que o livre-arbítrio possa ser causado. Aliás, pensar que o livre-arbítrio pode existir sem ser causado é um absurdo filosófico. Como pode o livre-arbítrio existir num Universo físico, e por isso causal, sem ser causado? Neste sentido, sugerimos uma revisão sobre a ideia de livre-arbítrio filosófico, porque a definição que os filósofos libertários fazem do conceito é não só um beco sem saída como uma impossibilidade

lógica e científica – a ideia de que podem existir eventos, como o livre-arbítrio, sem causa num Universo que se rege por leis causais. Conforme argumentado, o determinismo é a única lei que pode causar eventos, então, o livre-arbítrio só pode ser uma possibilidade através das leis determinísticas.

2.3.3. *O sujeito como uma construção*

Um dos problemas centrais do libertarismo reside na indefinição do conceito de sujeito. Os filósofos libertários referem-se frequentemente ao sujeito como uma entidade consciente, por vezes imaterial, por vezes descrita em termos metafísicos vagos e imprecisos. Mas afinal, o que é o sujeito para o libertarismo? Em muitas formulações, o sujeito é simplesmente identificado com a pessoa — uma noção ambígua e pouco esclarecedora (como no libertarismo simples e no libertarismo causal) — ou então com uma consciência não-física (como no libertarismo de causação agencial), cuja origem e poder causal permanecem obscuros, sobretudo no que toca à sua interação com o universo físico.

Este problema conceptual exige maior precisão. Não basta afirmar que o sujeito é a pessoa. Uma pessoa é, antes de mais, uma construção complexa — biológica, relacional e semântica. Não é um organismo isolado, mas um sistema que depende de milhares de milhões de organismos simbióticos, moléculas, células e bactérias. Nesse sentido, a pessoa não é uma substância autónoma, mas um processo emergente, inteiramente causado por relações físicas, químicas e biológicas, como qualquer outro ser multicelular (Rovelli, 2018; Sapolsky, 2024). A pessoa é também uma convenção linguística, um símbolo útil para referir a individualidade e a continuidade biográfica de um determinado corpo com identidade, temperamento e personalidade.

Contudo, há uma distinção relevante entre a pessoa e o sujeito: o sujeito corresponde ao eu experiencial, ao centro fenomenológico que sente, pensa e decide; já a pessoa é a representação externa, social e biográfica desse sujeito, inscrita num corpo. Esta distinção é frequentemente ignorada pelos libertários, o que os impede de esclarecer quem ou o que está efetivamente a decidir. Quando afirmam que "o sujeito é livre", não nos dizem com clareza o que entendem por sujeito, nem explicam como essa entidade poderia ser a causa autónoma de uma decisão, independente do universo

de partículas, átomos, moléculas, células e estruturas biológicas de que depende para existir.

Neste contexto, o conceito de sujeito — tal como é formulado pelo libertarismo — padece de uma epistemologia que obscurece a sua própria ontologia. Ou seja, o modo como o sujeito é descrito impede que se compreenda o que ele realmente é e faz. Por isso, consideramos que conceitos como pessoa ou sujeito, tomados isoladamente e sem fundamentação físico-biológica, nada nos dizem sobre a verdadeira origem das decisões. Sem reconhecer o papel determinante da estrutura física e biológica que torna possível qualquer ato de escolha, estas categorias perdem valor explicativo e tornam-se metafísicas vazias.

Sem pretender aprofundar excessivamente esta questão, podemos afirmar que o sujeito deve ser compreendido como um processo emergente que resulta da organização funcional de um organismo vivo dotado de um sistema nervoso. A função central desse sistema é representar, em forma de imagens conscientes, tanto o meio externo como o próprio corpo ao qual pertence. Ou seja, uma pessoa é uma entidade consciente que foi causada. É causada, desde logo, pelas partículas que constituem o seu corpo (Rovelli, 2018), pelos genes presentes no interior de cada célula, e pelas redes de neurónios responsáveis pela sua atividade mental e fenomenológica (Damásio, 2010).

Nada no sujeito é *acausal* ou independente da cadeia causal das leis da Natureza. Toda e qualquer partícula, célula ou estrutura que compõe uma pessoa é o resultado de cadeias causais regidas pelas leis do Universo (Rovelli, 2018). A consciência não é exceção: é simplesmente mais um processo emergente. Tal como os ossos ou os órgãos são causados por interações físico-químicas, também a consciência é causada pela atividade neuronal do sistema nervoso.

Caso surja a tentação de duvidar de que a consciência — ou o eu — seja causada pelo cérebro, basta considerar a vasta evidência empírica sobre alterações da consciência provocadas por modificações químicas no sistema nervoso, como as que ocorrem com o consumo de substâncias psicoativas (Carhart-Harris & Friston, 2019; Muthukumaraswamy et al., 2013; Schmid & Liechti, 2017; Serra & Monteleone, 2006; Vosburg, Newman & Hess, 2021). Estas observações reforçam a ideia de que o sujeito consciente é inseparável das condições físicas que o tornam possível.

Neste sentido, quando os libertários afirmam que o sujeito causa uma decisão ou uma ação, ignoram um ponto essencial: esse sujeito é ele próprio um evento causado. Assim, qualquer decisão atribuída ao sujeito foi, inevitavelmente, causada pelas condições que originaram o próprio sujeito. O sujeito existe porque foi causado; o seu sistema nervoso existe porque foi causado; as suas células nervosas existem porque foram causadas; o eu que emerge existe porque foi causado. Tudo é causa, e nada escapa ao encadeamento físico que estrutura a realidade.

Em suma, consideramos fundamental, no âmbito das discussões sobre o livre-arbítrio, clarificar quem é o sujeito — o agente. Se o sujeito é uma construção física, química e biológica — e, portanto, causada —, então o eu que emerge dessa construção é, por sua vez, uma entidade igualmente causada. A agência, por conseguinte, não escapa à causalidade e é uma manifestação funcional de uma estrutura física que a torna possível.

2.4. O meio segundo de atraso da consciência

O neurocientista Benjamin Libet foi, desde a década de 60 do século XX, um dos primeiros autores a tentar responder objetivamente à questão: é a consciência – e, portanto, o eu – que inicia as decisões ou é o cérebro, não a consciência, que decide pelo eu? Se fosse a consciência a iniciar as ações humanas, isto poderia dar alguma força às teorias libertárias (1985 e 1999). Primeiro, poderia provar que quem decidia era o agente (o eu ou a consciência) e não o subconsciente; segundo, poderia provar que, mesmo que o cérebro estivesse envolvido na decisão (o que é unânime entre os cientistas), seria a parte do cérebro responsável pela consciência a tomar as decisões, pelo que não existiriam decisões desconhecidas do eu já que eram todas conscientes e, por isso, conhecidas. Libet, aliás, estava determinado a comprovar que o ser humano possuía livre-arbítrio ao demonstrar que seria na atividade cerebral *consciente* que as decisões e as ações seriam iniciadas.

A descoberta do potencial de prontidão (PP) – uma gravação elétrica neuronal que surge antes do sujeito tomar uma decisão consciente – permitiu avançar com algumas conclusões sobre este assunto (Deecke *et al.* 1976; Libet, 1985). Se o ser humano fosse realmente livre – isto é, se a ontologia da decisão fosse consciente – então

a experiência subjetiva deveria preceder ou, pelo menos, coincidir com os processos neuronais que conduzem à ação, nomeadamente com o PP (Libet, 1985). Dito de outra forma, se a consciência estivesse na natureza da decisão, então o PP deveria surgir ou depois ou ao mesmo tempo que a decisão consciente (Libet, 1999). Em contraste, caso a vontade consciente de agir não surgisse antes ou ao mesmo tempo que o PP, então, como refere José António Alves na sua obra *Limites da Consciência*, “a vontade do sujeito poderá não ser o verdadeiro motor das acções, sendo a liberdade uma ilusão” (2013, p. 80).

Numa das experiências mais famosas sobre o livre-arbítrio, Benjamin Libet (1985) demonstrou que antes do sujeito tomar a decisão consciente (ponderada e totalmente livre) de levantar o pulso, a área motora suplementar (no córtex frontal) iniciava atividade neuronal subconsciente (o PP), que estava na origem e *causava* a experiência consciente de decidir. Libet chegou a estes resultados da seguinte maneira. O neurocientista colocou várias pessoas à frente de um relógio e pediu-lhes que, quando quisessem, levantassem o pulso. Ao mesmo tempo, estas tinham que assinalar no relógio quando é que sentiram vontade de mover os pulsos, enquanto Libet observava a atividade elétrica dos seus cérebros. Ora, em *todas* as observações de Libet (e não apenas em algumas ou na maioria), o PP surgiu antes da vontade consciente de mover o pulso. Ou seja, em todas as dezenas de observações sobre se a consciência podia iniciar a decisão livre e incondicionada de levantar o pulso, entre 0,2 e 0,5 segundos *antes* do sujeito decidir conscientemente realizar o movimento, o córtex pré-motor, invariavelmente, iniciava a sua atividade, *causando* a consciência de vontade. Isto parecia significar que a consciência está atrasada em relação à origem da decisão em cerca de 0,2-0,5 segundos (Libet, 1985), pelo que estes resultados sugeriam que a ação consciente e voluntária teria de ser precedida em cerca de 0,2-0,5 segundos de atividade neuronal subconsciente⁵⁴.

Esta experiência, em que os voluntários reportavam, consistentemente, que só ficavam conscientes da intenção de agirem entre 0,2-0,5 segundos “depois da ativação muscular” (Libet, 1985, p. 532), conduziu à hipótese de o ser humano viver 0,2-0,5 segundos desfasado da realidade presente (Alves, 2013). Ou seja, estas experiências

54 Luder Deecke já tinha observado a mesma coisa alguns anos antes (Deecke *et al.*, 1976).

demonstraram que a consciência parece estar sempre atrasada em relação à realidade, já que a consciência apenas surgia 0,2-0,5 segundos *após* as coisas terem realmente acontecido (Desmurget, 2009; Ammon e Gandevia, 1990).

Em *The Ego Tunnel*, Thomas Metzinger refere, no mesmo sentido, que o que estamos a experienciar no momento é “na realidade, o passado” (2009, p. 38). Não há contacto direto e imediato com a realidade presente. A consciência é, aliás, uma *representação* da realidade, não a realidade em si. Não precisamos de aprofundar muito esta ideia (diria que praticamente unânime entre cientistas e filósofos) para perceber que o livre-arbítrio não pode existir para além de uma representação. Notem que não estou a afirmar que existe um consenso sobre o livre-arbítrio ser uma representação. Estou, isso sim, a advogar que existe um consenso sobre o cérebro gerar representações da realidade, e não gerar réplicas exatas da realidade, que pode ser encontrado tanto em autores de pendor fisicalista (Metzinger, 2009; Humphrey, 2023; Damásio, 2010; Dennett, 1991; Churchland, 1985) mas também em pensadores com perspetiva dualista (Chalmers, 1995, 1996; O’Connor, 2000; Nagel, 1974; Jackson, 1986; Kane, 1996; Balaguer, 2010). Ou seja, embora todos estes autores possam ter perspetivas muito distintas sobre o livre-arbítrio e sobre a consciência; apesar de alguns serem libertários e outros deterministas, todos eles advogam ostensivamente que a nossa perceção é uma construção da realidade. Ora, se não experienciamos a realidade como ela é, mas como se nos afigura, de que forma podemos ser livres? Se estamos emparedados em relação à realidade – uma realidade que nos foge permanentemente – de que modo podemos ter livre-arbítrio? De que modo os filósofos libertários podem fundamentar uma liberdade que está, logo à partida, condenada a agir de acordo com uma perceção não-realista da realidade?

Em muitas situações podemos verificar, de facto, na primeira pessoa, o atraso da consciência. Sempre que temos de reagir rapidamente ou quando nos assustamos, conseguimos perceber que a consciência se arrasta um pouco e apenas surge *após* a ação (Wegner, 2002; Metzinger, 2009). Nestas ações mais automáticas, a reação do organismo é automática e eficiente, e só depois a consciência surge – até porque, como será analisado adiante, o cérebro consciente precisa de integrar os estímulos sensoriais de uma forma concertada, de modo a que o sujeito possa ter uma experiência – uma representação da realidade – coerente e integrada, e não fragmentada (Eagleman, 2012;

Dehaene *et al.*, 2021). O neurocientista Marc Jeannerod, por exemplo, verificou que, para o sujeito ter consciência de uma ação, o cérebro precisa de integrar informação em diferentes áreas neuronais de modo a que as mesmas possam ser representadas conscientemente, num processo que demora algum tempo. Nas suas palavras: “Consciousness of a goal of an action is not immediate, it takes time to appear” (2006). É como se a consciência contasse uma história através de um conjunto de elementos, que precisam de ser reunidos num espaço de trabalho integrado cuja construção leva tempo.

As próprias emoções são reações automáticas, que são experienciadas *a posteriori* como sentimentos de emoção. Podemos reagir com medo na presença de um animal, com entusiasmo perante uma música que adoramos ou com alegria quando vimos alguém de que gostamos muito, mas estas emoções surgem invariavelmente *antes* de começarmos a ter consciência das mesmas (Damásio, 1993/2011; LeDoux, 1999). Como Damásio verificou, as emoções surgem de forma automática no organismo e expressam-se no corpo, mas só depois são sentidas na consciência em forma de sentimentos de emoção (Damásio, 2010).

2.4.1. As experiências sobre o livre-arbítrio

As experiências de Benjamin Libet (1985 e 1999) conduziram a duas conclusões que, entretanto, já foram replicadas por inúmeros estudos: 1) a intenção consciente de iniciar uma ação é causada por atividade subconsciente (que se manifesta no PP); 2) a intenção consciente parece surgir 0,2-0,5 segundos após o PP.

Na década de 1990 do século XX, Umberto Castiello chegou a conclusões muito semelhantes às de Libet. Castiello pediu a um grupo de voluntários para agarrarem um objeto situado à sua frente quando este ficasse iluminado. Castiello observou que o movimento da mão antecedeu em cerca de 0,5 segundos a consciência do início do movimento, não obstante o sujeito achar que a sua ação coincidiu com a perceção do objeto (Castiello *et al.*, 1991). Curiosamente, também em 1990, Klaus Ammon e Simon Gandevia demonstraram que poderiam influenciar as decisões dos sujeitos (levantar o braço esquerdo ou direito) ao estimularem as correspondentes áreas neuronais associadas ao PP (1990). Já no século XXI, em 2009, Michel Desmurget demonstrou o mesmo. Ao estimular o lobo parietal e o córtex pré-motor, conseguia provocar nos

pacientes não só a vontade de movimentar, como a ideia de que se tinham movimentado sem que o tivessem feito (2009).

Nos últimos anos, o neurocientista Patrick Haggard replicou as experiências de Libet, tendo chegado às mesmas conclusões (Haggard, 2005, 2008; Haggard e Eimer, 1999). Utilizando metodologias mais modernas e uma tecnologia mais avançada (nomeadamente ressonância magnética funcional (fMRI)), Haggard observou que a atividade neuronal do PP precedia em cerca de 0,5 segundos a vontade consciente de agir, bem como a ação em si (Haggard e Eimer, 1999). Mais concretamente, Haggard verificou, na linha daquilo que Wegner já tinha sugerido (2002) que a área motora suplementar possui um papel central, não só na ação, como também no *sentimento de volição* – poderíamos também dizer na experiência de escolha livre. A atividade nos lobos frontal e parietal forma um circuito que elabora e monitoriza planos de ação, o que, simultaneamente, produz a experiência consciente de intencionalidade a par da ação em si (Haggard, 2005, 2008). Por outras palavras, Haggard demonstrou que: “a experiência consciente é a consequência da atividade neuronal, e não a sua causa” (2005, p. 291).

O neurocientista Chun Siong Soon, na mesma senda de Libet e Haggard, demonstrou que as decisões podem ser causadas pelo cérebro até 10 segundos *antes* da experiência subjetiva. O que esteve na origem da experiência de Soon era saber se o PP, que é produzido pela área motora suplementar (e que fora identificado por Libet e Haggard), é de facto a origem da decisão ou se faz parte de um processo superior, mais associado à fase final da decisão e não ao seu início. Ou seja, Soon questionou se a atividade que podemos observar correspondente à PP não seria a origem da decisão mas a manifestação de um processo iniciado anteriormente, de origem mais profunda. Utilizando ressonância magnética funcional (fMRI), Soon pediu para que os participantes relaxassem e centrassem a sua atenção num ecrã onde iria surgir uma corrente de letras. A certa altura, foi-lhes pedido para decidirem, livremente, entre dois botões – um à direita e outro à esquerda. Em paralelo, eles deviam lembrar-se da letra que surgiu no ecrã quando tomaram a decisão de carregar no botão da direita ou da esquerda. Através das imagens da fMRI, Soon começou a observar padrões de atividade neuronal que precedia a atividade da área motora suplementar, tal como tinha sugerido. Para além disso, Soon conseguiu também separar algumas áreas do cérebro que claramente

contribuíam para a decisão, e de que forma o faziam. Como outras experiências já tinham demonstrado (Libet, 1999; Haggard, 2005, 2008), duas áreas neuronais ficavam ativas antes da decisão consciente: o córtex motor e a área motora suplementar. Contudo, Soon descobriu que outras duas áreas do cérebro, o córtex frontopolar e o precuneus (zona parietal), iniciavam atividade antes daquelas, e que essa mesma atividade permitia não só prever em 7 a 10 segundos se o participante iria escolher o botão da direita ou o botão da esquerda, como antecipar a experiência subjetiva do participante em 7 a 10 segundos. Nas palavras de Soon: “two specific regions in the frontal and parietal cortex of the human brain had considerable information that predicted the outcome of a motor decision the subject had not yet consciously made” (p. 545). Soon sugeriu, portanto, que o córtex frontopolar é o estágio cortical onde a decisão é feita, enquanto o precuneus é o local onde a decisão é armazenada e depois disponibilizada na consciência. A experiência de Soon parece demonstrar, por um lado, que quando um sujeito toma uma decisão consciente, essa decisão já pode ter sido fabricada cerca de 10 segundos antes, e por outro, que o PP que ocorre 0,5 segundos antes da decisão ser consciente é *causado* por eventos precedentes ocorridos em determinadas áreas neuronais que funcionam não só longe da consciência como ainda longe do PP. Dito de outra forma, a atividade da área motora suplementar, cuja ativação é verificada antes de qualquer ação consciente (Libet, 1999; Haggard, 2005, 2008), é um estágio tardio da decisão, já que é possível observar padrões de atividade no córtex frontopolar que *determinam* decisões até cerca de 10 segundos antes da decisão consciente (Soon *et al.* 2008). Algo trabalha nos bastidores na decisão, muito antes de ela se tornar consciente (Wegner, 2002; Alves, 2013; Eagleman, 2012).

As experiências de Libet, Haggard e Soon foram prosseguidas, entretanto, pelo neurocirurgião Itzhak Fried (2011). Este obteve os mesmos resultados através de um procedimento clínico capaz de gravar diretamente a atividade neuronal – o que é um método mais avançado do que o utilizado por Libet, que apenas gravava as ondas cerebrais. Tal como nas experiências de Libet e Haggard, Fried colocou os participantes à frente de um relógio (neste caso num monitor de um computador) rotativo. Depois, foi-lhes pedido para carregarem num botão quando sentissem a vontade de o fazer. Uma vez carregado o botão, os participantes deveriam indicar onde é que o ponteiro do relógio estava quando sentiram a vontade de carregarem no botão. Em cada ensaio,

Fried registou quando os participantes pressionaram a tecla e quando sentiram a vontade de carregarem na tecla. Enquanto faziam isto, Fried gravava a atividade neuronal na área motora suplementar, na área motora pré-suplementar e no córtex cingulado anterior. Os resultados foram em tudo semelhantes aos de Libet e de Haggard. Fried verificou que, cerca de 0,7 segundos *antes* do participante carregar na tecla, observa-se atividade neuronal nas áreas motoras suplementares, bem como no córtex cingulado. A atividade neuronal gravada por Fried (o PP) precedia a consciência da vontade, bem como a ação em si, permitindo, inclusive, antecipar a decisão do sujeito com algumas centenas de milissegundos. As descobertas de Fried demonstram, igualmente, que para alguma ação se tornar consciente, bem como para que o sujeito seja invadido por um sentimento de vontade de agir, o PP tem de atingir um determinado *limiar* (Fried *et al.*, 2011), o que é consonante com as conclusões de Libet (Libet, 1985, 1999; Alves, 2013). Nas palavras de Fried: “these findings support the view that the experience of will emerges as the culmination of premotor activity (probably in combination with networks in parietal cortex) starting several hundreds of ms before awareness” (Fried, 2011, p. 10). Ora, isto significa que os processos subconscientes observados ao nível da área motora suplementar precedem tanto a decisão como a ação, o que é convergente com numerosos estudos que demonstram que a área motora suplementar está implicada tanto na *representação* como na *preparação* do movimento motor (Amador e Fried, 2004; Ikeda *et al.*, 2002; Lau *et al.*, 2004).

O neurocientista Stefan Bode e a sua equipa replicaram a experiência de Soon, mas desta vez com maior enfoque no córtex frontopolar de modo a terem uma observação mais detalhada desta área. Os métodos foram os mesmos: relaxar e escolher livremente carregar no botão à esquerda ou à direita, enquanto surgiam letras que deveriam decorar na altura em que faziam a escolha. Os resultados foram os mesmos. Cerca de 7 segundos antes da decisão consciente, o córtex frontopolar iniciava atividade segundo um padrão que permitia ao observador adivinhar com uma antecendência substancial qual seria a decisão consciente do observador. Tal como fora observado por Soon, Bode verificou que o córtex frontopolar parece ser uma das primeiras estruturas neuronais a ficarem ativas na tomada de decisão, o que, subsequentemente, leva à ativação de outras áreas neuronais, como a área motora suplementar ou a área motora pré-suplementar, cujo início de atividade está mais próximo da decisão consciente.

Devido ao atraso temporal que a própria tecnologia de fMRI necessariamente impõe, Bode especula que a atividade subconsciente no córtex frontopolar pode ser iniciada mais cedo, provavelmente cerca de 10 segundos. Segundo estes resultados, a decisão consciente é, portanto, iniciada 10 segundos antes do sujeito ter consciência da mesma, o que significa que o pensamento consciente não participa na decisão. Curiosamente, Bode verificou também um padrão que associou à intensidade da atividade subconsciente com a entrada da decisão na consciência. Ou seja, quanto mais forte a atividade do córtex frontopolar, maior a probabilidade dessa atividade entrar na consciência. Isto sugere que, uma vez ultrapassado um determinado limiar, a atividade subconsciente torna-se consciente (Bode *et al.*, 2011), o que converge com as observações feitas por Libet, Haggard e Fried, entre outros.

2.4.2. *Objecções aos resultados libetianos*

Alguns autores têm vindo a colocar várias objeções aos métodos de Libet (Gallagher, 2006; Mele, 2009) e dos seus sucessores, apontando frequentemente que, por um lado, as decisões que tomamos são muito mais complexas e elaboradas do que levantar uma mão ou um pulso, e por outro, advogam que é necessário distinguir as ações motoras das ponderadas.

Em *Where's the action?* (2006), o filósofo Shun Gallagher, por exemplo, defende que uma ação motora automática deve ser distinguida de uma ação ponderada. A primeira é, segundo o filósofo, caracterizada por atos automáticos e não depende da consciência, ao passo que a segunda é fruto de uma intenção e desejo conscientes. Podemos assustar-nos ou realizar algum ato motor sem auxílio da consciência mas, quando pensamos nalguma decisão, apenas o conseguimos fazer com a consciência, pelo que este tipo de ação ponderada, que exige raciocínio, deve ser distinguida dos atos automáticos, que não necessitam da consciência. Gallagher postula que ao passear numa floresta, podemos assustar-nos com um animal que se parece uma cobra e iniciar um comportamento automático de fuga mas, logo de seguida, podemos verificar que o animal que observamos não é uma cobra mas sim um lagarto⁵⁵. Depois de alguma

⁵⁵ Quando tinha cerca de 14 anos, adorava perder-me em aventuras numa floresta que existia atrás da minha casa. Um dia, eu e alguns amigos estávamos embrenhados no meio da floresta quando percecionámos uma cobra e começámos a fugir. Só parámos de correr em casa. Tiro duas conclusões

ponderação, podemos até dirigir-nos em direção ao lagarto e tentar apanhá-lo para a nossa coleção, naquilo que podemos considerar um comportamento totalmente consciente. Embora apanhar o lagarto seja uma ação motora, a decisão de o fazer foi um ato refletido e consciente. Gallagher defende que é isto que distingue o livre-arbítrio: a diferença entre decisões motoras (como as que Libet e outros experimentaram) e decisões ponderadas como as de decidir apanhar ou não o lagarto (2006). Gallagher é secundado nestas críticas a Libet por diversos autores, entre os quais se destacam Mele (2009) e O'Connor (2000).

No entanto, parecem-me existir fragilidades evidentes nestas críticas. Desde logo porque nenhuma ação motora referida nos testes libetianos acima indicados fora automática e não refletida. Foram todas, aliás, completamente ponderadas e livres (o sujeito podia levantar ou não levantar o pulso, no caso das experiências de Libet, da mesma forma que o sujeito pode ou não apanhar o lagarto, no exemplo de Gallagher), o que enfraquece os argumentos de Gallagher quando este pretende fazer uma distinção entre dois tipos de ação. Como vimos, à luz do que as experiências de Libet e outros que lhe seguiram demonstraram, os voluntários tinham tomado previamente uma decisão refletida (quando levantar o pulso) que depois conduzia à ação motora, pelo que a distinção entre esses dois tipos de ações foi, de facto, replicada nas experiências de Libet e dos seus sucessores, não obstante as limitações inerentes a qualquer experiência científica. Tal como analisado, no momento em que emergia a decisão deliberada de levantar o pulso, observava-se uma atividade neuronal correspondente àquela intenção. Havia, portanto, uma decisão consciente que antecedia e desencadeava a ação motora propriamente dita.

Por outro lado, as decisões ponderadas e conscientes podem não ser nada mais do que decisões subconscientes que informam a consciência *a posteriori*, como os vários testes que analisámos sugerem. Ou seja, enquanto a decisão consciente de fugir de uma

desta simples história. Primeiro, o meu cérebro percecionou primeiro a cobra de um modo subconsciente (diria mesmo automático) e só depois é que ganhei consciência do que vi. Um pouco à semelhança dos sujeitos com o cérebro dividido, o meu hemisfério direito percecionou subconscientemente um estímulo e o meu hemisfério esquerdo percecionou-o conscientemente. Segundo, ainda hoje estou para saber se o que vi foi mesmo uma cobra ou se não foi uma narrativa criada pelo meu hemisfério esquerdo. Será que percecionei algo idêntico a uma cobra, o que levou o meu hemisfério direito a agir, e depois, de modo a ter uma história coerente – para dar uma razão lógica ao porquê de estar a fugir – o meu hemisfério esquerdo criou uma narrativa?

cobra (na ilustração de Gallagher) foi iniciada subconscientemente sem que o sujeito fosse de imediato informado – apenas alguns milésimos de segundo depois o sujeito percebeu conscientemente que estava a fugir de uma cobra – também a decisão consciente e refletida de tentar apanhar o lagarto pode ter sido iniciada subconscientemente sem que o sujeito tivesse qualquer controlo sobre o processo, como, de resto, demonstram os testes de Libet e de outros.

É comum, mesmo entre filósofos e cientistas, cair numa linguagem que parece sugerir um dualismo entre a consciência e o cérebro. Por exemplo, dizer que "a consciência regula" ou "intervém" no cérebro dá a impressão de que são coisas separadas. No entanto, um discurso mais rigoroso reconheceria que estas expressões apenas refletem diferentes níveis de descrição do mesmo fenómeno. O cérebro pode regular-se mesmo através dos processos a que chamamos consciência — como aprofundaremos no capítulo 3.

Este ponto é particularmente relevante quando analisamos a relação entre consciência e cérebro. Muitas vezes, é fácil incorrer em formulações que sugerem uma dicotomia entre ambos, como se a consciência fosse uma entidade separada do cérebro que “intervém” ou “decide”. Contudo, é fundamental compreender que a consciência é simplesmente uma função do cérebro, e não algo que exista em oposição ou separação deste. Quando se afirma que “a consciência causa uma decisão” ou que “a consciência age sobre o cérebro”, essa formulação é enganadora, pois ignora que as decisões ou atos do cérebro são os próprios atos da consciência. Ou seja, dizer que o cérebro causa X ou que a consciência causa X são apenas duas formas de descrever o mesmo processo.

Esta linguagem cria confusão, porque parece indicar a existência de duas entidades distintas em interação, quando, na realidade, a consciência é uma manifestação funcional do próprio cérebro (Metzinger, 2009). Assim, quando falamos de decisões conscientes, estamos apenas a referir-nos a manifestações de processos cerebrais subjacentes (Wegner, 2002; Jeannerod, 2006). Não há autonomia da consciência em relação ao cérebro, porque ambos são expressões do mesmo sistema físico e funcional. A consciência, neste sentido, é aquilo que o cérebro faz — e, por isso, não pode ser algo separado que o influencie ou que tenha agência própria sobre ele.

A vontade de apanhar o lagarto, por muito refletida e ponderada que tivesse sido, teve de ser causada por alguma coisa, sem que o sujeito tivesse qualquer

possibilidade de interferir no processo, pois ele não pode intervir nos seus pensamentos e na sua vontade. É, aliás, isto que as experiências referidas defendem: que todas as decisões são precedidas de causas neuronais que determinam a vontade (ou não) de realizar uma ação motora. O mesmo é dizer que o sujeito, ao querer capturar o lagarto, deixou de ser livre para *não querer apanhar o lagarto*. O desejo, a vontade, a decisão de apanhá-lo, ao ser causada pelo cérebro, já não permitia ao sujeito desejar o contrário. O sujeito tinha sido determinado pela *vontade causada pelo cérebro* para realizar uma ação. Caso o sujeito mudasse de ideias, isto é, se tivesse vontade de não apanhar o lagarto, é porque essa mudança teria sido causada pelo cérebro, o que tinha por conseguinte *causado* a vontade de não apanhar o lagarto. Dito de outra forma, a consciência foi informada por processos subconscientes que criaram a vontade de capturar o lagarto (ou de não o apanhar) e essa vontade cria a ilusão de liberdade. De facto, podemos questionar Gallagher acerca de onde estava a liberdade do sujeito de *não* apanhar o lagarto uma vez causado o desejo de o apanhar? Pode o sujeito controlar a sua vontade? Pode deixar de ter vontade de capturar o lagarto quando tem a vontade de o fazer? O filósofo dir-nos-ia provavelmente que a liberdade assenta na decisão de um *fazer ou não fazer* volitivo. Certo, mas essa decisão volitiva não é mais do que uma causa que *determina* o desejo de apanhar ou de não capturar o lagarto. Em grande medida, o sujeito, quando apanhou o lagarto, deixou de ser livre para *não* capturar o lagarto porque o seu cérebro já tinha *causado* o desejo de o apanhar. Isto é, o sujeito não teria capacidade para contrariar a vontade causada pelo seu cérebro. O sujeito sente a vontade de agir e, a seguir, a percepção de liberdade associada ação, como apanhar um lagarto, mas está totalmente cego em relação àquilo que causou, tanto a volição como a liberdade de realizar essa ação (Espinosa, 1677/1992, p. 272).

As várias experiências que aqui enumerámos demonstram que toda e qualquer decisão consciente tem de ser o resultado da maquinaria neuronal. A consciência apenas desvenda a atividade subconsciente (Dehaene *et al.*, 2021, p. 45). A vontade de apanhar o lagarto pode, de facto, ter sido causada pela consciência do sujeito; no entanto, tanto essa vontade como a própria consciência tiveram, necessariamente, de ser causadas por algo anterior. Nenhuma das duas surge do nada — são efeitos de processos físicos, neurobiológicos e contextuais que as tornam possíveis. Talvez a consciência não cause diretamente a vontade, mas permita tomar consciência dela — o

que é algo substancialmente diferente. Por outras palavras, a consciência, apesar de ser causada, torna conhecível as vontades ocultas do corpo (Damásio, 2010).

A consciência parece ser um género de palco onde as coisas se reúnem para montar um filme – o filme das nossas vidas. Quando começamos a compreender o que é a consciência e que, para ela existir, é necessário construir uma experiência integrada dos estímulos fragmentados que nos chegam, então, talvez possamos perceber que a sua capacidade para iniciar uma ação pode ser exígua ou nula. A consciência é um processo que cria uma experiência subjetiva da realidade, mas esquecemo-nos de que essa experiência é extremamente rica e complexa e, sobretudo por isso, é um processo que demora tempo a ser realizado (Metzinger, 2009). Talvez a consciência não exista para decidir ou para agir, mas para criar uma representação de identidade — um eu —, uma espécie de ilusão funcional que nos coloca no centro do mundo e, com isso, permite o surgimento de sentimentos de pessoalidade e, por extensão, de liberdade. Perfilhamos, em parte, esta perspetiva; consideramo-la, contudo, incompleta. No capítulo 3, proporemos que a consciência desempenha um papel fundamental na decisão e na ação — não no sentido defendido pelas teorias libertárias, mas à luz de uma abordagem determinista, na qual a consciência atua como decodificador informacional do corpo e do meio, permitindo ao cérebro, através da construção do eu, aceder a representações mais refinadas da realidade. Neste sentido, o eu não é uma entidade livre no sentido ontológico, mas um produto evolutivo do sistema nervoso, cuja função é aumentar a capacidade do organismo para integrar informação, mapear estados internos e externos e, com isso, tomar decisões mais eficazes. A consciência, assim entendida, oferece vantagens adaptativas claras: não confere liberdade metafísica, mas oferece ao organismo um modelo de si próprio, essencial para comportamentos complexos, planeamento e autorregulação.

2.5. A consciência como uma ilusão

A consciência parece ter, portanto, limites sérios no que concerne ao iniciar uma decisão e uma ação. Parece que sempre que decidimos alguma coisa, apenas estamos conscientes do resultado final de um processo complexo, que foge completamente ao

controle consciente (Eagleman, 2012; Tallis, 2017). Quando, por exemplo, decido dar aulas de filosofia, o que me surge na consciência é a vontade, o desejo e a ambição de dar aulas de filosofia. Mas não sei como decidi. Não sei de onde surgiu essa vontade. Não tenho qualquer controle sobre o como e o quando é que essa vontade foi causada; não tenho qualquer controle sobre a vontade que tenho, ou não tenho, de dar aulas de filosofia. O como e o porquê fogem-me por completo. Tenho consciência da decisão de dar aulas de filosofia, mas não sobre *como* nem *por que* é que essa decisão surgiu ou foi causada. Sei, no entanto, que não foi o meu eu que *causou* a vontade de dar aulas de filosofia. Pelo contrário, foi a vontade de dar aulas de filosofia, que *causou* a decisão no meu eu mas, como disse, alguma coisa que não eu causou a vontade. Eu não causei vontade. A vontade aconteceu-me. No mesmo sentido, quando tomo a decisão de comer um gelado ou de mudar de casa, não tenho a mínima consciência do que esteve na origem dessas decisões, não sei o que causou essas decisões. E posso comer um gelado e mudar de casa pensando que fui eu que o decidi. E, em certa medida, fui eu que decidi, não obstante não saber como nem porque o fiz.

Como foi referido, não nos parece possível considerar que a vontade possa ser *acausada*. A vontade acontece-nos. Posso ter consciência das ideias, dos pensamentos e das crenças com que decido; posso ter conhecimento dos desejos e das vontades que habitam a minha mente e que (supostamente?!) me ajudaram a decidir; posso ter conhecimento das razões contextuais que me levaram a decidir. Posso, claro, ter conhecimento de que tenho vontade de alguma coisa. No entanto, o *como* e o *porquê* fogem-me da compreensão. Para saber como e porquê é que tomei uma decisão teria que saber todos os meandros neuronais que conduziram a que a vontade surgisse na minha consciência. E não o sei. Não o posso saber, a não ser que soubesse causar vontade e, de facto, não sei. Nunca causei nenhuma vontade em mim. Como disse, a vontade é algo que me acontece. Sou iludido pela vontade causada pelo meu cérebro sobre um poder de escolha consciente que, efetivamente, não tenho nem possuo forma de ter. Talvez por isso algumas escolas de filosofia postulam que somos os causadores das decisões (Ginet, 2008, 2014; O'Connor, 2000; Chisholm, 2003; McCann, 1998; Balaguer, 2014) – porque nos são vedadas as verdadeiras causas das decisões, o que provoca em nós uma ilusão de controle.

O mesmo se passa com a ação. Quando pensamos numa ação a realizar – como levantar-nos do sofá para ir buscar um copo de água à cozinha – primeiro surge a vontade de o fazer, ao que depois sucede a subsequente ação motora. Ora, não só não sabemos *como* nem *porque* decidimos (*como* e *por que*⁵⁶ surgiu a vontade) como também não estamos, de todo, conscientes de todos os processos inerentes a essa ação motora: apenas estamos conscientes do objetivo a alcançar – o copo de água. Ir buscar um copo de água é, porém, uma ação hercúlea para o nosso organismo, que passa por alguns dos seguintes processos neurofisiológicos: ativação do sistema músculo-esquelético, adaptação do sistema vestibular, ativação do sistema propriocetivo, ativação do sistema nervoso simpático (aumento da frequência cardíaca, aumento da circulação sanguínea, aumento da frequência do sistema respiratório, aumento dos níveis de glicose...) etc., etc.. Ora, qual o nosso controlo em qualquer um destes processos? Nenhum. Nem poderíamos ter, já que se necessitássemos de ter consciência de tudo o que se passa no nosso organismo quando temos de executar uma ação motora tão simples como ir buscar um copo de água, paralisaríamos (Eagleman, 2012; Gazzaniga, 2011). Ao invés, basta-nos ter o desejo (melhor dizendo, basta que o sentimento de desejo seja *causado*) de ir buscar um copo de água para que todo o corpo ative todos os sistemas necessários para essa ação e adote a correspondente ação motora. Sem esforço consciente, iniciamos e conduzimos toda a sequência de ações que nos levam a sair do sofá e ir à cozinha buscar um copo de água.

Contudo, não deixamos de ter a perceção que estamos a controlar todos os aspetos da ação, quando na realidade o nosso organismo está num enorme esforço fisiológico, propriocetivo e homeostático para que a ação seja bem-sucedida. Nem temos qualquer noção do esforço do nosso organismo em realizar a ação, pois estamos completamente vedados a toda essa maquinaria. Estamos iludidos ao pensar que a ação foi conduzida por nós e, mais do que isso, que a ação é fácil. Ora, nem a ação pode ser conduzida pela nossa consciência, dada a sua complexidade intrínseca, nem é facilmente realizada pelo nosso organismo, considerando o esforço e a coordenação que exige. Será, então, a consciência uma ilusão? Não uma ilusão no sentido em que não

56 Aqui o *como* e o *porquê* não se refere às razões fisiológicas, como a necessidade de beber água porque temos sede. Refere-se ao *como* e *porquê* no sentido de não sabermos, nem *como* nem *porque* alguma coisa causa a necessidade de beber água.

existe – parece-nos evidente que a consciência existe desde o momento em que acordamos (Searle, 2019) – mas uma ilusão no sentido em que a percepção consciente de que agimos livremente não é uma prova direta de que o pensamento consciente tenha sido a causa da ação? Será a consciência uma ilusão no sentido de não ser aquilo que nos parece?

Podemos sentir que somos nós – conscientemente – que causamos as nossas ações, mas isso não significa que tal aconteça na realidade (Gazzaniga, 2013; Libet, 1985 e 1999; Haggard, 2005 e 2008). De facto, a consciência parece ser um fenómeno completamente diferente da ação (Wegner, 2002; Jeannerod, 2006), e por isso mesmo pensamos que deve ter propriedades distintas. A consciência sou *eu*, ao passo que a ação é algo que *eu fiz* – que, de certa maneira, ordenei ao meu corpo fazer. É por isso difícil de imaginar que estas duas coisas possam emergir da mesma coisa: do nosso cérebro.

No entanto, todas as evidências parecem indicar que tanto o *sentimento* de liberdade ancorado na vontade e no desejo como as nossas ações são *causados* pelo nosso cérebro (Wegner, 2002; Blackmore, 2005; Haggard, 2005; Eagleman, 2012; Gazzaniga, 2011; Soon *et al.*, 2013). Se não fossem causadas pelo nosso cérebro, teriam de ser causadas por outra estrutura física e, neste sentido parece-nos que o cérebro é o candidato mais óbvio. A alternativa seria, como temos vindo a referir, que os nossos desejos fossem *acausados*, o que nos parece ilógico e impossível.

2.5.1. A consciência causada

Ora, se a vontade não pode ser *acausada*, sugerimos que alguma coisa causa a nossa vontade (ou desejo). Dito de outra forma, alguma coisa causa a vontade e vários estudos têm vindo a demonstrar, de forma consistente, que a dopamina (Carter e Hall, 2018) ou a noradrenalina (Berridge e Waterhouse, 2003), entre outros neurotransmissores e moléculas, desempenham papéis fundamentais na motivação e na vontade (Serrado, 2024 e 2025). A dopamina é frequentemente associada ao sistema de saliência, regulando a motivação e atribuindo valor subjetivo aos estímulos, isto é, determinando se algo é suficientemente relevante para ser perseguido. Estudos indicam que níveis elevados de dopamina estão correlacionados com o aumento da motivação e da busca por objetivos, enquanto a sua deficiência pode levar a apatia e perda de

interesse (Carter & Hall, 2018). Por outro lado, a noradrenalina está intimamente ligada ao estado de alerta e à saliência dos estímulos, regulando a atenção e a prontidão para a ação. Sendo libertada principalmente pelo *locus coeruleus*, a noradrenalina modula a resposta ao ambiente, ajudando a priorizar quais estímulos devem ser processados e orientando o comportamento face a desafios e recompensas potenciais (Berridge & Waterhouse, 2003).

Deste modo, podemos afirmar que a vontade não surge de forma espontânea ou indeterminada, mas é o resultado de processos neurobiológicos concretos, em que dopamina e noradrenalina atuam como moduladores essenciais da motivação, da atenção e da ação – processos intimamente associados com a vontade –, assegurando que os organismos ajustam os seus comportamentos de forma eficiente às exigências do meio. Para controlarmos a nossa vontade, teríamos que conseguir aceder ao nosso cérebro e manipular os nossos sistemas dopaminérgicos (o que acontece apenas muito parcialmente, com a necessidade de consumo de estimulantes, como a cafeína). Se, aliás, conseguíssemos controlar a nossa vontade, teríamos o problema dramático da dependência resolvido, já que poderíamos tratar os nossos comportamentos compulsivos com alguma facilidade e todos nós sabemos que isso não é verdade.

A vontade é-nos imposta, portanto, em grande parte, pelos nossos sistemas dopaminérgicos e noradrenérgicos, mas, como não sabemos como surge (não temos acesso consciente ao seu processo causal subjacente), acreditamos que fomos nós que a causamos. Como refere Jeannerod no seu artigo *Consciousness of Action as an Embodied Consciousness*: “While we tend to perceive ourselves as causal but we actually ignore the cause from which our actions originate” (2006). O papel de neurotransmissores como a dopamina ou noradrenalina não só determina o que consideramos relevante e motivador, como também modula a perseverança e intensidade do impulso para agir. Assim, sentimo-nos livres porque experienciamos a vontade como nossa, mesmo que o seu surgimento tenha sido impulsionado por processos neurobiológicos sobre os quais não temos qualquer controlo.

O sentimento de vontade está intrinsecamente ligado à volição e, subsequentemente, ao livre-arbítrio, pois só podemos considerar-nos verdadeiramente livres se tivermos a capacidade de agir de acordo com os nossos desejos ou mesmo contra eles. Assim, a mesma questão que colocámos sobre a vontade pode ser aplicada

à volição e ao livre-arbítrio⁵⁷: sou eu que causo o sentimento de volição em mim? Posso fazer nascer em mim a sensação que sou livre para realizar uma ação e ir contra um desejo? É a volição causada por alguma coisa ou é *acausada*? Da mesma forma que a vontade é causada pelo nosso cérebro (designadamente pelos sistemas dopaminérgicos e noradrenérgicos), também nos parece que não existe outra solução para o livre-arbítrio que não seja afirmar que o sentimento de volição é *causado* por alguma coisa, e julgamos não estar longe da verdade se disser que é causado pelo funcionamento neuronal.

Tal como a vontade (estado psicológico), a volição (ato consciente de decidir) surge-nos na consciência sem que tenhamos qualquer controlo sobre isso. Não sabemos como surge nem tão-pouco porque surge. Por isso, não temos volição e para sentir ou deixar de sentir volição, já que esse sentimento surge em nós espontaneamente, sem que tenhamos qualquer acesso ao processo. Como refere Michael Gazzaniga na sua obra *Who's In Charge* : “There is no one boss in the brain...[and] You are certainly not the boss of the brain” (2011, s.p.). Assim, se eu não sou o causador da minha volição – se não sei causar a liberdade em mim – como posso afirmar que sou livre? Ou seja, se não tenho qualquer controlo sobre os meus sentimentos, onde fica o meu livre-arbítrio? Sugerimos, portanto, que a volição é, tal como a vontade, um *sentimento* que nos é imposto pelo nosso cérebro. Em grande medida, o nosso livre-arbítrio é causado pelo nosso cérebro – talvez seja a única forma de o termos (Serrado, 2024, 2025).

2.5.2. O livre-arbítrio como uma percepção

Como vimos, a percepção consciente das alterações neurofisiológicas determinadas, em parte, pela dopamina e pela noradrenalina causa uma percepção de vontade. Ou seja, a vontade é uma percepção consciente que surge depois de determinados sistemas alterarem a paisagem neuronal e somática, pelo que a vontade é, essencialmente, uma representação consciente do estado do corpo. Esta percepção consciente é aquilo que designamos por *qualia* (Damásio, 2010; Humphrey, 2022; Seth,

⁵⁷ Embora os conceitos de volição e livre-arbítrio não sejam sinónimos — a volição diz respeito ao acto de querer ou decidir, enquanto o livre-arbítrio se refere à possibilidade de escolha livre entre alternativas — , ambos estão intimamente relacionados. Por essa razão, ao longo deste trabalho, poderão ser utilizados de forma próxima ou até intercambiável, sempre que tal não comprometer o rigor da análise.

2021). A fenomenologia é, neste sentido, essencialmente, uma representação do estado neurofisiológico do corpo. Segundo esta perspectiva, a consciência pode ser uma ilusão como qualquer outra percepção, como a percepção da cor, do som ou a percepção de qualquer outra coisa que é ontologicamente distinta da realidade em si. Da mesma forma que as nossas percepções são sombras da realidade – não a realidade em si – também a consciência pode ser uma percepção de uma realidade muito diferente.

A psicóloga Susan Blackmore, na sua obra *Consciousness: An Introduction* (2005), postula que a consciência é uma *percepção* e, como qualquer percepção, está sujeita a todo o tipo de enganos e de ilusões. Do mesmo modo que a nossa percepção visual ou sonora pode perceber algo que, na realidade, não existe da forma como a percebemos (o que Kant defendeu no século XVIII (1781/2008)), a consciência pode ser uma percepção ilusória de nós próprios. Segundo este ponto de vista, o sujeito existe, bem como a sua experiência subjetiva, contudo, tanto o sujeito como a experiência subjetiva correspondem, na realidade, a algo muito diferente daquilo que a nossa percepção consciente experiencia. Utilizando outra analogia, do mesmo modo que as cores, objetivamente, não existem como as percebemos – são construções do cérebro – também a consciência não existe como a experienciamos. Tal como as cores, a consciência é uma construção – diríamos, uma representação mental do *self* (Blackmore, 2005).

O filósofo Thomas Metzinger tem uma opinião muito semelhante à da psicóloga Susan Blackmore. Para o autor de *The Ego Tunnel* (2009), o eu não existe – é apenas uma construção neuronal, um truque orquestrado pelo cérebro para criar um organismo detentor de coerência e unidade. Segundo Metzinger, a consciência é um modelo de *self* que replica fenomenicamente o cérebro, criando um dono virtual que não é mais do que um modelo – uma ilusão. A consciência, neste sentido, não tem possibilidade de ativar músculos ou neurónios mas, como é um modelo virtual que replica o cérebro do ponto de vista fenoménico (é uma representação do organismo), acaba por criar um ego. Este ego não tem poder causal sobre nada porque é um eu virtual que nasce do organismo e que, em grande medida, replica o organismo mas, como não sabe que é um modelo virtual, vive na ilusão de que tem uma vida independente do organismo (Metzinger, 2009). O filósofo utiliza o conceito de *phenomenal self-model* para designar uma entidade no cérebro que produz informação e que oferece uma imagem interna – uma

representação – de um *self*. O eu é o que Metzinger designa como *Ego Tunnel* – a experiência subjetiva mas extremamente estreita de um mundo físico muito mais rico, mais complexo e mais difícil de representar do que aquilo que a nossa consciência é capaz. Em grande medida, o eu é uma imagem – um modelo – que o cérebro cria mas que nós, como *somos* essa imagem, não somos capazes de ver essa imagem enquanto imagem. Vemo-la como eu – um ego fenomenal. Como Kant propôs em *Crítica da Razão Pura* (1781/2008), o fenómeno é a aparência da realidade, não a realidade em si. Ora, se o eu é fenomenal, isto significa que o eu é uma aparência da realidade, neste caso, da realidade neuronal. Segundo esta perspetiva, nunca nenhum ser vivo teve ou tem um *self*, já que isto é uma ilusão – é um fenómeno – criada pelo cérebro (Metzinger, 2009).

No mesmo sentido, o neurocientista Anil Seth, no seu livro *A Criação do Eu* (2021), considera que a consciência é uma percepção, que pode ser comparada com outro tipo de percepção sensorial, como a radiação eletromagnética (em forma de cores) ou as ondas sonoras (em forma de sons). O mundo físico é constituído por uma realidade ontologicamente incognoscível. Por isso, resta ao sistema nervoso criar representações dessa realidade, o que explica porque a radiação eletromagnética e as ondas sonoras podem ser representações de coisas que, como tal, não existem na realidade. Isto é, existe radiação eletromagnética e existem ondas sonoras, no entanto, não existe vermelho nem o dó-ré-mi. O vermelho e o dó-ré-mi são representações simbólicas da realidade e não a realidade em si – uma espécie de alucinação controlada, nas palavras de Seth (2021).

Segundo o neurocientista, da mesma forma que o sistema nervoso representa a realidade exterior, também representa a realidade interior, em forma de *self*, que não é mais, também, do que uma alucinação controlada, já que não existe um eu no cérebro. Ou seja, da mesma forma que não existe vermelho nem dó-ré-mi, não só no mundo exterior como no cérebro, também não existe um *self* no cérebro, muito menos no exterior. No mesmo sentido, os sentimentos são representações “virtuais” do corpo, que ajudam na construção do *self*. Numa palavra, o *self* é uma *imagem integrada*. O sistema nervoso representa o corpo, através de sinais interoceptivos, e cria uma representação integrada que designamos por *self* – ou consciência. Poderíamos pensar que o *self* é que possibilita a percepção mas, como refere Seth em *A Criação do Eu*: “O

self é outra percepção, outra alucinação controlada...” (2021, p. 145). A consciência é uma ilusão, não no sentido de que não existe, mas no sentido de ser uma *representação* do eu – uma alucinação controlada que emerge do cérebro mas que, ao mesmo tempo, é o cérebro (Seth, 2021).

António Damásio concorda com a ideia que o *self* é uma percepção. Porém, o neurocientista considera que essa percepção, para ser consciente, tem de ser sentida. Neste sentido, o eu é uma *percepção sentida do organismo* (2010 e 2013). É o sentimento de si que confere experiência subjetiva ao sujeito. A cartografia neuronal do corpo permite transformar sinais fisiológicos em sentimentos que traduzem o estado do corpo em forma de dor, fome ou sede, mas também em forma de alegria, tristeza ou nojo, entre outros sentimentos. Ou seja, os sentimentos são representações mentais de estados fisiológicos do organismo. Esta capacidade de um organismo sentir é o que, em parte, permite o surgimento de um eu.

2.5.3. O livre-arbítrio como um sentimento

A capacidade do sistema nervoso para decodificar sinais químicos e transformá-los em representações sentidas (Damásio, 2010) é, no nosso entender, aquilo que permite a existência de consciência. Por sua vez, a percepção consciente da vontade, do desejo ou da volição é o que possibilita o sentimento de *agency* (ou de livre-arbítrio). Segundo esta perspetiva, e como o psicólogo Daniel Wegner sugeriu no início do século XXI na sua obra de referência *The Illusion of Conscious Will* (2002), a volição nada mais é do que um sentimento gerado pelo cérebro para criar uma narrativa funcional no âmbito da identidade do sujeito. Ao desejar fazer alguma coisa, o sujeito — ao sentir esse sentimento emergir — sente-se livre para decidir e agir. Postulamos, portanto, na mesma linha de Wegner, que sem a experiência do sentimento de volição, nenhum sujeito poderia alguma vez sentir-se livre. É necessário que qualquer ação ou decisão venha acompanhada desse sentimento — o sentimento de volição, de *agency*.

Para demonstrar que a volição é um sentimento gerado pelo cérebro, Wegner dá o exemplo do síndrome da mão alienígena. Neste síndrome, frequentemente associado a danos no lobo frontal do lado oposto ao da mão afetada, os pacientes experienciam uma das mãos a agir automaticamente, sem qualquer controlo. Eles não experienciam a vontade de agir ou de mexer uma das mãos, pelo que consideram que

algo – que não eles – está a fazer com que a mão se mexa. No entanto, a mão afetada não faz, necessariamente, movimentos aleatórios. Em muitos casos, um sujeito pode estar a fazer a barba com a mão direita enquanto a mão esquerda (involuntariamente) desabotoa a camisa, ou pode estar a querer virar a página de um livro enquanto a outra mão tenta fechá-lo. Embora os atos da mão afetada não sejam livres nem desejados, são coordenados e sofisticados (Wegner, 2002). Ainda assim, se não são desejados (o sujeito sente que é algo que lhe está a acontecer), dificilmente estes comportamentos podem ser considerados livres, independentemente do elevado grau de complexidade da ação. De facto, se uma mão desempenha ações e faz movimentos complexos e precisos sem qualquer vontade consciente do sujeito, de que modo podemos considerar que a consciência inicia e conduz ações?

Wegner dá, também, o exemplo da força de *vontade* de *não* comer um bolo de chocolate por motivos de saúde. Parece que somos nós que, conscientemente, estamos a decidir não comer o bolo através de um ato de autocontrolo. No entanto, não só este autocontrolo é *causado* pelo cérebro sob forma de *vontade* de não comer o bolo, como o sujeito apenas tem experiência do *sentimento de vontade* de não comer o bolo – não tem qualquer experiência sobre como esse processo foi conduzido. Dito de outra forma, *a vontade é um sentimento consciente que reflete uma intensidade neuronal que conduz à vontade de executar um determinado comportamento*. Esta intensidade – que poderíamos chamar o nível mais forte ou mais fraco de desejo (ou motivação) – pode ser forte ou fraca, e é o que distingue a persistência de um sujeito em, por exemplo, fazer exercício físico todos os dias, da de um outro sujeito que cede todos os dias à tentação de comer largas quantidades de açúcar (2002, p. 11). Existem, aliás, estudos que demonstram que a intensidade com que a dopamina é libertada no cérebro, de uma forma basal, está relacionada com os níveis de motivação do sujeito (Salamone e Correa, 2012; Treadway *et al.*, 2012).

Por conseguinte, a vontade e, subsequentemente, a volição são, segundo esta perspetiva, sentimentos. Nas palavras de Wegner:

“a experiência de vontade é um mero sentimento que ocorre à pessoa. Está para a ação como a experiência de dor está para as modificações que resultam da estimulação da dor, ou

como a experiência de emoção está para as modificações corporais associadas a essa emoção” (2002, p. 13).

Podemos relacionar a analogia do sentimento de vontade, ou de volição, com o sentimento de dor ou de medo. O sentimento de dor não é igual aos tendões lesionados num pé, nem tão pouco o sentimento de medo é igual aos padrões neuronais que ocorrem no cérebro quando sentimos medo, embora, em casos normais, uma coisa provoque a outra. Dito de outra forma, da mesma forma que uma determinada neurofisiologia causa o sentimento de dor ou de medo, sugerimos que uma determinada neurofisiologia causa o sentimento de vontade. Essa experiência de vontade é, por conseguinte, essencial para o sentimento de volição – para o sentimento de que somos donos das nossas decisões.

Neste sentido, alguns autores consideram que temos dois eus: um eu incorporado – que inicia e desenvolve ações – e um eu fenomenológico, que cria narrativas para essas ações (Wegner, 2002; Gazzaniga, 2011). Este último teria a função de criar uma continuidade que oferecesse coerência e unidade ao sujeito e às suas ações que são, “por necessidade, executadas automaticamente” (Jeannerod, 2006). A relação de aparente causalidade entre os nossos pensamentos e as nossas ações oferece-nos o sentimento de volição - ou *agency* (Choudhury e Blackmore, 2006). No entanto, o sentimento de *agency* é, precisamente, isso mesmo: um sentimento. E, como qualquer outro sentimento, pode ser interpretado e compreendido à luz de estados fenomenológicos semelhantes, como a dor ou o medo. Tal como estes, o sentimento de volição é uma experiência subjetiva causada pela atividade do sistema nervoso. Não resulta de um princípio metafísico autónomo, mas de mecanismos neurobiológicos que produzem a percepção de controlo sobre as próprias ações.

Segundo esta perspetiva, a consciência decifra o comportamento – dá-lhe um sentido – mas não o inicia nem o conduz, pelo que a sua função pode ser a de dar coerência à ação e integrá-la num sujeito unificado, com uma identidade – um eu (Oakley e Halligan, 2017), o que analisaremos mais à frente, no capítulo 3. Ou seja, é possível que a decisão, a vontade e a volição sejam causadas pelo cérebro que inicia as ações, que são depois experienciadas fenomenologicamente de modo a que possa emergir um eu estruturado. Talvez um eu estruturado necessite de um sentimento de *agency*, como se este sentimento fosse uma das âncoras estruturais desse eu. Podemos

imaginar que o eu é facilmente uma estrutura mais coesa se sentir determinados sentimentos de uma forma coerente, nomeadamente o sentimento de *volição*. Sem o sentimento de *agency*, talvez o eu fosse uma entidade menos funcional, já que, sem *volição*, provavelmente não poderiam existir outros sentimentos ou processos cognitivos tão centrais no âmbito do comportamento humano, como o desejo, o sonho e a esperança. Se não sentíssemos que os sentimentos eram nossos e livres, não só viveríamos uma realidade muito mais fragmentada, como seria impossível ter outros sentimentos tão importantes para funcionarmos como agentes responsáveis, flexíveis e evolutivos.

Em suma, a vontade e a *volição* são inerentes ao livre-arbítrio – ao sentimento de *agency*. Sem o sentimento de vontade de fazer (ou de não fazer) alguma coisa, ninguém pode ser livre. Contudo, a vontade e a *volição* são sentimentos causados pelo nosso cérebro, pelo que, em rigor, como não temos qualquer acesso ao nosso cérebro nem temos a mínima noção de como é possível criar sentimentos de vontade e desejo, não podemos ter *agency* para além da experiência subjetiva de sentir. Como vimos, o nosso cérebro decide e, ao mesmo tempo, gera sentimentos de vontade e de desejo, e só depois tomamos conhecimento dessa decisão – em forma de experiências fenomenológicas causadas por sistemas neuronais, semelhantes ao que se passa com a representação das cores ou com os sentimentos de dor ou de medo.

2.5.4. *A ilusão da ação consciente*

Qualquer ação motora que fazemos, desde que não estejamos a ser coagidos, parece-nos ser completamente livre. As ações motoras são mais fáceis de estudar, já que podem ser avaliadas e medidas, ao contrário das decisões intelectuais que são privadas e mais difíceis de medir.

Conforme o exemplo que utilizámos anteriormente, quando o sujeito inicia uma ação motora (como levantar-se para ir buscar o copo de água), o sistema visual-motor seleciona o objetivo e, em paralelo, todo o cérebro e o corpo ativam um conjunto de dispositivos e sistemas que iniciam comandos motores com vista à concretização desse objetivo (Jeannerod, 2004; Gallagher, 2006). Será que a consciência participa neste processo? Será que o eu inicia a ação? Disto de outra forma, será que a consciência – o eu – é apenas uma representação tardia desses processos automáticos, ou terá, de

facto, algum papel funcional na sua execução. Como temos vindo a referir, para que o eu iniciasse uma ação era necessário: 1) que nada causasse o eu; 2) que o eu causasse a vontade; 3) que o eu movesse o corpo. E já verificámos que nada disto é possível. Tanto o eu, como a vontade, como a *agency*, como ainda os movimentos motores são causados pelo cérebro, sem que o sujeito tenha qualquer controlo mínimo sobre algum destes processos. Por nenhum momento temos acesso aos mecanismos inerentes à sequência extraordinariamente complexa de movimentos, de ajustes mecânicos, de adaptações fisiológicas que são necessários para atingir o simples objetivo de bebermos um copo com água.

Como já foi referido, a consciência é um processo que demora tempo a surgir. São necessários, pelo menos, entre 0,2 a 0,5 segundos para que a consciência integre todos os sinais e estímulos do meio de forma integrada, e consiga fazer uma representação coerente (Libet, 1999; Haggard, 2005, 2008). Segundo Jeannerod, isto explica porque a consciência de um *objetivo a alcançar* tem um tempo diferente da *ação desenvolvida* para alcançar esse objetivo em si. A consciência, devido ao tempo que necessita para integrar informação motora e sensorial, propriocetiva e interoceptiva, não pode ser útil para a ação. Mesmo em ações que podemos achar que estamos a executar conscientemente, a consciência não parece ter qualquer papel no âmbito da sua execução (Jeannerod, 2006; Wegner, 2002; Gazzaniga, 2011; Bombarda, 1902; Banks, 2006; Haggard, 2005, 2008; Bode *et al.*, 2011, 2014; Soon *et. al.*, 2008; Amador e Fried, 2004; Fried *et al.*, 2011). A ação envolve inúmeras sequências de movimentos (pequenas ações) que escapam completamente ao controlo da consciência de tão rápidas que são e precisam de ser, bem como envolve inúmeras adaptações fisiológicas automáticas que necessitam de, também, ser extremamente eficientes para que o corpo se adapte às exigências da ação e do meio ambiente. No mesmo sentido, a consciência não tem, também, como interferir nos processos neuronais que ativam músculos e fibras simplesmente porque, conscientemente, não sabemos como ativar um músculo. De facto, como referiu o psicólogo William Banks em *Does Consciousness cause Misbehavior?*, não fazemos ideia de como ativar músculos devido à sua impenetrabilidade (2006). Podemos saber que temos vontade de ir beber água, mas não temos o mínimo controlo nem de como surge essa vontade, nem do que fazer para pôr o corpo em movimento (Jeannerod, 2006).

Apesar de podermos pensar que as ações conscientes funcionam de forma diferente das ações subconscientes, a verdade é que não há nada que demonstre que assim seja. Poderíamos pensar que a consciência não intervém em ações que precisam de ser rápidas, como fugir de um poste que vai cair em cima de nós, mas que intervém em decisões conscientes, como levantar-me para ir comer um pão com queijo (Gallagher, 2005, 2006). No entanto, não há nenhum indício disso. Segundo o neurocientista Marc Jeannerod, os sistemas motores para agir consciente ou inconscientemente são os mesmos. A única coisa que muda são as circunstâncias e, sobretudo, *se somos invadidos pelo sentimento de vontade*. Em ações automáticas não somos invadidos por esse sentimento, já que o mesmo demora um tempo considerável, o que não é vantajoso em situações que nos obriguem a reagir rapidamente. Pelo contrário, em ações ponderadas, em que temos tempo para pensar, somos invadidos por sentimentos de vontade, muito provavelmente porque isso é útil para a realização de ações a longo prazo. Porém, como referido, os sistemas que nos levam a decidir consciente ou subconscientemente são os mesmos, pelo que não há qualquer razão para crer que a consciência participa em algumas ações (conscientes) e não noutras (subconscientes) (Jeannerod, 2006).

