

# Metafísica Analítica

Pedro Merluzzi & Rhamon de Oliveira Nunes  
(Organizadores)

COLEÇÃO  
ANPOF  
XX  
ENCONTRO



Entre os dias 30 de setembro e 4 de outubro de 2024, a cidade de Recife recebeu o XX Encontro de Filosofia da ANPOF, um evento que reuniu pesquisadores, docentes e estudantes em torno das mais diversas questões filosóficas. Com a participação de 73 Grupos de Trabalho, o encontro consolidou-se como um espaço essencial para o intercâmbio acadêmico, promovendo debates aprofundados sobre temas que atravessam a história da filosofia e suas interfaces com outras áreas do conhecimento.

# **Metafísica Analítica**



# **Metafísica Analítica**

Pedro Merlussi

Rhamon de Oliveira Nunes

Organizadores



**Gerente Editorial**

Junior Cunha

**Conselho Editorial**

Ana Karine Braggio

Gustavo Rohte de Oliveira

Jaqueline Thais de Souza

José Francisco de Assis Dias

Júlio da Silveira Moreira

Pietra Maria Gulak Welter

Reginaldo César Pinheiro

Ronaldo de Oliveira

**Produção Editorial**

Ammy Lee Vitória

Daniela Valentini

José Luiz G. Mariani

Medéia Lais Reis

Mônica Chiodi

**Instituto Quero Saber**

[www.institutoquerosaber.org](http://www.institutoquerosaber.org)

[editora@institutoquerosaber.org](mailto:editora@institutoquerosaber.org)

**Dados Internacionais de Catalogação-na-Publicação (CIP)**

M587      Metafísica analítica / organizadores, Pedro Merlussi  
e Rhamon de Oliveira Nunes. 1. ed. (ebook) -  
Toledo, Pr.: Instituto Quero Saber, 2025.  
132 p. (Coleção do XX Encontro Nacional de Filosofia  
da ANPOF)

Modo de Acesso: World Wide Web:

[www.institutoquerosaber.org/editora](http://www.institutoquerosaber.org/editora)

ISBN: 978-65-5121-157-7

DOI: <https://doi.org/10.58942/eqs.158>

1. Filosofia.

CDD 22. ed. 190

Rosimarizy Linaris Montanhano Astolphi – Bibliotecária CRB/9-1610

Este livro foi editado pelo Instituto Quero Saber em parceria com a ANPOF.

O teor da publicação é de responsabilidade exclusiva de seus respectivos autores.

## **ANPOF – Associação Nacional de Pós-Graduação em Filosofia**

### **Diretoria 2025-2026**

Janyne Sattler (UFSC), presidente

Eduardo Vicentini de Medeiros (UFSM), secretário-geral

Wanderson Flor do Nascimento (UnB), secretário-adjunto

Francisca Galiléia Pereira da Silva (UFC), tesoureira-geral

Ester Maria Dreher Heuser (Unioeste), tesoureira-adjunta

Halina Leal (PUCPR), diretora de comunicação

Mariana Claudia Broens (Unesp), diretora editorial

### *Conselho Fiscal*

Érico Andrade (UFPE)

Susana de Castro (UFRJ)

Adriano Correia (UFG)

### **Diretoria 2023-2024**

Érico Andrade Marques de Oliveira (UFPE), presidente

Eduardo Vicentini de Medeiros (UFSM), secretário-geral

Tessa Moura Lacerda (USP), secretária-adjunta

Judikael Castelo Branco (PROF-FILO/UFT), tesoureiro-geral

Francisca Galiléia Pereira da Silva (UFC), tesoureira-adjunta

Georgia Cristina Amitrano (UFU), diretora de comunicação

Solange Aparecida de Campos Costa (UESPI), diretora editorial

### *Conselho Fiscal*

Taís Silva Pereira (PPFEN-CEFET/RJ)

Ester Maria Dreher Heuser (Unioeste)

Castor Bartolomé Ruiz (Unisinos)



## **Apresentação da Coleção do XX Encontro Nacional de Filosofia da ANPOF**

Entre os dias 30 de setembro e 4 de outubro de 2024, a cidade de Recife recebeu o XX Encontro de Filosofia da ANPOF, um evento que reuniu pesquisadores, docentes e estudantes em torno das mais diversas questões filosóficas. Com a participação de 73 Grupos de Trabalho, o encontro consolidou-se como um espaço essencial para o intercâmbio acadêmico, promovendo debates aprofundados sobre temas que atravessam a história da filosofia e suas interfaces com outras áreas do conhecimento.

Os trabalhos apresentados neste encontro refletem a pluralidade da pesquisa filosófica no Brasil, abordando desde questões clássicas da metafísica e da epistemologia até discussões urgentes sobre raça, gênero, decolonialidade e filosofia da deficiência.

Os livros que agora publicamos reúnem os trabalhos apresentados no XX Encontro da ANPOF, preservando a riqueza das reflexões desenvolvidas durante o evento. Cada artigo representa uma contribuição valiosa para o avanço do debate filosófico e para o fortalecimento das comunidades acadêmicas que buscam compreender e transformar a realidade por meio do pensamento crítico. Mais do que um registro, estas publicações são um convite à continuidade das discussões iniciadas no encontro, abrindo caminho para novas investigações e diálogos.

Agradecemos a todos os participantes que contribuíram para a realização deste evento e para a produção deste material. Que estas publicações possam servir de inspiração e referência para pesquisadores, estudantes e interessados na filosofia, reafirmando o papel fundamental da ANPOF no fomento à pesquisa e à circulação do conhecimento filosófico no Brasil.

*Solange Costa*  
Diretora editorial da ANPOF  
Biênio 2023-2024



# Sumário

|                                                                                                                                  |            |
|----------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Apresentação.....</b>                                                                                                         | <b>11</b>  |
| <b>Novas considerações sobre existência e composição.....</b>                                                                    | <b>13</b>  |
| Rhamon de Oliveira Nunes                                                                                                         |            |
| <b>Composição e <i>Qualia</i> .....</b>                                                                                          | <b>33</b>  |
| Jerônimo Erig Weiller                                                                                                            |            |
| <b>O tratamento supervaloracionista da vagueza existencial e seus desafios.....</b>                                              | <b>53</b>  |
| Felipe da Silva Avena                                                                                                            |            |
| <b>Modos de impossibilidade: uma nova taxonomia .....</b>                                                                        | <b>81</b>  |
| Pedro de Sá Torres Barbosa                                                                                                       |            |
| <b>Some remarks on mind-dependence in social ontology and its relationship to thought experiments in philosophy of mind.....</b> | <b>99</b>  |
| Valdenor Monteiro Brito Júnior                                                                                                   |            |
| <b>The subjective rate of time in artificial intelligence in light of the Metaphysics of Time .....</b>                          | <b>119</b> |
| Ana Maria Correa Moreira da Silva                                                                                                |            |



# Apresentação

Desde sua estreia oficial como grupo de trabalho da Associação Nacional de Pós-Graduação em Filosofia (ANPOF) em 2018, o GT Metafísica Analítica tem se esforçado para criar um ambiente de cooperação e produtividade entre os pesquisadores, professores e alunos de filosofia que se interessam pelas discussões em metafísica realizadas no contexto da filosofia analítica. Além dos encontros periódicos realizados pela ANPOF, nos últimos anos o GT também realizou uma sessão especial no Simpósio Principia, em 2023, e organizou um encontro internacional com o GT Ética no mesmo ano, bem como o Colóquio do GT de Metafísica Analítica, nos anos de 2022, 2023 e 2024. Todos esses eventos contaram com a participação de importantes pesquisadores nacionais e internacionais. No último encontro da ANPOF foram realizadas catorze apresentações no GT, duas a mais em relação à edição anterior. Algumas das apresentações se tornaram capítulos deste volume. Num momento em que os critérios de avaliação de qualidade das publicações estão passando por uma reformulação, acreditamos ser imperativo que a comunidade nacional tenha acesso ao que se é produzido por profissionais brasileiros. Nesse sentido, a Coleção ANPOF é um importante veículo para a disseminação das ideias que estão sendo discutidas em cada área. Não seria diferente no caso da metafísica analítica, que, apesar de tudo, ainda é uma área pouco conhecida no Brasil.

O presente volume possui seis capítulos, que versam sobre temas centrais da metafísica e ontologia contemporânea: o primeiro capítulo, escrito pelo Dr. Rhamon de Oliveira Nunes, discute a relação

entre o problema da composição mereológica e algumas das principais teorias da existência. Seguindo a temática em mereologia, o segundo capítulo, do Me. Jerônimo Erig Weiller, apresenta uma resposta ao problema da composição a partir da noção de *qualia*, um importante conceito da filosofia da mente. Por sua vez, Felipe da Silva Avena discute de maneira mais detalhada o problema da vagueza e suas tentativas de soluções a partir de uma semântica supervaloracionista no terceiro capítulo. No quarto capítulo, de autoria de Pedro de Sá Torres Barbosa, temos uma investigação acerca da natureza da impossibilidade, em que o autor apresenta uma nova taxonomia para classificar tais modalidades. O quinto capítulo, escrito pelo Dr. Valdenor Monteiro Brito Júnior, é dedicado a discutir a relação entre o conceito de “dependência mental” da ontologia social e os diversos experimentos de pensamento na filosofia da mente. Por fim, o sexto e último capítulo, escrito pela Dra. Ana Maria Correa Moreira da Silva, discute problemas de filosofia da inteligência artificial a partir da metafísica do tempo.

A coordenação do GT Metafísica Analítica agradece a todos os participantes do último Encontro da ANPOF e, em especial, aos autores que submeteram seus trabalhos para a composição deste volume. Acreditamos que o presente volume, assim como o anterior, é uma importante ferramenta para a divulgação e ampliação do debate especializado na área e esperamos que ele possa estimular e fortalecer o desenvolvimento da metafísica analítica no Brasil.

*Prof. Dr. Pedro Merlussi*  
Coordenador do GT

*Prof. Dr. Rhamon de Oliveira Nunes*  
Vice coordenador do GT

# Novas considerações sobre existência e composição

Rhamon de Oliveira Nunes<sup>1</sup>

DOI: <https://doi.org/10.58942/eqs.158.01>

## 1 Introdução

O problema da composição mereológica diz respeito a determinar quando objetos compostos (i.e., objetos com partes próprias) passam a existir. No cerne do problema, portanto, encontra-se a noção de existência, de modo que um passo importante para qualquer tentativa de solucionar o problema da composição envolve também a elucidação do conceito de existência. O conceito de existência, por sua vez, é um dos mais centrais da metafísica ocidental, tendo sido tratado de diversas maneiras ao longo da história. A partir do século XX, na filosofia analítica, a análise quantificacional da noção de existência foi tomada como ortodoxa. Ela é base da ideia quiniana de que ser é ser o valor de uma variável ligada e, portanto, de que a existência é o que o quantificador existencial expressa. Esta ideia, contudo, surge mais propriamente em Frege, sendo posteriormente desenvolvida por Russell. Mais recentemente, algumas alternativas à ortodoxia surgiram, como o neo-meinongianismo (que entende a existência como um tipo de predicado), o princípio eleata (segundo o qual apenas entidades com

---

<sup>1</sup> Bacharel em filosofia pela Universidade Federal Fluminense, mestre e doutor em filosofia pelo Programa de Pós-Graduação em Lógica e Metafísica da Universidade Federal do Rio de Janeiro. E-mail: [rhamon.oliveira@gmail.com](mailto:rhamon.oliveira@gmail.com)

poderes causais existem), as teorias mundiais da existência (onde a existência dos indivíduos é analisada em termos do seu pertencimento a um mundo possível) e, ainda, teorias identitárias da existência, onde a existência é entendida como sendo idêntica aos indivíduos que existem.

A relação entre as teorias da composição e as teorias da existência é um tópico que foi pouco explorado na literatura contemporânea. Em geral, assume-se que a existência pode ser analisada quantitativamente e, a partir daí, derivam-se diversos argumentos sobre a natureza da composição. Todavia, um mapeamento mais profundo (de preferência, exaustivo) das diversas teorias da existência precisa ser feito para se chegar à conclusão de se as diferenças entre elas são ou não relevantes para o problema da composição. O presente capítulo é uma tentativa de dar início a este mapeamento. De fato, uma análise completa das principais teorias da existência e suas consequências para o problema da composição mereológica seria impossível de ser feita em um espaço tão pequeno: cada uma das teorias mencionadas acima é tão complexa que seria possível fazer dezenas de capítulos sobre cada uma delas. Ainda assim, este estudo preliminar nos dá boas indicações de que um problema central envolvendo a noção de composição permanece presente para a maioria das teorias da existência, a saber, o problema da vagueza. O ponto é simples: a maioria das teorias da existência adotada pelos filósofos assume que a existência não é um fenômeno vago: ou as coisas existem ou não existem e ponto final. Todavia, se for possível apresentar uma teoria em que o conceito de existência possa ser reduzido a um fenômeno vago, então talvez esta teoria possa acomodar o fenômeno da vagueza da composição sem grandes problemas.

## **2 Existência como uma propriedade de primeira ordem**

Vamos começar considerando a teoria de que a existência é uma propriedade de primeira ordem, ou seja, uma propriedade dos



indivíduos que existem. Exemplos de propriedades de primeira ordem são atributos como massa, carga, volume, textura, cor, etc. A intuição por trás desta teoria, portanto, é a de que assim como nós podemos dizer de um certo indivíduo (por exemplo, um carro) que ele é azul, nós também podemos dizer que ele é existente. Ou seja, existir consiste em instanciar a propriedade da existência, assim como ser azul consiste em instanciar a propriedade de ser azul. Mas levando em consideração que muitos dos indivíduos que povoam a realidade passaram a existir em algum momento no tempo e, presumivelmente, vão deixar de existir em algum momento no tempo, parece plausível admitir que a existência seja uma propriedade de primeira ordem *contingente*<sup>2</sup>. Além disso, assim como há indivíduos que não foram e nunca serão azuis, é plausível imaginar que há objetos que não foram e nunca serão existentes, como Papai Noel ou Sherlock Holmes. Em suma, a tese de que a existência é uma propriedade de primeira ordem parece nos comprometer com a tese de que há objetos que não existem.

Se nós adotarmos uma posição meinongiana clássica, a realidade seria dividida em três grandes grupos: objetos existentes (em geral, entidades concretas, com eficácia causal e localização espaço-temporal, como organismos, artefatos, etc.), objetos subsistentes (em geral, entidades abstratas, sem eficácia causal e localização espaço-temporal, como números, proposições, etc.) e, por fim, objetos que não possuem ser em absoluto (objetos que nem existem e nem subsistem, como unicórnios, Papai Noel e o quadrado redondo)<sup>3</sup>. Agora, aceitemos ou não tal divisão, uma coisa é certa: estas são as únicas opções nas quais os objetos se dividem:

---

<sup>2</sup> Pelo menos para os objetos ordinários que povoam o mundo ao nosso redor. É irrelevante para os objetivos deste texto se existem ou não indivíduos cuja existência seja necessária, como Deus, por exemplo.

<sup>3</sup> A expressão usada por Meinong para caracterizar tais objetos é *ausserseiend*, que significa literalmente “fora do ser”. Assim, devemos ter em mente que esta classe de coisas não constituem um terceiro modo de ser ao lado da existência e da subsistência.

Seja *O* um objeto qualquer,

- (i) *O* é um objeto existente;
- (ii) *O* é um objeto subsistente;
- (iii) *O* é não possui ser em absoluto.

Assim, é impossível que *O* exista, subsista e não possua ser em absoluto ao mesmo tempo e é impossível que *O* nem exista, nem subsista nem não possua ser em absoluto ao mesmo tempo. Mas agora temos uma dificuldade, porque a questão da composição justamente parece nos levar à tese de que há indivíduos que violam este princípio de classificação. Se nós reformularmos o problema da composição em termos existenciais, nós podemos dizer, por exemplo, que a existência de um composto *C* consiste na existências das partes de *C* juntamente com a satisfação de algum critério de composição. Esta seria, por exemplo, a posição conhecida como *restritivismo mereológico*, a saber, a tese de que, embora existam objetos compostos, nem toda coleção arbitrária de objetos compõe um novo objeto: é necessário que as partes estejam arranjadas de um modo especial e não-trivial. O problema, apontado por Lewis (1986, 1991), é que para todo e qualquer critério que pudermos supor, haverá casos nos quais é vago se o critério é satisfeito ou não pelas partes. Mas como a existência de um composto consiste na satisfação do critério pelas suas partes (existentes), então a existência do composto será vaga. O que é absurdo, dado que a existência não é um fenômeno vago. Nas palavras de Lewis:

O problema da composição restrita é o seguinte. A questão de saber se uma determinada classe satisfaz nossos *desiderata* intuitivos para a composição é vaga. [...] Mas, se a composição seguisse uma restrição vaga, então às vezes seria uma questão vaga determinar se a composição ocorre ou não. E isso é impossível (Lewis, 1986, p. 212).

e

[...] [A] existência não pode ser uma questão de grau. [...] Existir significa simplesmente ser uma das coisas que há, nada mais. A linha difusa entre fusões menos peculiares e mais peculiares não pode, de maneira alguma, coincidir com o limite nítido onde a existência cessa e nada se encontra além (Lewis, 1991, p. 81).

Podemos reconstruir as ideias de Lewis a partir do seguinte argumento:

P1. Se a composição é restrita, então toda restrição em casos de composição é uma restrição vaga.

P2. Se toda restrição em casos de composição é vaga, então é vago quando a composição ocorre.

D1. Se a composição é restrita, então é vago quando a composição ocorre. [*Silogismo Hipotético*, P1 e P2]

P3. Se é vago quando a composição ocorre, então é vago quando um objeto composto passa a existir.

D2. Se a composição é restrita, então é vago quando um objeto composto passa a existir. [*Silogismo Hipotético*, D1 e P3]

P4. Não é o caso que: a existência é uma propriedade vaga.

P5. Se não é o caso que: a existência é uma propriedade vaga, então não é o caso que: é vago quando um objeto composto passa a existir.

D3. Não é o caso que: é vago quando um objeto composto passa a existir. [*Modus Ponens*, P4 e P5]

Cc. Não é o caso que: a composição é restrita. [*Modus Tollens*, D2 e D3]

Como a teoria de primeira ordem poderia nos ajudar aqui? Aparentemente, ela não pode. Se a existência for uma propriedade de primeira ordem, ela é uma propriedade que deve ser instanciada pelo composto. Mas o composto só pode instanciar esta propriedade na medida em que ele também satisfaz o critério de composição. Em suma, se o critério não é satisfeito, a propriedade também não é instanciada. Mas agora, se for indeterminado que o critério é satisfeito, então também é indeterminado se a propriedade é instanciada. Poderíamos, portanto, modificar P3 de modo que ela afirme que “Se é vago quando a composição ocorre, então é vago quando um objeto composto passa a instanciar a propriedade da existência”. Mas dado que a divisão entre objetos existentes, subsistentes absolutamente sem ser é conjuntamente exaustiva e mutuamente exclusiva, não é possível que seja indeterminado se um composto existe ou não e, portanto, ela não nos leva muito longe no que diz respeito ao problema da composição.

### **3 A existência como uma propriedade de propriedades**

Este problema também contamina outras teorias da existência. Considere, por exemplo, a tese de Frege-Russell, segundo a qual a existência não pode ser uma propriedade dos indivíduos. Na versão fregeana, a existência de cavalos, por exemplo, não consiste em cavalos terem alguma propriedade especial além das propriedades que eles ordinariamente possuem. Pelo contrário, ela consiste no conceito <Cavalo> possuir uma cardinalidade maior que zero. Nesse sentido, afirmar que cavalos existem é simplesmente dizer que o número de cavalos é maior que zero. Daí sua famosa afirmação de que “a afirmação da existência é, de fato, nada mais do que a negação do número zero.” (Frege, 1950, p. 65). Mas nós não precisamos adotar o vocabulário fregeano para sustentar esta posição. Nós poderíamos dizer, por exemplo, que a existência de cavalos consiste na instanciação da propriedade “Ser um cavalo” por alguma coisa ou outra. Assim, diferentemente da

posição anterior, a posição em questão trata a existência como sendo uma propriedade de propriedades (ou conceitos fregianos), sendo, portanto, uma teoria de segunda ordem. É claro que, neste caso, não estamos dizendo que a propriedade “Ser um cavalo” tem a propriedade de existir. O que nós estamos dizendo é que a existência é idêntica a uma outra propriedade que as propriedades de primeira ordem possuem, a saber, a propriedade de ser instanciada. Inspirado pela tese de Frege, Russell também defendeu que a existência não é algo que se possa aplicar a indivíduos. Segundo ele:

Quando você diz “Unicórnios existem”, você não está afirmando nada sobre qualquer coisa individual, e o mesmo se aplica quando diz “Homens existem”. Se você disser “Homens existem, e Sócrates é um homem, portanto, Sócrates existe”, isso é exatamente o mesmo tipo de falácia que seria dizer “Homens são numerosos, Sócrates é um homem, portanto, Sócrates é numeroso”, porque a existência é um predicado de uma função proposicional ou, de forma derivada, de uma classe (Russell, 1972, p. 67).

*Mutatis mutandis*, esta ideia está presente em sua famosa teoria das descrições definidas (Russell, 1905), que posteriormente seria reutilizada por Quine em seu famoso artigo *Sobre o Que Há* (Quine, 1948). Ali, Quine nos mostra que afirmações existenciais negativas, como “Pégaso não existe” não nos comprometem com a existência de objetos não-existentes. A ideia contra a qual Quine (e, em última análise, Russell) argumentam é a de que, ao afirmar “Pégaso não existe”, estamos afirmando que há algo (i.e., Pégaso), que carece da propriedade de existir. Teríamos então, o velho paradoxo sofista de que, ao negar a existência de algo, estamos fadados a automaticamente afirmar sua existência. Todavia, o que a análise russelliana tenta mostrar (pelo menos na leitura feia por Quine), é que este tipo de afirmação na verdade é uma descrição disfarçada, onde o nome ‘Pégaso’ pode ser substituído pela expressão “o cavalo alado capturado por Belerofonte”. Assim, teríamos algo como “Não há um  $x$ , tal que  $x$  seja o cavalo alado capturado por

Belerofonte”, o que por sua vez é uma forma mais sofisticada de dizer “nada/ninguém é o cavalo alado capturado por Belerofonte”. Mas então, temos a mesma consequência da posição de Frege: o que estamos afirmando, no final das contas, é que o número de coisas que satisfaz a descrição/instancia a propriedade acima é igual a zero. Em suma, dado que expressões descritivas são expressões predicativas, nós explicamos a existência em termos da instanciação de certas propriedades.

Mas dada esta caracterização geral da tese de Frege-Russell-Quine, o que nos interessa é saber se esta concepção acerca da existência nos deixa mais próximos de resolver o problema da composição. Ora, é uma posição bastante aceitável que a relação de instanciação não admite graus: ou  $x$  instancia  $F$  ou  $x$  não instancia  $F$ . É claro que existem exemplos clássicos de predicados vagos como “ser careca”, onde não sabemos se esta propriedade se aplica ou não. Mas casos como este normalmente mostram mais a nossa inabilidade de determinar quando uma propriedade é instanciada do que qualquer tipo de vagueza na relação de instanciação em si. Além disso, a identidade de um indivíduo é classicamente determinada em termos de suas propriedades. Se fosse possível que um indivíduo possuísse vagamente uma propriedade, então seria possível que sua identidade fosse vaga, dado que não seria possível determinar se o indivíduo possui ou não possui a propriedade em questão. Mas isso é impossível: todo indivíduo é auto idêntico, como o argumento de Evans (1978) demonstra (onde os símbolos ‘ $\Delta$ ’ e ‘ $\nabla$ ’ significam “é determinado que” e “é indeterminado que”):

|                                    |                           |
|------------------------------------|---------------------------|
| P1. $\nabla(a = b)$                | [Hipótese]                |
| P2. $\neg \nabla(a = a)$           | [Princípio de Identidade] |
| D1. $\lambda[\nabla(x = a)]b$      | [Abstração, P1]           |
| D2. $\neg \lambda[\nabla(x = a)]a$ | [Abstração, P2]           |
| D3. $\neg(a = b)$                  | [Lei de Leibniz, D1 e D2] |
| Cc. $\Delta(\neg(a = b))$          | [Introdução $\Delta$ ]    |

Como podemos ver, Cc. diz que é determinado que A não é idêntico a B, o que é inconsistente com P1, que diz que é vago que A é idêntico a B. Mas dado que a conclusão se segue logicamente das premissas, uma delas precisa ser descartada. A melhor candidata é justamente P1, visto que P2 é a afirmação trivial de que tudo é idêntico a si mesmo. Portanto, se a existência de um composto consiste na existência de suas partes satisfazendo um critério de composição e este critério é vago, então também será vago quando a propriedade relevante para a existência do composto terá pelo menos uma instância ou não. Para usar o exemplo acima: se é vago quando alguns simples compõem um dado cavalo, então será vago quando a propriedade “Ser um cavalo” terá uma dada instância, e, portanto, a cardinalidade da propriedade será vaga. Mas é claro que isto é impossível: sentenças numéricas como “há um número  $N$  de cavalos no mundo” não podem ser vagas.

## 4 A existência como uma propriedade do mundo

Passemos então à teoria mundial da existência e vejamos se elas conseguem resolver este problema. De acordo com estas teorias, a existência de um indivíduo como Sócrates, por exemplo, consiste no fato do mundo atual conter Sócrates como um de seus constituintes. A existência aqui não é uma propriedade de primeira ordem, porque ela não pertence a Sócrates, mas também não é uma propriedade de segunda ordem, porque ela não pertence a nenhuma das propriedades de Sócrates. Ela pertence, antes, ao mundo no qual Sócrates está contido. Em suma, estamos dizendo que o mundo é tal que Sócrates está presente nele. O principal expoente desta posição é Sommers. Em *Naturalismo e Realismo* (1994), ele afirma que:

Uma afirmação verdadeira também expressa uma caracterização (proposicional); por exemplo, “algum ateniense é sábio” expressa a proposição algum ateniense sendo sábio, a existência de um ateniense sábio, ou que algum ateniense é sábio. Ela significa uma característica

(mundial, existencial) — um fato ou estado de coisas: algum ateniense sendo sábio, a existência de um ateniense sábio ou que algum ateniense é sábio. E ela denota o mundo assim caracterizado.

Da mesma forma que termos que significam características diferentes (ser sábio, ser ateniense) podem denotar uma mesma pessoa, diferentes afirmações verdadeiras significam diferentes fatos, mas todas as afirmações verdadeiras denotam um único e mesmo mundo (Sommers, 1994, p. 31).

E, em *O Realismo Renascido de Putnam* (1997), ele defende que:

Desvendamos o véu quando respondemos corretamente à pergunta que Kant inadvertidamente nos deixou: “Se a existência não é um atributo das coisas que existem, de que é ela um atributo?” A resposta mais sólida é que existência e não existência específicas (por exemplo, a existência de leões, a não existência de grifos) são atributos mundiais, ou seja, atributos que pertencem ao mundo em virtude do que há nele, em oposição ao que não há. Uma afirmação como “Alguns tigres são dóceis” corresponde à existência de tigres dóceis, um atributo positivo do mundo; já “Nenhum tigre voa” corresponde à não existência de tigres que voam, um atributo negativo do mundo (Sommers, 1997, p. 471).

Todavia, deslocar a existência dos indivíduos para o mundo não parece ser uma saída viável para evitar o problema da vagueza da composição, por um motivo bastante semelhante ao caso anterior: do mesmo modo que é impossível que um indivíduo possua vagamente uma propriedade, também é impossível que um mundo contenha vagamente um indivíduo. Ou seja, se a existência de Sócrates é uma propriedade do mundo, então é claro que esta propriedade não pode ser vaga: ou Sócrates está no mundo, ou Sócrates não está no mundo! Assim como a identidade dos indivíduos é caracterizada pelas propriedades que eles possuem, a identidade de um mundo possível é caracterizada (pelo menos em parte), pelos indivíduos que o habitam. Dois mundos possíveis com indivíduos distintos são mundos possíveis distintos. Logo, a ocorrência de um indivíduo vago levaria à vagueza da identidade de um mundo possível, o que não pode ocorrer. Mas é



justamente isto que aconteceria se, por exemplo, o critério para a ocorrência de um composto fosse vaga, como o argumento abaixo deixa claro:

P1. A existência de um composto consiste no mundo sendo tal que suas partes satisfazem o critério de composição.

P2. Se o critério de composição é vago, então é vago que as partes satisfazem o critério de composição.

P3. Se a existência de um composto consiste no mundo sendo tal que suas partes satisfazem o critério de composição e é vago que as partes satisfazem o critério de composição, então é vago que o composto existe ou não.

P4. O critério de composição é vago.

D1. É vago que as partes satisfazem o critério de composição.  
[*Modus Ponens*, P2 e P4]

D2. A existência de um composto consiste no mundo sendo tal que suas partes satisfazem o critério de composição e é vago que as partes satisfazem o critério de composição. [*Conjunção*, P1 e D1]

Cc. É vago que o composto existe ou não. [*Modus Ponens*, D2 e P3]

P1 é simplesmente a reformulação do problema da composição em termos existenciais a partir da teoria mundial da existência. P2 é a tese lewisiana de que a vagueza composicional é derivada do critério usado para restringir a composição. Tudo isto já foi estabelecido anteriormente, de modo que o elemento novo aqui é P3, que dificilmente

poderia ser contestada. Segue-se, portanto, Cc. Mas novamente, permanecemos com o problema da vagueza existencial, nós apenas a jogamos para o andar de cima!

## 5 A existência como idêntica aos existentes

Consideremos, então, a teoria identitária da existência: segundo esta posição, a existência consiste simplesmente nos indivíduos que existem. Em suma, podemos dizer que a existência de Sócrates = Sócrates. Observe que isto é bem diferente de dizer que: (a) a existência de Sócrates = a instanciação da propriedade de existir por Sócrates; (b) a existência de Sócrates = a satisfação da propriedade “ser Sócrates” por alguma coisa ou outra; (c) a existência de Sócrates = o mundo atual ser tal que Sócrates está contido nele. O que a posição em questão afirma é que não há nada na existência de Sócrates além do próprio Sócrates. A existência é algo que se divide sem sobras entre os existentes e ponto final. Um dos defensores desta posição é Grossmann (1983), que afirmou que “Não há um indivíduo *A* separado de sua existência. Não podemos opor *A* à sua existência... uma entidade e (sua) existência são, de certo modo, uma e a mesma coisa”. Sem dúvidas, esta é uma tese curiosa: por um lado, ela possui certa plausibilidade, como apontado por Vallicella (2002). Afinal, se a existência não é algo distinto dos indivíduos, como uma propriedade de primeira ordem, segunda ordem ou do próprio mundo, então o que mais ela poderia ser? Além disso, parece indubitável que os indivíduos existem. Assim, parece que tudo o que nos resta é dizer que a existência consiste justamente nestes indivíduos diante de nós. Num certo sentido, esta posição se assemelha às posições nominalistas em que universais são reduzidos aos indivíduos caracterizados (*e.g.*, onde o universal vermelho, por exemplo, é reduzido à totalidade de coisas vermelhas). Por outro lado, esta tese parece um tanto quanto vazia: nossa pergunta inicial é sobre a natureza da existência, e a resposta que esta posição nos dá é simplesmente apontar

para os existentes. Parece que terminamos sem saber nada de novo sobre o que significa existir. Seja como for, meu objetivo aqui não é argumentar contra nenhuma teoria da existência *qua* teoria da existência. Sendo assim, vejamos como esta posição se relaciona com o problema da composição.

Devemos ter em mente que, se Sócrates é um composto, então sua existência dependeria, no contexto restritivista, da existência de suas partes satisfazendo um critério de composição. De acordo com a teoria existencial em questão, a existência de Sócrates é idêntica ao próprio Sócrates. Juntando estas duas posições, teríamos a tese de que a existência de Sócrates consiste em Sócrates ser idêntico a certas partes arranjadas de certa maneira a partir de um certo critério. Mas se Lewis estiver certo e todo critério de composição for vago, então temos um problema: se é vago que Sócrates satisfaz o critério de composição (digamos, antes de nascer, não havia um momento exato em que pudéssemos dizer que Sócrates existia ou não), então a existência de Sócrates será idêntica a quê? A existência de Sócrates não pode ser idêntica às suas partes *simpliciter*, dado que elas poderiam existir sem que Sócrates existisse. Mas ela também não pode ser idêntica ao próprio Sócrates, dado que é vago que ele satisfaz o critério de composição, e, portanto, é vago que haja tal coisa como Sócrates (*i.e.*, suas partes satisfazendo o critério necessário para o compor) para que sua existência seja idêntica a isto. Logo, não há determinadamente uma coisa ou outra à qual Sócrates seja idêntico. Mas isto é impossível! Ou Sócrates é idêntico a algo (e existe) ou ele não é idêntico a nada (e não existe!). Assim, a posição identitária também não parece nos ajudar no que diz respeito ao problema da composição.

## 6 Reduzindo a existência a algo ainda mais fundamental

Vejamos, por fim as teorias reducionistas. Poderíamos, por exemplo, supor que o que nós entendemos por existência na verdade

se refere a algo mais fundamental. Os principais candidatos são espaço-temporalidade e concretude. Assim, uma teoria reducionista da existência poderia defender que existir consiste em estar localizado no espaço-tempo ou ser concreto. Mas então, a existência de um composto consistiria nele ser espaço-temporal (ou concreto) e isto seria derivado do fato de que suas partes espaço-temporais (ou concretas) satisfazem um critério de composição. Mas é claro que estar localizado no espaço-tempo (ou ser concreto) não é um fenômeno vago. As coisas não são “meio espaço-temporais” e “meio não-espaço-temporais” (ou “meio concretas” e “meio abstratas”). Portanto, uma teoria reducionista da existência simplesmente não é capaz de evitar o problema central da vagueza mereológica/existencial.

A fim de completude, consideremos outra teoria reducionista da existência. De acordo com o *princípio eleata*, existir significa ter eficácia causal. Este princípio recebe este nome por ter sido introduzido pela primeira vez no diálogo *Sofista*, de Platão, pelo Estrangeiro de Eleia:

Declaro, então, que tudo o que possui uma determinada faculdade, seja de atuar de a algum modo sobre outra coisa, seja de sofrer a influência, embora mínima, do mais insignificante agente, mas que fosse uma única vez, é um ser real. Minha definição para explicar os seres é que não passam de capacidade ou força (Platão, *Sofista*, 247d-248e).

Na filosofia contemporânea, este princípio é mencionado por Armstrong como sendo um argumento contra a existência de entidades abstratas como classes:

Um argumento contra a distinção entre objetos e classes de objetos é o da falta de poder causal. Se o agregado de pessoas nesta sala e a classe de pessoas nesta sala fossem realmente duas entidades diferentes, deveríamos esperar que tivessem poderes causais distintos. Se não têm, então não parece haver motivo para postular ambas. Em especial, se a classe não exerce nenhuma função no mundo além daquela já desempenhada pelo agregado, por que postulá-la como algo adicional? Os fenômenos seriam exatamente os mesmos, com ou sem a classe (Armstrong, 1978, p. 31).

Posteriormente, o nome “princípio eleata” seria cunhado por Oddie em sua discussão com Armstrong (Assaturian, 2021). Oddie afirma que “Armstrong usa um princípio geral que poderia ser adequadamente chamado de “Princípio Eleata”, porque é inspirado pela “marca do ser” apresentada pelo Estrangeiro de Eleia no *Sofista* de Platão” (Oddie, 1982). Seja como for, a ideia é a de que tudo o que é capaz de interagir causalmente com outras coisas existe e tudo o que existe é capaz de interagir causalmente com outras coisas. Portanto, no caso dos objetos compostos, poderíamos dizer que a existência de um composto consiste em suas partes satisfazendo um critério de modo a gerar algo que tem eficácia causal. Mas aqui, ao contrário do que foi visto nas outras teorias mencionadas acima, talvez haja espaço para acomodarmos certa noção de vagueza existencial. A princípio, parece tentador afirmar que se é vago que as partes compõem alguma coisa, então é vago que esta coisa possui eficácia causal. Nesse caso em especial, seria impossível determinar, por exemplo, se a eficácia causal de um conjunto de simples pertence a eles de fato ou ao objeto que eles possivelmente estão compondo: se a ocorrência da composição for vaga, não é possível determinar quem causa o quê. Mas devemos observar que a causalidade é uma relação que se dá ao longo do tempo e, em geral, de forma *contínua*. Por ser contínua, não seria possível determinar um momento preciso no qual, por exemplo, *A* causa *B*. Considere o caso de duas bolas de bilhar que se chocam uma contra a outra: seria possível determinar um *momento exato* em que a bola *A* gera o movimento da bola *B*? Imagine que este evento tenha ocorrido em uma linha temporal que vai de 0 segundos a 1 segundo. Em qual fração de segundo a bola *B* começou a se mover em virtude de ter sido atingida pela bola *A*? Digamos que alguém defenda que isto ocorreu na marca dos 0,5 segundos e outra pessoa defenda que isto ocorreu na marca dos 0,49 segundos. Seria possível determinar qual das pessoas está correta? Esta é uma tarefa difícil, especialmente se considerarmos que quanto mais aumentarmos o “zoom temporal”, por assim dizer, mais indiscerníveis os dois

momentos se tornam: qual é diferença entre o estado da bola em 0,5 e 0,499999999...? O problema aqui é objetos macroscópicos não parecem dar saltos discretos entre seus estados físicos. Não é como se em um instante  $t$  a bola  $B$  estivesse em repouso e exatamente no instante posterior,  $t^*$ , ela estivesse em movimento.

O que esta digressão sobre a natureza da causalidade nos mostra é que o princípio eleata da existência afeta tanto os compostos quanto as suas partes: é verdade que a vagueza da composição gerará algo tal que é vago que esta coisa possui poderes causais. Mas se a marca da causalidade é o efetivo impacto causal das coisas no mundo, então as próprias partes sofrem com algum tipo de vagueza causal, dado que não é determinado quando eles afetam causalmente outras coisas. Aqui, talvez, pudesse ser objetado que as partes dos compostos são, em geral, outras coisas compostas e, assim sendo, o problema se aplicaria a elas naturalmente, mas talvez não aos simples de verdade. Talvez, no nível das partículas elementares, se elas puderem mesmo ser boas candidatas a simples mereológicos, toda relação de causalidade seja discreta. Esta é uma objeção válida, mas uma que depende de uma apreciação profunda das nossas melhores teorias sobre o comportamento das partículas elementares, incluindo a física quântica, o que escapa muito do escopo deste texto. Seja como for, a possibilidade de toda e qualquer relação de causalidade ser vaga abre a possibilidade de haver casos de existência vaga. Se isto for o caso, então posições restritivistas fundadas em critérios causais podem se tornar uma alternativa para responder ao problema da composição.

Antes de encerrarmos esta seção, é importante mencionar que, embora a discussão acima tenha mencionado a expressão “poder causal”, os exemplos usados por mim focaram apenas em casos clássicos de causalidade eficiente. A ideia de que os objetos possuem poderes causais encontra sua origem na noção aristotélica de potência (*dynamis*) e, na filosofia contemporânea, foi tratada por autores como Carnap (1928), Ryle (1949) e Quine (1960). No contexto da metafísica

propriamente dita, autores que defendem a teoria dos poderes/disposições como alternativas ao tratamento tradicional da causalidade incluem Mumford (1998, 2004) e Cartwright (1989). Em linhas gerais, um poder ou disposição é uma propriedade que um objeto possui, mas que se manifesta em determinadas circunstâncias. Exemplos clássicos são qualidades como fragilidade ou solubilidade. Um torrão de açúcar, por exemplo, é solúvel em água, mas esta qualidade só se manifesta quando ele é de fato submerso em água. Da mesma maneira, a fragilidade de uma taça de cristal só se manifesta quando ela se quebra ao se chocar contra outro objeto. De fato, se as teorias de poderes causais (ou disposições) forem plausíveis, então as coisas podem mudar de figura em relação ao que foi dito sobre a causalidade. Parece razoável assumir que os poderes causais de um objeto não são propriedades vagas: ou  $x$  é solúvel em água ou  $x$  não é solúvel em água. Seria estranho que, ao submergir um objeto em uma porção de água, ele se dissolvesse e não se dissolvesse ao mesmo tempo (ou, pior, nem se dissolvesse e nem não-se dissolvesse). É claro que certo material pode se dissolver parcialmente em água, mas disso não se segue que seja vago que ele se dissolveu ou não. Afinal, é determinado que que parte dele se dissolveu e que parte dele não se dissolveu. Neste caso, teríamos uma conclusão semelhante aos casos anteriores.

## 7 Considerações finais

Dado que foi apresentado, parece que a maioria das teorias vigentes da existência não lança luz sobre o problema da composição: a vagueza da composição continua sendo incompatível com a precisão da existência, salvo no caso peculiar em que a existência é reduzida a um fenômeno que, pelo menos a princípio, pode ser vago, como é o caso da relação de causalidade. Esta posição, portanto, pode ser uma rota de fuga para aqueles que querem defender uma posição restritiva sobre a composição e estão dispostos a aceitar casos de vagueza

existencial, embora ainda seja uma questão aberta se, em última análise, as relações causais que ocorrem no mundo sejam contínuas ou discretas. O problema, a meu ver, pode ser abordado a partir de outro ângulo: o questão da vagueza existencial está intimamente ligado à concepção objectual de composição, a saber, a ideia de que a relação de composição gera novos objetos no mundo (Nunes, 2020 e 2022). Se nós abrirmos mão desta ideia e aceitarmos a concepção predicativa de composição, segundo a qual ser um objeto composto é instanciar ou exibir uma estrutura, então nós poderemos evitar este problema que tem assolado as discussões de mereologia nos últimos 40 anos.

## Referências

- ARMSTRONG, David Malet. *Universals and Scientific Realism*. Cambridge: Cambridge University Press, 1978.
- ASSATURIAN, Sosseh. *What's Eleatic about the Eleatic Principle?* Archai 31, e-03122., 2021.
- CARNAP, Rudolf. *The Logical Structure of the World*. Berkeley: University of California Press, 1928.
- CARTWRIGHT, Nancy. *Nature's Capacities and Their Measurement*. Oxford: Clarendon, 1989.
- EVANS, Gareth. *Can there be Vague Objects?* Analysis. v. 38, p. 208, 1978.
- FREGE, Gottlob. *The Foundations of Arithmetic: A logico-mathematical enquiry into the concept of number*. Trad. J. L. Austin. Oxford: Blackwell, 1950.
- GROSSMANN, Reinhardt. *The Categorical Structure of the World*. Bloomington: Indiana University Press, 1983.
- LEWIS, David. *On the Plurality Of Worlds*. Oxford: Blackwell Publishers, 1986.
- LEWIS, David. *Parts Of Classes*. Oxford: Blackwell Publishers, 1991.
- MUMFORD, Stephen. *Dispositions*. Oxford: Oxford University Press, 1998.
- MUMFORD, Stephen. *Laws in Nature*. Oxford: Routledge. 2004.



- NUNES, R. Totalidades e estrutura mereológica: Um estudo sobre a natureza dos objetos compostos. Rio de Janeiro: Universidade Federal do Rio de Janeiro (UFRJ). Tese, 2020.
- NUNES, R. Quasi-niilismo mereológico: uma resposta ao problema da composição, In: MERLUSSI, Pedro; NUNES, Rhamon (Org.). *Metafísica Analítica*. Toledo: Instituto Quero Saber, 2022.
- ODDIE, Graham. *Armstrong on the Eleatic Principle and Abstract Entities*. Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, v. 41, n. 2, p. 285-295, 1982.
- PLATÃO. *Sofista*. Trad. NUNES, Carlos Alberto. Universidade Federal do Pará, 1980.
- QUINE, Willard van Orman. On What There is (1948). In: *From a Logical Point of View*. Cambridge, Mass: Harvard University Press, revised version, 1980.
- QUINE, Willard van Orman. *Word and Object*. Cambridge, Mass: MIT Press, 1960.
- RUSSELL, Bertrand. *On Denoting*. Mind. v. 14, p. 479-493, 1905.
- RUSSELL, Bertrand. *The Philosophy of Logical Atomism*. Fontana, 1972.
- RYLE, Gilbert. *The Concept of Mind*. London: Penguin, 1949.
- SOMMERS, Frederic. *Naturalism and Realism*. Midwest Studies in Philosophy, v. 19, n. 1, p. 22-38, 1994.
- SOMMERS, Frederic. *Putnam's Born-Again Realism*. The Journal of Philosophy, v. 94, n. 9, p. 453-471, 1997.
- VALLICELLA, William. *A Paradigm Theory of Existence*. Kluwer Academic Publishers, 2002.



# Composição e *Qualia*

Jerônimo Erig Weiller<sup>1</sup>

DOI: <https://doi.org/10.58942/eqs.158.02>

## 1 Introdução

Neste trabalho defenderei uma resposta ao que ficou conhecido como “a questão especial da composição”: como van Inwagen define, a questão de em que circunstâncias coisas compõem algo (Van Inwagen, 1990, p. 31). Esta resposta é um exemplo de composição restrita, isto é, uma resposta segundo a qual há situações em que certos objetos compõem algo e há também situações em que certos objetos não compõem algo. Embora isso possa soar como senso comum, duas das respostas mais populares à questão especial da composição são o niilismo mereológico (segundo o qual a composição nunca ocorre) e o universalismo mereológico (segundo o qual qualquer grupo de objetos sempre e necessariamente compõe outro objeto).

A resposta que eu proponho é esta: “objetos compõem outro objeto se e somente se eles geram *qualia* juntos”. Vou primeiro tentar esclarecer o que isso significa exatamente. Por *qualia* eu quero dizer sensações subjetivas, como as cores, sons e gostos que caracterizam um episódio de consciência particular. Eu não pressuponho que essas *qualia* tenham quaisquer poderes causais; visando parcimônia, tomo-as como sendo epifenômenos (como descritas em (Jackson, 1982)).

---

<sup>1</sup> Mestre em filosofia, bacharel em direito e bacharelando em filosofia pela Universidade Federal do Rio Grande do Sul. E-mail: [jeronimoeu@hotmail.com](mailto:jeronimoeu@hotmail.com)

Os objetos que eu acredito ser razoável dizer que geram *qualia* são alguns daqueles simples (isto é, objetos sem partes próprias, cuja existência eu pressuponho aqui) que o senso comum diria que são parte de um sistema nervoso. Eu suspendo o julgamento quanto a se isso inclui todos os simples que o senso comum diria que fazem parte de um sistema nervoso, apenas aqueles que o senso comum diria que são parte de um cérebro ou mesmo um subconjunto próprio destes últimos.

Pode-se perguntar a razão para restringir os objetos relevantes desta maneira. “Certamente,” pode ser dito, “quando eu experiencio a visão de uma cadeira, a cadeira em si é relevante para eu vê-la!” É importante, eu diria, porque causalmente faz com que o seu sistema nervoso esteja de tal modo que *qualia* sejam experienciadas. Mas a cadeira em si não é essencial para a experiência de vê-la, já que ilusões de ótica e alucinações podem gerar a mesma experiência. Quando eu digo que os objetos relevantes “geram *qualia*”, eu não quero dizer algum tipo de causalidade diacrônica; quero dizer que, em certo momento, o arranjo desses objetos faz com que, no mesmo momento, algumas *qualia* sejam experienciadas por uma consciência.

Até onde eu vi, esta resposta específica ainda não foi defendida na literatura. Irei, portanto, dar o meu melhor para defendê-la contra possíveis objeções, mas este trabalho inevitavelmente será muito tentativo e alguns contra-argumentos válidos podem não ter sido previstos.

Embora esta resposta seja um tipo de composição restrita, assim como o organicismo de van Inwagen (segundo o qual apenas seres vivos e simples existem), ela implica um afastamento radical das nossas ideias do senso comum sobre que tipos de objetos existem. Portanto, assim como os niilistas mereológicos, terei de defender o eliminativismo com respeito à maioria dos objetos cuja existência damos como certa, tendo que ao mesmo tempo argumentar contra um niilismo mereológico ferrenho.

Como resultado, este trabalho está dividido em duas seções principais: a primeira seção busca motivar o eliminativismo com respeito aos objetos do senso comum; a segunda seção explica por que, mesmo levando em consideração todos os argumentos da primeira seção, eu acredito que um niilismo ferrenho não possa ser aceito.

## **2 Considerações em favor do eliminativismo**

### **2.1 Colocalização**

Um quebra-cabeça filosófico famoso diz respeito à possibilidade de dois objetos estarem colocalizados na mesma região do espaço. O exemplo clássico é o de uma estátua e o amontoado de argila que a constitui: o monte poderia sobreviver ser deformado, ao passo que a deformação destruiria a estátua; por outro lado, a estátua poderia perder um pequeno pedaço seu e permanecer a mesma, enquanto o amontoado seria outro nessa situação. A estátua e o amontoado têm condições de persistência diferentes e seriam, portanto, objetos diferentes.

Essa história deixa filósofos desconfortáveis por diversas razões. Primeiramente, a maioria das pessoas parece ter uma forte intuição de que dois objetos nunca podem ocupar exatamente a mesma região do espaço. Isso também parece aumentar desnecessariamente a quantidade de objetos na situação. Afirmar que apenas a estátua existe e o amontoado passa a existir assim que ela é deformada parece muito arbitrário. Por fim, já que a estátua e o amontoado são compostos pelas exatas mesmas partículas subatômicas nas exatas mesmas relações umas com as outras, não é claro com base em que as suas diferentes condições de persistência sobreviriam (Merricks, 2001, p. 39-40).

Houve muitas tentativas de solução dessas questões sem negar a existência da estátua ou do amontoado. Alguns autores (aludidos em (Merricks, 2001, p. 42-44)) defenderam que eles são objetos diferentes ao adotar o perdurantismo e sustentar que eles possuem partes

temporais distintas; outros alegaram que “ser uma estátua” é um sortal de fase do amontoado, como “ser uma criança” é um sortal de fase de um humano: o humano continua a existir quando cresce, mas por um tempo ele é idêntico a uma criança.

Todas essas tentativas consideradas, contudo, deve ser admitido que é uma vantagem do eliminativismo com relação a objetos do senso comum que ele acaba com esse quebra-cabeça facilmente: nem a estátua, tampouco o amontoado existe, então não podem ocupar a mesma região do espaço ao mesmo tempo.

## 2.2 Vagueza

Um problema típico para respostas à questão especial da composição que admitem a existência de objetos do senso comum é o problema da vagueza. De modo geral, a maioria de tais respostas pede critérios para composição que satisfaçam nossos *desiderata* intuitivos com relação aos tipos de objetos que normalmente consideraríamos que existem; tomando carros como um exemplo, a maioria das pessoas diria que, para simples comporem um carro, o objeto resultante deve estar dentro de uma certa faixa de tamanho, ter partes de carro suficientes *etc.* A maioria desses *desiderata*, contudo, é vaga; não há um tamanho ou quantidade de partes de carro exatos que um carro precisa ter tais que, quando satisfeitos, o carro certamente existe e, quando não satisfeitos, o carro certamente não existe. Então parece que um defensor de composição restrita pode ter de sustentar que é uma questão vaga se alguns objetos existem ou não.

Essa é uma das razões que levaram David Lewis a adotar o universalismo mereológico (Lewis, 1986, p. 212-213). O universalismo mereológico evita o problema da vagueza ao admitir a existência de um objeto diferente composto por cada conjunto possível de simples. Então há objetos diferentes para, digamos, um carro e o mesmo carro menos um parafuso, assim como para qualquer outra combinação de suas partes que poderíamos imaginar. Mas esse problema também pode ser

evitado ao simplesmente negar a existência desses objetos compostos, o que é um ponto a favor do eliminativismo.

### 2.3 “Isso tudo não é um pouco absurdo?”

Algumas pessoas podem alegar que eliminar a maioria dos objetos do senso comum é inaceitável porque seria simplesmente absurdo. Se qualquer tese tem como consequência a negação de que carros, cadeiras e pedras sejam reais, deveríamos apenas realizar um *modus tolles* e chegar à conclusão de que a tese é falsa e algo de errado aconteceu no raciocínio que deveria defendê-la. Responder a essa acusação, portanto, é uma parte importante de defender o eliminativismo.

Se alguém vai raciocinar desse modo depende de quanto valor essa pessoa atribui às nossas crenças do senso comum. Poderíamos, é claro, repetir a questão perplexa de Sider: “Por que os preconceitos herdados de nossos antepassados deveriam valer *algo*?” (Sider, 2013, p. 246, trad. livre). Esta é uma pergunta razoável. Mas eu acredito que algo mais possa talvez ser dito contra atribuir muito valor às nossas crenças do senso comum.

Deixe-me começar a falar de ciência. A maioria das pessoas educadas estão dispostas a aceitar consequências contraintuitivas de teorias científicas bem estabelecidas. A teoria da relatividade de Einstein diz que é possível um gêmeo viajar a uma velocidade muito alta em uma espaçonave e ser mais jovem que o seu irmão ao voltar à Terra. A mecânica quântica (ao menos na interpretação de Copenhaga) diz que é possível um gato estar em uma superposição de um estado “vivo” e um estado “morto” e essa superposição apenas acabar quando uma observação é feita do sistema em que o gato está. Mesmo que nos restrinjamos à mecânica clássica, dada a pressuposição (razoável, embora não óbvia) de que o universo é infinito em tamanho e quantidade de matéria, é um fato estatístico que, muito longe daqui, há uma cópia minha fazendo exatamente tudo o que estou fazendo agora, assim como várias versões de mim levemente diferentes.