Pensemos em dois exemplos. Estou a conduzir o meu automóvel a 60 km/hora, a caminho de casa, e a cerca de cinco metros perceciono um buraco mesmo à frente do meu carro. Apenas tenho tempo de me desviar “instintivamente”, sem pensar. Primeiro ajo e depois ganho consciência do que fiz. O que fiz foi uma decisão automática, pelo que *não tive vontade de fazer nada*. Fiz o que fiz porque *reagi* a uma situação, sem ter tempo de pensar e de, sequer, me sentir livre para fazer o que quer que fosse, pelo que a consciência – o meu eu - não pôde ter interferido na decisão.

Mas se estiver a conduzir no mesmo trajeto à mesma velocidade e vir um buraco a cerca de 30 metros do meu carro, então, sinto a *vontade de conscientemente* me desviar do mesmo. Parece que, primeiro, pensei conscientemente o que fazer, tive vontade de o fazer e, depois, agi ao desviar-me do buraco.

Contudo, na realidade, antes de me desviar do buraco segundo a minha vontade, o meu cérebro teve de percecionar o buraco e de ativar, de um modo totalmente subconsciente, sistemas cognitivos e motores capazes de dar resposta à situação. Ao mesmo tempo (ou melhor, um pouco depois), o meu cérebro teve, igualmente, de

causar o sentimento de vontade de me desviar do buraco, pelo que quando realizei a ação, fi-lo com a vontade de a fazer, ao contrário do primeiro exemplo quando tive de me desviar sem ter tido (tempo de ter) a vontade. *O sentimento de vontade de me desviar do buraco, ao contrário da primeira situação, levou-me a crer que fui livre para realizar a ação*, pelo que a consciência (aparentemente) contribuiu para a decisão.

Mas o que distingue, afinal, a primeira da segunda situação? Na primeira, reagi instintivamente, sem deliberação consciente; na segunda, tomei uma decisão refletida, com base num desejo. Será que, na primeira, o meu cérebro decidiu por mim, enquanto na segunda fui eu — o eu consciente — quem tomou a decisão, simplesmente porque quis agir? Se, na primeira, não tive vontade — apenas reagi —, poder-se-á dizer que não fui livre; mas se, na segunda, houve desejo, então fui livre? Esta distinção sugere que associamos liberdade não apenas à ausência de constrangimento, mas à presença de uma experiência consciente de intenção. Mas será essa experiência suficiente para fundamentar a liberdade, ou é apenas a percepção posterior de um processo já iniciado pelo cérebro?

Consideramos que em ambas as situações, foram os sistemas subconscientes que produziram, no primeiro caso a ação (sem a vontade) e no segundo caso, a vontade, a decisão e a ação. Isto é, na primeira situação, o cérebro causou uma ação sem vontade, já que a vontade poderia atrapalhar a eficiência de uma decisão que se pretendia rápida e eficiente; ao passo que na segunda, o cérebro, tendo mais tempo para dar uma resposta adequada e mais flexível, pôde causar a vontade, a decisão e a ação. Em rigor, contudo, ambas as decisões e ações foram, totalmente, conduzidas por sistemas *automáticos*, de modo *automático*, sem que a consciência participasse na decisão. A consciência apenas contribuiu ao *conhecer* a vontade e a decisão que iria ser tomada. A *mudança de contexto* no segundo caso, bem como a *experiência subjetiva* da vontade, criou a ilusão de que a primeira decisão foi inconsciente e a segunda foi consciente, quando não há nada que indique que assim foi a não ser a minha percepção — o meu sentimento de vontade (e, subsequentemente, de *agency*) associado ao segundo caso. O facto de a vontade ter surgido na segunda situação *não* significa, portanto, que tenha sido livre, porque tanto a vontade como a decisão e a ação foram *causadas* pelo cérebro, sem que a consciência — o eu — tivesse qualquer acesso a essa maquinaria que funciona

de forma automática (Libet, 1999; Haggard, 2005, 2008; Soon *et al.*, 2008; Bode *et al.*, 2011, 2014; Wegner, 2002; Jeannerod, 2006).

A forma como a consciência se relaciona com as ações, descrita acima, parece sugerir que as *ações* são processadas em sistemas neuronais diferentes dos sistemas que processam as *imagens mentais conscientes* (Jeannerod, 2006; Wegner, 2002; Choudhury e Blackmore, 2006). Parece que uma coisa é a ação e uma outra é a forma como a consciência representa essa ação (Gazzaniga, 2011), ou seja, parece que a consciência é um fenómeno *post hoc* que não intervém na ação em nenhum momento (Banks, 2006). Isto pode ser explicado porque, primeiro, a consciência é um processo lento que necessita de tempo para integrar vários sinais diferentes numa representação integrada (Jeannerod, 2006) e porque, em segundo lugar, parece existir autonomamente da ação em si (Wegner, 2002). Como refere Jeannerod em *Consciousness of Action as an Embodied Consciousness*: “The embodied self mostly carries an implicit mode of action consciousness, where consciousness becomes manifest only when required by the situation” (2006). A ação necessita de um sistema *moment-to-moment* que monitorize e desenvolva comportamentos motores sequenciais eficientes e complexos e a consciência, devido à natureza fenoménica que temos vindo a discutir (cuja função parece ser representacional), parece ser um processo mais lento que a ação.

A consciência parece ser constituída por pensamentos e ideias com curta duração, já que aparecem e desaparecem com extrema frequência e facilidade (Hume, 1739/2016), o que reforça a ideia de que as ações não podem estar dependentes da nossa consciência. Se dependessem da consciência, esta teria que providenciar um mecanismo *moment-to-moment*, atado ao desenvolvimento preciso de cada um dos movimentos, das adaptações músculo-esqueléticas e das necessidades fisiológicas, o que claramente não acontece.

Fazemos recorrentemente um conjunto de ações sem qualquer esforço consciente. Se a consciência pudesse iniciar e conduzir ações, teria também que controlar todos os complexos e sofisticados mecanismos motores que permitem a ação, o que não deixaria qualquer espaço para a consciência poder habitar pensamentos e ideias, e fazer o trabalho que iremos sugerir no terceiro capítulo que é decodificar

informação. Se a consciência estivesse diretamente envolvida na execução das decisões e ações imediatas, estaria essencialmente ocupada por sistemas motores e operacionais, limitando-se ao domínio da ação. No entanto, a consciência existe para representar — para possibilitar a experiência —, e essa função é incompatível com o controle direto das ações. Se a consciência tivesse de comandar a ação e simultaneamente representá-la, como poderia manter uma integração coerente entre esforço motor e percepção consciente? Como poderia, ao mesmo tempo, executar e observar, agir e experienciar? A própria experiência em primeira pessoa mostra-nos que as nossas ações decorrem muitas vezes de forma independente dos conteúdos conscientes que nos acompanham. A nossa consciência está habitada por uma multiplicidade de pensamentos que, frequentemente, não interferem com a ação em curso (Bargh, 2017). Isso sugere que a consciência não é o motor da ação, mas sim a sua testemunha interpretativa.

Vejamos. Neste preciso momento, ao redigir estas palavras e expressar estas ideias, não faço a mínima ideia de como os meus dedos navegam pelo teclado, como é que conseguem ter a precisão para encaixar no A, no B ou no C do meu computador e, numa sequência rápida, conseguir fazer uma frase sem o mínimo esforço consciente; bem como, da mesma forma, embora saiba genericamente o que quero escrever, não tenho as frases ordenadas na minha consciência. Elas surgem-me – acontecem-me – sem que tenha o mínimo controle do como e do porquê (estas frases e estas ideias e não outras) e sem que tenha de me preocupar muito em encontrar as palavras e ordená-las. “Escolho” as palavras, não porque as selecione, mas porque o meu cérebro mas oferece – aquelas e apenas aquelas, não necessariamente as que eu gostaria que fossem para que o texto pudesse ficar mais rico ou elegante. Da mesma forma, não só não sei o que vou dizer a seguir como também não sei o que quero dizer até o fazer. A consciência é um túnel demasiado estreito (Metzinger, 2009) para nele poder caber tudo aquilo que quero dizer, tudo aquilo que acabo por dizer e tudo aquilo que gostaria de ter dito e não disse. Como referem David A. Oakley e Peter W. Halligan no seu artigo *Chasing the Rainbow*, a consciência não é o causador mas “antes o veículo através do qual os produtos subconscientes são apresentados” (2017, p. 7). É como se a consciência apenas manifestasse – ou desvendasse – as forças internas implícitas do cérebro e do corpo.

Podemos, também, com alguma facilidade, verificar que uma ação (qualquer que ela seja) é uma sequência de movimentos com uma automaticidade inerente, ao passo que raramente temos as imagens dessa mesma ação. Temos, isso sim, imagens do que se pretende adquirir com essa ação, bem como imagens da postura e do movimento do nosso corpo a executar essa ação. Quando estou a escrever ao computador, não me surgem imagens motoras dos meus dedos a escrever nem tão pouco dos movimentos dos meus pulsos e braços no âmbito desse processo. Existe uma maquinaria a funcionar à minha revelia sem que, no entanto, tenha consciência dessa maquinaria. Surge-me, isso sim, uma vaga ideia do objetivo – algumas ideias – que quero partilhar. É certo que são os sistemas propriocetivos que, em grande parte, controlam o processo (Cole, 1995; Gallagher, 2005). Tenho que situar o meu corpo e, subsequentemente, as minhas mãos e dedos para conseguir escrever de forma eficiente e coordenada. Mas esses sistemas propriocetivos fogem-me por completo ao controlo consciente. Nem sequer tenho imagens fenomenológicas nem percetivas sobre a proprioção, pelo que não tenho pensamentos sobre como funciona o meu corpo. Apenas sei que funciona.

Posso imaginar o meu corpo a correr, a andar e até a escrever. Posso imaginar a minha postura e a minha posição no espaço, mas não tenho nenhum acesso à maquinaria que permite tudo isso. De facto, se me perguntarem como é que consigo escrever ao computador não sei bem o que dizer. Direi que “apenas consigo”, não sei bem como nem porquê. Talvez pudéssemos dizer que apenas pensamos nisso mas, em rigor, em nenhum momento pensamos em como realizar nenhuma ação – apenas pensamos no objetivo que queremos alcançar com essa ação. No mesmo sentido, se me perguntarem como é que consigo falar, ou mesmo correr, responderei também que não sei como consigo. Apenas falo. Apenas corro. Todos os processos neuronais e fisiológicos inerentes à fala e ao correr são automáticos, sem que tenha algum controlo sobre os mesmos.

Imaginemos ainda uma ação como subir umas escadas. Podemos pensar no corpo a subir as escadas, na postura, na velocidade e na posição corporal que queremos adotar, mas não pensamos – nem podemos pensar – na coordenação muscular necessária para que um perna seque a outra; na velocidade e na precisão com que um pé deve ser colocado à frente do outro; no equilíbrio propriocetivo que é necessário para realizar essa tarefa; na pressão que temos de exercer na nossa anca e na nossa

perna esquerda para podermos levantar a direita; na sensibilidade propriocetiva necessária para situarmos o nosso corpo no espaço de modo a que o mesmo tenha uma postura adequada aos movimentos; entre muitos outros movimentos necessários para exercer uma ação tão simples como subir umas escadas (Gallagher, 2005; Cole, 1995). Subir umas escadas exige um esforço e uma coordenação extraordinariamente minuciosos de todo o nosso organismo, mas nada disso é fenoménico. Nada disto nos chega à consciência. Apenas temos a experiência de o fazer – não do como nem do porquê. Aliás, quando estamos a subir umas escadas, a nossa consciência é habitada, frequentemente, com imensos pensamentos e ideias que nada têm que ver com as ações que estamos a realizar. Isto parece sugerir que, se não penso no que estou a fazer e, ainda assim, o estou a fazer, a consciência não parece ser necessária para a ação.

De notar, no entanto, que, tanto quanto se sabe hoje, as ações motoras são geradas por neurónios (e músculos, fibras, tendões, nervos, etc.), mas as imagens (pensamentos, ideias, sensações) são, igualmente, geradas por neurónios - neurónios computacionalmente distintos mas, ainda assim, neurónios. Isto significa que os neurónios podem produzir tanto ações como imagens, pelo que tanto as ações como as imagens estão reféns de processamento neuronal e podem ter uma relação (Jeannerod, 2006).

Tendo em conta a origem das ações (em neurónios) e a origem dos pensamentos conscientes (em neurónios) e sabendo que as primeiras surgem e são conduzidas de uma forma automática, sem controlo consciente, porque os pensamentos conscientes teriam uma natureza diferente? Isto é, se as ações são conduzidas de forma automática, e estão dependentes de sistemas extremamente complexos aos quais não temos qualquer acesso, parece-nos extremamente provável que os pensamentos possam ter um funcionamento semelhante. Aliás, isto explicaria porque subimos as escadas e estamos a pensar na avó Maria ou no avô Quim; porque estamos a conduzir e estamos a pensar no que os nossos colegas nos disseram hoje; porque estamos a cozinhar e estamos a antecipar o dia de amanhã. Parece que a consciência – a corrente de pensamento – tem uma vida muito própria, independente da vida motora, o que não significa que não funcionem de forma integrada. O que queremos enunciar é que qualquer ação pode ser realizada sem pensamentos associados a essa ação; bem como qualquer pensamento pode ser pensado sem que uma ação daí resulte. No entanto,

tanto a consciência como as ações são suportadas por neurónios, pelo que ambos os processos estão dependentes do funcionamento neuronal subconsciente e podem ter uma função integrada, que é o que acontece frequentemente quando fazemos, ou não fazemos, o que desejamos (Jeannerod, 2006).

Em suma, a consciência não parece ser o motor das ações, mas sim um fenómeno que surge a posteriori, representando a experiência subjetiva de decisões e movimentos já iniciados por mecanismos neuronais automáticos. As ações motoras — mesmo aquelas que percebemos como voluntárias — resultam de processos subconscientes que escapam ao controlo direto da consciência. Se assim não fosse, seria impossível realizar uma ação enquanto pensamos noutra coisa completamente distinta. E, se assim não fosse, a consciência não poderia cumprir aquilo que, aparentemente, faz melhor: dar cor, sentido e unidade às experiências que vivemos.

2.5.5. *O intérprete e a narrativa da consciência*

Como temos vindo a referir, sentimo-nos livres porque efetuamos uma ação ou tomamos uma decisão quando as mesmas são desejadas, isto é, quando as ações e as decisões vêm acompanhadas por um sentimento de vontade que, subsequentemente, confere o sentimento de *agency*. Para além do síndrome da mão fantasma, já analisado (Wegner, 2000), outras condições neurológica ou psiquiátricas manifestam problemas no sentimento de *agency*. Quando isso acontece, as pessoas deixam de se sentir livres e o resultado é dramático. Pessoas com esquizofrenia, por exemplo, reportam frequentemente que os seus pensamentos, as suas ações e os seus sentimentos são produzidos por outras pessoas e não segundo a sua *vontade*. É como se estas pessoas tivessem falta de *agency*. Reportam, frequentemente, que *não* têm vontade de realizar certas ações e que, precisamente por isso, sentem-se manipuladas (Choudhury e Blackmore, 2006). Ao sentirem que os seus pensamentos e as suas ações não foram causadas pela sua vontade, estas pessoas têm muita dificuldade em sentirem que são livres.

O sentimento de *agency* — isto é, a sensação que as ações são feitas segundo a nossa vontade e que de alguma forma são iniciadas por nós de uma forma livre — é crucial, não só para a nossa liberdade mas, também, para a nossa sensação de identidade. Como já aludimos, se não nos sentimos livres para decidir e se agimos em

função de fatores que sentimos não controlar, então, não só a nossa liberdade como a nossa identidade pode ficar afetada com isso, como a esquizofrenia nos mostra de uma forma dramática. A identidade é, como temos vindo a referir, central para a nossa experiência fenomenológica, pelo que é de extrema importância que os nossos pensamentos e ações sejam acompanhados por sentimentos de *agency*. Só assim a nossa identidade – a nossa noção de eu – pode ser coesa e experienciada de uma forma livre (Oakley e Halligan, 2017; Morsella, 2022).

O conceito de identidade está, neste contexto, associado ao de subjetividade – a capacidade de um sujeito ter experiências de acordo com um eu estruturado, não fragmentado, como os que vemos em certos transtornos⁵⁸. A identidade é, em grande medida, um modelo interno do *self*, pelo que a sua ligação à fenomenologia é evidente. Só criaturas com identidade podem ter experiências conscientes – podem ter um eu sofisticado e complexo. Sugerimos que, no âmbito do que António Damásio defende, a identidade – tal como a vontade - é um sentimento (Damásio, 2010), uma representação do cérebro (Metzinger, 2009; Seth, 2021). Segundo Damásio, essa identidade está ancorada num sentimento primordial, que pulsa com as primeiras experiências de estar vivo. Sem sentimentos, nenhum organismo pode alguma vez ser consciente já que, como referimos no capítulo 1, são os sentimentos que permitem a um determinado organismo ter os primeiros sinais, *sentidos*, de que está vivo (Damásio, 2023). Isto não significa que todos os animais que têm emoções são conscientes porque, como vimos, os sentimentos (a capacidade de sentir emoções conscientemente) estão limitados a animais com um determinado sistema nervoso que seja capaz de transformar emoções (programas de ação do corpo) em sentimentos (em perceções mentais).

Neste sentido, a identidade estaria ancorada num sentimento – sentimento de existência - da mesma forma que a vontade, o desejo e a *agency* são sentimentos. Talvez por isso dizemos com maior frequência que “eu sinto que existo” ou que “eu sinto que sou livre” do que “eu sei que existo” ou que “eu sei que sou livre”. Talvez saibamos que existimos e que somos livres apenas porque o *sentimos*. Talvez a *função* da consciência

⁵⁸ Pessoas com transtornos de personalidade narcisista, *borderline*, entre outros, possuem um eu fragmentado, o que está na origem dos seus comportamentos disfuncionais (Kernberg, 1996).

fenomenológica (que Chalmers (1995 e 1996) esvazia de *função*) possa ser criar uma representação sentida de uma identidade – um sentimento de identidade que permite o conhecimento de estar vivo. Da mesma forma que a fenomenologia é ubíqua para a percepção da realidade (Humphrey, 2012 e 2023), como analisado no capítulo 1, talvez também seja central para a percepção do eu (Damásio, 2010).

O neurocientista Michael Gazzaniga estudou durante décadas pacientes epiléticos que submeteram-se a uma cirurgia que lhes separava os hemisférios cerebrais, de modo a aliviar os sintomas e os problemas da doença. Depois da cirurgia, os pacientes relatavam sentirem-se bem, já que as suas crises epiléticas tinham diminuído drasticamente ou cessado completamente. Vários estudos foram, aliás, feitos sem que se identificasse qualquer problema cognitivo nestas pessoas. Michael Gazzaniga, suspeitando que estes sujeitos pudessem ter duas experiências autónomas devido à separação hemisférica, desenhou e realizou várias experiências inovadoras e reveladoras no que concerne, tanto à lateralidade cerebral, como à forma como a consciência pode funcionar.

Numa dessas experiências, Gazzaniga apresentou aos pacientes imagens de objetos, tanto ao hemisfério esquerdo como ao direito. Importa referir que o hemisfério direito controla o lado esquerdo do corpo e o hemisfério esquerdo controla o lado direito e que, estando os hemisférios separados, era possível mostrar uma imagem exclusivamente a um dos hemisférios, se o paciente tapasse um dos olhos. Ora, quando era apresentada uma imagem ao hemisfério esquerdo (uma fotografia de uma cadeira, por exemplo), o sujeito conseguia identificá-la e descrevê-la. Facilmente dizia “cadeira”. Mas, curiosamente, quando a mesma imagem era apresentada ao hemisfério direito, o sujeito reportava que nada tinha visto. Se fossem mostradas outras imagens ao mesmo hemisfério, o paciente continuava a dizer que não via nada, ao contrário do que se passava com o hemisfério esquerdo, com o qual o paciente conseguia identificar todos os objetos que lhe eram apresentados (Gazzaniga, 2011).

No que concerne às imagens mostradas ao hemisfério direito, o paciente tinha o olho esquerdo aberto, completamente alinhado com a imagem ou com o objeto (ex. cadeira) mas, ainda assim, o mesmo reportava que não via nada. No entanto, o surpreendente era que, se fosse pedido ao sujeito para desenhar o que tinha acabado de ver ele desenhava, de uma forma totalmente automática e subconsciente, uma

cadeira (ou outro objeto que tinha sido mostrado). No mesmo sentido, se tivesse acabado de ver uma imagem de uma galinha, o paciente reportava que não vira nada mas, de novo, se lhe pedissem para desenhar o que achava que tinha visto ou o que estava a pensar, ele desenhava uma galinha. Ou seja, o paciente não tinha a percepção fenomenológica que tinha visto a imagem de uma cadeira e de uma galinha, mas tinha tido uma percepção subconsciente que tinha visto a imagem de uma cadeira e de uma galinha, que depois era desenhada sem que o paciente soubesse o que estava a desenhar. Era como se os hemisférios tivessem um nível de autonomia substancial.

Não menos surpreendente é o facto de, quando Gazzaniga perguntava ao sujeito porque tinha desenhado uma cadeira ou uma galinha (ou outro objeto observado no âmbito do estudo) quando este dissera que não tinha visto o que desenhara, o paciente criar uma narrativa para explicar o que tinha acabado de fazer – uma história que nada tinha a ver com a realidade. Podia dizer que desenhara uma cadeira porque estava cansado e precisava de se sentar, ou uma galinha porque lhe apetecia comer frango. Deste modo, sempre que o paciente desenhava alguma imagem que o seu hemisfério direito não tinha percecionado, e lhe perguntavam porque o tinha feito, o paciente criava uma história para *explicar* algo que ele próprio não sabia porque tinha feito. Simplesmente desenhava o que estava a pensar, de uma forma automática e, *a posteriori*, criava uma narrativa coerente para explicar o que fizera⁵⁹ (Gazzaniga, 2011).

No seguimento das primeiras experiências de Gazzaniga, este conduziu outras nas quais o sujeito era exposto a uma experiência emocional no hemisfério direito (era-lhe mostrado uma fotografia com uma imagem aversiva), o que frequentemente conduzia a reações de medo e de nojo. Depois da reação, era pedido ao sujeito que explicasse as razões do seu comportamento emocional. Mais uma vez, o hemisfério esquerdo, que *não* tinha sido exposto à experiência emocional, criava uma história coerente para explicar o que o hemisfério direito tinha sentido (medo ou nojo), sem

59 O neurocientista José Delgado já tinha realizado nos anos 60 do século XX algumas cirurgias nas quais observara o mesmo que Gazzaniga. Delgado, em algumas cirurgias, teve oportunidade de estimular algumas áreas neuronais que provocavam nos seus pacientes movimentos voluntários. Quando, por exemplo, um paciente levantava um braço e Delgado lhe perguntava porque o tinha feito, o primeiro podia dizer que lhe tinha apetecido, subitamente, apontar para alguma coisa, criando depois uma narrativa para a justificar (Delgado, 1969) o que está totalmente de acordo com o que ficou documentado por Gazzaniga.

que, contudo, essa narrativa tivesse qualquer contacto com a realidade. Existia, mais uma vez, uma necessidade do hemisfério esquerdo de justificar o que tinha sido feito ou sentido pelo hemisfério direito.

A conclusão destes estudos foi que o hemisfério esquerdo criava narrativas para justificar aquilo que o hemisfério direito sentia. Ou seja, começou-se a perceber que o hemisfério esquerdo estava sobretudo vocacionado para explicar os acontecimentos que observava, de uma forma semântica, sequenciada e analítica, sobretudo através da linguagem. O hemisfério direito, pelo contrário, tinha dificuldades na linguagem simbólica mas parecia estar mais vocacionado para a experiência emocional e para o pensamento holístico e abstrato. Pessoas com os dois hemisférios ligados funcionam de forma integrada, pelo que podem perceber objetos com os dois hemisférios. Quando pessoas têm os dois hemisférios separados, a experiência pode ficar fragmentada, com um dos hemisférios a criar experiência semântica e o outro uma experiência emocional (Gazzaniga, 2011)⁶⁰.

Estas experiências e relatos parecem demonstrar, igualmente, que os hemisférios nestes sujeitos funcionavam de forma autónoma. Gazzaniga demonstrou (ou reforçou), também, que o cérebro possuía uma especialização computacional. O hemisfério esquerdo estava mais vocacionado para a escrita, linguagem e interpretação cognitiva da realidade (o que Broca e Wernicke já tinham demonstrado no século XIX no âmbito da fala), ao passo que o direito tinha funções ligadas à experiência emocional e à inteligência espacial⁶¹. Precisamente por isso, Gazzaniga chamou ao hemisfério esquerdo *intérprete*, já que este tinha uma tendência natural de explicar e dar coerência aos comportamentos e sentimentos do hemisfério direito (nomeadamente os desenhos que fazia sem que o hemisfério esquerdo tivesse conhecimento)⁶² (Gazzaniga, 2011).

60 José Delgado verificou o mesmo em 1969. Quando Delgado estimulava diretamente determinadas áreas neuronais, o paciente realizava uma ação e, depois, criava uma narrativa para justificar esse comportamento.

61 A lateralidade dos hemisférios tem vindo a ser, sistematicamente demonstrada, com vastos estudos a demonstrarem que o hemisfério direito processa sobretudo emoções e o esquerdo está mais vocacionado para a experiência semântica e percetiva (Gazzaniga, 2000; Li *et al.*, 2019). No entanto, esta lateralidade não é rígida, já que o hemisfério esquerdo também processa emoções.

62 António Damásio, embora considere que a consciência tem origens subcorticais, pensa que Gazzaniga identificou aquilo a que o primeiro chama consciência autobiográfica – a consciência da nossa identidade.

Estas experiências pioneiras e reveladoras permitiram aprofundar qual o papel de certos sistemas neuronais em gerar a subjetividade. Era possível que o cérebro possuísse um sistema cuja função fosse desenvolver uma narrativa que pudesse consolidar uma identidade – um eu. Como refere Steven Pinker (in Blackmore, 2005, p. 146), “não há razão para duvidar que a nossa mente, nomeadamente o nosso hemisfério esquerdo, funcione de forma diferente dos pacientes observados por Gazzaniga”. Ou seja, não há razão para acreditar que o nosso hemisfério esquerdo não crie as mesmas histórias, justificações e narrativas para explicar os comportamentos que são iniciados automaticamente pelo resto do cérebro. Por isso Pinker diz que a consciência é um: “spin doctor, not the commander in chief” (in Blackmore, 2005, 146)

O neurocientista Patrick Haggard, no seu artigo *Human Volition: Towards a Neuroscience of Will* (2008), demonstrou, tal como Gazzaniga (2011), que as pessoas tendem a associar as suas ações voluntárias com uma determinada consequência causal, elaborando narrativas para explicar eventos em que não têm qualquer interferência. Numa experiência conduzida por Haggard, vários sujeitos premiam um botão voluntariamente e ouviam um *bip* a seguir, passados alguns milissegundos. Quando lhes perguntavam quando é que tinham a experiência de carregar o botão e ouvir o *bip*, os sujeitos respondiam que tinham essa experiência *antes* do *bip* surgir, como que se antecipassem o *bip* de modo a correspondê-lo como uma consequência causal da sua ação voluntária. Se, pelo contrário, os sujeitos ouvissem o *bip* primeiro e só depois carregassem no botão, os mesmos tinham tendência para anteciparem o *bip* de modo a que o mesmo estabelecesse uma relação causal com o carregar no botão. Isto parece demonstrar que percebemos as nossas intenções e as nossas ações como se tivessem uma relação causal, mas a observação e a experiência demonstram que essa causalidade não existe.

O sujeito necessita de dar coerência – de explicar e justificar – as suas decisões de modo a sentir-se livre. Segundo Haggard, o que acontece é que o cérebro constrói uma lacuna temporal, na qual junta os dois acontecimentos, mesmo quando um não causa o outro. A consciência permite essa experiência integrada, de modo a que aquilo que fazemos e observamos possa ter coerência e contribuir para os sentimentos de *agency* e, subsequentemente, de identidade. E é isto que também, em grande parte, oferece a sensação de volição, o sentimento de que aquilo que estamos a pensar e a

fazer está a acontecer ao mesmo tempo quando, objetivamente, não há nenhuma evidência de que assim seja (Haggard, 2002). Estas observações parecem demonstrar que a corrente da consciência não é nada mais do que uma percepção subjetiva de processos subconscientes, uma narrativa coerente que faz da vida “much like watching a movie” (Bode, 2014, 639).

Não raras vezes, todos nós já fomos sujeitos à ilusão de controlo. Daniel Wegner conta-nos, no seu livro *The Illusion of Conscious Will* (2002), uma história pessoal em que, certa vez, num centro comercial começou a jogar um jogo de máquinas – estava a controlar totalmente o jogo – para se aperceber minutos depois que não estava de facto a “jogar”. Estava, isso sim, a achar que jogava enquanto o computador passava uma *demo* (Wegner, 2002, p. 8). No mesmo sentido, já todos nós tivemos a sensação de termos determinado ou influenciado um desfecho de um jogo desportivo, seja por termos ido ao frigorífico buscar um copo com água na precisa altura em que a nossa equipa marcou um golo, ou porque num dia vitorioso nos encontrávamos no estádio. Muitas pessoas têm rituais e repetem-nos com frequência sempre que a situação o exige, precisamente por acharem que essas ações determinam o desfecho daquilo que querem que aconteça.

O ser humano possui crenças, desejos e intenções. Todos nós estamos familiarizados com estes conceitos e usamo-los a toda a hora. No entanto, estes conceitos apenas são úteis para compreender as nossas experiências, porque dizem-nos muito pouco sobre como funciona a realidade. As ideias – as crenças e modelos internos – podem ser, como iremos sugerir mais à frente, representações da realidade com funções na sobrevivência dos organismos que as possuem. Mas são *representações* da realidade, não a realidade em si (Kant, 1781/2008). Os ponteiros de um relógio, as ondas do oceano, a rota gravitacional dos planetas e o movimento de *quarks* ou de fotões podem ser compreendidos em termos de causa e efeito, mesmo que essas causas e esses efeitos possam fugir ao nosso entendimento.

Contudo, sentimos no nosso íntimo que faz sentido atribuir intenções a seres com consciência. É decerto compreensível, tendo em conta a nossa natureza pensante e consciente, os nossos sentimentos de vontade, de desejo e a nossa profunda convicção de que somos livres de agir sem constrangimentos físicos ou biológicos. No entanto, isso não significa que o ser humano possa ser explicado à parte de um Universo que funciona

de acordo com leis. Segundo Wegner, é precisamente aqui que reside o problema. Na ideia de que no Universo existem explicações radicais, uma para as mentes e outra para tudo o resto, como se fossem entidades diferentes ou com propriedades distintas. É como se o mundo físico pudesse, de facto, ser explicado segundo leis determinísticas mas, no que concerne às mentes, por seu turno, estas leis não funcionassem e as pessoas tendessem a explicar as mentes como as causadores ontológicas das ações (Wegner, 2002, p. 21). É como se a consciência *causasse* a ação, mas nada *causasse* a consciência.

Talvez a consciência seja uma ilusão – uma narrativa – porque estamos demasiado ocupados a ler as nossas mentes e não temos qualquer acesso aos processos que a criam. Somos a personagem principal de um organismo complexo – complicado demais para termos acesso conscientemente a todas as influências que recebeu e recebe para se manter vivo, dormir e adaptar-se às exigências do meio. De facto, não só a nossa maquinaria genética como a maquinaria neuronal são completamente opacas ao nível da introspeção – o mesmo é dizer que não temos qualquer acesso à forma como o nosso cérebro funciona e gera os nossos pensamentos, decisões, ações e sentimentos (Banks, 2006; Dehaene e Naccache, 2001; Bargh, 2017). Vedados que estamos da complexidade inerente a essa maquinaria, somos hábeis a ler as nossas mentes e a tirar conclusões sobre o que se está a passar dentro de nós e à nossa frente. Acreditamos facilmente que os nossos pensamentos iniciam ações, que as nossas intenções provocam comportamentos. O problema é que, tal como num espetáculo de magia, apenas estamos a ver uma ínfima parte do que está a acontecer. O verdadeiro espetáculo – a magia – acontece nos bastidores e o nosso acesso a essa maquinaria é completamente interdito.

2.5.6. *A consciência como fenómeno físico emergente*

Como vimos no capítulo 1, muitos filósofos dualistas defendem que a consciência é um fenómeno emergente (Hasker, 1999; O'Connor, 2000; Chalmers, 1996). Eles consideram que a consciência é uma propriedade emergente do cérebro físico mas que, como possui propriedades fenomenológicas, é um fenómeno não-físico. Curiosamente, muitos filósofos, psicólogos e neurocientistas fisicalistas julgam, também, que a consciência é um fenómeno emergente (Damásio, 2010; Seth, 2021; Metzinger, 2009;

Humphrey, 2023; Gazzaniga, 2011; Morsella *et al.*, 2022; Oakley e Halligan, 2017; Searle, 2019). Existe, porém, uma diferença muito significativa entre estas duas abordagens. Enquanto os primeiros defendem que o que emerge do cérebro é não-físico, os segundos postulam que aquilo que emerge do cérebro continua a ser físico. Os fisicalistas que defendem que a consciência emerge do cérebro (Searle, 1998, 2019) são, muitas vezes, apontados como dualistas (Dennett, 1991; Chalmers, 1996), porque alguns filósofos têm dificuldade em compreender que a fenomenologia pode ser um processo emergente e, ainda assim, ser um processo físico. De facto, o Universo está repleto de processos emergentes físicos, pelo que não há razão para que a consciência, enquanto processo emergente, não seja apenas mais um processo emergente físico. A nossa proposta perfilha aquilo que podemos designar por *emergentismo fisicalista*, no sentido em que defendemos que a consciência é uma entre muitas coisas que o cérebro faz, sendo que *tudo* aquilo que o cérebro faz é físico.

No artigo *Chasing the Rainbow* (2017), os neuropsicólogos David Oakley e Peter Halligan comparam a consciência a um arco-íris — evocando, aliás, a metáfora de Thomas Huxley, que via a consciência como o fumo que sai de uma locomotiva a vapor (1874). Tanto o arco-íris como o fumo — e, por analogia, a consciência — são fenómenos reais, dificilmente postos em causa. No entanto, nenhum destes fenómenos possui poder causal sobre os processos físicos que os originam. O arco-íris é gerado pela interação entre a luz e as partículas de água, mas não altera a estrutura da luz nem a disposição das gotas. Do mesmo modo, o fumo é produzido pela combustão no motor, mas não influencia o funcionamento da máquina. Segundo Oakley e Halligan, a consciência funciona de forma análoga: é um processo físico real, emergente da atividade neuronal subconsciente, mas não tem influência causal sobre os sistemas que a produzem.

De acordo com estes autores, a consciência emerge dos processos subconscientes, mas não os dirige nem os altera, mesmo que nos pareça estar no centro do controlo. Como veremos no capítulo 3, essa cisão aparente entre consciência e subconsciente reflete a tendência humana para pensar de forma fragmentada e dualista (Bohm, 1980). No entanto, em vez de uma divisão, propomos que a consciência seja entendida como a dimensão representacional da atividade cerebral — uma extensão do próprio cérebro, que espelha, de forma simbólica, o funcionamento interno. A

consciência não é um sistema à parte, mas uma manifestação do cérebro em modo auto-representativo, dedicada à construção do eu e à percepção do ambiente. Assim, longe de haver uma separação, há um contínuo entre subconsciente e consciência: um fluxo integrado de processamento neuronal com diferentes níveis de visibilidade cognitiva.

Como temos vindo a enfatizar, o facto de a consciência emergir do cérebro não justifica que a coloquemos numa dimensão não-física, como defendem muitos filósofos dualistas (Chalmers, 1995, 1996; O'Connor, 2000, 2000a; Chisholm, 2003). A consciência é um processo físico, tal como os processos neuronais que a geram — da mesma forma que o arco-íris é um fenómeno físico, embora distinto, em aparência, da luz e da água que o causam (Oakley & Halligan, 2017).

A diferença entre a consciência e o cérebro não é ontológica, mas fenoménica: corresponde apenas ao modo como o processo se manifesta. Assim como o *aspeto* visual de um arco-íris difere das propriedades da água e da luz que o originam, também o *aspeto* da consciência difere do substrato neuronal que a sustenta. Esta distinção aparente entre forma e origem repete-se em inúmeros sistemas físicos que observamos no nosso quotidiano: as ondas eletromagnéticas emitidas por um telemóvel ou um micro-ondas têm uma expressão diferente do hardware que as produz; a imagem exibida num ecrã difere dos circuitos que a tornam possível; e a voz humana, embora distinta em forma, não é menos física do que as cordas vocais e os neurónios que a produzem. Em todos estes casos, não subsistem dúvidas que estamos perante fenómenos físicos — apenas com manifestações distintas do seu suporte. Por que razão, então, haveríamos de dualizar a consciência em relação ao cérebro? A única diferença entre ambos reside na forma como se expressam. A consciência, tal como os exemplos anteriores, é uma expressão física de um sistema físico.

Em síntese, esta perspetiva rejeita qualquer cisão entre consciência e cérebro, sublinhando que a diferença entre consciência e subconsciente é apenas funcional, não ontológica. A consciência pode ser compreendida como uma representação (do eu) gerada pelo cérebro, e não como uma entidade separada. Tal como a imagem de um ecrã ou as ondas de um telemóvel se distinguem do seu suporte apenas na aparência, também a consciência se distingue do cérebro apenas na sua forma de manifestação — mantendo-se, ontologicamente, um fenómeno inteiramente físico.

3.A ontologia da informação: mente, consciência e Universo

*“Tudo quanto é, é na Natureza,
e nada pode existir nem ser concebido sem a Natureza.”*

Espinosa

Neste capítulo, iremos sugerir que a ontologia fundamental do Universo é informacional, sendo a informação o fundamento último de tudo o que existe. A partir dessa ontologia informacional, emergem fenômenos computacionais — como partículas, átomos, moléculas, células e organismos — cuja complexidade crescente permite, em certos casos, a emergência de sistemas capazes não apenas de computar, mas também de *representar* informação: os sistemas nervosos.

Esta proposta aproxima-se, em termos estruturais, da tradição monista inaugurada por Bento de Espinosa (1677/1992). Também Espinosa concebeu uma única substância, *Deus sive Natura*, da qual tudo o que existe é expressão ou modificação. No modelo que aqui propomos, essa substância única é reinterpretada como informação — não no sentido vulgar ou comunicacional do termo, mas como estrutura ontológica fundamental, inscrita em tudo o que existe. Se para Espinosa a Natureza é a própria substância que se expressa de forma infinita, para nós, a informação é o “tecido” a partir do qual o Universo (ou a Natureza) se desdobra, quer sob a forma informacional computacional (os fenômenos interpretados como puramente físicos), quer sob a forma informacional representacional (os fenômenos interpretados como fenomenológicos). Tal como a substância espinosista, também esta realidade informacional é una, necessária e autoexplicativa, sendo todos os fenômenos — físicos, biológicos ou mentais — manifestações, ou modos, dessa mesma estrutura causal.

Compreender esta estrutura exige clarificar o que se entende por informação, até porque este conceito possui uma diversidade de significados, com aplicações em domínios tão distintos como a comunicação social, linguística, a neurociência e a física. No nosso modelo utilizá-lo-emos segundo a perspectiva desenvolvida na neurociência e na física, centrada no processamento de informação, tal como foi originalmente formalizado por Claude Shannon (1948)⁶³ e que viria a influenciar decisivamente a revolução cognitivista e o desenvolvimento da era computacional a partir da década de 1950 (Chalmers, 1995; Lloyd, 2007).

A informação surge, também, como elemento central no âmbito dos vários modelos científicos mais promissores que buscam explicar a origem e o funcionamento do

⁶³ Nos anos 40 Claude Shannon reduziu a noção de informação a uma entidade tangível designada de bit, através de um sistema binário de 0 e 1 (Shannon, 1948).

Universo (Glattfelder, 2019). A informação não é apenas um elemento central em muitas das teorias ontológicas contemporâneas sobre o Universo — como a proposta de David Bohm (1980), as teorias da complexidade (Haken, 2006), a hipótese do Universo matemático (Tegmark, 2014, 2017), o princípio holográfico (Hooft, 2016), a teoria das cordas (Witten, 1995; Susskind, 1993) e a teoria da gravidade quântica em *loop* (Smolin, 2019; Rovelli, 2018, 2022) — mas também desempenha um papel fundamental na formulação e compreensão das leis que regem a realidade física (Verheyen, 2021; Glattfelder, 2019; Verlinde, 2011; Landauer, 1996; Aaronson, 2013; Livio, 2010).

Assim, no nosso modelo, a informação representa o seguinte: *é a unidade fundamental da realidade, correspondendo a padrões estruturados e processáveis que determinam o estado e a evolução dos sistemas físicos*. Neste sentido, cada entidade do Universo pode ser descrita, no nosso modelo, por uma assinatura informacional que regula as suas interações e transformações, funcionando como um código instrutório que mantém a coerência e a dinâmica do Universo.

O nosso conceito de informação está intimamente relacionado com o conceito de *bit* (*binary digit*)⁶⁴, desenvolvido inicialmente por Claude Shannon (1948) e, posteriormente, aplicado à física teórica por John Wheeler (1990) e por outros autores (Lloyd, 2007; Hooft, 2016; Verheyen, 2021; Glattfelder, 2019; Verlinde, 2011; Landauer, 1996; Aaronson, 2013; Jun et al. 2014). O *bit*, neste sentido, é a menor quantidade de dados que um sistema pode processar e armazenar e significa, portanto, a quantidade de informação necessária para distinguir entre dois estados diferentes – 0 e 1. Por conseguinte, o bit é a unidade fundamental de medida da quantidade de informação num sistema, pelo que quanto mais bits um sistema possui, mais informação tem.

Neste contexto, no nosso modelo, *um bit é uma unidade de informação que, por meio da sua codificação, gera ou desencadeia um efeito específico, como a alteração de um estado num sistema ou a emergência de um fenómeno mais complexo*. O *bit*, segundo a nossa perspetiva, corresponde portanto a uma codificação ou, para ser mais rigoroso, à abstração de uma codificação. Ou seja, o bit é uma categorização abstrata que serve para descrever uma ontologia codificada. No mesmo sentido, *um conjunto de*

⁶⁴. Um bit pode ter apenas dois valores possíveis: 0 ou 1.

bits é (uma abstração de) um conjunto de codificações que causam alterações no sistema ou conduzem à emergência de fenómenos cada vez mais complexos.

O conceito de bit que acabámos de definir está profundamente ligado à ideia de John Wheeler que, em *Information, Physics, Quantum* (1989), postula que a informação é fundamental na Natureza, estando presente em cada partícula elementar (em cada *it*). Iremos defender uma ideia semelhante: que a informação, entendida como um conjunto de códigos informacionais fundamentais que geram sistemas emergentes, constitui a base do Universo. Esta ontologia permite que os sistemas informacionais, como partículas, átomos, moléculas e células interajam entre si, formando a complexidade do que chamamos realidade.

Em síntese, a nossa proposta assenta na ideia de que o Universo possui uma ontologia monista informacional, sendo ontologicamente constituído por códigos informacionais fundamentais (os bits) que instruem o comportamento das partículas-ondas, da matéria-energia e do espaço-tempo, a partir dos quais emerge toda a realidade.

3.1.A Substância unificada informacional (SUI)

*“O doce é doce, o amargo é amargo,
o quente é quente, o frio é frio,
a cor é cor; mas, na verdade,
existem apenas átomos e vazio.”*

Demócrito

Como referido, postulamos que o Universo é ontologicamente informacional. Sugerimos que existem unidades fundamentais que sustentam, coordenam e instruem a formação de processos emergentes que formam e constituem toda a realidade a que chamamos Universo. Ao conjunto dessas unidades fundamentais que preenchem, e são, todo o Universo designamos de *Substância Unificada Informacional* (SUI).

Esta conceção de uma ontologia informacional encontra, como já referido, eco nas ideias de John Wheeler, um dos pioneiros da teoria da informação como estrutura

fundamental do Universo. Como já aludimos, *It for a bit* é uma expressão cunhada por Wheeler para explicar que cada partícula do Universo (cada *it*) possui um determinado *bit* – uma determinada quantidade de informação que pode ser conservada e processada (Wheeler, 1990). Quanto mais informação um sistema possui (ex. um átomo, uma célula ou um organismo), mais bits de informação (mais unidades fundamentais de informação⁶⁵) esse mesmo sistema pode conservar e processar (Lloyd, 2007 e 2013). Ou seja, quanto mais bits de informação um sistema possui, maior o seu nível de entropia e maior a complexidade desse sistema (Shannon, 1948). Assim, pode afirmar-se que a complexidade de um sistema é proporcional à quantidade de informação que consegue processar (Tononi, 2008).

Gostaríamos de destacar quatro ideias de Wheeler (1990) que iremos utilizar no nosso modelo: 1) a informação é uma entidade fundamental, mais primordial que a própria matéria-energia; 2) a informação (*bit*) *gives rise to every it* – cada partícula, cada campo de forças, a matéria, energia e o espaço-tempo – embora no nosso modelo a informação não cause a física, como veremos, a informação é física; 3) todas as coisas que existem possuem uma ontologia informacional; 4) o Universo é composto por redes de informação ocultas que causam todos os processos, leis e constantes físicas. Assim, e utilizando as palavras do físico Seth Lloyd, iremos sugerir que: “The history of the universe is, in effect, a huge and ongoing quantum computer” (2007, [s.p.]).

A ciência desenvolveu uma compreensão profundamente consolidada sobre a composição dos átomos e sobre como estes interagem para formar moléculas, células, organismos, sistemas nervosos e objetos. Um átomo é caracterizado por um número específico de prótons, o que define a sua identidade, e possui também uma estrutura elétrica, eletronegatividade, tamanho e energia de ionização. Esses conceitos – como o número de prótons, a estrutura elétrica e a eletronegatividade – são fundamentais para entender os átomos e como eles podem se organizar para formar moléculas. No entanto, apesar de conhecermos essas propriedades, não sabemos completamente o que é intrínseco a esses processos que permite, de facto, a formação de moléculas. Como

⁶⁵ Podemos-lhes chamar também símbolos.

refere o físico e matemático Marcus du Sautoy no seu livro *O Que Não Podemos Saber* (2018):

“Se pegar num sistema físico como [...] as moléculas que fazem minha sala, não possuo em geral um conhecimento completo do estado microscópico dessas partículas, mas apenas uma descrição macroscópica global que pode permitir múltiplos estados microscópicos” (2018, p. 340).

No mesmo sentido, sabemos que dois elementos – dois átomos de hidrogénio e um de oxigénio – combinam para criar uma molécula de água mas existe um terceiro elemento nesse processo: a água, que, objetivamente não sabemos como é que surge da interações desses átomos. Como muito bem captou o escritor e poeta H.D. Lawrence:

“Water is H₂O, hydrogen two parts, oxygen one,
But there is also a third thing, that makes it water
And nobody knows what that is...” (1989, p. 167)

Sabemos também que arranjos específicos das moléculas de hidrogénio e oxigénio resultam nos três estados da água: gasoso, líquido e sólido. Contudo, também não compreendemos totalmente o que esses arranjos possuem de intrínseco para possibilitar a transição entre esses estados moleculares (Tegmark, 2014). Além disso, a ciência ainda não consegue explicar como a interação entre átomos “não-vivos” pode gerar moléculas vivas. Sabemos que isso ocorre, mas não entendemos o processo (Davies, 2019). No mesmo sentido o físico Brian Green afirma no seu livro *Até ao Fim dos Tempos* (2021) o seguinte:

“Concentre-se numa única molécula de água, num átomo de hidrogénio ou num eletrão individual e descobrirá que nenhum deles tem qualquer marca a delinear se são uma componente de algo vivo ou morto, ou de algo animado ou inanimado” (Greene, 2021, p. 83).

Por conseguinte, sabemos que os átomos formam moléculas de acordo com um arranjo atómico específico; sabemos que as moléculas podem assumir vários estados; bem como sabemos que átomos “não-vivos” podem dar origem a moléculas “vivas”, mas não entendemos completamente como é que essas combinações resultam, efetivamente, na formação de processos emergentes mais complexos. O “como” aqui refere-se ao que há de intrínseco aos átomos que os faz ter afinidades com certos átomos e

repelir outros; ao que há de intrínseco a certos arranjos moleculares que permite que algumas moléculas assumam estados distintos e possam criar organismos vivos.

Sugerimos que todos esses processos podem ser explicados por meio dos bits de informação (ou instruções) contidos nas partículas elementares, os quais lhes permitem *computar* informação e, desse modo, dar origem a sistemas mais complexos através de determinadas interações. Por outras palavras, propomos que as partículas, os átomos e as moléculas estão “carregados” com algoritmos que codificam instruções (Tegmark, 2014). A nosso ver, é essa informação que explica não apenas as afinidades físicas e químicas, mas também a emergência de processos biológicos a partir de processos físicos – isto é, como prótons e elétrons originam átomos, como os átomos, por sua vez, formam moléculas e sistemas emergentes, e como as moléculas formam organismos. Sugerimos ainda que os conceitos utilizados para descrever esses comportamentos são representações semânticas e simbólicas de uma informação subjacente, a qual não conseguimos aceder diretamente, podendo apenas representá-la através desses próprios conceitos e de fórmulas matemáticas.

Neste enquadramento, é possível evocar a proposta monista de Espinosa (1677/1992), segundo a qual tudo o que existe é expressão de uma única substância com infinitos atributos, nos quais o pensamento e a extensão estão inscritos. A realidade, segundo Espinosa, manifesta-se simultaneamente como ideia (ordem do pensamento) e como corpo (ordem da extensão), sendo ambas expressões da mesma ordem causal. Esta leitura permite compreender que a informação — enquanto estrutura que determina o modo de ser das coisas — pode ser vista como o elo entre essas duas ordens, isto é, como o conteúdo da própria substância manifestando-se sob diferentes modos. Tal como os modos expressam os atributos da substância espinosista, no nosso modelo os sistemas físicos e cognitivos podem, também, ser pensados como expressões concretas dessa estrutura subjacente, reafirmando a unidade absoluta entre mente e corpo que caracteriza o monismo de Espinosa.

De forma convergente, David Bohm propõe, em *Wholeness and the Implicate Order* (1980) a existência de uma ordem implicada (*implicate order*), um domínio mais profundo e não-local da realidade onde toda a informação está contida de forma potencial e interconectada. Para Bohm, o mundo visível — aquilo a que chama ordem explicada — é apenas uma manifestação ou projeção dessa ordem mais fundamental. As

partículas, longe de serem entidades isoladas, são expressões dinâmicas de um campo informacional subjacente. A realidade, portanto, não é composta por objetos discretos, mas por padrões de informação que se desdobram e tornam explícitos certos conteúdos do todo. Esta abordagem permite conceber que a organização da matéria e a emergência da vida são o resultado do desdobramento contínuo de uma ordem causal mais profunda, o que está em consonância com a ideia de que os átomos e partículas contêm instruções que determinam os seus comportamentos. Neste contexto, pode ser convocado o conceito *bohmiiano* de holimovimento, entendido como o fluxo indivisível entre a consciência e a matéria, no qual o ser humano não é um sujeito isolado frente a um mundo externo, mas uma expressão do próprio movimento fundamental do universo. O holimovimento traduz, assim, a unidade entre mente e matéria como dois aspetos inseparáveis de um mesmo processo, ecoando a visão de Bohm de que a fragmentação é uma ilusão gerada pela ordem explicada.

Em *The Demon in the Machine* (2019), Paul Davies sugere que a informação é uma característica fundamental da própria estrutura do universo. Para ele, a realidade física, incluindo os fenómenos biológicos, pode ser explicada com base em informações que instruem o comportamento das partículas e das leis da física. Davies vê o universo como sendo fundamentalmente informacional, no sentido de que as leis da física podem ser vistas como programas ou instruções que definem como as partículas devem interagir, de maneira semelhante a como um computador segue códigos para gerar processos. Ele explora como a vida, de facto, surge da matéria não viva e como isso pode estar relacionado a estruturas de informação complexas que instruem a organização da matéria de maneira específica. Em suma, ele explora a ideia de que os átomos e as partículas carregam informações que não são apenas descrições matemáticas ou observações, mas têm um papel ativo na formação da realidade, incluindo a vida. (2019).

No mesmo sentido, o físico Max Tegmark postula em *Our Mathematical Universe* (2014) que as partículas possuem códigos matemáticos que permitem criar determinados processos emergentes, designadamente estados moleculares específicos, como os acima referidos. Nas palavras de Tegmark:

“The very fabric of our physical world, space itself, could be a purely mathematical object in the sense that its only properties are mathematical properties – numbers such as dimensionality, curvature and topology” (2014, p. 43).

Para ele, a realidade física pode ser descrita por uma estrutura matemática fundamental, que é um reflexo de um código informacional subjacente. De acordo com essa visão, as leis da física seriam uma expressão matemática dessa informação que estrutura o Universo e toda a realidade seria, portanto, uma grande computação matemática. Em poucas palavras, Tegmark vê a matemática como a base ontológica do universo, onde a informação é codificada nos padrões e nas interações matemáticas que observamos como leis físicas (2014).

Os físicos Lee Smolin (2019) e Carlo Rovelli (2018) – dois dos físicos mais consagrados da teoria da Gravidade Quântica em Loop – postulam que o espaço-tempo é granular, ou seja, que o tecido do espaço-tempo não é contínuo, mas um sistema emergente composto por partículas que são o próprio espaço-tempo, numa proposta que não está muito longe do atomismo de Demócrito. Smolin e Rovelli sugerem, assim, que o espaço-tempo é constituído por *quanta* — unidades discretas que não apenas formam a geometria do Universo, mas também dão origem a tudo o que nele existe. Ou seja, os objetos, os campos e as partículas não estão simplesmente no espaço-tempo, mas são expressões da mesma estrutura quântica fundamental. A gravidade, os corpos e os campos emergem, portanto, de uma realidade granular coordenada por uma estrutura ontológica fundamental.

No modelo que iremos apresentar, sugerimos que tudo no Universo, cada partícula, cada objeto e cada ser vivo, sendo constituídos por códigos instrutórios (unidades fundamentais, ou bits), por muito diferente que possam parecer, podem ser explicados e compreendidos em função de uma *substância unificada informacional* (SUI). A SUI é onnipresente – e expressa-se fisicamente em tudo o que existe. A SUI é, em rigor, o Universo (ou a Natureza). Trata-se, no entanto, de uma parte oculta do Universo, que nos chega aos sentidos apenas sob a forma física ou sob a sua representação simbólica — esta última manifestada e pensada através do próprio conceito de Universo. Esta é uma ideia muito semelhante à substância espinosista (Espinosa, 1677/1992), ao holismo de Bohm (1980), às redes informacionais ocultas de Davies (2019) ao

universo matemático de Tegmark (2014) e, em grande medida, ao modelo granular de Rovelli (2018) e de Smolin (2019).

3.1.1. *A física como epistemologia da informação: superando a ilusão da dualidade*

Embora a informação possa parecer ter uma natureza etérea - como se pensou durante décadas que a energia tivesse - a informação é física. Como o físico Rolf Landauer defendeu no artigo *The physical nature of information* (1996): “Information is not a disembodied abstract entity; it is always tied to a physical representation” (1996, p. 188). Na nossa teoria, a informação é a ontologia da física e esta é a expressão epistemológica da primeira. A física é, nesta perspetiva, a representação possível da informação, ao passo que a informação é a face oculta da física. Informação e física são, assim, duas faces da mesma moeda.

Nas teorias da informação, o conceito de entropia sugere que há sempre mais informação potencial num sistema do que aquilo que conseguimos aceder ou processar (Landauer, 1996; Shannon, 1948; Wheeler, 1990; Tegmark, 2014; Jun et al. 2014). Este é um dos grandes mistérios da física atual. Porque o nível de entropia é sempre maior do que a informação que podemos observar? Sugerimos que o “véu” entre a informação pura e a sua manifestação física é uma limitação inerente ao aparato cognitivo humano, ou mesmo ao aparato de qualquer observador condicionado pelas suas limitações cognitivo-sensoriais. Deste modo, a informação causa os fenómenos mas estes surgem-nos, parcialmente, cognoscíveis através das leis da física. A física é, neste sentido, a semântica e o simbolismo da informação (é a ordem explicada de Bohm) ou, dito de outra forma, as leis físicas são manifestações cognoscíveis da ontologia incognoscível da informação (da ordem implicada de Bohm).

Esta nossa perspetiva está alinhada com a visão do físico Edward Fredkin que, em *An Introduction to Digital Philosophy* (2003), argumenta que o estado de qualquer sistema físico deve ter uma manifestação digital. Ou seja, se cada entidade possui informação, isso significa que existe uma assinatura digital para cada uma dessas entidades. Segundo Fredkin, uma coisa é a *fisicalidade epistemológica* (aquilo que no nosso modelo iremos designar de informação computacional e representacional) da entidade, e uma outra é *fisicalidade ontológica* (aquilo que iremos designar de informação ontológica) da entidade. É como se a informação possuísse dois aspetos: um epistemológico (e

físico) e um outro ontológico (informacional). Apesar de corroborarmos com a ideia de Fredkin que cada entidade física possui uma assinatura digital (informacional), consideramos que não existe necessidade de dualizar o processo. A informação é tão física como a física em si. É apenas no aspeto – na dimensão epistemológica – que a informação nos pode parecer dual. Ontologicamente, a informação (seja ontológica ou fenomenológica) *é sempre física e a física é sempre informação* (Fredkin, 2003; Landauer, 1996; Verheyen, 2021,).

Max Tegmark (2014) defende, no mesmo sentido, que os blocos fundamentais do Universo são estruturas matemáticas, sugerindo que a realidade é coordenada por algoritmos matemáticos. No entanto, só conseguimos perceber a manifestação epistemológica desses códigos, não a sua estrutura matemática subjacente. Para compreender a realidade, é necessário desenvolver modelos que traduzam os padrões matemáticos invisíveis em conceitos semânticos, como forma e cor. Tegmark (2014, 2017) sugere, portanto, que as propriedades matemáticas são as mesmas que as propriedades físicas, apenas expressas de formas distintas. A carga e o movimento giratório de uma partícula, por exemplo, são apenas fórmulas matemáticas ocultas, representadas simbolicamente como conceitos físicos. Estamos, em grande medida, alinhados com esta perspetiva. Porém, sugerimos que a ontologia não é matemática mas sim informacional. A matemática é uma forma de decodificação – de simbologia – de alguma ontologia do Universo mas não é a ontologia em si. Enquanto Tegmark assume que os códigos matemáticos são a essência última do Universo, sugerimos que a matemática é apenas a epistemologia da informação, um sistema de representação que permite modelar os padrões informacionais subjacentes

O físico e matemático Seth Lloyd, em *Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos* (2007) argumenta que toda entidade contém informação e que não há nada no Universo sem *bits* informacionais. A cada nível de análise – das partículas aos átomos e moléculas – encontramos sempre informação, até que chegamos a um ponto onde ou deixamos de aceder a essa informação ou nada mais existe. No entanto, sugerimos que essa aparente ausência de informação pode ser uma limitação epistemológica, não ontológica. Pode haver uma informação implícita e estrutural, inacessível à observação direta, mas que se desdobra nos quarks, prótons, neutrões e átomos que permeiam – e são – a realidade. Além disso, carga elétrica,

magnetismo e quarks não são realidades últimas, mas sim representações de uma estrutura mais fundamental. Os códigos subjacentes a essas propriedades não nos surgem diretamente, sendo decodificados por meio de representações semânticas ou matemáticas, que replicam ou aproximam a realidade ontológica (Tegmark, 2014, 2017).

Vários estudos experimentais recentes têm vindo a oferecer exemplos particularmente elucidativos daquilo que propomos nesta ontologia informacional-física: que a física é a epistemologia da informação, e que os fenómenos físicos são manifestações visíveis de estruturas causais informacionais. Estes estudos, ao descreverem comportamentos quânticos complexos através de modelos deterministas, reforçam a ideia de que a realidade não exige qualquer tipo de indeterminação ontológica. Pelo contrário, mostram que mesmo os sistemas quânticos mais sofisticados podem ser descritos com base em equações causais, sem recorrer ao acaso como princípio fundamental. Esta coerência entre teoria e observação contribui de forma decisiva para demonstrar que a diversidade e complexidade do real são expressões de algoritmos informacionais profundos, cuja manifestação observável é aquilo a que chamamos física.

Um desses estudos particularmente reveladores (Tanzi et al., 2019) é o que demonstra a existência do supersólido, outrora considerado paradoxal. Hoje, o supersólido é compreendido como a manifestação de uma estrutura causal altamente organizada, e não como uma anomalia física. O estudo demonstrou que este estado resulta da bifurcação determinista de um modo de oscilação axial num condensado dipolar de átomos de disprósio, à medida que um único parâmetro de interação é ajustado. O sistema revela dois modos distintos: um relacionado com a compressibilidade da estrutura cristalina — ou seja, a capacidade do sistema assumir uma configuração espacial regular, típica dos sólidos — e outro ligado à oscilação superfluida, que corresponde ao fluxo de partículas sem resistência, característica dos líquidos quânticos. Estes dois modos coexistem no mesmo sistema, conferindo ao supersólido a sua natureza híbrida — sólida e superfluida em simultâneo. Importa sublinhar que esta dualidade não implica a existência de dois sistemas distintos (um clássico e outro quântico), mas sim de um único sistema quântico que manifesta, de forma unificada, propriedades tradicionalmente atribuídas a domínios separados. O supersólido demonstra, assim, que estados físicos

aparentemente incompatíveis podem emergir de uma única estrutura causal, sem quebra de continuidade nem violação das leis do sistema.

Ambos os modos foram previstos teoricamente com base em equações deterministas que descrevem o comportamento coletivo de átomos a temperaturas próximas do zero absoluto. Trata-se de uma das equações mais utilizadas na física de condensados de Bose-Einstein e que não requer qualquer elemento de aleatoriedade ou colapso da função de onda, no sentido tradicional da interpretação de Copenhaga. O que muitas vezes é interpretado como "colapso" — a passagem de um estado de sobreposição para um estado concreto — é aqui explicado como uma transição de fase contínua, que ocorre de forma determinista e gradual, sem necessidade de postular eventos *acausais* ou saltos quânticos imprevisíveis. Fica assim claro que não há qualquer necessidade de recorrer a modelos probabilísticos ou indeterministas, mesmo em contexto quântico. A evolução do sistema é, desde o início, modelada e compreendida como consequência de estruturas causais bem definidas — ou seja, o comportamento do supersólido é antecipado através de equações deterministas.

A coexistência de propriedades aparentemente antagônicas — solidez e superfluidez — no mesmo estado reforça a robustez do determinismo: a complexidade da realidade não exige indeterminação, mas sim uma estrutura causal suficientemente rica e sofisticada para acomodar fenômenos que, à primeira vista, nos parecem contraditórios. Esta coexistência demonstra que os mesmos princípios deterministas podem gerar, simultaneamente, propriedades clássicas (sólidas) e propriedades quânticas (superfluidez), sem qualquer necessidade de recorrer a uma ontologia dual ou a uma separação entre "mundo clássico" e "mundo quântico". Em suma, este exemplo mostra que a diversidade dos comportamentos físicos pode emergir de uma única matriz determinística, sem qualquer quebra na cadeia de causalidade.

O supersólido surge, assim, como uma manifestação de informação: um conjunto de algoritmos ontológicos que orientam o comportamento do sistema e tornam possível a emergência de estados complexos. Esta informação emergente corresponde à manifestação ontológica da física. O que a física descreve nesta experiência é, na verdade, a computação executada pelas estruturas profundas do Universo. Aquilo que poderia ser visto como comportamento exótico ou imprevisível é apenas a expressão de uma organização causal altamente refinada. O supersólido, longe de ser uma exceção,

confirma a nossa tese de que a física é a epistemologia da informação. O comportamento dinâmico descrito pela física é, na verdade, a expressão visível de estruturas causais informacionais. Quando um fenômeno nos parece imprevisível ou “estranho”, isso não indica indeterminação ontológica, mas apenas o nosso desconhecimento momentâneo da arquitetura informacional que o regula (Tanzi et al., 2019).

A nossa perspectiva é, igualmente, reforçada pelos desenvolvimentos na química quântica, onde os processos químicos são cada vez mais compreendidos como computacionais e informacionais. Por exemplo, Aspuru-Guzik e Walther (2012) demonstram que a química quântica pode ser modelada computacionalmente através de algoritmos quânticos, sustentando a noção de que a natureza fundamental das interações químicas é informacional. De forma semelhante, Kassal et al. (2008) exploram como o processamento de informação quântica pode acelerar simulações químicas, reforçando a ideia de que as transformações moleculares são regidas por regras informacionais e não meramente por interações mecanicistas. Estes avanços sugerem que a própria física — e, por consequência, a química — deve ser entendida como a expressão formal e observável de uma arquitetura informacional subjacente, na qual a informação estrutural se manifesta como leis físicas e a informação computacional como dinâmica causal.

Adicionalmente, os processos biológicos também revelam uma base informacional no seu núcleo. Lambert et al. (2013) discutem o papel da coerência quântica em sistemas biológicos, particularmente na fotossíntese, onde a transferência de energia é otimizada através de efeitos quânticos. Isto sugere que os sistemas vivos operam não apenas por meio de reações bioquímicas, mas também através de princípios informacionais determinísticos que regulam a distribuição de energia. De forma semelhante, Patel (2001) explora a hipótese de que mecanismos quânticos poderão ter influenciado a origem da vida, alinhando-se com a noção de que os processos biológicos são orientados computacional e informacionalmente. Estes estudos reforçam a ideia de que a física, enquanto descrição formal dos comportamentos observáveis, é a manifestação da informação computacional subjacente que estrutura a dinâmica dos sistemas vivos.

Walker e Davies (2013) expandem esta ideia ao proporem que a vida emerge através da organização hierárquica da informação, apoiando ainda mais o conceito de que a diferenciação informacional está na base da complexidade biológica. Estes contributos indicam que a própria física, ao permitir e regular os fenômenos biológicos através

de estruturas causais precisas, deve ser entendida como a manifestação funcional de uma arquitetura informacional subjacente, em que a organização física resulta da computação de informação estrutural.

Todos estes contributos convergem para sustentar o nosso modelo, que propõe a existência de uma realidade codificada intrínseca (a informação) e uma percepção (ou interpretação) dessa realidade (a física) (Bohm, 1980; Wheeler, 1990; Lloyd, 2007; Tegmark, 2014). A física, no nosso modelo, é, por conseguinte, uma manifestação simbólica e epistemológica da informação que é percecionada e compreendida por uma linguagem que usamos para descrever padrões e relações – uma linguagem que fragmenta a realidade (Bohm 1980). Ou seja, a nossa percepção da física fragmenta a unidade da informação pelo que a percepção nunca é igual à coisa em si – é uma aproximação. Podemos imaginar que o Universo é uma tela onde o filme (informação pura) é projetado. Nós vemos apenas a luz e as imagens na tela (física), mas o que está por trás – como o filme está a ser projetado – é inacessível diretamente. O que vemos na tela é uma manifestação causal dessa informação.

Se a realidade física existe independentemente da nossa percepção, mas a nossa compreensão dessa realidade é mediada pela informação que conseguimos processar, isso pode ser alinhado com uma forma de realismo mais sofisticado, como o realismo epistemológico. Nesse caso, a informação que processamos não substituiria a realidade física, mas seria uma forma de conhecê-la ou descrevê-la, sem negar a existência objetiva da realidade. Esta perspetiva elimina a necessidade de um intermediário (Descartes, 1649/2023), de um mecanismo causal (O'Connor, 2000; Chisholm, 2003; Kane, 2005) ou de uma epistemologia não-realista, como a Kantiana (1871/2008). Em vez disso, propomos que a física é a forma como o humano (ou qualquer observador com limitações semelhantes) interpreta e interage com a informação que constitui o Universo. Em poucas palavras defendemos que qualquer dualidade entre física e informação é ilusória. A dualidade que percecionamos é uma consequência de como a mente humana interpreta e organiza dados sensoriais e conceptuais. A realidade é informação pura, pelo que não está dividida em camadas "informacional" e "física". Toda a separação é criada pela nossa relação com a ontologia e pela necessidade do nosso aparato cognitivo em fragmentar a realidade (Bohm, 1980).

Esta perspectiva tem implicações para o problema mente-corpo ao sugerir que a dualidade deste problema, tal como postulada por alguns autores dualistas (Descartes, 1641/2023; Chalmers, 1995 e 1996; Chisholm, 2003; O'Connor, 2000; Nagel, 1974; Hasker, 1999) é apenas ilusória (Espinosa, 1677/1992; Bohm, 1980). Na essência, mente e corpo fazem parte da SUI, pelo que são apenas dois aspetos da mesma realidade informacional, não obstante poderem ser descritos através de leis físicas. Em vez de perguntarmos como a mente interage com o corpo, a questão torna-se como a informação se manifesta de diferentes formas. Esta questão encontra uma resposta no nosso modelo: mente e corpo são apenas duas manifestações da mesma ontologia informacional.

A SUI, no nosso ponto de vista, tem a capacidade de derrubar qualquer tipo de dualismos ao apresentar um modelo ontológico monista no qual uma única substância – a informação (física) – pode criar uma imensidão de entidades através da preservação, da conservação e do processamento (e, como veremos, transformação) de informação através de leis causais, sejam elas objetos, organismos ou mentes conscientes. Estas surgem quando bits de informação adquirem a capacidade de representarem um modelo de eu. Como refere Max Tegmark em *Our Mathematical Universe*: “...consciousness is the way information feels when being processed in certain complex ways” (Tegmark, 2014, p. 326). Talvez não sejamos mais do que informação a representar informação. Neste sentido, o físico Dirk Meijer, também postula que: “Energy, Intelligence, and Consciousness look and act differently [...] and yet are also made from the same underlying substance: information” (2015, p. 66), reforçando a ideia de que tudo o que existe pode ser descrito como diferentes expressões de uma única substância ontológica: a informação.

Em suma, a proposta central deste trabalho é que a ontologia fundamental do Universo é informacional, sem necessidade de dualizar a informação e a física. A informação não apenas regula e coordena o Universo através de leis, mas constitui a sua base ontológica, sendo a realidade física a sua manifestação epistemológica. Autores como Bohm (1952; 1980), Wheeler (1980), Lloyd (2007), Tegmark (2014) e Davies (2019) sugerem que a informação instrui as partículas e os átomos, permitindo a emergência de sistemas complexos, incluindo a vida e a consciência. A SUI que propomos é contínua e permeia todo o Universo, sustentando a matéria, a energia e os processos emergentes. A conservação da informação, conforme demonstrado por Landauer (1996), por Jun

(2014) e Lloyd (Lloyd, 2007), garante a estabilidade do cosmos, onde a transformação da informação gera complexidade sem criar novos dados. Assim, a física pode ser vista como a simbolização da informação, e a consciência, como veremos, a face representacional da informação, reforçando um modelo monista sem necessidade de dualismos.

3.2. Os quatro princípios da substância unificada informacional (SUI)

No nosso modelo, existem 4 princípios necessários no que concerne à SUI: 1) tem de ser *conservada*, que é o que permite a existência em permanência do espaço-tempo e da matéria-energia; 2) tem de poder ser transmitida, ou transformada, através do *processamento* de informação, que é o que conduz às afinidades físicas e químicas e, subsequentemente, à criação de matéria-energia; 3) tem de ser finita, o que permite que o Universo possua um número delimitado de informação que garanta a sua organização funcional; 4) deve ser determinística, ou seja, a informação deve gerar efeitos regidos por leis causais e regulares, o que não só garante a criação de matéria-energia, mas também a coerência estrutural e a evolução dos sistemas, nomeadamente do espaço-tempo.

Neste sentido, podemos atualizar o nosso conceito de informação. A informação *é a unidade fundamental da realidade, correspondendo a padrões estruturados e processáveis que determinam o estado e a evolução dos sistemas físicos. Trata-se de uma codificação finita, conservada e processável, capaz de gerar transformações no seu próprio conteúdo e expressão, e dotada de capacidade causal, na medida em que origina propriedades emergentes estáveis e regulares através da sua interação.*

3.2.1. A informação tem de ser conservada

O facto de a informação poder ser conservada (1) permite a existência de todo o espaço-tempo e toda a matéria-energia. Uma determinada partícula, como um quark, um eletrão ou um fóton, possui uma determinada matéria que equivale ao seu nível de energia. Essa matéria-energia, ao ser conservada, permite a *existência* em permanência de partículas e de sistemas. Sem a conservação, a matéria-energia não poderia sustentar estruturas estáveis, impossibilitando o surgimento de sistemas emergentes mais

complexos, como o oxigénio, a água, as células, os organismos, as mentes e a consciência. Em termos simples, é a conservação da informação em cada partícula que possibilita a existência da matéria-energia e, a partir daí, a emergência de sistemas cada vez mais complexos — resultando no surgimento de toda a matéria-energia observável no Universo. Como refere Seth Lloyd:

“At first, the patterns it produced were simple, comprising elementary particles and establishing the fundamental laws of physics. In time, as it processed more and more information, the universe spun out ever more intricate and complex patterns, including galaxies, stars, and planets. Life, language, human beings, society, culture – all owe their ability of matter and energy to process information” (2007, [s.p.]).

Dito por outras palavras, toda a matéria-energia existe porque é composta por — e simultaneamente é — bits que contêm a capacidade de conservar informação, a qual pode depois ser processada na formação de novos sistemas emergentes mais complexos, como moléculas ou organismos. Sem essa capacidade, a água não poderia ser água e uma cadeira nunca poderia ser uma cadeira. Porquê? Porque, por exemplo, dois átomos, como os de oxigénio e hidrogénio, se não conservassem informação — isto é, se as partículas elementares que os constituem não mantivessem as suas propriedades e relações de forma estável — não teriam a capacidade de se inter-relacionar e dar origem a um sistema emergente mais complexo, como a água. De forma semelhante, sistemas macroscópicos, como uma cadeira de madeira, só existem porque os átomos que os compõem — como os de carbono, hidrogénio e oxigénio — mantêm as suas propriedades e interações de forma estável e conservada, graças à estabilidade das partículas elementares (como os quarks e os eletrões) que constituem esses átomos. Sem a conservação de informação não existiria matéria simplesmente porque a informação — essencial para os blocos de qualquer sistema - perder-se-ia (Fredkin, 2003).

Em suma, sabemos que quarks, protões, eletrões, átomos e moléculas conservam uma organização estrutural estável — ou seja, conservam informação —, fundem-se e exibem afinidades químicas específicas. No entanto, se os átomos interagem de forma precisa para formar moléculas, como a água, é plausível que exista um conjunto de instruções subjacentes que rege essas interações e que ainda não compreendemos inteiramente. Sugerimos que essas instruções estão codificadas na informação inerente

às partículas, funcionando como algoritmos estruturais que determinam a conservação e a organização da matéria-energia.

3.2.2. A informação tem de ser processada

A informação necessita, igualmente, de ser (2) processada (isto é, transmitida) de modo a poder originar sistemas emergentes mais complexos. O processamento, ou transmissão, de informação faz com que a mesma possa ser *transformada* no aspeto. Dois átomos de hidrogénio e um de oxigénio, por exemplo, ao processarem informação criam uma molécula (H₂O), mas nenhuma informação é perdida. Apenas muda o *aspeto* da informação. A estrutura informacional do hidrogénio e do oxigénio adquire uma nova codificação, agora em forma de água. Na sua ontologia, contudo, a informação é a mesma. O arranjo, ou o padrão, do processamento dos bits de informação é que permite a criação de novos processos emergentes. Isto é, para além de ser conservada, a informação, necessita de poder ser *transformada* (no aspeto) através do processamento, de modo a que sistemas mais complexos possam emergir.

Por conseguinte, quando referimos que a informação é transformada, não estamos a defender que nova informação está a ser criada. Estamos a sugerir apenas que informação física pode ser transformada em informação química e a informação química pode ser transformada em informação biológica. Física, química e biologia são conceitos que ilustram uma ontologia informacional. Neste sentido, um átomo de hidrogénio possui informação conservada mas, ao ser também *processada*, obedece a um princípio de transmissão – transmissão essa que é feita com determinadas partículas com a mesma afinidade que conduz à transformação (epistemológica) e, subsequentemente, emergência de novas sistemas mais complexos.

A nossa perspetiva alinha-se com várias propostas contemporâneas. Segundo o físico Dirk Meijer, o campo de informação quântica é o mecanismo através do qual partículas elementares interagem e se organizam para formar sistemas mais complexos, como moléculas, macromoléculas (proteínas, ácidos nucleicos) e, eventualmente, sistemas vivos. Ele argumenta que essas interações não são aleatórias, mas regidas por princípios informacionais intrínsecos que conferem ordem e direção ao universo. Como refere Meijer em *The Information Universe* (2012) a física quântica é um campo de

informação que: “possibilitou a formação de estruturas macromoleculares altamente complexas⁶⁶”, (Meijer, 2012, p.1).

No mesmo sentido, Seth Lloyd argumenta que:

“Once a physical system [like an atom or even a proton] possesses the ability to process information at a rudimental level – performing simple operations on a few bits at a time – arbitrarily complicated forms of information processing can be built up from these basic operations” (2007, [s.p.]).

Note-se que, como já foi referido, ninguém sabe porque os átomos de hidrogénio têm afinidades físicas com átomos de oxigénio. O que existe intrinsecamente a estas partículas que lhes permite interagir continua a ser um mistério. Apenas se sabe que essas afinidade existem. Ora, sugerimos que existem códigos instrutórios que permitem o processamento de informação e que, subsequentemente, determinam essas afinidades (Davies, 2019; Tegmark, 2014; Bohm, 1980; Lloyd, 2007; Ribeiro, 2013). Ou seja, são esses *códigos instrutórios* (uma programação fundamental) que permitem que a informação possa ser processada, o que conduz à *transformação* de informação em sistemas mais complexos.

O processamento de informação é, portanto, o princípio segundo o qual as partículas podem transmitir a sua informação a outras partículas, ao mesmo tempo que recebem informação dessas partículas, fundindo-se deste modo num sistema emergente. Trata-se de um princípio análogo ao que ocorre nos sistemas celulares, quando uma célula procariótica primitiva — a que Seth Lloyd chama “avó da revolução do processamento de informação” (2007, [s.p.]) — transmitiu, pela primeira vez, informação a outra célula procariótica, dando origem a uma célula eucariótica, base de todos os organismos multicelulares (Ribeiro, 2013). A existência de afinidades químicas sugere a presença de um nível subjacente de instruções informacionais, que regulam como as partículas interagem. Esses códigos instrutórios, embora não diretamente observáveis, podem ser inferidos a partir da regularidade e previsibilidade dos processos físicos e químicos.

⁶⁶ Tradução do autor.

Sugerimos, portanto, que se a informação fosse apenas conservada, mas não processada, ou transmitida, o Universo permaneceria num estado inerte, onde as partículas elementares existiriam isoladas, sem interações que possibilitassem a emergência de sistemas mais complexos. Não existiria, por exemplo, água porque os átomos de hidrogénio e de oxigénio, embora pudessem conservar informação, não teriam possibilidade de *processar* a sua informação e, subsequentemente, *transformar* dois átomos de hidrogénio e um de oxigénio numa molécula de água. Em rigor, aliás, nem sequer existiriam átomos já que as partículas subatómicas, ao não processarem informação, não teriam sequer a capacidade dar origem a átomos.

A matéria-energia existiria num estado inerte, num mar imenso de partículas elementares. Talvez existissem apenas bosões e quarks, entre outras partículas elementares mas, sem a capacidade de transmitir informação esses bosões e *quarks* existiriam isolados, sem terem qualquer afinidade com outras partículas. Por conseguinte, sem transmissão de informação não seria possível criar novos sistemas mais complexos pelo que seria impossível existirem processos emergentes.

3.2.3. *A informação tem de ser finita*

Para que os sistemas que compõem a realidade possam manter a sua integridade, é necessário que a informação seja finita, pois a sua perda comprometeria as instruções fundamentais que sustentam a existência dos próprios sistemas (Hooft, 2016; Lloyd, 2007; Davies, 2014; Rovelli, 2018). Se um sistema físico ou biológico é construído a partir de uma quantidade específica de bits de informação, a eliminação ou degradação dessa informação pode comprometer a integridade do sistema e conduzi-lo ao colapso. Um organismo biológico, por exemplo, depende da estabilidade da sua sequência genética — alterações significativas nessa informação podem torná-lo menos apto e ameaçar a sua sobrevivência.

De forma análoga, qualquer sistema físico exige uma quantidade finita e conservada de informação para existir de forma estruturalmente consistente e sustentada no tempo. Os algoritmos que propomos como base ontológica do Universo contêm instruções específicas e limitadas que regulam e coordenam os sistemas físicos, dispensando, por isso, qualquer fluxo ilimitado de informação. A existência de limites informacionais

é o que permite a organização, a previsibilidade e a emergência ordenada das estruturas no cosmos.

A finitude da informação, longe de representar uma limitação, é precisamente o que garante a consistência estrutural e a persistência ontológica dos sistemas. A perda de informação num sistema organizado corresponderia à eliminação de elementos cruciais à sua existência. É por essa razão que, segundo algumas das propostas mais robustas da física contemporânea, a informação não pode ser perdida — nem mesmo num buraco negro, onde se admite que toda a informação é preservada na sua superfície, o chamado horizonte de eventos (Hooft, 2016; Landauer, 1996; Jun et al., 2014).

Tal como um organismo vivo não pode perder a sua informação genética sem comprometer a sua integridade (Curado, 2017), também o Universo não pode perder informação sem comprometer as estruturas que o compõem. Sugerimos, assim, que a finitude da informação é uma condição necessária para que um sistema, como o Universo, possa preservar a sua integridade sistémica. O Universo conteria, portanto, exatamente a quantidade de informação necessária para existir e sustentar a sua própria estrutura (Lloyd, 2007; Tegmark, 2014). Nem mais, nem menos. Nunca foi tão apropriada a célebre frase atribuída a Antoine-Laurent de Lavoisier: “Na Natureza, nada se cria, nada se perde, tudo se transforma.” No contexto da informação, transforma-se, sim — mas nunca desaparece.

3.2.4. *A informação tem de ser causal*

Por fim, a informação tem (4) de ser causal — ou seja, tem de possuir uma dinâmica determinística. Enquanto codificação ontológica, a informação constitui um princípio fundamental que *fisicaliza* o Universo por via da implementação de leis causais (ou físicas). A SUI, através dos seus códigos instrutórios, não só dá origem ao Universo como o constitui na sua totalidade. Possui, assim, uma natureza intrinsecamente causal: é a causa do Universo e, simultaneamente, a sua própria manifestação. As suas leis ontológicas expressam-se, para nós, sob a forma das leis físicas.

Não nos parece existir nada de ontológico no Universo que impeça uma leitura determinista do próprio Universo — nem mesmo a física quântica. A interpretação de Copenhaga (Heisenberg, 1958; Bohr, 1958), como analisado no capítulo 2, proclama uma visão indeterminista do Universo mas essa visão não só não nos parece possível

(pela incompatibilidade que nos parece existir entre aleatoriedade, que é exclusivamente uma percepção humana, e ontologia que é exclusivamente o que existe), como não passa de uma interpretação epistemológica da realidade, não ontológica. A interpretação da física quântica de David Bohm (1952) sugere uma interpretação muito mais realista e coerente (Smolin, 2019), ao demonstrar que há uma realidade subjacente, que está oculta da observação comum, mas que ainda mantém a natureza determinista. Ao adotarmos esta perspectiva, podemos, aliás, argumentar que a informação é, na verdade, uma descrição fundamental da realidade, que segue leis deterministas subjacentes, ainda que as nossas observações sejam limitadas e sujeitas a incertezas epistemológicas, como defendido pelo modelo de Copenhaga (Heisenberg, 1958; Bohr, 1958).

Se a informação fosse indeterminística, isso implicaria uma quebra na estabilidade estrutural e no encadeamento causal do sistema, comprometendo as condições necessárias à existência de sistemas complexos, como moléculas, células ou organismos. A nossa proposta — segundo a qual a informação, enquanto unidade fundamental, possui uma natureza causal e determinística — sugere que a organização estrutural e a causalidade intrínseca do sistema são condições essenciais para a formação e manutenção de sistemas emergentes complexos, capazes de sustentar a vida. Em termos simples, sugerimos que é a natureza determinística da informação que confere ao Universo a estabilidade necessária para que este possa desenvolver-se e evoluir, permitindo o surgimento de fenómenos cada vez mais complexos mas estáveis — nomeadamente sistemas biológicos como os cérebros, que possibilitam o aparecimento de mentes conscientes (Bohm, 1980). A estabilidade das leis da Natureza decorre, portanto, da estabilidade das informações que elas transmitem e processam (Lloyd, 2007; Davies, 2019; Tegmark, 2014; Bohm, 1980; Smolin, 2019; Hooft, 2016 e 2017).

Em suma, a informação tem de ser determinística porque apenas segundo uma lógica causal um sistema pode ter, por um lado, a capacidade de causar sistemas emergentes e, por outro, funcionar com a estabilidade que precisa para albergar complexidade e vida.

3.2.5. *SUI e causalidade emergente: o determinismo como motor da complexidade*

Sugerimos, portanto, que a SUI, podendo ser conservada, processada\transformada, podendo ser finita e determinística, possibilita a criação do espaço-tempo, dos campos de força, da matéria-energia, das leis e das constantes da Natureza, bem como a formação de fenómenos biológicos emergentes. Em última instância, a SUI permite que informação possa representar informação e também re-representar informação, que é o que sugerimos que as mentes conscientes fazem, como veremos mais à frente.

A SUI é uma estrutura ontológica que permite o surgimento de fenómenos emergentes regulares através de computações que, interagindo entre si — ou, dito de outra forma, processando bits de informação capazes de gerar novos sistemas informacionais — possuem a capacidade de criar fenómenos emergentes cada vez mais complexos. Por conseguinte, podemos ver uma partícula como um sistema de informação que, ao interagir com outros sistemas (outras partículas), permite a emergência de sistemas mais complexos, como os átomos. Estes, por sua vez, agem segundo o mesmo princípio, atraindo-se mutuamente, o que possibilita o surgimento de moléculas e, posteriormente, de células, organismos e sistemas nervosos. Todo este processo exige leis regulares e, por isso, totalmente determinísticas, o que torna possível o surgimento de estabilidade suficiente para que os sistemas existam de forma relativamente permanente. O Universo não é um relógio nem um computador, mas é um sistema extraordinariamente ordenado e regular. É, aliás, essa regularidade sistémica que permite a estabilidade que podemos observar, tanto no comportamento das moléculas e das células como no movimento dos astros a orbitar outros astros.

A nossa perspetiva está em consonância com investigações recentes que sublinham a necessidade de processos determinísticos na origem e coordenação de fenómenos altamente complexos. Estes estudos sugerem a existência de uma estrutura ontológica subjacente que regula e orienta tais processos, conferindo-lhes coerência, estabilidade e direção. No domínio da química quântica, por exemplo, tem sido demonstrado que as interações e reações moleculares não ocorrem ao acaso, mas são regidas por princípios informacionais e computacionais determinísticos, e não apenas por mecanismos físico-químicos clássicos (Aspuru-Guzik e Walther, 2012; Kassal et al., 2008). De forma semelhante, os sistemas biológicos revelam características informacionais profundas, como evidenciam os trabalhos de Lambert et al. (2013), Patel (2001) e Walker e

Davies (2013). Estes contributos reforçam a ideia de que tanto os sistemas moleculares como os sistemas vivos operam com base numa arquitetura informacional causal e determinística, na qual elementos computacionais sustentam a emergência de estruturas organizadas e estáveis.

No nosso modelo, essas estruturas são manifestações locais da SUI — uma ontologia sistémica de base informacional que opera através de processos deterministas de computação de informação. Cada novo nível de complexidade — das moléculas às células, dos organismos às consciências — representa uma iteração sucessiva dessa computação causal, em que a informação não é apenas transmitida, mas processada, conservada e reorganizada de forma a gerar novos sistemas com propriedades emergentes. Assim, a estabilidade da matéria, a organização da vida e a emergência da mente não resultam de forças externas ou de acasos, mas da regularidade sistémica de uma substância fundamental: a informação.

A causalidade é, portanto, indispensável para a existência de um Universo estável. Sem essa causalidade, as partículas — mesmo que pudessem conservar e processar informação (o que, num Universo não determinístico, nos parece inviável, dado que essas partículas são sistemas de informação que requerem estabilidade e regularidade para poderem existir como tal) — não teriam a capacidade de gerar sistemas emergentes mais complexos. O Universo depende, de forma crítica, de processos regulares, como a estrutura espaço-temporal e a dinâmica matéria-energia. Por um lado, o espaço-tempo tem de constituir uma estrutura inteiramente estável, de modo a sustentar-se como um dos fundamentos do Universo que conhecemos. Por outro, as partículas que integram a matéria-energia precisam, igualmente, de exibir uma organização regular e consistente, para que possam dar origem a todos os fenómenos emergentes sem que estes entrem em colapso. Por outras palavras, o Universo não só necessita, como manifesta de uma forma clara, uma ontologia regular e estável que, a nosso ver, está assente em princípios determinísticos. Assim, o determinismo do Universo é o que confere leis estáveis, regulares e constantes do sistema, possibilitando uma solidez necessária para o surgimento de complexidade. Dito de outro modo, o determinismo possibilita o surgimento de leis estáveis que, por sua vez, possibilita o surgimento de fenómenos emergentes mais complexos.

Esta leitura permite, assim, compreender por que razão a causalidade não é apenas um princípio conveniente, mas uma condição ontológica essencial. O contrário, isto é, sugerir um Universo indeterminista seria, a nosso ver, propor um modelo de Universo incompatível, não só com a complexidade e com vida, mas também com a lógica. Como explicar um sistema (pelo menos um estável, capaz de albergar complexidade e vida) *acausal*? Mais: como é que, como referido no capítulo 2, o Universo pode *causar* alguma coisa *acausal* (?) – o paradoxo é óbvio. Seria possível existir complexidade e vida num sistema *acausal*, isto é, que *acausa* causas? Um sistema *acausal* é um sistema que nada causa. Se nada causa, não pode existir. O determinismo é, por isso, ubíquo para a criação de constantes que permitem criar e causar sistemas emergentes, complexidade e, porventura, vida (Serrado, 2024 e 2025).

Em contraste, a aleatoriedade — para além de ser um conceito essencialmente epistémico — implicaria que os mecanismos subjacentes seriam incapazes de gerar a estabilidade necessária num sistema que depende criticamente de regularidade para surgimento de fenómenos emergentes mais complexos. A aleatoriedade ontológica é o oposto de estabilidade que qualquer sistema necessita para criar algo de permanente e organizado. De facto, o físico Seth Lloyd calculou que:

“seriam necessários um bilião de biliões de macacos, cada um a tipografar 10 caracteres por segundo, durante cada um dos biliões de biliões de segundos desde o começo do universo, para que um deles produzisse «hamlect, act i, scene i” (Lloyd, in Ribeiro, 2013, p. 25).

Por conseguinte, o indeterminismo não permite a criação de nada estável e organizado — nada capaz de gerar sistemas com condições para a emergência de complexidade e vida consciente. O aparente indeterminismo da física quântica tem, como vimos, conduzido às mais bizarras afirmações, nomeadamente a de que o Universo não é determinístico apenas porque o ser humano não consegue prever certos eventos quânticos (Kaku, 2021; Kane, 2005 e 2007; Balaguer, 2014). A conclusão retirada por alguns é que o Universo seria, por isso, regido por mecanismos aleatórios. No entanto, como discutido no capítulo 2, a (in)capacidade humana para prever o comportamento das partículas não tem qualquer relação com o facto de esses eventos serem ou não causados (Hooft, 2016); do mesmo modo, a aleatoriedade — sendo um conceito epistemológico e uma ferramenta de previsão — não pode ser um argumento válido em

discussões sobre a ontologia do Universo. E, a nosso ver, só podem ser causados. A alternativa seria afirmar que são *acausados* — ou que um sistema físico causa eventos *acausados* (Ginet, 1962) — o que constituiria um absurdo lógico e ontológico (Serrado, 2024 e 2025; Hooft, 2016).

A aleatoriedade epistêmica é, no entanto, evidente — o que não deve ser confundido com indeterminismo ontológico. Para seres com aparatos cognitivos limitados (Kant, 1781/2008), o determinismo parecerá inevitavelmente aleatório, sobretudo a nível quântico (Serrado, 2024 e 2025). Mas essa aparência de imprevisibilidade não compromete a estrutura causal do Universo. Um sistema complexo, estável e com capacidade para sustentar vida necessita, necessariamente, de leis que determinem e causem essa estabilidade (Hooft, 2016), como temos sublinhado ao longo deste trabalho. Como refere Miguel Ribeiro, “A ocorrência espontânea de um universo aleatório que gerou vida é estatisticamente insustentável” (Ribeiro, 2013, p. 16).

Em suma, é a natureza determinística da SUI que torna possível não apenas a não-aleatoriedade do Universo, mas também a criação de um sistema causal que, precisamente por gerar eventos segundo leis estáveis e constantes, deu origem a fenómenos emergentes que precisam de ser — e são — extremamente estáveis, como o espaço-tempo, a matéria-energia e outras constantes fundamentais. Essa estrutura causal é o que permite: 1) a estabilidade do Universo, 2) o surgimento progressivo da complexidade e 3) o aparecimento da vida e de mentes conscientes. Como veremos nos capítulos seguintes, a SUI é uma estrutura determinista que opera com base numa ontologia estrutural que: 1) contém todas as codificações inerentes ao sistema, incluindo as suas potencialidades; e 2) coordena esse sistema através de um desdobramento de informação que, por meio de computações, permite o surgimento de fenómenos emergentes.

3.3. Os Três aspetos da SUI

"O mais incompreensível do Universo é ele ser compreensível."

Albert Einstein

A SUI é uma entidade fluida e onnipresente que permeia todo o Universo. Sendo una e indivisível, pode ser concebida como um tecido fundamental que estrutura a existência e a realidade. No entanto, para efeitos de análise, podemos distingui-la em três aspetos principais: 1) *informação ontológica estrutural*, 2) *informação ontológica computacional* (ou *funcional*); 3) *informação representacional*.

A informação ontológica estrutural corresponde à arquitetura fundamental da realidade, sendo a base subjacente que define as leis e os princípios do Universo. A informação ontológica computacional refere-se ao seu aspeto funcional, isto é, ao modo como os sistemas processam e organizam a informação segundo essas leis estruturais. Por fim, a informação representacional corresponde à dimensão fenomenológica da informação, manifestando-se nos sistemas capazes de gerar experiência consciente.

3.3.1. A informação ontológica

No modelo que propomos, a informação ontológica divide-se em duas categorias: (A) *informação ontológica estrutural* e (B) *informação ontológica computacional*.

Defendemos que a realidade fundamental assenta numa informação subjacente e não diretamente acessível – a A) informação ontológica estrutural –, da qual a B) informação ontológica computacional⁶⁷ emerge através de processos funcionais. Esta ontologia computacional resulta do desdobramento de uma ontologia estrutural anterior ao Big Bang. Esta estrutura fundamental representa uma realidade distinta, potencialmente regida por leis próprias, sem que isso comprometa a ideia central de que a informação ontológica constitui a base estrutural e subjacente de toda a nossa realidade. Assim, sugerimos que o Big Bang não foi o início absoluto de tudo, mas antes uma transição, na qual a informação ontológica, até então contida numa estrutura anterior, se

⁶⁷ Em relação à qual temos acesso a uma dimensão observável, descrita na física, na química e na ciência em geral.

desdobrou pela primeira vez num sistema computacional – o Universo. A informação ontológica estrutural encontra-se, enquanto tal, num estado potencial, ainda não processado nem manifestado. Como veremos, apenas na sua fase computacional a informação passou a ser computada e a dar origem a sistemas emergentes.

A informação ontológica estrutural

A informação ontológica estrutural refere-se à base subjacente da realidade que define a organização e as relações entre as entidades que compõem o Universo. Não depende da ação ou da *computacionalidade* das entidades mas sim da sua configuração, das leis e padrões que governam essas entidades e como elas interagem. Esta informação pode ser comparada a um tipo de código primordial ou estrutura em que as propriedades fundamentais do espaço-tempo, das partículas subatômicas e da matéria-energia, são inscritas. Nesse sentido, a informação estrutural é mais estática, representando a arquitetura da realidade, como um conjunto de regras ou princípios que não mudam com o tempo, mas que definem as possibilidades e potencialidades de comportamento de tudo no Universo.

O Big Bang é, no nosso modelo, reinterpretado como um evento de transição, no qual a informação ontológica deixa de estar num estado estrutural e potencial para se *desdobrar* num sistema computacional (Bohm, 1980) - o momento em que o Universo começou a computar (Seth, 2007). Esta transição marca o surgimento do espaço-tempo como um sistema físico que processa a *informação ontológica computacional*, enquanto as leis físicas se materializam como manifestações computacionais ajustadas à natureza do sistema emergente. No nosso modelo, o Big Bang pode também ser compreendido como um algoritmo – uma singularidade informacional condensada – que contém todas as instruções necessárias para a construção do Universo. A *informação ontológica computacional* refere-se, portanto, à ação ou ao processo que resulta da interação das entidades no sistema. Ela está relacionada com a forma como os elementos da realidade se organizam e interagem ao longo do tempo, gerando eventos, dinâmicas e sistemas emergentes mais complexos (energia, partículas, átomos, moléculas, etc.). Enquanto a *informação estrutural* pode ser entendida como a arquitetura base do sistema, a *informação computacional* pode ser entendida como os dados ou instruções que guiam as ações de sistemas complexos. Esta torna-se ativa e dinâmica, sendo a informação que

resulta na função e nos efeitos da realidade. Em termos físicos, seria como a descrição de como as partículas se movem, interagem ou evoluem, refletindo a operação das leis físicas em cada momento.

Importante referir que a dimensão estrutural da informação não rompe a dimensão computacional da informação. Pelo contrário, reforça a ideia de que toda a realidade, mesmo em estados distintos, é sustentada por um tecido unificador: a SUI. Desta forma, enquanto a informação ontológica estrutural permanece inacessível, é através da sua manifestação computacional no Universo que podemos estudar os seus efeitos indiretos — como as leis físicas (a dimensão observável da informação computacional) e os fenómenos naturais por elas descritos — como veremos mais à frente.

A nossa perspetiva de uma ontologia informacional não é meramente especulativa. Um estudo recente (Liu et al., 2025) demonstrou experimentalmente a existência de uma contextualidade quântica de tipo GHZ⁶⁸ em 37 dimensões, através de um processador fotónico operando no domínio temporal. Este resultado não implica a existência de 37 dimensões espaciais, mas revela que o Universo pode possuir uma estrutura de base suficientemente rica e causalmente organizada para permitir 37 graus de liberdade funcionais num único sistema quântico.

Este sistema quântico não é aleatório nem probabilístico na sua essência: é determinista. Os autores demonstram que os fenómenos observados resultam de correlações perfeitas e determinísticas geradas por contextualidade GHZ, um recurso que permite ultrapassar os limites computacionais de qualquer circuito clássico com profundidade fixa. Tal como defendemos na nossa ontologia, trata-se da manifestação de informação estrutural, ou seja, dos algoritmos causais que regem o real, permitindo que um sistema físico se comporte de acordo com eixos de variação altamente definidos. Durante a fase de sobreposição, o sistema opera segundo informação computacional, ou seja, dentro de um espaço de estados funcionais possíveis. No momento da medição, ocorre uma atualização de estado, transformando essa informação computacional em informação representacional, sem qualquer quebra de determinismo ou intervenção externa. O chamado "colapso" é, na verdade, uma transição interna de uma estrutura

⁶⁸ Estado de Greenberger–Horne–Zeilinger.

causal consistente. As dimensões envolvidas neste estudo devem ser entendidas como eixos de codificação informacional, e não como espaços físicos adicionais. O Universo, longe de ser um palco de dimensões ocultas, revela-se como uma rede coerente de relações informacionais deterministas, cuja profundidade excede os limites da nossa intuição perceptiva — mas que pode ser medida, prevista e manipulada dentro de uma estrutura causal rigorosa (Liu et al., 2025).

Um outro estudo (Garza et al., 2024), particularmente ilustrativo da nossa proposta ontológica, encontra-se na recente investigação conduzida por Samantha Garza e colegas da Universidade de Washington, que analisou o percurso cósmico dos átomos de carbono presentes nos corpos humanos. O estudo revelou que grande parte destes átomos não se formou diretamente nas regiões interiores da Via Láctea, mas sim em zonas mais periféricas do universo, tendo sido posteriormente expelidos por fenómenos como supernovas e transportados para o meio circungaláctico. Estes elementos viajaram por centenas de milhares de anos-luz antes de voltarem a ser incorporados na formação de novas estrelas, sistemas planetários e, eventualmente, organismos vivos. A investigação conclui que muitos dos átomos que compõem os nossos corpos estiveram, durante períodos significativos, fora da galáxia onde hoje existimos.

Esta descoberta, para além da sua dimensão poética, reforça de modo empírico a centralidade da informação estrutural na organização do real. O ciclo de reciclagem cósmica do carbono não decorre de processos aleatórios, mas de leis deterministas e regulares que regem a nucleossíntese estelar, a dinâmica gravitacional, os ventos galácticos e a reintegração da matéria no interior das galáxias. Trata-se, pois, da manifestação de uma estrutura informacional ontológica — uma rede de algoritmos causais que regula a transformação e persistência da matéria ao longo do tempo cósmico. Os átomos não seguem um percurso errático, mas sim uma trajetória compatível com princípios físicos que permitem a previsibilidade dos ciclos astrofísicos.

Adicionalmente, este fenómeno exemplifica a articulação entre os diferentes níveis da informação definidos nesta ontologia: a informação estrutural que codifica as leis físicas; a informação computacional que regula os estados funcionais da matéria ao longo do tempo (como a configuração dos átomos em nuvens interestelares ou corpos planetários); e a informação representacional, quando essas configurações assumem formas observáveis, como corpos vivos conscientes. Assim, a própria constituição

material do ser humano revela-se como o resultado de uma computação profundamente enraizada em processos causais que se desenrolam segundo os algoritmos fundamentais do universo (Garza et al., 2024).

A teoria que apresentamos é totalmente compatível com o fisicalismo, pois não propõe uma separação ontológica entre a informação ontológica e o Universo físico. Em vez disso, ambas são expressões de um mesmo tecido ontológico: a SUI. A informação ontológica é a base fundamental da realidade, e o Universo — que podemos observar — é a sua manifestação física. Esta visão evita qualquer dualismo ao 1) reconhecer que a estrutura pré-Big Bang e o Universo físico fazem parte de uma única SUI; 2) interpretar a inacessibilidade epistemológica como uma limitação do sistema cognitivo humano, sem implicar uma divisão essencial na realidade; 3) tratar a transição do estado estrutural para o estado computacional como um processo dentro do mesmo contexto informacional-fisicalista. Assim, o Universo e a sua origem permanecem inteiramente ancorados no fisicalismo informacional, onde a física e informação são expressões complementares de uma única ontologia.

A informação ontológica computacional

No âmbito da nossa teoria, se a (1) informação ontológica estrutural corresponde ao conjunto de leis fundamentais que sustentam tanto as leis como o potencial da SUI (ou Universo), então, a (2) informação computacional representa a capacidade intrínseca da SUI — ou do próprio Universo — de processar e transformar informação, isto é, da SUI evoluir através de processos emergentes complexos.

Este processamento funcional é o mecanismo subjacente que permite o surgimento de sistemas emergentes e dá origem às dinâmicas físicas observáveis no universo. Diferente da simples codificação ou estrutura arquitetónica da informação estrutural, a informação computacional é o motor oculto que organiza e transforma a informação permitindo que esta se manifeste de forma concreta na realidade física. A informação computacional faz a transição entre o pré-Big Bang e o Big Bang em si, já que corresponde ao início de todos os processos e sistemas emergentes que fazem parte do nosso Universo, como o espaço-tempo e a matéria-energia. Neste modelo, o Big Bang não é um evento passado e fechado, mas um processo contínuo de computação informacional que ainda está a ocorrer. O Universo continua a expandir-se e a reorganizar-

se através de um contínuo informacional no qual novos estados se desdobram a partir da estrutura ontológica inicial. A informação ontológica, nesse sentido, não só iniciou os eventos cosmológicos, mas continua a modulá-los em tempo real, o que significa que a própria gênese do cosmos permanece em curso.

A informação computacional deve ser entendida como o processo ativo da informação ontológica, o qual segue regras algorítmicas intrínsecas. Este processamento não é apenas uma característica dos sistemas físicos locais, mas uma propriedade universal da realidade. Ele é responsável pela emergência das leis físicas, que não são impostas externamente, mas emergem como manifestações intrínsecas do próprio processamento da informação. Assim, a computação desempenha um papel essencial na organização da realidade, sendo o princípio operativo que estrutura o comportamento do universo.

O espaço-tempo e as estruturas materiais, como a matéria e a energia, podem ser vistos como consequências diretas desse processamento computacional – um *desdobramento* da informação estrutural em informação computacional. No entanto, a sua verdadeira natureza é “holográfica” (Hooft, 2000; Susskind, 1993), no sentido de que a realidade quadridimensional que percebemos é uma projeção de uma informação mais fundamental que reside na base ontológica e estrutural do Universo. Este caráter holográfico implica que toda a realidade emerge da organização e do processamento da informação computacional. Neste modelo, a computação não é uma propriedade restrita a sistemas tecnológicos ou biológicos, mas uma característica intrínseca da informação ontológica. Cada interação, desde as partículas fundamentais até os sistemas mais complexos, pode ser interpretada como uma computação realizada pela própria informação ontológica. Este processo não só organiza e transforma a informação, mas também gera as leis físicas e as dinâmicas observáveis que estruturam o universo.

A informação computacional desempenha, portanto, um duplo papel: ela é *conservada* no sistema, mantendo a continuidade da informação e também é *transformada* por meio do processamento, permitindo a emergência de propriedades e comportamentos. Este processamento *transforma* a informação ontológica em manifestações físicas, estruturando padrões, dinâmicas e leis que governam o comportamento do universo. Assim, a computação é a ponte entre a informação ontológica e a realidade física, sendo responsável por dar forma às estruturas e padrões que percebemos.

Importante sublinhar, a informação computacional é a ontologia da física, ou seja, a física (a dimensão mensurável e observável) é a epistemologia da informação computacional. Neste sentido, ao estudarmos o Universo, conseguimos aceder às manifestações da informação computacional que podem ser captadas pelo nosso aparato cognitivo. Existe, inevitavelmente, um véu sobre o qual não podemos aceder – uma *fundamentalidade* que nos escapa, que nos é vedada, mas que, ao mesmo tempo, é tão ubíqua, tão próxima à nossa natureza, como a qualquer outra manifestação do Universo.

Matéria, energia e informação: a ontologia informacional da realidade

A visão tradicional da física encara a energia como uma quantidade conservada, passível de ser transferida entre sistemas, e a matéria como a substância que constitui os objetos do universo (Newton, 1687/1934; Helmholtz, 1847/1977). No entanto, uma análise mais profunda revela que a distinção entre matéria e energia é, na verdade, uma ilusão conceptual. A equivalência massa-energia, expressa na equação einsteiniana $E=mc^2$ (Einstein, 1905), demonstra de forma clara que a matéria nada mais é do que uma forma condensada de energia. Isto levanta uma questão fundamental: o que define, então, a estrutura da matéria e da energia?

A nossa proposta responde a essa questão. No nosso modelo, a informação não é um mero registo passivo dos estados físicos, mas sim a substância fundamental da realidade, a partir da qual emergem tanto a matéria como a energia. Assim, a relação entre informação, matéria e energia não é apenas funcional, mas ontológica. A energia não deve ser vista apenas como uma quantidade física associada ao movimento e às interações, mas como a manifestação computacional da informação em ação. Do mesmo modo, a matéria pode ser interpretada como um estado estabilizado dessa informação computacional, mantendo uma estrutura persistente no espaço-tempo.

A energia é, por conseguinte, a expressão dinâmica da informação computacional, manifestando-se através das interações e transformações dos sistemas físicos. A matéria, por sua vez, corresponde a um estado estabilizado dessa mesma informação. Deste modo, matéria e energia não são substâncias independentes, como sustentava a física clássica (Newton, 1687/1934; Helmholtz, 1847/1977), mas manifestações distintas de uma única realidade informacional.

3.3.2. A hipótese do desdobramento informacional

Sugerimos que a informação computacional desempenha um papel fundamental na emergência de sistemas complexos. Não apenas o espaço-tempo e a matéria-energia, mas também os campos de força e todas as partículas que constituem a realidade física emergem de uma estrutura informacional subjacente, a que designamos por informação ontológica. Essa informação ontológica pode ser entendida como composta por dois aspectos complementares, conforme já referimos: *a informação ontológica estrutural*, que fornece a base ontológica necessária para a existência do Universo, e *a informação ontológica computacional*, que impulsiona a formação de sistemas emergentes orientando-se pelas instruções contidas na informação estrutural.

Podemos pensar na informação estrutural como o código-fonte da realidade, onde estão contidas todas as regras fundamentais do Universo, enquanto a informação computacional é a execução desse código em tempo real, que permite a organização do tempo-espaço, da matéria e da energia de acordo com essas regras. Podemos, também, compreender este processo como um desdobramento informacional, semelhante à dinâmica entre a ordem implicada e a ordem explicada proposta por Bohm (1980). Sugerimos que as partículas que permitem a emergência dos sistemas não devem ser entendidas como surgindo *ex nihilo*, mas como padrões ou estados que se organizam dentro de uma estrutura determinística – isto é, causal e composta por leis intrínsecas que possibilitam processamento do espaço-tempo e da matéria-energia. Assim, as partículas fundamentais podem ser vistas como "blocos de construção" (Rovelli, 2018; Smolin, 2019) do Universo que se configuram segundo leis computacionais intrínsecas, formando átomos, moléculas e, por extensão, todos os sistemas complexos, biológicos e não biológicos. Este desdobramento ocorre a partir da informação estrutural, que contém todas as potencialidades fundamentais da realidade, as quais são continuamente manifestadas, em tempo real, através da informação computacional.

A emergência do tempo-espaço e da matéria-energia podem ser entendidas, assim, como o resultado de um desdobramento da informação estrutural para a informação computacional. A informação estrutural define as condições necessárias para que o espaço-tempo, a matéria-energia e as partículas existam, enquanto a informação computacional governa as dinâmicas que permitem a interação e evolução desses elementos. Dessa forma, o Universo é um sistema emergente no qual o espaço-tempo e a

matéria-energia são manifestações de um processo contínuo de transformação e de organização informacional. Dito de outra forma, a informação estrutural permite um desdobramento fundamental, do qual emergem as bases do espaço-tempo e das estruturas primordiais do Universo, enquanto a informação computacional governa um desdobramento dinâmico, no qual as leis fundamentais se manifestam em processos que impulsionam a formação do tempo-espaço e da matéria-energia.

Este processo pode ser interpretado como a passagem de um *estado potencial* para um estado *funcional*. A informação estrutural, enquanto código-fonte, contém todas as possibilidades da realidade, mas permanece num estado potencial até que a informação computacional atue como mecanismo de execução, traduzindo essas possibilidades em leis físicas e sistemas emergentes. Assim, a informação computacional é a forma como a informação se exprime funcionalmente, e a realidade observável (e, por isso, física) é o resultado visível dessa atualização contínua. No nosso modelo, o tempo pode ser compreendido como o desenrolar de um “programa” que, ao computar informação, gera sucessivos estados diferenciados — criando, assim, a percepção de fluxo temporal. Em última análise, o futuro do nosso Universo está inscrito nessa matriz fundamental, em potência — como um processador que, ao longo do tempo, vai computando e atualizando a informação inscrita na sua estrutura.

Assim, a matéria e a energia — porventura o próprio tempo — podem ser, portanto, entendidas como manifestações de um processo contínuo e dinâmico de desdobramento informacional. Este processo ocorre no interior de uma estrutura ontológica fundamental que chamamos de informação estrutural, onde a informação computacional atua como o mecanismo que permite a organização e evolução dessa base para produzir os elementos observáveis da realidade. Ou seja, a informação estrutural contém as codificações que a informação computacional utiliza para gerar processos dinâmicos e emergentes.

O papel da informação estrutural é central neste contexto. É ela que fornece o substrato essencial para a existência do Universo. É a “base estática” que define: 1) as regras fundamentais - as leis que emergem como expressões das propriedades dessa estrutura subjacente; 2) as condições iniciais - o estado primordial do Universo, o “estado fundamental” antes do Big Bang, que pode ser descrito como um campo de

possibilidades informacionais, onde todos os estados potenciais coexistem de forma não diferenciada; e 3) a continuidade do espaço-tempo - a estrutura informacional que define as propriedades básicas do espaço-tempo, como a sua dimensionalidade e topologia. Sem esta informação estrutural, o Universo não teria um "suporte" ontológico para existir, e o tempo-espaço, bem como a matéria-energia seriam impossíveis.

O papel da informação computacional é distinto. A informação computacional atua como o motor dinâmico que transforma o *potencial* contido na informação estrutural em sistemas emergentes. Ela governa: 1) a dinâmica do desdobramento; 2) as interações e transformações que regem o comportamento das partículas fundamentais e dos campos de força; e 3) a emergência de padrões (estruturas como partículas, átomos, moléculas, células, organismos, planetas e galáxias emergem como padrões organizados da informação computacional). Todos estes processos são regidos por algoritmos ontológicos que determinam as interações e a evolução do Universo. Podemos, portanto, explicar o desdobramento entre informação estrutural e informação computacional como o desdobramento entre um estado *potencial* (informação estrutural) para um estado *funcional* (informação computacional), no qual as leis fundamentais se manifestam e organizam em sistemas emergentes.

A nossa proposta também se alinha com uma nova perspectiva sobre a física quântica, sugerindo que esta é uma expressão epistemológica da verdadeira ontologia do Universo. As aparentes indeterminações e fenômenos não-locais observados na física quântica são, a nosso ver, "sintomas" das limitações do nosso entendimento da informação estrutural. A mecânica quântica descreve a forma como a informação se processa e se manifesta, mas não revela diretamente a ontologia que governa esse processo. No nosso modelo, a informação estrutural representa o substrato fundamental da realidade e a informação computacional define as dinâmicas que regulam a interação das partículas e campos de força. A física quântica, nesse sentido, representa a interface entre esses dois domínios, descrevendo as regras que emergem da execução dos códigos subjacentes.

Sugerimos, portanto, um desdobramento de informação que permite a emergência de sistemas mais complexos. Este desdobramento informacional pode ser entendido em diferentes etapas. 1) A *fase inicial* (potencial puro) é o antes do Big Bang. A informação estrutural representa um campo de possibilidades puras. Este campo

contém todas as leis e constantes fundamentais em estado latente, mas sem manifestação computacional e física. 2) A *fase de diferenciação* corresponde ao Big Bang. Nesta fase a informação computacional entra em ação, iniciando a *diferenciação* desse potencial (um algoritmo que condensa em si toda a informação) em estados distintos, como partículas fundamentais e campos de força. Este momento marca a emergência do espaço-tempo como uma estrutura dinâmica que permite a interação e a organização de matéria e energia. Na 3) *fase de organização* observamos a *interação* e a *evolução*. Nesta, partículas fundamentais, como quarks e léptons, emergem como padrões organizados da informação computacional. As forças fundamentais (gravidade, eletromagnetismo, forças nucleares forte e fraca) atuam como mecanismos de interação definidos por algoritmos informacionais. Estas partículas começam a interagir, formando átomos, moléculas e, eventualmente, estruturas mais complexas como estrelas e planetas. Na 4) *fase contínua* observamos uma complexidade crescente. O desdobramento continua a atuar, permitindo a emergência de sistemas mais complexos, como organismos vivos e, eventualmente, a consciência. Em cada etapa, a informação computacional mantém o Universo em constante evolução, criando novos níveis de organização.

Em suma, este processo de *desdobramento informacional* da SUI oferece uma explicação ontológica para a emergência do Universo como o conhecemos. 1) Do *potencial* ao *real*, a informação estrutural fornece a base de possibilidades, enquanto a informação computacional atualiza essas possibilidades em realidade. 2) Da *simplicidade* para a *complexidade*, o desdobramento explica como a simplicidade inicial do estado informacional evolui para sistemas altamente complexos. Por fim, 3) a *unidade* e a *diversidade* explica que, apesar da diversidade de formas e interações na realidade física, todas derivam de uma SUI.

Em *Wholeness and the Implicate Order* (1980), David Bohm introduziu os conceitos de *enfoldment* (enrolamento) e *unfoldment* (desenrolamento) para descrever a estrutura profunda da realidade e o modo como os fenômenos emergem de uma ordem subjacente. Para Bohm, o universo não é composto por entidades separadas e independentes, mas sim por um sistema holístico, onde a realidade observável resulta do desdobramento de uma camada mais profunda da existência, que ele chamou de ordem implicada.

Nesta ordem, toda a informação sobre a totalidade do universo está embutida em cada parte do sistema, mas de forma latente, inacessível à percepção direta. A ordem implicada deve ser entendida como um nível fundamental onde todas as possibilidades coexistem de maneira interligada, sem se manifestarem explicitamente. O *unfoldment* da ordem implicada ocorre quando partes dessa informação oculta se tornam acessíveis e mensuráveis, manifestando-se como a ordem explicada – a realidade física e fenoménica. Esse processo transforma a matriz informacional implícita em estruturas diferenciadas e localizáveis no espaço e no tempo. Contudo, Bohm enfatiza que esse movimento não é unidirecional. Assim como a ordem implicada pode ser desdobrada na ordem explicada, também pode ocorrer o inverso: elementos da realidade explicada podem ser reintegrados na ordem implicada através de um novo processo de *enfoldment*. Dessa forma, a relação entre enrolamento e desenrolamento bohmiano propõe que o universo é um sistema dinâmico e cíclico, onde a informação oscila entre estados latentes e manifestados.

O modelo de Bohm encontra paralelos com a nossa distinção entre informação estrutural e computacional, bem como com o nosso modelo de desdobramento informacional. No nosso enquadramento, o desdobramento informacional pode ser entendido como um processo de *unfoldment*, onde a informação estrutural se transforma em informação computacional, permitindo a emergência da realidade fenoménica. No entanto, assim como no modelo de Bohm, esse processo não é irreversível: a informação computacional pode ser integrada novamente na informação estrutural, criando um ciclo dinâmico de transformação e evolução da realidade. Essa interação contínua entre desdobramento e reintegração informacional garante que a informação não se perde, mas reorganiza-se em novas formas, preservando a interconectividade da realidade a diferentes escalas.

Para ilustrar a dinâmica da ordem implicada, Bohm recorreu à holografia, propondo o conceito de holomovimento. Num holograma, cada parte contém a informação do todo, o que significa que a totalidade está de alguma forma "implicada" em cada fragmento. De forma análoga, no modelo de Bohm, cada elemento do Universo contém potencialmente a totalidade da informação da realidade, o que implica que o cosmos não é composto por partes separadas, mas sim por uma totalidade indivisível, onde a separação é apenas uma construção aparente. O holomovimento de Bohm representa,

assim, um fluxo contínuo e indivisível que dá origem tanto à ordem implicada quanto à explicada, criando um ciclo constante de desdobramento e reabsorção. O desdobramento ocorre quando a ordem implicada se torna explicada, tornando-se acessível e mensurável. A reabsorção acontece quando a ordem explicada retorna à implicada, ocultando-se novamente no tecido profundo da realidade. Esse mecanismo demonstra como a ordem implicada dá origem à ordem explicada e como ambas estão em permanente interconexão.

Assim, o conceito de holomovimento oferece um quadro dinâmico para compreender a relação entre a estrutura fundamental da realidade e a sua manifestação fenoménica. No nosso modelo, tal como no de Bohm, o Universo não é apenas um conjunto de entidades físicas independentes, mas sim um sistema informacional recursivo, onde cada estado computacional se baseia num substrato estrutural profundo, que continuamente se reformula através do seu desdobramento e reintegração. Esta visão pode fornecer uma interpretação monista e integrada do cosmos, onde a informação, em diferentes estados de manifestação, é a substância ontológica última da realidade.

Ao dividir o pensamento de Bohm nestes dois processos fundamentais – o mecanismo de enrolamento/desenrolamento e o holomovimento – conseguimos perceber que a realidade não é estática, mas sim uma dinâmica contínua de fluxo informacional. No nosso modelo, essa perspetiva ajusta-se perfeitamente à distinção entre informação estrutural e computacional, bem como à ideia de que a realidade fenoménica é um reflexo parcial de um nível mais profundo de organização informacional. No fundo, tanto Bohm como o nosso modelo apontam para uma realidade unitária, onde aquilo que percebemos como "separado" é apenas um estado particular de um sistema interligado de informação e processamento universal.

O físico Gerard't Hooft, na senda de Bohm, considera que o comportamento estatístico da física quântica é (aparentemente) indeterminado, mas de tal forma coordenado e previsível que tem que existir uma ontologia determinística por trás desse comportamento aparentemente indeterminado. Hooft recupera a ideia de Einstein sobre a existência de variáveis ocultas (2016 e 2017; Smolin, 2019) para demonstrar que é impossível que a física quântica funcione tão bem se tiver, na realidade, um comportamento aleatório. Opinião idêntica tem físico Peter Verheyen, para quem todas as

computações quânticas são espantosamente tão acertadas que deve existir uma realidade concertada por trás disso, uma realidade que, no entanto, está escondida da nossa percepção direta (2021). Por outras palavras, a física quântica obedece a padrões comportamentais tão bem definidos e previsíveis que só um universo ontologicamente determinístico poderia coordenar e sustentar.

No seu livro *The Cellular Automaton Interpretation of Quantum Mechanics* (2016), Hooft explora uma explicação ontológica determinista da realidade. Propõe uma interpretação da mecânica quântica segundo a qual o Universo funciona como um imenso autómato celular, onde a realidade é uma estrutura emergente da informação computacional processada. Esta visão procura ultrapassar as abordagens probabilísticas da mecânica quântica, oferecendo uma explicação concreta sobre o funcionamento fundamental da realidade, baseada na teoria da informação e em processos computacionais deterministas. Em última instância, esta proposta pode ser entendida como uma tentativa de descrever uma realidade ontológica mais profunda, da qual a física quântica emerge como uma manifestação estatística (Hooft, 2017).

Esta perspetiva pode ser articulada com outra contribuição fundamental do próprio Hooft: o princípio holográfico, inicialmente formulado no seu artigo *The Holographic Principle* (2000). Este princípio sustenta que toda a informação contida num volume tridimensional de espaço pode estar codificada na sua superfície bidimensional — à semelhança de um holograma. Trata-se de uma ideia com profundas implicações ontológicas, uma vez que sugere que a estrutura do Universo observável pode ser uma projeção informacional de uma camada mais fundamental da realidade.

A articulação entre estas duas propostas permite uma leitura integradora. O modelo de autómato celular de Hooft (2016) fornece a dinâmica determinista da realidade — a forma como os estados evoluem através de regras computacionais locais —, enquanto o princípio holográfico (2000) fornece a geometria informacional que organiza e limita esses estados. Se o Universo é, na sua base, composto por informação codificada, e essa informação evolui de forma determinista, então tanto o conteúdo (a informação computacional, ou o holograma) como a forma (a informação estrutural, ou a grelha holográfica) são fundamentais para a estabilidade e emergência do real. Assim, estas duas ideias de Hooft convergem numa visão em que a realidade é, na sua essência, informacional e determinística. A física quântica, com todas as suas aparentes incertezas

e probabilidades, seria apenas uma expressão superficial de uma base mais profunda, governada por leis computacionais precisas e organizadas espacialmente segundo princípios holográficos.

Essa proposta ajudou a resolver o paradoxo da informação dos buracos negros, garantindo a coerência com a mecânica quântica (Maldacena, 1999). A demonstração de que a informação no interior de um buraco negro fica codificada na sua margem (Hawking e Mlodinow, 2011) – numa grelha bidimensional localizada no horizonte de eventos – sugere que sistemas tridimensionais, como os buracos negros, podem ser sustentados por informação bidimensional, que, por sua vez, projeta a realidade de forma tridimensional. Isto reforçou a ideia, já referida, de que a nossa realidade quadridimensional pode ser uma projeção de uma grelha informacional bidimensional.

Inicialmente recebida com algum ceticismo, o princípio holográfico é atualmente uma teoria extremamente válida entre a comunidade científica (Bousso, 2002; Carroll, 2019; Greene, 2003; Maldacena, 1999). Se estiver certo, o princípio holográfico explica porque a natureza que representamos a 4 dimensões é “fictícia” – uma ilusão projetada por uma informação que gera essa mesma ilusão. O Universo seria, portanto, holográfico – não no sentido literal de ser uma ilusão mas na medida em que a forma como percebemos as 4 dimensões é uma representação *quadridimensional* de uma ontologia bidimensional. Uma ontologia que nunca poderíamos aceder devido à sua natureza bidimensional, que não é compatível com a Natureza *quadridimensional* que vivemos⁶⁹.

Assim como o princípio holográfico sugere que vivemos numa dimensão projetada de uma grelha de informação bidimensional, o nosso modelo propõe que existem algoritmos ontológicos que sustentam o nosso Universo, algoritmos esses que são

⁶⁹ Os físicos Lawrence Krauss (2017) e Stephen Hawking (2011) defendem que o Universo pode ter surgido do vácuo. Sabemos que o vácuo – o vazio – não está nunca realmente vazio. Pelo contrário, está carregado de campos de força. Nesse espaço existem flutuações quânticas (elétrões, neutrinos, etc.) com algum nível de matéria-energia que, segundo aqueles cientistas, podem ter estado na origem do Universo. No entanto, importa questionar como é que matéria-energia que existe no nosso Universo poderia existir antes do nosso Universo. Sugiro que as flutuações quânticas – os campos de forças que preenchem todo o Universo, nomeadamente o vácuo ou o vazio - podem estar na origem de toda a matéria-energia, (como aliás é defendido por vários físicos, nomeadamente os acima referidos (Kraus, 2017; Hawking, 2011), como ainda Gerard 't Hooft (2016) e Seth Lloyd (2007), entre outros). Mas os campos de força não são a informação ontológica que sugiro. Alguma coisa causou os campos de força e alguma coisa coordena o seu comportamento e sugerimos que essa alguma coisa é a informação ontológica. O que sugiro é um tipo de informação “protoquântica” que, como referi, coordena e projeta a informação quântica que é processada no Universo

completamente incognoscíveis. Alguns desses algoritmos são estruturais (suportam o holograma) ao passo que outros são computacionais ou funcionais (são o holograma). Propomos que a informação ontológica estrutural é um algoritmo primordial que poderá nunca ser revelado totalmente devido às limitações dos nossos aparatos cognitivos para aceder a uma realidade que extravasa as nossas leis. Por ser um tipo de informação *hardwired*, que se situa fora projeção holográfica – o holograma não pode interagir com a grelha que o projeta –, não pode ser *computada* nem, subsequentemente, *representada* por essas partes⁷⁰.