Por que aceitamos esses afastamentos do senso comum? Presumivelmente porque essas teorias científicas fazem um ótimo trabalho em explicar e prever experiências, livrando-se ao mesmo tempo de entidades desnecessárias, como o éter. Eu argumentarei depois que a teoria que estou defendendo faz um bom trabalho com relação à simplicidade, sem contradizer qualquer experiência. Se alguém resistir a consequências contraintuitivas de teorias filosóficas, essa pessoa deve sustentar uma assimetria entre a ciência e a filosofia. Mas há uma dificuldade notória em definir filosofia, definir ciência e estabelecer fronteiras claras entre as duas: discussões sobre a natureza do tempo parecem envolver noções e argumentos que são normalmente considerados científicos e noções e argumentos que são normalmente considerados filosóficos; o verificacionismo não funciona como uma demarcação para a ciência já que tornaria leis físicas universais não científicas; e a teoria das cordas, que muitos consideram ser uma teoria científica, não parece ser falsificável. Muitos filósofos até mesmo desistiram de encontrar uma demarcação entre ciência e filosofia, adotando a sugestão de Quine de encarar a filosofia apenas como “ciência natural treinada sobre si mesma e com o uso livre permitido de descobertas científicas” (Quine, 1981, p. 125, trad. livre) ou focando, como Laudan, no que torna crenças bem fundamentadas, ao invés de científicas (Laudan, 1983, p. 125). Sem uma boa defesa dessa demarcação e suposta assimetria entre os dois campos, um afastamento do senso comum em teorias filosóficas não é sempre uma razão para negar tais teorias sem levar em consideração outros *desiderata* teóricos.

## 2.4 Redundância causal

Outra consideração que pode ser levantada em favor de eliminar a maioria dos objetos do senso comum é o fato de que eles parecem ser causalmente redundantes. Merricks nos dá o exemplo da “água na piscina” (Merricks, 2001, p. 30-32). Alguns filósofos diriam que ela é um objeto distinto de todos os simples na piscina. Algumas motivações



para isso (para além de adotar o universalismo mereológico clássico) incluem: “a água é o que impede o mergulhador de se aleijar quando ele pula”, “nós podemos ver a água” e “a água é o que determina o valor de verdade de frases como ‘a água na piscina hoje não estava na piscina ano passado’”.

A resposta oferecida por Merricks nesse caso, que eu acredito ser adequada, é que tudo isso pode ser feito pelos simples na piscina.<sup>2</sup> Eles impedem o mergulhador de se aleijar; eles causam a experiência visual que descrevemos como “ver a água”; e eles não serem os mesmos simples que estavam na piscina há um ano torna aquela frase verdadeira.<sup>3</sup>

Já que a água na piscina não parece necessária para explicar coisas que normalmente diríamos que dependem dela, passa-se a imaginar por que deveríamos postular a sua existência. E, é claro, um ponto similar pode ser levantado sobre muitos outros objetos cuja existência nós admitiríamos mais prontamente que “a água na piscina”. Por exemplo: se alguns simples onde normalmente diríamos que há uma estátua pudessem explicar coisas como “o peso que a estátua exerce sobre o pedestal”, “a experiência visual da estátua”, “as emoções causadas por ver a estátua” e a verdade de “aquela estátua é do formato de um ser humano”, talvez postular a estátua seja simplesmente um aumento desnecessário da nossa ontologia.

Merricks também desenvolveu um esquema geral para argumentos em favor de eliminar objetos macroscópicos inanimados.<sup>4</sup> Ele é o seguinte (Merricks, 2001, p. 79-80):

---

<sup>2</sup> Mais estritamente, ele se refere a moléculas de H<sub>2</sub>O, mas eu me refiro a simples aqui porque não defenderei a existência de moléculas de H<sub>2</sub>O.

<sup>3</sup> Ou, se quisermos ser muito escrupulosos, poderíamos dizer que torna a frase “correta”, para usar a terminologia de Sider (2013, p. 249), já que estritamente falando não há água lá, mas agir como se houvesse é bom o suficiente para questões práticas.

<sup>4</sup> Ele restringe o seu argumento deste modo para defender que seres humanos existem por não serem causalmente redundantes. Infelizmente, rebater esses argumentos tomaria demasiado espaço deste trabalho.

- 1) O objeto O — se O existe — é causalmente irrelevante quanto a se suas partes P1 ... Pn, agindo em conjunto, causam o efeito E.
  - 2) P1 ... Pn causam E.
  - 3) E não é sobredeterminado.
- Portanto,
- 4) Se O existe, O não causa E.

Vamos tentar esclarecer isso com o exemplo de uma pedra. Suponha que uma pedra é jogada e quebra um vaso. A premissa um diz que, se a pedra existe, ela é causalmente irrelevante<sup>5</sup> quanto a se suas partes, agindo em conjunto, causam a quebra do vaso. Isso parece razoável, já que um monte de partículas subatômicas arranjadas em formato de pedra de fato quebraria um vaso se jogado, mesmo que não compusesse nada. A premissa dois simplesmente afirma que aquelas partículas subatômicas quebram o vaso, e a premissa três afirma que a quebra do vaso não é sobredeterminada (isto é, não é o caso que a quebra é causada por um objeto ou objetos, esse(s) objeto(s) é(são) causalmente irrelevante(s) quanto a se outro(s) objeto(s) causa(m) a quebra e o(s) outro(s) objeto(s) de fato causa(m) a quebra). Disso tudo podemos concluir que, se a pedra existe, ela não causa a quebra do vaso.

Essa conclusão, junto do princípio de que não deveríamos acreditar na existência de objetos macroscópicos que não exerçam quaisquer poderes causais, motiva o eliminativismo com relação à maioria dos objetos do senso comum. As principais linhas de crítica contra este argumento consistem em questionar esse princípio e a premissa três.

Alguns filósofos talvez neguem a premissa três com base na alegação de que sobredeterminação causal entre um objeto e suas partes não é objetável. Eu concordo com Merricks que essa crítica pode ser

---

<sup>5</sup> Merricks apresenta quatro condições, que eu também adoto aqui, para um objeto A ser “causalmente irrelevante” quanto a se um objeto ou grupo de objetos B causa um evento: A não é um membro de B; não é o caso que B causa o evento somente quando combinado com a contribuição causal adicional de A; A não é uma causa intermediária em uma cadeia causal entre B e o evento; e A não causa que B cause o evento (Merricks, 2001, p. 57).

respondida por meio de duas considerações: nós sempre temos uma razão para resistir à sobredeterminação<sup>6</sup> e não é o caso que ser um objeto composto implica, por sua própria natureza, ser causalmente redundante (Merricks, 2001, p. 70) (já que alguns filósofos poderiam querer alegar que, porque ser um objeto composto é por sua própria natureza ser causalmente redundante, objetar à sobredeterminação causal seria objetar a o que é (em parte) ser um objeto composto, e composição enquanto tal não é objetável).

Merricks defende essa última consideração ao argumentar que há certos objetos compostos com poderes causais não redundantes, como seres humanos. Eu não acredito que serem humanos possuam poderes causais para além do que já é determinado pelos simples que os constituem; mas eu acredito que objetos compostos geram consciências,<sup>7</sup> algo que os seus simples constituintes não fazem, o que vai contra a ideia de que é da própria natureza dos objetos compostos ser causalmente redundante.

O esquema sendo válido, parece que muitos supostos objetos não teriam quaisquer poderes causais, isto é, seriam epifenômenos. Isso é suficiente para que neguemos a sua existência? É se supusermos que não é razoável postular a existência de objetos macroscópicos tais que não possuam quaisquer poderes causais. Merricks diz que esses objetos são exatamente o tipo para o qual esse requisito causal para a existência parece ser verdadeiro (Merricks, 2001, p. 81). Eu não sei se esse requisito causal para objetos macroscópicos é tão intuitivo, mas eu também negaria tais objetos epifenomênicos simplesmente com base em que uma teoria que não os postule e explique os fenômenos igualmente bem deve ser preferida. O que nos leva à próxima consideração em favor de eliminar objetos do senso comum.

---

<sup>6</sup> O que me motiva a evitar sobredeterminação em particular é o valor atribuído à simplicidade em teorias, que será abordado na próxima seção.

<sup>7</sup> Isso será argumentado na segunda parte deste trabalho.

## 2.5 Simplicidade

Outra coisa que motiva o eliminativismo é a ideia de simplicidade. Eu primeiro apresentarei a defesa de Sider do niilismo mereológico baseada em “parcimônia ideológica” e, posteriormente, defenderei a minha visão, que difere daquela ao focar em simplicidade com relação a leis.

De acordo com a terminologia de Quine, uma ontologia é os objetos que uma teoria postula, enquanto a sua ontologia é as noções não definidas que ela emprega, lógicas ou extra-lógicas (Quine, 1951). Já que “parte” não é um objeto, mas sim uma noção usada em algumas teorias, contaria como ideologia. Se adotarmos o niilismo mereológico, nossa teoria não precisaria fazer referência a “partes”, o que a tornaria ideologicamente mais simples.

Isso deveria motivar-nos a adotar o niilismo mereológico, diz Sider, devido ao princípio de que teorias ideologicamente mais simples são mais provavelmente verdadeiras (Sider, 2013, p. 238-243). Isso, ele aponta, é válido somente com relação a “teorias fundamentais”, tais como, na sua visão, física, matemática e metafísica fundamental. Não temos razão, portanto, para buscar simplicidade ideológica em nossas teorias biológicas ou sociológicas, assim como em nossos negócios do dia a dia. Então o argumento de que o niilista é obrigado a complexificar sua ideologia com noções como “arranjados em formato de estátua” (Bennett, 2009, p. 64) não funcionaria, já que estátuas não são parte da nossa descrição da realidade fundamental, que é com o que o niilista mereológico lida.

É claro, simplicidade ideológica não é a única virtude à qual uma teoria deve visar. Mas, quando duas teorias têm o mesmo poder preditivo e se adéquam igualmente bem às nossas evidências, deveríamos escolher a com menos ideologia. E esse parece ser o caso, à primeira vista, do niilismo mereológico, já que as predições que a física faz com relação a partículas subatômicas não parecem demandar referência à noção de “parte”.

Mas eu não estou defendendo o niilismo mereológico. Acredito que ele seja falso. Portanto, minha posição não seria capaz de simplificar a ideologia ao remover a noção de “parte”. Mas o que motiva esse desejo por teorias ideologicamente mais simples é um princípio mais geral: simplicidade é um guia para a verdade (creio que esse princípio possa ser defendido com base em argumentos epistemológicos, mas isso nos afastaria demais do assunto deste trabalho, então eu o tomo como certo aqui). E simplicidade pode ser obtida de mais de uma maneira. O próprio Sider comenta que outro tipo de simplicidade desejável é simplicidade com respeito a leis (Sider, 2013, p. 241). Minha tese faz um bom trabalho nesse quesito, já que mantém as leis da composição o mais simples que podem ser, pressupondo que o niilismo mereológico é falso e há tal coisa como leis da composição: há apenas uma situação em que a composição ocorre, e ela ocorre sempre que essa situação é o caso. Podemos dizer que, intuitivamente, há apenas uma lei:<sup>8</sup> “objetos compõem algo se e somente se eles geram *qualia* juntos”.

### 3 Por que admitir alguns objetos compostos?

Vamos agora abordar a razão pela qual eu admito um caso de composição. Isso torna a minha teoria da realidade fundamental, como foi apontado, menos simples que o niilismo mereológico. Eu faço isso, porém, porque não consigo entender uma negação deste caso específico de composição. Eu consigo entender a posição segundo a qual estátuas não existem, apenas simples arranjos em formato de estátua, e acho que essa posição está em harmonia com nossas observações empíricas e demanda menos complexidade das leis da natureza que a sua rival, segundo a qual há leis sobre as situações em que estátuas passam

---

<sup>8</sup> Embora, em outro sentido, as coisas não sejam tão simples, já que diferentes arranjos de simples parecem gerar diferentes tipos de *qualia*. Mas abordarei essa questão mais tarde.

a existir. Mas, quando se trata de *qualia*, eu acho que a crença na composição é inevitável, contanto que admitamos algumas premissas básicas.

Irei, portanto, listar as premissas que o meu argumento pressupõe e explicar o meu raciocínio quanto a como elas nos comprometem com um certo tipo de composição. Elas são:

1 – Eu experiencio *qualia*

2 – Há consciências para além da minha, também experienciando *qualia*

3 – Há um mundo físico, externo à minha consciência

4 – *Qualia* são geradas por objetos físicos

Sou cético quanto à possibilidade de bons argumentos poderem ser dados para essas premissas. A premissa número um me parece autoevidente, se qualquer coisa o é, e não consegui até agora encontrar muito sentido na posição segundo a qual *qualia* não existem ou são em algum sentido “uma ilusão”. Quanto às outras premissas, eu acredito nelas e não acho que tenha bons argumentos a seu favor, mas consigo pelo menos entender a posição que as nega.

Vamos agora ao raciocínio. Das premissas um e dois obtemos uma certa imagem das consciências que existem. Eu, a qualquer momento, experiencio uma gama diversa de *qualia*: muitas cores, sons, cheiros, sensações de toque etc. Há, contudo, muitas *qualia* que eu não experiencio, mas que são experienciadas por outra consciência; outro “ponto de vista de *qualia*”, por assim dizer.

Vamos tentar esclarecer isso com um exemplo. Eu estou em uma sala em frente a um amigo. Entre nós há uma bola que é vermelha de um lado e verde do outro. Eu só posso ver o lado vermelho, enquanto que ele só pode ver o lado verde. Além disso, eu estou com fones de ouvido e ouço uma boa música clássica, ao passo que ele está com os seus e escuta um *podcast*. Neste caso é intuitivo dizer, creio eu, que *qualia* vermelhas e *qualia* de som de violino são experienciadas juntas pela mesma consciência, o mesmo ponto de vista, que não experiencia as

igualmente existentes *qualia* verdes e *qualia* de voz de anfitrião de *podcast*. E a recíproca, é claro, também é verdadeira: há outra consciência que experiencia as *qualia* verdes e de voz de anfitrião de *podcast*, sem experienciar as *qualia* vermelhas e de som de violino.

Cada uma dessas consciências tem muitas *qualia* diferentes como elementos de um certo episódio de consciência; diferentes *qualia* como “partes” do episódio, formando a totalidade que ele é, alguém poderia se sentir inclinado a dizer. Uma coisa (um episódio de consciência) formada por muitas coisas diferentes (*qualia*): isso é composição? Bem, para dar uma resposta perfeita a essa questão, teríamos de saber o que é composição, respondendo a o que van Inwagen denominou a “questão geral da composição” sem usar quaisquer termos mereológicos. Sou cético quanto à ideia de que algum dia encontraremos uma boa resposta a essa questão. Mas há um modo pelo qual a noção de composição foi caracterizada que dá crédito, creio eu, à ideia de que isso poderia ser composição: van Inwagen se referiu a composição como “unidade surgindo da pluralidade” (Van Inwagen, 1990, p. 31, trad. livre). Isso certamente é um pouco vago, mas acredito que seja uma explicação razoável do que é composição (considerando que não temos explicações muito melhores). E aqui parecemos ter uma unidade (o episódio de consciência) surgindo de uma pluralidade (de *qualia*).

Mas, deixando de lado a questão sobre se um episódio de consciência particular é um tipo de composto, e quanto à composição de objetos físicos? As outras premissas que eu apresentei nos dizem que *qualia* são geradas por objetos físicos. E nós vimos que há muitas consciências diferentes, experienciando diferentes tipos de *qualia*. Parece natural, até mesmo forçoso, sustentar que diferentes consciências são geradas por diferentes conjuntos de objetos físicos (é difícil imaginar o que explicaria o mesmo conjunto gerar, por exemplo, duas consciências completamente diferentes).

Se consciências são geradas por grupos de objetos físicos, deve haver algum tipo de relação real entre esses objetos que explique por

que eles geram uma consciência. Imaginemos dois cérebros, um ao lado do outro, cada um gerando uma consciência. Deve haver uma relação que os simples de cada cérebro mantêm entre si, mas não com os simples do outro cérebro; do contrário, nenhum fato poderia explicar por que há *duas* consciências lá, cada uma gerada por um cérebro. Bem, os simples de ambos os cérebros interagem gravitacionalmente; eles também podem, quando muito próximos uns dos outros, interagir eletromagneticamente. O único fato que poderia explicar a unidade de um episódio de consciência, até onde eu consigo ver, é os simples de um cérebro também formarem uma unidade a partir da pluralidade devido ao seu arranjo; em outras palavras, compõem aquele objeto complexo que é responsável por uma dada consciência.

Alguém pode ainda se sentir um pouco desconfortável com isso. Pode ser pensado que, embora existam consciências experienciando *qualia*, essas talvez não existam devido a unidades físicas surgindo de pluralidades, mas meramente devido a essas pluralidades em si, não sendo tomadas como uma coisa. Essa é, acredito, a posição de Sider. Ao tentar encontrar o raciocínio que leva van Inwagen a defender sua própria existência enquanto um objeto composto por meio da natureza da mentalidade, Sider escreve o seguinte (Sider, 2013, p. 269, trad. livre):

Talvez a crença de van Inwagen em uma singularidade metafísica tenha algo a ver com o caráter da experiência consciente? As experiências simultâneas de um sujeito são experienciadas por esse sujeito como sendo em algum sentido parte de um episódio consciente, e como experienciadas por um único sujeito. Mas não é claro por que esses aspectos da fenomenologia não poderiam ser devidos, metafisicamente, a estados de partículas.

A isso eu preciso responder: não me é claro como esses aspectos da fenomenologia poderiam ser devidos, metafisicamente, a estados de partículas, sem que isso implique composição. Vamos recapitular: alguns objetos físicos geram uma consciência que experiencia certas



*qualia*; há outras *qualia* que essa primeira consciência não experiencia, mas que são experienciadas por outra consciência, gerada por outros objetos físicos. Portanto, a natureza fundamental da realidade (para usar uma expressão que soa um pouco misteriosa, mas que com sorte é suficientemente intuitiva) reconhece alguns objetos físicos como “aqueles que geram a consciência que experiencia as *qualia* v, w e z” (vamos chamá-los de os xs) e outros objetos como “aqueles que geram a consciência que experiencia as *qualia* p, q e r” (vamos chamá-los de os ys).

Se os xs não fossem tomados pela realidade como sendo uma coisa, uma totalidade, isso significaria que eles não são fundamentalmente mais conectados entre si do que são com relação aos ys. Mas, se eles não fossem fundamentalmente mais conectados entre si de um modo que não são com relação aos ys, o que explicaria as *qualia* geradas pelos ys (p, q e r) não serem experienciadas junto com as *qualia* geradas pelos xs (v, w e z)? Se uma conexão fundamental dessas não existisse, então ou todas essas *qualia* seriam experienciadas juntas por uma consciência, ou haveria uma consciência diferente para cada *quale* individual.

É um pouco difícil antecipar objeções quando se trata de questões tão abstratas. Acho que um crítico poderia argumentar, de maneira humeana, que não há razão para acreditar que as nossas percepções têm algum tipo de unidade e que elas são apenas um fluxo. Não tenho problemas com isso quando se trata de diferentes episódios de consciência ao longo do tempo; na verdade, eu tendo a concordar com essa visão, embora não haja espaço aqui para defender isso. Mas, quando se trata de um episódio de consciência particular em um momento particular, não consigo compreender alguém que diz que as suas *qualia* não formam algum tipo de totalidade e não têm qualquer relação especial umas com as outras que não têm com *qualia* do que normalmente chamaríamos de “outros episódios de consciência”.

Não consigo compreender tal ideia porque parece óbvio para mim que as *qualia* que eu experiencio enquanto digito e escuto música são sentidas juntas de uma maneira que não inclui as *qualia* que outra pessoa experiencia enquanto faz outra coisa. Infelizmente, quando a discussão chega ao assunto de experiências subjetivas, argumentos tornam-se difíceis e parecemos depender um pouco de intuições. Pode-se observar toda a controvérsia com relação à inteligibilidade de zumbis filosóficos como um exemplo.<sup>9</sup> Posso apenas esperar que, ao dar exemplos como os acima, o leitor concorde comigo que um episódio de consciência particular forma um tipo de totalidade, que inclui muitas *qualia* diferentes.

“Mas você estava disposto antes a negar a existência de objetos físicos compostos sem poderes causais! Por que não negaria também a existência de *qualia*, já que disse que as considera epifenômenos?” A existência de *qualia* epifenomênicas é uma premissa na argumentação acima. Mas, se eu tivesse que defendê-la, diria que há uma assimetria entre *qualia* e objetos físicos. Assim como Quine afirma, acredito que a existência de objetos físicos seja postulada para ajudar a explicar nossas experiências conscientes (Quine, 1948, p. 36); essas experiências em si são exatamente o que deve ser explicado e não podem, consequentemente, ser negadas.

Se esta argumentação for sólida, todos os filósofos que aceitarem as premissas que listei acima teriam de reconhecer esse caso de composição. Esse sendo o caso, mesmo se as leis da composição que explicassem as diferentes combinações de *qualia* experienciadas em um único episódio de consciência fossem muito complexas, desde que esses filósofos admitissem a existência de quaisquer objetos para além

---

<sup>9</sup> Essa dificuldade com relação a discussões sobre experiências subjetivas pode ser encontrada até mesmo em Hume, que diz que, se alguém ainda alegar ter a experiência de um “self” simples e contínuo após reflexão séria, ele não pode raciocinar com essa pessoa e talvez ela seja realmente diferente dele quanto a isso (Hume, 1960, p. 252 (1.4.6)).

dos que eu admiti aqui, as suas teorias sobre as leis da composição seriam mais complexas do que a teoria apresentada aqui. Um compromisso com a simplicidade maior que o desejo de defender objetos do senso comum os compeliaria a adotar a teoria deste trabalho.

## 4 Considerações finais

Embora esta resposta à questão especial da composição vá contra as nossas crenças do senso comum, deve ser notado que ela possui muitas vantagens quando comparada com outras respostas restritas. A maioria dessas respostas tentam acomodar objetos cuja existência é reconhecida pelo senso comum, o que leva a problemas como a colocação de uma estátua e o amontoado de argila que a constitui. A resposta que eu apresentei não reconhece tais objetos; apenas aqueles objetos que geram *qualia*. Esses últimos não compartilham espaço com nenhum outro suposto objeto cuja existência alguns poderiam querer defender, então o problema da colocação é completamente evitado.

Ela também evita, assim como o niilismo e o universalismo me-reológicos, o problema da vagueza. Isso porque ela nega a existência da maioria dos objetos do senso comum e, quanto se trata da questão de em que situações o arranjo de simples gera *qualia*, não parecemos ter muitas convicções intuitivas prévias; a maioria das pessoas não tem certeza absoluta quanto a se insetos experienciam *qualia*, por exemplo. Não só isso; pode não ser tão contraintuitivo que a presença de um único simples, ou o seu rearranjo, pudesse fazer toda a diferença quanto a se o conjunto de simples compõe algo e um episódio de consciência é gerado. Isso porque consciências parecem ser o tipo de coisa que ou está lá, ou não está. Certamente diferentes episódios de consciência podem variar quanto à sua complexidade, mas ou há *qualia* sendo experienciadas por uma consciência, ou não há. O caráter “tudo ou nada” da consciência foi destacado por Hawthorne ao defender a composição restrita do problema da vagueza contra Sider, embora ele não

escreva muito em defesa da resposta à questão especial da composição apresentada neste trabalho, apenas trazendo-a como uma possibilidade que enfraqueceria o argumento da vagueza de Sider (Hawthorne, 2006, p. 107-109).

A resposta à questão especial da composição apresentada aqui deve ser levada em consideração. Ela tem razões positivas a seu favor, como ter as leis mais simples possíveis enquanto ainda é capaz de reconhecer algum fato físico que explique a unidade da experiência consciente. Ela evita, ao mesmo tempo, problemas que tipicamente assolam outras respostas restritas. Certamente vai contra nossas crenças do senso comum, que poderiam ser tomadas por alguns filósofos como uma razão para rejeitá-la, mas, contanto que reconheçamos que o senso comum humano em 2.025 D.C. não é um guia infalível para a verdadeira natureza da realidade, esta resposta se apresenta como uma possibilidade séria.

## Referências

- BENNETT, Karen. Composition, Colocation and Metaontology. In: *Metametaphysics: New Essays on the Foundations of Ontology*. CHALMERS, David; MANLEY, David; WASSERMAN, Ryan (Eds.). Oxford: Clarendon Press, 2009.
- HAWTHORNE, John. Three-Dimensionalism. In: *Metaphysical Essays*. Oxford: Clarendon Press, 2006.
- HUME, David. *A Treatise of Human Nature*. Ed. L. A. Selby-Bigge. Oxford: Clarendon Press, 1960.
- JACKSON, Frank. *Epiphenomenal Qualia*. The Philosophical Quarterly, vol. 32, no. 127. Hoboken, New Jersey: Wiley, 1982.
- LAUDAN, Laurens Lynn. The Demise of the Demarcation Problem. In: *Physics, Philosophy and Psychoanalysis*. Eds. Robert Sonné Cohen and Laurens Lynn Laudan. Dordrecht, Holland: D. Reidel Publishing Company, 1983.
- LEWIS, David Kellogg. *On the Plurality of Worlds*. Oxford: Basil Blackwell Ltd., 1986.

- MERRICKS, Trenton. *Objects and Persons*. Oxford: Clarendon Press, 2001.
- QUINE, Willard van Orman. Ontology and Ideology. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, v. 2, n. 1. Berlim: Springer, 1951.
- QUINE, Willard van Orman. On What There Is. *The Review of Metaphysics*, v. 2, n. 5. Washington, D.C.: Philosophy Education Society Inc., 1948.
- QUINE, Willard van Orman. *Theories and Things*. Cambridge, Massachusetts: Harvard University Press, 1981.
- SIDER, Theodore. Against Parthood. In: *Oxford Studies in Metaphysics*, v. 8. BENNETT, Karen; ZIMMERMAN, Dean W. (Eds.). Oxford: Oxford University Press, 2013.
- VAN INWAGEN, Peter. *Material Beings*. Ithaca: Cornell University Press, 1990.