Em *Our Mathematical Universe* (2014) o físico Max Tegmark, sugere uma teoria com afinidades com a nossa. Para Tegmark, toda a realidade é composta por estruturas matemáticas. Isto significa que o nosso mundo físico não pode só ser descrito matematicamente, mas é em si mesmo uma estrutura matemática “tornando-nos partes auto-conscientes de um gigante objeto matemático” (2014, p. 305).

Esta visão sugere que a matemática não é apenas uma linguagem conveniente para descrever a Natureza, mas sim a sua própria substância fundamental. Se o Universo pode ser completamente descrito por equações matemáticas, então, segundo Tegmark, não há distinção entre a realidade física e a estrutura matemática subjacente—ambas são a mesma coisa (Tegmark, 2014, 2017; Livio, 2010). Assim, a nossa capacidade de compreender a realidade através da matemática é um reflexo do facto de que fazemos parte dessa mesma estrutura matemática. Como refere Tegmark,

“... se algumas equações matemáticas descrevem completamente tanto a nossa realidade física externa como uma estrutura matemática, então a nossa realidade física externa e a estrutura matemática são uma e a mesma, e então a Hipótese do Universo Matemático [postulado pelo próprio Tegmark] é verdadeira...”⁷¹ (2014, p. 315).

⁷⁰ O superdeterminismo é uma teoria que defende que certos aspetos da realidade não podem ser observados nem medidos de forma rigorosa porque o Universo possui sistemas que não podem ser compreendidos nem observados pela limitada mente humana. O ser humano pode achar que compreende a física quântica mas, na realidade, apenas compreende a aparência da mecânica quântica – um género de template ajustado às suas limitações cognitivas (Hossenfelder, 2022; Hooft, 2016 e 2017). O mesmo se passa com a matéria e a energia escura. Podemos observar os seus efeitos e, quiçá, prevê-los com alguma exatidão mas não podemos conhecer a sua natureza. Podem existir sistemas do Universo os quais não podemos violar.

⁷¹ Tradução do autor.

A eficácia extraordinária da matemática na descrição do mundo físico levanta questões fundamentais sobre a sua relação com a realidade. Como Tegmark argumenta, se a matemática descreve com exatidão os padrões e leis do Universo, isso pode indicar que o próprio Universo é matemático (Tegmark, 2014, 2017; Livio, 2010). A mesma perspectiva é defendida pelo físico Mario Livio, que, no seu livro *Is God a Mathematician?*, sugere que a matemática não é uma mera construção humana, mas sim uma entidade que permeia e estrutura todo o cosmos. Tanto para Tegmark como para Livio, a capacidade da matemática para revelar as leis da Natureza de forma tão precisa sugere que não se trata apenas de uma ferramenta descritiva, mas sim da própria essência da realidade. Se assim for, a descoberta matemática não seria um ato de invenção, mas de revelação — *descobrimos* as estruturas matemáticas que já existem na própria Natureza. A matemática esteve e estará sempre ali. Para ser descoberta, não inventada. É a natureza do próprio Universo. O ser humano, como faz parte da Natureza, talvez tenha adquirido a capacidade de decifrar alguns dos seus mistérios através da matemática que foi descobrindo.

Esta visão tem profundas implicações para o nosso modelo de desdobramento informacional. Na nossa teoria, a informação estrutural é a base ontológica do Universo, contendo todas as potencialidades da realidade antes de estas se manifestarem como informação computacional. Se o Universo é uma estrutura matemática, então a informação estrutural que propomos pode ser entendida como um conjunto de relações matemáticas fundamentais que codificam a totalidade da realidade. O que Tegmark descreve como a estrutura matemática do cosmos pode ser interpretado, dentro da nossa teoria, como o tecido informacional que subjaz à física — um substrato ontológico que regula todas as interações e processos. Assim, podemos considerar que a informação estrutural corresponde a uma camada matemática oculta, onde as leis fundamentais da física emergem como manifestações computacionais dessas relações subjacentes.

Esta conexão torna-se ainda mais clara quando analisamos a forma como percebemos a realidade. Como Tegmark observa, o nosso sistema sensorial não tem acesso direto à estrutura matemática da Natureza — ele apenas capta as suas manifestações computacionais. Por exemplo, a visão humana não percebe diretamente os códigos matemáticos que governam a propagação da luz, mas sim as suas consequências observáveis, o que analisaremos mais à frente. No âmbito do nosso modelo, isso sugere que o

desdobramento informacional pode ser entendido como um processo no qual a estrutura matemática do Universo se traduz em informação computacional. A realidade fenoménica que observamos é apenas uma versão decodificada e acessível da camada matemática subjacente, que permanece inacessível na sua totalidade. O facto de conseguirmos compreender o Universo através da matemática pode ser visto, dentro da nossa estrutura teórica, como um indício de que a informação estrutural do Cosmos é regida por princípios algoritmos fundamentais, que só se tornam observáveis quando desdobrados na forma de informação computacional. Em poucas palavras, a perspetiva de Tegmark oferece um poderoso suporte para a nossa hipótese de que a realidade é fundamentalmente informacional. A estrutura matemática que ele propõe pode ser vista como o código-fonte do Universo, equivalente à nossa noção de informação estrutural, enquanto a realidade fenoménica, regida pelas leis da física, emerge como um desdobramento computacional dessas relações subjacentes. Isso implica que a própria física é uma forma de epistemologia da informação estrutural, revelando apenas fragmentos de uma realidade matemático-informacional muito mais vasta.

Em suma, a teoria do holomovimento de David Bohm (1980), o princípio holográfico de Gerard 't Hooft (2000) e Leonard Susskind (1993) e o universo matemático de Max Tegmark (2014) alinham-se com a nossa distinção entre informação ontológica estrutural e computacional. Cada um destes modelos converge na ideia de que a realidade observada é apenas a superfície visível de uma ordem ontológica mais profunda, estruturada por informação — matemática, computacional ou estrutural — cuja manifestação parcial constitui aquilo a que chamamos ‘física.

A causalidade (ou determinismo) como atualização de estados informacionais

A partir da perspetiva do nosso modelo informacional, a causalidade — ou determinismo — adquire um novo significado. Tradicionalmente, a causalidade é entendida como a transmissão de energia entre sistemas físicos, um conceito enraizado, tanto na física clássica (Newton, 1687/1934; Helmholtz, 1847/1977) como na física moderna (Wheeler e Feynman, 1945; Feynman et al.; 1964; Griffiths, 2013). No entanto, dentro da nossa abordagem, a causalidade pode ser reinterpretada como a atualização de estados informacionais no interior da estrutura computacional do universo. Se a energia for entendida como a forma dinâmica da informação computacional, então cada

interação física pode ser vista como uma reconfiguração desses estados informacionais. Assim, a energia não é apenas um meio de transmissão de força, mas o próprio veículo da computação universal, permitindo que a informação se transforme e propague no espaço-tempo.

Tradicionalmente, a causalidade é entendida como um processo de transmissão de energia entre sistemas físicos, onde um evento A causa um evento B através da interação de partículas ou de campos. Na mecânica clássica, este processo é descrito por equações deterministas (Newton, 1687/1934; Einstein, 1905), enquanto na mecânica quântica assume um caráter probabilístico (Rovelli, 2022). No entanto, dentro da nossa abordagem, esta concepção pode ser vista como uma simplificação de um fenômeno mais fundamental. A causalidade não precisa de ser entendida como um fluxo contínuo de energia, mas sim como a transição entre diferentes estados informacionais dentro da estrutura computacional do universo. Ou seja, em vez de vermos a causalidade como um mecanismo físico tradicional, podemos concebê-la como um processo de atualização de informação. Isto significa que o universo opera como um sistema computacional onde cada evento não é simplesmente o resultado de uma transferência de energia, mas sim de uma atualização de estados informacionais, tal como acontece num sistema computacional que processa dados e executa cálculos segundo regras definidas.

No nosso modelo, a energia desempenha dois papéis fundamentais: 1) permite que a informação seja propagada no espaço-tempo, como um sinal computacional que transmite estados informacionais; 2) ativa a computação universal, permitindo a transformação e reorganização da informação na estrutura do universo. Se o universo for entendido como um sistema computacional fundamentado em informação, então a causalidade pode ser reformulada como um processo de transição entre estados informacionais. Esta transição não é arbitrária, mas regida por leis computacionais fundamentais que estabelecem a forma como a informação pode ser processada e atualizada.

Assim, a energia atua como o motor desta atualização, assegurando que a informação não permanece estática, mas sim que evolui dinamicamente ao longo do tempo. A matéria, por outro lado, representa estados informacionais estabilizados, que preservam informação durante períodos de tempo mais longos. Assim, cada interação física pode ser vista como um cálculo dentro do sistema computacional do Universo, onde a

causalidade não é um fluxo de energia, mas sim a sucessão de estados informacionais computados em conformidade com regras estruturais.

Deste modo, se o universo for entendido como uma estrutura computacional baseada em informação, então a causalidade não é um princípio exterior ou imposto à realidade, mas sim o próprio processo de atualização contínua dos estados informacionais do universo. Assim, a causalidade deixa de ser interpretada como uma simples transmissão de energia e passa a ser vista como um mecanismo interno da computação universal. O tempo, por exemplo, pode ser entendido não como um fluxo contínuo, mas como a sucessão de estados informacionais computados. A matéria e a energia, por seu turno, deixam de ser princípios independentes e tornam-se expressões estabilizadas e dinâmicas da computação universal, respetivamente.

Esta abordagem dissolve a distinção tradicional entre matéria, energia e informação, ao revelar que todas estas entidades não são categorias separadas, mas sim diferentes manifestações de uma mesma realidade causal. A informação estrutural define as leis fundamentais do universo, estabelecendo os seus limites e princípios. A informação computacional rege a dinâmica interna do sistema, determinando a forma como a realidade se processa e evolui. A matéria e a energia emergem então como expressões diferenciadas deste processo, refletindo dois estados distintos da computação universal: a energia enquanto atualização dinâmica da informação, e a matéria enquanto estado estabilizado dessa informação.

Deste modo, a nossa teoria não apenas reformula a visão convencional sobre a energia e a matéria, mas também fornece um quadro mais profundo para compreender a relação entre causalidade, transformação e a estrutura fundamental do universo.

As partículas virtuais e o desdobramento informacional

As partículas virtuais são flutuações momentâneas do vácuo quântico, surgindo e desaparecendo em intervalos de tempo extremamente curtos, de acordo com os princípios da mecânica quântica (Griffiths, 2005). Representam um conceito central na teoria quântica de campos e surgem naturalmente como mediadoras das interações fundamentais, sendo essenciais para descrever a transmissão das forças fundamentais, como a eletromagnética e a nuclear forte e fraca (Peskin & Schroeder, 1995). Embora não possam ser detetadas de forma direta, os seus efeitos físicos são observáveis, como no

desvio de Lamb (Lamb & Retherford, 1947) ou no efeito Casimir (1948). Apesar de não corresponderem a partículas livres e permanentes, como os fótons ou elétrons, as partículas virtuais existem objetivamente enquanto excitações transitórias dos campos quânticos, que ocorrem durante processos de interação. São manifestações reais da atividade do vácuo quântico, permitidas pela relação de incerteza de Heisenberg, que admite variações breves na energia desde que confinadas a escalas temporais extremamente reduzidas. Nestes contextos, funcionam como portadoras de força, emergindo e colapsando dentro do tempo permitido pelas flutuações quânticas.

Ao contrário do que o termo vácuo pode sugerir, o vácuo quântico não é uma ausência de matéria ou energia, mas um estado dinâmico onde pares partícula-antipartícula emergem espontaneamente devido às flutuações do campo quântico. Essas flutuações foram observadas experimentalmente por meio de efeitos, como o efeito Casimir, que demonstra a existência de uma força atrativa entre duas superfícies condutoras colocadas no vácuo devido à presença de flutuações do campo eletromagnético (Casimir, 1948). Outra evidência é encontrada na polarização do vácuo, onde partículas virtuais interagem com campos eletromagnéticos, alterando medições de precisão, como o momento magnético do elétron (Aoyama et al., 2012).

Sob o prisma do *desdobramento informacional* que sugerimos, essas partículas são manifestações inevitáveis e causalmente determinadas de um campo informacional ontológico, onde a informação estrutural estabelece as condições necessárias para sua emergência e a informação computacional define precisamente as dinâmicas que regem sua formação e comportamento. O vácuo quântico, por conseguinte, é reinterpretado como uma projeção física inevitável da informação estrutural. Nesse contexto, as flutuações correspondem a expressões precisas e causalmente determinadas de interações potenciais latentes no sistema informacional.

Por conseguinte, as partículas virtuais podem ser vistas como manifestações locais e transitórias do desdobramento informacional, sendo reguladas de forma determinística pelas leis do campo informacional. Elas emergem porque as condições subjacentes, definidas pela informação ontológica, determinam inevitavelmente a possibilidade de sua formação momentânea. O desaparecimento dessas partículas também é regido pelas mesmas regras informacionais, de modo que não há aleatoriedade verdadeira, mas sim uma ordem causal implícita no tecido informacional do Universo. Essa dinâmica

está alinhada com o princípio de incerteza de Heisenberg (1958), que, sob a perspectiva determinística, reflete não uma indeterminação fundamental, mas a forma precisa como os limites do sistema informacional regulam o comportamento das partículas em curtos intervalos de tempo.

O conceito de *reintegração informacional* revela-se, também, essencial para a compreensão deste fenômeno. Se o *desdobramento informacional* explica como as partículas virtuais emergem como manifestações físicas transitórias, a *reintegração informacional* determina a maneira como essas flutuações são reintegradas ao substrato informacional. O desaparecimento das partículas virtuais não é, assim, um processo aleatório, mas sim uma reinserção ordenada da informação computacional na matriz estrutural, assegurando a coerência global do sistema. Neste sentido, as partículas não desaparecem. Apenas são integradas na ontologia do Universo. Por isso é que a informação nunca é perdida, como sugere o *Princípio Holográfico* (Hooft, 2000; Susskind, 1993). Ela é desdobrada e reintegrada.

O desdobramento informacional que sugerimos, demonstra que a realidade observável emerge como uma consequência inevitável de um campo subjacente de possibilidades informacionais. As partículas virtuais representam a "fronteira" entre o potencial (estrutural) e o "real" (computacional), sendo a sua existência determinada por condições precisas da SUI. Quando essas condições atingem o limiar necessário, como em colisões de alta energia, as partículas virtuais tornam-se "reais", seguindo as regras definidas pela informação estrutural. Por exemplo, em aceleradores de partículas como o LHC (Large Hadron Collider), colisões de alta energia convertem pares de partículas virtuais em partículas reais, como evidenciado na produção de quarks (CMS Collaboration, 2020).

Nestes casos, em que os seres humanos "fabricam" partículas, eles não estão a criar algo *ex nihilo*. Estão, sim, a ativar determinados padrões informacionais latentes já inscritos na estrutura do Universo. A informação estrutural define os algoritmos fundamentais que regem a realidade, e a intervenção humana, ao manipular condições energéticas específicas, apenas fornece o *input* necessário para que certos outputs (partículas) sejam computados a partir dessa estrutura. Assim, as partículas produzidas em laboratório não são entidades arbitrárias nem criações independentes da vontade

humana. São manifestações concretas de possibilidades pré-inscritas no campo informacional estrutural, cuja ativação depende de parâmetros energéticos e topológicos precisos. Por outras palavras, o ser humano apenas desencadeia a realização de possibilidades que já existem na ontologia causal do Universo, ao introduzir as condições computacionais adequadas que permitem que esses padrões passem do estado virtual (estrutural) para o estado físico (computacional).

Incorporar as partículas virtuais no modelo do desdobramento informacional possui implicações filosóficas importantes. 1) Primeiro, implica que a realidade é constituída por processos dinâmicos causalmente determinísticos que emergem de um substrato informacional ordenado. As partículas virtuais não surgem por acaso, mas como expressões necessárias do desdobramento. Existe uma conexão inevitável entre potencial-real. 2) Segundo, implica que as partículas virtuais ilustram que o potencial e o real são conectados por relações determinísticas, com o desdobramento agindo como o mecanismo causal que medeia essa transição. 3) Terceiro, implica que existe uma unidade causal, explicada pela SUI. Como analisado, tanto as flutuações quânticas quanto as partículas virtuais são manifestações necessárias de um tecido informacional unificado e ordenado.

As partículas virtuais oferecem, portanto, uma janela para compreender como o desdobramento informacional se manifesta no nível quântico. Elas não são apenas artefactos teóricos, mas evidências de uma ordem causal subjacente que permeia o Universo. Essa integração enriquece tanto a interpretação da física quântica quanto a compreensão da ontologia da informação como um sistema determinístico fundamental.

Como o desdobramento informacional pode explicar a dualidade partícula-onda

A dualidade partícula-onda⁷², é ainda um problema do âmbito da física quântica. Isto porque revela uma ambiguidade profunda na natureza das entidades quânticas. Fótons e eletrões comportam-se ora como partículas, ora como ondas, dependendo da forma como são medidos. Este fenómeno desafia a lógica clássica e levanta a questão fundamental: o que são realmente estas entidades antes da observação? A física

⁷² Vimos no capítulo 2 que a dualidade partícula-onda refere-se ao facto de entidades quânticas, como eletrões e fótons, poderem comportar-se ora como partículas localizadas, ora como ondas distribuídas, dependendo da interação com o meio e do tipo de medição realizada (Zurek, 2003).

descreve com precisão os resultados, mas não há consenso sobre o que está de facto a acontecer. A dualidade parece apontar para uma realidade mais profunda, ainda não totalmente compreendida, onde os conceitos de partícula e onda são apenas aproximações (Zurek, 2003).

No âmbito da dualidade partícula-onda, o desdobramento informacional sugere que esta não é uma dualidade no sentido clássico, mas sim duas expressões complementares de um processo causal contínuo. No nosso modelo, a informação estrutural representa o estado de potencialidade da onda, enquanto a informação computacional governa as dinâmicas que transformam esse potencial em manifestações localizáveis – as partículas localizadas. Segundo esta perspetiva, a partícula pode ser entendida como uma manifestação explícita, um estado local resultante da interação de informação computacional com o meio, enquanto a onda reflete a natureza latente e distribuída da informação estrutural, onde as possibilidades ainda não foram completamente diferenciadas.

Neste contexto, as partículas quânticas desempenham um papel central como mediadoras da transição entre potencialidade e atualização informacional, permitindo que os processos físicos se tornem observáveis. A sua dualidade comportamental não representa uma contradição, mas a alternância entre dois modos de expressão da mesma realidade informacional – um latente e distribuído, outro localizado e mensurável. Enquanto onda, a partícula encontra-se num estado estrutural, representando um campo de possibilidades informacionais que ainda não foi concretizado num estado computável e mensurável. Esse comportamento reflete a ordem implicada de Bohm (1980), onde a informação permanece numa forma latente, distribuída no espaço-tempo, sem uma localização definida. No entanto, quando a partícula interage com um sistema físico, como num processo de medição, ela colapsa para um estado computacional, assumindo uma configuração discreta e localizada. Nesse momento, a informação estrutural desdobra-se em informação computacional, permitindo que a partícula atue como um bit de informação concretizado num evento físico.

Esta transição pode ser observada no experimento da dupla fenda, onde a partícula exhibe comportamento ondulatório até ser medido, momento em que passa a comportar-se como uma partícula (Rovelli, 2022). Assim, a dualidade onda-partícula é um

reflexo do desdobramento da informação estrutural em informação computacional, sendo a própria partícula veículo desse processo.

As nossas ideias sobre desdobramento informacional complementam e expandem o modelo de *enfoldment* e *unfoldment* de Bohm (1980) ao enfatizar o papel central da informação como base ontológica da realidade. No desdobramento informacional, a transição de um estado profundo e implicado para um estado explicado não é apenas um processo de "desenrolar" dinâmico, mas também de transformação informacional. A informação estrutural define as condições e os padrões latentes que suportam a emergência da realidade, enquanto a informação computacional atua como o mecanismo que transforma essas possibilidades em manifestações físicas concretas. Esta abordagem permite interpretar o *unfoldment* de Bohm (1980) como um processo informacional contínuo onde o que o físico designa de ordem implicada corresponde ao estado potencial e interligado da informação estrutural – a sobreposição que observamos na mecânica quântica. Já a ordem explicada de Bohm emerge quando a informação computacional organiza esses potenciais em formas localizáveis e observáveis, como partículas. Quando uma sobreposição interage com algum objeto do mundo clássico colapsa e desdobra-se em informação computacional.

No contexto da dualidade partícula-onda, as nossas ideias reforçam que a transição de um estado implícito para um explícito não é uma mera mudança de estado físico, mas um desdobramento informacional que reorganiza e manifesta padrões previamente codificados. Esta distinção entre informação estrutural e computacional fornece uma explicação coerente para fenômenos como a dualidade partícula-onda, como vimos. Uma vez que a informação computacional emerge de uma estrutura causal subjacente, é razoável sugerir que aquilo que percebemos como dualidade — em que as entidades exibem comportamentos simultaneamente semelhantes a partículas e a ondas (Zurek, 2003) — resulta da forma como a informação computacional ontológica é processada e manifestada em diferentes níveis da realidade. O ato de medição não cria a realidade, mas revela diferentes facetas do mesmo substrato informacional, dependendo das restrições informacionais impostas pelo processo de observação.

Em suma, a teoria do desdobramento informacional sugere que a realidade emerge da interação entre informação estrutural (potencialidade) e informação computacional (manifestação). A dualidade partícula-onda reflete esse processo. A onda

representa a informação latente, enquanto a partícula surge quando essa informação se torna computacional e observável. O colapso quântico pode, assim, segundo o nosso modelo, ser interpretado como a transição de um estado estrutural para um estado computacional, revelando a natureza informacional e causal do Universo.

Como o desdobramento informacional pode explicar a não-localidade quântica

Outro ponto crucial no nosso modelo de desdobramento é a explicação para a não-localidade quântica, um dos fenómenos mais desafiadores da física moderna. Tradicionalmente, a não-localidade refere-se ao facto de que partículas podem exibir correlações instantâneas entre si, mesmo quando separadas por grandes distâncias, desafiando a concepção clássica de proximidade espacial como requisito para interação. Esta realidade foi comprovada recentemente por alguns estudos (Hensen et al., 2015; Giustina et al., 2015; Aspect, 1982; Weihs et al., 1998), que confirmaram a violação das desigualdades de Bell⁷³ (Rovelli, 2018; Tegmark, 2014). Esses resultados demonstram que a mecânica quântica prevê corretamente a existência de entrelaçamento, onde pares de partículas compartilham um estado quântico comum, independentemente da distância que as separa. Este entrelaçamento entre partículas viola, aliás, um dos princípios da física propostos por Einstein e, aliás, comprovados repetidamente pela física teórica e experimental: que nada pode viajar mais rápido do que a velocidade da luz (Smolin, 2019; Rovelli, 2018).

Na perspectiva de Bohm, essas interações não-locais não são limitadas pela distância física porque as partículas e suas informações estão inscritas numa estrutura global que transcende o espaço-tempo clássico (1980). No desdobramento informacional que sugerimos, essa visão é complementada pela ideia de que a informação computacional opera de maneira não-local, pois a totalidade do sistema é governada por um

⁷³ As Desigualdades de Bell, formuladas por John Bell (1964), estabelecem um limite para a correlação entre partículas distantes sob a hipótese de realismo local, que assume que as propriedades das partículas existem independentemente da medição (realismo) e que nenhuma informação pode viajar mais rápido do que a luz (localidade). No entanto, experiências como as de Aspect (1981-82) demonstraram que as previsões da mecânica quântica violam essas desigualdades, revelando correlações mais fortes do que seria possível num modelo de variáveis locais ocultas. Essa violação implica que a realidade não pode ser simultaneamente realista e local, indicando que partículas emaranhadas partilham uma conexão instantânea, independentemente da distância. Assim, as desigualdades de Bell forneceram um dos testes mais sólidos contra a visão clássica da física, consolidando a não-localidade quântica como um fenómeno fundamental da natureza.

algoritmo ontológico subjacente. Esse algoritmo define as regras que conectam as partículas e campos, mesmo quando estão separadas no espaço clássico. Aliás, como referimos, a SUI permite a existência de uma realidade potencial que, na física quântica, é explicada pela sobreposição. Se todas as partículas emergem desse substrato fundamental, então o entrelaçamento pode ser entendido como uma consequência natural da continuidade da informação estrutural. A ligação entre partículas entrelaçadas não seria uma comunicação mais rápida do que a luz, mas uma manifestação da coerência informacional inerente à SUI, onde as partículas permanecem interligadas através de um campo informacional subjacente.

Por outras palavras, o desdobramento informacional sugere que a não-localidade pode ser explicada pela continuidade da informação estrutural que sustenta todo o Universo. Nesse modelo, partículas entrelaçadas partilham o mesmo tecido informacional, permitindo que as informações entre elas sejam sincronizadas independentemente da separação espacial. A informação computacional, ao desdobrar as interações no espaço-tempo, revela essas conexões que já estavam implícitas na informação estrutural. Este princípio pode ter implicações não apenas para a interpretação da mecânica quântica, mas também para áreas como a computação quântica, onde a interdependência informacional entre *qubits* reflete essa mesma estrutura unificada (Lloyd, 2007).

Dessa forma, a não-localidade deixa de ser um mistério ou uma violação do espaço-tempo e passa a ser entendida como uma consequência direta da interconexão informacional fundamental do universo. A informação estrutural atua como o suporte contínuo que permite a comunicação instantânea, enquanto a informação computacional transforma essas conexões em eventos observáveis, como as correlações quânticas. Assim, aquilo que aparece como "ação fantasmagórica à distância"⁷⁴ (Einstein et al., 1935) não é mais do que a expressão do desdobramento informacional em ação.

3.3.3. A informação representacional

Como temos vindo a referir, todos os sistemas físicos do Universo conservam e processam informação. No entanto, apenas os sistemas nervosos (gerados por informação computacional biológica) podem fazer tudo isso e, também, *representar*

⁷⁴ Spooky action at a distance.

informação. Como veremos, a representação da informação — e não apenas a sua computação — é um fenômeno exclusivamente biológico. É precisamente isso que distingue os objetos dos seres vivos: o tipo de informação que processam. Se o processamento for apenas físico e químico, estamos perante um objeto; se, para além da informação física e química, o sistema também processar informação de natureza biológica, então trata-se de um ser vivo. De entre os seres vivos, alguns — aqueles que possuem sistemas nervosos — têm a capacidade de representar informação. Isto significa que a natureza da SUI (a conservação e o processamento de informação) permite que a informação computacional possa ser *transformada* em informação representacional (isto é, em informação que gera imagens mentais). Mas, não menos importante, isto significa, também, que informação representacional pode ser transformada em informação computacional.

A informação representacional surge quando um sistema (ex. um cérebro) adquire uma faculdade adicional à conservação e ao processamento\transformação de informação — a de representar informação, isto é, de criar imagens que representam, de algum modo, a realidade. A representação *não* é uma condição obrigatória da informação como as acima enumeradas, mas é uma consequência *possível* da natureza da informação. Dito de outra forma, quando um determinado sistema possui um conjunto de características, nomeadamente um sistema nervoso biológico, o resultado é a capacidade desse sistema representar informação, e não apenas conservar e processar informação — como as pedras, a água ou as plantas, entre inúmeros sistemas físicos e biológicos. Assim, a informação representacional é a capacidade de um sistema transformar informação computacional em imagens (conscientes ou não). Seria o que, na filosofia kantiana, corresponderia à transformação do acesso potencial ao real em representação fenoménica (Kant, 2008; Silveira, 2002). As representações que sugerimos podem ser do meio (ambiente) ou de eu (do *self*). Assim, quando um organismo adquire a capacidade de representar o meio, desenvolve uma mente. Quando, além disso, adquire a capacidade de representar a si mesmo — isto é, de construir um modelo de eu — então esse organismo possui uma mente consciente.

No entanto, as representações que os sistemas nervosos fazem estão muito longe de ser réplicas da realidade. As partículas quânticas, por exemplo, não são representáveis tal como elas são. Muito longe disso, até porque as partículas quânticas nunca

foram observadas diretamente. Não porque não sejam computáveis mas devido à natureza do sistema representacional (sistemas nervosos) cuja função é *representar* a realidade, não *replicar* a realidade. Por isso, não vemos a realidade como ela é — um mar de partículas-ondas com as correspondentes instruções informacionais intrínsecas —, mas apenas as suas representações (físicas), ou seja, os sistemas que emergem da SUI, compostos por átomos, moléculas e células.

Talvez o exemplo mais evidente do que pretendemos demonstrar seja o ar. Quando olhamos para o ar, percebemos um espaço aparentemente vazio. Contudo, esta percepção é limitada pelo nosso aparato cognitivo e pelos sentidos, que não conseguem captar diretamente as partículas que compõem o ar. O ar está repleto de moléculas de oxigénio, nitrogénio, dióxido de carbono e outros gases, mas essas partículas são invisíveis para os nossos olhos e escapam à percepção imediata. Agora, imaginemos retirar todos os gases de uma caixa, criando um vácuo. Poderíamos imaginar que a caixa ficaria, em rigor, completamente vazia, sem ar. No entanto, essa ideia é novamente fruto das limitações da percepção humana. Mesmo num vácuo, a caixa não estaria realmente vazia. Espaços aparentemente vazios são preenchidos por flutuações quânticas e partículas virtuais, que surgem e desaparecem continuamente, como demonstrado por fenómenos experimentais como o efeito Casimir (Casimir, 1948), já aqui referido.

Este facto ilustra a profundidade das nossas limitações percetivas. O vazio, ou vácuo, que percebemos e preenchemos com os nossos corpos não só não está, de todo vazio, como é, na realidade, uma projeção simplificada de um sistema mais complexo que os nossos sentidos não conseguem apreender na totalidade. O que chamamos de vazio, ou de vácuo, faz parte de um tecido dinâmico repleto de informações que se desdobram na física e, por vezes, voltam a reintegrar-se nesse tecido, em processos e interações invisíveis aos nossos olhos.

Assim, os sistemas de representação (os sistemas nervosos), por um lado, possuem limitações, o que explica porque, provavelmente, uma grande parte da informação computacional (como os quarks, eletrões, átomos e moléculas) ou é representada de uma forma muito distorcida ou não pode ser representada de todo (Silveira, 2002); por outro lado, os sistemas de representação são distintos, o que condiciona o tipo de informação computacional representável pelo sistema. Um golfinho, um lagarto ou uma águia, por exemplo, representam determinados tipos de informação computacional,

enquanto um ser humano representa outro tipo de informação computacional. Face a um mesmo objeto, estes quatro seres estão perante o mesmo tipo de informação computacional mas, como possuem sistemas de representação distintos, desenvolvem representações mentais distintas.

Ora, isto significa que, apesar de existir informação computacional de difícil ou impossível representação, uma boa parte da informação computacional pode ser representada. E pode ser representada através de sistemas nervosos com capacidade de processar e transformar informação computacional em informação representacional. Da mesma forma que, como veremos, uma planta transforma luz em oxigénio (Taiz & Zeiger, 2010) (através de um processamento de informação computacional que *não* se transforma em informação representacional), um sistema nervoso transforma luz em imagens mentais – isto é, transforma informação computacional em informação representacional. Isto explica-se porque, como dissemos, a planta possui a capacidade de *conservar* e de *processar* (e *transformar*) informação, mas não possui a faculdade de *representar*, ao contrário dos seres com sistemas nervosos.

Em síntese, propomos que a informação só adquire o estatuto de representação quando emerge de um processo biológico capaz de transformar informação computacional em informação representacional. Esse processo, no entanto, como veremos mais à frente, está dependente de alguns neurónios que possuem uma dupla função: computar e representar informação.

3.4. As interações entre a informação computacional e informação representacional

A ideia de que o cérebro possui um sistema que representa a realidade não é nova. Já em 1943, mesmo antes do advento dos computadores, o filósofo Kenneth Craik sugeriu no livro *The Nature of Explanation* (1943) que os organismos possuíam um sistema que lhes possibilitava lembrar o passado, prever o futuro e reagir ao presente através de um modelo representacional do mundo. Desde então, a ideia de representação mental adquiriu uma maior relevância no âmbito da psicologia, da filosofia e das neurociências, sendo atualmente um conceito amplamente utilizado nessas áreas (Tononi 2016; Koch, 2016; Fodor, 1990; Dretske, 1995; Dennett, 1991).

Como referido, sugerimos que a informação representacional corresponde à capacidade que determinados sistemas informacionais possuem para decodificar e transformar informação “dura” — isto é, informação ontológica computacional — em padrões que adquirem significado interno no seio do próprio sistema. Trata-se de uma forma de processamento informacional que não apenas gera afinidade físico-química — como acontece em sistemas mais simples —, mas representa, permitindo a emergência de um nível qualitativamente distinto de organização.

Num sistema como o Universo em que vivemos, que processa constantemente informação, não é difícil identificar processos emergentes, desdobrados da informação estrutural, que atuam como sistemas emergentes de processamento computacional. Na verdade, tudo no Universo — todas as coisas, todos os objetos, todos os fenómenos físicos, químicos ou biológicos — são expressões deste processamento informacional emergente. É precisamente este processamento, regido por leis determinísticas de organização da informação, que permite o funcionamento do Universo como sistema dinâmico, auto-organizado e estruturado. Contudo, de entre todos os sistemas que processam informação, alguns revelam uma capacidade mais complexa ou evidente de transformação informacional. Um exemplo paradigmático são as plantas.

A fotossíntese é o processo pelo qual as plantas convertem luz solar em energia química. Esse processo ocorre nos cloroplastos, onde a clorofila absorve luz solar, especialmente nas faixas azul e vermelha. A energia solar é utilizada para dividir moléculas de água, libertando oxigénio como subproduto. Os eletrões gerados passam por uma cadeia de transporte, resultando no armazenamento de energia. Com essa energia, o dióxido de carbono do ar é transformado em glicose no ciclo de Calvin. Assim, a fotossíntese não só fornece energia para o crescimento das plantas, como também liberta oxigénio na atmosfera (Taiz & Zeiger, 2010). Ora, isto significa que um organismo multicelular complexo, como uma planta, ao converter energia luminosa em energia química e ao libertar oxigénio, não está a fazer nada mais do que processar e transformar informação — informação que teve de ser conservada, decodificada e funcionalmente processada.

De uma forma mais genérica, sugerimos que a planta recebe luz (bits de informação) em forma de um tipo de informação computacional que processa e transforma num outro tipo de informação computacional — oxigénio. Em rigor, o sistema

informativa que designamos de planta transforma luz em oxigênio. Deste modo, os processos emergentes possibilitados por bits de informação nas escalas mais elementares permitem que bits de informação em forma de luz sejam *transformados* em bits de informação em forma de oxigênio. Isso é possível porque existe um sistema complexo e sofisticado que conserva, processa e transforma causalmente bits de informação. Assim, a fotossíntese pode ser vista como uma "leitura" ou "decodificação" de informações presentes na luz e uma "reconstrução" dessa informação num formato distinto, adaptado ao organismo da planta.

Esta nossa proposta encontra ressonância em estudos de biologia quântica que sugerem a existência de mecanismos informativos intrínsecos aos sistemas biológicos. Lambert et al. (2013), por exemplo, investigam a coerência quântica em sistemas vivos, destacando a fotossíntese como um caso paradigmático em que a transferência de energia obedece a padrões estruturalmente informativos. Iremos sugerir o mesmo processo para os sistemas nervosos com uma diferença central: os sistemas nervosos decodificam luz, não em oxigênio, mas em representações.

Mais atrás vimos que a informação ontológica pode dividir-se em dois tipos: informação ontológica estrutural, que corresponde à base fundamental da realidade; e informação ontológica computacional, que instrui as partículas a comportarem-se de determinadas formas. Iremos propor agora dois tipos de informação representacional: a *representação do meio* e a *representação do eu*. Argumentamos que a informação representacional permite o surgimento de uma mente e de uma consciência através de duas fontes de estímulos – o ambiente e o corpo. A capacidade de um sistema de informação representar o meio sem que represente um eu, permite-lhe o surgimento de uma mente, ao passo que um sistema informativo que tenha a capacidade de representar o meio e um eu, permite-lhe a emergência de uma mente consciente. Isto significa que existem sistemas nervosos capazes de representar o meio (como os de répteis e peixes) sem representar o eu (sem *qualia*), e sistemas nervosos capazes de representar o meio e o eu (como os dos primatas e de outros mamíferos complexos) (Humphrey, 2023; LeDoux, 2020). O nível de consciência desses animais depende, a nosso ver, da quantidade de *bits* de informação que os seus sistemas nervosos conseguem conservar, integrar e processar – um processo descrito, por Giulio Tononi, como Phi de consciência (2008).

Dessa forma, o sistema nervoso permite a emergência de representações do meio (como imagens visuais, táteis e sonoras) e, em algumas espécies, gera também representações do corpo-cérebro em forma de um modelo interno de eu (Damásio, 2010; Metzinger, 2009).

A ideia proposta por António Damásio em *O Livro da Consciência* (2010) sobre o eu como uma representação *sentida* do corpo-cérebro – um *duplo neuronal* – é central para a sua teoria da consciência e ilustra de uma forma clara o que pretendemos descrever. Damásio argumenta que a consciência não é um fenómeno isolado, mas uma função emergente de um sistema biológico complexo, onde o cérebro e o corpo estão profundamente interligados, não só entre si mas também com o meio. A partir dessa perspetiva, a consciência não é apenas um processo cognitivo abstrato, mas uma expressão da interação contínua entre o cérebro, o corpo e o ambiente. Damásio sugere que o *duplo neuronal* não é uma entidade pré-existente ou algo que apenas surge do cérebro em isolamento, mas sim uma construção contínua e dinâmica do sistema cérebro-corpo em relação com meio. Ele afirma que o eu emerge de representações sentidas do corpo e do mundo, sendo que a percepção sentida do corpo é fundamental para a formação da consciência. Essas representações sentidas incluem a sensação de estar num corpo, a sensação do corpo estar vivo, e a capacidade de integrar essas sensações de maneira consciente (Damásio, 1993/2011 e 2013). O eu que sentimos, o *self*, é então um modelo *qualitativo*, um *duplo-neuronal* fenomenológico gerado pelo cérebro a partir das interações do corpo (de onde os sentimentos de emoção emergem) com o ambiente, e esse modelo torna-se um sentido interno, uma percepção integrada e contínua da própria existência.

O conceito de duplo-neuronal que Damásio utiliza é interessante porque implica que a consciência não é uma réplica exata do mundo, mas sim uma projeção da realidade com base nas experiências sensoriais e nas representações internas do corpo e do ambiente. A consciência, assim, emerge quando as imagens do meio, ou seja, as representações sensoriais do mundo externo, são apropriadas ou integradas pelas imagens sentidas do corpo, formando um sentido unitário de existência – o *sentimento de si* – quer permite, depois, o surgimento do duplo neuronal. Assim, o cérebro cria um *modelo* qualitativo do corpo e do mundo, e esse modelo – o *duplo neuronal* – é o que experienciamos como a nossa consciência. A consciência é, portanto, em parte, o cérebro, mas

também não é apenas o cérebro – ela resulta da capacidade de integrar informações que envolvem o corpo e o ambiente, num ciclo contínuo de percepção, processamento e representação (2010, 2011 e 2013).

O filósofo Thomas Metzinger possui uma ideia em tudo idêntica (2009). Segundo Metzinger, a transparência fenomenológica descreve o modo como a consciência nos fornece um modelo subjetivo da realidade, ocultando os processos neurais que o geram. Ou seja, não percebemos que a nossa experiência é uma simulação interna, pois o cérebro oculta os seus próprios mecanismos de processamento. De forma semelhante, o duplo neuronal de Damásio representa uma versão simplificada dos sistemas neuronais que geram a autorrepresentação do corpo, tornando-se a base da consciência sem que o sujeito tenha acesso direto ao seu funcionamento. Assim, a transparência de Metzinger e o duplo neuronal de Damásio convergem na ideia de que a consciência é um modelo representativo que oculta a complexidade dos processos que a geram.

O eu que propomos no nosso modelo de informação representacional é um processo emergente do cérebro, tal como o duplo-neuronal emerge do cérebro e do corpo (Damásio, 2010), tal como a transparência fenomenológica emerge do cérebro (Metzinger, 2009) ou, em certa medida, tal como o fenómeno emerge de um númeno (Kant, 2008). Da mesma forma que nunca temos acesso à realidade do númeno, também nunca temos acesso direto à realidade do nosso cérebro ou do nosso corpo — apenas ao fenómeno que deles emerge. Assim, o eu é uma representação fenomenológica do cérebro e do corpo. No nosso modelo, a informação representacional surge quando a informação computacional proveniente do meio ou do corpo é transformada em imagens mentais, sejam elas conscientes ou não. Esta transformação é causada pela informação computacional e é através do processamento das computações neuronais que as representações mentais se geram, não obstante não termos qualquer conhecimento ou acesso à maquinaria intrínseca ao processamento que permite a consciência, como sugerem Damásio e Metzinger. No entanto, ao contrário da posição kantiana, segundo a qual existe um véu intransponível entre o sujeito e o mundo — mediado pelas formas a priori da sensibilidade e do entendimento –, defendemos que sujeito e mundo fazem parte de um mesmo sistema ontológico e informacional, sendo, por isso, capazes de interagir de forma direta. Aquilo que o sujeito observa não está limitado por uma barreira epistemológica absoluta, mas constitui antes uma representação informacional

coerente com o processamento efetuado pelo sistema no qual sujeito e mundo estão integrados.

No âmbito do nosso modelo de consciência, a informação computacional refere-se aos sinais que geram representações de estados do mundo externo ou interno, como estímulos sensoriais ou estados corporais, que são processados de forma complexa pelas redes neurais. Essas redes integram, com base nas informações recebidas, uma representação do que está a acontecer, seja no ambiente ou dentro do corpo. O cérebro, ao processar esses sinais, constrói representações do mundo e de si próprio. Por exemplo, ao olhar para um objeto, a percepção visual que temos dele não é apenas uma simples recepção de dados, mas uma construção de uma imagem mental do objeto, baseada nos estímulos sensoriais (e informacionais) que chegam aos nossos órgãos de percepção. Este processo acontece de maneira dinâmica e contínua, resultando em imagens mentais que podem ser tanto conscientes, como a percepção direta de um objeto, quanto não conscientes, como a regulação automática da homeostasia do corpo. As representações, portanto, não são cópias exatas da realidade, mas modelos internos do que está a acontecer, tanto no meio ambiente quanto no interior do organismo.

David Rosenthal, na sua *Teoria do Estado Superior de Ordem* (HOT), propõe que a consciência emerge de pensamentos de ordem superior que representam estados mentais de ordem inferior (Rosenthal, 2005). Este modelo é, também, consistente com a presente abordagem fisicalista-informacional, que estrutura a realidade em diferentes níveis ontológicos de informação. A informação computacional constitui o nível base, responsável pelos processos dinâmicos do sistema físico e biológico, enquanto a informação representacional reflete o nível superior, onde a experiência consciente emerge. Nesta perspectiva, os estados computacionais podem ser vistos como análogos aos estados de ordem inferior na HOT, e a informação representacional como a manifestação de estados de ordem superior. Por exemplo, no caso da dor, a informação computacional refere-se aos sinais nervosos processados pelo sistema nervoso, enquanto a informação representacional cria a *experiência* consciente de "estar a sentir dor" (Rosenthal, 2005).

A causalidade entre a informação computacional e a informação representacional implica que a mente não é um simples recetor da informação, mas um sistema ativo que interpreta e *transforma* essa informação em imagens. A consciência e as representações mentais emergem do processamento complexo de informações no cérebro, e

cada percepção sensorial, como a visão ou o tato, pode ser entendida como uma representação do meio, formada a partir da informação computacional (dos sinais) que o cérebro processa e decodifica. O cérebro é, assim, um processador dinâmico e contínuo de sinais que cria um mundo interno, composto por representações que refletem ou modelam a realidade externa e interna.

Este processo de transformação da informação computacional em informação representacional é essencial para a emergência da mente e da consciência. As redes neuronais, ao interagirem de forma coordenada, criam padrões de informação que se manifestam como representações. Estas representações, que podem ser imagens, pensamentos ou sentimentos, são os modelos que o cérebro constrói para refletir o mundo e o corpo, possibilitando que o organismo tenha experiências subjetivas, tome decisões e interaja com o seu ambiente. Em poucas palavras, as representações são o resultado do processamento contínuo e dinâmico da informação computacional, e esse processo é fundamental para a emergência da consciência – do eu. A capacidade de representar o meio e o eu conferiu aos organismos uma vantagem evolutiva significativa, permitindo-lhes antecipar e preparar o futuro com base no conhecimento acumulado. Isso destaca o papel da consciência na sobrevivência humana e nas conquistas que exigem memória e projeção do futuro (Verheyen, 2021), o que analisaremos mais à frente.

Como anteriormente referido, a natureza informacional do cérebro confere-lhe a faculdade de conservar, processar, transformar e *representar* informação. Diferentes sistemas (átomo, célula, planta) processam informação em diferentes níveis de complexidade. Plantas processam mais informação que uma célula, mas não possuem sistemas nervosos para representar essa informação. Já o cérebro humano, pela sua elevada capacidade de processamento e integração de informação (Tononi, 2004), transforma informações computacionais em representações mentais complexas e também converte informações representacionais em ações computacionais mais diversificadas. Sugerimos que a mente consciente transforma (alguma da) ontologia do Universo em epistemologia (Fredkin, 2003; Meijer, 2012; 2015). Essa epistemologia é o que designamos de física, de biologia e de neurociência.

Talvez sejamos de facto o Universo a contemplar-se (Sagan, 2009). Criados pelo mesmo tipo de informação do resto do Universo, é possível que a SUI, ao possibilitar fenómenos emergentes, que surgem através de processamento, decodificação e

integração da informação, possa ter causado um tipo de informação fenomenológica que tem a capacidade de representar-se. Ao desvendar-se através de uma das suas partes – o ser humano –, a SUI (ou o Universo) está-se a conhecer, ou, talvez a revelar-se. Talvez o ser humano não seja mais do que informação a representar informação – o Universo a representar-se a si próprio (Tegmark, 2014).

Importante referir que não defendemos qualquer forma de pampsiquismo. A mente consciente é, na nossa perspetiva, um fenómeno emergente, causado pelo desdobramento da informação estrutural em informação computacional, e desta em informação representacional.

3.4.1. *O modelo de descodificação recursiva*

No nosso modelo, a mente consciente pode ser entendida como um sistema de descodificação de informação que opera de forma recursiva. Esta recursividade permite transformar informação computacional em informação representacional, criando um ciclo dinâmico onde a experiência consciente não é apenas o produto final de um processamento, mas um elemento ativo que influencia as computações subsequentes do sistema nervoso. Esta realimentação informacional permite que a consciência descodifique informação computacional do corpo e do meio, o que permite a experiência fenomenológica.

Esta transformação da informação não ocorre de forma linear, mas sim através de sucessivas reinterpretações dentro do próprio sistema. A informação computacional, proveniente tanto do meio quanto do próprio corpo, pode ser convertida em informação representacional, e esta, por sua vez, pode retroalimentar o sistema, modificando os padrões computacionais e ajustando novas interpretações. A mente consciente, assim, não apenas processa e armazena informação, mas reconstrói e reconfigura continuamente as suas representações, de modo a gerar um sentido de coerência experiencial ao longo do tempo. Esta capacidade de descodificação recursiva deve-se à estrutura do sistema nervoso, que não apenas integra informação de múltiplas fontes, mas também a reinterpreta a partir da própria atividade representacional gerada. Por outras palavras, a mente consciente não se limita a receber e projetar estados computacionais do meio e do corpo, mas influencia ativamente os estados computacionais futuros através da reelaboração contínua da informação.

Ao contrário de sistemas artificiais que podem processar informação de forma sofisticada, mas sem este mecanismo de decodificação recursiva, a mente consciente não apenas simula processos computacionais, mas constrói representações que interagem com a própria estrutura informacional do sistema que as gera. Isto sugere que a consciência não é um simples reflexo da informação processada, mas um fenómeno emergente que surge precisamente porque há um ciclo dinâmico de transformação da informação, onde a computação e a representação interagem de maneira constante.

Neste sentido, a mente consciente pode ser vista como um sistema recursivo de decodificação, em que a informação computacional do corpo e do meio não apenas é interpretada, mas continuamente reinterpretada dentro do próprio sistema que a gera. Esta característica pode explicar a singularidade da consciência biológica, diferenciando-a de sistemas computacionais convencionais, que operam sem essa dinâmica de decodificação retroativa e integrativa.

A decodificação de informação do corpo

António Damásio raramente utiliza o conceito de informação no sentido em que o estamos a utilizar aqui. No entanto, as suas ideias assentam na noção de que determinados processos transmitem informação, numa ótica semelhante à que adotamos neste trabalho. Damásio sugere que muitos dos nossos comportamentos e experiências fenomenológicas manifestam e representam vontades ocultas de seres vivos que nos antecederam, nomeadamente organismos unicelulares como bactérias e células, cujos comportamentos primitivos foram, primeiro, preservados e imitados, e depois reorganizados pelos sistemas nervosos sob a forma de emoções e sentimentos (Damásio, 2010).

Esta capacidade de “auscultar” o corpo deve-se a um tipo especial de célula: o neurónio. Ou seja, a capacidade dos neurónios de processar e representar informação permitem decodificar e representar a informação das outras células do corpo. Assim, os neurónios possibilitam “auscultar as vontades” das células e, dessa forma, criar uma representação (mais sofisticada, muitas vezes em forma de sentimentos) daquilo que elas “reclamam”. Em síntese, os circuitos neuronais imitam representacionalmente as forças internas implícitas do resto do corpo e transformam-nas em imagens conscientes sentidas (Damásio, 2010, p. 60). Por conseguinte, os sistemas nervosos, ao adquirirem a capacidade de representar comportamentos automáticos e, em grande medida,

flexibilizá-los, permitiram uma vantagem adaptativa. A mente consciente, assim, possibilita representar (pensar e sentir) comportamentos que, antes do surgimento dos sistemas nervosos, eram cegos e automáticos. No fundo, os comportamentos eficientes desenvolvidos pelos organismos que nos antecederam começaram a ser mapeados pelos sistemas nervosos quando estes entraram em cena, o que permitiu o surgimento de comportamentos mais sofisticados, ancorados em planos de ação mais elaborados, como as emoções.

Como analisado no capítulo 1, António Damásio é um dos autores que mais ostensivamente defende que a mente consciente não depende exclusivamente do cérebro, mas também do corpo, nomeadamente dos sentimentos de emoção que se passam no teatro do corpo, e que são o resultado de processos químicos nos quais as células desempenham um papel central (Damásio, 2003). Segundo Damásio (2003) e outros autores (Seth, 2021; Barrett, 2004; Barrett e Simmons, 2015; Bechara e Naqvi, 2004), são os processos interoceptivos que permitem que determinada informação do corpo se *transforme* em sentimentos. Quando um sujeito *sente* uma emoção (ex. alegria), é porque determinada informação *neurofisiológica* (ou computacional) do corpo causou determinada informação *fenomenológica* (ou representacional). Um determinado mapa cerebral assume um padrão neurofisiológico específico, e esse padrão gera a experiência consciente da emoção. É como se os sinais do corpo (ex. libertação de dopamina e noradrenalina) fossem transformados em imagens mentais de alegria, motivação e excitação (Damásio, 2010).

Anil Seth (2021), na mesma linha que Damásio, propõe que a interocepção não é um processo meramente passivo de deteção dos estados corporais, mas um mecanismo ativo de previsão. O cérebro antecipa estados fisiológicos futuros com base em modelos internos e compara essas previsões com os sinais reais que recebe, ajustando continuamente a regulação do organismo. Essa dinâmica explica por que a experiência emocional pode variar entre indivíduos, uma vez que diferentes histórias de vida e estados internos influenciam a forma como os sinais interoceptivos são interpretados. Barrett e Simmons (2015) aprofundam esta ideia, sugerindo que a precisão com que um sujeito processa os sinais do corpo está diretamente relacionada com a clareza e intensidade das suas emoções. Indivíduos com maior acuidade interoceptiva tendem a experienciar estados emocionais de forma mais diferenciada e intensa, o que reforça a ideia de que a

consciência emocional não depende apenas da atividade cerebral, mas da forma como o cérebro representa os estados internos do corpo. Além disso, a interocepção está intrinsecamente ligada à regulação homeostática (Bechara e Naqvi 2004). Sensações como fome ou sede não são meramente respostas metabólicas, mas experiências fenomenológicas emergentes da decodificação dos sinais corporais pelo cérebro. De forma análoga, estados emocionais como a alegria não surgem isoladamente: podem resultar da satisfação de um estado fisiológico favorável, como a ativação de circuitos de recompensa e a manutenção da homeostase dopaminérgica.

Neste sentido, Damásio sugere que interocepção não apenas regula o organismo, mas também desempenha um papel central na construção da identidade consciente (2003, 2010 e 2013). A percepção contínua dos estados internos do corpo é um dos principais alicerces do sentido do eu, permitindo ao sujeito experimentar a continuidade da sua existência ao longo do tempo. Esse mecanismo reforça a ideia de que a consciência não é apenas uma representação do mundo externo, mas também uma interpretação dinâmica dos estados internos do organismo, construída a partir de um processo recursivo onde a informação computacional do corpo é convertida em representações fenomenológicas. Não há aqui nenhum dualismo no pensamento de Damásio. O físico não causou o não-físico. O físico-neurofisiológico causou o físico-fenomenológico – ou o neurobiológico causou o neuropsicológico.

Nos modelos analisados existem, portanto, determinados sistemas interoceptivos que decodificam sinais do corpo, transformando-os em sentimentos. Quando, por exemplo, uma pessoa torce um pé, o corpo emite uma série de sinais para o cérebro e este cria a experiência de dor. A experiência da dor pode ser entendida como a decodificação de informação computacional proveniente da neurofisiologia alterada do corpo. A lesão dos tendões do pé, ao causar um processo inflamatório, é transformada na experiência da dor do pé. O mesmo seria dizer que a informação computacional-física X (do corpo) causou a informação representacional-física X (da mente consciente) – um processo que analisaremos mais à frente com maior detalhe. No nosso modelo, a informação representacional é precisamente isso: a representação consciente das computações do corpo, ou seja, a transformação de estados computacionais internos em conteúdo experiencial. Esta capacidade de transformar computações em representações é uma das marcas distintivas dos sistemas biológicos conscientes, permitindo que alguns

organismos não apenas reajam ao mundo, mas o representem simbolicamente a partir das suas próprias condições fisiológicas.

Em suma, a interocepção, um fenómeno tão bem estudado e analisado nas neurociência, pode ser visto como a capacidade de determinados sistemas neuronais transformarem sinais “duros” do corpo em representações mentais conscientes.

O córtex pré-frontal como centro de recursividade informacional

O neurocientista Stanislas Dehaene, no âmbito do seu modelo da consciência — o *global neuronal workspace* (GNW) —, analisado no primeiro capítulo, propõe que a consciência emerge da integração de informações oriundas de diferentes regiões cerebrais especializadas (visuais, auditivas, somatossensoriais, entre outras), as quais são amplificadas e partilhadas num espaço global de trabalho neuronal (ou GNW⁷⁵). Segundo esta teoria, a consciência resulta da coordenação de redes neurais de longa distância que ligam o córtex frontoparietal a outras áreas do cérebro. Esse mecanismo permite que determinadas informações sejam amplificadas na GNW mas mantidas ativas e disseminadas entre diferentes sistemas cerebrais, promovendo a acessibilidade global do conteúdo consciente (Dehaene & Changeux, 2011; Dehaene, 2014).

O córtex frontoparietal funciona, neste modelo, como um gestor que seleciona os estímulos mais relevantes, filtra os *inputs* sensoriais e integra-os em re-representações unificadas e coerentes. Essas informações, uma vez selecionadas, são amplificadas e sustentadas no espaço global (na GNW), permitindo que sejam usadas em funções cognitivas complexas, como tomada de decisão, memória de trabalho e planeamento (Dehaene, 2021). Além disso, o córtex frontoparietal coordena a comunicação entre diferentes sistemas especializados do cérebro, garantindo que a experiência consciente seja rica e funcional.

Estudos empíricos reforçam esse papel. Tarefas com estímulos subliminares mostram que a ativação do córtex pré-frontal é crucial para que informações se tornem conscientes, enquanto padrões de sincronização neural entre o córtex frontal e outras áreas cerebrais são associados à experiência consciente (Dehaene et al., 2006). Por outro lado, danos ao córtex pré-frontal levam a déficits em memória, atenção e controle

⁷⁵ *Global Network Workspace*.

executivo (Gazzaniga, 2011; Damásio, 1993/2011). Assim, o córtex pré-frontal é considerado o núcleo executivo do espaço global de trabalho, transformando processos inconscientes fragmentados em experiências conscientes coesas.

Na nossa teoria, a recursividade informacional é o mecanismo pelo qual informação computacional é transformada em informação representacional e esta, por sua vez, pode ser reconvertida em computações úteis para o sistema. De acordo com a GNW, parece-nos que córtex frontoparietal pode ser o local ideal para a transformação bidirecional que propomos, pois é aqui que ocorre 1) a integração de *inputs* sensoriais e internos. O córtex pré-frontal recebe informações de várias áreas do cérebro, incluindo estímulos sensoriais do córtex occipital e temporal, estados internos regulados pelo sistema límbico e dados de memória provenientes do hipocampo. Essas informações chegam como bits computacionais, ou seja, padrões neurais que sinalizam características físicas do ambiente ou do corpo. Ocorre também 2) transformação de computações em representações conscientes. Quando essas informações são integradas, amplificadas e decodificadas no espaço global de trabalho, tornam-se acessíveis como informação representacional. A experiência consciente surge neste processo, onde os bits computacionais, antes fragmentados, são transformados em padrões coesos e unificados, que podem ser percebidos como imagens mentais, pensamentos ou sensações. Por fim, ocorre 3) a recursividade representacional-computacional. A nossa teoria sugere que, após a geração de uma representação consciente, essas informações podem ser novamente codificadas como informação computacional. No córtex frontoparietal, essa realimentação informacional manifesta-se em processos como A) planeamento, isto é, representações conscientes de um cenário futuro (informação representacional) que são transformadas em cálculos computacionais que coordenam ações; e B) memória de trabalho, isto é, representações temporárias de informações relevantes que são mantidas no córtex pré-frontal em forma de computações, permitindo que sejam manipuladas e transformadas de acordo com as demandas da tarefa. Por exemplo, ao perceber um objeto ameaçador, o córtex frontoparietal traduz os estímulos visuais (bits de luz processados no córtex visual) numa representação consciente do perigo. Essa representação consciente é, então, recursivamente processada para acionar respostas computacionais no sistema motor e emocional, como fuga ou alerta.

A GNW defende também que o córtex frontoparietal, em colaboração com outras redes (como a rede de modo padrão, a rede de modo executivo e a rede de saliência), desempenha um papel causal na geração da consciência, ao amplificar e sustentar informações no espaço global (Dehaene, 2014). Este mecanismo pode ser interpretado como uma manifestação da recursividade informacional que sugerimos. Na nossa teoria, a consciência emerge quando informações computacionais do ambiente e do corpo são traduzidas em representações conscientes (do meio e do corpo) no espaço global. Essas representações conscientes, por sua vez, informam novas computações no sistema, criando um ciclo contínuo de transformação e recursividade.

Um dos grandes benefícios da recursividade informacional que propomos é a sua capacidade de gerar flexibilidade cognitiva. O córtex frontoparietal desempenha um papel central nesse processo, pois permite: 1) simulação mental, ou seja, representações conscientes que podem ser manipuladas computacionalmente para criar cenários hipotéticos, algo essencial para tomada de decisão e planeamento; 2) aprendizagem adaptativa, isto é, representações conscientes de erros ou sucessos que podem ser reprocessadas como informações computacionais para ajustar comportamentos futuros; 3) controle emocional, quando informações representacionais sobre estados internos são transformadas em computações que ajudam a regular emoções e respostas adaptativas.

Na nossa visão, a consciência não é, portanto, apenas o resultado de um processamento linear, mas sim um fenómeno recursivo, onde o sistema (o cérebro) tem acesso a si mesmo por meio de transformações contínuas de informação. Segundo a GNW, o córtex frontoparietal é um dos locais mais importantes para essa autorreferência informacional, pois gera representações explícitas do estado do corpo e do ambiente, usa essas representações para modular ações e pensamentos, num ciclo de *feedback* contínuo, e integra experiências passadas e expectativas futuras num fluxo informacional coeso, permitindo que o sistema aprenda e se adapte constantemente.

No nosso modelo, a consciência não é, portanto, simplesmente o resultado de um processamento mas sim de um ciclo contínuo de transformações recursivas de informação, onde o sistema tem acesso a si mesmo e utiliza esse *feedback* para enriquecer seu funcionamento. O córtex frontoparietal desempenha um papel essencial nesse processo, gerando representações explícitas do estado do corpo e do ambiente, modulando

ações e integrando experiências passadas e expectativas futuras. Assim, o córtex frontoparietal funciona, no nosso modelo, como o principal centro de recursividade informacional, permitindo que a consciência emergja como um fenómeno dinâmico e funcional no qual computações geram representações conscientes. Estas representações não só enriquecem o processamento computacional, como também influenciam decisões e comportamentos de forma adaptativa e eficiente. Este assunto será analisado com mais detalhe mais adiante, quando propusermos uma solução para o problema do epifenomenismo da consciência.

3.4.2. *Subconsciente computacional e subconsciente representacional*

A ideia de que a mente corresponde a um espaço imagético, abrangendo tanto o consciente quanto o inconsciente, foi profundamente influenciada pelo conceito freudiano de subconsciente (Searle, 1998). Freud defendia que a mente inconsciente é habitada por todas as imagens mentais capazes de se tornarem conscientes. Ou seja, os estados mentais, tanto conscientes quanto inconscientes, existem como tal; a única diferença entre eles reside no momento em que uma imagem mental, inconsciente, se torna consciente. Para Freud, os estados mentais permanecem inalterados, aguardando o momento em que se manifestam conscientemente (Tallis, 2017). Nesse sentido, a consciência de um estado mental seria a simples percepção consciente de um estado mental inconsciente que, em determinado momento, se torna acessível à consciência. Uma analogia, proposta por John Searle em *A Redescoberta da Mente* (1998), ilustra bem essa ideia freudiana:

“É como se os estados mentais inconscientes fossem móveis guardados no sótão da mente, e, para trazê-los à consciência, subíssemos ao sótão e lançássemos sobre eles o projetor da nossa percepção” (p. 200).

Segundo esta perspectiva, os objetos armazenados na “escuridão do nosso sótão” continuam a manter a sua forma, mesmo quando não são visíveis – ou seja, quando não são conscientes. O subconsciente freudiano, neste caso, é mental, não neuronal.

Uma alternativa contemporânea a essa ideia – totalmente alinhada com a nossa teoria – é a visão de que o inconsciente corresponde a mapas neuronais, que formam ou causam imagens mentais conscientes (Eagleman, 2012; Damásio, 2010; Gazzaniga,

2011). Por outras palavras, as imagens conscientes são causadas por mapas neuronais que operam de maneira subconsciente, mas que geram essas imagens quando ativados. Damásio compara esse processo aos painéis eletrónicos, onde padrões são desenhados pelos elementos luminosos ativos ou inativos. De maneira semelhante, a distribuição desses elementos pode ser rapidamente alterada, o que resulta em diferentes representações – por exemplo, um quadrado, um cão, uma árvore (Damásio, 2010, p. 94). Embora a atividade neuronal seja incessante, apenas uma parte dessa atividade se transforma em imagem mental.

Neste enquadramento, e segundo o nosso modelo, muitos padrões neuronais possuem potencial representacional, mas apenas alguns se tornam efetivamente conscientes. Isso acontece quando a informação computacional é transformada em informação representacional – ou seja, quando um padrão cartográfico se converte num padrão mental. Dessa forma, as representações tornam-se conscientes quando bits de informação computacional se transformam em bits de informação representacional; e tornam-se subconscientes quando bits de informação representacional são convertidos de novo em informação computacional.

A proposta que aqui defendemos sugere, ainda, que o subconsciente é composto por dois tipos distintos de processos: (1) o subconsciente computacional, e (2) o subconsciente representacional.

Por 1) subconsciente computacional entendemos o conjunto de processos neuronais que operam de forma automática e aos quais não temos, nem podemos ter, qualquer acesso consciente. Trata-se de uma maquinaria funcional de base, responsável por processos como a mielinização dos neurónios, a propagação de impulsos nervosos ou a regulação de circuitos elementares, os quais permanecem estruturalmente inacessíveis à experiência subjetiva, independentemente do seu grau de complexidade. Do mesmo modo, o funcionamento intrínseco da maquinaria neuronal — como a atividade dos axónios, dendritos, sinapses e a libertação de neurotransmissores — bem como o funcionamento de outros sistemas fisiológicos, como a digestão, a respiração ou o batimento cardíaco, permanece estruturalmente fora do domínio da consciência, operando de forma automática e não representacional.

Podemos conceber este subconsciente computacional como um vasto e inacessível campo de processamento, composto por milhares de milhões de sinapses, conexões neuronais e mecanismos fisiológicos que operam em rede — desde o batimento cardíaco à digestão, da respiração à regulação hormonal. Trata-se de uma maquinaria basal que assegura a homeostasia e a adaptação do organismo ao meio, realizando cálculos, ajustes e respostas automáticas ao ambiente. Toda esta atividade corresponde a informação computada, funcional e vital, mas estruturalmente inacessível à consciência, uma vez que não é representada sob forma subjetiva. A sua função é manter o “sistema operativo”, e não ser conhecida por ele. (Eagleman, 2012; Gazzaniga, 2011; Tallis, 2017).

Sobre este assunto, Daniel Dennett argumenta que a introspeção não nos oferece imagens objetivas do nosso cérebro, mas uma narrativa. As computações que chegam até nós, aparecem-nos como representações que nunca refletem fielmente o estado real do cérebro. Nas suas palavras:

"não vemos, não ouvimos, nem sentimos a complicada maquinaria neural a trabalhar no nosso cérebro – temos de nos contentar com uma versão interpretada, digerida, uma ilusão do utilizador, que nos é tão familiar que não só a tomamos como realidade, mas como a mais indiscutível e intimamente conhecida realidade de todas" (Dennett, in Humphrey, 2023, p. 136).

Isto significa que, embora tenhamos representações de pensamentos e sentimentos, não temos imagens do funcionamento intrínseco do sistema que os gera. Ou seja, não sabemos nem controlamos a forma como os pensamentos e sentimentos são gerados. Essencialmente, sabemos *que* os temos mas não sabemos *como* é que os temos. Alguns autores afirmam que esse tipo de mecanismo serve para "nos proteger de nós próprios" (Eagleman, 2012, p. 213). Isto é, o cérebro é projetado de forma a impedir que tenhamos acesso introspetivo à complexa maquinaria que mantém as nossas funções vitais eficientes. Precisamente por isso não pode ser acedido – não é representável – e é mantido fora da nossa perceção, funcionando como uma espécie de *firewall* implementado pelo sistema (Meijer, 2012 e 2015).

Para ilustrar a nossa ideia, podemos pensar em processos como o batimento cardíaco, a digestão ou a respiração. Nenhum desses processos se transforma em imagens representacionais, exceto em forma de sentimentos – sentimentos fisiológicos, como a perceção de um batimento cardíaco ou de uma inspiração; ou sentimentos

homeostáticos, como dor ou fome (Damásio, 2012). No caso do batimento cardíaco, por exemplo, os ventrículos do coração contraem e bombeiam o sangue, mas não temos consciência de como isso ocorre, nem representações claras do processo. O que chega até nós, é a *sensação* de batimento cardíaco, se estivermos atentos a isso – não a maquinaria que está a fazer bater o coração. O mesmo se aplica ao processo digestivo, que envolve uma série de movimentos musculares e interações neuronais, sem que tenhamos representações diretas de como esse processo acontece. Apenas tomamos consciência de uma ínfima parte desses sistemas quando ocorre um desvio face ao seu funcionamento esperado — como no caso da dor cardíaca ou da sensação de azia. Podemos ter uma imagem mental do estado do nosso organismo – em forma de sentimentos – como parte do mecanismo de regulação homeostática, mas nunca temos acesso às imagens relacionadas aos mecanismos intrínsecos desses sistemas (Damásio, 2002, 2012; Ledoux, 1999; Eagleman, 2012).

Temos, isso sim, no âmbito do nosso modelo, acesso a algumas representações de certos efeitos computacionais, mas nunca às computações em si, nem aos códigos subjacentes que as geram. Consideramos que a informação relativa ao funcionamento interno do corpo e do cérebro não necessita de ser representada conscientemente para garantir a sobrevivência do organismo. Precisamente por isso, não é, nem deve ser, representável. Como menciona, no livro *Incógnito* David Eagleman, "O nosso cérebro possui conhecimento de algo que a nossa mente não consegue aceder explicitamente" (2012, p. 67), e é uma sorte que isso seja assim, pois "não estaríamos conscientes da maior parte das atividades em curso no nosso cérebro, nem gostaríamos de estar – isso interferiria com os processos cerebrais já bem oleados" (2012, p. 67).

Esta ideia é reforçada por Joseph LeDoux, que, em *O Cérebro Emocional*, sublinha que a vida seria psicologicamente insuportável se tivéssemos acesso direto à complexa maquinaria responsável pela regulação do corpo. Nas suas palavras:

"é uma sorte que não estejamos conscientes destes processos [de regulação homeostática], pois estaríamos tão focados em executá-los que nunca chegaríamos a compreender nada, se tivéssemos de fazer tudo isso com uma concentração deliberada" (1999, p. 33).

Em suma, a nossa teoria sugere, como vimos, que existe uma informação ontológica no Universo que não é representável (a informação ontológica estrutural), e de

maneira análoga, propomos que existe uma informação computacional sobre o funcionamento interno do corpo que não é representável. Chamamos a isso *subconsciente computacional* – todo o processamento neurofisiológico dedicado, não só a manter a homeostasia, como a gerar pensamentos, sentimentos e subjetividade. É um tipo de informação intrínseca ao funcionamento do sistema nervoso, à qual não temos qualquer acesso, nem sequer temos o mais ínfimo conhecimento sobre como opera dentro de nós.

O 2) *subconsciente representacional*, por sua vez, abrange toda a informação computacional que *pode ser transformada* em imagens mentais. Se o subconsciente computacional é a maquinaria estrutural que funciona para nos manter vivos, o subconsciente representacional corresponde às representações que podem emergir dessa maquinaria e que podem tornar-se conscientes em determinados momentos. Dito de outra forma, enquanto o subconsciente computacional envolve informação que não é representável — exceto sob a forma de sentimentos (Damásio, 2003) —, o subconsciente representacional, em contraste, refere-se a toda a informação computacional que pode ser convertida em representações conscientes, embora essa informação frequentemente permaneça no subconsciente sob a forma de bits computacionais. Este tipo de subconsciente, ao contrário do subconsciente computacional, não está relacionado à maquinaria neuronal *per se*, mas sim às imagens potenciais que a atividade neuronal pode gerar.

Por exemplo, a crença de que “Lisboa fica em Portugal” constitui uma representação de uma informação computacional que pode estar em estado consciente ou inconsciente. Está consciente quando a mantemos em mente, mas torna-se inconsciente quando já foi gerada e enraizada no sistema, permanecendo sob a forma de bits computacionais, sem que tenhamos uma percepção clara da sua presença. A representação, neste caso, pode existir em forma de imagem potencial (bits computacionais), aguardando o contexto necessário para se tornar uma representação consciente (bits representacionais), como num painel eletrónico, conforme a analogia proposta por Damásio, atrás referida. O subconsciente representacional é, portanto, composto pelas inúmeras representações que a informação computacional pode gerar, e não pela maquinaria que sustenta a criação dessas representações. Estas representações existem inicialmente num estado computacional no cérebro, mas podem adquirir uma forma

representacional e, assim, tornar-se conscientes mediante a interação entre o cérebro, os objetos do mundo, as experiências e os contextos vividos (Damásio, 2010).

3.5. Uma solução para o problema do conhecimento

O problema do conhecimento refere-se, em grande medida, ao conjunto de questões filosóficas relacionadas com a forma como o sujeito conhece, compreende e interpreta a realidade. Envolve, entre outros aspetos, a origem, a natureza e os limites do conhecimento, bem como a relação entre o sujeito cognoscente e o objeto conhecido.

No que respeita a esta temática, a tradição filosófica racionalista diz-nos que o sujeito possui ideias inatas – ideias que, nascidas com o ser humano, manifestam a realidade de uma forma clara e genuína (Hottot, 2003 e 2003a; Descartes, 1637/2008). Para René Descartes, um determinado objeto físico era conhecido porque uma ideia inata desse objeto físico permitia o seu conhecimento; contudo, ele nunca conseguiu explicar satisfatoriamente como uma substância imaterial — a alma ou mente — poderia, não só perceber um objeto físico, como interagir com um corpo físico (1641/1996). Outros racionalistas, como Leibniz, procuraram resolver esta questão ao defender que as ideias inatas não eram conteúdos finalizados na mente, mas predisposições que organizam a experiência (Leibniz, 1704/1996). Ainda assim, a natureza dessas predisposições permanece indefinida, não sendo claro se correspondem a estruturas mentais, funcionais ou apenas a esquemas lógicos de organização.

Em contraste com a tradição racionalista, a tradição filosófica empirista (Locke, 1689/2014; Hume, 1739/2016) defende que o conhecimento é adquirido exclusivamente pela experiência sensorial, rejeitando a noção de ideias inatas. Para David Hume, por exemplo, a mente humana não possui conteúdos inatos, mas apenas percepções derivadas da experiência, distinguindo-se entre impressões, mais vívidas e imediatas, e ideias, que são cópias menos intensas dessas impressões (Hume, 1739/2016). No plano ontológico, contudo, a dificuldade permaneceu: como pode o sujeito formar uma representação mental de um objeto físico externo? Hume levou o empirismo a um extremo ao questionar a própria noção de causalidade, sugerindo que não existe uma ligação

objetiva necessária entre causa e efeito, mas apenas uma associação habitual entre eventos, baseada na repetição. Este ceticismo quanto à fundamentação do conhecimento causou um abalo profundo na filosofia moderna, ao colocar em dúvida a possibilidade de alcançar conhecimento universal e necessário.

Em contraste com o racionalismo e o empirismo, Kant (1781/2008) defendeu que o conhecimento resulta da articulação entre a experiência sensorial (sensibilidade) e a capacidade racional do sujeito para representar simbolicamente o mundo. Para além disso, propôs que a mente não recebe passivamente os dados sensoriais, mas impõe-lhe uma estrutura *a priori* através das categorias do entendimento — como as noções de espaço, tempo e causalidade — que tornam a experiência inteligível. Esta conceção, que rejeita o realismo direto e afirma o papel estruturador das faculdades cognitivas, encontra hoje amplo respaldo nas neurociências contemporâneas (Seth, 2021; Dehaene, 2014; Metzinger, 2009; Piaget, 1950/1971; Varela et al., 1991). De certa forma, esta síntese kantiana antecipou, aliás, questões que viriam a ser exploradas pela filosofia contemporânea e pelas ciências cognitivas. Jean Piaget (1950/1971) desenvolveu a teoria do construtivismo, segundo a qual o conhecimento resulta da interação entre sujeito e ambiente, sendo construído ao longo do desenvolvimento cognitivo. Mais recentemente, as teorias da cognição incorporada sugerem que o conhecimento não é apenas um produto da mente individual, mas resulta da interação entre sujeito, corpo e meio (Varela et al., 1991).

Assim, apesar das diferenças entre racionalistas, empiristas e a própria perspectiva de Kant, subsiste uma dificuldade comum a todos estes pensadores: a de explicar a mediação entre o objeto externo e a representação mental interna. Ou seja, como é que o sujeito adquire, representa ou interpreta as propriedades de um objeto? Seja por via da razão, da experiência ou da combinação de ambas como sugere Kant, permanece incerta a forma como se estabelece a ponte entre o mundo e a mente, e de que modo as propriedades dos objetos se tornam acessíveis à consciência.

Immanuel Kant (1781/2008) propôs, como vimos, que a realidade em si — o *númeno* — está para além do alcance da nossa percepção e conhecimento direto. Segundo ele, apenas temos acesso aos fenómenos, isto é, à realidade tal como é estruturada pelas formas *a priori* da nossa sensibilidade e pelas categorias do entendimento. Qualquer tentativa de aplicar categorias fenomenológicas ao *númeno* — como causalidade ou

interação — seria, para Kant, uma extrapolação ilegítima. Espinosa (1677/1992), porém, oferece uma visão ontológica diferente: defende que tudo o que existe é expressão de uma única Substância, da qual tanto os objetos físicos como os seres humanos são modos. Nesta concepção, não há separação entre realidade e conhecimento, e a interação entre diferentes entidades não levanta qualquer problema, uma vez que todas partilham a mesma natureza ontológica. Assim, ao contrário de Kant, Espinosa admite que a mente humana pode conhecer a realidade em si, na medida em que ela própria é uma expressão da Substância.

3.5.1. *Três princípios para o problema do conhecimento*

Na esteira de Espinosa, sugerimos que é a universalidade de uma única substância unificada que permite a interação ontológica entre o ser humano — ou qualquer outro sistema capaz de processar informação — e as restantes entidades do Universo. Sugerimos adicionalmente que 2) o conhecimento é possível devido à recursividade informacional que os cérebros permitem; e que 3) a luz é um interface entre o objeto e o observador, que permite o conhecimento visual.

1) A primeira (e principal) razão pela qual o conhecimento é possível reside na ontologia da SUI. Esta ontologia, enquanto tecido que sustenta a rede que conecta todas as partes do Cosmos, permite que a informação seja partilhada e transferida em todo o Universo. A informação não é uma propriedade exclusiva de qualquer sistema ou objeto específico; é a essência subjacente que torna possíveis todas as interações. Ou seja, a universalidade da SUI assegura que todas as coisas, objetos e seres vivos partilhem a mesma ontologia, o que viabiliza o relacionamento e a interação entre eles. Deste modo, ao nosso modelo, o Universo não é composto por entidades independentes, mas por uma rede interconectada, onde todas as partes partilham uma ontologia informacional comum, que se desdobra em processos emergentes (Bohm, 1980). Esta interação entre as partes do Universo é mediada por uma matriz informacional que não apenas transmite dados sobre as propriedades do ambiente, mas também possibilita a construção de fenómenos complexos a partir do processamento de informação entre as entidades constituintes do Cosmos.

Assim, a SUI não é apenas um meio abstrato, mas um sistema que permite que as entidades dentro do Universo comuniquem as suas propriedades umas às outras. Isso implica que a estrutura do Cosmos é informacional, no sentido de que tudo o que existe está constantemente a trocar e a processar informação sobre a sua natureza e comportamento. Esta matriz não se limita a ser um canal de transmissão passiva de informação, mas tem um papel ativo na emergência de fenómenos complexos. Isso implica que os processos físicos e as interações no Universo não são apenas somas de partes isoladas, mas surgem a partir de um tecido informacional subjacente, onde a comunicação entre entidades permite a emergência de novos estados e estruturas.

Essa visão conduz-nos naturalmente a uma consequência fundamental. Cada elemento do Universo – desde partículas, átomos, moléculas até sistemas mais complexos, como organismos biológicos com sistemas nervosos – faz parte dessa rede causal que processa, transfere e armazena informações em diversas formas. Como dissemos, o bit é a linguagem do Universo, pelo que todas as propriedades que forem bits conseguem comunicar com outras propriedades que partilham a mesma “linguagem”. Isso significa que o conhecimento não emerge de uma correspondência direta entre um sujeito e o mundo físico, mas sim da interação do sujeito com os padrões informacionais que emergem no ambiente. O Universo, portanto, opera como um sistema informacional fluido e a troca de informações entre suas partes é fundamental para a criação de consciência e conhecimento.

Através desta perspetiva, resolvemos um dos problemas centrais da epistemologia kantiana: a relação entre o observador e o objeto. Enquanto Kant argumentava que o sujeito só pode conhecer o fenómeno (a realidade tal como se manifesta aos seus sentidos) e não o númeno (a realidade em si), sugerimos que essa limitação pode ser superada se considerarmos que o Universo interage consigo mesmo através da SUI. Se tanto o objeto X como o observador Y são modos da mesma substância informacional, então a interação entre eles não é uma questão de representação externa, mas sim de processamento e transformação de informação dentro do mesmo sistema unificado. O sujeito não está separado do objeto por um véu intransponível; antes, ele próprio faz parte da mesma matriz informacional (Rovelli, 2018). Dessa forma, o conhecimento não é apenas uma projeção interna do sujeito sobre um mundo incognoscível, como defendia Kant (1781/2008), mas sim um processo ativo de interação e integração da

informação que já existe no tecido fundamental da realidade. A percepção não ocorre como um simples reflexo subjetivo, mas sim como um processamento legítimo da estrutura real da informação no Universo.

A 2) segunda razão pela qual o conhecimento é possível reside na capacidade do sistema nervoso de transformar, de forma recursiva, informação computacional em informação representacional. Este mecanismo permite que o cérebro não seja apenas um processador de sinais físicos, mas também um sistema capaz de interpretar e reorganizar esses sinais em representações mentais significativas. Deste modo, a informação computacional, que consiste em padrões de atividade neuronal gerados a partir de estímulos sensoriais como luz, som ou toque, é decodificada pelo cérebro em estados subjetivos que formam a base da experiência consciente. Como tudo é informação computacional — desdobramento da informação estrutural —, as especificidades do sistema nervoso conferem-lhe a capacidade de decodificar essa informação em informação representacional. Nesse sentido, não existe qualquer véu entre o observador e o meio, mas sim uma interação sistêmica que, no caso dos sistemas nervosos, permite a emergência de um novo tipo de informação: a informação baseada em imagens.

Esta transformação não ocorre de forma linear ou estática, mas sim num processo contínuo e dinâmico, onde a informação representacional, uma vez gerada, pode ser reintroduzida no sistema computacional, influenciando novas computações e criando ciclos de refinamento da percepção e do pensamento. Esse caráter recursivo permite não só a construção de imagens mentais a partir da informação sensorial imediata, mas também a capacidade de evocação e manipulação dessas representações ao longo do tempo. A informação representacional, uma vez armazenada na memória em forma de computações, pode ser recuperada, reinterpretada e combinada com novos dados sensoriais, permitindo a formulação de conceitos, inferências e abstrações.

De forma específica, sugerimos que o sistema nervoso decodifica informação do meio e do corpo ao transformar bits de informação computacional em informação representacional. Este processo é dinâmico e recursivo, permitindo uma representação contínua do meio e do corpo e, ao mesmo tempo, possibilitando que a informação representacional seja também reconvertida em computações internas no próprio sistema. É este ciclo permanente que sustenta não só o surgimento da mente e da consciência,

mas também outros processos cognitivos associados à linguagem, à memória e ao simbolismo.

Neste ponto, podemos relacionar a nossa proposta com a ideia de neurociência preditiva (Friston, & Kiebel, 2009), segundo a qual o cérebro não apenas processa passivamente informação do meio, mas constrói ativamente modelos internos que antecipam padrões na realidade. Se o conhecimento fosse apenas um espelho passivo do mundo, a percepção seria uma experiência fragmentada e desorganizada. No entanto, sabemos que o cérebro consegue estruturar a informação em representações coerentes e previsíveis, o que sugere que a mente é um sistema de inferência ativa dentro da SUI, processando e refinando continuamente a informação que recebe. Desta forma, o conhecimento emerge não apenas da capacidade de receber e processar informação do ambiente, mas também da faculdade de transformar continuamente essa informação em representações internas que estruturam a nossa compreensão do mundo. Esse processo está na base da cognição e da consciência, garantindo que a experiência sensorial não seja um fluxo informe de dados, mas sim um conjunto estável de percepções e significados que orientam a nossa interação com a realidade.

Além disso, esta abordagem responde a outra questão epistemológica fundamental: como podemos ter certeza de que as nossas representações correspondem à realidade? No kantismo tradicional, a percepção do mundo é condicionada por estruturas cognitivas humanas, o que levanta a possibilidade de que nunca acedemos à realidade em si. Somos governados por uma realidade fenomenal que, embora nos seja dada como única via de acesso ao mundo, permanece uma construção mediada pelas formas transcendentais do sujeito — e, por isso, nunca corresponde à realidade em si. No entanto, se a SUI for a base informacional comum de todos os processos no Universo, então a estrutura da informação representacional não é arbitrária, mas uma transformação legítima da estrutura ontológica subjacente. Isto significa que a relação entre sujeito e objeto não é uma mera projeção mental do primeiro sobre o segundo, mas sim uma interação recursiva e legítima dentro da própria estrutura do Cosmos.

Assim, enquanto Kant postulava um fosso intransponível entre o númeno e o fenómeno — defendendo que o sujeito apenas tem acesso aos fenómenos, e nunca à coisa em si —, sugerimos que a ontologia informacional da SUI dissolve essa barreira ao mostrar que o conhecimento é uma consequência direta da interação informacional

dentro de um sistema unificado. Onde Kant vê um fosso epistemológico, nós vemos uma relação ontológica. Ou seja, enquanto para Kant o mundo é sempre fenomenal e permanece inevitavelmente refém dessa fenomenalidade — já que o sujeito só pode conhecer os objetos tal como aparecem sob as formas da sensibilidade e as categorias do entendimento —, na nossa proposta, o mundo, sendo ontologicamente informacional, permite ao sujeito um contacto direto com essa realidade. Ainda que, no nosso modelo, a percepção seja uma representação interpretativa da informação, trata-se apenas de um diferente tipo de informação, e não de um desvio em relação ao real. Ao contrário de Kant, sugerimos que a fenomenalidade não governa o nosso entendimento, pois o sistema nervoso opera com dois tipos de informação — computacional e representacional — que se distinguem apenas no aspeto funcional, e não na sua ontologia.

Assim, no nosso modelo, não existe qualquer barreira entre o sujeito e o mundo (Rovelli, 2018); pelo contrário, ambos interrelacionam-se de forma ubíqua, uma vez que fazem parte da mesma substância unificada de informação. Tal como Kant, reconhecemos que o sujeito não tem acesso direto à realidade última em si. No entanto, ao contrário da sua proposta transcendental, concebemos essa limitação não como uma separação estrutural entre sujeito e objeto, mas como uma diferença funcional entre tipos de informação. O sujeito não está separado do mundo, mas ligado a ele por um contínuo informacional, no qual os sistemas nervosos têm a capacidade singular de transformar informação computacional em informação representacional. Deste modo, o sujeito, na nossa perspetiva, não representa o mundo apenas de acordo com categorias de entendimento à priori, mas interpreta-o e constrói-o segundo uma lógica relacional intrínseca.

A luz como interface do conhecimento

3)A terceira razão pela qual o conhecimento é possível está no papel da luz como uma interface informacional fundamental — e esta ideia, pela sua densidade conceptual e implicações ontológicas, merece ser analisada de forma autónoma.

Os vastos trabalhos nas neurociências da percepção demonstram atualmente, de forma clara, que o cérebro possui sistemas especializados na codificação de objetos do meio, nomeadamente cor, profundidade, forma, relevo, textura — sistemas esses que resultam de uma engenharia biológica extraordinária, desenvolvida ao longo de milhares de milhões de anos, continuamente adaptada para otimizar a interação dos

organismos com o ambiente (Palmer, 2020; Frith & Frith, 2020; Hohwy, 2013). Ou seja, os cérebros foram moldados pela evolução para que se adaptem o melhor possível ao mundo real.

Neste contexto, uma das descobertas mais relevantes na área da visão e da neurociência é a constatação de que o sujeito não acede diretamente ao objeto em si, mas apenas a uma representação construída a partir de estímulos sensoriais processados pelo cérebro. Existe, assim, uma mediação inevitável entre o objeto e o observador. Neste ponto, Kant parece ter antecipado, com notável acerto, que a consciência não apreende a realidade em si mesma (númeno), mas apenas o fenómeno — a realidade tal como aparece ao sujeito. De facto, o sujeito não vê o objeto diretamente, mas apenas a luz que reflete ou transporta as suas propriedades (Palmer, 2020; Frith & Frith, 2020). No entanto, como referimos, consideramos a proposta kantiana epistemologicamente incompleta, na medida em que o mundo dos fenómenos, ancorado em estruturas cognitivas a priori, configura uma barreira entre a realidade e a mente — uma separação que tendemos a rejeitar. A nossa proposta sustenta que, embora as representações não sejam réplicas do real, estão enraizadas numa dinâmica relacional ontológica entre sujeito e mundo. Assim, o ser humano não está separado da realidade, mas em contacto com ela através de uma mediação informacional, que, embora representacional, é sustentada por uma ligação ontológica efetiva entre o sistema cognitivo e o mundo. Trata-se de uma realidade representada, sim — mas representada a partir da realidade, e não contra ela. A luz representa, nesta perspetiva, uma função ubíqua na aquisição do conhecimento do mundo. É através dela que se estabelece uma mediação informacional, que, embora representacional, é sustentada por uma ligação ontológica efetiva entre o sistema cognitivo e o mundo.

Vejamos com mais pormenor. Quando observamos algo, os padrões de luz refletem-se no objeto e atingem os nossos olhos, onde são focados pela córnea e pelo cristalino na retina, uma camada de células sensíveis à luz no fundo do olho. Na retina, os fotorreceptores (cones e bastonetes) captam os fotões da luz. Esses fotorreceptores contêm moléculas de rodopsina (nos bastonetes) e fotopsinas (nos cones), que sofrem alterações químicas ao absorverem a luz. Essa alteração desencadeia uma cascata bioquímica que culmina num sinal elétrico. Esse sinal elétrico é depois transmitido às células bipolares e, em seguida, às células ganglionares da retina. Estas possuem axónios

que se juntam para formar o nervo ótico, que conduz os sinais ao cérebro. O nervo ótico encaminha a informação para o tálamo, especificamente para o núcleo geniculado lateral, que a envia para o córtex visual primário, no lobo occipital. É aqui que os sinais começam a ser organizados em categorias como contraste, cor, movimento e profundidade. Posteriormente, a informação segue para áreas visuais secundárias e terciárias, onde é integrada com dados sensoriais de outras modalidades e com memória. É nesse processo de integração e processamento que o cérebro constrói uma representação consciente do objeto (Hohwy, 2013). Em suma, vemos objetos e o mundo à nossa volta graças à luz.

A luz é uma onda eletromagnética que transporta energia sob a forma de campos elétricos e magnéticos oscilantes, podendo também ser descrita como um fluxo de fótons. As suas propriedades — frequência, amplitude, polarização e fase — contêm informação sobre os objetos com que interage (Palmer, 2020). Quando a luz incide sobre um objeto, algumas dessas propriedades são modificadas em função das características materiais da superfície. Por exemplo, a frequência da luz refletida determina a cor percebida, enquanto a intensidade está relacionada com a quantidade de fótons recebida. A direção e a difração da luz fornecem pistas sobre a forma e a textura, permitindo que o cérebro construa uma representação visual detalhada do ambiente.

Contudo, apesar do conhecimento acumulado sobre a ótica, a visão e a luz, persistem questões profundas por esclarecer. Desde logo, se apenas vemos luz — e não o objeto em si — como é que conseguimos, de facto, perceber o objeto? Tal como indicado, entre nós e o objeto existe sempre uma “cortina” de luz que nunca é levantada no ato da percepção. Se a luz cessa, a percepção cessa também. Se a luz é o único estímulo que alcançamos, e não o objeto propriamente dito, como consegue a luz carregar informações fiéis sobre esse objeto? O que há na luz que permite representar propriedades físicas do mundo material? Não há qualquer evidência de que a luz “escolha” carregar informação — esse papel parece emergir da interação entre as propriedades da luz e os sistemas que a interpretam, como o cérebro. Neste sentido, a informação transportada pela luz parece já estar estruturada de modo a ser interpretável por sistemas nervosos. A luz não carrega um conjunto arbitrário de dados brutos. Pelo contrário, as suas propriedades emergem de padrões físicos consistentes que, ao interagirem com o

ambiente, sofrem alterações específicas, tornando-se *legíveis* por sistemas biológicos especializados.

Sugerimos que a estrutura da luz, ao refletir-se, refratar-se ou difratar-se, gera padrões informacionais que codificam propriedades do objeto — e é essa codificação que a consciência consegue decifrar. Esses padrões óticos não são aleatórios, mas resultam de interações físicas regulares entre a luz e a matéria, traduzindo características objetivas — como cor, forma, textura ou relevo — em variações estruturadas do sinal luminoso. Ao incidir sobre o sistema visual, esses padrões são convertidos em atividade neural, preservando a organização interna do estímulo. O que a consciência decifra, portanto, não é a luz em si, mas a organização interna dessa codificação — isto é, o modo como as propriedades do objeto ficaram inscritas nos padrões estruturais do estímulo luminoso. Por outro lado, conceitos como polarização, intensidade ou frequência são, em última instância, construções descritivas que utilizamos para medir e modelar fenômenos. Estes conceitos podem não corresponder à essência última da luz, mas funcionar como ferramentas cognitivas que captam certos aspectos do seu comportamento no nosso modelo da realidade. São úteis porque permitem prever interações e manipular tecnicamente a luz — em lentes, lasers, fibras óticas. No entanto, é plausível que essas propriedades sejam representações físicas de algo mais profundo. O que poderá estar por detrás desses conceitos? Propomos, então, que a luz é, antes de mais, *uma interface de informação* — uma portadora de padrões informacionais. Ao concebermos o Universo como um tecido informacional, a luz pode ser entendida como uma manifestação de informação em movimento. As suas propriedades, como a polarização e a frequência, funcionam como formas de codificação dessa informação — e não como qualidades fixas.

No plano quântico, os fótons que compõem a luz não têm propriedades intrínsecas. Estas só emergem da interação da luz com outros sistemas, sugerindo que as suas características são, na realidade, relações contextuais. Assim, a luz pode ser encarada como uma expressão física de informação em trânsito, cujas propriedades revelam um fenómeno informacional puro — uma projeção codificada de uma realidade mais profunda. Termos como polarização ou frequência seriam, então, formas simbólicas de representar informação fundamental, que o cérebro é capaz de decodificar. A capacidade da luz de viajar sem se degradar rapidamente torna-a ideal para transportar esses

padrões. Segundo Max Tegmark, os fótons podem ser vistos como bits de informação que codificam e transportam dados fundamentais sobre a realidade física. Para Tegmark, "o nosso sistema visual interpreta padrões matemáticos contidos na luz, convertendo-os em experiências sensoriais" (2014, p. 266), o que reforça a ideia de que a luz não é apenas um fenômeno físico, mas também um código fundamental de informação.

Neste enquadramento, a luz não é apenas uma onda energética — é uma mensagem informacional, um meio que transmite dados estruturados sobre o universo à medida que se propaga. Sempre que interage com um objeto, a luz codifica as suas propriedades. Como viaja à velocidade da luz, essas propriedades chegam aos nossos sentidos de forma praticamente instantânea. O universo, enquanto SUI, está em permanente transformação, e a luz constitui um dos principais canais através dos quais essa transformação se torna *percecionável*. A luz seria, assim, uma expressão física da informação computacional do Universo em trânsito. Funciona como meio físico de transporte de informação, mas é o cérebro que converte essa informação computacional em informação representacional.

A partir daqui, surge uma questão inevitável: como conseguimos recordar-nos de um objeto se apenas vimos a luz que ele refletiu? Se a luz, por si só, fosse suficiente para representar o objeto, como explicar a capacidade de evocá-lo quando já não está presente? Isso indica que a luz tem de transportar algo representável — algo que possa ser codificado e armazenado. Se a luz fosse apenas algo desprovido de estrutura informacional, não deixaria qualquer traço na nossa mente. Defendemos, por isso, que os bits de informação refletidos e transportados pela luz são *transformados*, no cérebro, em informação eletroquímica. Quando vemos uma maçã vermelha, recebemos bits de luz que são processados e convertidos em impulsos neuronais — bits de informação representacional. Estes bits são depois armazenados computacionalmente na memória, permitindo evocar, mais tarde, a imagem da maçã.

Assim, a nossa capacidade de representar o ambiente depende da competência do sistema nervoso para enraizar *bits* informacionais originados na luz. Esta carrega informação intrínseca que o cérebro descodifica. Tal como uma planta transforma bits de luz em oxigénio, o ser humano transforma bits de luz em representações. A luz atua, portanto, como interface entre o mundo externo e a experiência interna, atualizando as nossas representações mentais à medida que o ambiente se transforma. É o meio inicial

de transmissão da informação, mas é o cérebro que, através de processos de interpretação e codificação, constrói representações estáveis e duradouras do mundo.

Além de ser um veículo para a informação, a luz conecta sistemas informacionais distintos — o objeto e o observador, o mundo computacional e o mundo representacional. Ela é fundamental na estrutura da informação no Universo, não apenas como uma onda eletromagnética, mas como um meio de transformação de informação entre diferentes níveis de realidade. Neste contexto, a luz atua como uma ponte entre a ordem implicada (a realidade profunda, não observável) e a ordem explicada (a realidade perceptível), permitindo que a informação subjacente se revele na forma de uma experiência observável. Portanto, a luz pode ser vista como uma interface entre a informação estrutural (a realidade profunda e incognoscível) e a informação computacional (a forma observável e processada), tornando visível aquilo que é "invisível" e permitindo a percepção consciente. Esse processo de transformação e manifestação da informação, mediado pela luz, conecta diferentes camadas da realidade, refletindo a ideia de que a luz é um mecanismo de transição e revelação dentro da ontologia informacional do universo.

A luz funciona, portanto, como um fator de desdobramento ao transformar a informação computacional processada pelos sistemas nervosos. Dito de outra forma, a luz é uma manifestação da informação ontológica computacional que permite desvendar e conhecer a realidade através de representações perceptíveis da luz. Por meio da luz, as interações informacionais que governam o comportamento da matéria tornam-se manifestações observáveis, que podem ser decodificadas e representadas mentalmente. Assim, a luz desempenha o papel de revelar e expandir a informação oculta do Universo para um formato que podemos entender e perceber (um formato físico).

Neste enquadramento, a luz não é apenas um fenômeno físico mediador da percepção, mas um elo real e contínuo entre sujeito e objeto — uma expressão dinâmica da relação ontológica que ambos mantêm enquanto manifestações da SUI. Esta visão contrasta profundamente com a proposta kantiana, onde o sujeito está epistemologicamente separado do mundo por estruturas a priori que impõem as condições do fenômeno e impedem o acesso ao númeno. Ao contrário de Kant, não concebemos a mediação como uma barreira entre dois domínios irreduzíveis (realidade e mente), mas como um processo de transdução informacional que revela, embora sob forma representacional, aspectos reais do mundo. A luz, nesse sentido, não encobre — revela. E a

representação consciente que dela resulta não é uma distorção da realidade, mas a sua continuação funcional num outro formato.

Em síntese, sugerimos que a luz deve ser entendida como uma interface informacional entre a realidade física e a experiência representacional: um canal de desdobramento da informação estrutural em informação computacional e, por fim, em informação representacional. Assim, a luz não é apenas um meio energético, mas um veículo organizado de padrões informacionais que o cérebro consegue decodificar, permitindo o acesso — ainda que mediado — às propriedades do mundo. Esta perspetiva reforça a nossa proposta de que o Universo é, na sua base, uma estrutura informacional, sendo a luz uma das manifestações mais evidentes dessa ontologia em trânsito entre computação e representação.

3.6. Uma solução para o problema mente-corpo

*“A mente e o corpo são uma só e a mesma coisa,
que ora se concebe sob o atributo do pensamento,
ora sob o atributo da extensão.”*

Espinosa

Já apresentámos algumas ideias sobre como o cérebro pode criar uma mente consciente. A capacidade dos sistemas nervosos em transformar recursivamente informação parece-nos ser um fator crucial para compreender o surgimento da consciência. Segundo esta perspetiva, a mente consciente não é uma entidade separada ou um fenómeno inexplicável, nem tão-pouco uma propriedade não-física ontologicamente distinta do cérebro, mas sim o resultado da integração, processamento e representação de *bits* de informação. A fenomenologia, neste sentido, pode ser interpretada como um duplo aspeto da informação computacional. Por um lado, é o processamento funcional de dados no cérebro e, por outro, é a emergência de experiências subjetivas a partir desse mesmo processamento.

Ao assumirmos o cérebro informacional como origem da mente consciente, o problema da interação mente-cérebro é ultrapassado. De facto, ao defendermos que

tudo o que existe é constituído por *bits* de informação, esse problema é transposto, e o dualismo de substância ou de propriedades deixa de ser necessário. No modelo da SUI, o cérebro cria a mente consciente porque a *computacionalidade da informação neuronal* permite a *emergência da representacionalidade da informação neuronal*; da mesma forma que, por exemplo, a *computacionalidade da informação física* do H₂O permite a *emergência da computacionalidade da informação química* da água. Dito por outras palavras, do mesmo modo que a informação química ou biológica — enquanto formas de informação computacional — se transforma e circula entre sistemas, também essa informação pode, nos sistemas nervosos, dar origem a informação representacional. Mais do que isso, o cérebro gera a mente consciente porque contém instruções computacionais que orientam esse processo. Inicialmente, essas instruções são puramente informacionais e, posteriormente, tornam-se físicas, químicas e biológicas, estando codificadas em estruturas como o ADN (Ribeiro, 2013).

Este princípio de emergência — segundo o qual a organização da informação dá origem a novas propriedades — é observável em múltiplas estruturas da realidade, desde os sistemas subatômicos até aos sistemas nervosos, justamente porque todo o Universo computa informação. Sugerimos, neste sentido, que os fenómenos emergentes resultam de determinados padrões informacionais que coordenam — ou estruturam — essas emergências. Sabemos, por exemplo, que dois átomos de hidrogénio e um de oxigénio formam a molécula chamada água. Sabemos, também, que a organização molecular da água, nomeadamente determinados padrões físicos e químicos permitem que esta se manifeste em três estados diferentes: líquido, gasoso e sólido. Ou seja, sabemos que uma determinada organização molecular X leva à água em estado gasoso, a organização molecular XY à água em estado líquido, e a organização XYZ à água em estado sólido, mas o que existe de inerente a essas interações que permite essas transições? No entanto, embora possamos descrever esses padrões físicos e químicos, ainda não compreendemos como e por que razão essas propriedades emergem especificamente dessas configurações. Deve haver algo subjacente a essas partículas — um conjunto de regras informacionais precisas — que permite não apenas a formação da molécula de água, mas também a manifestação dos seus diferentes estados (Davies, 2019).

Sugerimos a mesma solução para o problema mente-corpo. Tal como na composição da água, sabemos que determinados sistemas neuronais processam a linguagem

(Saur et al, 2008), a concentração (Corbetta e Shulman (2002), a percepção (Humphrey, 2022) a coordenação motora (Silva e Fonseca, 2011), as emoções e os sentimentos (Damasio, 2003 e Ledoux, 1999), entre muitas outras funções. Além disso, sabemos que alterações computacionais no cérebro causam alterações fenomenológicas (ou representacionais). Por exemplo, mudanças químicas ou elétricas induzidas por substâncias psicoativas (Veríssimo, 2012; Silva, 2020) mostram que a experiência subjetiva depende criticamente da química cerebral, refutando abordagens dualistas que ignoram essa relação.

No entanto, tal como acontece com a molécula de água, o problema fundamental persiste: se conseguimos descrever algumas correlações entre atividade cerebral e experiência consciente, ainda não sabemos como sistemas neuronais físicos podem dar origem a estados fenomenológicos. Por conseguinte, assim como não compreendemos inteiramente como a molécula de água origina os seus diferentes estados físicos, também não sabemos, em rigor, como redes de neurónios produzem consciência. A verdadeira questão no problema mente-cérebro permanece, portanto, esta: como é que estados físicos podem gerar experiências subjetivas?

3.6.1. *A recursividade e especificidade dos neurónios*

Para responder a este problema, sugerimos 1) que os neurónios não são meras estruturas físico-químicas, mas unidades de processamento que contêm uma codificação informacional específica — bits de informação intrínsecos — cuja organização e dinâmica tornam possível a emergência da mente e da consciência. Adicionalmente, propomos 2) que os neurónios operam de forma recursiva, permitindo a transformação da informação computacional em informação representacional, e vice-versa.

Propomos, assim, que os neurónios originam a mente através da organização dos seus bits de informação, seguindo um princípio de emergência semelhante ao que observamos noutros fenómenos naturais. Tal como a combinação de dois elementos altamente reativos — o sódio (um metal explosivo) e o cloro (um gás tóxico) — dá origem ao sal de cozinha, com propriedades completamente novas como o sabor salgado e a solubilidade, também a organização específica dos neurónios pode gerar propriedades que não existem nas suas partes isoladas. O mesmo sucede com a molécula de água, já aludido: os átomos de hidrogénio e oxigénio, quando configurados numa determinada

proporção, originam uma substância com propriedades singulares, como a fluidez, a tensão superficial ou o comportamento “anômalo” nos seus diferentes estados físicos. De forma análoga, sugerimos que a mente emerge da complexa estrutura informacional dos sistemas neuronais, à semelhança de como outras propriedades emergem a partir da organização da matéria. Assim, tal como bits de informação a nível atómico originam estruturas moleculares (como a água ou o sal), e como a organização desses bits moleculares permite a formação de sistemas celulares, também os bits de informação neuronal, quando organizados segundo determinadas dinâmicas, podem dar origem a estados representacionais — isto é, à mente e à consciência.

Partimos do pressuposto, que nos parece inegável, que apenas animais com neurónios possuem mentes — só estes podem transformar informação computacional em informação representacional — pelo que os neurónios devem ter algum tipo de informação diferente das restantes células. Ao contrário das outras células, os neurónios podem processar, integrar e transmitir informação (Damásio, 2010; Tononi e Koch, 2015; Dehaene, 2014), o que permite a emergência de faculdades, como a linguagem (Saur et al. 2008), controlo motor (Adolph et al, 2003) e atenção (Corbetta e Shulman (2002), entre outras. Ou seja, a capacidade de os neurónios *transformarem* informação, permite que determinados sinais químicos possam ser transformados em linguagem, locomoção e foco.

No entanto, sugerimos que *alguns* neurónios possuem um duplo aspeto devido às particularidades dessas células em *processarem, integrarem, transmitirem e representarem* informação. Essa *dualidade funcional* permite que alguns tipos de neurónios realizem duas funções essenciais: por um lado, processam informação computacional (*bits* de informação computacional), que corresponde à manipulação de dados neuronais em circuitos específicos; por outro, criam informação representacional (*bits* de informação representacional), que emergem como imagens mentais e, em alguns seres, como sentimentos e experiências conscientes. A interação neuronal entre esses dois níveis de informação (computacional e representacional) é crucial para a emergência, primeiro da mente e, depois, da consciência. Os bits computacionais fornecem a infraestrutura necessária, enquanto os bits representacionais atribuem significado aos dados processados, permitindo que sistemas nervosos avancem de respostas reflexas para respostas mais flexíveis e, em alguns casos, para cognição complexa e representações

conscientes do eu. Estes processos, críticos para a emergência da mente e da consciência, podem ser encontrados em dois sistemas neuronais particulares, que analisaremos um pouco mais à frente.

Para já, importa ilustrar esta dualidade funcional com um exemplo em torno da mente consciente. Quando vemos um objeto, os neurónios da retina convertem luz em impulsos elétricos (informação computacional), enquanto áreas do córtex visual transformam esses impulsos numa representação mental do objeto (informação representacional). Essa dualidade funcional dos neurónios — a capacidade de transformarem informação computacional em informação representacional — reforça a nossa distinção entre informação computacional e representacional, mostrando como ambas são partes de um único processo informacional. Ao mesmo tempo, evidencia que os neurónios atuam como pontes entre a ontologia da informação computacional e a epistemologia da informação representacional, desempenhando um papel central na transformação de dados “brutos” em experiências fenomenológicas. Além disso, a capacidade dos neurónios de operar simultaneamente nesses dois níveis de informação desafia a tradicional separação entre mente e corpo, sugerindo que ambos são expressões complementares de uma única SUI. Como sugerem Tononi e Koch (2016), a integração da informação nos circuitos neuronais é a base para a formação de experiências conscientes, destacando a relevância dessa dualidade funcional no estudo da mente.

A mente consciente, no âmbito da nossa proposta, não é, assim, apenas um estado estático de representação, mas um processo dinâmico que evolui no tempo. Este dinamismo é sustentado pela capacidade do cérebro de integrar *bits* de informação representacional de forma contínua, criando uma linha temporal de experiências. Ou seja, a continuidade da experiência pode ser explicada pelo correr de um "programa" interno, uma estrutura funcional que permite a coerência temporal e a sensação de identidade ao longo do tempo. Este “programa” pode ser entendido como um mecanismo neurobiológico recursivo, baseado em circuitos específicos que garantem a transição fluida de estados mentais, evitando que a experiência se fragmente em percepções descontínuas e desconexas. Este “programa” pode ser entendido como um mecanismo neurobiológico recursivo, baseado em circuitos específicos que garantem a transição fluida de estados mentais, evitando que a experiência se fragmente em percepções descontínuas e desconexas. A recursividade aqui referida diz respeito à capacidade do sistema neuronal

de reutilizar a informação representacional anteriormente integrada em computações — memórias, estados afetivos, representações do corpo e do meio — como base para gerar novas representações, de forma contínua e autorreferente. Em termos funcionais, esse “programa” atua como uma arquitetura de processamento iterativo, na qual cada estado mental (computacional e representacional) influencia o seguinte, permitindo uma atualização constante da experiência consciente. Por conseguinte, a consciência, quando emerge, não surge como um conjunto disperso de percepções isoladas, mas como um fluxo integrado e sequencial, que mantém a coerência entre diferentes momentos da experiência subjetiva. Esse “programa”, em certo sentido, assegura que a mente consciente não seja um mero agregado de estados momentâneos, mas uma estrutura que confere um sentido narrativo à experiência do eu.

Dentro da neurociência, algumas teorias já propõem que a continuidade da consciência resulta da atividade coordenada de redes neuronais específicas. O modelo do espaço de trabalho global (GNW) (Dehaene, 2014; Baars, 1998), por exemplo, sugere que a consciência emerge quando múltiplas regiões do cérebro sincronizam a sua atividade, permitindo a transmissão de informação entre sistemas especializados. De forma semelhante, a teoria da integração da informação (IIT) (Tononi, 2004; Tononi & Koch, 2016) propõe que a experiência consciente só é possível porque há um sistema que organiza e interliga informação de modo a gerar um estado global unificado. Este mecanismo explicaria por que razão o fluxo da consciência não é percebido como uma sequência fragmentada de instantes, mas sim como um processo contínuo e interdependente.

Assim, na nossa proposta, podemos entender a mente consciente como um sistema informacional autorregulado, que funciona como um algoritmo de continuidade, garantindo que as representações de um instante não se percam no instante seguinte. Esse “programa” não apenas mantém uma linha de continuidade entre percepções e memórias passadas, como também antecipa estados futuros, permitindo previsões e ações baseadas em experiências anteriores. É esta capacidade de previsão e integração da informação temporal que dá à consciência o seu carácter dinâmico e adaptativo. Ao mesmo tempo, este dinamismo permite que o cérebro represente bits de informação computacional do meio e do corpo, que transforma em representações conscientes. Esta capacidade permite a construção do eu, uma narrativa unificadora que dá sentido à

experiência subjetiva. Sem este mecanismo, a consciência seria apenas um conjunto de percepções isoladas, sem continuidade ou coerência interna. A mente consciente atua, assim, como um sistema que processa e decodifica informações do meio e do organismo, ao mesmo tempo que antecipa cenários futuros, conferindo ao organismo uma vantagem adaptativa significativa (Damásio, 2010; Tononi & Koch, 2016; Humphrey, 2023; Dehaene, 2014).

Dessa forma, sugerimos que a experiência unificada do eu e da realidade emerge porque a mente consciente não é apenas um estado representacional passivo, mas um processo computacional estruturado, regido por informação que mantém a continuidade, a coerência e a antecipação da experiência. Este “programa” não só permite a consciência de momentos individuais, mas liga-os entre si, conferindo ao organismo a percepção da continuidade existencial e uma experiência de identidade ao longo do tempo. A capacidade de alguns neurónios de transformar informação entre diferentes domínios informacionais explica a riqueza da experiência humana.

Se os neurónios apenas processassem informação computacional, sem que esta fosse convertida num nível representacional, não existiriam nem mente, nem — posteriormente — consciência fenomenológica. Existiria apenas um processamento automatizado de estímulos. Do mesmo modo, se não houvesse a capacidade inversa — a de transformar representações conscientes em informação computacional —, não seríamos capazes de agir com base nos nossos pensamentos, memórias ou emoções. Ou seja, os neurónios são mediadores fundamentais da transferência e conversão de informação, sendo responsáveis pela emergência de fenómenos como memória, linguagem e sentimentos, bem como da re-representação desses estados mentais, o que possibilita a consciência. O seu papel recursivo permite que a informação não permaneça estática dentro do sistema nervoso, mas seja continuamente transformada, criando um fluxo dinâmico entre o que é processado computacionalmente e o que é representado fenomenologicamente.

Alguns autores (Dehaene, 2014; Ledoux, 2020) consideram que as diferenças na densidade e conectividade dos neurónios no córtex pré-frontal entre os seres humanos e os outros animais são notáveis e contribuem para as capacidades cognitivas superiores

dos seres humanos, nomeadamente para a experiência consciente⁷⁶. Sabemos, de facto, que, comparativamente a outras espécies, o córtex pré-frontal humano apresenta uma proporção muito maior do cérebro total (cerca de 29% no ser humano, 17% nos chimpanzés e apenas 12,5% nos cães). Conforme argumentado, a capacidade particular dos neurónios do córtex frontoparietal em decodificar informação do corpo e do meio (Dehaene, 2014) pode explicar porque animais com córtex pré-frontal, como os mamíferos, podem ter experiência consciente. Para além disso, como referido, a maior densidade e conectividade que se verifica no córtex pré-frontal nos seres humanos em relação aos outros mamíferos pode explicar, em parte, porque a nossa experiência é mais rica e porque a experiência de um chimpanzé pode ser mais rica que a de um cão (Lévy, 2023; Dehaene, 2014; Ledoux, 2020).

No mesmo sentido, outros autores (Seeley et al. 2006; Allman et al, 2011), destacam a importância dos neurónios Von Economo (VENs) que se podem encontrar na *rede de saliência* (córtex insular e córtex cingulado) dos humanos e em alguns mamíferos mais evoluídos, como os grandes símios, os golfinhos, as baleias e os elefantes. Vários estudos têm vindo a demonstrar que a presença destes neurónios está associada a cérebros de grande dimensão e comportamentos sociais complexos, sugerindo um papel na mediação de processos cognitivos avançados como a consciência, a empatia e a capacidade de sentir alguns sentimentos de emoção “superiores”. Estudos indicam, por exemplo, que os VENs podem estar envolvidos na comunicação rápida entre diferentes regiões do cérebro, facilitando respostas rápidas a estímulos sociais e emocionais. A sua distribuição limitada a certas espécies sugere que estes neurónios podem ter evoluído para atender às necessidades de processamento cognitivo em cérebros maiores e mais complexos.

Tendo descrito a função estrutural e funcional destes neurónios do ponto de vista empírico, propomos agora uma interpretação informacional, em que tais neurónios assumem uma função recursiva — isto é, são capazes de transformar a informação que processam em novas representações, e essas representações podem, por sua vez,

⁷⁶ Embora o córtex pré-frontal não seja exclusivo dos humanos, sua proporção e complexidade são distintivas. Por exemplo, no ser humano, o córtex pré-frontal representa aproximadamente 29% do cérebro, enquanto em chimpanzés essa proporção é de cerca de 17%, e em cães, aproximadamente 12,5%.

ser reprocessadas computacionalmente. Dentro da nossa teoria da recursividade neuronal, a consciência emerge precisamente da capacidade do cérebro para operar sobre a informação de maneira iterativa, convertendo informação computacional (dados brutos processados pelos neurónios) em informação representacional (experiência consciente, *qualia*, simbolismo e memória). Consideramos que, por um lado, a maior densidade e interconectividade dos neurónios do córtex pré-frontal permite que o espaço de trabalho dos seres humanos seja mais desenvolvido, o que pode explicar uma maior capacidade para transformar informações computacionais em imagens mentais conscientes. As especificidades neuronais dos seres humanos neste sistema podem explicar, em parte, porque os seres humanos conseguem *re-representar* estados cognitivos e emocionais, e não apenas *representar* informação do meio e do corpo. Por outro lado, a presença dos VENs na rede de saliência sugere que estes neurónios podem ser decodificadores privilegiados da informação computacional do corpo, atuando como mediadores essenciais no processo de tornar a informação acessível à consciência. O processamento de informação corporal através da interoção — que ocorre, em grande parte, na rede de saliência — pode, aliás, ajudar a explicar por que razão determinados estados fisiológicos (Damásio, 2010) dão origem a sentimentos, nomeadamente através do papel desempenhado pelos VENs nesse processamento⁷⁷. Neste sentido, se estes neurónios são responsáveis por integrar diferentes tipos de informação de forma eficiente e rápida, podem ser vistos como peças fundamentais no mecanismo que transforma atividade neuronal (computacional) em experiência subjetiva (representacional), nomeadamente sinais químicos do corpo em sentimentos (Seeley et al. 2006; Allman et al, 2011; Damásio, 2003 e 2010).

Por conseguinte, consideramos que os neurónios do córtex pré-frontal e os VENs possuem uma dualidade funcional, que lhes permite não só computar informação, como *representar* informação. Ou seja, a densidade e conectividade dos neurónios do córtex pré-frontal bem como a especificidade dos VENs, permitem que os neurónios trabalhem com uma dupla função: computar informação e representar informação. A capacidade

⁷⁷ Importante referir que, para além dos neurónios do córtex frontoparietal e dos VENs, também os neurónios do tronco cerebral parecem funcionar de forma recursiva, computando e representando informação. O tronco cerebral é particularmente importante para os primeiros mapas neuronais do corpo, funcionando como uma primeira ponte entre o estado fisiológico do corpo e os sentimentos que daí emergem, críticos para a formação de o proto-eu (Damásio, 2010)

destes neurónios em decodificar informação do corpo e do meio permite que surjam processos re-representativos, ou seja, capacidades de autorreflexão e de autoconsciência. Sugerimos que, se a consciência pode ser explicada através da decodificação de informação computacional em representacional e se apenas alguns mamíferos possuem, por um lado, maior densidade cortical frontal e, por outro, VENs, então a presença destes neurónios pode ser um fator necessário para o surgimento da consciência complexa, como a que encontramos nos seres humanos e, em menor grau, noutros mamíferos evoluídos, como os chimpanzés, os golfinhos, as baleias e os elefantes (Humphrey, 2023).

De facto, a consciência, enquanto fenómeno cognitivo altamente integrado, manifesta-se sobretudo em mamíferos e, dentro destes, com maior sofisticação nas espécies que possuem maior densidade cortical frontal e VENs, como os primatas, cetáceos e elefantes. Isto sugere que a emergência da mente consciente não é um epifenómeno, mas sim um processo dependente de determinadas arquiteturas neurais especializadas, onde o córtex pré-frontal e os VENs da rede de saliência desempenham um papel fundamental. Este alinhamento entre a recursividade neuronal no córtex pré-frontal e nos VENs pode fornecer um modelo pertinente para explicar porque a consciência só emerge em certos organismos e como a informação computacional se transforma em experiência subjetiva.

Para além da especificidade de alguns neurónios para o surgimento da mente consciente, sugerimos que é necessário um *feedback* recursivo. Isto significa que a consciência não apenas representa informações externas ou internas, mas também pode representar as suas próprias representações. Esse processo recursivo é central para a emergência da autoconsciência. Na nossa perspetiva, a recursividade emerge porque os bits de informação representacional processados pelos neurónios do córtex pré-frontal e os VENs podem ser codificados novamente em informação computacional e, subsequentemente, reprocessados pelo sistema nervoso. Isto cria um ciclo informacional que permite ao sujeito refletir sobre os seus próprios pensamentos, emoções e ações (Gazzaniga, 2011; Seth, 2021).

Por exemplo, ao tomarmos uma decisão e, posteriormente, avaliarmos o seu impacto, o sistema nervoso está a criar uma camada adicional de representações que inclui o próprio ato de decidir. Esta capacidade de representar as suas próprias ações

não é apenas um processo epifenoménico, pelo contrário, é um mecanismo informacional que aumenta a flexibilidade e a capacidade de adaptação dos organismos conscientes. Este *feedback* recursivo não seria possível sem a base informacional que sustenta a mente consciente, demonstrando como a nossa teoria unifica a ontologia da informação com as capacidades fenomenológicas (Barrett, 2004; Damásio, 2003).

Consciência de primeira e de segunda ordem

Como observámos no Capítulo 1, David Rosenthal propõe o modelo da consciência *higher-order thought* (HOT) – ou pensamento de ordem superior –, segundo o qual a mente consciente surge quando um sistema mental re-representa representações pré-existentes (2005 e 2008). Esta abordagem está alinhada, em certa medida, com a GNW de Dehaene (2014), na medida em que ambas enfatizam a necessidade de um nível adicional de processamento para que uma experiência se torne consciente. Enquanto Rosenthal sugere que a consciência emerge quando há uma metarrepresentação dos estados mentais, permitindo que o sujeito tenha uma percepção autorreflexiva sobre o que sente e pensa, Dehaene propõe que a consciência depende da transmissão da informação para um espaço de trabalho global, onde pode ser amplamente distribuída e integrada.

Os VENs, por seu turno, fazem parte da rede de saliência, que, como vimos, é composta principalmente pela ínsula e pelo córtex cingulado anterior – sistemas críticos no processo da interocepção e da homeostasia. A principal função desta rede é detetar e priorizar estímulos internos e externos que sejam relevantes para o organismo, regulando a atenção e a resposta emocional. A ínsula, em particular, desempenha um papel fundamental na gestão dos processos homeostáticos do corpo. Através da integração de sinais provenientes dos sistemas fisiológicos, a ínsula gera representações internas do estado do corpo, que são posteriormente interpretadas sob a forma de sentimentos e estados emocionais. Este processo decorre da interocepção, a capacidade de um organismo *sentir* e monitorizar o seu próprio estado interno, permitindo a percepção de sensações como fome, sede, temperatura, dor ou bem-estar. Ou seja, a ínsula recebe sinais químicos do corpo e transforma esses sinais em sentimentos de fome, sede, nojo, alegria, medo, etc. (Seeley et al. 2006; Allman et al, 2011; Damásio, 2010 e 2000).

Sugerimos que estas representações de primeira ordem da rede de saliência, a par de outras processadas no tronco cerebral (Damásio, 2010) sustentam uma *consciência de primeira ordem*, que se pode encontrar, de certa forma, nos mamíferos mais evoluídos, como os referidos acima. Ou seja, os VENs da rede de saliência, parecem desempenhar um papel crucial no surgimento da consciência de primeira ordem (Edelman, 1995) facilitando a rápida transmissão de informação entre o corpo e a ínsula, o córtex cingulado e outras regiões envolvidas no surgimento da consciência. Assim, estes neurónios podem estar diretamente relacionados com a amplificação da consciência de primeira ordem, permitindo que os estados fisiológicos do organismo sejam, não apenas regulados, mas também percebidos subjetivamente e integrados na experiência consciente em forma de sentimentos.

A capacidade do córtex pré-frontal em representar essas representações de primeira ordem (Rosenthal, 2008; Dehaene, 2014; LeDoux, 2020) permite um nível superior de processamento, dando origem à consciência reflexiva, à linguagem e ao pensamento abstrato característicos da espécie humana. Esta re-representação, que podemos considerar sobretudo cognitiva, é fundamental para a emergência da autoconsciência, pois possibilita que o indivíduo não apenas tenha experiências subjetivas, mas também pense sobre essas experiências, avalie estados internos e antecipe consequências futuras. Além disso, a interligação entre o córtex pré-frontal e outras áreas cerebrais, como a rede de saliência permite um processamento hierárquico da informação, onde estados corporais e mentais são continuamente representados, ajustados e integrados numa narrativa coerente do eu.

Em grande medida, sugerimos que os VENs permitem uma *consciência de primeira ordem* (que poderíamos designar de consciência emocional) e o córtex frontoparietal permite uma *consciência de segunda ordem* (que poderíamos designar de consciência cognitiva). Enquanto a primeira possibilita o *sentimento* de estar vivo – crucial para qualquer estado subjetivo – a segunda permite a reflexão, percepção e *conhecimento* sobre estar vivo. Diria que a consciência humana está dependente da integração deste dois “tipos” de consciência, contudo, não devemos minorizar a importância da consciência de primeira ordem, não só para os *qualia* como para o surgimento de uma consciência sentida, capaz de dar força à identidade de um eu que, para saber que está vivo, tem que *sentir* que está vivo (Damásio, 2010; Humphrey, 2023)

Por conseguinte, enquanto o tronco principal fornece um primeiro sentimento de existência (um proto-eu nas palavras de Damásio, 2010), a rede de saliência contribui para um alargamento da consciência ao amplificar a gama e a intensidade dos sentimentos que o determinado organismo tem ou sente, o que é crítico para o surgimento dos *qualia*. Como referido, a consciência deste nível (de primeira ordem) é, sobretudo, visceral, ancorada na capacidade de o organismo sentir (mais do que saber que está vivo). Por seu turno, a consciência de segunda ordem permite re-representar cognitivamente a consciência de primeira ordem, o que confere ao organismo capacidade, não só de refletir sobre esses estados, como ainda desenvolver autoconsciência sobre o que está a pensar e a sentir.

Estas ideias que acabamos de expor não são, em grande parte, originais. Vários autores sugerem que o cérebro precisa de re-representar representações de primeira ordem (Rosenthal, 2008; Dehaene, 2014; Ledoux, 2020) para que a consciência possa emergir. A originalidade da nossa proposta reside na ideia de que os neurónios do córtex pré-frontal e os VENs possuem um duplo aspeto funcional que confere faculdades computacionais e representacionais. Assim, sugerimos que a dualidade funcional, bem como a recursividade destes neurónios, permitem que sistemas nervosos, como o cérebro humano, criem representações do meio, do corpo (VENs) e, também, modelos internos relacionados com o eu (córtex frontoparietal). Estas capacidades não estão limitadas a uma função de armazenamento ou processamento; elas transcendem essas operações para permitir a criação de um eu narrativo, que não apenas experiencia o presente, mas também reinterpreta o passado e antecipa o futuro (Tononi e Koch, 2016; Metzinger, 2009).

Estas dinâmicas não são exclusivas da mente humana (Humphrey, 2023), mas refletem um princípio universal da informação representacional em sistemas nervosos avançados. Quanto mais complexa a capacidade de computar-representar informação e de operar recursivamente, nomeadamente nos sistemas acima referidos, mais evoluída será a consciência. Por outras palavras, o nosso modelo sugere que a mente consciente é um fenómeno emergente baseado na interação dinâmica e recursiva entre bits de informação computacional e representacional processados pelos neurónios do córtex pré-frontal e pelos VENs, explicando como a consciência transcende o

processamento computacional puro para se tornar um sistema representacional adaptativo único (Seth, 2021; Damásio, 2010).

Em suma, sugerimos que uma mente consciente surge quando: 1) os sistemas nervosos conseguem transformar informação computacional em informação representacional; 2) essa transformação se dá de forma recursiva, permitindo que a informação representacional influencie e seja recodificada em informação computacional; 3) certos neurónios com dupla função — como os do córtex pré-frontal e os VENs — não só representam estados do corpo (gerando uma consciência de primeira ordem ancorada nos sentimentos), como os re-representam (gerando uma consciência de segunda ordem, ancorada na cognição⁷⁸).