# O tratamento supervaloracionista da vagueza existencial e seus desafios

Felipe da Silva Avena<sup>1</sup>

DOI: <https://doi.org/10.58942/eqs.158.03>

## 1 Introdução

Encontramos uma relação íntima entre existência, vagueza, identidade e composição muito bem ilustradas em dois experimentos mentais bastante conhecidos, o paradoxo de sorites e o problema dos muitos. Na tentativa de responder aos paradoxos envolvendo existência vaga que são suscitados, foram propostas muitas abordagens de tratamento semântico da noção de “vagueza”. A mais bem aceita foi denominada supervaloracionismo.

Baseado no tratamento supervaloracionista da vagueza, Sider (Sider, 2003) elabora um argumento detalhado, o qual tem como objetivo provar a inconsistência de vagueza existencial. Partindo deste resultado, seria possível extrair conclusões de ordem mereológica ou quanto à natureza da existência por *modus tollens* (se  $\langle \text{não há existência vaga} \rangle$  e  $\langle P \text{ implica em existência vaga} \rangle$ , então  $\langle \text{não } P \rangle$ ). Se uma tese mereológica ou ontológica mais geral resulta em vagueza existencial, por negação do consequente como absurdo negamos também o

---

<sup>1</sup> Bacharel em Filosofia pela Universidade Federal do Rio de Janeiro. Mestrando no Programa de Pós-Graduação Lógica e Metafísica da mesma universidade. Interessado em Filosofia Analítica, Metafísica, Ontologia, Mereologia e Lógica.

E-mail: [felipe.s.avena@gmail.com](mailto:felipe.s.avena@gmail.com)

antecedente. Torza (2017), por sua vez, levanta questões contundentes contra a validade do argumento de Sider, buscando resgatar a possibilidade de um tipo de vagueza existencial, a qual é denominada existência super-vaga. De forma colateral, termina por resgatar aquelas teorias ameaçadas pela impossibilidade de vagueza existencial. Na sequência, Loss (2018) concorda com a validade do modelo proposto por Torza de uma super-metalinguagem. Tal modelo seria capaz de regimentar de modo coerente o operador de existência super-vago. Entretanto, Loss adverte que tal modelo tem como pressuposto a vagueza ôntica, o que é negado como ponto de partida por Sider e de outros filósofos signatários da mesma opinião, como Lewis.

Assumida a possibilidade de vagueza ôntica, são suscitados desafios profundos e de caráter metafilosófico. As relações entre vagueza, existência, identidade e composição estão longe de serem facilmente identificadas e completamente escrutinadas. Um exemplo disso pode ser visto no caso do argumento de Evans (1978), que serviu de motivo para desvelar uma relação íntima, mas nada evidente, entre identidade vaga e existência vaga. Além disso, seria aceitável, ou mesmo consistente, considerar que há distintos tipos de existência? ou ainda, graus de existência? Hawley (2002) argumenta em favor de que aceitar um tipo de existência mais fraco não implica em um comprometimento imediato com vagueza de identidade e nem mesmo com a aceitação de entes não existentes. Hershenov (2014), por sua vez, argumenta em sentido contrário, defendendo que vagueza de existência implicaria em identidade vagas, acarretando compromissos com algumas teses mereológicas (aquelas mesmas que seriam fortalecidas pelos argumentos de Sider quando aplicado ao *modus tollens*). De modo geral, o que se pode obter no debate destas distintas posições é que a aceitação de entidades compostas envolvendo processos graduais efetivamente implicaria em vagueza ôntica e identidade vaga.



## 2 Paradoxo de sorites: problemas entre persistência, composição e vagueza de existência

O paradoxo de sorites<sup>2</sup> é bastante antigo, que recebeu diversas formulações e tentativas de resposta (Hyde; Raffman, 2018). Sendo que todas as abordagens terminam por levantar outras questões filosóficas profundas envolvendo vagueza.

De modo mais geral, o paradoxo de sorites é formado pela aplicação de termos vagos ou difusos levando a conclusões absurdas e contraditórias, inspirando ou a tese de que predicados vagos implicam em absurdos ou que deva haver algo de errado na formulação do paradoxo. A forma mais comum do paradoxo pode ser expressa da seguinte maneira:

(1) Um grão de areia não caracteriza um montinho

(2) Se um grão não caracteriza um montinho, então dois grãos também não caracterizam.

(3) Se dois grãos não caracterizam um montinho, então três também não caracterizam.

....

(999999) Se 999999 grãos não caracterizam um montinho, então 1 milhão também não caracterizam.

Logo:

1 milhão de grãos não caracterizam um montinho.

Enquanto aceitarmos as premissas e a validade do argumento, linha a linha, devemos aceitar a conclusão absurda. De outro modo temos que aceitar que

Há um número  $n$  qualquer em que caracteriza um montinho, e que  $n-1$  já não mais caracterizaria um montinho.

---

<sup>2</sup> O nome “sorites” vem do termo grego “soros” que significa “pilha” ou “monte”. O paradoxo foi proposto primeiramente por Eubulides no quarto século AC e posteriormente usado pelos cétricos contra os estoicos, só foi resgatado por neo-hegelianos marxistas e ao fim das doutrinas em favor das linguagens ideais.

Mas isso também parece absurdo, já que não parece ser plausível nem justificável supor que haja um número determinado de grãos de areia em que um mero coletivo de grãos passa a ser, determinada-mente, um montinho de areia.

Desta maneira, o paradoxo se estabelece, uma vez que ficamos sem conseguir apontar um erro nos passos do argumento nem em suas premissas. Mas ao mesmo tempo não podemos aceitar sua conclusão. Além disso, não podemos aceitar uma forma de barrar o paradoxo, já que não podemos argumentar em favor de um limite precisamente determinado de aplicação do conceito “montinho”.

Podemos dividir em 4 grandes categorias de abordagens possíveis de resposta ao paradoxo de sorites: (1) Propondo uma linguagem ideal. De acordo com essa resposta as expressões soríticas são intratáveis pela lógica a tentativa de elaborar o problema pela lógica acabaria levando a absurdos. Outro tipo de resposta possível é a que aceita uma (2) Teoria epistêmica da vagueza. Esta teoria não considera a vagueza como um fenômeno propriamente ôntico, nem mesmo semântico. Aqueles que apostam em uma resposta epistêmica ao paradoxo de sorites consideram que a vagueza deve ser interpretada estritamente como um grau de ignorância quanto aos limites que não são conhecidos por nós. Para eles, ao menos algum passo do argumento deve ser rejeitado por ser falso (invalidando o paradoxo), ainda que seja impossível determinar qual é este passo. Um outro tipo de abordagem possível é (3) Assumir a validade do paradoxo de sorites. Essa abordagem não é exatamente uma resposta, pois se caracteriza por abrir mão de buscar tentar “responder” ou, mais precisamente, “desmontar” o paradoxo de sorites. Ela implica na conclusão radical de que o paradoxo é insolúvel. É assumir “de frente” uma radical incoerência no mundo. Pode-se dizer que por isso é a posição mais rejeitada. A quarta grande categoria de respostas ao paradoxo é aceitar a noção de (4) Vagueza semântica. Esta proposta aposta na implementação de lógicas não clássicas e suas teorias semânticas derivadas. Vale ressaltar que essa

proposta se baseia nas teorias mais aceitas que buscam caracterizar a vagueza. De acordo com essa abordagem, a caracterização correta da vagueza resultaria em uma invalidação da estrutura argumentativa do paradoxo. O presente artigo irá focar nesta categoria de respostas.

Nesta grande categoria de respostas que aceitam vagueza semântica podemos citar 4 abordagens distintas. Duas focam em aspectos extrasemânticos relacionados ao uso e aplicação de regras gramaticais, como as (1) Teorias contextualistas e outras familiares. Estas respostas são obrigadas a assumir uma semântica não padrão para quantificadores, ao apostar em uma mudança disposicional do uso de conceito de acordo com o contexto em que o conceito é aplicado. Ou ainda (2) A Teoria dos múltiplos intervalos, que é característica por manter lógica clássica e a bivalência, o que aumenta seu caráter atrativo. Tais intervalos propostos são nada mais do que abstração dos modos permissivos de uso de um predicado. O engano produzido pelo paradoxo de sorites seria justificado por uma confusão entre distintas regras pragmáticas que orientam o uso de palavras vagas.

As outras duas abordagens serão tratadas no presente artigo. Elas consideram estritamente questões semânticas, sem considerar aspectos epistêmicos ou contextualistas do uso da linguagem. Em primeiro lugar citamos a (3) A abordagem supervaloracionista. Isso inclui não apenas uma formulação do supervaloracionismo, mas também inúmeras outras teses aparentadas ao supervaloracionismo. A força da abordagem semântica supervaloracionista é ser capaz de assumir uma lógica não bivalente e, ao mesmo tempo, manter as leis da lógica clássica. Para isso lança mão de noções de super-verdade e super-falsidade, mantendo a lei do terceiro excluído, por exemplo. Essa característica importa prestígio por parte daqueles que prezam pela manutenção de tais leis da lógica clássica.

Ainda dentro da abordagem semântica, encontramos a proposta de solução do paradoxo de sorites pela abordagem de (4) teorias multivaloracionistas. Neste grupo podemos encontrar dois grandes

subgrupos: o subgrupo das teorias multivaloracionistas com número finito de valores de verdade<sup>3</sup> que as proposições podem apresentar e o subgrupo das teorias com número infinito de valores de verdade que as proposições podem apresentar, caracterizando a proposta “fuzzy” de resposta ao paradoxo de sorites. Tanto as teorias com número finito quanto infinito podem dar origem a inúmeras versões, dependendo da definição de preservação de verdade. Cada uma destas versões encontra problemas particulares.

É possível ver como estes possíveis meios de oferecer respostas ao paradoxo, em especial o supervaloracionismo e teorias multivaloracionistas como a lógica fuzzy, estão intimamente ligadas às questões envolvendo persistência. Podemos imaginar o mesmo argumento constituído de forma inversa, em que os grãos de areia, em vez de serem adicionados passo a passo poderiam ser retirados. Partindo da certeza que 1 milhão de grãos caracterizam um montinho de areia, quantos grãos de areia deveriam ser retirados para que esta pluralidade de grãos a qual denominamos “monte de areia” deixe de perdurar? Sabemos que apenas um grão não constitui um monte de areia, e que a retirada de 999999 grãos garante que o monte de areia não perdure. Também sabemos que um monte de areia não deixa de existir ao perder apenas um grão de areia. Mas nesse caso, há algum momento intermediário em que o monte de areia perdura “vagamente”? Ora, essa questão quanto a persistência é também uma questão quanto à composição, já que introduz a estrutura de argumento sorítica em questões diretamente relacionadas à composição do “monte de areia”. Faz sentido pensar que uma pluralidade de objetos compõem vagamente uma totalidade? O mesmo ocorre quanto a vagueza de existência. Se em um dado momento sabemos que um monte de areia definitivamente existe

---

<sup>3</sup> Por exemplo, uma lógica trivalente, tetravalente, pentavalente... n- valente. com n sendo sempre um número natural maior que 2.

e em outro não? Há um momento intermediário em que um monte de areia existe vagamente? Se sim, o que exatamente isso significa?

### **3 Problema dos muitos como argumento niilista: é melhor não existir a assumir dialetéias?**

Peter Unger elabora um argumento (Weatherson, 2023) com o intuito de fortalecer uma versão de niilismo mereológico. Ou seja, um argumento que fortalece a tese de que a própria noção de objetos compostos leva a algum tipo de contradição. À primeira vista parece um argumento relativamente simples, mas é muito mais difícil oferecer um tratamento para tal questão do que faz parecer, ameaçando teses de composição restritivistas<sup>4</sup>. No núcleo de seu argumento podemos identificar o que ficou conhecido como “o problema dos muitos”<sup>5</sup>. Esse argumento encontra sua força lógica na seguinte estrutura argumentativa em forma de uma *reductio* (Van Inwagen, 1995, p. 214-216):

Vamos supor que existe um homem M. Tal homem possui uma parte dispensável x (onde x é qualquer átomo mereológico dispensável na constituição de M.). Disso, podemos dizer que é tão adequado dizer que M-x é um homem quanto é adequado dizer que M é um homem. Ou seja, de qualquer átomo meteorológico a,b,c... podemos dizer que M-a, M-b, ou M-c é um homem quanto é adequado dizer que M é um homem. Seria completamente arbitrário selecionar, da enorme quantidade de átomos mereológicos que compõem um homem, alguns

---

<sup>4</sup> Teses de composição restritivistista são aquelas teses mereológicas que fornecem uma condição necessária e suficiente para que as partes venham a compor um todo. Por exemplo, o organicismo é uma tese restritivistista no sentido que fornece uma condição para que haja composição: um coletivo de objetos vêm a compor algo se e somente se se constitui um organismo. Pode-se dizer que teses de composição restritivististas são aquelas que não se encaixam nem no campo de teses niilistas nem no campo das teses universalistas.

<sup>5</sup> Esse problema pode ser encontrado em sua obra (Unger, 1980) mas também foi elaborado de maneira distinta em (Geach, 1997).

átomos que poderiam (ou que não poderiam) faltar na composição de um homem. Disso se seguiria que há uma quantidade enorme de homens distintos como M, M-a, M-b, M-c... dentro da mesma sala. Seríamos forçados a aceitar tal conclusão. Não sendo aceitável tal conclusão, somos obrigados a negar a existência de M.

Alguém poderia argumentar, dizendo que é ridículo defender a ideia de que M não existe, e que precisaríamos apenas de um “princípio de seleção” dentre tantos homens-candidatos distintos para decidir qual deles é o homem que realmente existe (vamos supor, o homem M-s). Mas não há nenhuma distinção significativa entre M e qualquer um dos M-x, e qualquer princípio que proponha distinguir cada um destes “homens-candidatos” de um supostamente mais “real” seria intoleravelmente arbitrário.

Uma vez que não é possível negar nenhum dos passos e das premissas de onde eles advêm, restam apenas duas soluções: negar a existência de M ou aceitar contradições verdadeiras, aplicando modelos semânticos que possam assumir vagueza como o supervaloracionismo ou o multivaloracionismo. Ora, ao utilizar esses modelos semânticos as mesmas questões que surgem no tratamento do paradoxo de sorites podem surgir. Retirando cada um dos átomos mereológicos, sob que circunstâncias o homem (ou a pluralidade de homens) persiste? Ele pode persistir de modo vago, ser composto vagamente ou ainda existir vagamente em conjunto com outros candidatos, apresentando uma espécie de identidade parcial e vaga? Se tudo isso é possível dentro de alguns modelos, que real significado possui quando interpretamos estes resultados sob a ótica de uma ontologia?

Como vimos anteriormente, a relação entre persistência, composição e existência estão intimamente amarradas à questão da vagueza. Vamos apresentar em seguida uma das formas mais bem aceitas de compreender o fenômeno da vagueza, o supervaloracionismo.

## 4 O tratamento supervaloracionista da vagueza

O tratamento supervaloracionista da vagueza foca no modo de determinar o valor da verdade das sentenças, além da dependência dos valores de verdade com o modo de precisificação. Se há uma sentença vaga, ela não possui valor de verdade se não for relacionada com uma dada precisificação. Entretanto, se uma sentença é verdadeira/falsa em *todas* as precisificações então podemos dizer que é um caso de super-verdade/falsidade. Desta forma, ao mesmo tempo que a ideia de um intervalo de valor de verdade é aceita, são conservados teoremas da lógica padrão. Por exemplo, não podemos determinar o valor de verdade da sentença “João é alto”  $A(j)$ , dado que a propriedade “ser alto” é vaga.

A noção de precisificação é central, e pode ser mais facilmente entendida através de um exemplo. Enquanto Ana se arruma para a festa da faculdade, recebe uma ligação de sua amiga Maria. Nesta ligação, recebe a notícia que um amigo de Maria estará presente na festa, um rapaz belo e alto. Chegando na festa, Ana se decepciona. De fato João é um rapaz bonito, mas no máximo tem uma estatura média. Tendo morado muitos anos em uma cidade com a altura média bastante elevada, Ana interpreta o predicado “ $x$  é alto”  $=_{\text{def}}$  “ $x$  têm uma altura  $\geq 1,80$ ”. Já Maria, vinda de uma cidade com uma estatura média mais baixa, interpreta o predicado “ $x$  é alto”  $=_{\text{def}}$  “ $x$  têm uma altura  $\geq 1,70$ ”. Vemos assim a precisificação como uma definição de um termo lançando mão de uma linguagem mais precisa. Maria não mentiu para Ana, apenas utilizou uma precisificação distinta de Ana. Em  $P_a$   $A(j)$  é falso, enquanto que em  $P_m$   $A(j)$  é verdadeiro. Podemos ver como a noção de super-verdade opera para manter a validade de teoremas da lógica clássica no seguinte exemplo: se pensarmos que em ambas as precisificações disponíveis  $P_a$  e  $P_m$ , a sentença  $A(j) \vee \neg A(j)$  é verdadeira. Ser

---

<sup>6</sup> Isto é, Precisão utilizada por Ana

verdadeira/falsa em todas as precisificações é por definição ser uma super-verdade/super-falsidade. É importante observar que as motivações que determinam as precisificações dependem de aspectos extralinguísticos como o uso aceitável de um termo por uma comunidade e um certo contextualismo linguístico. Por isso, podemos dizer que a determinação de uma precisificação carrega um certo grau de arbitrariedade. Por que não assumir que “x é alto” como =def “x têm uma altura  $\geq 1,79$ ”? Não parece plausível que Maria iria se opor à tal precisificação<sup>7</sup>.

Esta caracterização do supervaloracionismo mostra que seu tratamento opera estritamente sobre a questão das propriedades. Mas não está claro qual seria o tratamento para o caso da existência vaga.<sup>8</sup>

## 5 Sider: contra a existência vaga

Utilizando-se do supervaloracionismo, Sider investe em argumentos com o intuito de mostrar que a existência não pode ser vaga<sup>9</sup>. Ele apresenta duas grandes oposições aos defensores da vagueza existencial, que serão apresentadas a seguir. Sua argumentação inicia ao retomar uma crítica feita por Katherin Koslick (2003), de que o ponto de partida de que a existência não poderia ser vaga é pobremente sustentada. Concordando com tal crítica, Sider se propõe a apresentar

---

<sup>7</sup> É interessante observar que este tipo de questionamento dá origem a argumentos em defesa da tese de que a aceitação de termos vagos implicaria também na emergência de vagueza de segunda ordem e assim sucessivamente.

<sup>8</sup> Lembrando que esta questão não é nada trivial, uma vez que a existência, dentro da lógica de predicado de primeira ordem, não pode ser tratada como uma propriedade, já que é expressa (ou mais aproximadamente expressa) pelo uso de quantificador existencial “ $\exists$ ”, e dizer que “x existe” não pode ser expresso por “ $\exists(x)$ ”, que é uma fórmula mal formada.

<sup>9</sup> Sua motivação parte do fato de que a defesa de sua tese tetradimensionalista depende de tal impossibilidade.



argumentos em favor da tese fundamental de que a existência só pode ser precisamente determinada, nunca vaga:

(EP) Existência não pode ser tomada vagamente.

A maneira como o argumento tetradimensionalista se conecta com a negação de existência vaga é apresentado partindo da aceitação de que há problemas de persistência colocados em termos de fusão dos mínimos diacrônicos <sup>10</sup>. Assumir no interior de uma ontologia tais tipos de fusão poderia levar a uma resposta universalista e aceitação de objetos de existência instantânea, o que não é de forma alguma ponto pacífico. Além, vagueza quanto a ocorrência ou não de vagueza de mínimos diacrônicos implica na aceitação de um ou mais símbolos lógicos. Mas essa é uma questão secundária. Sider parte de um argumento de sorites capaz de concluir que “Nunca é vago se uma fusão de mínimos diacrônico existe.”<sup>11</sup> Ou seja, é necessário aceitar que haja um corte preciso na fusão de mínimos diacrônicos. Caso contrário, há uma indeterminação do número de objetos no domínio, caso seja feita uma formalização em lógica de primeira ordem de sentenças envolvendo o número de objetos. Koslick discorda, defendendo a vagueza do quantificador existencial, assumindo que quantificadores irrestritos podem sim ser tomados de maneira vaga.

Em vistas de defender a impossibilidade de vagueza existencial, são apresentados argumentos em favor de duas teses:

---

<sup>10</sup> Em resumo, tomando fusão como composição, Sider propõe a seguinte definição para mínimos diacrônicos: i) Se pensarmos em um conjunto de objetos  $S_1, S_2, \dots, S_n$  os quais ii) são referentes a cada momento  $T_1, T_2, \dots, T_n$  e iii)  $x_i$  é composto apenas por  $S_i$  no instante  $T_i$ , sendo assim definido como uma “fusão mínima diacrônica”.

<sup>11</sup>No original: “it is never vague whether a minimal diachronic fusion exists” (Sider, 2003, p. 30)

- (1) É insustentável defender existência vaga, já que isso implicaria em aceitar uma precisificação do quantificador “ $\exists$ ”, o que leva ou a contradição ou a um regresso ao infinito.<sup>12</sup>
- (2) A Existência irrestrita não é sujeita à precisificação, uma vez que é um tipo natural. Uma vez que a definição de vagueza pressupõe precisificação, é impossível que existência seja vaga.

A defesa da primeira tese toma como ponto de partida que casos paradigmáticos de vagueza podem ser originados a partir de múltiplas precisificações em uma linguagem de fundo precisa. Esta mesma característica de formação da vagueza acaba forçando que seja pressuposto os quantificadores como “bits”<sup>13</sup> fundamentais da linguagem, não podendo ser por isso precisificados. Dado que o supervaloracionismo interpreta a vagueza em termos de precisificação, quantificadores com aplicação irrestrita não poderiam ser vagos.

Restam assim apenas três opções ao defensor da vagueza existencial frente ao desafio de oferecer uma precisificação do quantificador existencial (Sider, 2003, p. 6-9): (1) Rejeitar a necessidade de precisificação não vaga (2) Usar quantificadores vagos para realizar precisificação não vaga (3) Oferecer uma linguagem que seja ao mesmo tempo não vaga e não quantificacional para a precisificação de quantificadores

A primeira opção termina por ser desinteressante e classificada por Sider como “quietista”, uma vez que aceitaria injustificadamente a vagueza de existência e ao mesmo tempo se comprometeria com a impossibilidade de sua completa inteligibilidade. A terceira opção não parece ser possível de ser levada à frente, ou ao menos ninguém parece

---

<sup>12</sup> O ponto que não é explicitamente dito por Sider, mas parece constituir o núcleo de seu argumento: a aceitação da vagueza na noção de existência em metalinguagem implicaria (exclusivamente) a alguma das duas possibilidades.

<sup>13</sup> Uso aqui o termo utilizado por Sider, mas poderia ser substituído pelo termo “unidades”.

ter sido capaz de apresentar um projeto de tal linguagem de maneira minimamente convincente.

A segunda opção é a mais promissora, ainda que à primeira vista pareça contraintuitiva. Entretanto há Impasses para aqueles que defendem a noção de existência vaga: (i) A inevitável presença de quantificadores existenciais vagos nas precisificações. Se aceitarmos a presença do conectivo vago “ $\exists$ ” em sua própria precisificação, então trata-se de uma circularidade. Se o conectivo fosse outro, digamos  $\exists^*$ , então ele necessitaria também de uma precisificação  $\exists^{**}$  e assim sucessivamente, produzindo uma cadeia de regresso ao infinito. Outro desafio é a (ii) Ausência de regras para eleição de uma precisificação em detrimento de outras. Desta forma a tese de uma vagueza de existência seria solapada por uma arbitrariedade impossível de ser eliminada.

O segundo argumento de Sider é bem sucinto, se baseando no fato de que, se o “ $\exists$ ” expressa um tipo natural, então ele não poderia ser vago, já que não pode possuir múltiplas precisificações. A estrutura do argumento segue da seguinte forma (Sider, 2003, p. 9): (1) Vagueza requer múltiplas precisificações. (2) Qualquer que seja o tipo natural, então ele não é sujeito à distintas precisificações. (3) Existência irrestrita é um tipo natural. Conclusão: “ $\exists$ ” não pode ser vago.

Focando na primeira oposição à possibilidade de existência vaga apresentada por Sider, Torza oferece uma resposta aos problemas com o tratamento de possíveis quantificadores existenciais vagos. Para isso, ele critica o argumento de Sider e propõe um modelo semântico especificamente feito para regimentar a noção de vagueza existencial, eliminando qualquer inconsistência.

## 6 Torza: contra “contra a existência vaga”

A crítica ao argumento de Sider poderia ser resumida da seguinte maneira: com o intuito de provar a impossibilidade de existência vaga, Sider defende, por redução ao absurdo, que a vagueza

existencial é inaceitável, já que eleva a contradições na metalinguagem. Mas qualquer tentativa de realizar essa prova exige uma outra prova em uma meta-metalingagem que seja absolutamente determinada, e assim consecutivamente. Isso torna o argumento em forma de *reductio* apresentado por Sider invalidado. Frente à tal conclusão, Torza se põe a apresentar um modelo capaz de renderizar um tipo de vagueza existencial.

O argumento de Sider consiste em uma defesa de que “se o quantificador  $\exists$  de uma linguagem-objeto é vago e absoluto [no sentido de possuir aplicação irrestrita], uma contradição é obtida da sua metalinguagem.”<sup>14</sup>

Tal argumento pode ser reconstruído da seguinte maneira (Torza, 2017, p. 7-8):

(1)  $\langle \exists x\varphi \rangle$  é vaga

Dada a interpretação da vagueza em termos de precisificação temos:

(2) Em ao menos uma precisificação  $\langle \exists x\varphi \rangle$  é verdadeiro e em ao menos outra  $\langle \exists x\varphi \rangle$  é falso.

Por separação da conjunção:

(3) Há ao menos uma precisificação em que  $\langle \exists x\varphi \rangle$  é verdadeiro

Dado que as condições de verdade de afirmações quantificadas são determinadas por extensionalidade temos que:

(4) Há ao menos alguma coisa em uma dada precisificação que pertence ao domínio  $\exists$  e também satisfaz  $\varphi$

Por separação da conjunção temos que:

(5) Há ao menos alguma coisa que, em uma dada precisificação, satisfaz  $\varphi$ .

---

<sup>14</sup> No original: “if an object-language quantifier  $\exists$  is vague and absolute, a contradiction can be derived in its metalanguage” (Torza, 2017, p. 6).

<sup>15</sup> Onde “ $\varphi$ ” representa uma sentença qualquer na linguagem objeto sobre a qual o argumento se detém.