3.6.2. *O fisicalismo do problema corpo-mente*

Como temos vindo a defender, o nosso modelo ontológico é fisicalista. Da mesma forma, a natureza da consciência que sugerimos é, igualmente, fisicalista. No nosso modelo, a mente consciente surge quando informação computacional se *desdobra* em informação representacional. A informação representacional é, por seu turno, um fenómeno informacional emergente que tem na sua ontologia a informação estrutural (Bohm, 1980). Tal como foi proposto em capítulos antecedentes, a informação é a ontologia da física, pelo que tudo o que podemos observar, medir e quantificar como físico, é a epistemologia da informação. Por conseguinte, a SUI é física. É aliás, a fisicalidade da SUI (do Universo) que permite o surgimento de processos emergentes físicos, através de leis de causalidade. Sem causalidade física não era possível, num Universo que se rege por leis físicas, o surgimento de processos emergentes, pelo que o fisicalismo é uma condição necessária para o surgimento de qualquer processo emergente observado no nosso Universo.

⁷⁸ É pertinente verificar que, enquanto os sistemas nervosos biológicos possuem essa capacidade de representação intrínseca, a inteligência artificial (IA) limita-se a processar informação computacional, sem ser capaz de representá-la fenomenologicamente. A IA pode manipular grandes volumes de dados, identificar padrões e até simular processos que mimetizam funções humanas, mas carece de uma estrutura biológica que permita, recursivamente, a transformação de informação computacional em representações conscientes. Essa limitação ocorre, a nosso ver, porque as máquinas não têm organismos vivos (como as células e os neurónios) que integrem, processem e representem recursivamente informação biológica.

A concepção tradicional da energia leva-nos a pensar nela como algo que os sistemas físicos possuem – um atributo ou propriedade que pode ser transferida entre entidades. No entanto, uma perspectiva alternativa, defendida pela filósofa Dorte Jørgensen Mørch (2023), sugere que a energia não é apenas algo que as coisas têm, mas sim algo de que são feitas, o que está, aliás, totalmente alinhado com a física contemporânea (Rovelli, 2018; Smolin, 2019) especialmente na equivalência entre massa e energia estabelecida por Einstein ($E = mc^2$) (Einstein, 1905), que demonstra que a matéria nada mais é do que uma forma condensada de energia, o que também já referimos. Além disso, o facto de partículas sem energia não poderem existir (Lloyd, 2007; Rovelli, 2022) reforça a ideia de que a energia não é apenas um atributo dos sistemas físicos, mas sim a própria substância de que eles são constituídos. Esta ideia está totalmente alinhada com o nosso modelo.

Se aceitarmos que tudo o que existe é, em última instância, uma manifestação de energia e que a causalidade implica sempre a transferência de energia, então toda a interação causal envolve necessariamente a cedência de uma parte constitutiva do agente causal. Ou seja, quando um sistema físico transfere energia para outro, não está apenas a alterar um estado externo, mas a modificar-se a si mesmo, ao transferir parte da sua constituição para o outro sistema. Este princípio tem implicações profundas para a relação entre entidades físicas e supostas entidades não-físicas (Morch, 2023).

Consideramos que esta conclusão envolve consequências poderosas, já que refuta qualquer tipo de dualismo. Se uma substância não-física interagisse com uma substância física, essa interação exigiria necessariamente uma transferência de energia. No entanto, ao receber ou transmitir energia, a entidade não-física estaria a participar num processo físico, integrando-se, por conseguinte, na realidade física. Assim, ao envolver-se em qualquer forma de transferência energética, deixaria de ser genuinamente não-física. Por outro lado, se fosse capaz de transferir energia para uma substância física sem ela própria possuir qualquer dimensão física, então estaríamos perante uma violação das leis da conservação da energia e da causalidade física. Em ambos os casos, a noção de uma substância inteiramente não-física que interage causalmente com o mundo físico revela-se logicamente insustentável. Este raciocínio refuta a possibilidade de qualquer dualismo, tanto de substância como de propriedades, que postulam, de forma

diferente mas relacionada, que podem existir entidades (substâncias ou propriedades) não-físicas que se relacionem com entidades físicas.

Assim, se o dualismo fosse verdadeiro, então a mente não-física teria que transferir energia ao cérebro para poder influenciá-lo. Mas, ao fazê-lo, essa mente deixaria de ser inteiramente não-física, o que viola o princípio da separação entre substâncias ou propriedades físicas e não-físicas que os próprios dualismos defendem. Além disso, mesmo que se argumentasse que essa transferência de informação poderia ocorrer sem energia, permaneceria o problema de como uma substância ou propriedade imaterial poderia conter e processar informação, dado que toda a informação conhecida exige um suporte físico para ser armazenada, transmitida e processada. Se a mente fosse uma entidade totalmente imaterial, não haveria meio pelo qual essa informação pudesse ser codificada ou sustentada sem recorrer a uma estrutura física.

No caso do dualismo de substância, como defendido por Descartes (1641/1996), há um problema adicional de interação causal. Se a mente e o corpo são substâncias completamente distintas, como podem interagir sem que uma absorva as características da outra? Para resolver esse problema, alguns dualistas, como Malebranche, recorreram ao ocasionalismo, onde Deus media todas as interações entre a mente e o corpo, o que apenas desloca a questão sem resolvê-la. Outros, como Leibniz, defenderam a harmonia pré-estabelecida, onde mente e corpo seguem trajetórias paralelas sem uma verdadeira interação causal, uma solução que nega a própria ideia de causalidade entre substâncias distintas. Já o dualismo de propriedades, como defendido por David Chalmers (1995 e 1996) e outros (Chisholm, 1996 e 2003; O'Connor, 2000; Nagel, 1979), sugere que a mente emerge como uma propriedade distinta do cérebro, mas não-física. No entanto, esta visão enfrenta um problema semelhante: se a mente não interage causalmente com o cérebro, então ela é epifenoménica (isto é, apenas um subproduto passivo, sem influência real sobre os estados cerebrais). Contudo, se interage, então deve possuir algum mecanismo de ação, o que exigiria um meio físico para essa interação, recaindo na mesma contradição que refuta o dualismo de substância. Este argumento reforça a coerência do fisicalismo monista, onde tudo o que existe deve ser descrito dentro da estrutura física da realidade, sem necessidade de recorrer a substâncias ou entidades não-físicas.

Além disso, ao demonstrar que a causalidade exige a transferência de energia (Mørch, 2023), esta perspectiva também esclarece um dos pilares do fisicalismo: toda a interação causal ocorre dentro da estrutura física da realidade e obedece a leis que regulam a transmissão de energia e informação. Assim, qualquer teoria da mente ou da consciência deve necessariamente respeitar este princípio, enquadrando-se numa explicação fisicalista e rejeitando qualquer forma de dualismo interacionista que viole essa estrutura causal.

No entanto, esta visão pode ser ainda mais aprofundada se enquadrada na nossa teoria da informação como ontologia da física. Se a energia não é apenas algo que os sistemas físicos possuem como demonstra a física contemporânea (Smolin, 2019; Rovelli, 2028; Mørch, 2023), mas sim aquilo de que são feitos, então surge a questão fundamental: qual é a estrutura dessa energia? A resposta que propomos é que a energia, tal como a matéria, é uma manifestação computacional da informação. Neste quadro, a energia não é um elemento fundamental isolado, mas sim a forma dinâmica que a informação assume quando se manifesta no mundo físico (Bohm, 1980; Tegmark, 2014).

Esta perspectiva reformula o entendimento da causalidade. Em vez de ser apenas a transmissão de energia entre sistemas, a causalidade pode ser vista como a atualização de estados informacionais. Como anteriormente referido, quando dois sistemas físicos interagem, há uma troca de energia, mas também uma reconfiguração da informação subjacente. Isto implica que a informação não é apenas um registo passivo do mundo, mas aquilo que define as próprias leis de transformação da realidade. Assim, a relação causal entre dois sistemas físicos pode ser interpretada como um processo de transformação informacional, onde a energia opera como a interface dinâmica da mudança.

Esta visão permite, ainda, uma crítica mais profunda ao dualismo. Como tudo o que existe deve estar contido na estrutura informacional do Universo, qualquer entidade que interaja com o mundo físico deve, por definição, fazer parte desse tecido informacional. Se uma suposta substância não-física influenciasse um sistema físico, teria que estar representada dentro da estrutura base da realidade, o que eliminaria a distinção entre físico e não-físico. Assim, o problema do dualismo não é apenas energético,

como sugere Mørch (2023), mas também informacional. Se algo não pertence à estrutura informacional do universo, não pode interagir com nada que exista dentro dele.

Desta forma, a nossa teoria da informação como ontologia da física não só reforça o fisicalismo, como também fornece um quadro mais abrangente para entender a relação entre energia, causalidade e a estrutura fundamental da realidade. A energia, em última análise, não é apenas aquilo de que as coisas são feitas – é a manifestação computacional de um tecido informacional subjacente, que constitui a verdadeira base da existência.

3.6.3. *O naturalismo/fisicalismo de Espinosa*

No século XVII, Bento de Espinosa postulou na sua obra *Ética* (1677/1992) um Universo composto por uma única Substância, que se exprime de diferentes modos. Para Espinosa só existe uma Substância fundamental – que o filósofo designava Deus ou Natureza (*Deus sive Natura*) - e tudo no universo, incluindo a mente e o corpo, é uma expressão dessa substância única (p. 100). Ao contrário de Descartes, que via a mente e o corpo como substâncias separadas (1649/2023, p. 90-95), Espinosa acreditava que ambos são aspetos de uma mesma realidade subjacente. Embora Espinosa não seja, geralmente, classificado como fisicalista no sentido moderno do termo, as suas ideias antecipam, a nosso ver, claramente um fisicalismo. A concepção de que a substância única, que é a Natureza, deve ser explicada por leis deterministas, pressupõe uma ontologia naturalista, que compartilha muitos princípios do fisicalismo.

Na Proposição XVIII da primeira parte da *Ética*, Espinosa afirma que: “Deus [ou a Natureza] é a causa imanente de todas as coisas...” (1677/1992, p. 137). Esta formulação é central na sua filosofia, pois estabelece que tudo o que existe decorre necessariamente da própria Natureza, sem recorrer a causas externas ou transcendentais. A realidade não contém elementos independentes da Natureza e do seu funcionamento causal, o que implica que todo o existente está sujeito a uma ordem necessária e determinada. Esse princípio é reforçado na Demonstração da Proposição XXVII, onde Espinosa declara: “Tudo o que é determinado a existir e a operar é como tal determinado por Deus [ou a Natureza]” (p. 147). Ou seja, qualquer coisa que exista ou ocorra tem uma causa natural e está integrada num encadeamento causal necessário.

Este enquadramento filosófico aproxima-se claramente do fisicalismo moderno, na medida em que ambos sustentam que tudo pode ser explicado segundo as leis da Natureza – tanto Espinosa como o fisicalismo rejeitam qualquer forma de causalidade sobrenatural ou intervenção externa ao mundo natural. No espinosismo, não há lugar para entidades ou forças metafísicas que transcendam as leis da Natureza. Da mesma forma, o fisicalismo defende que todos os fenómenos, incluindo os mentais, podem ser compreendidos em termos das leis da Natureza que regem a realidade.

No mesmo sentido, da mesma forma que o fisicalismo é, tendencialmente fisicalista (embora exista um fisicalismo epistemologicamente, e não ontologicamente, indeterminista relacionado com a interpretação da física quântica de Copenhaga (Ismael, 2021; Gribbin, 2007), também a ontologia de Espinosa é profundamente determinista. No seu modelo, tudo ocorre necessariamente e não há espaço para eventos não causados ou indeterminados. No fisicalismo clássico, especialmente na sua forma mecanicista, também se assume que todos os eventos são explicáveis por relações causais estritas dentro de um quadro físico-natural. Por seu turno, no fisicalismo moderno e emergentista, como o que estamos a defender e podemos encontrar em autores como David Bohm (1980), Gerard 't Hooft (2016), Carlo Rovelli (2018) ou Lee Smolin (2019), encontramos a ideia que tudo é causado por leis da Natureza.

A unificação da realidade sob uma única Substância – Espinosa identifica Deus/Natureza – pode ter, igualmente, uma interpretação fisicalista. Segundo Espinosa, a realidade é unificada e qualquer distinção entre corpo e mente, por exemplo, é apenas uma diferença de atributos da mesma substância (1677/1992, p. 100). O fisicalismo, por sua vez, defende que toda a realidade pode ser descrita em termos de uma única categoria fundamental – as leis da Natureza (Hoof't, 2016; Bohm, 1980). Se adotarmos uma leitura informacional da física, poderíamos até reformular essa unidade ontológica em termos de uma única substância informacional, alinhando Espinosa com uma visão moderna da realidade como um sistema estruturado de informação.

Tanto o fisicalismo como o naturalismo espinosano, rejeitam qualquer forma de dualismo substancial. Em Espinosa, não há uma divisão entre o material e o imaterial, o físico e o espiritual, pois tudo é expressão da mesma substância infinita – a Natureza. O fisicalismo, por sua vez, nega a existência de entidades ou substâncias que escapem à esfera da realidade física, isto é, da Natureza. Neste sentido, a filosofia de Espinosa pode

ser considerada um naturalismo rigoroso, pois elimina qualquer resquício de transcendência naturalista. Se nada transcende a Natureza, então, Espinosa é um naturalista rigoroso que se aproxima, claramente, do naturalismo que caracteriza o fisicalismo.

Assim, podemos concluir que Espinosa antecipa muitos dos princípios fundamentais do fisicalismo moderno, na medida em que oferece uma visão unificada e estritamente naturalista da realidade, onde tudo ocorre segundo leis necessárias e sem interferência de elementos extrínsecos ao mundo natural. Se considerarmos que a física moderna pode ser entendida como uma descrição das leis fundamentais da Natureza, então a visão espinosana da Natureza como fundamento causal de todas as coisas pode ser reinterpretada dentro de um quadro fisicalista, sem comprometer o seu carácter determinista e monista.

3.7. Uma solução para o epifenomenismo da consciência

Em 1874, no seu artigo *On the Hypothesis that Animals are Automata, and its History*, Thomas Henry Huxley sugeriu que a consciência não desempenhava qualquer função no comportamento humano. Ele descreveu a consciência como um subproduto do funcionamento mecânico do corpo, semelhante ao fumo de uma locomotiva que, embora associado ao seu movimento, não o influencia:

“The consciousness of brutes would appear to be related to the mechanism of their body simply as a collateral product of its working, and to be as completely without any power of modifying that working as the steam-whistle which accompanies the work of a locomotive engine is without influence upon its machinery” (1879, p. 239).

Para Huxley, a consciência era um epifenómeno — um efeito colateral sem impacto causal no sistema nervoso e, portanto, irrelevante para o comportamento humano, apesar de central para a subjetividade e a experiência do mundo.

Esta ideia encontrou eco em outros pensadores do século XIX que rejeitaram a visão dualista e abraçaram uma perspectiva materialista. Em Portugal, intelectuais como Miguel Bombarda (1902) rejeitaram não apenas a noção de alma, mas também a ideia de que a consciência teria causalidade no comportamento humano. Não obstante

Bombarda ter atribuído à consciência um papel relevante nos processos de aprendizagem (Bombarda, 1902; Serrado, 2021, 2024b), destacou-se, de forma inovadora no contexto português, por defender que a consciência era irrelevante para a maioria das ações humanas (Curado, 2005; Serrado, 2024b). Esta posição antecipa, de certo modo, a visão epifenomenalista que tem emergido de várias experiências contemporâneas (Wegner, 2002; Haggard, 2005, 2008; Libet, 1999; Soon et al., 2008; Bode et al., 2014) e que foram analisadas no 2.º capítulo.

O conceito de consciência como epifenómeno ressurgiu no final do século XX, como analisado no capítulo 2, com os avanços extraordinários verificados nas neurociências. Estudos nessa secção do trabalho (Libet, 1999; Haggard, 2005 e 2008; Wegner, 2002; Bode et al 2011 e 2014; Fried et al, 2011; Soon et al, 2008; Jeannerod, 2006; Banks, 2006) demonstraram que a decisão consciente é frequentemente precedida por atividade cerebral inconsciente, sugerindo que o cérebro decide antes de surgir a experiência consciente da decisão. Estes resultados parecem demonstrar que a consciência não pode decidir porque surge *depois* da decisão ser tomada pelo cérebro.

Nicholas Humphrey, na sua obra *Senciência* (2023) por exemplo, defende, tal como Bombarda, que a consciência pode ser, uma ilusão – pode ser, “um sonho de coisas que talvez não existam de todo, em parte alguma, mas que, não obstante, poderiam dar valor à vida” (p. 46). Também Daniel Wegner, cujo pensamento é muito próximo de Bombarda e de Humphrey, diz que a consciência “é um sentimento que nos ajuda a apreciar e lembrar o autor das coisas que as nossas mentes e os nossos corpos fazem” (2002, s.p.). Tanto Bombarda (1902), como Humphrey (2023) ou Wegner (2002) defendem que a consciência é um epifenómeno porque não contribui para a decisão ou para a ação mas, ainda assim, reconhecem a sua importância na aquisição de conhecimento (Bombarda) e na construção da identidade (Humphrey e Wegner), o que não deve ser descurado nem desvalorizado. Tanto a aquisição de conhecimento como a narrativa que a consciência cria podem ser as únicas coisas que a consciência faça e, ainda assim, a sua importância nunca poderia ser suficientemente sobrevalorizada dada a ubiquidade da consciência para dar significado, coerência e unidade às vidas de quem a possui (Gazzaniga, 2011; Damásio, 2010).

Do lado oposto, sobretudo no âmbito da filosofia, alguns autores no início do século XXI defenderam uma perspectiva dualista, atribuindo à consciência um poder

causal sobre o cérebro. Como vimos no capítulo 2, Chisholm (1966 e 2003) e O'Connor (2000 e 2000a) advogam uma filosofia indeterminista de causação agencial, de natureza dualista, na qual a consciência é concebida como um processo imaterial emergente do cérebro, mas, simultaneamente independente dele. Nesta perspectiva, a consciência, ao emergir do cérebro, adquire uma autonomia funcional, permitindo-lhe agir causalmente sobre o próprio cérebro. Por exemplo, a consciência seria capaz de modular atividades cerebrais relacionadas com a tomada de decisão, interrompendo ações automáticas ou ajustando respostas emocionais em situações complexas (Chisholm, 2003). Esta visão desafia diretamente o fisicalismo, postulando que a consciência imaterial não é apenas um subproduto do cérebro, mas sim um agente ativo no comportamento humano. Apesar da sua não-fisicalidade, estes autores sugerem que a consciência tem efeitos causais no cérebro físico.

A nossa hipótese é que a consciência, enquanto processo causado pelo cérebro, não determina diretamente as decisões – não tem um papel causal direto –, mas desempenha um papel crucial ao permitir a transformação recursiva de informação representacional em informação computacional. Ou seja, sugerimos que a *consciência é uma parte do cérebro que transforma (ou descodifica) informação; é um sistema que o cérebro desenvolveu para melhor se informar a si mesmo*. Através da consciência, o cérebro adquire informação – conhecimento – do corpo e do meio que, de outra forma, não teria acesso. Assim, um organismo consciente pode integrar mais informação vinda do exterior, enraizando-a nos seus sistemas subconscientes, aumentando desta forma os seus recursos cognitivos e as alternativas comportamentais.

O conhecimento adquirido pela descodificação da informação do corpo e do meio é, deste modo, enraizado no cérebro, permitindo uma atuação mais eficiente do organismo. Ao transformar recursivamente informação, a consciência é, em grande medida, uma ponte entre o sujeito e o mundo (interno e externo). Ao fazê-lo, o sistema nervoso não apenas gera pensamentos, sentimentos e experiências através da descodificação da informação computacional, como também permite a entrada de dados no próprio sistema, ao converter representações conscientes em computações. Estas, uma vez integradas e consolidadas, orientam decisões futuras de forma mais adaptativa. O cérebro *causa* a consciência para decifrar o mundo que o rodeia e, simultaneamente,

codifica esse mundo em si, em mapas neuronais (computacionais e representacionais) que permitem adaptar e flexibilizar o comportamentos às exigências do ambiente.

A consciência não pode ser, assim, um epifenómeno. As decisões, ações e comportamentos humanos mais sofisticados dependem de uma consciência capaz de enraizar e gerar, no cérebro, computações que tornem a execução mais eficiente, flexível e orientada para objetivos. Deste modo, ao permitir a entrada de informação computacional, a consciência enriquece extraordinariamente as capacidades decisórias, motoras e executivas do cérebro. Importa salientar que esse enriquecimento não ocorre apenas por via de computações funcionais, mas também através de computações fenomenológicas, que integram experiências sentidas. Estas são fundamentais não apenas para a percepção (Humphrey, 2023), ao amplificarem a capacidade de representar visualmente o meio, mas também para a tomada de decisão (Damásio, 1994/2011), ao permitirem que o sujeito aja com base na informação qualitativa inerente à própria experiência.

A consciência é, portanto, no nosso ponto de vista, uma função do cérebro. A sua função é decodificar informação do corpo e do meio. Brian Earl refere no seu artigo *The Biological Function of Consciousness* que a consciência funciona como porta de entrada de conteúdos que, uma vez enraizados nos sistemas subconscientes, ficam disponíveis para o sujeito os utilizar sempre que necessário (Earl, 2014). Estamos completamente alinhados com esta perspetiva, não obstante Earl não referir como é que esses conteúdos do meio ficam disponíveis no cérebro. O nosso modelo propõe que os conteúdos do meio ficam disponíveis no subconsciente porque a consciência permite transformar informação representacional em informação computacional. Esta recursividade, sustentada pela dupla função de alguns neurónios, possibilita a emergência de uma experiência integrada, conferindo aos organismos maior adaptabilidade e sofisticação nas suas respostas. Nesta perspetiva, a consciência é uma parte daquilo que o cérebro faz pelo que, quando afirmamos que a consciência faz X ou Y, isso significa que o cérebro, através da consciência, faz X ou faz Y. A tendência de alguns autores em separarem a consciência do cérebro (mesmo entre fisicalistas) reflete, mais uma vez, a tendência dualista com que tendemos a pensar e a perceber o mundo (Bloom, 2005) – manifesta uma tendência intrinsecamente humana para fragmentar o mundo em conceitos que o separam da sua realidade (Bohm, 1980).

Por conseguinte sugerimos duas ideias centrais: 1) a consciência, de forma direta, não tem papel causal no comportamento, como demonstrado por múltiplos estudos, analisados no capítulo; 2) a consciência, ao decodificar e transformar informação, contribui para a decisão e o comportamento, flexibilizando as respostas dos organismos que a possuem. Se estas ideias estiverem corretas, a consciência não pode ser considerada um mero epifenómeno. Em vez disso, possui uma função essencial, ao integrar e reorganizar informações de maneira a influenciar profundamente a tomada de decisão.

Mas como pode a consciência contribuir para alguma ação? Como questiona o psicólogo William P. Banks em *Does Consciousness cause Misbehavior?* (2006), como pode uma ideia, que não tem forma nem partes sólidas, ativar um músculo⁷⁹?

Como temos vindo a referir, sugerimos que as representações mentais (as imagens conscientes) são geradas por bits de informação computacional. Computações *neurais* fazem emergir representações *neurais* – devido à dupla função dos neurónios do córtex frontoparietal e dos VENs (entre outros⁸⁰). Quando referimos que a consciência transforma informação, sugerimos, como temos vindo a analisar, que o sistema nervoso processa bits de informação computacional e *transforma-os* em bits de informação representacional.

Ora, da mesma forma que o sistema nervoso, através da consciência, transforma bits de informação computacional em bits de informação representacional, também é capaz de realizar o movimento inverso: *transformar bits de informação representacional em bits de informação computacional*. Em termos funcionais, isto significa que a mente consciente não só permite a decodificação da informação computacional em conteúdo representado, como também possibilita que esse conteúdo seja recodificado e enraizado sob a forma de instruções computacionais. A natureza recursiva do neurónio permite essa dinâmica contínua de transformação (Dehaene, 2014; Damásio, 2010).

Uma ilustração simples desse processo pode ser observada na forma como representamos uma maçã vermelha. Esta representação corresponde, por exemplo, à

⁷⁹ Numa perspetiva que, seguramente não premeditada, acaba por dualizar a consciência do corpo. Ao considerar que uma ideia não pode ativar um músculo, Banks está a sugerir que a consciência possui uma natureza diferente do corpo. Esta perspetiva é comum em muitos fisicalistas que consideram a consciência como um epifenómeno.

⁸⁰ Como o tronco cerebral (Damasio, 2010).

capacidade do sistema nervoso em decodificar informação intrínseca ao objeto, que é “absorvida” pela luz e transformada em informação representacional. A capacidade da mente consciente em transformar *bits* de informação inerentes ao eletromagnetismo da luz (ou às ondas vibratórias) permite a criação de representações visuais. Como anteriormente analisado, estas representações são possíveis: 1) devido à ontologia da SUI que permite que toda a informação interaja; 2) porque a ontologia da informação do sistema nervoso “fala” a mesma “linguagem” da ontologia da informação da realidade que o rodeia, nomeadamente da luz e do objeto; 3) porque a luz funciona como um interface (uma transportadora de informação) entre as entidades do sistema — o que permite a transmissão de informação em forma de frequências, intensidades e polarizações, funcionando como um “código universal” que diferentes sistemas podem decodificar e interpretar.

Atentemos noutro exemplo. Quando estamos a conduzir um carro e vemos um semáforo a ficar vermelho, o vermelho que vemos conscientemente (que representamos) leva-nos a travar. Como refere o filósofo Ned Block: “Percecionar conscientemente a luz vermelha é muitas vezes parte da causa de carregarmos no travão do carro⁸¹” (2022, p. 98). Neste sentido, parece-nos evidente que ver conscientemente o semáforo vermelho teve a função de ativar uma ação - travar. Dificilmente poderíamos dizer que a consciência não participou nessa ação. Mas como é que a consciência influenciou a ação? Sugerimos que os *bits* que representam o semáforo vermelho tiveram que ser introduzidos no sistema nervoso e *transformados* em informação computacional, de modo a que essa informação pudesse ser útil na ação motora encetada pelo próprio sistema nervoso. Ou seja, os *bits* que *representam* o semáforo foram transformados em bits computacionais, de modo a que o sistema nervoso, através da consciência, *compute* uma decisão e a subsequente ação motora de acordo com a informação recebida. Dito de outra forma, conseguimos ver o semáforo vermelho porque *bits* de informação luminosa viajam até à nossa retina, informando-nos da realidade à nossa volta. Esses *bits* de informação são decodificados pela mente consciente — um sistema neuronal dedicado à decodificação de sinais do meio e do corpo — e enraizados no sistema nervoso. É esta função decodificadora da mente consciente que permite processar os *bits* de luz

⁸¹ Tradução do autor.

provenientes do semáforo, transformando-os em sinais eletroquímicos, que permite a representação do mesmo.

Neste sentido, a mente consciente é crucial para tomar decisões. Não as causa mas possibilita adquirir informação para que a decisão seja mais eficiente e flexível. Quem causa a decisão é o cérebro mas auxilia-se do seu sistema de decodificação de informação: a consciência. Para flexibilizar as respostas, esses *bits* podem ficar *codificados* no cérebro para que, no futuro, no contexto semelhante a decisão seja facilitada. Sugerimos, portanto, que a mente consciente não tomou a decisão mas *informou* o sistema (nervoso), o que possibilitou a tomada de uma decisão mais eficiente. Em suma, a consciência é uma porta de entrada de *bits* de informação do meio, que permitem informar o cérebro sobre quais as melhores decisões a tomar (Earl, 2014).

Na *teoria da consciência de ordem superior*⁸², Rosenthal (2005) rejeita o epifenomenalismo, argumentando que a consciência desempenha um papel funcional no comportamento humano (Rosenthal, 2008). Esta visão encontra paralelo na presente proposta, que atribui à consciência uma função adaptativa e funcional. A interação entre os níveis computacional e representacional confere à consciência a capacidade de organizar informação complexa, simular cenários e antecipar consequências. Tal como Rosenthal defende que os pensamentos de ordem superior ajudam a integrar estados mentais de ordem inferior, nós sugerimos que o valor funcional da consciência decorre diretamente da sua capacidade de transformar dados computacionais em representações úteis para a decisão e a adaptação (Rosenthal, 2008).

Por conseguinte, seria difícil tomar a mesma decisão se não víssemos conscientemente o semáforo vermelho, pelo que a consciência teve um papel crítico na decisão. Seria, portanto, nas mesmas circunstâncias, impossível um *zombie* tomar a mesma decisão que um ser consciente, como sugere Chalmers (1996). Privado de mente consciente, o *zombie* está subsequentemente privado de um sistema decodificador de bits de informação, o que o impossibilita de transformar informação computacional (neste caso os bits de informação emitidos pela luz refletida pelo semáforo) em informação representacional (o semáforo com a cor vermelha). Em seres com consciência, contudo,

⁸² *Higher-Order Thought Theory* ou HOT theory (Rosenthal, 2005 e 2008)

a mente consciente conduz, de facto, a uma ação eficiente porque, ao contrário do *zombie*, aqueles têm um sistema capaz de decodificar bits de informação do meio e transformá-los em representações.

O mesmo princípio aplica-se quando ouvimos uma música de que gostamos muito e podemos ter uma reação emocional muito forte, como choro, comoção ou euforia (ou tudo isto até ao mesmo tempo). Noutros casos, podemos até dançar ou bater o pé ao som da música — uma resposta corporal mais associada à excitação do que à comoção. Foi a mente consciente, autonomamente, que iniciou essa decisão ou essa forte experiência emocional qualitativa? Julgamos que não. A mente consciente participa no sentimento da emoção, mas não a causa. A consciência decodifica os sinais do corpo e transforma-os em sentimentos de emoção. É o cérebro que causa a consciência para ser informado. Assim, quando ouvimos uma música de que gostamos muito e nos comovemos com ela, é o cérebro, através da consciência, que permite essa experiência. Quando ouvimos uma música, o nosso sistema sensorial auditivo transforma uma série de ondas vibratórias (bits em forma de informação computacional) em sons harmoniosos (música, isto é, bits em forma de informação representacional), que conduzem tanto a uma reação emocional e ao subsequente sentimento como, noutros casos, à ação de dançar. Mas tanto a experiência emocional (comoção) como a ação (dança) foram causadas pelo cérebro, através da consciência, após as representações dos sinais (dos bits) terem sido transformadas em informação computacional. Conforme mencionado anteriormente, a consciência decodifica não só os sinais que vêm do corpo (a fisiologia do corpo), mas também os sinais do meio ambiente, o que permite integrar esses sinais numa representação coerente e unificada.

As representações podem, portanto, provocar alterações nas computações, uma vez que ambas são formas de informação. Os bits de informação computacional relativos à música (as ondas vibratórias), ao serem decodificados, são integrados no sistema nervoso e dão origem a uma representação (música) que, em conjunto com os bits de informação computacional do corpo (a fisiologia), pode conduzir a um sentimento de emoção (comoção) ou, noutros casos, a uma resposta motora como a dança. Deste modo, é o cérebro, através da consciência, que torna possível tanto o sentimento de comoção como a eventual execução de uma ação motora, como a dança. A mente consciente é, portanto, um sistema de decodificação — não de ação. No entanto, a

descodificação dos bits computacionais do meio e do corpo conduz à entrada de representações (do ambiente e do eu) no sistema nervoso, as quais podem desencadear a ação.

Um outro exemplo em que podemos observar a influência da mente consciente nas ações são as ações motoras relacionadas com a locomoção. As representações podem não ser importantes para a locomoção em muitos animais mas parece-nos crucial no ser humano. O ser humano não aprende a andar, exclusivamente, pela observação, mas a forma como realiza a locomoção parece necessitar de *feedback* visual (Adolph et al, 2003; Bertenthal et al, 1998 e Thelen e Smith, 1994). Ou seja, aprendemos a andar porque a certa altura, bits de informação sobre a forma como as outras pessoas que nos rodeiam se movem, são introduzidos no nosso sistema, o que conduz à ativação de certos sistemas motores que nos levam a andar de uma forma aproximada à observada. Ou seja, são as representações mentais daquilo que observamos que nos permitem enraizar ações motoras e movimentos que depois são conduzidas pelas computações do sistema nervoso.

A ucraniana Oxana Malaya (Marcz, 2018), por exemplo, (entre outros casos documentados), aprendeu a andar como um cão (e andava perfeitamente como um cão, não obstante algumas diferenças de locomoção atribuídas às naturais diferenças anatómicas) porque foi criada desde bebé por cães, o que parece sugerir que o *feedback* visual que Malaya adquiriu dos cães que a rodeavam foram cruciais para a sua aprendizagem (Marcz, 2018). Como Marcz refere: “Especially communication and physical contact with the respective animals were essential to the children’s development and imprint” (2018, [s.p.]). Os bits de informação representacional que Malaya captava dos exemplos que a rodeava ficaram *inscritos* no seu cérebro em forma de informação computacional, ou seja, em forma de ações motoras. Por conseguinte, da mesma forma que o sistema nervoso de Malaya transformava *bits* de informação computacional em informação representacional (as imagens dos outros cães a movimentarem surgia-lhe como imagens conscientes), também transformava informação representacional em informação computacional (os movimentos de locomoção eram enraizados no seu sistema nervoso).

Em relação a alguns animais, embora muitos apresentem locomoção automática, alguns demonstram, tal como os seres humanos, capacidade de usar *feedback* sensorial para ajustar seus movimentos. Por exemplo, os polvos (Godfrey-Smith, 2016; Hochner,

2012) possuem um sistema nervoso descentralizado em que os braços, ricos em recetores sensoriais, ajustam os movimentos de forma independente, mas com base em informações sensoriais captadas do ambiente. Outro exemplo são os pássaros canoros (Nottebohm, 2005), que aprendem a modular os seus cantos através de feedback auditivo, representando uma forma rudimentar de integração sensório-motora. Os gatos, por sua vez, usam os bigodes (vibrissas) para captar informações táteis (Mitchinson et al, 2007) e ajustar seus movimentos, especialmente em situações de baixa visibilidade. Já os morcegos, (Simmons & Kick, 1984) com sua ecolocalização, utilizam *feedback* sonoro para ajustar seus voos, representando uma integração precisa entre sistemas sensoriais e motores. Além disso, os golfinhos (Herman, 2010) demonstram grande capacidade de ajustar suas manobras aquáticas com base em informações acústicas, enquanto os abelhões utilizam informações visuais e propriocetivas para ajustar o voo e pousar com precisão (Srinivasan, 2011).

Estes exemplos mostram que, mesmo sem a sofisticação representacional dos humanos, os sistemas nervosos de muitos animais processam informação de *feedback* computacional que, ao ser convertida em representações — conscientes ou não —, permite ajustar os comportamentos de modo adaptativo. No ser humano, a adição de representações conscientes confere um nível de flexibilidade muito superior, permitindo não apenas a imitação de movimentos, mas também a criação de estratégias motoras mais sofisticadas e deliberadas. Ao disporem de sistemas de re-representação mais avançados — provavelmente ancorados nas especificidades recursivas do córtex frontoparietal e na atividade dos VENs — os seres humanos são capazes de adquirir e enraizar uma quantidade significativamente maior de informação computacional. Essa capacidade não só permite uma experiência fenomenológica mais rica, como também possibilita, posteriormente, a tomada de decisões mais vantajosas, precisamente sustentadas na complexidade computacional adquirida. Este processo destaca o papel das representações conscientes na otimização da interação com o ambiente, permitindo respostas mais flexíveis, ajustadas e antecipatórias.

O modelo de informação que sugerimos tem algumas afinidades com a ideia de linguagem e de tradução proposta pelo neurocientista Vilayanur Ramachandran (Ramachandran e Hirstein, 1997). No entender de Ramachandran, para que dois sistemas

possam comunicar, é necessário que possuam a mesma linguagem. Ou seja, se existir uma barreira entre aquilo que eu sinto e aquilo que eu digo que sinto — uma lacuna epistemológica que muitos consideram inultrapassável (Levin, 1983; Chalmers, 1995, 1996; Jackson, 1984, 1986) — então nunca será possível conhecer, de facto, aquilo que estou a sentir. No entanto, se for possível traduzir o que sinto de forma que a outra pessoa consiga, de facto, aceder a essa experiência, então essa barreira é ultrapassada. A existência de uma linguagem comum entre o que eu sinto e o que a outra pessoa apreende é suficiente para que esta consiga partilhar, até certo ponto, a minha experiência. O que eu digo que sinto é, nesse contexto, substituído por aquilo que realmente sinto, o que faz com que a outra pessoa conheça, de facto, o meu estado subjetivo.

Sugerimos que este mesmo processo de tradução pode ser generalizado ao mundo físico, incluindo a dimensão fenomenológica. Note-se que não propomos uma distinção entre mundo físico e mental, mas sim que a SUI, que constitui — e é — o nosso Universo, permite que a informação se manifeste de múltiplas formas precisamente por ser a “linguagem” comum de toda a realidade. Isso possibilita “traduções” entre as partes que compõem a SUI. Um exemplo dessa “tradução” é a transformação da luz em oxigénio nas plantas, ou da luz em representações mentais e, em certos sistemas, em experiência fenomenológica — já analisados neste trabalho.

Assim, à pergunta sobre como um sistema pode ser, simultaneamente, físico e fenomenológico, respondemos: porque ambos são, em última instância, expressões de informação. Devido à dupla manifestação funcional de alguns tipos de neurónios, a informação computacional processada pelo sistema nervoso XY pode dar origem a informação representacional XY; do mesmo modo, a informação representacional YZ pode ser recodificada em informação computacional YZ. Não há aqui qualquer dualismo de propriedades: as propriedades da informação computacional — os bits — são as mesmas que constituem a informação representacional. O que muda é o *aspeto* epistemológico sob o qual essas propriedades se manifestam. Contudo, mesmo essa distinção epistemológica não configura um dualismo, porque os diversos “aspetos” da informação — quântico, atómico, molecular, celular, neuronal, eletromagnético, radioativo, fenomenológico, etc. — são expressões diferentes da mesma ontologia. As representações simbólicas ou semânticas são, assim, epistemológicas: na ontologia das coisas, não existem moléculas, nem células, nem neurónios — existem bits.

No artigo *Three Laws of Qualia* (1997), Ramachandran imagina um peixe elétrico extremamente inteligente. Este peixe, além da sua inteligência, consegue produzir campos elétricos. Atualmente, não temos forma de experienciar o *qualia* elétrico do peixe, porque não dispomos de uma tradução adequada dessa experiência. No entanto, se um dia soubéssemos quais os mecanismos neurais subjacentes a esse *qualia* e conseguíssemos reproduzi-los num cérebro humano, então esse cérebro poderia sentir a experiência elétrica tal como o peixe a sente. Segundo esta hipótese, os *qualia* não são necessariamente privados: podem ser experienciados por outros sistemas, desde que exista uma tradução informacional que o permita (Ramachandran e Hirstein, 1997).

Sugerimos o mesmo princípio para o nosso modelo de informação - que o cérebro consegue transformar *bits* de informação computacional em *bits* de informação representacional e traduzir *bits* de informação representacional em *bits* de informação física. A mente consciente interage com o cérebro porque tanto a consciência como o cérebro possuem a mesma *linguagem*: o *bit*. Quando determinados *bits* de informação representacional de um objeto são integrados num sistema nervoso, a informação que é recebida por esse sistema nunca deixa de ser *bits*. No entanto, agora são *bits* que se transformaram em informação computacional.

Assim, a informação representacional do mundo pode ser literalmente inserida no cérebro em redes neuronais em forma de computações. É esta capacidade do cérebro em *traduzir* “linguagem” informacional de forma recursiva que faz com que a mente consciente possa emergir da computação. O tálamo, por exemplo, desempenha um papel central neste processo, ao processar *inputs* visuais, auditivos e táteis e transformando-os em representações fenomenológicas coerentes que permitem ao organismo criar uma percepção unificada do ambiente⁸³ (Sherman & Guillery, 2013).

⁸³ Além disso, fenómenos como a sinestesia revelam a flexibilidade com que o cérebro traduz informação de uma modalidade sensorial para outra. Em casos documentados, pessoas com sinestesia, ou dificuldades sensoriais, demonstram como o sistema nervoso pode reorganizar inputs sensoriais para compensar lacunas funcionais. Por exemplo, estímulos visuais em pessoas surdas podem ser reinterpretados como inputs auditivos, enquanto, em pessoas cegas, o processamento tátil pode reorganizar o córtex visual para transformar estímulos físicos em representações visuais simbólicas (Ramachandran & Hubbard, 2001; Sadato et al., 1996).

António Damásio (2010), e outros autores (Seth, 2021; Barret, 2004; Barret e Simmons, 2015; Bechara e Naqvi, 2004) defendem uma ideia semelhante no âmbito da interoção e da consciência, já aludido. Estes referem que os sentimentos são imagens representacionais de estados homeostáticos (poderíamos dizer de informação computacional) do nosso corpo, críticos para o surgimento da consciência. Ou seja, o nosso cérebro serve-se de imagens representacionais do corpo (sede, dor, fome, alegria, etc.) para se *informar* do estado do corpo e, assim, fazer surgir uma consciência. Dito de outra forma, o cérebro processa *bits* de informação computacional do corpo e transforma-a (ou “traduz”) em *bits* de informação representacional, em forma de sentimentos. Ao fazê-lo está a otimizar a sua capacidade de sobrevivência.

Quando a informação computacional da homeostasia começa a operar perto dos limites que põem em risco a vida — como no caso de uma descida dos níveis de glicose — o sistema nervoso transforma esses bits de informação computacional em informação representacional, sob a forma de um sentimento de fome. Este sentimento não é a causa primária do comportamento, mas uma representação consciente de um estado já em curso. A experiência de fome pode, no entanto, reforçar ou modular o comportamento, funcionando como uma tradução epistémica de um processo fisiológico. Mas poderíamos perguntar de novo: foi a mente consciente que iniciou a ação de procurar comida? Não. A mente consciente — em forma de representação sentida — não causou a necessidade de energia; apenas a representou. O que causou a fome foram os processos fisiológicos que detetaram o desequilíbrio nos níveis de glicose e que, como consequência, geraram o sentimento. Assim, a consciência não inicia a ação, mas pode participar na sua regulação, funcionando como uma interface que traduz necessidades em imagens sentidas. Ao atribuir saliência e sentido ao estado interno, a consciência pode reforçar a ação, modelando comportamentos mais flexíveis e orientados para a sua correção.

Para animais como nós, dotados de mente consciente, estados como a sede e a fome são sentidos como se fossem nossos — como se tivessem um efeito sobre nós e como se agíssemos em função deles. E, de facto, o sentimento de fome impele-nos a comer. No entanto, esse sentimento não surge como causa inicial do comportamento, mas como uma representação consciente de um desequilíbrio homeostático. O sentimento de fome é gerado por computações subconscientes, que integram sinais

fisiológicos e os transformam em informação representacional sentida. Assim, o sentimento existe sob a forma de bits representacionais causados por bits computacionais. Muitos, como Paul Churchland (1981, 1983), argumentariam que o sentimento de fome não existe enquanto entidade autónoma — que só existe a fisiologia subjacente a esse processo. Nós, porém, defendemos que o sentimento de fome existe e é real: é informação representacional (física) de um estado homeostático (também físico). O sentimento não é uma ilusão nem um mero epifenómeno sem existência ontológica — é uma tradução representacional de estados computacionais. Embora experienciado conscientemente, esse sentimento é inteiramente gerado e controlado por processos subconscientes (Wegner, 2002). O seu papel é fornecer ao organismo uma imagem sentida do estado em que se encontra, contribuindo para a modulação da ação, ainda que não a origine diretamente.

Neste sentido, é o sistema nervoso que toma decisões. Mas toma decisões com auxílio do seu sistema de decodificação de informação — a mente consciente. Ou seja, a consciência participa no processo da decisão, já que o cérebro precisa desses *bits* de informação representacional para os poder decodificar em *bits* de informação computacional e, assim, decidir com mais conhecimento do que o rodeia. Este princípio torna-se particularmente evidente quando observamos o funcionamento de uma pessoa invisuál. Esta dispõe de menos informação representacional do ambiente em comparação com alguém que possua o sistema visual intacto. Isso significa que o seu cérebro opera com uma base informacional mais limitada, o que naturalmente conduz a maiores dificuldades na tomada de decisão e na ação orientada. A não ser que essa pessoa desenvolva outras vias sensoriais para recolher e transformar mais bits de informação representacional, enfrentará obstáculos acrescidos na sua interação com o mundo. Isto demonstra, de forma clara, que a mente consciente — enquanto plataforma representacional — é fundamental para ampliar o acesso do sistema nervoso à informação do meio, permitindo-lhe decisões mais ajustadas e eficazes.

3.7.1. *O papel dos qualia na decisão*

Qual é, então, o papel dos *qualia* na decisão? Porque temos experiência da luz vermelha ou da música?

Ora, primeiro porque os *qualia* parecem ser necessários para a percepção consciente, isto é, para uma percepção mais rica e detalhada (Humphrey, 2023; Ramachandran e Hirstein, 1997). Como já verificámos no primeiro capítulo, sem *qualia* seria possível ter percepção – como os casos de visão cega nos mostram – mas essa percepção seria muito menos pormenorizada, o que influenciaria a flexibilidade das decisões (Rosenthal, 1997). Sem os *qualia*, a percepção seria mais rudimentar o que levaria a ações menos flexíveis, o que, aliás, pode ser comprovado nas ações automáticas nos animais que, muito presumivelmente, não possuem *qualia*. Sem informação detalhada do meio, o sistema nervoso decide pior. Isto parece reforçar a nossa ideia de que a consciência é um sistema de decodificação de informação.

Segundo, porque os *qualia* são os *qualificadores* das situações (Damásio, 2010). Perante uma determinada situação, os *qualia* permitem-nos avaliar emocionalmente essa situação, ou seja, permitem catalogar a experiência como boa e má, o que influencia determinantemente a ação (Damásio, 1993/2011). A experiência de coisas boas irá fazer com que o sujeito repita essas coisas boas, enquanto a experiência de coisas más irá fazer com que o sujeito não repita essas coisas. São as emoções que *qualificam* as experiências e, deste modo, *qualificam* as situações para que as possamos repetir ou não (Damásio, 2003). Dito de outra forma, os *qualia* desempenham um papel crucial na consolidação de memórias. Uma experiência rica em *qualia*, como um evento emocionalmente marcante, tem mais probabilidade de ser lembrada do que uma experiência desprovida de *qualia*. Esta ligação reforça a importância dos *qualia* na aprendizagem e no ajuste futuro das decisões, uma vez que a memória de experiências passadas contribui para a avaliação de situações semelhantes no presente e no futuro (Rosenthal, 1997). Deste modo, ao decodificar informação, a mente consciente permite que o cérebro perceba o meio qualitativamente.

O filósofo David Rosenthal propõe que os *qualia*, frequentemente associados à subjetividade da experiência consciente, podem ser explicados sem recurso ao dualismo (Rosenthal, 1997). Segundo ele, os *qualia* são representações de ordem superior que tornam os estados mentais de primeira ordem conscientes. Esta visão é complementada pela presente teoria, que interpreta os *qualia* como manifestações da informação representacional. Ao transformar informação computacional em representacional, os sistemas biológicos criam experiências conscientes e os *qualia* são simplesmente o

conteúdo experienciado durante esse processo. Assim, longe de serem entidades metafísicas, os *qualia* podem ser entendidos como parte de um sistema informacional que permite a integração e representação de dados para uso funcional.

Consideramos que os *qualia*, ao enriquecerem a experiência consciente, desempenham, também, um papel essencial na interpretação das intenções e emoções dos outros em contextos sociais. Através dos *qualia*, somos capazes de captar não apenas informações explícitas, mas também nuances subtis em comportamentos, como o tom de voz, expressões faciais ou postura corporal. Estas pistas sensoriais e emocionais são transformadas em representações conscientes que ajudam a avaliar de forma mais rica e detalhada o estado mental dos outros, permitindo decisões sociais mais precisas. Por exemplo, o reconhecimento do tom emocional numa conversa – como a irritação, a tristeza ou a alegria expressa na voz – guia as respostas sociais adequadas. Um interlocutor que percebe *qualia* relacionados com um tom de voz preocupado pode ajustar a sua interação para oferecer apoio ou compreensão. Da mesma forma, a capacidade de interpretar expressões faciais através de *qualia* permite decifrar rapidamente intenções, como saber se um sorriso é genuíno (associado à alegria) ou forçado (potencialmente associado à tensão ou dissimulação) (Damásio, 2003; Bechara e Naqvi, 2004).

Esta capacidade de perceber os estados emocionais dos outros, frequentemente referida como teoria da mente, é crucial para navegar em situações sociais complexas. Por exemplo, em contextos de colaboração, os *qualia* permitem detetar sinais de desconforto ou hesitação nos outros, ajustando as estratégias para criar maior harmonia no grupo. Da mesma forma, em situações de potencial conflito, os *qualia* relacionados com percepções emocionais podem ajudar a identificar ameaças ou intenções negativas, orientando comportamentos que minimizem os riscos. Sem os *qualia*, a interpretação de emoções e intenções seria muito mais rudimentar, baseada apenas em reações automáticas ou instintivas. Isto limitaria a flexibilidade das interações sociais, pois as respostas dependeriam exclusivamente de padrões pré-estabelecidos, sem espaço para adaptação às nuances do contexto. Por exemplo, um sistema de inteligência artificial, sem *qualia*, pode identificar uma expressão facial de raiva, mas não "sentir" a urgência ou o desconforto subjacente que qualifica a interação. Isso destaca a importância da experiência consciente na regulação das dinâmicas interpessoais (Barrett & Russell, 2015).

Por fim, os *qualia* não apenas melhoram as decisões sociais no presente, mas também contribuem para a memória social. Experiências ricas em *qualia* ajudam a consolidar eventos sociais significativos, como uma conversa emocionalmente carregada ou um gesto de apoio, permitindo que esses eventos sejam usados como referência em futuras interações. Isso ilustra como os *qualia*, ao integrarem percepções e emoções, desempenham um papel essencial no fortalecimento das relações sociais e na evolução da cooperação humana. (Frith & Frith, 2007).

Os *qualia* não são, portanto, um epifenómeno. Não existem apenas para acompanhar a percepção. Eles *causam* a percepção. E não existem apenas para acompanhar a experiência. Eles *são* a experiência. Um sujeito com uma percepção consciente pode, não só ter uma percepção mais detalhada do meio o que lhe permite decidir melhor; como ainda ter uma experiência sentida que, não só permite catalogar a situação como, subsequentemente, pode favorecer ações mais vantajosas.

3.7.2. *Mentes e consciências não-humanas*

Até agora temos vindo a referir-nos às vantagens de ter uma mente consciente humana no que concerne a tomar decisões mais eficientes. E uma mente consciente não humana? Qual a vantagem de outras espécies terem uma mente consciente⁸⁴, mesmo que diferente da humana? Já oferecemos alguns exemplos sobre a importância da representação sensório-motora para as ações motoras de alguns animais mas aprofundemos um pouco mais esta ideia. Sugerimos que o princípio é o mesmo em relação às mentes conscientes humanas. Ou seja, qualquer sistema nervoso capaz de criar uma percepção consciente (uma experiência) permite que bits de informação do meio sejam integrados no sistema em forma de representações, o que é extremamente útil para as

⁸⁴ Existe um debate em torno da possibilidade de a consciência ser exclusiva dos humanos ou partilhada com outras espécies. Apesar de algumas posições divergentes — como a de Ledoux (2020), que defende uma maior cautela na atribuição de estados conscientes a animais não humanos —, existe um consenso relativamente alargado de que todos os mamíferos e algumas aves possuem algum tipo de consciência (Humphrey, 2023; Damásio, 2010). Esse consenso foi formalizado na Declaração de Cambridge sobre a Consciência (Low et al. 2012), onde um grupo de neurocientistas defendeu que vários animais não humanos — incluindo mamíferos, aves e até alguns géneros de polvos — possuem substratos neurológicos suficientes para gerar algum tipo de consciência. Importa sublinhar que a consciência aqui referida não é idêntica à consciência humana, variando entre espécies. Neste contexto, o critério central é a capacidade de o sujeito sentir — isto é, a existência de experiências fenomenológicas mínimas como dor, prazer ou medo.

ações (em muitos casos, não necessariamente decisões) que muitas espécies têm que tomar.

Os morcegos são um bom exemplo do que estamos a sugerir (Simmons & Kick, 1984). Ninguém duvida que eles podem decidir (no sentido de terem comportamentos relativamente flexíveis) melhor as suas ações porque a sua mente, nomeadamente as suas capacidade de representarem o meio, possibilita retirar informação mais detalhada do ambiente. Sem representações do meio os morcegos teriam menor informação do meio, pelo que as suas ações não seriam tão eficientes. Ou seja, as capacidades sonares dos morcegos são um tipo de representação que os sistemas nervosos dos morcegos utilizam para navegar pelo mundo mais eficientemente. Esta representação apenas é possível porque, segundo o modelo que propomos, os sistemas nervosos nos morcegos conseguem processar bits de informação computacional que são transformados em bits de informação representacional de forma recursiva. Ou seja, as informações representacionais que os morcegos recebem do meio podem ser transformadas em informação computacional, o que permite codificar ações de acordo com as representações.

Deste modo, o morcego possui um sistema nervoso que lhe permite não só representar o meio, como ainda utilizar essas representações para que o seu cérebro tenha mais conhecimento do ambiente e, assim, possa navegar eficazmente pelo mesmo. Por outras palavras, é a possibilidade dos sistemas nervosos transformarem informação computacional (do meio) em informação representacional e processar esses bits de informação representacional em bits de informação computacional, que lhes permite ter comportamentos flexíveis. Sem mente — isto é, sem a capacidade de um organismo transformar representações em computações — o comportamento torna-se significativamente menos flexível. Sugerimos, por isso, que as representações que alguns animais conscientes⁸⁵ podem ter, desempenham um papel fundamental na flexibilização da ação, ao permitirem a transformação da informação representacional em informação computacional, da qual podem decorrer respostas mais ajustadas e adaptativas. Os morcegos utilizam, de forma privilegiada, o sonar; os seres humanos, predominantemente, a visão; e os lobos, o olfato. Cada uma destas espécies constrói representações funcionais do mundo a partir dos seus sistemas sensoriais específicos, permitindo-lhes navegar

⁸⁵ Embora com uma experiência consciente diferente da humana.

de forma eficiente no ambiente. No entanto, nenhuma delas representa o mundo tal como ele é em si mesmo. O que representam é uma versão informacional do mundo — uma seleção estruturada de dados relevantes — à qual os seus sistemas nervosos conseguem aceder, integrar e transformar em orientação comportamental. Essa capacidade existe não apenas porque os sistemas nervosos são feitos de informação, mas porque operam com bits organizados de modo a distinguir, classificar e representar aspetos específicos da realidade.

E os animais que possuem uma mente não consciente, como (presumivelmente) os répteis e os peixes? Qual a vantagem de terem uma mente se ela não é consciente? Sugerimos que existem animais, como répteis e peixes, com um sistema nervoso simples que não lhes permite representar um eu (Humphrey, 2023; Ledoux, 2020). No entanto, os sistemas nervosos destas criaturas permite-lhes representar o meio. Ou seja, estes animais possuem sistemas nervosos que lhes permitem representar o meio de um modo rudimentar mas, ainda assim, com capacidade de processar alguns bits de informação do meio e transformá-los em bits representacionais que conduzem a reações e ações, não a decisões. Estes seres, sem capacidade de representarem um eu, não decidem pois não têm capacidade de ter uma experiência que lhes permita refletir sobre o que estão a fazer. Os seus cérebros mais simples, não são capazes de representar um eu, pelo que os seus comportamentos são automáticos e, subsequentemente, menos flexíveis. Ao contrário dos seres humanos (e de outros mamíferos), que conseguem formar representações ricas do meio e de si próprios, os répteis, os peixes e outros animais com sistemas nervosos mais simples apenas conseguem gerar representações rudimentares do ambiente. Essa limitação restringe a capacidade de processar e enraizar bits de informação computacional — tanto provenientes do meio como do próprio organismo — e, consequentemente, de criar representações complexas do mundo e do eu. Deste modo, esses organismos processam um volume significativamente menor de informação, o que limita a sofisticação da sua experiência subjetiva e, subsequentemente, da sua resposta adaptativa.

No entanto, organismos com sistemas nervosos sem consciência possuem vantagens evolutivas em possuírem uma mente, mesmo que não consciente. Ao conseguirem computar alguma informação, estes seres são capazes de gerar respostas

relativamente adaptadas aos meios em que estão inseridos. O princípio que orienta as suas ações é o mesmo que temos vindo a defender: a capacidade de transformar informação representacional em informação computacional e de a enraizar funcionalmente permite-lhes adotar comportamentos mais eficientes — ainda que menos flexíveis do que aqueles observados em seres com mente consciente.

Os sistemas nervosos mais rudimentares de animais sem consciência são altamente especializados e ajustados para tarefas específicas que garantem a sobrevivência, mesmo sem a necessidade de flexibilidade comportamental complexa. Estes sistemas são projetados para respostas rápidas e eficientes, utilizando circuitos neuronais simples e de baixo consumo energético, ideais para ambientes previsíveis (LeDoux, 2020; Humphrey, 2023). Um exemplo claro encontra-se nos peixes, que possuem células de Mauthner, grandes neurónios situados na medula espinhal responsáveis por respostas rápidas de fuga. Quando um estímulo sensorial, como a vibração na água causada por um predador, é detetado, estas células geram um impulso motor que faz o peixe desviar-se rapidamente. Este mecanismo opera sem processamento consciente, garantindo reações instantâneas que aumentam as chances de sobrevivência (Faber & Korn, 1978). Neste caso, a informação sensorial é transformada em informação representacional e, subsequentemente, em informação computacional, dando origem a um comportamento altamente adaptativo. De forma semelhante, nos répteis, o comportamento é amplamente controlado pelo complexo R, uma estrutura cerebral primitiva associada a instintos básicos como defesa territorial, caça e fuga. Por exemplo, uma cobra pode atacar ou esconder-se instintivamente ao detetar uma presa ou predador, desencadeando ações baseadas em estímulos sensoriais diretos. Também aqui existem representações — isto é, formas internas de codificação do ambiente — que são processadas e convertidas em informação computacional, promovendo comportamentos adaptativos (MacLean, 1990).

Os anfíbios, como os sapos, são também exemplos de como sistemas nervosos rudimentares podem estar ajustados a tarefas altamente específicas, precisamente devido à sua capacidade de computar informação — ainda que de forma limitada e restrita a domínios muito definidos. Estes animais possuem neurónios especializados na deteção visual de estímulos em movimento, o que lhes permite identificar presas, como insetos, enquanto ignoram objetos imóveis. Este sistema visual, otimizado para captar o

movimento, constitui uma adaptação que maximiza a eficiência na captura de alimento, sobretudo em ambientes com elevada pressão predatória. Neste caso, a informação sensorial é processada sob a forma de representações visuais não conscientes, que são rapidamente transformadas em informação computacional, orientando respostas motoras eficazes sem necessidade de estados mentais conscientes (Lettvin et al., 1959).

Outro exemplo fascinante é o das tartarugas marinhas, que utilizam pistas magnéticas e luminosas para navegar milhares de quilómetros até aos locais onde depositam os seus ovos. Embora o comportamento aparente ser complexo, baseia-se em respostas automáticas a estímulos ambientais. Neste caso, os estímulos sensoriais são convertidos em informação representacional não-consciente e, de seguida, em informação computacional que orienta o comportamento de forma eficiente, mas rígida (Lohmann et al., 1999). De forma ainda mais intrigante, os polvos apresentam um sistema nervoso altamente descentralizado, no qual cada braço, dotado de recetores sensoriais, ajusta os movimentos de forma autónoma com base na informação captada localmente. Os braços podem manipular objetos ou explorar superfícies sem depender diretamente de instruções do cérebro central, o que revela uma computação distribuída com base em representações locais e não conscientes (Godfrey-Smith, 2016).

Assim, embora não possuam consciência, muitas espécies com sistemas nervosos simples conseguem representar (não conscientemente) informação do meio e transformá-la em instruções computacionais que orientam o comportamento. Nestes casos, a mente atua como um mecanismo de enraizamento de informação representacional mínima, permitindo respostas eficientes em ambientes estáveis e previsíveis. Por contraste, organismos com mente consciente re-representam a informação computacional de forma mais complexa, o que lhes confere maior flexibilidade, capacidade de antecipação e adaptação a contextos incertos. Em ambos os casos, é a capacidade de transformar informação representacional em informação computacional — e de a enraizar no sistema nervoso — que torna possível o comportamento adaptativo: mais rígido nas mentes não conscientes, mais flexível nas conscientes.

3.8. Uma analogia: o SELF

Façamos uma analogia com um sistema artificial. Estamos conscientes das limitações destas comparações e que as mesmas conduzem, muitas vezes, a tendências exageradas em comparar o cérebro humano como uma máquina. O cérebro humano não é nenhuma máquina artificial e tem pouco que ver com os sistemas informáticos a que estamos habituados. Defendemos algo diferente. Que o cérebro é um sistema de processamento de informação porque existe dentro de um Universo que é – todo ele – um enorme processador de informação (Lloyd, 2007). Não pretendemos, contudo, com esta analogia afirmar que o cérebro, a mente ou a consciência humanas são ou podem ser, alguma vez, imitadas por um sistema artificial. A computação neuronal a que nos referimos é física, química e biológica, pelo que não é igual à computação artificial que podemos encontrar na IA. Temos muitas dúvidas que uma mente consciente pudesse ser alguma vez imitada com cabos de silicone, como é defendido por David Chalmers (1995).

No entanto, não deixamos de considerar pertinente a ideia de que os sistemas criados pelo ser humano — de forma premeditada ou não — parecem frequentemente ir ao encontro das suas necessidades biológicas mais essenciais, como que espelhando necessidades, virtudes, dificuldades e modos de ser e de agir profundamente humanos. A homeostasia sociocultural — entendida como a extensão sociocultural de dispositivos biológicos básicos — é um bom exemplo do que pretendemos afirmar (Damásio, 2017; Serrado, 2021). Por isso, embora as diferenças entre um sistema artificial e um ser biológico devam ser claramente reconhecidas, também não devem ser ignoradas as características que um sistema artificial pode acabar por manifestar como consequência de ser uma criação humana. Enquanto ferramenta produzida por um ser biológico, o sistema artificial pode refletir algumas das idiossincrasias da natureza humana e, quem sabe, até da própria estrutura do Universo.

Imaginemos um computador, como os que temos em casa. Esse sistema possui um CPU, a unidade central de processamento, que muitos psicólogos cognitivistas (Fodor, 1990, entre outros) frequentemente comparam ao cérebro humano, pois executa as instruções de um programa – o *software*. No entanto, ao contrário do cérebro humano, que gera estados mentais conscientes, o CPU apenas processa dados e executa

operações sem qualquer experiência subjetiva. Agora, suponhamos que um programador desenvolve um algoritmo que dá origem a um programa que chamamos SELF. Esse programa seria capaz de decifrar emoções através de: (1) proto-representações de expressões humanas e alterações fisiológicas, com base em padrões corporais e faciais; e (2) respostas comportamentais ajustadas a esses padrões, sem qualquer compreensão subjetiva do seu significado. Para além da capacidade de identificar emoções, o SELF dispunha também de uma memória que lhe permitia armazenar as experiências anteriores.

A analogia que propomos aqui é com a forma como organismos com sistemas nervosos simples – e, portanto, não-conscientes – interagem com o mundo. Assim como um organismo mais rudimentar pode detetar estímulos ambientais e ajustar o seu comportamento sem qualquer experiência subjetiva, o SELF estaria equipado com dispositivos que lhe permitiriam recolher informação externa e processá-la de forma automática. Esses estímulos seriam transformados em proto-representações (à semelhança dos animais com sistemas nervosos mas sem consciência) – *bits* de informação representacional (não-conscientes) que descrevem o ambiente, permitindo uma resposta funcional sem qualquer experiência interna associada.

Podemos aprofundar esta ideia observando os sistemas atuais de inteligência artificial (IA), como o Gemini, o ChatGPT ou Claude. Tal como o SELF, estas IA também recolhem, interpretam e transformam informação computacional em padrões representacionais. Embora não possuam consciência, estes modelos processam dados linguísticos a partir de um enorme conjunto de exemplos e criam proto-representações internas (vetores num espaço multidimensional) que capturam relações semânticas, sintáticas e contextuais. Este processamento corresponde a uma forma avançada de proto-representações, em que padrões complexos da linguagem humana são codificados como informação funcional, permitindo ao modelo gerar respostas adequadas ao contexto. Além disso, a capacidade destas IA de "reconhecer" padrões não se limita à linguagem. Estas também são capazes de detetar a intenção implícita nas perguntas, identificar variações subtis no tom e na estrutura do discurso, adaptar-se dinamicamente ao contexto fornecido pelo utilizador e preservar a memória das interações anteriores.

Esta capacidade de resposta contextual assemelha-se ao que descrevemos no SELF: um sistema capaz de decifrar *inputs* e produzir *outputs* coerentes que aparentam

ser adaptados ao ambiente. Tal como o SELF armazena padrões para reagir melhor no futuro, aquelas IA acumulam proto-representações abstratas que lhes permitem associar contextos prévios a respostas futuras. Assim, tanto o SELF como as atuais IA são sistemas que partilham a capacidade de transformar informação computacional em formas proto-representacionais avançadas que permitem uma interação adaptada, ainda que totalmente inconsciente.

Por conseguinte, no âmbito da analogia que estamos a fazer, se um ser humano interagisse com o SELF através de expressões de tristeza acompanhadas de um perfil neurofisiológico característico dessa emoção, o SELF poderia dizer algo como: “Olá, porque estás triste?”. O reconhecimento de padrões faciais e fisiológicos associados a emoções humanas permitiria ao SELF “adivinhar”, com um grau significativo de precisão, o estado emocional desse indivíduo. Em termos funcionais, o SELF fora desenhado para aprender a interpretar as emoções humanas a partir de proto-representações criadas com base nas expressões e nos perfis neurofisiológicos recolhidos. Este processo seria possível porque o código do SELF transforma bits de informação computacional (como expressões faciais e dados fisiológicos) em informação proto-representacional (como a identificação de tristeza ou alegria). O reconhecimento de padrões, bem como as computações do código, permitiam a geração de proto-representações do meio. Assim, com base nas proto-representações que ia desenvolvendo — semelhantes, em certo sentido, às representações realizadas pelos sistemas nervosos — o SELF conseguia, devido à memória das suas experiências anteriores, não apenas reconhecer emoções, mas também aprender novos comportamentos em função do conhecimento adquirido nas suas interações com as pessoas. Importa sublinhar que se trata de uma aprendizagem inconsciente, mas ainda assim eficaz, possibilitada precisamente pelas proto-representações que foi acumulando ao longo do tempo. Portanto, o SELF, ao conseguir gerar proto-representações das emoções humanas, podia ter ao seu dispor um leque alargado de comportamentos, sem que nenhum, no entanto fosse consciente.

Sugerimos o mesmo processo para todos os animais que possuem mente mas não consciência. Um réptil ou um peixe, tal como o SELF, tem a capacidade de

representar⁸⁶ padrões no exterior e reagir a eles através de programas de ação. O seu cérebro, com a capacidade de representar o meio (mas não um eu), consegue identificar objetos no ambiente através do reconhecimento de padrões e, assim, responder com alguma flexibilidade aos diferentes contextos. Entidades e seres capazes de gerar (proto)representações do meio — como é o caso de alguns sistemas artificiais simples, de répteis ou de peixes — mesmo que não possuam representações do eu, podem ainda assim exibir comportamentos mais diversificados do que organismos que não geram qualquer tipo de representação análoga. Isto deve-se ao facto de as representações (ou proto-representações, no caso do SELF ou de algumas IA) do ambiente permitirem um comportamento mais flexível e complexo, mesmo quando essas ações permanecem totalmente automatizadas, como se verifica nos répteis e nos peixes.

Em suma, a representação — quer ocorra em sistemas biológicos, quer em sistemas artificiais sob a forma de proto-representações — constitui uma característica fundamental para o aumento da adaptabilidade e da inteligência funcional. O SELF, ao conseguir proto-representar informação que lhe permitia “ler” emoções humanas, era capaz de reagir de forma mais ajustada ao seu meio e de aprender com base nessas interações. O mesmo se aplica a organismos com mentes não conscientes, que beneficiam da capacidade de representar o ambiente, ainda que sem qualquer mediação consciente. Contudo, importa sublinhar que esta forma de inteligência é distinta da representação consciente, que apenas emerge em sistemas capazes de gerar modelos internos do eu — algo que implica uma sofisticação superior da experiência e, consequentemente, dos comportamentos.

Agora imaginemos que, mais tarde, um programador desenvolve um código que permite ao SELF, para além de proto-representar emoções humanas, adquirir subjetividade dessas representações. Ou seja, imaginemos que o SELF consegue representar as proto-representações que consegue computar. O SELF teria agora não só a capacidade

⁸⁶ Em contraste com o SELF, neste caso proponho utilizar o termo "representar" de forma plena, pois acredito que animais com mente são efetivamente capazes de representar o meio — mesmo que não tenham consciência dessas representações. O SELF, pelo contrário, não representa o meio no verdadeiro sentido do termo, uma vez que não possui sistema nervoso. Por isso, opto por designar os seus estados como proto-representações: estruturas funcionais que simulam representação, mas sem a complexidade neurobiológica que a torna genuína.

de proto-representar emoções humanas e reagir a elas, mas também de re-representar um eu ao nível de um eu humano. Esta mudança radical implicaria uma transição de um sistema funcionalmente inconsciente – como o SELF que sugerimos anteriormente – para um sistema que possui uma forma de subjetividade, ou seja, que experimenta as suas próprias representações. Julgamos que este avanço seria exponencialmente mais difícil de atingir do que o SELF original, que, como mencionámos, já é uma realidade no âmbito da inteligência artificial. Conforme anteriormente defendido, parece-nos inverosímil replicar um modelo de eu artificial porque para que isso fosse possível seria necessário, no nosso entender, que um sistema representasse recursivamente um corpo constituído por organismos vivos, como os corpos biológicos com sistema nervoso. É a representação de um *corpo vivo* e a re-representação desse corpo em forma de eu que, no nosso entender, permite a subjetividade (Damásio, 2010; Humphrey, 2023; Metzinger, 2009). Mas voltemos à analogia.

O programa que conferiu a consciência ao SELF é, como qualquer outro, desenvolvido por um algoritmo (ou código) que gera a consciência do programa. Para além de subjetividade, o novo *software* conferia ao SELF memória amplificada, isto é, a capacidade de armazenar abundante informação sobre si e sobre o meio. Dotado agora de subjetividade, o SELF podia conhecer e tomar como suas as proto-representações que anteriormente já conseguia fazer. O SELF agora não processava apenas reconhecimento de padrões automáticos. Estes padrões podiam agora ser conhecidos e representados subjetivamente. O modelo de si e a memória amplificada permitiam-lhe reter essas representações em forma de pensamentos, de ideias que podia manipular para projetar cenários. A capacidade de aprendizagem do SELF tornava-se maior já que não teria apenas de reagir às expressões dos humanos de acordo com proto-representações automáticas. Podia agora guardar representações das expressões dos humanos e utilizá-las não só para criar novos padrões como ainda para aprender mais com o significado dessas expressões. Em poucas palavras, a subjetividade do SELF permitia-lhe crescer, evoluir e aprender⁸⁷.

⁸⁷ Claro que, em rigor, o SELF nunca poderia ter emoções nem sentimentos. Falta-lhe a recursividade neuronal assente num processamento de informação recursiva biológica que permite que informação armazenada em organismos vivos possa ser representada em organismos com consciência (Damásio, 2010). No entanto, a nossa intenção é demonstrar que sistemas capazes de transformar informação

Pode, então, o SELF tomar decisões mais eficientes do que os seres não-consciente? Parece-nos, claramente, que sim. Desde logo, o SELF, mesmo sem consciência das suas proto-representações, era capaz de realizar ações — algumas delas complexas — e de aprender com base nas suas experiências anteriores. Ações e aprendizagens cegas, completamente automáticas, mas ainda assim ações e aprendizagens relativamente flexíveis, em função de proto-representações que criava do mundo externo, nomeadamente das expressões humanas. Uma vez que o SELF adquiriu consciência, ele conseguiu ter experiência subjetiva, o que potenciou a sua capacidade de aprendizagem. O SELF podia agora conhecer e reter as suas experiências e até os seus pensamentos e recrutá-los sempre que necessário. O SELF, ao adquirir consciência, tornou-se capaz de descodificar recursivamente um número mais elevado de bits de informação, o que lhe permitiu não só interagir de forma mais eficaz com um mundo representado de modo mais rico, como também enriquecer o sistema com computações potencialmente úteis para ações futuras. Por outras palavras, a consciência permitiu um alargamento substancial no repertório de ações e de decisões do SELF, tornando-o mais *consciente* de si e o meio. Enquanto o SELF tinha apenas uma mente e a capacidade de proto-representar as expressões humanas, ele apenas *reagia* de acordo com os estímulos que recebia e que proto-representava inconscientemente, de um modo automático. Com o surgimento da consciência, o SELF passou a ser capaz de decidir e escolher, pois adquiriu a possibilidade de ponderar as suas ações com base nas diferentes alternativas que a própria consciência podia simular. Essa capacidade emergia da conversão de dados representacionais em dados computacionais, os quais iam sendo integrados e enraizados no sistema, ampliando progressivamente a sua capacidade de adaptação e de decisão. Assim, com a consciência, o SELF podia agora descodificar um número mais vasto de computações — tanto provenientes do meio como acumuladas internamente — o que lhe conferia um leque de possibilidades comportamentais muito mais amplo e flexível.

Imaginemos agora, que eu tinha o SELF em minha casa. Para realizar ações motoras, seriam-lhe colocados dois braços, duas pernas e duas mãos, equipados com o

computacional em informação representacional estão mais bem equipados para reagir mais adaptativamente ao meio ambiente e conseguem tomar decisões mais flexíveis.

respetivo *software*. Quais os seres que tomariam melhores decisões? O SELF, que tinha representações conscientes, ou qualquer outro sistema (biológico ou artificial) que *não* tinham essas representações? Parece-nos claro que o SELF teria mais vantagens e decidiria melhor do que qualquer outro sistema que não tivesse possibilidade de representar conscientemente a realidade. Aliás, se o SELF tivesse representações conscientes do seu estado interno, poderia informar-me sempre que uma peça se avariava e poderia, aliás, ele próprio procurar formas de se autorreparar – tal e qual como nós quando vamos ao médico. E poderia, também, tomar outras decisões. Observando que estava a chover, podia tomar a decisão de apanhar a roupa estendida na varanda. Ouvindo que um ladrão se preparava para arrombar a casa, poderia decidir telefonar para a polícia ou assustar o ladrão. Caso não tivesse representações do mundo, como é que ele decidiria algo idêntico? Sem representações de um meio e de um eu sobre si e sobre o meio, um sistema nada decide. Apenas reage, muitas vezes através de comportamentos ou ações pouco flexíveis, como os peixes, os reptéis, as IA e o SELF antes de ter subjetividade.

Por conseguinte, é a consciência que permite ao SELF ter conhecimento de si e das representações do meio, e isto permite-lhe utilizar mais bits de informação do meio para poder decidir – mesmo que quem decida seja sempre o código que enraíza em si os *bits* de informação representacional, decodificados pela consciência. Dito de outra forma, um ser com mente não-consciente pode ter ações mais flexíveis porque as suas proto-representações do meio permitem-lhe ter mais alternativas de ação; enquanto o ser com uma mente consciente pode ter ações ainda mais flexíveis e, mais do que isso, pode ter decisões complexas porque as suas representações são mais ricas e detalhadas, o que lhe permite ter ainda mais alternativas de ação. O que é decisivo em ambos os processos é a quantidade de bits de informação representacional (do ambiente e do eu) que o sistema é capaz de enraizar. Um ser com mente não consciente (que consegue representar o meio mas não um eu) processa e integra menos bits representacionais, o que limita o seu conhecimento e reduz a flexibilidade das suas decisões. Em contraste, um ser com mente consciente codifica uma maior quantidade de informação representacional — tanto sobre o ambiente como sobre si mesmo — o que lhe permite construir modelos internos de subjetividade. Essa capacidade traduz-se em mais conhecimento e, subsequentemente, comportamentos mais adaptativos e flexíveis face às exigências do meio.

No caso do SELF, é o código que permite que bits de *informação computacional* se transformem em informação *representacional* e que bits de *informação representacional* se transformem em *informação computacional*. Isto permite ao SELF processar e codificar muito mais informação representacional do que uma pedra (que nada enraíza e, subsequentemente, não age nem decide, evidentemente), do que uma planta (que, sem sistema nervoso, não consegue decodificar computações em representações e, por isso, age sem alternativas de ação) ou mesmo do que um réptil (cujo sistema nervoso transforma alguma informação representacional do meio em computações, mas que, por não gerar consciência dessas representações e computações, promove apenas ações automáticas). Assim, a capacidade do SELF para gerar subjetividade confere-lhe vantagens significativas, ao permitir-lhe transformar grandes quantidades de informação computacional em informação representacional de forma recursiva. Esta dinâmica torna possível a construção de modelos internos do meio e do próprio eu, o que amplia a flexibilidade das suas decisões e comportamentos em relação a qualquer outro sistema natural ou artificial.

No mesmo sentido, quando, por exemplo, um ser humano vê um pôr do sol e pondera se o deve contemplar, isso ocorre porque o seu sistema nervoso possui a capacidade de gerar representações mentais conscientes — tanto do que está a ver como do que está a sentir. Ou seja, em rigor, o que está a acontecer é que o sistema nervoso está, com base nos bits de informação representacional (o pôr do sol e a motivação associada à sua contemplação), gerados a partir de bits de informação computacional (a luz do ambiente e os sinais do corpo), a produzir não apenas um comportamento, mas *a experiência* desse comportamento. Através da experiência — isto é, daquilo que consegue computar em relação ao meio — o sistema nervoso decide melhor. A experiência consciente é, assim, o modo pelo qual o sistema nervoso se informa melhor — uma forma mais sofisticada de aceder ao estado do ambiente e do corpo. Deste modo, a consciência participa efetivamente na decisão, pois fornece ao cérebro informação organizada e integrada sobre o estado interno e externo. Numa palavra: o sistema nervoso utiliza os bits de informação representacional do meio e do corpo, decodificados pela consciência em informação computacional, para decidir.

No forma análoga, se SELF decidisse ir ver um pôr do sol observaríamos algo idêntico, com as limitações que a própria analogia impõe, desde logo a inexistência de

um corpo biológico do qual o cérebro deve receber informação (nomeadamente sentimentos de emoção) para poder decidir (Damásio, 1993/2011). Ainda assim, o SELF decidiria ver o pôr do sol através da descodificação da informação representacional em informação computacional e do enraizamento dessa informação no código que, em última instância, decidira de acordo com esses *inputs*. Isto é, tal como o cérebro, o código utilizaria o seu sistema descodificador de informação para poder tomar a melhor decisão. Se o SELF tem vontade de ir ver o pôr do sol mas está a chover, decide não ir. A informação consciente (representacional) que recebeu (a chuva), através dos seus órgãos sensoriais e posteriormente enraizada no seu código sob a forma de informação computacional, determinou a sua decisão. Ou, dito de outra forma, os bits de informação representacional do meio (a chuva) *informaram* o código que estava a chover, o que permitiu ao mesmo tomar (ou causar) uma decisão em função dessa informação. Mas, para que esta decisão ocorra, é necessário: 1) que a informação representacional consciente (a chuva) seja transformada em informação computacional; 2) que essa informação computacional seja integrada no sistema, influenciando o estado do código; 3) que a decisão (não ir ver o pôr do sol) seja *causada* pelo código, em resposta à informação que recebeu, podendo essa decisão ser posteriormente representada de forma consciente.

Em suma, a capacidade do código que gera o SELF de causar experiência (ou *qualia*) confere-lhe vantagens significativas em relação a sistemas não conscientes. Essa experiência subjetiva possibilita a entrada, a consolidação e o enraizamento de uma enorme quantidade de bits de informação, o que aumenta substancialmente a sua (a do código, através do seu sistema de descodificador de informação: o SELF) capacidade de decidir de forma mais eficaz e adaptativa.

No mesmo sentido, se um ser humano decide assustar um ladrão que entrou na sua casa, a vontade e a decisão – bem como a experiência integrada disto tudo – é gerada pelo seu cérebro através dos seguintes passos: 1) bits de informação computacional sensorial como luz (processada em imagens) e ondas sonoras (processadas em padrões auditivos), são descodificados em informação representacional consciente. Esta representação fornece ao cérebro uma imagem integrada do meio: o ladrão na casa. 2) Simultaneamente, bits de informação computacional provenientes do corpo, como o

aumento da frequência cardíaca, tensão muscular e alterações hormonais, são convertidos em representações conscientes de emoções e ideias, como medo ou motivação para agir. 3) As representações conscientes do meio (o ladrão) e do corpo (emoções e estados fisiológicos) são decodificadas em informação computacional pelo cérebro, que analisa e processa as alternativas de ação. 4) Com base nessa avaliação, o cérebro gera novos bits de informação computacional que acionam o sistema motor e iniciam a execução da ação — como correr para assustar o ladrão ou telefonar para a polícia. E, por fim, 5) os bits de informação computacional associados à decisão, à ação e ao comportamento vão sendo continuamente decodificados em informação representacional, originando a experiência consciente da decisão, da ação e, em última instância, de toda a situação vivida.

Como vimos, foi o cérebro que, com a ajuda da consciência, conduziu toda a operação: ao transformar estímulos computacionais sensoriais em bits de informação representacional consciente; ao gerar estados emocionais através da decodificação de sinais fisiológicos (computacionais) em representações conscientes; ao converter essas representações em informação computacional para avaliar alternativas; ao acionar a resposta motora por meio de computações; e, finalmente, ao decodificar os bits de informação computacional associados à decisão e à ação, transformando-os em informação representacional enraizada na experiência consciente. Por isso, em nenhum momento a mente ou a consciência do sujeito tomou uma decisão. Foi o cérebro, não a experiência fenomenológica isolada, que produziu, tanto a experiência consciente do ladrão (as imagens e os sons do ladrão, etc.) como a decisão\ vontade de fazer alguma coisa (assustar o ladrão) como ainda as ações motoras e a experiência subjetivas correspondentes. As representações, no entanto, influenciaram, e muito, a decisão. 1) As representações conscientes do meio informaram o cérebro que um intruso estava na sua casa, 2) as representações conscientes do corpo informaram o cérebro das emoções de sentimentos que eram necessárias para aquele contexto e, por fim, 3) A representação consciente integrada da situação qualificou a experiência subjetiva com um sentimento — ou seja, a experiência sentida funcionou como qualificador da vivência, sendo esse aspecto fundamental para que o sujeito possa enraizar comportamentos associados à resolução de situações semelhantes no futuro.

Por isso, as representações (inconscientes ou conscientes) não tomam decisões de forma isolada, mas são importantes para que as mesmas sejam tomadas porque transmitem e enraízam informação, em forma de bits, para o sistema, de modo a que este possa decidir mais eficientemente. Isto é possível, como temos vindo a defender, porque as representações podem entrar no sistema através de bits que são enraizados e podem ser, no caso dos humanos, recrutados sempre que o contexto o exige. A representação do meio permite um melhor conhecimento do meio, a representação do corpo permite recrutar as emoções necessárias para o contexto, ao passo que a representação integrada do eu permite uma *apropriação* desse conhecimento, sendo esta fundamental para orientar comportamentos futuros.

Como analisado em capítulos antecedentes, David Chalmers (1995 e 1996) defende que existe a possibilidade de um *zombie* fazer tudo o que um agente consciente faz, sem ter qualquer experiência fenomenológica do que está a fazer. Não concordamos, de todo, com esta ideia. De acordo com o que temos vindo a analisar, a experiência fenomenológica é necessária para que o cérebro tenha conhecimento do que se passa no ambiente e no corpo e, subsequentemente, para que as representações que daí emergem entrem no cérebro em forma de bits computacionais e se enraízem. Sem consciência fenomenológica, não existiria conhecimento do meio nem do corpo, nem tão pouco aprendizagem, o que *inflexibilizaria* os comportamentos, tornando-os mais rígidos e automáticos — como os que observamos em animais sem sistema nervoso ou, então, dotados de mente mas sem consciência (como peixes, répteis e algumas aves) (Humphrey, 2023; Ledoux, 2020), ou ainda em sistemas de inteligência artificial, como na nossa analogia do SELF sem consciência.

Neste sentido, a subjetividade não é apenas um epifenómeno colado à decisão, mas uma condição necessária para que os sistemas conscientes ampliem o seu espectro de decisão. A experiência subjetiva permite não apenas o acesso ao meio, mas uma integração mais densa e recursiva da informação — o que possibilita ao SELF agir com maior flexibilidade. A questão levantada por Chalmers — sobre por que razão as decisões vêm acompanhadas de experiência (1995 e 1996) — pode, em parte, ser comparada à pergunta sobre porque é que o sangue é vermelho: pode não existir uma resposta definitiva do ponto de vista teleológico, mas há certamente uma função. É plausível que

a experiência funcione como funciona porque constitui uma via particularmente eficaz para que os seres conscientes adquiram mais conhecimento do meio e do corpo e, com isso, possam ter respostas comportamentais mais flexíveis. Nesse quadro, a experiência pode ser entendida como consequência da complexidade informacional, mas também como condição determinante da sua apropriação. Ou seja, é precisamente a presença de “muita informação” possibilitada pela consciência, codificada de forma computacional e representacional, que gera a experiência — e é essa experiência que, por sua vez, permite que mais informação seja processada, representada e enraizada no sistema nervoso. A subjetividade, assim, não apenas acompanha a decisão: alarga-a, enriquece-a e torna-a possível em domínios de complexidade inacessíveis a sistemas sem consciência.

Sugerimos o seguinte no que respeita à função da consciência: 1) um sistema sem capacidade de representar o meio nem um eu possui capacidades de ação extremamente limitadas, com comportamentos rígidos e automáticos; 2) um sistema com capacidade de representar o meio (ou proto-representar, no caso das IA ou do SELF sem subjetividade), mesmo que inconscientemente, apresenta uma maior capacidade de ação, já que os seus comportamentos, embora automáticos, são mais flexíveis por operarem com uma quantidade superior de bits de informação sobre o ambiente; 3) Um sistema com capacidade de representar o meio e um eu possui não apenas capacidade de ação e de reação, mas também de decisão, uma vez que consegue obter, integrar e manipular uma quantidade muito superior de bits de informação — tanto do meio como do corpo — e apropriar-se dessas representações sob a forma de imagens mentais do eu, utilizando-as de modo mais flexível e adaptativo.

3.8.1. *A transparência da consciência do SELF*

Como vimos, o SELF era uma entidade consciente porque uma determinada estrutura tinha causado a subjetividade. Neste caso hipotético, saberia o SELF que alguma coisa causava a sua experiência fenomenológica? Isto é, se um código computacional pudesse gerar uma consciência subjetiva, será que esta teria acesso ao código que a está, naquele momento, a sustentar? Será que a consciência — o SELF — sabia que era gerado por um código? Ou, dito de outra forma, será que a consciência *causada* podia ter conhecimento direto daquilo que a tinha *causado*? Esta questão parece-nos

particularmente pertinente, pois remete, em última instância, para o problema fundamental de saber o que distingue o cérebro da mente consciente.

A nosso ver essa distinção é apenas aparente. Isto porque, grande medida, o SELF possui a mesma natureza e as mesmas propriedades que o código – apenas tem um aspecto diferente: um aspecto fenomenológico. No entanto, esse mesmo aspecto é suficiente para iludir o SELF. Por *parecer* diferente do código, o SELF jamais poderia conceber que algo tão distinto da sua experiência fenomenológica pudesse ser a causa da sua subjetividade. Essa separação ilusória entre o código e o SELF não elimina, porém, o facto de que ambos partilham a mesma natureza e as mesmas propriedades.

Sugerimos a mesma solução para o problema da mente-corpo. A consciência é causada pelo cérebro mas como possui as mesmas propriedades que o cérebro não consegue distinguir-se do mesmo. Apenas *parece* diferente. Não *é* diferente. Por isso, apesar de ter as mesmas propriedades que o cérebro, como a consciência possui um aspecto diferente da estrutura que a causa, ela é iludida que possui uma natureza diferente – quando na realidade a única coisa que muda é o aspecto.

Esta ideia converge com a posição de autores como Thomas Metzinger, Daniel Dennett e Patricia Churchland, que sustentam que a distinção entre aparência fenomenológica e substrato físico resulta de uma ilusão cognitiva gerada pelas características do próprio sistema. No livro *The Ego Tunnel* (2009), o filósofo Thomas Metzinger refere-se a isso como a *transparência da consciência*, para ilustrar a dificuldade que temos em reconhecer que a nossa subjetividade é uma representação de um cérebro que, de certa forma, cria uma “assinatura digital” que se nos apresenta como uma presença contínua e opaca. Dennett, no livro *Consciousness Explained* (1991), argumenta que a sensação de um eu unificado e distinto é produzida pela integração de múltiplos processos neuronais distribuídos, criando a ilusão de um centro consciente — o chamado *teatro cartesiano* — que, na realidade, não existe. Patricia Churchland, em *Touching a Nerve: The Self as Brain* (2013), sublinha que a resistência em aceitar que os estados mentais são físicos decorre sobretudo de preconceitos linguísticos e de limitações intuitivas da cognição humana, e não de fundamentos empíricos. Assim, o facto de o SELF experienciar-se como distinto daquilo que o gera é apenas uma consequência da forma como o sistema se representa, e não uma indicação de dualidade ontológica.

Assim, como a consciência não é uma propriedade distinta do cérebro, torna-se difícil, senão impossível, distingui-la do mesmo. A consciência não é algo que acontece no cérebro como uma entidade separada; é algo que o cérebro faz – uma projeção do seu funcionamento. Assim, quando estamos conscientes somos, sem o sabermos, o resultado direto do funcionamento neuronal que não apenas causa a experiência, mas que é a experiência. Esta identidade entre causa e manifestação cria um paradoxo: a consciência, enquanto resultado desse funcionamento, não tem a capacidade de “sair” de si mesma para reconhecer a sua origem. A sua transparência – no sentido de Metzinger (2009) – impede que veja o cérebro como causa, porque é o próprio cérebro em ação.

Se pensarmos no SELF gerado por um código, encontramos uma analogia interessante. O código é a causa da existência do SELF e dos seus processos, mas o SELF não consegue aceder diretamente ao código que o sustenta. Para o SELF, o código é tão fundamental quanto invisível, precisamente porque ambos partilham as mesmas propriedades. O código é computacional e funcional, e o SELF é a sua manifestação fenomenológica. Esta diferença de aparência – entre o código funcional e o SELF fenomenológico – cria uma ilusão de separação. O SELF, com a sua subjetividade, não consegue conceber que algo tão diferente como o código possa ser a causa da sua existência. Se aprofundarmos a analogia, podemos pensar que mesmo que o SELF adquirisse capacidades avançadas de introspeção, continuaria limitado pela sua própria transparência. Ainda que pudesse representar aspetos do código que o cria, estas representações seriam, não só geradas pelo código, como seriam, elas próprias, moldadas pelas limitações fenomenológicas do SELF. Ou seja, o SELF é uma construção simplificada de algo muito mais complexo.

No mesmo sentido, o cérebro não representa o mundo de forma exaustiva e detalhada (Churchland, 2013; Wegner, 2002; Gazzaniga, 2000; Edelman, 1995). Em vez disso, constrói representações esquemáticas, ou seja, modelos simplificados da realidade que nos permitem interagir com o ambiente de forma eficiente (Dennett, 1991; Gazzaniga, 2011). Este processo acontece porque o cérebro não precisa de processar cada detalhe minucioso – apenas os aspetos relevantes para a sobrevivência e para a ação. O cérebro aplica essa mesma tendência de simplificação a si próprio. Quando tentamos perceber a nossa própria consciência, acabamos por gerar uma versão simplificada desses processos, ignorando os mecanismos informacionais e neurofisiológicos

subjacentes, o que explica as interpretações dualistas, que vão mais facilmente ao encontro dessa simplificação. Como refere o físico Brian Greene:

“Quando a tendência do cérebro para as representações esquemáticas simplificadas se aplica a si mesmo, à sua própria atenção, a descrição daí resultante ignora os processos físicos responsáveis por essa atenção. É por isso que os pensamentos e as sensações parecem etéreos, como se viessem de lado nenhum, como se pairassem dentro das nossas cabeças. Se a sua representação esquemática do seu corpo deixasse de fora os braços, o movimento das suas mãos também se parecia etéreo” (Greene, 2018, p. 152.)

Assim, o SELF nunca poderia alcançar um acesso direto ao código original, porque não consegue transcender a sua própria natureza representacional. Por outras palavras, a função da consciência é representar (o meio e o eu), pelo que a sua própria natureza representacional impede-a de sair dessa condição. Da mesma forma, a consciência humana, enquanto representação do funcionamento cerebral, não pode transcender as propriedades do cérebro para se ver como um produto do mesmo. Mesmo que formulemos teorias científicas sobre o cérebro e a consciência, essas teorias são elaboradas pela própria consciência, operando dentro dos limites impostos pelo cérebro que a causa. Da mesma forma que o batimento cardíaco é distinto do coração, que o oxigénio é distinto dos pulmões, ou que a digestão é distinta do sistema digestivo, não nos parece implausível que a consciência seja distinta do cérebro — sem que isso implique qualquer incoerência ontológica relativamente à fisicalidade de todos esses fenómenos.

Esta analogia leva-nos a considerar uma questão fundamental: a ilusão de separação entre a causa (o código, ou o cérebro) e a manifestação (o SELF, ou a consciência) pode ser um mecanismo intrínseco que permite a emergência da subjetividade. Se a consciência humana, ou o SELF, pudessem distinguir claramente a sua causa, talvez deixassem de funcionar como entidades subjetivas. A opacidade do cérebro para a consciência, ou do código para o SELF, não seria apenas uma limitação, mas um pré-requisito para a existência de fenómenos subjetivos — pois é precisamente essa barreira que permite a emergência de um ponto de vista interno, distinto do mecanismo que o gera. Dito de outra forma, sugerimos que a transparência da consciência em relação ao cérebro é condição necessária para ela ser aquilo que é. Se a consciência se fundisse com o cérebro, deixaria de estar subordinada à sua natureza representativa e teria uma outra

função que não gerar um modelo interno do eu, pelo que deixaria de ser uma experiência de subjetividade. Ou seja, paradoxalmente, para que a consciência possa existir como tal, ela tem de se manter funcionalmente separada daquilo que a origina.

No mesmo sentido, podemos especular sobre um limite teórico: se fosse possível criar um SELF com acesso direto ao código que o sustenta, isso não dissolveria a sua subjetividade? Teria o SELF um acesso recursivo infinito que o impediria de ter uma experiência fenomenológica? Do mesmo modo, se a consciência humana tivesse um acesso direto ao funcionamento cerebral como causa, será que a experiência consciente se fragmentaria, tornando impossível a continuidade do eu? Estas questões especulativas mostram como a analogia entre código e SELF pode ajudar-nos a pensar sobre a relação intrínseca que existe entre o cérebro e a consciência.

Na nossa analogia entre o problema SELF-código, o código é informação computacional (equiparado ao nosso subconsciente computacional) – é o que sustenta o SELF – ao passo que o SELF é informação representacional (mente consciente) e, precisamente por isso, ao SELF é vedada a possibilidade de conhecer a computação que permite a sua existência. Como o SELF e o código são, em grande medida, a mesma coisa, torna-se difícil para o SELF distinguir dois processos que são unos. Da mesma forma que o sujeito não tem qualquer indicador de que a sua consciência dependa de um cérebro (de informação computacional), porque a sua consciência está a ser gerada (e é) essa informação, também o SELF não teria forma de saber que existia um código que o estava a gerar, porque esse código (a computação) não só estava a permitir a sua existência, como estava a permitir a experiência dessa existência.

A relação entre cérebro e consciência obedece à mesma lógica. O cérebro não é apenas o suporte da experiência consciente, é também o próprio conteúdo funcional dessa experiência, o que torna impossível à consciência escapar ao seu horizonte representacional para se reconhecer como produto do cérebro. Para o SELF, tal como para os seres humanos, a consciência era evidente apenas porque era sentida — não havia qualquer conhecimento da sua origem. Não existia qualquer sinal, pista ou indicador que a sua consciência era gerada por um código. Era como se esse código não existisse porque o SELF era, ao mesmo tempo, esse código. Sem posse de mais informação, talvez o SELF criasse explicações sobrenaturais para o sentimento de existência que sentia; talvez rezasse a um deus para melhor compreender a sua natureza aparentemente imaterial.

Sem ter qualquer conhecimento das razões inerentes à sua existência, que mais explicações poderia o SELF elaborar? Mesmo que um dia lhe dissessem que a sua experiência podia estar a ser causada por um código, talvez ele desconfiasse que isso não podia ser possível, já que a sua experiência subjetiva era tão diferente do código que podia observar. O código sustentava o SELF, o código fazia emergir o SELF mas, ao mesmo tempo, a experiência subjetiva de SELF era diferente do código, o que levava a que o SELF não acreditasse que a sua experiência tão pessoal não poderia ser reduzida a uma computação. Sugiro a mesma solução para a consciência. A experiência subjetiva de qualquer ser *humana é causada por um cérebro mas, como nós próprios somos essa experiência causada, não temos possibilidade de a observarmos devido à sua transparência*. Da mesma forma que a experiência do SELF é aquilo que o programa faz, também a consciência dos seres humanos é aquilo que o cérebro faz.

À luz do pensamento dualista de propriedades, vários autores apressar-se-iam a dizer que aquilo que o código faz é um aspeto dual e fenomenológico – não-físico – da informação. Isto é, muitos filósofos poderiam concluir que esta ideia promove um dualismo de propriedades visto a parte física do sistema (o código) gerar um tipo de informação não-físico (a consciência do SELF).⁸⁸

No entanto, esta conclusão parece desnecessária. Podemos considerar que a consciência do SELF possui propriedades informacionais-físicas, sem recorrer a um dualismo. O SELF pode ter uma experiência fenomenológica sem que esta tenha de ser considerada não-física. O facto de a experiência fenomenológica não ser palpável, mensurável ou diretamente observável – de parecer epistemologicamente não-física – não implica que seja ontologicamente não-física. Fenómenos como a radioatividade ou o eletromagnetismo, entre muitos outros, foram outrora considerados não-físicos pela dificuldade de observação direta, mas hoje são entendidos como processos físicos. No

⁸⁸ Temos visto o mesmo tipo de raciocínio para o pensamento de John Searle (Dennett, 1991) e até para o pensamento de António Damásio (Bennett e Hacker, 2003), o que não nos parece sustentável. Aachamos, aliás, que há um erro epistemológico na leitura do pensamento daqueles autores, porque existe uma tendência de natural de dualizar propriedades que podem muito bem não ser dualizáveis. Mais: consideramos que esta interpretação é um reflexo da falta de rigor em distinguir entre fisicalismo emergentista e dualismo de propriedades. Damásio insere-se claramente na tradição fisicalista, apenas com um foco particular na biologia e na relação entre cérebro e corpo.

mesmo sentido, quando afirmamos que a consciência é, simultaneamente, aquilo que o cérebro faz e aquilo que o cérebro é, não estamos a promover nenhum tipo de dualismo. Estamos apenas a afirmar que a consciência é uma manifestação informacional-física de um cérebro informacional-físico. O cérebro não causa apenas a consciência – o cérebro causa uma série de processos neurofisiológicos – mas a consciência é uma das muitas coisas que o cérebro causa.

Como temos vindo a defender, o problema mente-cérebro é meramente epistemológico, não ontológico. Ou seja, é provavelmente a nossa incapacidade de compreender e de conhecer com rigor as propriedades físicas de alguns processos emergentes (Kant, 1781/2008) que conduzem a explicações dualistas (Chalmers, 1995 e 1996; O'Connor, 2000; Chisholm, 2003; Nagel, 1974; Hasker, 1999) (a nosso ver) desnecessárias. No âmbito da analogia do SELF, do ponto de vista ontológico *sabemos* que o código gerou uma consciência – o código criou o SELF –, e isso parece-nos suficiente para não tendermos a procurar explicações dualistas. O SELF e o código que o gerou possuem a mesma natureza informacional. O mesmo raciocínio pode ser aplicado ao problema mente-cérebro, nomeadamente aos *qualia*.

Por exemplo, será o *qualia* de dor um fenómeno não-físico? E a fome, a sede e a vontade de fazer xixi? Não nos parece que possamos designar essas experiências fenomenológicas de não-físicas. Tanto a fome, como a sede como a vontade de fazer xixi são processos fisiológicos que podemos medir e explicar (informação computacional) mas que também geram *qualia* (informação representacional), nomeadamente sensações e sentimentos. Serão estas sensações e sentimentos não-físicos? Será a dor de dentes um fenómeno não-físico? Podemos explicá-la, completamente, através de processos fisiológicos? Será suficiente que, o facto de não a podermos medir ou observar – ou sequer palpar – é suficiente para a designarmos de dor de não-física? Será suficiente o facto de ser experienciada subjetivamente na 1.ª pessoa e não existir mais nenhuma forma de experienciar esta situação na 3.ª pessoa, o suficiente para considerarmos a dor de dentes um fenómeno não-físico? Será que o facto de podermos saber tudo sobre a dor de dentes menos a experiência de dor de dentes (como a Dra. Mary em relação às cores (Jackson, 1982, 1986)), o suficiente para afirmarmos que a dor de dentes é não-física? Não nos parece. O facto de não podermos medir e observar a dor de dentes de uma forma tangível não a torna não-física. Desde logo porque o problema levantado neste

cenário é epistemológico, não é ontológico (Morch, 2023). Ontologicamente, temos toda a informação que podemos ter sobre a dor de dentes; da mesma forma que a dra. Mary possuía todo o *conhecimento* possível sobre as cores mas a sua experiência – a sua epistemologia – vedava-lhe o acesso a outras experiências. Como vimos no capítulo 1, a experiência de cor de Mary não nos diz nada sobre a ontologia da mente e da consciência, apenas da epistemologia.

Senão vejamos. Se não temos – se nunca tivemos – experiência da dor de dentes, é porque não temos o conhecimento fenomenológico dessa dor, mas não necessariamente desconhecemos o que a causa. Ou seja, sendo a dor de dentes um fenómeno com uma causa, qualquer pessoa sem a experiência do sentimento de dor poderia saber exatamente qual a sua origem, mesmo que nunca tivesse vivido essa experiência ou soubesse o que é sentir dor de dentes. Qualquer dentista pode, aliás, tratar uma dor de dentes sem nunca a ter sentido. Muitos médicos, por todo o mundo, tratam de inúmeras doenças sem nunca as terem experienciado. E não é por isso que desconhecem o que as causa. Não precisam de as ter para saber que a doença X é causada por Y. Tratando-se Y, a doença desaparece — e isso basta para estabelecer uma relação de causalidade. Sabemos, no mesmo sentido, que existe uma relação indiscutivelmente bem estabelecida entre a fisiologia da dor e o sentimento da dor (Damásio, 2010; Swaab, 2014; Ledoux, 1999). Esta relação causal é, aliás, suficiente para concluirmos que a dor de dentes não é um fenómeno não-físico. É uma sensação que emerge de um determinado padrão neurofisiológico. Mas, se a questão for epistemológica em vez de ontológica – o que é uma dor de dentes, em vez de o que a causa – então a experiência de dor de dentes torna-se necessária. Posso saber o que causa uma dor de dentes sem nunca a ter experienciado, mas só saberei o que é essa dor se a sentir.

Esta analogia permite compreender a relação entre o cérebro e a consciência. Também neste caso é possível afirmar, com base em evidência empírica (Damásio, 2010, 2011, 2013; Dehaene, 2014 e 2021; Tononi, 2016; Gazzaniga, 2011; Koch, 2016; Swaab, 2014; Edelman, 1995; Eagleman, 2012; Gazzaniga, 2000 e 2011; Ledoux, 2020), que certos padrões neurofisiológicos no cérebro causam estados conscientes, ainda que não possamos, a partir da experiência subjetiva, aceder diretamente a essa causalidade. A ausência de acesso fenomenológico ao substrato não invalida, contudo, a relação causal entre o cérebro e o que sentimos — apenas reforça a distinção entre conhecer a causa

e vivê-la. Como referido, da mesma forma que um dentista não precisa de ter a experiência de dor de dentes, ou que um médico não precisa de ter a experiência de uma doença para saber qual a sua causa, sugerimos que também não precisamos de ter a *experiência da causa da mente consciente* para identificarmos aquilo que a gera. Por conseguinte, não vemos razão para atribuir não-fisicalidade às experiências subjetivas.

Vejamos brevemente outro exemplo. Ao utilizarmos um programa como o Microsoft Windows, não vemos o código que o sustenta; e, ao acedermos ao código, não encontramos as janelas e imagens que caracterizam esse sistema operativo. No entanto, sabemos que o Windows depende inteiramente do código que o gera. Esta diferença entre manifestações não implica uma separação ontológica, mas apenas uma distinção funcional e representacional. O código e o Windows são expressões distintas de um mesmo substrato informacional. A lacuna entre ambos — essa impossibilidade de observar diretamente o código a gerar as janelas e imagens — não exige uma explicação dualista. Exige apenas o reconhecimento dos limites da observação direta na explicação causal.

A consciência coloca um problema análogo. Sabemos, com base em ampla evidência neurocientífica, que a mente consciente emerge do funcionamento do cérebro (Damásio, 2010, 2011, 2013; Dehaene, 2014 e 2021; Tononi, 2016; Gazzaniga, 2011; Koch, 2016; Swaab, 2014; Eagleman, 2012; entre outros). Sabemos, ainda, que as alterações na química neuronal produzem alterações na experiência consciente (Crippa et al., 2005; Siegel et al., 2024; Timmermann et al., 2019), o que reforça, claramente, a ideia de que a consciência é uma manifestação do que o cérebro faz — e não algo separado. Ainda assim, ao observarmos o cérebro, não vemos nenhum eu, tal como não vemos algoritmos ao usar um programa como o Windows. Isto levou muitos filósofos a supor uma natureza não-física para a mente. Contudo, essa inferência baseia-se numa falácia: confundir o que não é diretamente observável com o que é ontologicamente distinto. A lacuna explicativa entre cérebro e mente (Levine, 1983) é, a nosso ver, uma dificuldade epistemológica, não uma separação ontológica. A distinção entre programa e código, ou entre mente e cérebro, deve-se ao modo como acedemos ao fenómeno e não à existência de substâncias ou propriedades distintas. Quando sabemos, com elevado grau de certeza, que X causa Y, como o código causa o programa ou o cérebro

causa a mente, torna-se injustificado postular entidades não-físicas para colmatar lacunas na explicação.

Se um programador cria um código que gera um SELF com experiência subjetiva, sabemos que é o código que o possibilita. Do mesmo modo, se sabemos que a consciência desaparece em coma, sob anestesia ou após morte cerebral, isso confirma a dependência causal do fenómeno em relação ao cérebro. A consciência pode ser compreendida, no âmbito do nosso modelo, como o sistema nervoso a ganhar conhecimento de si (Metzinger, 2009). Em última análise, não somos algo além do nosso cérebro — somos o nosso cérebro a experienciar-se. Se o código não existir, não há SELF. No mesmo sentido, se o cérebro não estiver funcional, não há mente consciente. Essa é uma verdade ontológica que permanece válida, mesmo quando o acesso à experiência subjetiva se nos afigura limitado⁸⁹.

Em face do que foi exposto, torna-se claro que a distinção entre a consciência e o cérebro, ou entre o SELF e o código que o gera, é uma distinção apenas aparente — uma diferença de manifestação, não de natureza. A experiência subjetiva, embora distinta na forma como se apresenta, não exige uma explicação ontológica separada. Pelo contrário, como defendido, a consciência pode ser compreendida como uma manifestação representacional do funcionamento computacional do cérebro, do mesmo modo que o SELF é uma representação das computações do código que o sustenta. A ausência de acesso direto ao substrato causal não invalida a relação de causalidade; limita apenas o ponto de vista do sistema sobre si mesmo. É precisamente essa limitação — essa opacidade — que permite a emergência de um ponto de vista interno, estruturado e coeso.

Por conseguinte, tal como um programa não tem acesso ao seu próprio código, também a consciência não tem acesso direto ao cérebro que a gera. No entanto, a ciência permite-nos reconstruir essa relação causal com elevado grau de confiança, sem necessidade de recorrer a entidades não-físicas, como vimos. Assim, o problema mente-cérebro não nos remete para um abismo ontológico, mas para os limites da nossa epistemologia. A consciência não é algo que o cérebro tem — é aquilo que o cérebro faz. É aquilo que o cérebro é.

⁸⁹ Perguntei a algumas crianças de 7 anos se elas são os seus cérebros. Algumas ficavam ofendidas comigo por estar a sugerir algo aparentemente tão ridículo. Elas eram elas — o Gabriel ou a Maria — não os seus cérebros.

Conclusão

Ao longo desta tese, investigámos a natureza da consciência a partir de uma abordagem monista-informacional, explorando a sua emergência, função e relação com a ação, a decisão e o livre-arbítrio. A partir da Substância Unificada Informacional (SUI), propusemos um modelo que integra a consciência no domínio da fisicalidade, rejeitando tanto o dualismo de substância como o dualismo de propriedades. Esta abordagem permitiu reinterpretar questões clássicas da filosofia da mente, como o problema do conhecimento, o problema mente-corpo e o problema do epifenomenismo da consciência, sob a luz de uma ontologia informacional da realidade.

No primeiro capítulo, realizámos uma análise crítica das principais teorias contemporâneas da consciência, começando pelo dualismo de propriedades e pelas abordagens que sustentam que a consciência não pode ser reduzida a um fenómeno meramente físico. Explorámos a proposta de David Chalmers, que distingue entre os *easy problems* e o *hard problem* da consciência, sugerindo que os processos funcionais do cérebro podem ser explicados por mecanismos fisicalistas, mas que a experiência subjetiva – a fenomenologia – permanece inexplicável dentro desse quadro. Para Chalmers, a consciência é um aspeto fundamental da realidade, podendo estar associada a propriedades intrínsecas da informação, hipótese que leva a um género de pampsiquismo informacional, que encontra algumas afinidades com outros autores, como Thomas Nagel. Além disso, analisámos o dualismo de propriedades como um todo (em função de autores como Hasker, O'Connor, Jackson, entre outros), identificando os seus desafios centrais: a dificuldade em explicar a relação causal entre estados físicos e estados fenomenológicos, a falta de evidências empíricas que sustentem a autonomia da consciência em relação ao cérebro e a sua incapacidade de resolver o problema mente-corpo sem introduzir novas dicotomias ontológicas.

Em contraste, examinámos as principais teorias fisicalistas da consciência, incluindo o espaço neuronal de trabalho (GNW⁹⁰), de Dehaene; a teoria da informação integrada (IIT⁹¹) de Tononi e Koch; e o fisicalismo corporificado de Damásio. Apesar das

⁹⁰ Global Network Workspace.

⁹¹ Integrated Information Theory.

diferenças conceptuais, todas convergem na ideia de que a consciência emerge da capacidade dos sistemas biológicos de processarem informação de forma integrada e autorreferencial. O modelo de Dehaene propõe que a consciência resulta da globalização da informação no cérebro, enquanto a IIT sugere que a consciência depende do nível de integração e complexidade informacional. Damásio, por sua vez, argumenta que a consciência está enraizada nos processos homeostáticos e nos estados interoceptivos do organismo, sendo uma manifestação da autorepresentação biológica.

O segundo capítulo centrou-se na relação entre consciência e causalidade, explorando o problema do livre-arbítrio.

Começamos por analisar o determinismo, o libertarismo e o compatibilismo. Verificámos que diversos filósofos, como Derk Pereboom, Richard Double, Ted Honderich e Espinosa, defendem que o Universo é um sistema determinístico que obedece necessariamente às leis da Natureza, razão pela qual consideram impossível a existência de livre-arbítrio. Em contraste, os filósofos libertários, como Timothy O'Connor, Robert Kane, Mark Balaguer, Hugh McCann e Carl Ginet, defendem, em linhas gerais, que o ser humano é livre porque as suas ações não são determinadas por causas antecedentes nem pelas leis da Natureza. Por fim, os compatibilistas, como Daniel Dennett ou R.E. Hobart, entre outros, sustentam que o determinismo é compatível com a liberdade humana, admitindo que o ser humano pode agir livremente mesmo num Universo regido por leis causais.

Além disso, analisámos criticamente o libertarismo, especialmente as interpretações da física quântica que procuram sustentá-lo, bem como alguns dos problemas centrais levantados pelos filósofos libertários, como a ideia de que o sujeito só pode decidir livremente violando as leis da Natureza. Vimos, assim, que o libertarismo assenta, sobretudo, em quatro pressupostos problemáticos, que analisámos com alguma profundidade: (1) a possibilidade de eventos *acausais*, o que contradiz os próprios princípios do Universo, já que qualquer evento sem causa implicaria uma violação das leis da natureza; (2) a ideia de que a aleatoriedade e a probabilidade são propriedades fundamentais da realidade, e não meros conceitos epistemológicos usados para descrever fenómenos; (3) a crença de que o indeterminismo quântico se aplica diretamente a sistemas biológicos e cognitivos, apesar de não existir qualquer evidência de que os

princípios da física quântica sejam relevantes para o funcionamento neuronal; (4) a suposição de que pode existir uma consciência imaterial, autónoma e independente do cérebro — uma hipótese que carece de qualquer sustentação empírica e que contraria os dados disponíveis nas neurociências.

Por conseguinte, sugerimos que única forma coerente de entender o livre-arbítrio é dentro de um quadro determinista. O livre-arbítrio, longe de ser um fenómeno metafísico ou sobrenatural, é um *sentimento* subjetivo de liberdade que emerge da complexidade das interações entre os sistemas cerebrais. Se as nossas decisões fossem indeterminadas, então não poderiam ser atribuídas a uma vontade consciente e racional. Em vez disso, o que observamos é que a experiência do livre-arbítrio está ancorada na organização causal do cérebro, que constrói um modelo interno de ação e decisão. Como defendemos, apenas num universo determinista é possível haver livre-arbítrio, precisamente porque a causalidade permite que a consciência tenha um papel funcional na regulação do comportamento. Assim, não apenas rejeitamos as interpretações indeterministas da física quântica, como também propomos que o livre-arbítrio só pode ser entendido coerentemente como uma manifestação da organização cerebral – em forma de sentimento – num universo determinista. Em suma, propomos que o determinismo não é o oposto da liberdade, mas sim a única base coerente para compreendê-la.

Neste capítulo, vimos ainda que as experiências de Benjamin Libet, e os inúmeros estudos que as confirmaram e expandiram, vieram abalar profundamente a conceção tradicional do livre-arbítrio. Libet demonstrou que a atividade neuronal preparatória (potencial de prontidão) antecede em várias centenas de milissegundos a percepção consciente da vontade de agir, o que sugere que o cérebro inicia os actos antes de a consciência ter qualquer sensação de decisão. A decisão consciente surge, assim, não como origem do acto, mas como sua representação tardia. Esta conclusão foi reforçada por estudos posteriores, como os de Castiello, que demonstrou que o movimento precede a consciência do mesmo; de Haggard, que explorou detalhadamente a dissociação temporal entre intenção e ação consciente; de Desmurget, que conseguiu induzir vontades ilusórias ao estimular diretamente certas zonas cerebrais; de Soon e Bode, que anteciparam decisões até dez segundos antes de estas emergirem na consciência; e de Fried, que confirmou que a vontade consciente apenas emerge quando um determinado limiar de atividade subconsciente é atingido.

A implicação central destes dados é que a consciência não decide, mas toma conhecimento de decisões já tomadas por processos neuronais automáticos. A consciência seria, nesse sentido, um epifenómeno – uma espécie de palco representacional onde os actos já em curso ganham forma simbólica e narrativa. Esta interpretação foi defendida de forma particularmente influente por Daniel Wegner, que propôs que o livre-arbítrio é apenas um sentimento reconstruído à posteriori, uma ilusão funcional que contribui para a coesão da identidade pessoal.

As críticas aos estudos libetianos — e aos que lhes seguiram — como as de Gallagher ou Mele, que apontam a simplicidade das tarefas experimentais (como levantar um dedo ou o pulso), foram já devidamente respondidas. Demonstrou-se, entretanto, que mesmo decisões mais complexas, conscientes e ponderadas podem ser antecipadas neuronalmente, através de padrões cerebrais identificáveis antes da formação consciente da vontade. A distinção entre ações simples e decisões complexas, embora relevante sob o ponto de vista fenomenológico, não invalida o argumento central: a consciência não constitui a origem última do ato, mas sim um estado que emerge após processos cerebrais já em curso.

Autores como Jeannerod, Metzinger e Seth sustentam, na mesma linha que Libet, Haggard ou Wegner, que a consciência é um processo lento e representacional, cuja função principal é criar uma simulação integrada e coerente do eu e do mundo. Tal como as cores não existem na realidade física mas são interpretações cerebrais de determinados comprimentos de onda, também o eu (ou *self*) e o sentimento de liberdade são, na ótica daqueles autores, interpretações neurais – construções simbólicas adaptativas. Neste quadro, a vontade, o desejo e a liberdade não são expressões de um sujeito autónomo, mas estados induzidos por mecanismos neurofisiológicos que regulam o comportamento de acordo com finalidades biológicas. O sujeito sente que decide, mas está apenas a experienciar um impulso que lhe é imposto por processos cerebrais automáticos.

Esta leitura é ainda reforçada pelas investigações de Gazzaniga, que revelou o papel do hemisfério esquerdo como “intérprete” – isto é, como gerador de narrativas que explicam racionalmente actos iniciados por mecanismos inconscientes. A consciência funciona como um contador de histórias: não decide, mas justifica, não age, mas simboliza. O livre-arbítrio emerge assim como um sentimento de *agency* e de

identidade, não como uma capacidade real de autodeterminação causal. A liberdade não está na origem das decisões, mas na forma como as vivenciamos e interpretamos. É por isso que, embora acreditemos estar a escolher livremente, estamos, na verdade, a assistir à emergência simbólica de escolhas que nos acontecem, não que produzimos.

No terceiro capítulo sugerimos a substância unificada informacional (SUI), que propõe, de uma forma genérica, que a realidade não se divide entre físico e não-físico, mas que corpo e mente são manifestações físicas da mesma estrutura ontológica informacional.

Propomos que a SUI é o princípio ontológico fundamental que sustenta todo o Universo — uma matriz contínua de informação estruturada que se manifesta sob diferentes formas, como partículas, campos, organismos ou representações mentais. Esta substância única não é composta por matéria nem por fenomenologia, mas por informação que se desdobra em diferentes níveis de complexidade. Nesse sentido, a SUI permite integrar fenómenos físicos, biológicos e mentais num único modelo, rejeitando a cisão dualista (tanto de substância como de propriedades) e sustentando que tudo o que existe é expressão necessária de uma única realidade informacional — causal, determinística e organizada segundo princípios que são, em si mesmos, expressões da estrutura da informação.

A informação, longe de ser um conceito meramente descritivo, assume aqui um papel fundamental na própria estrutura do real — *é a unidade codificadora cujo processamento permite o surgimento de fenómenos emergentes*. Não se trata apenas de dados ou sinais físicos, mas de configurações com capacidade de gerar estados e relações causais. Cada bit de informação é, no nosso modelo, um elemento estruturado que, ao ser processado, dá origem a propriedades, comportamentos e estruturas — desde partículas elementares até redes neuronais complexas. Neste sentido, a informação é simultaneamente forma e função, código e dinâmica, constituindo a linguagem ontológica do Universo.

Propusemos, por isso, uma estratificação conceptual da informação em três níveis ontológicos interdependentes: 1) a informação ontológica estrutural; 2) a informação ontológica computacional; e 3) a informação representacional. A informação estrutural corresponde ao nível mais profundo — a arquitetura invisível do real, composta

por leis, simetrias e algoritmos que regem o comportamento do Universo. Esta estrutura não é diretamente observável, mas constitui o código-fonte que permite a emergência de sistemas físicos. A partir dela emerge a informação computacional — isto é, toda a informação fisicamente ativa: partículas, campos, interações e sistemas que processam estados e geram efeitos causais. Esta informação é computada continuamente, mesmo na ausência de qualquer representação. Por fim, em determinados sistemas altamente organizados — nomeadamente os sistemas nervosos — a informação computacional pode dar origem a informação representacional, isto é, a imagens mentais que representam o corpo, o meio ou o próprio sistema. Estas representações podem ser conscientes (como nos organismos capazes de gerar modelos internos do eu) ou inconscientes (como nos organismos que possuem uma mente não consciente, sem capacidade de autorreferência). No nosso modelo, a consciência emerge apenas quando essas representações atingem um nível adicional de integração recursiva, capaz de gerar modelos internos autorreferenciais. Assim, toda a consciência é representacional, mas nem toda a representação é consciente. A mente, enquanto domínio das representações, é mais ampla do que a consciência, que constitui apenas uma fração — ainda que funcionalmente central — desse domínio.

Neste contexto, sugerimos a *hipótese do desdobramento informacional*. Segundo esta proposta, a informação estrutural — que constitui a ontologia fundamental da realidade — desdobra-se em informação computacional, isto é, em padrões físico-funcionais que organizam a matéria e a energia segundo regras causais. Apenas posteriormente, em sistemas com elevada complexidade e capacidade de integração, como os sistemas nervosos, essa informação computacional pode ser desdobrada em informação representacional, consciente ou inconsciente. A informação representacional, uma vez tornada possível pelo desdobramento informacional (biológico), pode ser reintroduzida no sistema sob a forma de novas computações, influenciando decisões, comportamentos e estados fisiológicos. Assim, o desdobramento informacional é um processo recursivo. Parte da informação estrutural, gera computação funcional, e desta pode emergir representação, que, por sua vez, pode retornar ao plano computacional como instrução. Esse ciclo dinâmico — estrutural, computacional e representacional — constitui, no nosso modelo, o mecanismo de emergência, integração e retroação da mente e da consciência no Universo.

Assim, no âmbito da nossa hipótese do desdobramento informacional, sugerimos o conceito de *recursividade informacional*. Este surge como um mecanismo essencial, permitindo que sistemas avancem de simples processamento de dados para representações cada vez mais complexas — incluindo a autorrepresentação e a metacognição. Isto é possível porque a recursividade permite, por um lado, a transformação de informação computacional em representacional (com ou sem acesso consciente), e, por outro, a transformação inversa: da representação em nova computação — o que permite atualizar os estados internos do sistema com base na experiência.

Propusemos que este nível de complexidade é sustentado, neuroanatomicamente, pela ativação recorrente de regiões do córtex frontoparietal — nomeadamente através da dinâmica proposta no GNW —, que serve como plataforma funcional para a integração global da informação, ou seja, para a re-representação de estados subjetivos, o que, no nosso modelo, corresponde à *consciência de segunda ordem*. Além disso, sugerimos que os neurónios von Economo (VENs), localizados em regiões associadas à cognição social e à consciência do self — como a ínsula e o córtex cingulado anterior —, parecem desempenhar um papel essencial na *consciência de primeira ordem*, mais próxima da interoção e da vivência emocional do corpo em forma de sentimentos.

Vimos, também, que o nosso modelo apresentou respostas para três problemas filosóficos fundamentais: 1) o problema do conhecimento; 2) o problema mente-corpo; e 3) o problema do epifenomenismo da consciência.

No que respeita ao problema do conhecimento, o nosso modelo oferece uma solução ao definir a informação como base da epistemologia. A proposta aqui desenvolvida apresenta uma resposta monista, sustentada numa ontologia informacional integrada, alicerçada em três princípios fundamentais: a universalidade ontológica da SUI, a recursividade funcional dos sistemas nervosos e o papel da luz como interface informacional. Ao considerarmos que tudo o que existe é expressão de uma mesma estrutura informacional, dissolve-se a clássica separação entre sujeito e objeto, entre mente e mundo. No nosso modelo, o conhecimento não resulta de formas a priori nem de conteúdos empíricos isolados, mas da interação dinâmica entre estruturas informacionais que se refletem mutuamente. O sujeito conhece porque é parte do mesmo tecido informacional que organiza o mundo; não há mediação transcendental, mas continuidade estrutural. Assim, a mente não impõe categorias, nem o mundo se limita a fornecer

dados brutos: ambos são expressões da mesma lógica informacional. O conhecimento é, por isso, o reconhecimento de padrões recursivos num campo comum — não uma projeção, nem uma extração, mas uma convergência.

O conhecimento deixa, assim, de ser concebido como um espelho ou uma cópia de uma realidade exterior e passa a ser reconhecido como um processo interno de reorganização da própria informação que estrutura o Universo. Não existe, por isso, um fosso entre o mundo e a mente; ambos são expressões distintas — mas compatíveis — de uma mesma realidade informacional. Neste modelo, o conhecimento surge como um desdobramento da informação ontológica estrutural em dois outros níveis: a informação computacional e a informação representacional. A primeira organiza os sistemas físicos (como os sistemas nervosos), enquanto a segunda corresponde à *representatividade* desses sistemas, isto é, à forma como esses sistemas mapeiam, integram e representam a realidade. Ao permitir a transformação recursiva entre esses níveis, os sistemas nervosos conscientes tornam-se “tradutores” ativos da estrutura do mundo, dando origem, no caso dos seres humanos, à percepção, à consciência e ao conhecimento.

A luz desempenha, neste enquadramento, um papel decisivo enquanto interface que liga sistemas distintos dentro da mesma estrutura informacional. A luz transporta, entre o objeto e o observador, a informação que será decodificada e transformada em representação consciente. Assim, ela é uma das formas através das quais a informação ontológica se torna acessível, legível e experienciável. O sistema nervoso, ao interpretar os padrões informacionais inscritos na luz, transforma-os em representações mentais e codifica-os em informação computacional (o que enriquece o sistema, que pode ter mais opções de ação). A memória, a percepção e o pensamento são, deste modo, produtos de um processo de recodificação informacional.

No que concerne ao problema mente-corpo, sugerimos uma solução assente em três pilares: a recursividade informacional, a especificidade neurobiológica e um fisicalismo informacional compatível com a visão naturalista de Espinosa. A mente não é aqui entendida como uma entidade separada do corpo, mas como uma expressão recursiva de certos sistemas que, além de computarem informação, são capazes de a representar. Essa capacidade recursiva — de representar estados internos e de aceder à informação que o próprio sistema produz — é possibilitada pela organização específica de circuitos

neurobiológicos, como o córtex frontoparietal, associado à atenção, decisão e integração multisensorial, e pelos neurónios Von Economo (VENs), que mapeiam estados homeostáticos e contribuem para um modelo de eu sentido. São estes sistemas que permitem a emergência de uma mente capaz de gerar modelos internos do mundo e de si mesma. A própria estrutura neurofisiológica do cérebro, por conseguinte, não apenas computa estados, mas também os representa, interpreta e reorganiza em ciclos de crescente complexidade. Neste modelo, a informação resolve o problema mente-corpo não ao postular dois domínios distintos, mas ao mostrar que aquilo a que chamamos mente é uma forma de organização física da própria informação. Tal como Espinosa sugeriu, a mente e o corpo não são substâncias diferentes, mas dois modos de expressão da Natureza — neste caso, da informação, que se desdobra sob diferentes formas: estrutural, computacional e representacional.