De acordo com a interpretação precisificacional de vagueza, o fato de  $\varphi$  ser preciso significa que a interpretação em todas as precisificações coincide. Disso e de (5) obtemos:

(6) Há ao menos alguma coisa que, em todas as precisificações, satisfaz  $\varphi$

dado que  $\exists$  é irrestrito, podemos então afirmar que:

(7) Há ao menos alguma coisa que, em todas as precisificações, satisfaz  $\varphi$  e pertence ao domínio  $\exists$

Mas, considerando as condições de verdade de afirmações quantificadas temos:

(8) Em toda precisificação  $\langle \exists x \varphi \rangle$  é verdadeira

Esta afirmação no framework precisificacional significa que

(9)  $\langle \exists x \varphi \rangle$  é verdadeiro

Ainda na interpretação precisificacional da vagueza podemos dizer:

(10)  $\langle \exists x \varphi \rangle$  não é vaga

Por conjunção de (1) e (10) temos a seguinte conclusão:

(11)  $\langle \exists x \varphi \rangle$  é vaga e não é vaga

Terminando assim aprova por *reductio ad absurdum*.

Torza mostra que a validade deste argumento depende que (1) não seja vaga. Entretanto, a prova da metalingagem como não vaga exigiria prova da não vagueza da meta-metalinguagem, e assim sucessivamente em um regresso ao infinito. Isso pode ser visto ao tentar dar uma prova de que (1) é não vaga através de uma *reductio* tendo como linguagem objeto a metalinguagem do argumento dado acima.

A segunda *reductio* tem como objetivo garantir a não vagueza da metalinguagem necessária a uma prova referente a uma dada linguagem objeto. Vamos considerar que a precisificação-n se refere ao nível em que se encontra a linguagem objeto sobre a qual o argumento é aplicado. Por exemplo, para a meta-linguagem teríamos uma

precisificação-2. Para uma meta-metalinguagem teríamos uma precisificação-3, e assim sucessivamente.

Assim, é possível reproduzir de forma resumida a estrutura do argumento de Sider para a metalinguagem da seguinte maneira (Torza, 2017, p. 11-12):

$\langle\langle\exists x\varphi\rangle\text{é vaga}\rangle\text{é vaga}$

Em ao menos uma precisificação-2  $\langle\langle\exists x\varphi\rangle\text{é vaga}\rangle$  é verdadeiro e em ao menos outra  $\langle\langle\exists x\varphi\rangle\text{é vaga}\rangle$  é falso.

Por separação da conjunção:

Há ao menos uma precisificação-2 em que  $\langle\langle\exists x\varphi\rangle\text{é vaga}\rangle$  é verdadeiro

do que precisificação-2 e clássica

Em todas as precisificações-2  $\langle\langle\exists x\varphi\rangle\text{é vaga}\rangle$  é falsa

O sendo assim é obtida uma contradição por redução ao absurdo:

$\langle\langle\exists x\varphi\rangle\text{é vaga}\rangle$  e não é vaga

Terminando assim aprova por *reductio ad absurdum* em segunda ordem.

Ora, sobre a mesma estrutura podemos lançar a mesma crítica: É necessário garantir que a meta-metalinguagem seja perfeitamente precisa para que a *reductio* aplicada em metalinguagem tenha validade garantida. Mas isso só poderia ser dado garantindo a perfeita precisão de uma meta-metalinguagem. Fica claro que o uso desta estrutura argumentativa leva a uma iteração cíclica, resultando em um regresso ao infinito.

Torza apresenta um modelo teórico para uso do quantificador existencial vago. Para isso ele toma como ponto de partida a formalização do supervaloracionismo proposta por Fine, o modelo precisificacional mais popular. Tal modelo é capaz de descrever um comportamento mais geral e desejado da vagueza, como a invalidade de  $\Delta p \vee \Delta \neg p$ , ainda que sejam válidas tautologias como  $p \vee \neg p$ . Entretanto,

este mesmo sistema popular precisificacional exclui a possibilidade de existência vaga. Desta maneira, é necessário superar essa limitação, propondo todo um quadro de domínio variado. Diferente do modelo de Fine e da formalização pressuposta por Sider, o modelo de Williamson admite múltiplos conjuntos de precisificação, permitindo vagueza e ordens superiores e atendendo à duas propriedades (1) determinância-de-Sider: Para qualquer  $n$ , definitivamente não é o caso que a existência é vaga na  $n$ -ésima ordem.<sup>16</sup> (2) Anti-determinância-de-Sider: Para qualquer  $n$ , não é definitivamente o caso que existência não é vaga na  $n$ -ésima ordem<sup>17</sup>

Em conjunto, as duas condições apresentadas caracterizam o que Torza denomina uma existência super-vaga. Mas isso ainda não é o suficiente, já que, mesmo na semântica de Williamson, há um domínio constante de aplicação do quantificador, eliminando a possibilidade de vagueza existencial. Para uma generalização completa, seria preciso formular uma semântica de Williamson capaz de oferecer um quadro para domínios variados e permitindo vagueza ontológica. Que é exatamente o que Torza desenvolve em seu trabalho.

## **7 Loss: contra “contra ‘contra a existência vaga’”, mas nem tanto**

Loss (2018) denomina a proposta semântica de Torza como uma super-metalinguagem. De acordo com sua crítica, Torza teria fracassado em seu intuito de desarmar a *reductio* de Sider. Entretanto, apesar da crítica, Loss assume que a proposta de uma super-metalinguagem é consistente, ainda que exija aceitação da ocorrência de vagueza ontológica com conjunto com uma teoria de vagueza linguística. Desta

---

<sup>16</sup> No original: “for all  $n$ , it is not definitely the case that existence is  $n$ -th order vague” (Torza, 2017, p. 18).

<sup>17</sup> No original: “for all  $n$ , it is not definitely the case that existence is not  $n$ -th order vague” (Torza, 2017, p. 18).

maneira, a crítica de Loss foca no fato de que a existência super-vaga assume vagueza não semântica, algo que Sider nega como ponto de partida. Entretanto, Loss assume o sucesso de Torza em fornecer uma teoria de super-metalinguagem coerente, ainda que partindo de um pressuposto da vagueza ontológica que é negado como ponto de partida por teóricos como Lewis e Sider.

Há alguns desafios no projeto de realização de uma super-metalinguagem (Loss, 2018, p. 3-7). Em primeiro lugar, determinar precisificações para existência leva a uma generalização para ordens superiores, em uma ascendência hierárquica indefinida de meta-linguagens. Fica assim em aberto se a existência seja de fato super-vaga ou não: Como o próprio Torza argumenta, não está provado que não pode haver existência vaga. Mas esta conclusão é uma faca de dois gumes, pois também não está provado que pode. Outro desafio é demonstrar que aceitando o tratamento padrão da vagueza, fica difícil determinar se de fato casos de super-vagueza são mesmo casos de vagueza. Além disso, a proposta de uma super-metalinguagem para sustentar a existência super-vaga é realizada com o objetivo de operar sentenças que seriam válidas para qualquer ordem  $n$  de metalinguagem e ao mesmo tempo inexprimíveis por nenhuma destas ordens. Entretanto, tal linguagem é ameaçada por uma possível versão da *reductio* de Sider para a super-metalinguagem.

Apesar das dificuldades referidas, há uma relação de suporte entre a aceitação de indeterminação do mundo e a existência super-vaga. Dito de forma sucinta, a aceitação de entes ontologicamente vagos — quer seja em termos de conjuntos difusamente existentes ou mundos possíveis — permite a consistência de existência super-vaga.

## 8 Desafios que a vagueza ôntica suscita

O percurso realizado até então nos mostra uma íntima relação entre vagueza de identidade, existência e composição. O deslocamento



para o tema mais geral quanto a vagueza ôntica nos coloca de frente com os desafios gerados pela aceitação de vagueza ôntica. Noonan (Noonan, 2004) trata deste tema. Para ele, ainda que não seja inconsistente aceitar a validade do argumento de Evans e ao mesmo tempo assumir a vagueza ôntica, isso levanta uma série de questões desafiadoras.

O argumento de Evans-Salmon (Nunes, 2020, p. 78) da impossibilidade de identidade vaga funciona como uma redução ao absurdo:

- 1)  $\forall I(a,b)$
- 2)  $\lambda[\forall I(x,a)]b$
- 3)  $\neg \forall I(a,a)$
- 4)  $\neg \lambda[\forall I(x,a)]a$
- 5)  $\neg I(a,b)$
- 6)  $\neg \forall I(a,b)$

De onde se mostra que de (1) se obtém (6), resultando assim em uma contradição.

Ainda que a tentativa do argumento de Evans seja provar a impossibilidade de identidade vaga, ele no máximo tangencia esse objetivo, uma vez que “Mesmo que o argumento tenha sucesso, tudo o que ele prova é que todo objeto vago é determinadamente distinto de todo objeto preciso e de qualquer outro objeto vago.”<sup>18</sup> Ou seja, mesmo que esteja correto o argumento de Evans, ele permite que haja relação de identidade precisa e objetos vagos. Tal posição não parece conflitar diretamente com o senso comum, que pressupõe objetos espacialmente e temporalmente vagos.

É bem verdade que é possível negar indeterminação ôntica aos moldes de Lewis (1986, p. 213), que se opõe às intuições sobre o assunto, defendendo que a vagueza não se encontra nas coisas mesmas, e que forçar um limite arbitrariamente definido seria uma tolice. Por outro lado, é possível ou assumir que há referência sem vagueza à uma

---

<sup>18</sup> No original: “Even if the argument succeeds all it proves is that every vague object is determinately distinct from every precise object and every other vague object” (Noonan, 2004, p. 131).

entidade vaga como aponta a intuição de denotação precisa de Tye (2000, p. 196), que argumenta que o termo “Evereste” se refere precisamente a uma montanha que é um objeto vago. Entretanto, aceitando o argumento de Evans, enquanto Evereste é um objeto vago, então é determinadamente falso que será idêntico a qualquer pedaço de terra preciso nem a outro objeto vago. Mas não é justamente este o caso? Ao assumir a visão de Tye de que o termo “Evereste” se refere precisamente a um objeto vago levanta questões muito desafiadoras aos defensores da vagueza ôntica.

Para alguém que defenda a validade do argumento de Evans e ao mesmo tempo a indeterminação ôntica, é possível aceitar que há uma indeterminação semântica nos termos. Por exemplo, no termo “O maior bretão” identificamos uma ausência de critério no uso desse designador. Mas ao mesmo tempo é consistente pensar que “O maior bretão” é designador preciso de um objeto vago, ainda que essa ideia pareça absurda. Nesse sentido, assim como o termo “Evereste”, o termo “O maior bretão” não poderia designar nenhum homem específico, tal qual “Evereste” não poderia ser igual a nenhum pedaço de terra específico.

Consideremos dois experimentos mentais. o primeiro deles é (1) o edifício de Shoemaker e Swinburne (1984, p. 146), onde duas grandes salas (alfa e beta), conectadas por um vão que impede que determinar definitivamente se as duas grandes salas são partes de um só edifício ou se são dois edifícios distintos conectados. Se  $x$  está em alfa e  $y$  está em beta, é indeterminado se  $x$  e  $y$  estão no mesmo edifício pois é indeterminado se  $\alpha = \beta$ . Isso é consistente para o defensor de vagueza ontológica, uma vez que o argumento de Evans não o ameaça, posto que “o prédio em que  $x$  está” e “o prédio em que  $y$  está” são designadores precisos de objetos vagos distintos. Está determinado que  $x$  está em alfa, mas indeterminado se está em beta. Enquanto que está determinado que  $y$  está em beta enquanto está indeterminado se está em alfa. Ao mesmo tempo cada um desses locais alfa e beta são

determinantemente distintos uns dos outros (2) O mesmo tipo de questão poderia ser colocado pelo experimento mental da Montanha de Edginton (2000), onde seria indeterminado se há duas montanhas ou dois picos de uma única montanha.

Dois outros experimentos mentais estudam questões envolvendo adição e subtração de partes. (3) A casa de Lewis (1993) perdura sendo a mesma casa C adicionando ou subtraindo uma garagem, isso é C+ e C-. Alguém que defenda a indeterminação ôntica pode defender que C designa um objeto com limites vagos, e por isso distinto tanto de C+ como de C-. O mesmo vale para o experimento mental do (4) Rato de Hawley (2002, p. 129), que poderia ter a cauda retirada ou acrescida. Esse rato R seria um objeto que teria uma indeterminação em sua massa, e distinto de R- e R+ que teriam ambas massas determinadas.

Ou seja, tanto nos experimentos da casa de Lewis ou do rato de Hawley, há uma saída para aqueles que defendem a vagueza ôntica, uma vez que a casa (ou o rato) é distinta quanto aos ambientes, uma vez que pode incluir a garagem ou mesmo excluí-la (ou incluindo a cauda ou excluindo), uma vez que a casa (ou o rato) é um objeto com limites vagos (Enquanto a casa apresenta vagueza de limite espacial, o rato apresentaria limites vagos de massa).

Tal qual nos casos anteriores, a defesa da indeterminância ôntica precisa assumir “Evereste” como designador preciso de um objeto vago distinto de qualquer agregado determinado que venha a constitui-lo. Se for assim, há uma série de questões que precisam ser respondidas. Por exemplo, o caso de designação que ocorre no caso de Everest é o mesmo que ocorre na casa de Lewis? E de designação que ocorre nos experimentos mentais do Edifício de Shoemaker ou ainda da Montanha de Edginton, são os mesmos tipos de designação de objetos vagos? E no caso de “O maior bretão”, temos um designador do mesmo tipo?

E suma, ainda que não seja inconsistente aceitar o argumento de Evans e ao mesmo tempo assumir a vagueza ôntica, isso levanta

uma série de questões desafiadoras. Um defensor da vagueza ôntica terminaria desafiado a explicar se (e o porquê) os distintos casos de designação precisa de objetos vagos são iguais ou distintos. Todas estas questões reafirmam o que foi elaborado anteriormente, uma íntima relação entre vagueza existencial e de identidade.

## 9 A vagueza de existência modesta como modo de evitar a vagueza de identidade

Katherine Hawley (2002) argumenta que a aceitação de vagueza existencial não implica em pressupor entes não existentes<sup>19</sup>. Por isso, as teorias que assumem vagueza existencial não precisam ser imediatamente descartadas. Ela mostra que há dois modos distintos de vagueza existencial. Uma versão de vagueza existencial não modesta, que implica na aceitação de entes não existentes, sendo a concepção que usualmente é assumida pelos críticos da noção de vagueza existencial. Entretanto, haveria um tipo de vagueza de existência modesta capaz de se esquivar do compromisso com entes não existentes.

Aqueles que aceitam ocorrência de vagueza podem ser divididos em dois grupos (1) epistemicistas, que consideram vagueza originada em nossa ignorância, e (2) não-epistemicistas, que disputam a origem de vagueza como ontológica ou por indecisão semântica (linguística). Entretanto, um tipo de vagueza não é capaz de determinar o outro. Por exemplo, se assumirmos como ponto de partida que a *palavra* <sup>20</sup> “existir” é não vaga, isso não implica que não haja vagueza existencial e vice-versa. Por isso a precisão do uso da palavra “existe” não seria capaz de esgotar as questões quanto à possibilidade ou não de vagueza ôntica. No mais, uma vez aceita a possibilidade de vagueza ontica de

---

<sup>19</sup> Remetendo ao embate envolvendo referência de entes não existentes como ocorre entre descritivismo e o meinongianismo.

<sup>20</sup> Mais precisamente, o conceito “existir”.

algun tipo (se propriedades ou de existência), disto não deriva diretamente que haja vagueza de identidade. Ainda que a vagueza de identidade seja incoerente e por isso deva ser descartada, tal qual propõem os argumentos Evans-Salmon, isso não implica que não há vagueza de existência.

A vagueza de existência modesta parte da questão especial da composição<sup>21</sup>, em específico na tese mereológica organicista proposta por Inwagen (1995). Uma vez que a composição ocorre pela instanciação de uma propriedade de maneira coletiva pelas partes, é possível identificar dois tipos de situação que invocam casos de vagueza ontológica: vagueza de instanciação de propriedade e vagueza de existência. Ainda que a vagueza de existência implique em vagueza numérica, não obriga o comprometimento com objetos não existentes, desde que tal vagueza seja tomada de maneira modesta.

Há uma distinção entre vagueza de existência e vagueza de identidade. Ainda que ambas sejam tipos de vagueza ontológica, assim como vagueza de instanciação de propriedades, elas não podem ser confundidas. Os argumentos de Evans-Salmon se baseiam na identidade de Leibniz, e por isso sua validade testemunha contra a viabilidade da vagueza de identidade. Para que o argumento ameace a vagueza de existência seria necessário que a indeterminação de existência implicasse em indeterminação de identidade, o que não é verdade.

Assumir que vagueza de existência implica em vagueza de identidade é impedida pela noção modesta de vagueza. Uma crítica possível é que vagueza de existência modesta pode implicar em vagueza de identidade entre conjuntos, obrigando a abrir mão do princípio da extensionalidade. Mas essa questão não obriga a aceitação de vagueza de identidade de objetos.

---

<sup>21</sup> A questão especial da composição se refere às condições necessárias e suficientes para que uma pluralidade caracterize uma totalidade, isso é, venha compor uma totalidade.

Por outro lado, é possível argumentar em sentido oposto, que a vagueza existencial implica em vagueza de identidade.

## **10 Vagueza de existência implica em vagueza de identidade**

Hershenov argumenta em favor da impossibilidade que haja vagueza de existência sem que se aceite vagueza de identidade (Hershenov, 2014). Por *modus tollens*<sup>22</sup>, argumentos contra a possibilidade de vagueza de identidade terminam por operar também contra a vagueza de existência. Além de ser impossível que algo exista determinadamente mas que tenha indeterminadamente partes. Seu argumento encontra implicações de teses mereológicas e também em favor do epistemicismo.

Objetos vagos poderiam ser ou não inteligíveis. Além disso, pode-se argumentar que possuem identidade ou existência vaga. Alguns filósofos são contra a possibilidade de identidade vaga, outros assumem a inteligibilidade de vagueza existencial apesar de sua possibilidade, enquanto outros não. São muito diversas as opiniões. Por exemplo, Morreau (2002) defende que existência vaga não implica em composição vaga, pois vaga é a constituição, não a identidade. Lançando mão de um contraexemplo utilizando organismos, Hershenov discorda de Morreau ao considerar inteligível coisas vagamente existentes que não sejam constituídas por partes determinadamente existentes. Já para Baker (2007), só é possível uma vagueza existencial se este mesmo ente existir futuramente de forma determinada. Tye (1990), por sua vez, defende que vagueza de composição não implica em vagueza existencial nem de identidade. Salmon, mesmo tomando existência como propriedade e objeto como determinados, não escapa da identidade vaga ao assumir objetos com bordas difusas.

---

<sup>22</sup> Como aquele que encontramos no argumento de Sider, no início deste artigo.

Em seguida, Hershenov nos fornece alguns experimentos mentais como o intuito de relacionar partes com seus objetos compostos, chegando à conclusão de que negar a identidade indeterminada obrigaria a aceitar repentina passagem entre duas substâncias. Mas isso seria “insuportavelmente arbitrário”<sup>23</sup>. Ou seja, aquele que desejasse defender a identidade indeterminada seria obrigado a escolher entre um essencialismo mereológico ou deveria adotar a posição epistemicista. Dentre essas duas possibilidades a postura epistemicista parece ser muito mais defensável.

Em resumo, evitar a vagueza de identidade que pode levar a conclusões absurdas envolve também a negação da vagueza existencial e de composição. Uma segunda via de argumentação poderia abordar a possibilidade de inteligibilidade de identidades vagas. Imaginar uma identidade vaga parece ser tão ininteligível que mesmo a experiência que esse tipo de evento causaria a um sujeito individual deveria ser “parcial”. Entretanto, parece ser completamente ininteligível que uma experiência possa ser parcial. É claro, esse tipo de crítica apoia-se na questionável tese de que qualquer que seja o evento no mundo, ele deve não apenas inteligível, mas também criar em nós experiências inteligíveis.

## 11 Considerações finais

Aqueles que desejarem defender a tese de que a existência vaga é coerente precisa apresentar um modelo semântico capaz de suportar um quantificador que esteja de acordo com tal tese. Além disso, terá que apresentar uma resposta contra o argumento de Evans-Salmon, mostrando que identidade vaga não incorre em contradição. Mesmo

---

<sup>23</sup> Assim como seria arbitrário decidir um momento de passagem nos paradoxos de sorites ou mesmo no problema dos muitos, onde não é possível fornecer um parâmetro para eleger um homem verdadeiramente existente dentre todos os outros homens candidatos.

aceitando a ideia de uma “existência modesta” tal qual Hawley interpreta a proposta organicista de Inwagen, o modelo supervaloracionista parece encontrar dificuldades insuperáveis envolvendo exigências de determinação de extensionalidade dos conjuntos, além de pressupor sempre um domínio determinado. Ou seja, a possibilidade aberta de vagueza ôntica aprofunda os desafios de um tratamento supervaloracionista da vagueza para além do que concerne às propriedades. Ciente destes desafios, Inwagen abandona a possibilidade de um tratamento supervaloracionista em detrimento de uma abordagem multivaloracionista<sup>24</sup>. Assim ele defende que é possível expressar todos os tipos de vagueza ôntica ao mesmo tempo que barra a validade irrestrita do princípio de identidade de Leibniz e torna inválido o argumento Evans-Salmon.

## Referências

- BAKER, L. R. *The metaphysics of everyday life: An essay in practical realism*. Cambridge: Cambridge, 2007.
- EDGINGTON, D. Indeterminacy de re. *Philosophical Topics*, v. 28, n. 1, p. 27-44, 2000.
- EVANS, G. Can there be vague objects. *Analysis*, v. 38, n. 4, p. 208, 1978.
- GEACH, P. T. *Reference and generality* (selections). Material constitution: A reader, 1997.
- HAWLEY, K. Vagueness and Existence. *Proceedings of the Aristotelian Society* (Hardback), v. 102, n. 1, p. 125-140, jun. 2002.
- HERSHENOV, D. B. Vague Existence Implies Vague Identity. Em: AKIBA, K.; ABASNEZHAD, A. (Eds.). *Vague Objects and Vague Identity*. Logic, Epistemology, and the Unity of Science. Dordrecht: Springer Netherlands, 2014. v. 33, p. 283-303.
- HYDE, D.; RAFFMAN, D. Sorites Paradox. *Metaphysics Research Lab*, Stanford University, 2018.

---

<sup>24</sup> Em específico ele utiliza a lógica fuzzy.



- KOSLICKI, K. The crooked path from vagueness to four-dimensionalism. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, v. 114, n. 1/2, p. 107-134, 2003.
- LEWIS, D. Many, but Almost One. In: CASTELLANI, E. (Ed.). *Interpreting Bodies*. Princeton University Press, 1993. p. 30-45.
- LEWIS, D. *On the Plurality of Worlds*. 1986.
- LOSS, R. Against 'Against "Against Vague Existence"'. *Oxford Studies in Metaphysics*, v. 11, p. 278, 2018.
- MORREAU, M. What vague objects are like. *The Journal of Philosophy*, v. 99, n. 7, p. 333-361, 2002.
- NOONAN, H. W. Are there vague objects? *Analysis*, v. 64, n. 2, p. 131-134, 2004.
- NUNES, R. *Totalidades e estrutura mereológica: Um estudo sobre a natureza dos objetos compostos*. Tese (Doutorado) – Universidade Federal do Rio de Janeiro, 2020.
- SHOEMAKER, S.; SWINBURNE, S. *Personal Identity: Great Debates in Philosophy*. Oxford, England: Blackwell, 1984.
- SIDER, T. Against vague existence. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, v. 114, n. 1/2, p. 135-146, 2003.
- TORZA, A. Vague existence. *Oxford Studies in Metaphysics*, v. 10, p. 201-233, 2017.
- TYE, M. Vague objects. *Mind*, v. 99, n. 396, p. 535-557, 1990.
- TYE, M. Vagueness and reality. *Philosophical Topics*, v. 28, n. 1, p. 195-209, 2000.
- UNGER, P. The Problem of the Many. *Midwest Studies in Philosophy*, 1980.
- VAN INWAGEN, P. *Material beings*. United Kingdom: Cornell University Press, 1995.
- WEATHERSON, B. The Problem of the Many. *Metaphysics Research Lab*, Stanford University, 2023.



# Modos de impossibilidade: uma nova taxonomia

Pedro de Sá Torres Barbosa<sup>1</sup>

DOI: <https://doi.org/10.58942/eqs.158.04>

## 1 Introdução

O que se pretende é examinar uma questão fundamental: a natureza mesma dos objetos impossíveis e do discurso significativo sobre eles, compreendendo como o conceito de “impossibilidade” admite variações e nuances que, uma vez propriamente apreendidas, podem auxiliar a entender melhor os limites e possibilidades do pensamento humano.

O que está em questão, aqui, não é um problema meramente técnico em lógica modal, mas a própria noção de modalidade está em jogo, se a metafísica modal se trata sobre o que pode ser o caso, então o que podemos, de fato, dizer sobre o que não pode ser o caso? Não pode em que sentido? Em que contexto? Como compreender as diferentes manifestações do discurso acerca do impossível na linguagem natural? Quais são os limites do que pode ser dito?

Para enfrentar algumas destas questões, propomos uma análise que se desdobra em duas dimensões fundamentais: as categorias e as

---

<sup>1</sup> Pedro de Sá Torres Barbosa é bacharel em filosofia pela Universidade Federal da Paraíba. Atualmente no programa de pós-graduação do PPGF da UFPB na linha de Lógica e Epistemologia no nível de mestrado. Desenvolve pesquisa acerca do Meinongianismo Modal. Email: [pstb@academico.ufpb.br](mailto:pstb@academico.ufpb.br)

forças da impossibilidade. Por categorias, entendemos os diferentes modos pelos quais algo pode ser dito como sendo ‘impossível’ — seja uma impossibilidade lógica, como no caso do círculo quadrado, uma impossibilidade ontológica, como um objeto que existe e não existe simultaneamente, esta diz respeito à sua progressiva dissociação com as categorias ontológicas como leis naturais, lógicas, de coerência conceitual, de existência e assim por diante. Por força, compreendemos o grau de resistência que cada impossibilidade oferece à própria possibilidade de pensamento e representação, esta é a dimensão, por assim dizer, epistemológica e cognitiva da impossibilidade.

Nossa proposta é demonstrar que uma compreensão adequada destas diferentes modalidades e forças da impossibilidade não apenas responde às críticas ao meinongianismo, mas nos permite identificar com maior precisão os verdadeiros limites do pensamento e do discurso significativo. Em última análise, argumentaremos que o verdadeiro limite não se encontra no impossível em si — pois sobre este ainda podemos pensar e falar significativamente, e sim no limite ultimo do ser (e, portanto, do pensável): O Nada.

## **2 Impossibilidades em metafísica modal**

Na lógica modal clássica, uma proposição impossível é aquela que é necessariamente falsa, ou seja, é verdadeira em nenhum mundo possível. Isso pode ser expresso em termos de quantificação: em todos os mundos possíveis, o quantificador existencial falha para objetos ficcionais, resultando na proposição de que sua existência é universalmente falsa. Mas isto não nos diz nada sobre o que significa dizer, de fato, que algo não é o caso. Aqui faremos uma pequena jornada em torno da literatura acerca de mundos impossíveis, focando particularmente em Takashi Yagisawa. Contraste e compare como David Lewis e Yagisawa definem, respectivamente, mundos possíveis e impossíveis:

I believe, and so do you, that things could have been different in countless ways. But what does this mean? Ordinary language permits the paraphrase: there are many ways things could have been besides the way they actually are. I believe that things could have been different in countless ways; I believe permissible paraphrases of what I believe; taking the paraphrase at its face value, I therefore believe in the existence of entities that might be called ‘ways things could have been.’ I prefer to call them ‘possible worlds’ (Lewis, 1973, p. 84).

There are other ways of the world than the way the world actually is. Call them ‘possible worlds.’ That, we recall, was Lewis’ argument. There are other ways of the world than the ways the world could be. Call them ‘impossible worlds’ (Yagisawa, 1988, p. 183)

Então vemos como, ao menos para Yagisawa, pensar em mundos impossíveis é tão natural quando pensar em mundos possíveis, afinal, se precisamos de mundos possíveis para dar sentido para o discurso modal, então precisaríamos de mundos impossíveis como uma forma de ‘anti-extensão’ para mundos possíveis, como poderíamos, afinal, dar sentido à maneira como as coisas não poderia, de alguma forma, ser? Mas esta dicotomia é limitada, afinal, há muitas maneiras como poderíamos descrever como estados de coisas poderiam ser diferentes, e o próprio Yagisawa reconhece:

*Some impossible worlds are worlds at which a logical impossibility, for example, a proposition expressed by a sentence of the form ‘P and not-P’, is true. These are logically impossible worlds. Some impossible worlds are worlds at which a metaphysical impossibility, for example, the de re proposition concerning this table — which is actually made of wood — that it is made of ice, is true. These are metaphysically impossible worlds. There are other kinds of impossibility, each corresponding to a kind of possibility, and they give rise to their own distinct kinds of impossible worlds (Yagisawa, 2010, p. 176).*

O problema, aqui, é que ‘impossibilidade’, como conceito, não é uma categoria estanque, então parece que precisamos dar conta de distinguir sistematicamente a taxonomia destes ‘mundos impossíveis’, Yagisawa propõe que pensemos em termos de “K-possibilidades”:

I suggest that we start with the idea of a way for things to be, and postulate worlds as modal indices at which things are various ways, rather than identifying worlds with these ways themselves. We should then say that for each type of K-relevant facts, some of these worlds are K-possible and the others are K-impossible. For each K, K-impossible worlds should be conceptually as uncontroversial as K-possible worlds. Unfortunately, impossible worlds are controversial when understood in terms of sufficiently strong K-relevant facts, such as the laws of mathematics, metaphysics, or logic (Yagisawa, 2010, p. 179).

Isto é, não existe algo como uma possibilidade “nua” — toda possibilidade é sempre relativa a algum conjunto K de fatos relevantes (K-possibilidade). Apresenta então três princípios fundamentais:

1. Todo mundo é possível em algum sentido (K-possível para algum K)
2. Para qualquer tipo K de possibilidade, existem mundos K-impossíveis
3. Para qualquer tipo K de possibilidade, existe um tipo mais amplo K' onde alguns mundos K-impossíveis são K'-possíveis

Ao meu ver, penso, que (1) e (2) são relativamente seguras de um ponto de vista meinongiano<sup>2</sup>, porém (3) é algo problemático: se para qualquer K-possibilidade, existem uma maneira de ampliar K de tal forma que algumas impossibilidades ‘tornem-se’ possíveis, então como sabemos qual é a extensão de alguns? Não me aparece razoável estender indefinidamente a noção de possibilidade de modo a trivializar a nossa semântica modal, mas não me parece que Yagisawa, apesar de sugerir a questão, nos oferece uma taxonomia satisfatória de como

---

<sup>2</sup> Yagisawa declara-se ‘realista’ (ainda que mais moderado que David Lewis), e não se comprometeria, penso, com a ‘carga’ filosófica do meinongianismo, mas também esta se trata de uma limitação fundamental do seu tratamento sobre impossibilidades, pois não é capaz de tratar, adequadamente das condições mesmas em que o discurso predicativo pode se dar.

classificar e delimitar a extensão do conceito de impossibilidade, mas sua sugestão de realizar uma tal taxonomia é, penso, fundamental.

Yagisawa argumenta que a possibilidade (e consequentemente a impossibilidade) é sempre relativa a um determinado conjunto de fatos ou informações relevantes, que ele denomina “K-relevantes”. Esta relativização manifesta-se em diferentes tipos de possibilidade:

- Possibilidade legal
- Possibilidade física
- Possibilidade metafísica
- Possibilidade lógica
- Possibilidade conceitual

Em geral eu concordo com esta categorização (ainda que eu ache que postular um espaço de ‘possibilidades legais’ é um tanto discutível), mas isso não nos interessa tanto quanto nosso desacordo fundamental: a análise de Yagisawa limita-se, fundamentalmente, ao nível da predicação, onde mesmo as ditas ‘possibilidades metafísicas’ podem ser relativizadas segundo uma ampliação conceitual de K-relevância. Eu argumento, entretanto, que a ampliação não pode dar-se indefinidamente. Vejamos como ele trata as ‘possibilidades metafísicas’. Se eu digo, por exemplo, que há ‘solteiros casados’, eu estou articulando uma possibilidade lógica em um espaço onde contradições são aceitáveis, porem uma impossibilidade metafísica onde os predicados contradizem-se entre si, algo não pode ser ambos solteiro e casado, entretanto seria plausível expandir o tipo K para permitir u mundo tal que seja possível dizer que alguém seja ambos casado e solteiro ao mesmo tempo, ou mesmo tornar a distinção trivial pela eliminação da categoria de matrimônio. Parece, então, que mesmo as tais ‘impossibilidades metafísicas’ tornam-se tão domesticadas, por assim dizer, quanto as categorias modais tradicionais, como um criminoso que sai de uma jurisdição para outra onde seus atos não são considerados criminosos, o problema é que, assim como o crime não é eliminado pela

fuga, o problema fundamental da impossibilidade não é resolvida: se algo não pode ser o caso, ele não pode ser o caso em contexto algum, isto é, ele não admite relativização ou enfraquecimento por uma manobra conceitual filosófica, é necessário ir mais fundo se algo é predicável, então em função de ser algo, antes, para quem, somente após isto, seja predicável. Na taxonomia de Yagisawa, portanto, não há espaço para uma posição como ‘é’ ou ‘não é’, em sentido fundamentalmente meta-ontológico (utilizando o jargão de Inwagen), eis a limitação fundamental de qualquer taxonomia que se limite a análise apenas sentenças de predicação.

### 3 Força das impossibilidades

A investigação sobre a força das impossibilidades emerge de uma intuição fundamental acerca da própria natureza da *objetividade*: existe uma relação inversamente proporcional entre o grau de impossibilidade de um objeto e sua proximidade com o domínio do inteligível. Esta relação não é meramente accidental, mas revela algo essencial sobre a estrutura mesma do pensamento e sua capacidade de apreensão objetiva.

É crucial notar que a força de uma impossibilidade, embora seja uma categoria epistêmica, está fundamentalmente subordinada à sua categoria ontológica. Esta subordinação não é accidental, mas reflete uma hierarquia metafilosófica fundamental: o ser precede e condiciona o conhecer. Assim, quanto mais fundamental a categoria de impossibilidade (isto é, quanto mais ela se afasta do ser enquanto ser), maior tende a ser sua força (isto é, maior seu afastamento do cognoscível). Esta correlação entre categoria e força não é um pressuposto arbitrário da taxonomia, mas emerge da própria natureza da relação entre ser e conhecimento, ainda que o leitor possa desconfiar desta relação hierárquica, penso que a intuição subjacente à taxonomia que estou propondo, ainda assim, será frutífera, para gregos e troianos.



Quando nos movemos progressivamente do domínio do atual em direção ao impossível, observamos um gradual obscurecimento da inteligibilidade que corresponde a diferentes níveis de afastamento das condições de possibilidade que estruturam nossa compreensão do real. Este movimento não é uniforme, mas apresenta gradações significativas que nos permitem compreender melhor a própria natureza da impossibilidade enquanto tal.

As impossibilidades podem variar em força precisamente na medida em que violam, com diferentes graus de profundidade, os princípios fundamentais da lógica e da realidade. Esta variação não é meramente quantitativa, mas qualitativa, revelando diferentes níveis de afastamento da própria estrutura do pensável. Podemos, assim, distinguir dois modos fundamentais desta variação:

As impossibilidades fracas são aquelas que, embora contradigam o estado atual de coisas, mantêm ainda uma relação relativamente próxima com o domínio do inteligível. São impossibilidades que dependem essencialmente de convenções ou circunstâncias particulares, como exemplifica o caso paradigmático da “montanha sem vale”. A força reduzida destas impossibilidades manifesta-se precisamente na possibilidade de sua relativização ou contorno dentro de sistemas conceituais alternativos. Quando afirmamos a impossibilidade de uma montanha sem vale, esta impossibilidade deriva não de uma contradição lógica fundamental, mas de nossa maneira contingente de conceitualizar e categorizar formações geológicas.

Por outro lado, as impossibilidades fortes caracterizam-se por uma violação mais radical da própria estrutura do pensável. O caso paradigmático do “círculo quadrado” não representa apenas uma contradição com o estado atual de coisas, mas uma violação direta dos princípios formais de identidade e não-contradição que estruturam nossa capacidade mesma de pensamento. Estas impossibilidades não podem ser simplesmente relativizadas ou contornadas por mudanças em nossos sistemas conceituais — sua “força” deriva precisamente do fato de

que sua própria concepção exigiria uma reformulação completa dos fundamentos mesmos de nossos sistemas lógicos.

Há ainda a impossibilidade de força *absoluta*, que representa um completo descolamento entre o que se quer dizer e o que é inteligível, isto é, há um descolamento intencional, onde o objeto que queremos atingir já não pode ser sob nenhuma circunstância, pois ele viola algo fundamental que danifica irreparavelmente nossa capacidade de representação, ela está para fora do pensável em sentido absoluto.

Esta distinção fundamental entre impossibilidades fracas e fortes não é meramente taxonômica, mas revela algo essencial sobre a própria estrutura da *objetividade*. Quanto mais “forte” a impossibilidade, mas ela se afasta não apenas do atual, mas das próprias condições de possibilidade que estruturam nossa capacidade de pensamento e representação. Neste sentido, a força de uma impossibilidade pode ser compreendida como uma medida do quanto ela desafia não apenas nossas concepções contingentes, mas os próprios fundamentos que possibilitam qualquer concepção.

O que esta análise revela sobre a estrutura mesma da objetividade é que ela não se apresenta como um fenômeno de tudo-ou-nada, mas admite gradações que correspondem a diferentes níveis de inteligibilidade e apreensibilidade. Um objeto impossível não deixa de ser objeto por ser impossível, mas sua objetividade se manifesta em diferentes graus de clareza e determinação, conforme o nível de sua impossibilidade. Esta compreensão graduada da objetividade não apenas enriquece nossa compreensão do impossível, mas nos permite mapear com maior precisão os próprios limites do pensável.

Em última análise, quanto mais “forte” a impossibilidade, mais ela se aproxima daquele limite último além do qual nenhum pensamento pode aventurar-se — o domínio do “nada absoluto”, onde nenhum pensamento pode se dar. Mas, é importante notar que a ‘força’ da impossibilidade é uma categoria epistêmica, não propriamente metafísica, ela não é uma categoria do ser, mas da relação do agente

epistêmico em relação aos conceitos, portanto, ela é derivada e subsidiária da metafísica, afinal, não se pode conhecer qualquer coisa que seja sem que, antes, haja algo a ser conhecido. Dizendo de outra forma, com uma terminologia mais sugestiva: a ‘força’ de uma impossibilidade diz respeito à sua *abertura*, por assim dizer, aos poderes de representação do cognoscente de maneira inversamente proporcional.

## 4 Categorias de impossibilidade

A investigação sobre as categorias fundamentais da impossibilidade exige-nos transcender a mera classificação taxonômica para alcançar uma compreensão mais profunda da própria estrutura do ser e do pensamento. Se a análise prévia das forças da impossibilidade nos revelou uma gradação na inteligibilidade dos objetos impossíveis, o exame de suas categorias fundamentais nos conduzirá a uma compreensão mais radical da própria natureza da impossibilidade enquanto tal.

A necessidade de distinguir diferentes categorias de impossibilidade não emerge de uma conveniência metodológica, mas da própria estrutura ontológica da realidade e do pensamento. Assim como o ser, como afirma Aristóteles na Metafísica Livro IV, se diz de múltiplos modos, também a impossibilidade se manifesta em diferentes modalidades que correspondem a diferentes níveis de violação da estrutura fundamental da realidade. Esta multiplicidade categorial não é acidental, mas revela distinções essenciais na própria natureza das impossibilias<sup>3</sup>.

As impossibilidades nomológicas nos confrontam com os limites das leis que governam a realidade física como a conhecemos. Os

---

<sup>3</sup> Deixarei de lado aqui as tais ‘impossibilidades legais’ postuladas por Yagisawa pois entendo que elas não tratam-se, de fato, da categoria geral de impossibilidade que nos interessa aqui, isto é, aquelas que dizem respeito, em maior ou menor medida à questões de metafísica. Categorias legais pertencem ao âmbito da tal ‘ontologia social’ e, portanto, é desinteressante para os propósitos deste artigo.

paradoxos temporais, por exemplo, embora não violem necessariamente princípios lógicos fundamentais, desafiam nossa compreensão das leis causais que estruturam o universo físico. Outro exemplo seria a noção de bilocação, um dos milagres que São Francisco teria realizado, onde ele pareceu estar em dois lugares ao mesmo tempo, claramente as leis naturais macroscópicas não permitem este ato, portanto é nomologicamente impossível e é, portanto, um evento milagroso (para aqueles que acreditam).

As impossibilidades conceituais, por sua vez, emergem no interstício entre a lógica pura e a experiência concreta. Uma “montanha sem vale” representa uma impossibilidade que, embora não viole diretamente leis da lógica, viola nosso entendimento intuitivo de como conceitos estão interligados, pensamos naturalmente que montanhas implicam vales pela sua própria definição, ou no jargão kantiano, dizemos que são verdadeiras analiticamente e apriori, também não trata-se de uma afirmação metafisicamente impossível, trata-se de uma limitação epistêmica, fruto, talvez, de como o ‘maquinário conceitual’ de nossas mentes operam, entretanto, não passa disso.

As impossibilidades lógicas constituem, talvez, o caso mais imediatamente apreensível desta categorização. Quando nos deparamos com um “círculo quadrado”, não estamos meramente diante de uma impossibilidade factual, mas de uma violação das próprias condições de possibilidade do pensamento geométrico. O círculo quadrado não é apenas contingentemente impossível — sua impossibilidade deriva de intuições lógico-matemáticas cuja relativização exige a reconstrução destas intuições de modo a torná-las cognoscíveis, talvez em alguma estranha geometria sejamos, de fato, capazes de adequar nossas intuições ao círculo-quadrado. Como Yagisawa (2010) sugeriu, a expansão do nosso espaço de relevância nos permite relativizar certos tipos de impossibilidade.

Em seguida podemos considerar o que gostaria de postular como sendo impossibilidades ônticas, isto é, aquelas que dizem

respeito à categoria de existência como instanciação concreta, por exemplo, o gato em minha casa não pode ser, ao mesmo tempo, o gato na casa do meu vizinho, e ele também não pode estar e não existir, também deve ser ‘aquele’ gato e não outro, senão não seria o meu gato. Tais impossibilidades estão intimamente ligadas com impossibilidades nomológica, mas diferenciam-se no sentido que, aqui, estamos tratando das condições mesmas de possibilidade de que uma ‘jurisdição’ nomológica possa se dar, portanto, é mais fundamental que aquela.

Mais fundamental ainda, são as impossibilidades ontológicas, se até então estávamos pensando em como as coisas poderiam ou não, apresentar-se em um domínio de coisas existentes, ou conceituais (existentes ou não<sup>4</sup>), ou mesmo possíveis em certos regimentos de leis naturais, lógico-matemáticas e etc...aqui estamos falando de impossibilidades que se dão a nível do ser enquanto ser, aquele mesmo que Parmênides, em seu poema, pretendia discursar: o ser é e o ser não é e nem pode ser, se não fosse, afinal, nada seria, e nada poderia ser, portanto, quando estamos falando nesse nível, estamos falando das condições mesmas de qualquer predicação possível, afinal, como poderemos falar que algo é ou não pertencente a esta ou aquela categoria sem antes entendermos que algo é<sup>5</sup>? Então, qual seria o exemplo emblemático? Ora, ‘algo não é’, em sentido propriamente ontológico, é fundamentalmente impossível, e impossível de maneira que não pode ser capturada por

---

<sup>4</sup> Com isso quero dizer que podemos conceptualizar, intuitivamente, sobre diferentes coisas, sejam elas existentes ou não, trata-se de outro nível de discurso, com a distinção que ele ainda está condicionado pela predicação.

<sup>5</sup> Esta é minha leitura da dita ‘interpretação não-predicativa’ de Parmênides como proposta por José Gabriel Trindade dos Santos, infelizmente não tive acesso, ainda, aos artigos originais, porém tive acesso a uma dissertação de mestrado feita por Antonio Carlos da Silva Pinheiro da UFCE que descreve a proposta. Como ele descreve: “não-predicatividade é um termo que Santos utiliza para se referir a uma leitura dos caminhos que não lê as formas verbais usadas como cópulas de enunciados virtuais, em que um sujeito implícito é ligado a um predicado elidido: “[S] é/não é [P]”. A sugestão, portanto, é que devemos ler o inai de Parmênides como uma categoria intencional. Particularmente eu simpatizo com esta proposta, ainda que a área não seja minha especialidade.

intuição alguma, trata-se, portanto, do ‘nada enquanto nada’<sup>6</sup>, dela nada se diz, não pode ser dito, não é sequer objeto: de uma perspectiva meinongiana, está fora do domínio do *aussersein*.

Esta categorização, longe de ser um mero exercício classificatório, revela-nos diferentes níveis de penetração na estrutura fundamental da realidade. As impossibilidades não são todas equivalentes precisamente porque correspondem a diferentes níveis de violação da ordem do ser e do pensamento. Uma impossibilidade ontológica é, em certo sentido, mais fundamental que uma impossibilidade meramente nomológica, pois atinge a própria estrutura do ser, e não apenas suas manifestações contingentes.

Mais ainda, esta análise categorial permite-nos compreender melhor a própria natureza do objeto impossível no quadro do pensamento meinongiano. Se diferentes categorias de impossibilidade correspondem a diferentes níveis de violação da ordem do ser, então diferentes objetos impossíveis mantêm diferentes relações com a própria estrutura da objetividade. Um objeto logicamente impossível, como o círculo quadrado, relaciona-se com a objetividade de maneira fundamentalmente distinta de um objeto apenas nomologicamente impossível.

Esta compreensão categorial da impossibilidade não apenas enriquece nossa taxonomia do impossível, mas nos permite mapear com maior precisão os diferentes modos pelos quais o pensamento pode aproximar-se — e eventualmente ultrapassar — seus próprios limites. Em última análise, são estas distinções categoriais que nos permitem

---

<sup>6</sup>A noção de ‘nada’ tem um grande peso filosófico, mas quando dizemos, na linguagem ordinária, que ‘isto não vai dar em nada’ ou ‘encontramos nada’, sobre o que estamos falando? Se estivéssemos, de fato, falando sobre o nada enquanto tal, enquanto ausência de ‘algo’, então as declarações acima não fazem sentido, então é claro que, ao menos nesses exemplos, não é sobre o ‘nada’, enquanto tal, que estamos falando. Mesmo quando desejo falar deste tal ‘nada-enquanto-tal’, não posso realmente falar sobre ela, nossa referência a este conceito é sempre oblíqua, nós temos alguma ‘intuição’ do que estamos falando, mas ao tentar enunciá-la, vemos que esta intuição imediatamente se desfaz e notamos que ‘algo’ nos escapa, e que linguagem é inadequada.

compreender não apenas o que é impossível, mas por que e como algo pode ser impossível.

## 5 Exemplos

### 5.1 Círculo quadrado

- Categoria: Impossibilidade lógica
- Força: Forte

Um “círculo quadrado” é impossível porque sua definição envolve uma contradição direta com os princípios da lógica formal. Um círculo, por definição, é uma figura geométrica cuja distância do centro a qualquer ponto da borda é constante, enquanto um quadrado tem lados retos e ângulos de 90 graus. Essas duas propriedades são mutuamente exclusivas. A impossibilidade aqui é lógica e conceitual, já que a tentativa de combinar essas propriedades viola o princípio da não-contradição. Portanto, a força dessa impossibilidade é muito forte, pois não pode ser contornada ou resolvida por qualquer modificação empírica, ou de reinterpretação.

### 5.2 Objeto que existe e não existe simultaneamente

- Categoria: Impossibilidade ôntica
- Força: Forte

A afirmação de que um objeto “existe e não existe simultaneamente” é uma impossibilidade ôntica, pois afeta a noção de existência, quando dizemos que algo existe e não existe ao mesmo tempo, dizemos dele que este é e não é um objeto ôntico, o que claramente viola a sua possibilidade como objeto concreto e existente.

### 5.3 Montanha sem Vale

- Categoria: Impossibilidade conceitual
- Força: Fraca

Uma “montanha sem vale” é uma impossibilidade conceitual. Embora não envolva contradições formais, esse conceito desafia as condições normais de aplicação de nossos conceitos. Uma montanha, por definição, exige uma elevação em relação a uma depressão ao redor, o que implica a existência de um vale. No entanto, a impossibilidade aqui é considerada fraca, pois o conceito de montanha e vale é contingente, dependente de nossa maneira de classificar formas geomorfológicas. Em outro sistema conceitual, essa distinção poderia ser reformulada.

### 5.4 Algo simultaneamente vermelho e não vermelho

- Categoria: Impossibilidade conceitual
- Força: Forte

A afirmação de que um objeto pode ser “simultaneamente vermelho e não vermelho” é uma impossibilidade conceitual, pois atribui ao mesmo objeto duas propriedades mutuamente exclusivas ao mesmo tempo. Isso é uma violação clara do princípio da não-contradição, que estabelece que um objeto não pode ter duas qualidades contraditórias simultaneamente. A força dessa impossibilidade é forte, uma vez que a contradição surge diretamente da aplicação de categorias perceptuais e lógicas fundamentais, e não pode ser resolvida empiricamente ou reinterpretada de maneira significativa.

### 5.5 Paradoxo do avô (viagem no tempo)

- Categoria: Impossibilidade nomológica (causal)
- Força: Forte (mas contingente)

O “Paradoxo do Avô” é um exemplo clássico de impossibilidade nomológica e causal. Ele descreve uma situação em que uma pessoa viaja de volta no tempo e impede a existência de seu próprio avô, tornando impossível sua própria existência. Esse paradoxo cria uma



contradição causal, já que a pessoa não poderia existir para fazer a viagem no tempo se o avô não tivesse existido. A impossibilidade aqui é fortemente nomológica, pois desafia as leis da causalidade como as conhecemos. Embora a força seja forte no sentido de que a causalidade não pode ser violada, há uma dependência das leis físicas atuais, o que implica que a impossibilidade pode ser, em princípio, contingente se teorias científicas futuras redefinirem o conceito de causalidade.

## 5.6 Objeto em dois lugares simultaneamente

- Categoria: Impossibilidade nomológica
- Força: Fraca a moderada (potencialmente contingente)

A impossibilidade de um “objeto estar em dois lugares ao mesmo tempo” é uma impossibilidade física que decorre das leis da física clássica, que estabelecem que um objeto ocupa um lugar único no espaço-tempo. No entanto, essa impossibilidade é enfraquecida por avanços na mecânica quântica, onde partículas subatômicas podem exibir comportamentos que parecem desafiar essa noção (como no caso da superposição quântica). Assim, a força dessa impossibilidade varia de fraca a moderada, dependendo do nível em que se considera as leis físicas: no nível clássico, a impossibilidade é mais forte, mas no nível quântico, pode ser relativizada.

## 5.7 Algo não é

- Categoria: Impossibilidade ontológica
- Força: Absoluta

A afirmação ‘algo não é’, quando é tratada não-predicativamente, trata-se de uma afirmação incompreensível, portanto de força absoluta, visto que se trata do afastamento completo de sua inteligibilidade, pois trata-se de uma impossibilidade igualmente absoluta. O que quero dizer é que não podemos nem sequer colocar ‘o que não é’ em palavras, quando eu coloco ‘nada absoluto’ em uma sentença, eu não posso sequer pressupor sua objetificação, trata-se de uma

referência-limite: podemos inferir obliquamente que para além do ser em sentido total, nada há, mas é uma referência oblíqua, semanticamente vazia e, portanto, não é objeto, não pode ser objeto, pois ela simplesmente não é.

## 6 Considerações finais

Ao distinguirmos entre diferentes categorias e forças de impossibilidade, podemos ver um caminho que, penso, será frutífero para uma investigação mais detalhada e sistemática da natureza da modalidade e dos próprios limites da ontologia.

A análise desenvolvida demonstra que a impossibilidade não é um fenômeno monolítico, mas apresenta gradações e modalidades que correspondem a diferentes níveis de afastamento da estrutura fundamental do ser e do pensamento. Das impossibilidades nomológicas, que apenas contradizem as leis naturais conhecidas, até as impossibilidades ontológicas fundamentais, que desafiam a própria estrutura do ser enquanto ser, observamos um progressivo obscurecimento da compreensão do ser em sentido fundamental. Por outro lado, vemos outra progressão, subsidiária à anterior, que estabelece um crescente obscurecimento da sua inteligibilidade.

## Referências

- BERTO, Francesco; JAGO, Mark. *Impossible Worlds*, The Stanford Encyclopedia of Philosophy (Summer 2023 Ed.), ZALTA, Edward N.; NODELMAN, Uri (Eds.), Disponível em: <https://plato.stanford.edu/archives/sum2023/entries/impossible-worlds/>.
- CORDERO, Nestor Luiz Sendo Se È: *A Tese de Parmênides*. Odysseus, 2011.
- DA SILVA PINHEIRO, Antonio Carlos. O poema de Parmênides: uma leitura não-predicativa do argumento. Dissertação (Mestrado) – Universidade Federal do Ceará, 2023. Disponível em: <http://repositorio.ufc.br/handle/riufc/75650>. Acesso em: 15 jan. 2025.