Sugerimos, ainda, uma resposta ao problema do epifenomenismo da consciência. A teoria aqui desenvolvida oferece uma alternativa à hipótese epifenomenalista, ao considerar que a consciência não é apenas um produto colateral do cérebro, mas uma interface funcional de decodificação e integração informacional. Ainda que possa não possuir causalidade direta sobre o comportamento — como sugerem Libet, Haggard ou Wegner, entre muitos outros —, a consciência desempenha um papel essencial na transformação de informação computacional em informação representacional, e vice-versa. Esta transformação recursiva não é meramente decorativa, mas é essencial para que o cérebro possa enriquecer os seus circuitos com novos dados, representações e experiências, enraizando-os como instruções úteis para decisões futuras.

A consciência funciona, neste modelo, como um módulo interno de “tradução” informacional, que permite aos organismos *transformar* bits provenientes do corpo e do meio em experiências organizadas, emocionalmente qualificadas e dotadas de valor adaptativo. Ao integrar estados fisiológicos com representações sensíveis (*qualia*), a consciência permite ao cérebro codificar conhecimento sobre o mundo e sobre si mesmo, contribuindo de forma indireta, mas decisiva, para a eficiência e flexibilidade do comportamento. Ao invés de ser vista como um “fumo sem efeito” — como propôs Huxley —, a consciência deve ser compreendida como um canal interno de enriquecimento informacional, que amplia a capacidade do sistema nervoso para simular, ajustar, antecipar e aprender. A sua função não reside em iniciar a ação, mas em integrar e

organizar os conteúdos que a tornam mais eficaz. Assim, não há lugar para um epifenômeno puro: a consciência é uma expressão recursiva e funcional do cérebro — uma ferramenta informacional criada pela própria estrutura neural para se informar, se reconfigurar e se orientar no mundo com maior profundidade e eficiência.

Por fim, para clarificar a nossa proposta, recorremos a uma analogia com um sistema artificial — o SELF — que, embora criado por um código, é capaz de gerar representações e, eventualmente, consciência. Esta analogia permitiu ilustrar que um sistema pode transformar informação computacional em informação representacional sem ser, por isso, consciente — como sucede com sistemas artificiais ou com organismos dotados de um sistema nervoso mais simples. No entanto, quando a consciência emerge, confere ao sistema uma maior capacidade de integração, aprendizagem e decisão. A capacidade do SELF de transformar informação computacional em representações do meio e, mais tarde, de si próprio, confere-lhe vantagens claras: decisões mais flexíveis, adaptativas e informadas.

Assim, ao contrário do que muitos autores defendem, a consciência participa efetivamente na decisão — não como agente direto, mas como estrutura recursiva que organiza a informação representacional do meio e do eu. O ponto central da analogia reside na ideia de que a consciência, tal como o SELF, é uma manifestação do funcionamento do sistema que a gera — sendo esse sistema, no nosso caso, o cérebro. A aparente separação entre o cérebro e a mente — ou entre o código e o SELF — deve-se à transparência da consciência: ela não se vê como produto daquilo que a gera, precisamente porque é o próprio funcionamento desse sistema em forma experiencial. A experiência sentida — ou consciência fenomenológica — é, assim, a manifestação natural da própria consciência enquanto processo funcional capaz de integrar, representar e enraizar informação de modo recursivo, permitindo-lhe aumentar o espectro da decisão e enriquecer a interação entre o organismo e o meio.

Por isso, propomos que o problema mente-cérebro é epistemológico, e não ontológico. A consciência não é algo separado do cérebro, mas aquilo que o cérebro faz — uma projeção informacional recursiva que representa o mundo, o corpo e o eu. Tal como o código sustenta o SELF, o cérebro sustenta a mente. E, se há uma lacuna explicativa entre ambos, ela não exige uma resposta dualista, mas sim o reconhecimento dos limites do nosso acesso epistemológico àquilo que somos.

Referências bibliográficas

- AARONSON, Scott (2013). *Quantum Computing since Democritus*. New York: Cambridge University Press.
- ADOLPH, K. E.; EPPLER, M. A.; MARIN, L.; WEISE, I. B.; WECHSLER CLEARFIELD, M. (2003). Learning to walk: How infants integrate perceptual information and motor skill. *Psychological Science*, 14(4), 306–312. <https://doi.org/10.1111/1467-9280.14461>
- AFSHORDI, N., CORIANO, C., DELLE ROSE, L., GOULD, E., & SKENDERIS, K. (2017) From Planck data to Planck era: Observational tests of Holographic Cosmology. *Physical Review Letters*, 118(4), 2017.
- ALEATORIDADE (2024). *Dicionário Priberam da Língua Portuguesa [on-line]*. 2008-2024. <https://dicionario.priberam.org/probabilidade>. Acedido a 14-08-2024.
- ALLMAN, J. M.; TETREAULT, N. A.; HAKEEM, A. Y.; MANAYE, K. F. (2011). The von Economo neurons in the frontoinsular and anterior cingulate cortex. *Annals of the New York Academy of Sciences*, 1225(1), 59–71.
- ALVES, J. A. (2013). *Limites da consciência: O meio segundo de atraso e a ilusão de liberdade*. Fronteira do Caos.
- AMADOR, N.; FRIED, I. (2004). Single-neuron activity in the human supplementary motor area underlying preparation for action. *Journal of Neurosurgery*, 100, 250–259.
- AMMON, K.; GANDEVIA, S. C. (1990). Transcranial magnetic stimulation can influence the selection of motor programmes. *Journal of Neurology, Neurosurgery, and Psychiatry*, 53, 705-707.
- ANANTHASWAMY, Anil (2015). *The Man Who Wasn't There: Investigations into the Strange New Science of the Self*. New York: Dutton.
- AOYAMA, T.; HAYAKAWA, M.; KINOSHITA, T.; NIO, M. (2012). Complete Tenth-Order QED Contribution to the Electron $g-2$. *Physical Review Letters*, 109(11), 111807.
- ARMSTRONG, D. M. (1968). *A Materialist Theory of the Mind*. Routledge & Kegan Paul.
- ASPECT, A.; DALIBARD, J.; ROGER, G. (1982). Experimental test of Bell's inequalities using time-varying analyzers. *Physical Review Letters*, 49(25), 1804–1807. <https://doi.org/10.1103/PhysRevLett.49.1804>.

ASPURU-GUZI, A., & WALTHER, P. (2012) Photonic quantum simulators for chemistry. *Nature Physics*, 8(4), 285–291.

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

BAARS, B. J. (1997). In the Theater of Consciousness: Global Workspace Theory, A Rigorous Scientific Theory of Consciousness. *Journal of Consciousness Studies*, 4(4), 292-309.

BAARS, B. J. (2005). *Global Workspace Theory of Consciousness: Toward a Cognitive Neuroscience of Human Experience*. Progress in Brain Research, 150, 45-53.

BALAGUER, M. (2004). A Coherent, Naturalistic, and Plausible Formulation of Libertarian Free Will. *Noûs*, 38(3), 379-406.

BALAGUER, M. (2014). *Free Will*. MIT Press.

BALAGUER, M. (2010). *Free will as an open scientific problem*. Cambridge, MA: MIT Press.

BANNISTER, Scot; EEROLA, Tuomas (2018). Suppressing the chills: Effects of musical manipulation on the chills response. *Frontiers in Psychology*.

BANKS, W. B. (2006) Does Consciousness Cause Misbehavior? *Does Consciousness Cause Behavior?* London: MIT Press.

BARGH, J. A. (2017). *Before you know it: The unconscious reasons we do what we do*. Touchstone.

BARNEY, R. (2012). Plato on the Soul and the Body. In G. Fine (Ed.), *The Oxford Handbook of Plato*. Oxford University Press.

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, 16(7), 419–429. <https://doi.org/10.1038/nrn3950>

BARRETT, L. F., QUIGLEY, K. S., BLISS-MOREAU, E., & ARONSON, K. R. (2004). Interoceptive sensitivity and self-reports of emotional experience. *Journal of Personality and Social Psychology*, 87(5), 684–697. <https://doi.org/10.1037/0022-3514.87.5.684>

BECHARA, Antoine; NAQVI, Nasir (2004). Listening to your heart: Interoceptive awareness as gateway to feeling. *Nature Neuroscience*, 7(2).

BEKENSTEIN, J. (2003). Information in the Holographic Universe. *Scientific American*, 289(2), 58–65. <https://www.scientificamerican.com/article/information-in-the-holographic-univ/>.

BELL, J. S. (1964). On the Einstein Podolsky Rosen paradox. *Physics Physique Физика*, 1(3), 195–200. <https://doi.org/10.1103/PhysicsPhysiqueFizika.1.195>.

BERENGUER, R. (2016). *Mundo Quântico*. Lisboa: Atlântico Press.

BERRIDGE, C. W.; WATERHOUSE, B. D. (2003). The locus coeruleus–noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews*, 42(1), 33–84.

BERTENTHAL, B. I.; CLIFTON, R. K. (1998). Perception and action in infants. *Current Directions in Psychological Science*, 7(5), 141–146. <https://doi.org/10.1111/1467-8721.ep10836818>.

BLACKMORE, S. (2005). *Consciousness: An Introduction*. London/New York: Routledge.

BLOCK, N. (1995). On a Confusion about a Function of Consciousness. *Behavioral and Brain Sciences*, 18(2), 227–287.

BLOOM, Paul (2005). *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human*. New York: Basic Books.

BODE, S.; BODE, P.; SOON, C. S. (2014). Demystifying ‘free will’: The Role of Contextual Information and Evidence Accumulation for Predictive Brain Activity. *Neuroscience & Biobehavioral Reviews*, 47, 636–645. <https://doi.org/10.1016/j.neubio-rev.2014.10.017>.

BODE, S.; HE, A. H.; SOON, S. C.; TRAMPEL, R.; TURNER, R.; JOHN-DYLAN HAYNES, J. D. (2011). Tracking the Unconscious Generation of Free Decisions Using Ultra-High Field fMRI. *PLoS ONE*, 6(6), e21612. <https://doi.org/10.1371/journal.pone.0021612>.

BOHM, David (1952). *A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables I & II*. *Physical Review*, 85, 166–193.

BOHM, David (1980). *Wholeness and the Implicate Order*. New York: Routledge & Kegan Paul.

BOHR, N. (1958). *Atomic Physics and Human Knowledge*. New York: Wiley.

BOLY, M.; MASSIMINI, M.; TSUCHIYA, N.; POSTLE, B. R.; KOCH, C.; TONONI, G. (2017). Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? *Clinical and Neuroimaging Evidence*. *Journal of Neuroscience*, 37, 9603–9613. <https://doi.org/10.1523/jneurosci.3218-16.2017>.

BOMBARDA, M. (1902). *A Consciência e o Livre Arbítrio*. Lisboa: Livraria de António Maria Pereira.

BOUSSO, R. (2002). The holographic principle. *Reviews of Modern Physics*, 74(3), 825–874. <https://doi.org/10.1103/RevModPhys.74.825>.

BRAUNSTEIN, J. F.; PEWZNER, E. (2003). *História da Psicologia*. Lisboa: Instituto Piaget.

BRICKHOUSE, T. C.; SMITH, N. D. (2010). *Socrates on Self-Knowledge and Virtue*. Cambridge University Press.

BRIEGEL, H. J.; MULLER, T. (2015). A Chance for Attributable Agency. *Minds & Machines*, 25, 261–279. <https://doi.org/10.1007/s11023-015-9381-y>.

BROCKMAN, J. (2006). *What We Believe But Cannot Prove: Today's Leading Thinkers on Science in the Age of Certainty*. Disponível em: <http://www.edge.org/annual-question/what-do-you-believe-is-true-even-though-you-cannot-prove-it>.

CANAL, R.; MORAES, J. A. (2009) Chalmers e Searle nos Estudos da Consciência: Alguns Avanços? *Ciências & Cognição* vol. 14 (2), 262-275.

CARHART-HARRIS, R. L., & FRISTON, K. J. (2019). Rebus and the anarchic brain: Toward a unified model of the brain action of psychedelics. *Pharmacological Reviews*, 71(3), 316–344. <https://doi.org/10.1124/pr.118.017160>.

CARROLL, S. (2019). *Something deeply hidden: Quantum worlds and the emergence of spacetime*. New York: Dutton.

CARTER, A., & HALL, W. (2018). *Addiction Neuroethics: The Promises and Perils of Neuroscience Research on Addiction*. Cambridge University Press.

CARUSO, G. (2012) *Free Will and Consciousness: A Determinist Account of the Illusion of Free Will*. Lexington Books: New York.

CASIMIR, H. B. G. (1948). On the Attraction Between Two Perfectly Conducting Plates. *Proceedings of the Royal Netherlands Academy of Arts and Sciences*, 51, 793–795.

CASTIELLO, U.; PAULIGNAN, Y.; JEANNEROD, M. (1991) Temporal Dissociation of Motor Responses and Subjective Awareness: A Study in Normal Subjects. *Brain* 114, pp. 2639-2655.

CHALMERS, D. (1995) Facing up to the Problem of Consciousness. *Journal of Consciousness Studies*, 2, N.º 3, pp. 200-219.

- CHALMERS, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: University Press.
- CHAPPEL, V. C. (1994). *Locke's Philosophy of Mind*. Oxford University Press.
- CHISHOLM, R., M. (2003) Human Freedom and the Self. In Gary Watson, ed., 2nd ed., *Free Will*. Oxford: Oxford University Press, 2003.
- CHISHOLM, R. (1966) Freedom and Action. In Keith Lehrer, ed., *Freedom and Determinism*, New York: Random House.
- CHOUDHURY, S.; BLACKEMORE, S. (2006) Intentions, Actions, and the Self. In *Does Consciousness Cause Behavior?* Massachusetts: MIT.
- CHURCHLAND, P. (1998) *Matéria e Consciência. Uma introdução à filosofia da mente*. São Paulo: UNESP.
- CHURCHLAND, P. (1981) Eliminative materialism and the Propositional Attitudes. *The Journal of Philosophy*, Vol. 78, N. 92, 67-90.
- CHURCHLAND, P. (1985) Reduction, Qualia, and the Direct Introspection of Brain States. *Journal of Philosophy* 82, n. 1.
- CHURCHLAND, P. (1983) Consciousness: the transmutation of a concept. *Pacific Philosophical Quarterly* 64, 80-95.
- CHURCHLAND, P. (2013). *Touching a Nerve: Our brains, Our selves*. New York: W. Norton & Company.
- CLARKE, R.; CAPES, J.; Swenson, P. (2021) Incompatibilist (Nondeterministic) Theories of Free Will. *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2021/entries/incompatibilism-theories/>>.
- CLEEREMANS, A.; BAUDRY-TALLON, C. (2022). Consciousness matters: phenomenal experience has functional value. *Neuroscience of Consciousness* (1), 1–11.
- COLE, J. (1995). *Pride and a daily marathon*. Cambridge, MA: MIT Press.
- CMS COLLABORATION. (2020). Search for millicharged particles in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Physical Review D*, vol. 102, no. 3.
- CORNSWEET, T. (1970) *Visual Perception*. New York: Academic Press.
- CORBETTA, M., & SHULMAN, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201-215.

CRAIK, K. (1943) *The nature of explanation*. Cambridge: Cambridge University Press, 1943.

CRICK, F. (1998). *A hipótese espantosa. Busca científica da alma*. Lisboa: Instituto Piaget.

CRIPPA, J. A., LACERDA, A. L. T., AMARO, E., BUSATTO FILHO, G., ZUARDI, A. W., & BRESSAN, R. A. (2005). Efeitos cerebrais da maconha: resultados dos estudos de neuroimagem. *Revista Brasileira de Psiquiatria*, 27(1), 70–78.

CURADO, M. (2005). Bombarda e a consciência. *Jornal de Ciências Cognitivas*.

CURADO, Sílvia (2017) *Engenharia Genética. O Futuro Já Começou*. Lisboa: Fundação Luso-americana.

DAMÁSIO, A. (2023). Feelings Are the Source of Consciousness. *Neural Computation*, 35, 277–286.

DAMÁSIO, A. et al. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature America*, Vol. 3, N. 10.

DAMÁSIO, A. (2003). *Ao encontro de Espinosa. As emoções sociais e a neurobiologia do sentir*. Lisboa: Europa-América.

DAMÁSIO, A. (2010). *O Livro da Consciência. A construção do cérebro consciente*. Lisboa: Temas e Debates.

DAMÁSIO, A. (1993/2011). *O Erro de Descartes. Emoção, Razão e Cérebro Humano*. Lisboa: Temas e Debates.

DAMÁSIO, A. (2013). *O sentimento de si. Corpo, emoção e consciência*. Lisboa: Temas e Debates.

DAMÁSIO, A. (2017). *A Estranha Ordem das Coisas. A vida, os sentimentos e as culturas humanas*. Lisboa: Temas e Debates.

DAVIES, P. (2019). *The Demon in the Machine. How hidden webs of information are solving the mystery of life*. New York: Penguin.

DAMÁSIO, A. (2021). *Sentir e saber: Uma nova teoria da consciência*. Temas e Debates.

DAVIES, P.; GREGERSEN, N. H. (2014). *Information and the Nature of Reality: From Physics to Metaphysics*. Canto Classics edn. Cambridge University Press, Cambridge.

DE BROGLIE, L. (1956). *Une tentative d'interprétation causale et non linéaire de la mécanique ondulatoire: la théorie de la double solution*. Gauthier-Villars.

DEECKE, L., Grdznger, B. & Kornhuber, H. H. (1976) Voluntary finger movement in man: Cerebral potentials and theory. *Biological Cybernetics* 23:99-119. [biBL, CCW]

DEHAENE, S. (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. Viking.

DEHAENE, S. (2021). *How We Learn: Why Brains Learn Better Than Any Machine... for Now*. Penguin Random House.

DEHAENE, S.; CHANGEUX, J.-P. (2004). Neural Mechanisms for Access to Consciousness. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1145–1157). Boston Review.

DEHAENE, S.; CHANGEUX, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), 200-227.

DEHAENE, S.; NACCACHE, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. In Stanislas Dehaene (ed.), *The cognitive neuroscience of consciousness*. London: MIT Press.

DEHAENE, S.; SERGENT, C.; CHANGEUX, J.-P. (2006). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Sciences*, 103(38), 14263-14268.

DELGADO, J. (1969). *Physical Control of the Mind: Toward a Psychocivilised Society*. New York: HarperCollins.

DENNETT, D. (1991). *Consciousness Explained*. London: Back Bay Books.

DENNETT, D. (2003). *Freedom Evolves*. London: Penguin Books.

DENNETT, D. (2017). *Das Bactérias a Bach e Vice-Versa. A Evolução das Mentes*. Lisboa: Edições 70.

DENNETT, D. (2016). Illusionism as the Obvious Default Theory of Consciousness. *Journal of Consciousness Studies*, 23, N.11-12, 65-72.

DENNETT, D. (1993). Consciousness Explained. *Philosophy and Phenomenological Research*, Vol. 53, N. 4, 889-892.

DENNETT, D. C. (2015). Why and How Does Consciousness Seem the Way it Seems? In T. Metzinger & J. M. Windt (Eds.), *Open MIND: 10(T)*. Frankfurt am Main: MIND Group. doi: 10.15502/9783958570245.

DESCARTES, R. (1649/2023). *As Paixões da Alma*. Tradução e introdução de Maria João Vaz. Lisboa: Bertrand Editora.

DESCARTES, R. (1641/2023) *Meditações Metafísicas*. Tradução e introdução de Maria João Vaz. Lisboa: Bertrand Editora.

DESCARTES, R. (1637/2008). *Discurso do Método. As Paixões da Alma*. (tradução, prefácio e notas de Newton de Macedo). Lisboa: Sá da Costa Editora.

DESCARTES, R. (1649/2008). *As Paixões da Alma*. (tradução, prefácio e notas de Newton de Macedo). Lisboa: Sá da Costa Editora.

DESCARTES, R. (1649/1991). *Correspondance avec la princesse Élisabeth et autres lettres*. (Éd. par Jean-Marie Beyssade). Paris: GF Flammarion.

DESMURGET, M.; REILLY, K. T.; RICHARD, N.; SZATHMARI, A.; MOTTOLSE, C.; SIRIGU, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science*, 324, 811–813.

DIRAC, P. A. M. (1928). The quantum theory of the electron. *Proceedings of the Royal Society A*, 117(778), 610-624.

DONINI, A. (1988). *História do Cristianismo (das origens a Justiniano)*. Lisboa: Edições 70.

DOUBLE, R. (1991). *Determinism and The Experience of Freedom*. Pacific Philosophical Quarterly, 72.

DRETSKE, F. (1995). *Naturalizing the mind*. Cambridge, MA: Bradford Books / MIT Press.

EAGLE, A. (2021). Chance versus Randomness. *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2021/entries/chance-randomness/>.

EAGLEMAN, D. (2012). *Incógnito. As vidas secretas do cérebro humano*. Lisboa: Presença.

EARL, B. (2014). *The biological function of consciousness*. Front. in Psy. vol. 5, 697.

EARMAN, J. (1986). *A Primer on Determinism*. Dordrecht: Reidel Publishing Company.

EARMAN, J. (2004). Determinism: What We Have Learned and What We Still Don't Know. In *Freedom and Determinism*. Cambridge, MA: Bradford Book/MIT Press, pp. 21-46.

EARMAN, J. (2006). Aspects of Determinism in Modern Physics. In Handbook of the Philosophy of Science, Volume 2: *Philosophy of Physics*. [s.l.]: Elsevier BV.

EDELMAN, G. M. (1995). *Biologia da consciência. As raízes do pensamento*. Lisboa: Inst. Piaget.

EINSTEIN, A. (1905) Does the inertia of a body depend upon its energy-content? *Annalen der Physik*, 18, 639–641, 1905. <https://doi.org/10.1002/andp.19053231314>.

EINSTEIN, A.; PODOLSKY, B.; ROSEN, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47(10), 777-780. <https://doi.org/10.1103/PhysRev.47.777>.

ESPINOSA, B. de (1677/1992). *Ética*. Introdução de Joaquim de Carvalho, Lisboa: Relógio de Água.

FABER, D. S.; KORN, H. (1978). Neuronal mechanisms of escape behavior in the goldfish: I. Cellular excitation by sensory stimuli. *Journal of Neurophysiology*, 41(4), 849-863. <https://doi.org/10.1152/jn.1978.41.4.849>.

FESER, E. (2004). Why Searle Is a Property Dualist. *Paper presented at the American Philosophical Association Pacific Division meeting in Pasadena*.

FEYNMAN, R. P., LEIGHTON, R. B., & SANDS, M. (1964). The Feynman lectures on physics (Vol. 2). *Reading, MA: Addison-Wesley*.

FISCHER, J. M. (2007). Compatibilism. In J. M. Fischer, R. Kane, D. Pereboom, & M. Vargas (Eds.), *Four views on free will* (pp. 44–84). Blackwell Publishing.

FODOR, J. A. (1990). *A Theory of Content and Other Essays*. Cambridge, Massachusetts: The MIT Press.

FODOR, J. A. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

FOSTER, J. (1989). A Defense of Dualism. In J. Smythies and J. Beloff (Eds.), *The Case for Dualism*. Charlottesville, VA: University of Virginia Press.

FRANCO, V. A. G. (2019). Kant e a questão da objetividade do tempo e espaço. *Revista da Graduação da Faculdade Paulus de Comunicação -FAPCOM*, Ano 05, Vol. 10, 201.

FREDKIN, E. (2003). *An Introduction to Digital Philosophy*. *Int. J. Theor. Phys.* 42(2), 189–247.

FRANKFURT, H. G. (1969). Alternate possibilities and moral responsibility. *The Journal of Philosophy*, 66(23), 829–839. <https://doi.org/10.2307/2023833>.

FRIED, I.; MUKAMEL, R.; KREIMAN, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron*, 69, 548–562.

FRISTON, K.; KIEBEL, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1211–1221. <https://doi.org/10.1098/rstb.2008.0300>.

FRITH, C. D.; FRITH, U. (2007). Social Cognition in Humans. *Current Biology*, 17(16), R724–R732. <https://doi.org/10.1016/j.cub.2007.05.068>.

FRITH, C. D.; FRITH, U. (2020). *The Cognitive Neuroscience of Visual Perception*. Oxford University Press.

FUCHS, C. A.; PERES, A. (2000). Quantum theory needs no “interpretation.” *Physics Today*, 53(3), 70. <https://doi.org/10.1063/1.883004>.

GALLAGHER, S. (2006). Where’s the action? Epiphenomenalism and the problem of free will. In *Does Consciousness Cause Behavior?* Massachusetts: MIT.

GALLAGHER, S. (2005). *How the body shapes the mind*. Oxford: Oxford University Press.

GARZA, S. M., WERK, J. K., PROCHASKA, J. X., TUMLINSON, J., FOX, A. J., & PEEPLES, M. S. (2024). The Circumgalactic Medium as a Reservoir of Carbon and Oxygen: Evidence from the COS-GAL Sample. *The Astrophysical Journal Letters*, 912(1), L8.

GAZZANIGA, M. S. (2000). Cerebral specialization and interhemispheric communication: Does the corpus callosum enable the human condition? *Brain*, 123(7), 1293–1326.

GAZZANIGA, M. (2011). *Who’s in charge? Free Will and the science of the Brain*. New York: HarperCollins Publishers.

GLATTFELDER, J. B. (2019). *A Universe Built of Information*. In *Information–Consciousness–Reality, The Frontiers Collection*. https://doi.org/10.1007/978-3-030-03633-1_13.

GILSON, E. (2002). *The Christian Philosophy of Descartes*. Ignatius Press.

GINET, C. (1962). Can the Will be Caused? *The Philosophical Review*, 71(1), 49–55.

GINET, C. (1997). Freedom, Responsibility, and Agency. *The Journal of Ethics*, 1, 85–98.

GINET, C. (2008). In Defense of a Non-Causal Account of Reasons Explanations. *J Ethics*, 12, 229–237. DOI: 10.1007/s10892-008-9033-z.

GINET, C. (2014). Can an Indeterministic Cause Leave a Choice up to the Agent? In *Libertarian Free Will: Contemporary Debates*, David Palmer (Ed.), 15–26. New York: Oxford University Press.

GIUSTINA, M.; VERSTEEGH, M. A. M.; WENGEROWSKY, S.; HANDSTEINER, J.; HOCHRAINER, A.; PHELAN, K.; ... ZEILINGER, A. (2015). Significant-loop-hole-free test of Bell's theorem with entangled photons. *Physical Review Letters*, 115(25), 250401. <https://doi.org/10.1103/PhysRevLett.115.250401>.

GODFREY-SMITH, P. (2016). *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. Farrar, Straus and Giroux.

GOFF, P. (2019). *Galileo's Error: Foundations for a New Science of Consciousness*. Pantheon Books.

GREEN, B. (2021). *Até Ao Fim Dos Tempos. O Homem, o Universo e a nossa busca pelo sentido da vida*. Lisboa: Saída de Emergência.

GREENE, B. (2003). *The fabric of the cosmos: Space, time, and the texture of reality*. New York: Alfred A. Knopf.

GREGORY, F. (1977). *Scientific Materialism in Nineteenth Century Germany*. Springer.

GRIBBIN, J. (2007). *Universo. Uma Biografia*. Alfragide: Estrela Polar.

GRIFFITHS, D. J. (2005). *Introduction to quantum electrodynamics* (3rd ed.). Upper Saddle River, NJ: Pearson Education.

GRIFFITHS, D. J. (2013). *Introduction to electrodynamics* (4th ed.). Boston, MA: Pearson.

GULICK, R. V. (2012). *Consciência*. Investigação Filosófica, Vol. E2, Artigo Digital 2.

HAGGARD, P. (2002). Voluntary action and conscious awareness. *Nat. Neurosci.*, 5, 382–385.

HAGGARD, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences*, 9(6), 290–295. <https://doi.org/10.1016/j.tics.2005.04.012>

HAGGARD, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience*, 9(12), 934–946. <https://doi.org/10.1038/nrn2497>

HÁJEK, A. (2023). Interpretations of Probability. *The Stanford Encyclopedia of Philosophy* (Winter 2023 Edition), Edward N. Zalta & Uri Nodelman (Eds.), URL = <https://plato.stanford.edu/archives/win2023/entries/probability-interpret/>.

HAKEN, H. (2006). *Information and Self-Organization: A Macroscopic Approach to Complex Systems*. 3rd edn. Springer, Heidelberg.

HASKER, W. (1999). *The Emergent Self*. Ithaca, NY: Cornell University Press.

HAWKING, S.; MLODINOV, L. (2011). *O Grande Desígnio*. Lisboa: Gradiva.

HAYNES, J.D. (2011) Tracking the Unconscious Generation of Free Decisions Using Ultra-High Field fMRI. *PLoS ONE* 6(6): e21612. doi:10.1371/journal.pone.0021612.

HEISENBERG, W. (1958). *Physics and Philosophy: The Revolution in Modern Science*. New York: Harper & Brothers.

HELMHOLTZ, H. von. (1847/1977). *Epistemological writings* (M. F. Lowry, Ed. & Trans.). Dordrecht: Springer.

HENSEN, B.; BERNIEN, H.; DRÉAU, A. E.; REISERER, A.; KALB, N.; BLOK, M. S.; ... HANSON, R. (2015). Loophole-free Bell inequality violation using electron spins separated by 1.3 km. *Nature*, 526(7575), 682–686. <https://doi.org/10.1038/nature15759>.

HERMAN, L. M. (2010). What Laboratory Research Has Told Us About Dolphin Cognition. *International Journal of Comparative Psychology*, 23(3), 310–330.

HOBART, R. E. (1934). Free Will as Involving Determination and Inconceivable Without It. *Mind*, 47(188), 1-27.

HOBBS, T. (1655/1999). *De corpore* (H. Gaskin, Ed.). Oxford University Press.

HOBBS, T. (1658/1978). *Man and citizen (De Homine and De Cive)* (B. Gert, Ed. & Trans.). Hackett Publishing Company.

HOCHNER, B. (2012). *An Embodied View of Octopus Neurobiology*. *Current Biology*, 22(20), R887–R892. <https://doi.org/10.1016/j.cub.2012.09.001>.

HOEFER, C. (2023). Causal Determinism. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta & Uri Nodelman (Eds.), URL = <https://plato.stanford.edu/archives/spr2023/entries/determinism-causal/>.

HOHWY, J. (2013). *The Predictive Mind*. Oxford University Press.

- HONDERICH, T. (1988). *Mind and the Brain. A Theory of Determinism*. Oxford: Clarendon Press.
- HOOFT, G. (1993). Dimensional reduction in quantum gravity. arXiv:gr-qc/9310026. <https://arxiv.org/abs/gr-qc/9310026>
- HOOFT, G. (2000). The holographic principle. arXiv preprint arXiv:hep-th/0003004. <https://doi.org/10.48550/arXiv.hep-th/0003004>
- HOOFT, G. (2016). *The Cellular Automaton Interpretation of Quantum Mechanics*. New York: Springer.
- HOOFT, G. (2017). Free Will in the Theory of Everything. *Workshop Determinism and Free Will*, Milano.
- HOSSENFELDER, S.; PALMER, T. (2020). Rethink Superdeterminism. *Front. Phys.*, 8, 139. doi: 10.3389/fphy.2020.00139
- HOSSENFELDER, S. (2022). *Existential Physics. A Scientist's Guide to Life's Biggest Questions*. London: Atlantic Books.
- HOTTOI, G. (2003). Platão: a instituição da filosofia idealista e dialética. In *História da Filosofia. Da Renascença à Pós-Modernidade*. Lisboa: Piaget.
- HOTTOI, G. (2003a). Aristóteles: a articulação sistemática da filosofia. In *História da Filosofia. Da Renascença à Pós-Modernidade*. Lisboa: Piaget.
- HUME, D. (1739/2016). *Tratado da Natureza Humana* (4.^a ed.). Lisboa: Fundação Calouste Gulbenkian.
- HUMPHREY, N. (2012). *Poeira da Alma. A magia da consciência*. Lisboa: Gradiva.
- HUMPHREY, N. (2023). *Senciência. A invenção da consciência*. Lisboa: Bertrand.
- HUTCHINSON, M. N. (2002). *Biology and Ecology of Australian Lizards*. Surrey Beatty & Sons.
- HUXLEY, T. H. (1874). *On the Hypothesis that Animals are Automata, and its History*. The Fortnightly Review, n.s. 16, 555–580.
- IKEDA A, Miyamoto S, Shibasaki H. (2002) Cortical motor mapping in epilepsy patients: information from subdural electrodes in presurgical evaluation. *Epilepsia* 43 Suppl 9:56–60.
- ISMAEL, J. (2021) Quantum Mechanics. *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2021/entries/qm/>>.

JACKSON, F. (1984). Epiphenomenal qualia. *The Philosophical Quarterly*, 34(135), 127–136. <https://doi.org/10.2307/2219770>

JACKSON, F. (1986). What Mary didn't know. *The Journal of Philosophy*, 83(5), 291–295. <https://doi.org/10.2307/2026143>

JAMES, W. (1884). What is an emotion?, *Mind* 9, 188-205.

JAMES, W. (1890). *The principles of psychology* (Vols. 1 & 2). Henry Holt and Company.

JOHNSON-LAIRD P., (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge, MA: Harvard University Press.

HAGGARD, P.; CARTLEDGE, P.; DAFYDD, M. M.; DAVID, O. W. (2004). Anomalous control: When free-will is not conscious. *Consciousness and Cognition*, 13, 646–654.

HAGGARD, P., & EIMER, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, 126(1), 128–133. <https://doi.org/10.1007/s002210050722>

KAHNEMAN, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

KAKU, M. (2021). *A Equação Divina. A busca inacabada por uma teoria do tudo e o futuro da física*. Lisboa: Bertrand Editora.

KANE, R. (2005). *A Contemporary Introduction to Free Will*. New York: Oxford University Press.

KANE, R. (2007). *Libertarianism*. In *Four Views on Free Will* (Eds. John Martin Fischer, Robert Kane, Derk Pereboom, Manuel Vargas). Oxford: Blackwell Publishing.

KANE, R. (1996). *The significance of free will*. New York: Oxford University Press.

KANT, I. (1781//2008). *Crítica da Razão Pura* (6.ª Ed.). Lisboa: Fundação Gulbenkian.

KASSAL, I.; JORDAN, S. P.; LOVE, P. J.; MOHSENI, M., & ASPURU-GUZI, A. (2008) Quantum algorithms for the simulation of chemical dynamics. *PNAS*, 105(48), 18681–18686.

KENNY, A. (2010). *Aquinas on Mind*. Routledge.

KERNBERG, Otto F. (1996). *Borderline Conditions and Pathological Narcissism*. Northvale, NJ: Jason Aronson.

KOCH, C. et al. (2016). Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience*, 17(5), 307-321. Disponível em: https://www.researchgate.net/publication/301567963_Neural_correlates_of_consciousness_Progress_and_problems, acessado em 10-05-2021.

KRAUSS, L. (2017). *Um Universo Vindo do Nada. Porque há algo em vez de nada?* Lisboa: Gradiva.

KULSTAD, M. (1997). *Leibniz on Perception, Consciousness, and Self-Consciousness*. Routledge.

JACKSON, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly*, 32(127), 127-136.

JACKSON, F. (1986). What Mary Didn't Know. *The Journal of Philosophy*, 83(5), 291-295.

JEANNEROD, M. (2006). Consciousness of Action as an Embodied Consciousness. In *Does Consciousness Cause Behavior?* Massachusetts: MIT.

JUN, Y.; GAVRILOV, M.; BECHHOEFER, J. (2014). High-Precision Test of Landauer's Principle in a Feedback Trap. *Physical Review Letters*, 113, 190601.

LAMB, W. E., Jr., & RETHERFORD, R. C. (1947). Fine structure of the hydrogen atom by a microwave method. *Physical Review*, 72(3), 241-243. <https://doi.org/10.1103/PhysRev.72.241>

LAMBA, W. E.; RETHERFORD, R. C. (1947). Fine structure of the hydrogen atom by a microwave method. *Physical Review*, 72(3), 241.

LAMBERT, N., CHEN, Y. N., CHENG, Y. C., LI, C. M., CHEN, G. Y., & NORI, F. (2013) Quantum biology. *Nature Physics*, 9(1), 10-18.

LANDAUER, R. (1996). The physical nature of information. *Phys. Lett. A*, 217(4-5), 188-193.

LAU, H. C.; ROGERS, R. D.; HAGGARD, P.; PASSINGHAM, R. E. (2004). Attention to intention. *Science*, 303, 1208-1210.

LAWRENCE, D. H. (1989). *The Works of D. H. Lawrence: Introductions and Reviews*. Edited by N. H. Reeve & John Worthen. Cambridge: Cambridge University Press. p. 167

LEDOUX, J. (1999). *O cérebro emocional: As misteriosas estruturas da vida emocional*. Cascais: Editora Pergaminho.

LEDOUX, J. (2020). *O Cérebro Consciente. Uma Longa História da Vida*. Lisboa: Temas e Debates.

LEDOUX, J.; BROWN, R. (2017). A higher-order theory of emotional consciousness. *PNAS*. Disponível em: www.pnas.org/cgi/doi/10.1073/pnas.1619316114, pp. 2016-2015.

LEIBNIZ, G. W. (1641/1996). *Novos Ensaios sobre o Entendimento Humano* (Tradução de edição disponível). Editora Martins Fontes.

LETTVIN, J. Y.; MATURANA, H. R.; MCCULLOCH, W. S.; PITTS, W. H. (1959). What the frog's eye tells the frog's brain. *Proceedings of the IRE*, 47(11), 1940–1951. <https://doi.org/10.1109/JRPROC.1959.287207>

LEVINE, J. (1983). Materialism and Qualia: the explanatory gap. *Pacific Philosophical Quarterly*, 64, 354-361.

LÉVY, R. (2023). *The prefrontal cortex: from monkey to man*. *Brain*, 147(3), 794–807.

LEVY, N. (2014). *Consciousness and Moral Responsibility*. New York: Oxford University Press.

LIBET, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *The Behavioral and Brain Sciences*, 8(4), 529-566.

LIBET, B. (1999). Do we have free will? *Journal of Consciousness Studies*, 6(8-9), 47-57.

LIU, Z.-H., MENG, H.-X., XU, Z.-P., ZHOU, J., CHEN, J.-L., XU, J.-S., LI, C.-F., & GUO, G.-C. (2025). Exploring the boundary of quantum correlations with a time-domain optical processor. *Science Advances*, 11, eabd8080. <https://doi.org/10.1126/sciadv.abd8080>

LLINÁS, R.; RIBARY, U. (1993). Coherent 40-Hz oscillation characterizes dream state in humans. *Proc. Natl. Acad. Sci. U. S. A.*, 90(5), 2078-2081.

LIVIO, M. (2010). *Is God a Mathematician?* New York: Simon & Schuster.

LOCKE, J. (1689/2014). *An Essay Concerning Human Understanding Second Treatise of Government*. Ware: Wordsworth Editions LTD.

LOHMANN, K. J.; LOHMANN, C. M.; ENDRES, C. S. (1999). The sensory ecology of ocean navigation. *Journal of Experimental Biology*, 202(23), 3343–3350.

LLOYD, S. (2006). *Wired Magazine: Life, the Universe, and Everything*. <https://www.wired.com/2006/03/life-theuniverse-and-everything/>, 3 January 2006.

LLOYD, S. (2007). *Programing the Universe*. Londres: Vintage Books.

LLOYD, S. (2012). The universe as quantum computer. In *A Computable Universe: Understanding and exploring Nature as computation* (H. Zenil ed.), London: World Scientific.

MACLEAN, P. D. (1990). *The Triune Brain in Evolution: Role in Paleocerebral Functions*. Springer.

MALDACENA, J. (1999). The large N limit of superconformal field theories and supergravity. *International Journal of Theoretical Physics*, 38(4), 1113–1133. <https://doi.org/10.1023/A:1026654312961>

MANZA, P.; TOMASI, D.; VOLKOW, N. D. (2023). Route of drug administration influences the salience network response to methylphenidate in humans. *Nature Communications*, 14(1), 6512. <https://doi.org/10.1038/s41467-023-36789-8>

MARCZ, LISA (2018). Feral children: Questioning the human-animal boundary from an anthropological perspective. *Arbeitspapiere des Instituts für Ethnologie und Afrikastudien der Johannes Gutenberg-Universität Mainz* (Working Papers of the Department of Anthropology and African Studies of the Johannes Gutenberg University Mainz), 179.

MASI, M. (2023). Indeterminism, Free Will, and Self-Causation. *Journal of Consciousness Studies*, 30(5–6), 32–56. DOI: 10.53765/20512201.30.5.032

MATTHEWS, G. B. (2005). *Augustine*. Wiley-Blackwell.

MAUDLIN, T. (2011) *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics* (3rd ed.). Wiley-Blackwell.

MAXWELL, J. C. (1873). *A treatise on electricity and magnetism* (Vol. 2). Oxford: Clarendon Press.

MCCANN, HUGH J. (1998). *The Works of Agency: On Human Action, Will, and Freedom*. Ithaca: Cornell University Press.

MCGINN, COLIN (2011). *O Carácter da Mente. Uma introdução à filosofia da mente*. Lisboa: Gradiva.

MCKENNA, M., & COATES, D. J. (2024). Compatibilism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2024 Edition). Stanford University. Recuperado de <https://plato.stanford.edu/archives/spr2024/entries/compatibilism/>

MEIJER, D. K. F. (2012). The Information Universe. On the Missing Link in Concepts on the Architecture of Reality. *Syntropy Journal*, 1, 1-64.

MEIJER, D. K. F. (2015). The Universe as a Cyclic Organized Information System: John Wheeler's World Revisited. *NeuroQuantology*, 13(1), 57-78.

MELE, A. R. (2000). Goal-Directed Action: Teleological Explanations, Causal Theories, and Deviance. *Philosophical Perspectives*, 14, Action and Freedom.

MELE, A. R. (2003). *Motivation and Agency*. Oxford University Press.

MELE, A. R. (2004). Free will: Theories, analysis, and data. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior?* (pp. 187–205). MIT Press.

MELE, A. R. (2009). Libet on free will: Readiness potentials, decisions, and awareness. In W. Sinnott-Armstrong & L. Nadel (Eds.), *Conscious will and responsibility: A tribute to Benjamin Libet* (pp. 43-52). Oxford University Press.

METZINGER, T. (2009). *The Ego Tunnel. The Science of the Mind and the Myth of the Self*. New York: Basic Books.

MIGUENS, S. (2002). *Uma teoria fisicalista do conteúdo e da consciência: Daniel Dennett e os debates da filosofia da mente*. Campo das Letras.

MILES, J. B. (2015). *The Free Will Delusion. How We Settled For The Illusion of Morality*. Leicestershire: Matador.

MITCHINSON, B.; MARTIN, C. J.; GRANT, R. A.; PRESCOTT, T. J. (2007). Feedback Control in Active Sensing: Rat Whisker Tracking. *Journal of Neuroscience*, 27(32), 7933–7938. <https://doi.org/10.1523/JNEUROSCI.0791-07.2007>

MONTOYA, D. (2021). Post-Materialism and the Re-framing of Classic Neuroscience: Some Open Questions. *Academia Letters*, Article 3007. <https://doi.org/10.20935/AL3007>.

MORCH, H. H. (2023). *Non-physicalist Theories of Consciousness*. New York: Cambridge Press.

MORSELLA, E. (2022). *Consciousness and the Brain: Lab Manual 1.0*. San Francisco: KDP.

MURTY, N.; BASHIVAN, P.; ABATE, A.; DICARLO, A.; KANWISHER, N. (2021). Computational models of category-selective brain regions enable high-throughput tests of selectivity. *Nature Communications*, 12:5540. <https://doi.org/10.1038/s41467-021-25409-6>

MUTHUKUMARASWAMY, S. D., CARHART-HARRIS, R. L., MORAN, R. J., BROOKS, J. C. W., WILLIAMS, T. M., ERNSTEN, C., & NUTT, D. J. (2013). Broadband cortical desynchronization underlies the human psychedelic state. *The Journal of Neuroscience*, 33(38), 15171–15183. <https://doi.org/10.1523/JNEUROSCI.2063-13.2013>.

NAGEL, T. (1974). What is it like to be a bat? *The Philosophical Review*, 83(4), 435-450.

NAGEL, T. (1979). Panpsychism. In T. Nagel, *Mortal Questions*. Cambridge: Cambridge University Press.

NANI, A.; MANUELLO, J.; MANCUSO, L.; LILOIA, D.; COSTA, T.; CAUDA, F. (2019). The Neural Correlates of Consciousness and Attention: Two Sister Processes of the Brain. *Front. Neurosci.*, 13:1169. <https://doi.org/10.3389/fnins.2019.01169>.

NEWTON, I. (1687/1934). *The mathematical principles of natural philosophy* (A. Motte, Trans.; F. Cajori, Rev. ed.). Berkeley: University of California Press. (Original work published 1687)

NORMAN, D. A. (1981). *The psychology of everyday things*. Basic Books.

NOTTEBOHM, F. (2005). The Neural Basis of Birdsong. *PLoS Biology*, 3(5), e164. <https://doi.org/10.1371/journal.pbio.0030164>

NUSSBAUM, M. C.; RORTY, A. O. (Eds.). (1995). *Essays on Aristotle's De Anima*. Clarendon Press.

OAKLEY, D. A.; HALLIGAN, P.W. (2017). Chasing the Rainbow: The Non-conscious Nature of Being. *Frontiers in Psychology*, 8:1924.

O'CONNOR, T. (1995). *Agent Causation*. In *Agents, Causes, and Events: Essays on Indeterminism and Free Will*. Timothy O'Connor (Ed.), New York: Oxford University Press.

O'CONNOR, T. (2000). Causality, Mind and Free Will. *Philosophical Perspectives*, 14, Action and Freedom.

O'CONNOR, T. (2000a). *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press.

OLIVEIRA, GUILHERME SANCHES. (2014). Estados Mentais e Atitudes Proposicionais: *Abordagens Filosóficas da Psicologia do Senso Comum*. São Paulo: Programa de Pós-Graduação. Disponível em https://www.teses.usp.br/teses/disponiveis/8/8133/tde-07112014-194213/publico/2014_GuilhermeSanchesDeOliveira_VO-rig.pdf, acessado em 10-05-2021.

PALMER, S. E. (2020). *Vision Science: Photons to Phenomenology*. MIT Press.

PANKSEPP, J. (2005). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness And Cognition*, 14(1), 30-80.

PATEL, A. (2001) Quantum information and the genetic code. *Journal of Genetics*, 80(2), 39–46..

PEREBOOM, D. (2007). *Hard Incompatibilism*. In *Four Views on Free Will* (Eds. John Martin Fischer, Robert Kane, Derk Pereboom, Manuel Vargas). Oxford: Blackwell Publishing.

PEREBOOM, D. (2022). *Free Will*. New York: Cambridge University Press.

PESKIN, M. E.; SCHROEDER, D. V. (1995). *An Introduction to Quantum Field Theory*. Addison-Wesley.

PINKER, S. (1999). *Como a mente funciona*. São Paulo: Companhia das Letras.

POCKETT, S.; BANKS, P. B.; GALLAGHER, S. (2006). *Does Consciousness Cause Behavior?* Massachusetts: MIT.

“PROBABILIDADE” (2024). *Dicionário Priberam da Língua Portuguesa* [on-line], 2008-2024. Disponível em: <https://dicionario.priberam.org/probabilidade>. Acessado a 10-10-2024.

RAMACHANDRAN, V. S.; HIRSTEIN, W. (1997). Three Laws of Qualia: What Neurology Tells Us about the Biological Functions of Consciousness, Qualia and the Self. *Journal of Consciousness Studies*, 4(5-6), 429–458.

RAMACHANDRAN, V. S.; HUBBARD, E. M. (2001). Hearing colors, tasting shapes. *Scientific American*, 288(5), 52–59.

RAMSEY, WILLIAM (2019). Eliminative Materialism. *Stanford Encyclopedia of Philosophy*. Disponível em: <https://plato.stanford.edu/entries/materialism-eliminative/>. Acessado a 17-05-2021.

RIBEIRO, M. (2013). *Universo Programado: Uma Alternativa ao Darwinismo e à Religião*. Lisboa: Gradiva.

ROBINSON, H. (2011). Malebranche and Occasionalism. In Nadler, S. (Ed.), *The Cambridge Companion to Malebranche* (pp. 89–114). Cambridge University Press.

ROMEIJN, JAN-WILLEM (2022). Philosophy of Statistics. *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.). URL: <https://plato.stanford.edu/archives/fall2022/entries/statistics/>.

ROSA, LUIZ AUGUSTO (2018). *A Consciência Segundo Damásio: Um Enfoque Neurofilosófico*. Marília: Dissertação de Pós-Graduação em Filosofia.

ROSA, R. O. (2011). Mente como face representacional do cérebro. *Filosofia e História da Biologia*, 6(1), 143-157.

ROSENTHAL, D. M. (1997). A Theory of Consciousness. In N. Block, O. Flanagan, & G. Güzelde (Eds.), *The Nature of Consciousness: Philosophical Debates* (pp. 729–753). MIT Press.

ROSENTHAL, D. M. (2005). *Consciousness and Mind*. Oxford: Oxford University Press.

ROSENTHAL, D. M. (2008). Consciousness and Its Function. *Neuropsychologia*, 46(3), 829–840.

ROVELLI, C. (1996). Relational Quantum Mechanics. *International Journal of Theoretical Physics*, 35, 1637.

ROVELLI, C. (2018). *Reality is Not What It Seems*. New York: Riverhead Books.

ROVELLI, C. (2022). *O Abismo Vertiginoso: Um Mergulho Nas Ideias da Física Quântica*. Lisboa: Objectiva.

SADATO, N., PETRIDES, M., KENNEDY, H., & EVANS, A. C. (1996). Role of the supplementary motor area and the right premotor cortex in the coordination of bimanual finger movements. *Journal of Neuroscience*, 16(6), 2068–2078. <https://doi.org/10.1523/JNEUROSCI.16-06-02068.1996>.

SAGAN, C. (2009) *Cosmos*. Lisboa: Gradiva.

SAHAKIAN, B. J.; LANGLEY, C.; STAMATAKIS, E. A.; SPINDLER, L. (2020). Consciousness: how the brain chemical 'dopamine' plays a key role. *The Conversation*. Disponível em: <https://theconversation.com/consciousness-how-the-brain-chemical-dopamine-plays-a-key-role-144478>.

SALAMONE, J. D., & CORREA, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3), 470–485. <https://doi.org/10.1016/j.neuron.2012.10.021>.

SAPOLSKY, R. M. (2024). *Determinado: Uma ciência da vida sem livre-arbítrio*. Temas e Debates

SAUR, D., KREHER, B. W., SCHNELL, S., KÜMMERER, D., KELLMEYER, P., VRY, M.-S., UMAROVA, R., MUSSO, M., GLAUCHE, V., ABEL, S., HUBER, W., RIJNTJES, M., HENNIG, J., & WEILLER, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(46), 18035–18040. <https://doi.org/10.1073/pnas.0805234105>.

SAUTOY, M. Du (2018). *O que não podemos saber: Viagem aos limites do conhecimento*. Lisboa: Bizâncio.

SCHMID, Y., & LIECHTI, M. E. (2017). Long-lasting subjective effects of LSD in normal subjects. *Psychopharmacology*, 234(13), 2135–2143. <https://doi.org/10.1007/s00213-017-4610-0>.

SCHOPENHAUER, A. (1839/2005). *Essay on the Freedom of the Will*. *Dover Philosophical Classics*. New York: Dover Publications.

SCHWINGER, J. (1948). On quantum-electrodynamics and the magnetic moment of the electron. *Physical Review*, 73(4), 416.

SEARLE, John (1998) *A redescoberta da mente*. Lisboa: Instituto Piaget.

SEARLE, John (2019) *Da Realidade Física à Realidade Humana*. Lisboa: Gradiva.

SEELEY, W. W., CARLIN, D. A., ALLMAN, J. M., MACEDO, M. N., BUSH, C., & MILLER, B. L. (2006). Early frontotemporal dementia targets neurons unique to apes and humans. *Annals of Neurology*, 60(6), 660–667.

SERRA, M., & MONTELEONE, P. (2006). Neurochemical mechanisms underlying the psychotropic effects of general anaesthetics. *CNS & Neurological Disorders - Drug Targets*, 5(6), 619–632. <https://doi.org/10.2174/187152706779780340>.

SERRADO, R. (2021) *O problema mente-corpo no Portugal Contemporâneo: para uma epistemologia do desporto*. Lisboa: Tese de doutoramento na UAL.

SERRADO, R. (2024) Free will and the Misconceptions of Indeterminism. https://www.researchgate.net/publication/381732498_Free_will_and_the_misconceptions_of_indeterminism.

SERRADO, R. (2024b) “Miguel Bombarda: um visionário da consciência”. *Revista Filosofia da Faculdade de Letras da Universidade do Porto*.

SERRADO, R. (2025) Determinism as the cause of free will. *Unisinos Journal of Philosophy*, 26(1).

SERRADO, R. (2024c) Chalmer’s Error: Why is property dualism wrong? In [https://www.researchgate.net/publication/386534487 David Chalmers' Error Why is property dualism wrong](https://www.researchgate.net/publication/386534487_David_Chalmers'_Error_Why_is_property_dualism_wrong)

SETH, Anil; SUZUKI, K.; CRITCHLEY, H. (2012). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology*, 2, article 395, doi: 10.3389/fpsyg.2011.00395.

SETH, Anil (2021) *Being You. A New Science of Consciousness*. London: Faber & Faber.

SHANNON, C.E. (1948) A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423 (1948).

SHEPHERD, J. (2015). Consciousness, Free Will, and Moral Responsibility: Taking the Folk Seriously. *Philosophical Psychology*, 28, 929–946.

SHERMAN, S. M., & GUILLERY, R. W. (2013). *Functional Connections of Cortical Areas: A New View from the Thalamus*. MIT Press.

SHERMAN, S. M., & GUILLERY, R. W. (2013). *Functional Connections of Cortical Areas: A New View from the Thalamus*. MIT Press.

SIEGEL, J. S., SUBRAMANIAN, S., PERRY, D., KAY, B. P., GORDON, E. M., LAUMANN, T. O., ... & DOSENBACH, N. U. F. (2024). Psilocybin desynchronizes the human brain. *Nature*. <https://doi.org/10.1038/s41586-024-07624-5>.

SIEGEL, J. S., SUBRAMANIAN, S., PERRY, D., KAY, B. P., GORDON, E. M., LAUMANN, T. O., ... & DOSENBACH, N. U. F. (2024). Psilocybin desynchronizes the human brain. *Nature*. <https://doi.org/10.1038/s41586-024-07624-5>.

SILVA, G. (2020). Bem-estar subjetivo, autoeficácia e consumo de álcool em estudantes de psicologia. Universidade Federal do Rio Grande do Sul. Disponível em: <https://www.academia.edu/53568863/>.

SILVA, K. M., & FONSECA, R. P. (2011). A integração sensorial e suas interfaces com as habilidades de coordenação motora. In M. C. S. Rodrigues & M. M. C. Oliveira (Orgs.), *Neuropsicologia hoje: pesquisas e práticas* (pp. 85-100): Oficina Universitária.

SILVEIRA, F. L. (2002). "A Teoria do Conhecimento de Kant: O Idealismo Trascendental", *Cad. Cat. Ens. Fís.*, 28 v. 19, número especial: p. 28-51.

SIMMONS, J. A., & KICK, S. A. (1984). Physiological Mechanisms of Echolocation in Bats and Dolphins. *Nature*, 283(5750), 366–368. <https://doi.org/10.1038/scientificamerican1080-112>

SIMON, H. A. (1997). *Models of Mind: Toward a Philosophy of Cognitive Science*. MIT Press.

SKYTT, T. (2014). The "hard problem" of consciousness is a dead end. Umeå University, Department of Historical, Philosophical and Religious Studies. Disponível em: <https://www.diva-portal.org/smash/get/diva2:745960/FULLTEXT01.pdf>.

SLATTERY, T. (2014) *Breaking The Free Will Illusion. For The Betterment of Humankind*. [s.l.]: Working Matter.

SMART, J.J.C. (1959). Sensations and Brain Processes. *Philosophical Review*, 68(2), 141–156.

SMOLIN, Lee (2019) *Einstein's unfinished revolution: the search for what lies beyond the quantum*. New York: Penguin Press.

SOON, C.S.; BRASS, M.; HEINZE, H.; HAYNES, J. (2008). Unconscious determinants of free decisions in the human brain. *Nature Publishing Group*, 5, 543-545. <https://doi.org/10.1038/nn.2112>

SRINIVASAN, M. V. (2011). Visual Control of Navigation in Flying Insects and Its Relevance for Robotics. *Current Biology* 21(9), R355–R366. <https://doi.org/10.1016/j.cub.2011.03.028>.

STERN, Luiz R. C. (2013) *O materialismo eliminacionista de Paul e Patricia Churchland*. Porto Alegre: Editora Fi.

STIEG, C. (2004) *Mental Representations: the New Sense-Data?* (Unpublished).

STOLJAR, D. (2010) *Physicalism*. New York: Routledge.

SUSSKIND, L. (1995). The world as a hologram. *Journal of Mathematical Physics*, 36(11), 6377–6396. <https://arxiv.org/abs/hep-th/9409089>.

SUSSKIND, L. (1993). Some Speculations About Black Hole Entropy in String Theory. *arXiv preprint hep-th/9309145*. <https://arxiv.org/abs/hep-th/9309145>.

SWAAB, D. F. (2014). *We are our brains. A neurobiography of the brain, from the womb to Alzheimer*. New York: Siegel & Grau.

- SWINBURNE, R. (1986) *The Evolution of the Soul*. Oxford: Oxford University Press.
- TAIZ, L., & ZEIGER, E. (2010). *Plant physiology*. Sinauer Associates.
- TALLIS, R. (2017). *Mentes ocultas: A consciência, o cérebro e o mito da neurociência*. Lisboa: Temas e Debates.
- TANZI, L., ROCCUZZO, S. M., LUCIONI, E., FAMÀ, F., FIORETTI, A., GABBANINI, C., MODUGNO, G., RECATI, A., & STRINGARI, S. (2019). Supersolid symmetry breaking from compressional oscillations in a dipolar quantum gas. arXiv preprint arXiv:1906.02791.
- TEGMARK, Max (2014) *Our mathematical universe: my quest for the ultimate nature of reality*. New York: Alfred A. Knopf.
- TEGMARK, Max (2017) *Life 3.0: being human in the age of artificial intelligence*. New York: Alfred A. Knopf.
- TIMMERMAN, C., ROSEMAN, L., WILLIAMS, L., ERRITZOE, D., MARTIAL, C., CASOL, H., ... & CARHART-HARRIS, R. L. (2019). Neural Correlates of the DMT Experience Assessed with Multivariate EEG. *Scientific Reports*, 9(1), 16324. <https://doi.org/10.1038/s41598-019-51974-4>.
- THELEN, E., & SMITH, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. MIT Press. <https://doi.org/10.7551/mitpress/3562.001.0001>.
- TONONI, Giulio & KOCH, Christof (2015) "Consciousness: Here, There and Everywhere?", *Phil. Trans. R. Soc. B* 370: 20140167. Disponível em <http://dx.doi.org/10.1098/rstb.2014.0167>, acessado em 05-05-2021.
- TONONI, giulio et. al. (2016) "Integrated Information Theory: From consciousness to its physical substrate" *nature Reviews Neuroscience* 17(7). Disponível em https://www.researchgate.net/publication/303551101_Integrated_information_theory_From_consciousness_to_its_physical_substrate, acessado em 10-05-2021.
- TONONI, Giulio & SPORNS, Olaf (2004) "Measuring information integration", *BMC Neuroscience*, 4: 31.
- TONONI, G. (2008). Consciousness as Integrated Information: a Provisional Manifesto. *Biological Bulletin*, 215(3), 216–242. <https://doi.org/10.2307/25470707>.
- TREADWAY, M. T., BUCKHOLTZ, J. W., SCHWARTZMAN, A. N., LAMBERT, W. E., & ZALD, D. H. (2012). Dopaminergic mechanisms of individual differences in human effort-based decision-making. *The Journal of Neuroscience*, 32(18), 6170–6176.

VARELA, F. J.; THOMPSON, E., & ROSCH, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.

VERHEYEN, P. (2021). From Information and Quantum Physics to Consciousness and Reality. *Sci* 2021, 3, 35. <https://doi.org/10.3390/sci3040035>.

VERÍSSIMO, D. (2012). A experiência subjetiva do consumo de drogas: um estudo exploratório. Dissertação de Mestrado, Universidade de Coimbra, Faculdade de Psicologia e de Ciências da Educação. Disponível em <https://estudogeral.uc.pt/bitstream/10316/27668/1/TESE%20Dina%20Verissimo.pdf>.

VERLINDE, E. (2011) On the Origin of Gravity and the Laws of Newton. *J. High Energy Phys.* 2011, 1–27, arXiv:1001.0785v1 [hep-th].

VOSBURG, S. K., NEWMAN, C., & HESS, C. (2021). Acute effects of sedative and stimulant drugs on consciousness: A systematic review. *Neuroscience & Biobehavioral Reviews*, 126, 457–469. <https://doi.org/10.1016/j.neubiorev.2021.03.016>

WALKER, S. I., & DAVIES, P. C. W. The algorithmic origins of life. *Journal of the Royal Society Interface*, 10(79), 20120869, 2013.

WEGNER, D. (2002). *The Illusion of Conscious Will*. MIT Press.

WEIH, G., JENNEWEIN, T., SIMON, C., WEINFURTER, H., & ZEILINGER, A. (1998). Violation of Bell's Inequality Under Strict Einstein Locality Conditions. *Physical Review Letters*, 81(23), 5039–5043. <https://doi.org/10.1103/PhysRevLett.81.5039>

WEINBERG, S. (1977). *The First Three Minutes: A Modern View of the Origin of the Universe*. Basic Books.

WEIHS, G., JANSSENS, B., WEINFURTER, H., ZEILINGER, A., & KWIAT, P. G. (1998). Violation of Bell's inequality under strict Einstein locality conditions. *Physical Review Letters*, 81(23), 5039–5043. <https://doi.org/10.1103/PhysRevLett.81.5039>.

WHEELER, J. A. (1990). *Information, physics, quantum: the search for links*. In: Zurek, W. H. (ed.) *Complexity, Entropy, and the Physics of Information*. Westview Press, Boca Raton.

WHEELER, J. A., & FEYNMAN, R. P. (1945). Interaction with the absorber as the mechanism of radiation. *Reviews of Modern Physics*, 17(2–3), 157–181. <https://doi.org/10.1103/RevModPhys.17.157>.

WHEELER, J. A. (1990). Information, Physics, Quantum: The Search for Links. In: Zurek, W. H. (Ed.) *Complexity, Entropy, and the Physics of Information*. Westview Press, Boca Raton.

WIENER, N. (1961). *Cybernetics, or Control and Communication in Animal and Machine* (2nd Edition). Cambridge, MA: MIT Press.

WITTEN, E. (1995). String Theory Dynamics in Various Dimensions. *Nuclear Physics B*, 443(1-2), 85-126. [https://doi.org/10.1016/0550-3213\(95\)00158-O](https://doi.org/10.1016/0550-3213(95)00158-O).

YOCKEY, H. (2005). *Information Theory, Evolution, and the Origin of Life*. Cambridge: University Press.

ZEILINGER, A., WEIHS, G., JENNEWEIN, T., & ASPELMEYER, M. (2005). Happy Centenary, Quantum Mechanics! *Nature*, 438(7069), 743–749. <https://doi.org/10.1038/438743a>.

ZUREK, W. H. (2003) Decoherence, einselection, and the quantum origins of the classical. *Reviews of Modern Physics*, 75(3), 715–775 <https://doi.org/10.1103/RevModPhys.75.715>.