- FITTING, Melvin; MENDELSON, Richard. *First-Order Modal Logic*. Springer, 1998. 301 p.
- LEWIS, David. *Counterfactuals*. Malden, Massachusetts: Blackwell Publishers, 1973. Disponível em: <https://perso.uclouvain.be/peter.verdee/counterfactuals/lewis.pdf>. Acesso em: 15 jan. 2025.
- MAREK, Johann; MEINONG, Alexius. The Stanford Encyclopedia of Philosophy (Summer 2021 Ed.), ZALTA, Edward N. (Ed.). Disponível em: <https://plato.stanford.edu/entries/meinong/#PriIndPriInd>.
- MEINONG, Alexius. *The Theory of Objects, parte da coletanea Realism and The Background of Realism*. CHISHOLM, Rocderick M. (Ed.). The Free Press of Glencoe, 1960.
- MEINONG, Alexius. *Três Aberturas em Ontologia: Frege, Twardowski e Meinong*. Trad. Celso Braída. Florianópolis: Rocca Brayde, 2005.
- PARMENIDES. *Da Natureza*. Trad. José Gabriel Trindade dos Santos. Loyola, 2002.
- YAGISAWA, Takashi. Beyond possible worlds. *Philosophical Studies* 53, 175-204, 1988. Disponível em: <https://doi.org/10.1007/BF00354640>.
- YAGISAWA, Takashi. *Worlds and Individuals, Possible and Otherwise*. Oxford, Nova Iorque: Oxford University Press, 2010.



# Some remarks on mind-dependence in social ontology and its relationship to thought experiments in philosophy of mind

Valdenor Monteiro Brito Júnior<sup>1</sup>

DOI: <https://doi.org/10.58942/eqs.158.05>

## 1 Introduction

‘Mind-dependence’ is an elusive term. Both in the specific case of social ontology and in the general case respective to many philosophical disciplines, it is not uncommon to see the notion of ‘mind-dependence’ mentioned without definition or barely defined. However, in social ontology we have nowadays some attempts to clarify what the mind-dependence thesis asserts for the case of social kinds (Khalidi, 2016; Mason, 2016; Brito Jr., 2021<sup>2</sup>).

The mind-dependence thesis in social ontology postulates that social entities ontologically depend on minds or mental states.

---

<sup>1</sup> Substitute Professor at Federal University of Paraná. PhD in Philosophy from Federal University of Santa Catarina. My main research interests are in Metaphysics and Social Ontology, with intersections spanning across Action Theory, Theoretical Biology, Philosophy of Social Sciences and Legal Theory. E-mail: [valdenormb@hotmail.com](mailto:valdenormb@hotmail.com)

<sup>2</sup> The last reference amounts to my doctoral degree thesis, written in Portuguese. In this paper, I use a little part of the discussion I made in the thesis’ second chapter. Here I had translated, revised and largely expanded my PhD discussion, articulating in more detail a topic which I was quite unable to tackle in the thesis: how different theories in philosophy of mind would impact the acceptance or rejection of the mind-dependence thesis for social entities, with special attention to the zombie thought experiment.

However, someone can ask: what are ‘minds’ and ‘mental states’ in this context? It is an extremely relevant question because the mentalistic vocabulary related to ‘mind’, ‘mental’ and so on is used in many ways. It is important to be very careful about which of these uses is in play here.

The above question is related to another: does the mind-dependence thesis have the same consequences across different theories about the nature of the mental, or is the theoretical meaning itself modified according to different views about the mind?

Therefore, in this paper I shall examine these two questions. In the first section, I analyze different phenomena which fall on some concept of ‘mind’ or ‘mental’. It is possible to assert that there are two general domains of the mental in the philosophy of mind: intentionality (cognitive and attitudinal); consciousness (qualitative/experiential). However, I shall argue that the relevant “mental states” for the mind-dependence thesis are intentional states. That is, the mind-dependence thesis asserts that social entities ontologically depend on intentional mental states. In fact, the mind-dependence theoreticians in social ontology also commit themselves with the assertion that social entities ontologically depend on sensorimotor states (related to perception and behavior), however, they understand that these sensorimotor states need to be attached to intentional mental states in order to generate social entities.

In the second section, I will discuss why we should expect that differences among philosophical accounts of mind makes a difference for what consequences the mind-dependence thesis has. That is, the mind-dependence thesis’ theoretical implications are not invariant under variations of theories about the nature of mind and intentionality. It seems to me that the social ontology literature assumes that the mind-dependence thesis for social entities is neutral in relation to different philosophical accounts of mind: someone’s account of mind is not relevant to acceptance (or rejection) of the mind-dependence thesis for

social entities. However, I think it is an untenable position. My aim is to show that the mind-dependence thesis is affected by what minds are.

Here I will focus on a specific case: the philosophical zombie thought experiment. Are there social entities in a zombie world? Of course, it depends on whether zombie worlds are possible and what mental states zombie worlds are devoid of.

## **2 What is “mind” for the mind-dependence thesis in social ontology?**

In the first place, we could notice that the social ontology literature uses the term ‘mind’ in the context of mind-dependence in an ambiguous way. ‘Mind’ is not defined or at least entirely circumscribed. However, specific mental states are listed and employed in theoretical constructions. For example, in relation to the mainstream account of social ontology, for which the notion of collective intentionality is decisive for making the social world:

The philosopher Francesco Guala, in fact, has gone so far as to call this second view the “Standard Model of Social Ontology.” The idea is this: the social world is a kind of projection of our thoughts, or attitudes, onto the world. We, as a community, make the social world by thinking of it in a particular way. The bills in my pocket are money because we all think of them as money. The president has the powers he does because we grant him those powers. America is a nation because we think of it as such. The social world, quite generally, is the social world in virtue of our beliefs about it (Epstein, 2015, p. 50; it is important to note that Epstein criticizes this position).

It seems that most writers in the field of social ontology agree that social institutions are created by we-attitudes, or at least this thesis has not been challenged (Ylikoski; Makela, 2002, p. 3; they challenge this assumption).

To put the point succinctly: ontologically speaking, to create a minimal institutional reality you need exactly three things: (1) human beings (or some sort of being with the relevantly similar cognitive capacities) with (2) intentionality, including collective intentionality including the capacity to impose functions on objects and people, and (3) a language capable of Declarational speech acts—that is, speech acts with the double direction of fit (Searle, 2010, p. 109).

Our aim is to explain human social ontology. Because that ontology is created by the mind, we have to begin with that property of the mind that creates the reality we are trying to analyze. We have to begin with intentionality (Searle, 2010, p. 25).

In the context of the dissidents from mainstream social ontology, which refuse the claim that all social kinds depend on collective intentionality, the emphasis concentrates on intentional mental states too:

In this paper, I have distinguished three kinds of social kinds, all of which are mind-dependent, though in significantly different ways. The first kind are dependent on human propositional attitudes though not about the kind itself, while the second kind are dependent on attitudes about the kind itself, though the existence of individual members of the kind may not be so dependent. The third kind of social kind depend [sic] for their existence on propositional attitudes about the kind itself, and individual members of the kind depend for their existence on someone having the relevant attitudes towards them. This third kind of social kind coincides with the most conventional of social kinds. The reason that instances of the third kind of social kind may not be natural kinds is not that they are mind-dependent but that the properties associated with them are conventionally rather than causally linked (Khalidi, 2013, p. 16-17).

In some sense, it is surely correct that social entities partly depend on people, society, and human intentionality — otherwise, they would not be “social” at all. However, the role of intentions and other attitudes is often overestimated, making prevailing views excessively intellectualist (Epstein, p. 2013, p. 54).



If we brainstorm on kinds of facts that might serve to anchor a constitutive rule for some social entity Y, it is easy to come up with a long list of candidates. Without attempting to be at all exhaustive, here are a few candidates roughly in order, from the most stringently intentional or intellectualized to the least:

- (A1) Explicit collective agreement regarding what it takes to be a Y
- (A2) Collective acceptance of the constitutive rule for Y
- (A3) Collective acceptance of something other than a constitutive rule for Y
- (A4) Widespread common (but not collective) intentions toward Y
- (A5) Intentions of one or a few individuals toward Y, with practices spread by intentional reproduction
- (A6) Intentions of one or a few individuals toward Y, with practices spread by mere causal transmission
- (A7) Patterns or regularities in practices and so on (Epstein, 2013, p. 61-62).

As I think it has turned clear, the point of the mind-dependence thesis concerns a specific kind of ‘mental’: intentional mental states, generally called ‘propositional attitudes’<sup>3</sup>, such as beliefs, desires, intentions and so on (beyond intentional mental states, Epstein also talks about practices). Social ontologists are primarily concerned with intentional mental states.

This domain-restriction could generate a little terminological misunderstanding. In fact, the mind-dependence thesis in social ontology is a ‘propositional-attitude-dependence’ thesis or ‘intentional-states-dependence-thesis.’ In this sense, the mind-dependence thesis is not about dependence on sensorial or motor states, although this dependence is also accepted. The reason is that dependence on sensorimotor states is too thin for counting as a relevant substantive thesis about mind-dependence. It is a trivial claim insofar social entities involve (animal) living things: all animals exhibit sensory and motor

---

<sup>3</sup> Searle is an exception: for him, intentional mental states are not propositional attitudes. In his framework, intentional mental states are not directed to propositions. See discussion in Searle (2010).

features independently of whether they have higher-order cognition or not.

A survey on books about the philosophy of mind reveals that the bulk of discussion revolves around two general themes: the nature of intentionality and the nature of consciousness. And consciousness and intentionality are distinct issues, even when their occurrences are intertwined with each other (for example, when we have a specific belief, and we have awareness of this belief).

In the intentionality domain, we are concerned with cognition and other intentional notions, generally interpreted as propositional attitudes. Examples are beliefs, desires, intentions, and so on. In the consciousness domain, we are concerned with conscious states, degrees of awareness and the distinction between conscious and unconscious states. Examples are phenomenal consciousness, access consciousness, and so on<sup>4</sup>. The distinction is helpful because intentional mental states do not need to be conscious, at least not every time (you have various beliefs which are not being consciously accessed at this moment), and

---

<sup>4</sup> Christopher Hill (2007) provides a list of sorts of consciousness discussed in the specialized literature: “It is customary to distinguish five forms of consciousness: agent consciousness (which is what we have in mind when we say of an agent that he is “losing consciousness” or “regaining consciousness”), propositional consciousness (which is expressed by the “conscious that” construction), introspective consciousness (which is what we have in mind when we say things like “His affection for me is fully conscious, but his hostility is not”), relational consciousness (which is expressed by the “conscious of” construction), and phenomenal consciousness (which is a property that mental states possess when they have a phenomenological dimension — that is, when they present us with such qualitative characteristics as pain and the taste of oranges). [...] In addition to the five forms of consciousness that have just been identified, will also discuss what take to be a sixth form — a form that can reasonably be called experiential consciousness. [...] Finally, it is sometimes maintained that there is a seventh form of consciousness, access consciousness, which is possessed by all and only those higher mental states that are poised for control of speech, reasoning, and intentional action. There will be some consideration of this proposal in the present chapter, but it will not figure prominently in later discussions. For reasons that will emerge, the standard characterizations of access consciousness are best seen as approximations to a correct account of experiential consciousness” (Hill, 2007, p. 2-3)

there are phenomenal conscious states which are non-intentional (such as pain states).

In support of my contention that it is relevant to quote some work in philosophy of mind in support of my contention that qualitative consciousness and intentionality are marks of the mental, that is, we attribute mind when we have attributed intentional states or conscious states for a person, animal or robot (and we do not attribute mind if we have verified sensory or motor states *simpliciter*<sup>5</sup>). For example, Churchland (1981) comments that intentionality and qualia are the two great questions involving philosophy of mind, under the rubric of physicalism and the nature of the mind:

My purpose in this paper is to explore these projections, especially as they bear on (1) the principal elements of common-sense psychology: the propositional attitudes (beliefs, desires, etc.), and (2) the conception of rationality in which these elements figure. This focus represents a change in the fortunes of materialism. Twenty years ago, emotions, qualia, and “raw feels” were held to be the principal stumbling blocks for the materialist program. With these barriers dissolving, the locus of opposition has shifted. Now it is the realm of the intentional, the realm of the propositional attitude, that is most commonly held up as being both irreducible to and ineliminable in favor of anything from within a materialist framework (Churchland, 1981, p. 67).

---

<sup>5</sup> Nevertheless, these sensory and motor states can be used as evidence of intentionality or consciousness if they pass some functional or structural threshold. My contention is not that perception and bodily movement are entirely unrelated to mental phenomena. I assert only the weaker claim that perception and bodily movement are not sufficient to attribute mind for a person, animal or robot — or at least a non-trivialized attribution of mind. That is, sensory and motor notions are relevant to the philosophy of mind, I do not negate it. However, they are not properly mental, and their discussion in the philosophy of mind always intersects the two proper mental domains mentioned: intentionality and consciousness. My point is that these sensory and motor notions are properly biological notions and their relation to the mental domain depends on some relation with intentionality or qualitative experience, for example, when sensory notions are related to the qualitative experience in talking about conscious perceptual data.

Note that Churchland makes use of the distinction I made. On one side of the question, we have propositional attitudes such as beliefs and desires. On the other hand, we have emotions, qualia and raw feels. The problem of the mind in relation to the physical occurs because these mental elements seem irreducible to the physical. Similar questions simply are not raised from the case of sensorial and motor states whose domain is biological life. The biological domain is not populated by any seemingly extra mysterious entity since vitalism falls out of fashion.

It is important to note that the different kinds of conscience discussed in the philosophy of mind literature have different relations to intentional mental states. Phenomenal conscience is related to perception and sensory states, although access conscience is related to thought and intentional action. It means that access conscience is a notion of consciousness strongly related to propositional attitudes. In fact, access conscience is defined as an access relation to someone's own mental states (which include intentional mental states): is the sort of consciousness that consists in a mental state's being accessible to (or accessed by, as preferred by Burge, 2010b) an individual through his or her rational, cognitive powers. Nevertheless, I agree with Tyler Burge (2010a) that phenomenal consciousness is the core notion for understanding consciousness. There is no point in attributing access consciousness for a being without phenomenal consciousness at all. It does not mean that access consciousness is always accompanied by phenomenal qualities or elements, however, if a being never has phenomenal consciousness, this being cannot be conscious from the start. This assertion can be disputed, however, for the sake of this paper, I assume it is right.

The mind-dependence thesis discussed in social ontology is related to the intentionality domain. All social ontologists think that social entities depend on the sensorimotor states, however, it is trivially true as I said before. The talk about mind-dependence is a stronger

claim: social entities depend on intentional/cognitive states, which are mental *proper*.

In this sense, the rejection of the mind-dependence thesis in social ontology would mean that social entities do not depend on intentional mental states. However, these social entities would ontologically depend on sensorimotor states yet, such as perception and behavior. That is, social entities would exist and depend on sensorimotor states even if intentional mental states are absent (in a counterfactual scenario) and independently of causal relations between sensorimotor states and intentional mental states (in the actual scenario). If sensorimotor notions are called 'mental', then mind-dependence for social entities is trivial and inevitable. However, the mind-dependence thesis (in its usual formulation) asserts a more substantive point about the ontological dependence of social entities, establishing this dependence on intentional mental states such as intentions and beliefs, in addition to sensorimotor states which are rendered ontologically insufficient to the ontological constitution of the social world.

And about the consciousness domain of mental? In this sense the social ontology literature is very ambiguous. Maybe some social metaphysicians accept this dependence because they think all social beings have qualitative experience, or because qualitative experience and intentional states are necessarily associated (in fact, Searle is a case in point). However, it is possible to argue that the qualitative experience dimension of the human mental life is simply irrelevant for matters about the metaphysical building of the social world. These various positions about consciousness are compatible with the mind-dependence thesis as described in social ontology literature.

In the face of the discussed distinction, it is adequate to say that if a theory about the social world refuses the claim that social entities depend on intentional mental states, then this theory refuses the mind-dependence thesis. This assertion can be associated with the idea that social entities also depend on consciousness, but it is not an obligatory

claim. It includes claims about conscious access to intentional mental states such as collective intentions related to a social kind. Even claims such as that are not obligatory requirements of the mind-dependence thesis in the context of social ontology. For asserting the mind-dependence thesis, it suffices to assert that social entities depend on intentional mental states, independently of presence or absence of consciousness (e.g. phenomenal consciousness). On the other hand, if someone refuses the claim that social entities depend on intentional mental states although accepting the claim that social entities depend on phenomenal consciousness (e. g. some phenomenal qualities involved in perception), I count this person as refusing the mind-dependence thesis too. It can seem like odd terminology. However, in the current formulation of the mind-dependence thesis ‘mind’ stands by ‘intentional mental states’, not by ‘consciousness’ (especially ‘phenomenal consciousness’). In reason of that I find my terminology is justified. However, if you prefer to understand this case as a modified mind-dependence thesis, I do not have any problem so long being clear that this alternate mind-dependence thesis is a non-intentionality version of mind-dependence thesis, and that it amounts to a rejection of the intentionality version of mind-dependence thesis (which it is the standard formulation in social ontology). For simplicity, I had maintained the abbreviated name ‘mind-dependence thesis’ for the intentionality version.

### **3 The relationship between social ontology and philosophy of mind**

Another question we can ask in the context of the mind-dependence thesis is how different theories about the nature of the mind can affect the theoretical underpinnings of the mind-dependence thesis for social entities. That is, is it possible that the theories of mind had a major impact on social ontology than previously thought. Mind-

dependence is about “mind”, but different philosophical views conceive minds in diverse ways, making contradictory claims.

Note that it goes beyond definitions of intentionality or consciousness or even perception and action. Here the point concerns conceptions of the mental domain, not its concepts. Is consciousness conceived as a physical entity? Is intentionality a right description of our mental lives? And so on. These challenging questions are formulated and discussed in contemporary philosophy of mind.

In this work I only intend to talk about a very preliminary assessment of this question. It seems to me that a very generic understanding of the mental for the mind-dependence is employed in social ontology literature. Searle in fact is an exceptional case: he explicitly commits himself to a specific vision about how the mind constitutes language and action alike. And his approach to these matters is opposite to the computational and cognitivist conventional paradigm which prevails in analytical philosophy of mind nowadays.

An interesting thought experiment we can discuss here is that of the zombie. This thought experiment is related to the controversy between property dualism and physicalism (including non-reductive physicalism). Its original formulation was made by Robert Kirk (1974a; 1974b) and much contemporary discussion had employed the characterization made by David Chalmers (1996; 2010). The zombie thought experiment invites us to conceive of a physically identical world as our actual world, with the same physical profile but without conscious experiences. That is, in the zombie world, there are beings like us in all physical respects, our physical and functional duplicates, however they do not have any phenomenal experience at all. Your zombie duplicate has the same microphysical profile, neurobiology, behavioral responses, etc. as well you do. However, your zombie duplicate lacks qualitative experience entirely. In situations where you experience a visual quale, the zombie experiences nothing. This zombie world seems to be conceivable and, if conceivability is a guide to metaphysical

possibility, this world is also metaphysically possible. And if a zombie world is metaphysically possible, physicalism is false.

Here I will not argue that zombie worlds have social entities. It is a task I want to reserve for a future paper. Nevertheless, I think it is a plausible *prima facie* claim about zombie worlds. For example, we can conceive of monetary transactions occurring among zombies simply why they are physically indiscernible to us and therefore they behave in an equal fashion. The material they use seems to be the same as what we use: pieces of paper, bits on a computer screen, cigarettes, salt, and so on. That is, their currency has the same physical composition, and it is used in the same way. However, in a zombie world no one is a conscious agent. In this sense, they have the social kind money, but they are not conscious of it.

At first glance, zombie worlds do not seem relevant to the mind-dependence thesis for social entities, even if someone agrees that zombie worlds have social kinds, since zombie worlds are devoid of conscious experiences, which is not the mental feature relevant to the mind-dependence thesis. By the structure of the thought experiment, zombies do not have phenomenal consciousness, and, following Burge, they do not have any sort of consciousness, including access consciousness.

I hold with Block that access to central rational powers and rational activity of an individual is a necessary condition for a kind of consciousness having to do with rational access. But I believe, contrary to Block, that poised accessibility to central rational activity is not a sufficient condition for any sort of consciousness. No matter how accessible and no matter how poised for use in reasoning, rational control of action, or rational control of speech a thought may be, it can still fail to be conscious in every way. As noted, I think that a phenomenal zombie, no matter how freely rational, is not conscious in any sense. A being that lacks phenomenal consciousness altogether could not be conscious in any way. It would not, for example, have imageless conscious thought. Phenomenal consciousness need not be part of a thought or of its articulation in order for the thought to be rational-



access conscious. But there must be some phenomenal consciousness—some sensed or imaged what-it-is-like quality—in the individual for a thought to count as conscious in any sense. The individual may close his or her eyes and think imageless thoughts with a surround of subliminal, phenomenally conscious blackness. Or the individual might have some sort of super blindsight but have some other sensations or a capacity to articulate thoughts through phenomenal verbalizations. But lacking some such frame of phenomenal consciousness, one's thinking could not be conscious in any way. A phenomenal zombie has no consciousness—no matter how efficiently rational its behavior, verbalizations, and reasoning (Burge, 2007, p. 385-386).

It would be problematic for the mind-dependence thesis if we found possible worlds where social kinds obtain without intentional mental states. However, zombies can be understood as being intentional agents, whereas nonconscious ones. Therefore, it seems *prima facie* possible to extend the standard analysis of social entities in terms of collective intentionality, shared attention or individual intentions to zombie worlds.

However, I think it is not the last word about it. There are at least two ways to use zombie worlds in arguments against the mind-dependence thesis. In both, we will suppose that social entities occur in zombie worlds. And we will employ one premise blocking the analysis in terms of intentional mental states.

The first way is to block the analysis in terms of collective intentionality by stating that collective intentionality is not possible without a subjective perspective. Zombies do not have a first-person perspective because there is no what it is like to be a zombie. We maybe could use the analysis of individual intentionality for zombies accepting this caveat. 'Individual' stands for 'one-zombie thought'. However, it seems that collective intentionality has a strong disadvantage here: collective intentions seem to presuppose some awareness of other agents *qua* first-perspective agents. However, if a zombie does not conceive himself as "I", what is the meaning of saying that a zombie conceives

himself and other zombies as “we”? I think more work is needed before we can say if this line of reasoning works or not, among different accounts of collective intentionality (for example, Searle’s account is quite distinct from Margaret Gilbert’s and Tuomela’s accounts; see Gilbert, 1989; Tuomela, 2007).

The second way has a more general character, comprehending analysis both from collective and individual intentionality perspectives, however, this version depends on a controversial premise. If someone already agrees with a principle that consciousness and intentionality are strongly linked, then zombie worlds are also devoid of intentionality altogether. Searle’s connection principle (discussed in Searle, 1993) states that intentional states must, at least in principle, be accessible to consciousness. In Searle’s own words:

Only a being that could have conscious intentional states could have intentional states at all, and every unconscious intentional state is at least potentially conscious. This thesis has enormous consequences for the study of the mind. It implies, for example, that any discussion of intentionality that leaves out the question of consciousness will be incomplete. It is possible to describe the logical structure of intentional phenomena without discussing consciousness—indeed, for the most part, I did so in *Intentionality* (Searle 1983), but there is a conceptual connection between consciousness and intentionality that has the consequence that a complete theory of intentionality requires an account of consciousness (Searle, 1992, p. 138).

This idea, that all unconscious intentional states are in principle accessible to consciousness, I call the connection principle [...] (Searle, 1992, p. 161).

Zombies’ candidate intentional states cannot be accessible to consciousness since zombies do not have phenomenal states, then zombies do not really have *bona fide* intentional states either (by Searle’s connection principle). It follows that if zombie worlds have social kinds, then social kinds are not mind-dependent and the mind-dependence thesis for social entities is false.

It is important to note that it is a very controversial principle. For example, Barbara Montero (2006, p. 159, n. 10) states that the central problem discussed by Fodor and Lepore (1994) against Searle's connection principle concerns crossworld identification: while the connection principle must be about token mental states, it seems that there are token mental states that are essentially nonconscious in as much as the addition of consciousness makes them into different states.

In the specific context of zombie thought experiment, Searle's connection principle is discussed by Katalin Balog:

Contrary to Searle, a good case can be made that zombie-Jackson does have intentional states. Zombie-Jackson communicates with his colleagues: he answers questions, his utterances convey information, his actions are made intelligible by the assumption that he has beliefs and desires, etc. His cognitive organization seems to be essentially the same as Jackson's. On a more technical note: on all the extant theories of meaning, zombie-Jackson will count as a thinker (Balog, 1998, p. 39).

Although a necessary connection between intentionality and consciousness is not universally accepted, it is important to note its relevance for the debate of the mind-dependence thesis in social ontology. If someone accepts the connection principle, the mind-dependence thesis in social ontology can suffer a decisive blow whether there is at least the metaphysical possibility of nonconscious beings having social kinds. By the connection principle, nonconscious beings do not have intentional states and then social kinds do not depend on intentional mental states.

A weaker claim than the Searle's connection principle also works in this context. Someone could argue that, although the connection principle is not generally valid, it is valid for some cases, such as of the collective intentionality (in this case, the argument amounts to the previous one I mentioned) or of individual intentions related to social phenomena (for example, shared attention, mutual beliefs, and so

on, which seems to presuppose some awareness of other agents). Then, endorsing Searle's connection principle is not strictly necessary for the argument. It is enough that intentional states related to the social phenomena do not comfortably adjust to scenarios entirely devoid of consciousness.

## 4 Concluding remarks

In this paper I have stressed that the philosophy of mind can have more impact on social ontology than previously thought. I think the mind-dependence thesis for social entities could be at stake if we look closely into thought experiments in philosophy of mind. However, a close examination of the mind-dependence thesis never was undertaken in English language (whereas in Portuguese language, see Brito Jr., 2021). This paper provides an initial sketch for the purpose of addressing this gap in social ontology literature.

I think there is a non-explored theoretical room for non-mentalistic social ontology. Social ontology can be reconceived by refusing or reformulating the mind-dependence thesis. One case in which a non-mentalistic formulation is helpful was discussed here. Given a non-mentalistic social ontology, it is expected that zombie worlds have social entities. What seems to be problematic in a mentalistic social ontology is made intelligible by dropping the mind-dependence thesis (in its standard formulation).

However, some caveats and limitations should be appointed. The zombie thought experiment makes the point against the mind-dependence thesis given very specific conditions. That is, someone needs to accept that the zombie world is metaphysically possible and that the Searle's connection principle or a similar although weaker principle which pl to collective intentionality or individual intentionality related to social phenomena. Both are controversial assertions. And it means that this style of argument works only for the dualist about

consciousness. However, that is exactly what I wanted to discuss in this paper. Someone's theory of mind affects what the mind-dependence thesis does at the theoretical landscape. Arguments against mind-dependence thesis for physicalists must use zombie thought experiments acceptable from a physicalist perspective (e.g. Dretske's zombie, behavioral zombies, and so on), a possibility I intend to explore in a future paper. And insofar its being limited in scope, Searle's perspective about social ontology is vulnerable to this argument, since Searle accepts Searle's connection principle. In fact, Searle (2010, p. 135-137) made a discussion about robots behaving as if they could make monetary transactions which resemble the possibility of the zombie argument I sketch here<sup>6</sup>. For him, apparent social robots without consciousness cannot be social at all. Then, Searle refuses the status of social entities for behavioral patterns of nonconscious agents. But can he (and social ontologists in general) refuse this status for nonconscious agents in a principled way? Accepting that these robots imagined by Searle do not constitute social reality, what is the reason for it? Is it lacking consciousness or are there other grounds for it? A full examination of these issues

---

<sup>6</sup> "What goes for promising goes for institutional reality generally. Consider property. We might program robots so that for each robot there is a set of objects to which it stands in a special relation. It interacts with its objects in a way that it does not interact with the objects of other robots, and the other robots are programmed to respond similarly in ways that distinguish between their objects and objects of the other robots. Does this ersatz give us private property? I don't think so. What we have left out is any sense of property rights. The robots respond to objects as they might respond to rain. They are programmed to respond in certain ways and that's that. Similar remarks could be made about money. The robots, we will suppose, are given a set of counters that they can use to satisfy their needs. When they run short of fuel, they go to the fuel provider and exchange counters for fuel. But the difficulty in the way that I have described it is that the features do not yet have the structure of a genuine exchange. There is no buying or selling, because there is no sense of property rights, duties, or obligations involved. In a completely mechanistic universe, buying and selling would be exactly like the operation of a machine that spits out four quarters when a dollar bill is inserted" (Searle, 2010, p. 136-37).

is not available yet. I think it is time for bringing non-mentalistic social ontologies into the spotlight.

## Referências

- ARSTILA, Valtteri; LLOYD, Dan (Eds.). *Subjective Time. The Philosophy, Psychology, and Neuroscience of Temporality*. Cambridge: The MIT Press, 2014.
- BOSTROM, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford, UK: Oxford University Press, 2014.
- BOSTROM, Nick; YUDKOWSKY, Eliezer. The ethics of artificial intelligence. In: *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press, p. 316-334, 2014.
- BRENTANO, Franz. *Psychology from an Empirical Standpoint*. Transl. A.C. Rancurello, D.B. Terrell, and L. McAlister, London: Routledge, 1973.
- BRINGSJORD, Selmer; GOVINDARAJULU, Naveen. Artificial Intelligence, *The Stanford Encyclopedia of Philosophy* (Fall 2018 Ed.). ZALTA, Edward N. (Ed.). Disponível em: <https://plato.stanford.edu/archives/fall2018/entries/artificial-intelligence/>.
- HUSSERL, Edmund. *Logical Investigations*. Transl. J. N. Findlay. London: Routledge, 1973.
- KURZWEIL, Ray. *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Penguin USA, 2000.
- KURZWEIL, Ray. *The Singularity Is Near: When Humans Transcend Biology*. New York: Penguin USA, 2006.
- PROUDFOOT, Diane. Software Immortals: Science or Faith? In: EDEN, Amnon; MOOR, James; SORAKER, Johnny; STEINHART, Eric (Eds.). *Singularity Hypotheses – A Scientific and Philosophical Assessment*. Berlin: Springer-Verlag, p. 367-394, 2012.
- RUSSELL, Stuart J.; NORVIG, Peter. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited, 2016.
- SANDBERG, Anders. The Physics of Information Processing Superobjects: Daily Life Among the Jupiter Brains. *Journal of Evolution and Technology* 5 (1) (1999).

- WEARDEN, John; O'DONOGHUE, Alan; OGDEN, Ruth; MONTGOMERY, Catharine. Subjective Duration in the Laboratory and the World Outside. In: ARSTILA, Valtteri; LLOYD, Dan (Eds.). *Subjective Time. The Philosophy, Psychology, and Neuroscience of Temporality*. Cambridge: The MIT Press, p. 287-305, 2014.
- YAMPOLSKIY, Roman; FOX, Joshua. Artificial Intelligence and the Human Mental Model. In: EDEN, Amnon; MOOR, James; SORAKER, Johnny; STEINHART, Eric (Eds.). *Singularity Hypotheses – A Scientific and Philosophical Assessment*. Berlin: Springer-Verlag, p. 129-146, 2012.





# The subjective rate of time in artificial intelligence in light of the Metaphysics of Time

Ana Maria Correa Moreira da Silva<sup>1</sup>

DOI: <https://doi.org/10.58942/eqs.158.06>

## 1 Introduction

In his book *Superintelligence: Paths, Dangers* (2014), Nick Bostrom analyzes the possibility of creating three different forms of superintelligence which, in a sense, can be considered equivalent: *speed superintelligence*, *collective superintelligence*, and *quality superintelligence*. The first form (*speed superintelligence*) is a system that can perform all tasks that the human intellect can, but at a much faster pace; the second (*collective superintelligence*) is a system composed of a large number of smaller intellects, whose performance substantially exceeds that of any current cognitive system; and the third (*quality superintelligence*) is a system that is at least as fast as a human mind and vastly smarter qualitatively.

For our present purposes, we will focus here on the concept of *speed superintelligence*, an example of which would be a whole brain emulation running on fast hardware. This example assumes the

---

<sup>1</sup> Doutora em Filosofia pela PUC-Rio e Pesquisadora Colaboradora Plena junto ao Programa de Pós-Graduação em Filosofia da Universidade de Brasília.

Email: [ana.maria@unb.br](mailto:ana.maria@unb.br)

possibility of a hypothetical future technology which could enable an organic brain to be uploaded to a digital computer. The idea is that if the uploading process is successful, then the computer program will replicate the essential functional characteristics of the original brain. This process might give rise to some exotic properties shared by artificial minds that operate faster than human minds, one of which is the *subjective rate of time*, as introduced by Nick Bostrom and Eliezer Yudkowsky in their paper *The Ethics of Artificial Intelligence* (2014).

Preliminarily, the *subjective rate of time* should be distinguished from two other related concepts: the *objective rate of time* and the *subjective perception of time*. The *objective rate of time* belongs to the external world with all its physical laws, including the concept of Spacetime introduced by Einstein's Theory of Special Relativity. Diversely, the *subjective perception of time* corresponds to the different ways in which people sense the so-called "passage of time" or "time flow". Accordingly, one can perceive time flowing in a faster or slower manner depending on a range of external factors, such as the use of certain drugs, the kind of activities one performs (exciting or boring), a person's age or health conditions, and so on.

By its turn, the *subjective rate of time* represents a third approach to the possibility of measuring time, which is situated between the pure subjectivity and the mere objectivity. In this sense, it is neither a *subjective perception* nor an *objective rate*, but it is a *subjective rate* (as its name indicates), combining aspects of both concepts. Roughly speaking, the subjective rate of time is associated with the speed of information processing in artificial minds: the faster a mind operates, the slower it perceives the external world. This is a novel concept which raises a host of challenging questions within the debate on the Philosophy of Artificial Intelligence (AI).

An ethical issue raised by Bostrom and Yudkowsky is based on the hypothesis that a fast intelligence possesses *Sentience* — understood as the capacity for phenomenal experience or *qualia*, such as the ability

to feel pain and suffer. There is another capacity which, together with Sentience, characterizes the possession of a moral status by an AI, which is *Sapience* — understood as a set of capacities associated with higher intelligence, such as self-awareness and being a reason-responsive agent. The common view is that many animals have *qualia* and therefore some moral status, but that only human beings have sapience, which gives them a higher moral status than non-human animals.

Let us consider the case of a sentient AI, which is defined as a system with the capacity of suffering. And let us consider that it processes information one thousand times faster than a biological brain. In this scenario, we realize that one objective year of suffering for a human being could correspond to a thousand subjective years of suffering for this fast artificial mind. Faced with this much longer duration of the painful experience for the AI, it is ethically relevant to determine whether the duration of an experience should be measured in objective or subjective time. Bostrom and Yudkowsky propose that, in cases where the duration of an experience is of moral significance, it is the experience's subjective duration that should count, measured by the *subjective rate of time*, and not the experience's objective duration, measured by the objective rate of time.

The aim of this paper is to analyze the concept of *subjective rate of time* possessed by a fast AI (or a speed superintelligence) in light of some issues in the Metaphysics of Time. Firstly, we examine this concept in comparison with some conceptions about the variability in temporal perception in human beings, related to the duration of events and the speed of time passage. This will highlight both similarities and differences between the subjective rate of time and human subjective perception of time, based on the notions of clock speed and information processing. Secondly, we depart from Bostrom and Yudkowsky's example of the perception of an external event (the falling of a cup) by a fast AI, in order to examine how its subjective rate of time relates to concepts within the phenomenology of temporal subjectivity, focusing

on the distinction between *duration of experience* and *experience of duration*. This will confirm the already mentioned distinction between the two types of subjective time, based on the necessity or contingency of their causal production. We conclude by suggesting that this concept of how the AI experiences the world encompasses not only a quantitative nature but also a qualitative one. This observation has important consequences to the ethical problem of how to measure the duration of a painful experience in a fast sentient AI — whether objectively or subjectively.

## 2 Making sense of the concept of subjective rate of time

As we have seen, the concept of subjective rate of time presupposes the existence of a whole brain emulation running on fast hardware. The resulting upload may inhabit a simulated virtual reality or, alternatively, may interact with the external physical reality through some robotic mechanism. Despite the many questions that arise in the context of such a scenario — related to personal identity construed by means of memories and personality — it is feasible to assume that this being (or Upload) would be sentient and, as such, subject to ethical concerns. Bostrom and Yudkowsky have put the issue as follows:

Suppose that an upload could be sentient. If we run the upload program on a faster computer, this will cause the upload, if it is connected to an input device such as a video camera, to perceive the external world as if it had been slowed down. For example, if the upload is running a thousand times faster than the original brain, then the external world will appear to the upload as if it were slowed down by a factor of thousand. Somebody drops a physical coffee mug: The upload observes the mug slowly falling to the ground while the upload finishes reading the morning newspaper and sends off a few emails. One second of objective time corresponds to 17 minutes of subjective time. Objective and subjective duration can thus diverge (Bostrom; Yudkowsky, 2014, p. 327).

This variable subjective rate of time is grounded in the concept of *Superintelligence*, conceived as an AI sufficiently intelligent to understand its own design and to redesign itself or create a successor system, more intelligent, which could then redesign itself again to become even more intelligent, and so on, almost *ad infinitum*. In their words:

Superintelligence may also be achievable by increasing processing speed. The fastest observed neurons fire 1000 times per second; the fastest axon fibers conduct signals at 150 meters/second, a half-millionth the speed of light (Sandberg 1999). It seems that it should be physically possible to build a brain which computes a million times as fast as a human brain, without shrinking its size or rewriting its software. If a human mind were thus accelerated, a subjective year of thinking would be accomplished for every 31 physical seconds in the outside world, and a millennium would fly by in eight and a half hours. Vinge (1993) referred to such sped-up minds as “weak superintelligence”: a mind that thinks like a human but much faster (Bostrom; Yudkowsky, 2014, p. 330).

The possibility of increasing the processing speed is raised by Bostrom and Yudkowsky with basis on some data provided by Anders Sandberg in the paper “The Physics of Information Processing Superobjects: Daily Life Among the Jupiter Brains” (1999), where he links all types of intelligence — either as the act of thinking or as human abilities like economy, art and emotion — to *information processing*, with the consequence that the physics of information processing imposes limits on what can be achieved by any intelligent civilization. The speed at which information may be processed is affected by the fact that nothing can communicate faster than light. The faster the processing speed of a system, the longer the delays will appear from an internal subjective view. The finitude of light speed inevitably introduces delays in the signals sent across an extended system. The general idea is that some transition or “clock cycle” takes a time to occur depending on the length scale of the system and the speed of signals sent through it: the more clock cycles occur, the larger the “subjective” delay appears.

In the book *Subjective Time. The Philosophy, Psychology, and Neuroscience of Temporality*, Valtteri Arstila and Dan Lloyd (2014) present several essays on the subjectivity of temporality, which can be useful in our analysis of the subjective rate of time related to processing speed.

In this sense, Wearden, O'Donoghue, Ogden and Montgomery (2014, p. 287) defined the concept of *subjective duration* as judgments of various sorts about how long *stimuli* and events last, or judgments about how fast time seems to pass. These two kinds of judgments can be based on either *direct experience* of the event judged, or some kind of *inference* about the event. Regarding this latter difference, in general, when “fast time” is reported, the report is generated on the basis of an inference, often prompted by an external time marker, without any actual “feel” of fast time during the event. In contrast, “slow time” seems to have been directly experienced during the event itself, when attention is paid to time during it.

Judgments about time flow raise questions about the phenomenal “speed” of time during some event. Accordingly, there are several questions to be posed: would the estimate of the interval duration be longer if “time flew,” or shorter? If time “goes faster” during some event, is there “more time” in the event, or is the event shortened in time, just as journey time is shortened if we go faster?

In order to clarify these matters, the authors introduce the theory of internal clock as follows:

[...] Suppose that the estimated duration of some event is ordinarily related to the number of “ticks” that occur in it: more ticks therefore give rise to longer estimates. The effect of increasing the rate of ticks (“clock speed”) would be to increase duration estimates (as more ticks accumulate than before), so external stimuli would seem to last longer, presumably giving rise to the sensation of time “dragging.” So, if the clock “speeds up,” perceived time would appear to “slow down,” and the reverse is also true of clock slowing: slower clocks make events seem shorter, so (at least presumably) give rise to time “speeding up” (Wearden; O'donoghue; Ogden; Montgomery, 2014, p. 288).

A similar analysis is achieved through the relation between the amount of information stored in memory and the retrospective time estimate: the more information processed during a time period, the longer the retrospective duration judgment of that period. Additionally, longer events are associated with slower passage of time.

These are some remarks about the subjectivity of temporality that could shed new light on the concept of subjective rate of time in a fast AI.

We mentioned earlier the importance of distinguishing the AI's subjective time from a human's estimate or perception of how fast time flows, since the latter case can involve a distorted time perception rather than a shift in the rate of subjective time. The main difference seems to lie in the *necessity* of the subjective rate of time in AIs, contrasting with the *contingency* of the subjective perception of time in humans. In other words, while the fluctuations in human perception of time are influenced by external factors, the AI's perception of time is inherent to its internal structure (more precisely, to the increasing processing speed).

Nevertheless, we may observe some similarities between the two types of subjective time (in AIs and humans). In both cases, we notice that time seems to pass slower when an observed external event appears longer in its duration — which can happen when both the inner clock (in humans) and the information processing (in AIs) speed up. Conversely, time appears to flow faster when the perceived duration of an event seems shorter due to slowing the rate of clicks (the clock speed) or diminishing the rate of information processing, respectively.

This somewhat closeness in the subjectivity of time in AIs and humans also relates to the way the brain (organic or emulated) processes temporal information. As we shall see, the true nature of this temporal subjectivity remains so far unknown. There are models according to which time is inherent to the neural dynamic of the brain, as

an emergent property of its functioning. The working of brain mechanisms (neurons, cell-assemblies, neural networks, etc.) is inherently and essentially based on time, which is not simply an outcome of such mechanisms. Time judgements and their fluctuations result from complex interactions between different processes. In short, “the neural basis of time perception remains shrouded in mystery”, and “the basic question as to *how* and *where* time is processed in the brain remains unresolved” (Arstila; Lloyd, 2014, p. 482, 517).

### 3 Experience of duration and Duration of experience

As previously indicated, it is important to distinguish between the concept of *subjective rate of time* and the concept of *subjective perception of time*. The latter pertains to experiencing various temporal properties of events, such as their order, duration, time of occurrence, relations of simultaneity, anteriority, and posteriority. These temporal properties are attributed to both internal events, like introspective thoughts, and external events, such as the fall of a cup from a table. It's essential not to conflate them. The perception of the temporal properties of both the external and the internal events can be considered a subjective experience possessed by human beings and by other superior animals.

We are particularly interested in the subjective experience of the duration of events and the relation between internal events, like thoughts, and external events, like the falling of objects, from the perspective of their duration. Consider this specific scenario: a particular thought and the falling of a cup each last one second, indicating they share the same duration. We assume that the duration of these two events, whether internal or external, is measured in objective time. Objective time, or objective rate of time, refers to the temporal properties of internal and external events as measured by Physics.



To state that these two events last one second might imply they are simultaneous, but not necessarily so. Consider a scenario where someone thinks “it winds” for one second, and then, one second later, the cup in front of them falls. Despite both events having the same duration, they are not simultaneous as one occurs one second after the other.

Now, let’s imagine they are indeed simultaneous: the internal event of a thought and the external event of the cup falling. In this case, one possibility is that the internal event is the perception of the external event. For instance, someone witnesses a cup falling from a table and simultaneously thinks “the cup is falling.” At first glance, it appears that the duration of their perception (“the cup is falling”) is nearly the same as the duration of the fall itself, which is indeed possible. However, it is also plausible that the perception of the external event has a much longer duration than the event itself.

This is the case concerning the above-mentioned *speed superintelligence*, which processes information much faster than normal observers do, and, for this reason, experiences a slower passage of time. In Bostrom and Yudkowsky’s example, while the fall of the cup lasts one second, the experience of this event for a fast AI may endure 17 minutes. Conversely, if a human being is present in the same room, her experience of the fall may last just one second, coinciding with the actual duration of the fall itself.

This disparity between the objective rate of time and the subjective rate of time arises from the rapid rate of information processing in a brain or processor. Earlier, we provided some insights into this phenomenon: the faster the brain/processor runs, the slower the time passes for it. What does it mean to say that time passes slower for someone, slower than what, or in relation to what? The answer is: slower than the objective rate of time, which gives a kind of “universal” measure for the duration of internal and external events.

Thus, for Observer A, with a normal human brain, the fall of a cup lasts one second, while for Observer B, with speed superintelligence, the same external event lasts 17 minutes. The difference between the two cases can be found not only in the duration of the external event but also in the phenomenological experience of it. In this sense, there is a tension between the concepts of *duration of the experience* and *experience of duration*, as developed by the philosopher Franz Brentano and his student Edmund Husserl, who made significant contributions to our understanding of the subjective experience of temporality.

For Brentano (1973), the *duration of the experience* refers to the temporal extension of the experience, understood as a conscious mental act which possesses an intrinsic temporal character by enduring through time. He regarded mental acts as having a certain temporal span during which they persist in consciousness.

The *experience of duration*, as explored by Husserl (1973), involves how individuals perceive and experience the flow of time in their conscious awareness. He analyzed how the present moment is experienced as a temporal “now,” and how it is continuously shifting from the past to the future.

Comparing the two approaches, we notice that while Brentano focused on the temporal duration of mental acts themselves to build the concept of *duration of experience*, Husserl delved into the subjective experience of time and how individuals consciously perceive the passage of time in their lived experiences to construe the concept of *experience of duration*.

In our case, we suggest correlating the *duration of experience* with the *subjective rate of time*, as well as the *experience of duration* with the *subjective perception of time*.

We have previously argued that the difference between the *subjective rate of time* and the *subjective perception of time* may be rooted in the necessity and contingency of their causal production, respectively. A similar reasoning can be applied to the distinction between *duration*

*of the experience* and *experience of duration*. To explain: only contingently can a human being have a variable time perception and, thus, a different *experience of the duration* of an external event, because, in order to occur, this subjective perception needs some external factor, such as chemical agents, environmental contexts, the nature and complexity of the activities performed, and so on. Concerning the subjective rate of time, the speed superintelligence necessarily has a different *duration of its experience*, since it internally processes information much faster than human beings, and no external factor is needed for it.

To summarize, while the cause of a subjective perception of time lies in something external that contingently produces a peculiar experience of duration, the cause of a subjective rate of time lies in something internal that necessarily produces a peculiar duration of experience. In this sense, what distinguishes a human brain from a speed superintelligence is not only the experience of the duration of external events but also the duration of each experience itself.

## 4 Concluding remarks

Based on the analysis presented in this paper, we intend to conclude by defending the idea that, beyond the variability of the subjective rate of time, which is a quantitative property of fast artificial minds, the exponential increase in the speed of information processing qualitatively affects the way in which AIs experience the world. This intuition finds support in Sandberg's assertion mentioned above (1999) — that all forms of intelligence, including thoughts, feelings, and emotions, are inherently influenced by the mechanism and speed of *information processing*.

As illustrated by the example provided by Bostrom and Yudkowsky, the fact that a fast AI reads the newspaper in one second entails more than just having more time than we do for other activities or possessing vastly more knowledge than the smartest person in the

world. Due to its exceptionally rapid information processing, the act of reading a newspaper holds an entirely different significance for the AI compared to us, encompassing potential thoughts, impressions, and feelings elicited by this seemingly simple experience.

Assuming the existence of an essential qualitative aspect in the way the AI experiences the world has important implications for the ethical question of how to measure the duration of a painful experience in a fast sentient AI — whether objectively or subjectively. Bostrom and Yudkowsky's proposed solution to this problem appears controversial, since it relies solely on the contentious assumption that quantitatively reducing this duration would suffice to balance the amount of suffering between humans and AIs. One should consider the possibility that the suffering of the fast AI may involve more intense pain — or a distinct type of pain — not only due to the prolonged duration of the experience but also because of its unique nature resulting from the increased processing speed, which can deeply affect this complex phenomenon.

The claim that fast AIs can perceive the world both quantitatively and qualitatively also underscores the observation that the subjective experience of temporality, whether in organic or emulated brains, remains mysterious. The subjective rate of time represents just one of the many perplexing implications of the concept of speed superintelligence.

## Referências

- ARSTILA, Valtteri; LLOYD, Dan (Eds.). *Subjective Time. The Philosophy, Psychology, and Neuroscience of Temporality*. Cambridge: The MIT Press, 2014.
- BOSTROM, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford, UK: Oxford University Press, 2014.
- BOSTROM, Nick; YUDKOWSKY, Eliezer. The ethics of artificial intelligence. In: *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press, p. 316-334, 2014.

- BRENTANO, Franz. *Psychology from an Empirical Standpoint*. Transl. A.C. Rancurello, D.B. Terrell, and L. McAlister, London: Routledge, 1973.
- BRINGSJORD, Selmer; GOVINDARAJULU, Naveen. Artificial Intelligence, *The Stanford Encyclopedia of Philosophy* (Fall 2018 Ed.). ZALTA, Edward N. (Ed.). Disponível em: <https://plato.stanford.edu/archives/fall2018/entries/artificial-intelligence/>.
- HUSSERL, Edmund. *Logical Investigations*. Transl. J. N. Findlay, London: Routledge, 1973.
- KURZWEIL, Ray. *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Penguin USA, 2000.
- KURZWEIL, Ray. *The Singularity Is Near: When Humans Transcend Biology*. New York: Penguin USA, 2006.
- PROUDFOOT, Diane. Software Immortals: Science or Faith? In: EDEN, Amnon; MOOR, James; SORAKER, Johnny; STEINHART, Eric (Eds.). *Singularity Hypotheses – A Scientific and Philosophical Assessment*. Berlin: Springer-Verlag, p. 367-394, 2012.
- RUSSELL, Stuart J.; NORVIG, Peter. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited, 2016.
- SANDBERG, Anders. The Physics of Information Processing Superobjects: Daily Life Among the Jupiter Brains. *Journal of Evolution and Technology* 5 (1) (1999).
- WEARDEN, John; O'DONOGHUE, Alan; OGDEN, Ruth; MONTGOMERY, Catharine. Subjective Duration in the Laboratory and the World Outside. In: ARSTILA, Valtteri; LLOYD, Dan (Eds.). *Subjective Time. The Philosophy, Psychology, and Neuroscience of Temporality*. Cambridge: The MIT Press, p. 287-305, 2014.
- YAMPOLSKIY, Roman & FOX, Joshua. "Artificial Intelligence and the Human Mental Model. "In: EDEN, Amnon, MOOR, James, SORAKER, Johnny & STEINHART, Eric (Eds.). *Singularity Hypotheses – A Scientific and Philosophical Assessment*. Berlin: Springer-Verlag, p. 129-146, 2012.





**Instituto Quero Saber**  
[www.institutoquerosaber.org](http://www.institutoquerosaber.org)  
[editora@institutoquerosaber.org](mailto:editora@institutoquerosaber.org)

### **Informações técnicas**

capa projetada com ativos de *Freepik.com*

formato: *16 x 23 cm*

tipografia: *Palatino Linotype*



Os livros que agora publicamos reúnem os trabalhos apresentados no encontro, preservando a riqueza das reflexões desenvolvidas durante o evento. Cada artigo representa uma contribuição valiosa para o avanço do debate filosófico e para o fortalecimento das comunidades acadêmicas que buscam compreender e transformar a realidade por meio do pensamento crítico. Mais do que um registro, estas publicações são um convite à continuidade das discussões iniciadas no encontro, abrindo caminho para novas investigações e diálogos.

Desde sua estreia oficial como grupo de trabalho da Associação Nacional de Pós-Graduação em Filosofia (ANPOF) em 2018, o GT Metafísica Analítica tem se esforçado para criar um ambiente de cooperação e produtividade entre os pesquisadores, professores e alunos de filosofia que se interessam pelas discussões em metafísica realizadas no contexto da filosofia analítica. Além dos encontros periódicos realizados pela ANPOF, nos últimos anos o GT também realizou uma sessão especial no Simpósio Principia, em 2023, e organizou um encontro internacional com o GT Ética no mesmo ano, bem como o Colóquio do GT de Metafísica Analítica, nos anos de 2022, 2023 e 2024. Todos esses eventos contaram com a participação de importantes pesquisadores nacionais e internacionais. No último encontro da ANPOF foram realizadas catorze apresentações no GT, duas a mais em relação à edição anterior. Algumas das apresentações se tornaram capítulos deste volume. Num momento em que os critérios de avaliação de qualidade das publicações estão passando por uma reformulação, acreditamos ser imperativo que a comunidade nacional tenha acesso ao que se é produzido por profissionais brasileiros. Nesse sentido, a Coleção ANPOF é um importante veículo para a disseminação das ideias que estão sendo discutidas em cada área. Não seria diferente no caso da metafísica analítica, que, apesar de tudo, ainda é uma área pouco conhecida no Brasil.

